



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada

TESIS DOCTORAL

Contribuciones al proceso de
Súper-Resolución mediante
técnicas de filtros selectivos,
topología de Macro-Bloques
adaptable y sistemas Multi-Cámara

Eduardo Quevedo Gutiérrez
Las Palmas de Gran Canaria, abril de 2015



**D. PEDRO PÉREZ CARBALLO SECRETARIO DEL INSTITUTO
UNIVERSITARIO DE MICROELECTRÓNICA APLICADA DE LA
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,**

CERTIFICA,

Que el Consejo de Doctores del Departamento en su sesión de fecha **siete de abril de 2015** tomó el acuerdo de dar el consentimiento para su tramitación, a la tesis doctoral titulada ***“Contribuciones al proceso de Súper-Resolución mediante técnicas de filtros selectivos, topología de Macro-Bloques adaptable y sistemas Multi-Cámara”*** presentada por el doctorando D. Eduardo G. Quevedo Gutiérrez y dirigida por los Doctores D. Gustavo I. Marrero Callicó, y D. Félix B. Tobajas Guerrero.

Y para que así conste, y a efectos de lo previsto en el Artº 6 del Reglamento para la elaboración, defensa, tribunal y evaluación de tesis doctorales de la Universidad de Las Palmas de Gran Canaria, firmo la presente en Las Palmas de Gran Canaria, a **catorce de abril de 2015**.



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada

Instituto: **Instituto Universitario de Microelectrónica Aplicada**

Programa de doctorado: **Ingeniería de Telecomunicación Avanzada**

Título de la Tesis

**“CONTRIBUCIONES AL PROCESO DE SÚPER-RESOLUCIÓN MEDIANTE
TÉCNICAS DE FILTROS SELECTIVOS, TOPOLOGÍA DE MACRO-
BLOQUES ADAPTABLE Y SISTEMAS MULTI-CÁMARA”**

Tesis Doctoral presentada por D. Eduardo G. Quevedo Gutiérrez

Dirigida por el Dr. D. Gustavo I. Marrero Callicó

Codirigida por el Dr. D. Félix B. Tobajas Guerrero

El Director,

El Codirector

El Doctorando,

Las Palmas de Gran Canaria, a 14 de abril de 2015



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada

DIVISIÓN DE DISEÑO DE SISTEMAS INTEGRADOS

TESIS DOCTORAL

Contribuciones al proceso de
Súper-Resolución mediante
técnicas de filtros selectivos,
topología de Macro-Bloques
adaptable y sistemas Multi-Cámara

Eduardo Quevedo Gutiérrez
Las Palmas de Gran Canaria, abril de 2015

Dedicado a mi novia, Patricia (tris)

Agradecimientos

Mantengo vivo el recuerdo de un día, sentado en el despacho de uno de mis directores de Tesis, que me preguntó: ¿tú por qué quieres hacer la Tesis? No recuerdo exactamente lo que le respondí en aquel entonces; sin embargo, sí me queda el sentimiento de que sabía que iba a contar con el apoyo de muchas personas en su desarrollo. En primer lugar, quiero agradecer a mis directores Gustavo I. Marrero Callicó y Félix B. Tobajas Guerrero, la confianza en mí para la elaboración de esta Tesis Doctoral y el haber estado ahí desde muchos años atrás. La combinación del pragmatismo de Félix y la imaginación de Gustavo han sido fundamentales para poder llevar a buen término esta Tesis Doctoral. Detrás de ellos se asienta además la División de Diseño de Sistemas Integrados (DSI) del Instituto Universitario de Microelectrónica Aplicada (IUMA) de la Universidad de Las Palmas de Gran Canaria (ULPGC), que conforma un amplio grupo de personas, a las que también he de agradecer su labor para que podamos estar hoy aquí. Un agradecimiento especial a Jesús de la Cruz Estévez y a Luis Sánchez Crespo por todo lo que me han ayudado a conformar las conclusiones que se extraen de esta Tesis Doctoral.

Quiero dar las gracias, además, a mi novia, Patricia, a quien dedico esta Tesis Doctoral al igual que hice con el Proyecto Fin de Carrera de la titulación de Ingeniero de Telecomunicación, y el de la titulación de Ingeniero en Electrónica, porque es quién más se lo merece (otra vez), ya que se ha volcado conmigo desde el primer momento y me ha apoyado en todo el sacrificio que supone realizar una Tesis Doctoral al tiempo que he desarrollado mi actividad laboral. Ha sido también primordial el apoyo con el que he contado desde el consorcio donde he trabajado durante los últimos cuatro años, la Plataforma Oceánica de CANarias (PLOCAN), para poder desarrollar mi Tesis Doctoral, con el beneplácito de su Director, Octavio Llinás González, y de su Gerente, Joaquín Hernández Brito.

Gracias, además, a mis padres, que siempre me han apoyado en todas mis decisiones, y me han inculcado unos valores y una cultura sin los que no podría haber acabado esta Tesis Doctoral. Espero haber heredado alguna de sus abundantes virtudes; de mi madre, la pasión por la escritura, de mi padre, el interés por la ciencia y la tecnología y su divulgación. No me puedo olvidar de mi hermano, su mujer y el bebé que han tenido hace poco, que me hacen ver que todo esto tiene un sentido más allá del trabajo. He tenido la inmensa suerte de contar con una gran familia, que forma parte de lo que soy, y con unos amigos con los que he podido compartir buenos y malos momentos, especial mención a las acertadas revisiones de muchos de los artículos presentados a congresos y revistas por parte de Dácil Sánchez y David Horat, y a la portada diseñada por Graciela Gutiérrez Diepa y Carolina Cantero Márquez. A ellos también, gracias.

Resumen

En la mayoría de los sistemas electrónicos de captación de imágenes se requieren imágenes de alta resolución. Las imágenes de alta resolución resultan fundamentales en áreas relacionadas principalmente con dos bloques de aplicación: la mejora de información visual para la interpretación humana, y la ayuda a la representación para sistemas de percepción automática. La resolución de imagen es un término que describe su nivel de detalle. Cuanto mayor sea la resolución, más detalles contendrá la imagen. La resolución de una imagen digital puede clasificarse atendiendo a diferentes criterios: resolución de píxel, resolución espacial, resolución espectral, resolución temporal, y resolución radiométrica. En este contexto, la presente Tesis Doctoral se centra en contribuir a la mejora de la resolución espacial.

La técnica basada en la utilización de múltiples observaciones de una misma escena para generar una imagen de alta resolución se conoce como Súper-Resolución (*Super-Resolution*). Existen otras técnicas de Súper-Resolución tomando otro tipo de información como punto de partida, si bien la planteada en la presente Tesis Doctoral es una de las más comunes, conociéndose generalmente en la literatura como Súper-Resolución por fusión. Aunque se han propuesto diversas aproximaciones diferentes desde que se introdujo el concepto de Súper-Resolución, la mayoría de ellas opera adecuadamente en imágenes de *test* y no tanto en situaciones reales. Para realizar de forma práctica comparativas adecuadas entre secuencias de vídeo o imágenes, se han determinado una serie de métricas objetivas y subjetivas; entre éstas se utilizan en la presente Tesis Doctoral el PSNR (*Peak Signal-to-Noise Ratio*) y el índice SSIM (*Structural SIMilarity*). Por otra parte, a la hora de concebir un sistema de SR práctico, existen muchos retos para que éste pueda aplicarse en un amplio rango de aplicaciones. Los desafíos clave en el desarrollo y aplicación de las técnicas de Súper-Resolución han propiciado las aportaciones de la presente Tesis Doctoral (centradas en técnicas de Súper-Resolución por fusión), que se resumen en:

- **Eficiencia en el cómputo:** Estudio de cómo disminuir el cómputo del algoritmo considerando un pre-procesamiento en la información de baja resolución de entrada, determinando los *frames* más adecuados a considerar en el proceso de Súper-Resolución (Filtro Selectivo de Ventana) haciendo uso de la métrica objetiva de similitud estructural SSIM.
- **Registro de la imagen:** Demostración de cómo el uso de una estructura basada en múltiples cámaras (*array* de cámaras o sistema Multi-Cámara) puede influir de forma positiva en el problema planteado por la Súper-Resolución en relación con el registro de la imagen, ya que al disponer de múltiples observaciones de una misma secuencia, y considerando que el desplazamiento presente entre los *frames* capturados sea subpíxel, el resultado de la imagen de alta resolución procesada mejorará las prestaciones frente a una situación en la que se disponga de una única cámara.

- **Robustez:** Introducción de un filtro orientado a cada uno de los Macro-Bloques (Filtro Selectivo de MB) en los que se divide la secuencia de entrada para realizar el proceso de Súper-Resolución. Se mostrará que un estudio particular para cada uno de los Macro-Bloques proporciona una mayor robustez al algoritmo BSR usado como punto de partida, ya que al habilitar un tamaño variable de Macro-Bloque y determinar el grado de movimiento local de cada uno de ellos frente a una aproximación basada únicamente en movimiento global, la calidad global se incrementa. Además, se plantea un caso real orientado a secuencias de vídeo submarino grabadas por vehículos ROV (*Remotely Operated Vehicle*), presentando cómo el cálculo de la velocidad media del vehículo partiendo de información de georreferenciación proporciona información clave a la hora de determinar el número de *frames* a seleccionar en el proceso de Súper-Resolución, lo que proporciona un mayor grado de robustez al algoritmo.
- **Límites en las prestaciones:** Se han diseñado bancos de prueba específicos ajustados para proporcionar información fidedigna de los resultados obtenidos. Al mismo tiempo, se han desarrollado sistemas que permiten determinar los máximos teóricos de calidad objetiva que se podrían alcanzar con los algoritmos propuestos.

Summary

In most digital imaging systems, high-resolution images are required. High-resolution images are essential in those areas which are mainly related to two main applications: improvement of pictorial information for human interpretation, and helping representation for automatic machine perception. Image resolution describes the details contained in an image, the higher resolution of an image, the more details it will contain. The resolution of a digital image can be classified in many different ways: pixel resolution, spatial resolution, spectral resolution, temporal resolution, and radiometric resolution. In this context, this PhD Thesis focuses on contributing to the improvement of spatial resolution.

Using multiple observations of the same scene to generate a high resolution image is known as Super-Resolution. There are other Super-Resolution techniques which take other information as a starting point, although the one considered in this PhD Thesis is one of the most common, generally known in the literature as fusion Super-Resolution. Although many different approaches have been proposed since the concept of Super-Resolution was introduced, most of them operate properly in test images rather than in actual situations. In order to determine appropriate comparatives between video sequences or images, a number of objective and subjective metrics have been identified; among these the PSNR (*Peak Signal-to-Noise Ratio*) and the SSIM index (*Structural SIMilarity*) metrics have been used in this PhD thesis. Moreover, when conceiving a practical Super-Resolution system, there are many challenges in order to operate in a wide range of applications. The key challenges in the development and application of Super-Resolution techniques have led to the contributions of this PhD Thesis (focused on fusion Super-Resolution techniques), which can be summarized as follows:

- **Computation efficiency:** Study of how to reduce the computation time of the proposed algorithms considering pre-processing steps applied to the Low-Resolution input information, determining the most appropriate frames to be considered in the Super-Resolution process (*Window Selective Filter*) using the structural similarity objective metric.
- **Image registration:** Demonstration of how the use of a structure based on multiple cameras (array of cameras or Multi-Camera system) may have a positive impact on the problem posed by Super-Resolution in relation to image registration. This is due to the fact that, when counting with multiple observations of the same sequence, and considering that the displacement between frames is subpixel displacement, the result of the processed High-Resolution image improves the performance compared to a situation with a single camera.

- **Robustness aspects:** Introduction of a filter for each Macro-Block (*Block Selective Filter*) in which the input sequence is divided in the Super-Resolution process. It will be shown that a particular study for each Macro-Block provides greater robustness to the algorithm used as a reference, and that enabling a variable Macro-Block size and determining the level of local motion of each Macro-Block against an approach based solely on global motion, the overall quality is increased. Additionally, a real case oriented to underwater video sequences recorded by an ROV (*Remotely Operated Vehicle*) presents how the calculation of the average speed based on georeferencing information provides key information when determining the number of frames to be selected in the Super-Resolution process, which provides greater robustness to the algorithm.
- **Performance limits:** Specific test benches have been designed to provide reliable information on the obtained results. At the same time, systems have been developed to determine the maximum theoretical objective quality to be achieved with the proposed algorithms.

Índice de Contenidos

1.- INTRODUCCIÓN	1
1.1.- PLANTEAMIENTO DEL PROBLEMA.....	2
1.2.- MOTIVACIÓN DE LA TESIS DOCTORAL.....	4
1.3.- OBJETIVOS DE LA TESIS DOCTORAL.....	6
1.4.- ORGANIZACIÓN DE LA TESIS DOCTORAL.....	7
2.- ESTADO DEL ARTE	9
2.1.- INTRODUCCIÓN.....	10
2.2.- ESTADO DEL ARTE EN TÉCNICAS DE SÚPER-RESOLUCIÓN.....	11
2.3.- ESTADO DEL ARTE DE SISTEMAS MULTI-CÁMARA.....	22
2.4.- MÉTRICAS SOBRE CALIDAD DE IMAGEN.....	37
2.5.- DESAFÍOS DE LA SÚPER-RESOLUCIÓN Y CONTRIBUCIONES.....	44
2.6.- CONCLUSIONES.....	48
3.- CONTRIBUCIONES ALGORÍTMICAS	49
3.1.- INTRODUCCIÓN.....	50
3.2.- ALGORITMO BASE DE SÚPER-RESOLUCIÓN.....	51
3.3.- FILTROS SELECTIVOS Y TOPOLOGÍA DE MACRO-BLOQUES ADAPTABLE.....	65
3.4.- ENTORNO MULTI-CÁMARA.....	75
3.5.- CONCLUSIONES.....	84
4.- RESULTADOS Y ANÁLISIS: FILTROS SELECTIVOS Y TOPOLOGÍA DE MACRO-BLOQUES ADAPTABLE	85
4.1.- INTRODUCCIÓN.....	86
4.2.- ENTORNO DE SIMULACIÓN Y ANÁLISIS.....	87
4.3.- ANÁLISIS DE SIMULACIONES.....	92
4.4.- CONCLUSIONES.....	125

5.- RESULTADOS Y ANÁLISIS: ENTORNO MULTI-CÁMARA	127
5.1.- INTRODUCCIÓN.....	128
5.2.- ENTORNO DE SIMULACIÓN Y ANÁLISIS.....	129
5.3.- ANÁLISIS DE SIMULACIONES.....	138
5.4.- CONCLUSIONES.....	176
6.- ANÁLISIS DE SISTEMA GLOBAL	177
6.1.- INTRODUCCIÓN.....	178
6.2.- RESULTADOS DE SISTEMA GLOBAL.....	179
6.3.- APLICACIÓN EN ENTORNO SUBMARINO.....	191
6.4.- <i>TEST</i> DE CALIDAD SUBJETIVA	202
6.5.- CONCLUSIONES.....	204
7.- CONCLUSIONES Y LÍNEAS FUTURAS	205
7.1.- INTRODUCCIÓN.....	206
7.2.- CONCLUSIONES	207
7.3.- LÍNEAS FUTURAS	209
ANEXOS	211
A.1.- SECUENCIAS DE VÍDEO UTILIZADAS.....	213
A.2.- ENTORNO SUBMARINO: CARACTERÍSTICAS Y MEJORA DE IMAGEN.....	219
A.3.- <i>TEST</i> DE CALIDAD SUBJETIVA: CUESTIONARIO.....	229
A.4.- PUBLICACIONES.....	231
BIBLIOGRAFÍA	235

Índice de Figuras

1.- INTRODUCCIÓN	1
FIGURA 1.1.- IMAGEN DE <i>TEST (TEST TARGET)</i> 1951 USAF	2
2.- ESTADO DEL ARTE	9
FIGURA 2.1.- PROCESO DE SR: RECONSTRUCCIÓN DE IMAGEN HR A PARTIR DE VARIAS IMÁGENES EN LR	10
FIGURA 2.2.- MODELO DE OBSERVACIÓN DE UN SISTEMA DE IMAGEN	12
FIGURA 2.3.- SR BASADA EN ALINEAMIENTO Y POS-PROCESAMIENTO DEL DESENFUQUE	14
FIGURA 2.4.- EL MODELO MRF PARA SR DE UN ÚNICO <i>FRAME</i>	17
FIGURA 2.5.- <i>ARRAY MC</i> CON GEOMETRÍA SEMIESFÉRICA	23
FIGURA 2.6.- PROPUESTA DE WEINMANN <i>ET AL.</i> [WSR+11]	23
FIGURA 2.7.- DISPOSICIÓN DE CÁMARAS (A Y C) Y MODELO 3D (B Y D) DE DINOSAURIO Y RUINAS [ZCI+08]	24
FIGURA 2.8.- <i>ARRAY MC</i> DE 64 CÁMARAS DISTRIBUIDAS EN UNA MATRIZ CUADRADA DE 8x8	24
FIGURA 2.9.- TRES VISTAS DONDE SE ENFOCA A LA PERSONA DE LA IZQUIERDA, CENTRO Y DERECHA	25
FIGURA 2.10.- PROPUESTA DE SMITH <i>ET AL.</i> [SZJ+09]	25
FIGURA 2.11.- SISTEMA MODULAR PROPUESTO POR BAKER Y TANGUAY [BT06]	26
FIGURA 2.12.- IMAGEN CON GRAN ÁNGULO DE VISIÓN TOMADA MEDIANTE UN <i>ARRAY</i> DE DIMENSIONES 2x9	26
FIGURA 2.13.- PROPUESTA DE BELLOTTO <i>ET AL.</i> [BSB+09]	26
FIGURA 2.14.- PROPUESTA DE WILBURN <i>ET AL.</i> [WJV05]	27
FIGURA 2.15.- <i>ARRAYS MC</i> DE LA UNIVERSIDAD DE STANFORD	27
FIGURA 2.16.- SECUENCIA DE VÍDEO PROCESADA CON FOTOGRAFÍA DE APERTURA SINTÉTICA NO LINEAL	27
FIGURA 2.17.- <i>ARRAY MC</i> IMPLEMENTADO MEDIANTE CÁMARAS <i>GoPro</i> ®	28
FIGURA 2.18.- CÁMARA PANONO® PARA OBTENER PANORÁMICAS EN 360°	28
FIGURA 2.19.- SISTEMA <i>MC</i> DE LA EMPRESA <i>IDS</i> ® A TRAVÉS DE <i>USB 3.0</i>	29
FIGURA 2.20.- SISTEMA DE SR BASADO EN ESTABILIZACIÓN ÓPTICA DE LA IMAGEN [BT14]	29
FIGURA 2.21.- INTEGRACIÓN PREVISTA DE SISTEMA DE SR EN CÁMARA DE <i>iPhone</i> ® [BT14]	29
FIGURA 2.22.- PROPUESTA GENÉRICA DE KANG <i>ET AL.</i> [KLH08]	30
FIGURA 2.23.- <i>ARRAY MC</i> DE DIMENSIONES 3x3 DEL INSTITUTO UNIVERSITARIO DE YONSEI [HMG08]	31
FIGURA 2.24.- POSIBLES CONFIGURACIONES DEL <i>ARRAY MC</i> PROPUESTO POR LA UNIVERSIDAD DE ROCHESTER ...	33
FIGURA 2.25.- SOLAPE ENTRE LAS IMÁGENES CAPTURADAS POR LAS CÁMARAS	35
FIGURA 2.26.- <i>ARRAY MC</i> DE DIMENSIONES 2x2	35
FIGURA 2.27.- MODELO DE PRESENTACIÓN DE ESTÍMULO PARA EL MÉTODO ACR SEGÚN ITU-T P.911	43
FIGURA 2.28.- NÚMERO DE ARTÍCULOS DE SR DESDE EL AÑO 1984 HASTA EL AÑO 2013	46

3.- CONTRIBUCIONES ALGORÍTMICAS	49
FIGURA 3.1.- ESQUEMA GENERAL DE LAS CONTRIBUCIONES DE LA PRESENTE TESIS DOCTORAL	50
FIGURA 3.2.- DIAGRAMA DE BLOQUES DE ALGORITMO DE SR DINÁMICA TOMADO COMO REFERENCIA	51
FIGURA 3.3.- PROCESO DE BSR DESCRITO EN PSEUDO-CÓDIGO	52
FIGURA 3.4.- PROCESO DE ESTIMACIÓN DE MOVIMIENTO	52
FIGURA 3.5.- PROCESO DE COMPENSACIÓN DE MOVIMIENTO	53
FIGURA 3.6.- EJEMPLO DE COLISIONES EN LA MALLA DE VHR	54
FIGURA 3.7.- COMPENSACIÓN DE MOVIMIENTO POR MB	54
FIGURA 3.8.- MALLA DE VHR	55
FIGURA 3.9.- PARÁMETROS DE SR	57
FIGURA 3.10.- EJEMPLO DE VENTANA DE TRABAJO DE +/- 5 (#WBF = 4 Y #WFF = 5) CON CF = 6	59
FIGURA 3.11.- RESULTADOS CON SECUENCIA <i>GALDAR</i>	60
FIGURA 3.12.- RESULTADOS CON SECUENCIA <i>MOBCAL</i>	61
FIGURA 3.13.- RESULTADOS CON SECUENCIA <i>REEF</i>	62
FIGURA 3.14.- INFORMACIÓN ESPACIAL E INFORMACIÓN TEMPORAL	64
FIGURA 3.15.- APLICACIÓN DEL FILTRO DE VENTANA CORTANTE A VENTANA DE TRABAJO DE FIGURA 3.10	66
FIGURA 3.16.- APLICACIÓN DEL FILTRO SELECTIVO DE VENTANA A VENTANA DE TRABAJO DE FIGURA 3.10	67
FIGURA 3.17.- CONDICIÓN PARA CONSIDERAR UN MB EN EL PROCESO DE SR (FILTRO SELECTIVO DE MB)	68
FIGURA 3.18.- CONDICIONES PARA DIVIDIR MBS EN EL PROCESO DE SR USANDO VBS	70
FIGURA 3.19.- SELECCIÓN DE MBS PARA APLICAR SR EN UN <i>FRAME</i> DE LA SECUENCIA <i>SUZIE</i> UTILIZANDO FBS	73
FIGURA 3.20.- DIVISIÓN DE MBS EN FRAMES DE LAS SECUENCIAS <i>MOBILE</i> (A) Y <i>SUZIE</i> (B) USANDO VBS	73
FIGURA 3.21.- PSEUDOCÓDIGO INCLUYENDO FILTROS SELECTIVOS Y TOPOLOGÍA DE MBS ADAPTABLE	74
FIGURA 3.22.- DIAGRAMA DEL MÉTODO TEMPORAL-ESPACIAL USANDO LA CÁMARA 2 COMO REFERENCIA	76
FIGURA 3.23.- PROCESO DE REORDENAMIENTO DE <i>FRAMES</i>	76
FIGURA 3.24.- DIAGRAMA DEL MÉTODO ESPACIAL-TEMPORAL	77
FIGURA 3.25.- DIAGRAMA DEL MÉTODO MIXTO	78
FIGURA 3.26.- <i>ARRAY MC</i>	80
FIGURA 3.27.- DIVISIÓN DE LOS <i>FRAMES</i> DEL <i>ARRAY MC</i> EN BORDES Y SOLAPE	81
FIGURA 3.28.- PSEUDOCÓDIGO INCLUYENDO MÉTODOS MULTI-CÁMARA	83
FIGURA 3.29.- DIAGRAMA GLOBAL DE LAS APORTACIONES ALGORÍTMICAS DESARROLLADAS	84
4.- RESULTADOS Y ANÁLISIS: FILTROS SELECTIVOS Y TOPOLOGÍA DE MACRO-BLOQUES ADAPTABLE	85
FIGURA 4.1.- DIAGRAMA GLOBAL DE LAS APORTACIONES ALGORÍTMICAS DESARROLLADAS DESTACANDO FILTROS ...	86
FIGURA 4.2.- DISEÑO DEL BANCO DE PRUEBAS (<i>TESTBENCH</i>)	88
FIGURA 4.3.- FORMATOS DE MUESTREO (A) Y REPRESENTACIÓN DE PÍXELES EN FORMATO 4:2:0 (B)	89
FIGURA 4.4.- DIFERENCIA ENTRE WSFSR Y BSR RESPECTO A WSFTHR – SECUENCIA <i>FOREMAN</i>	94
FIGURA 4.5.- DIFERENCIA ENTRE WSFSR Y BSR RESPECTO A WSFTHR – SECUENCIA <i>SUZIE</i>	95
FIGURA 4.6.- DIFERENCIA ENTRE WSFSR Y BSR RESPECTO A WSFTHR – SECUENCIA <i>GALDAR</i>	95
FIGURA 4.7.- DIFERENCIA ENTRE WSFSR Y BSR RESPECTO A WSFTHR – SECUENCIA <i>MOBILE</i>	95
FIGURA 4.8.- TIEMPO DE EJECUCIÓN RESPECTO A WSFTHR – SECUENCIA <i>FOREMAN</i>	96
FIGURA 4.9.- VARIACIÓN PORCENTUAL DE TIEMPO RESPECTO AL UMBRAL	96
FIGURA 4.10.- SECUENCIA <i>FOREMAN</i> PARA UMBRAL ÓPTIMO = 80%	97
FIGURA 4.11.- SECUENCIA <i>GALDAR</i> PARA UMBRAL ÓPTIMO = 20%	97
FIGURA 4.12.- SECUENCIA <i>FLOWER</i> PARA UMBRAL ÓPTIMO = 30%	98
FIGURA 4.13.- VENTANA DE SR, PSNR Y SSIM CON RESPECTO A NÚMERO DE <i>FRAME</i> – SECUENCIA <i>FOREMAN</i>	99
FIGURA 4.14.- VENTANA DE SR CON RESPECTO A NÚMERO DE <i>FRAME</i> - SECUENCIA <i>GALDAR</i>	100
FIGURA 4.15.- VENTANA DE SR CON RESPECTO A NÚMERO DE <i>FRAME</i> - SECUENCIA <i>SUZIE</i>	100
FIGURA 4.16.- ANÁLISIS DEL <i>FRAME</i> 20 DE LA SECUENCIA <i>FOREMAN</i> QCIF PARA WSFSR	101
FIGURA 4.17.- DIFERENCIA ENTRE BSFSR Y BSR RESPECTO A BSFTHR – SECUENCIA <i>FOREMAN</i>	104

FIGURA 4.18.- TIEMPO DE EJECUCIÓN RESPECTO A BSFTHR – SECUENCIA <i>FOREMAN</i>	104
FIGURA 4.19.- SECUENCIA <i>FOREMAN</i> PARA UMBRAL ÓPTIMO = 4%	105
FIGURA 4.20.- NÚMERO DE MBS CON RESPECTO A NÚMERO DE <i>FRAME</i> , PSNR Y SSIM – SECUENCIA <i>FOREMAN</i> ..	106
FIGURA 4.21.- MEJORA DE PSNR CON UMBRAL ADAPTATIVO <i>FRAME A FRAME</i>	107
FIGURA 4.22.- DIFERENCIAS WSFSR-BSR Y WBSFSR-BSR RESPECTO A WSFTHR – SECUENCIA <i>FOREMAN</i>	107
FIGURA 4.23.- SECUENCIA <i>FOREMAN</i> PARA INT, BSR Y UMBRALES ÓPTIMOS DE WSFSR, BSFSR Y WBSFSR	108
FIGURA 4.24.- SECUENCIA <i>FOREMAN</i> PARA UMBRALES ÓPTIMOS DE WSFSR, BSFSR Y WBSFSR	108
FIGURA 4.25.- ANÁLISIS DEL <i>FRAME</i> 20 DE LA SECUENCIA <i>FOREMAN</i> QCIF PARA BSFSR	109
FIGURA 4.26.- DIFERENCIA ENTRE VBSSR Y BSR RESPECTO A VBSTHR1 Y VBSTHR2 – SECUENCIA <i>FOREMAN</i> ...	112
FIGURA 4.27.- SECUENCIA <i>FOREMAN</i> CONSIDERANDO VBSSR FRENTE A BSR E INT	113
FIGURA 4.28.- PORCENTAJE DE PÍXELES CLASIFICADOS EN UN TAMAÑO DE MB – SECUENCIA <i>FOREMAN</i>	113
FIGURA 4.29.- ANÁLISIS DEL <i>FRAME</i> 20 DE LA SECUENCIA <i>FOREMAN</i> QCIF PARA WSFSR	114
FIGURA 4.30.- PRESTACIONES DEL FILTRO VBSFSR CON RESPECTO A BSFSR	115
FIGURA 4.31.- PRESTACIONES DEL FILTRO VBSWSFSR CON RESPECTO A WSFSR	116
FIGURA 4.32.- PRESTACIONES DE FILTROS RESPECTO A WSFTHR – SECUENCIA <i>FOREMAN</i>	117
FIGURA 4.33.- TIEMPOS DE EJECUCIÓN RESPECTO A WSFTHR – SECUENCIA <i>FOREMAN</i>	117
FIGURA 4.34.- RESUMEN DE APORTACIÓN DE WSFSR, BSFSR Y VBSSR	119
FIGURA 4.35.- <i>FRAME</i> 70 DE LA SECUENCIA <i>FOREMAN</i> QCIF	120
FIGURA 4.36.- <i>FRAME</i> 20 DE LA SECUENCIA <i>FOREMAN</i> QCIF	120
FIGURA 4.37.- RESULTADOS DE PSNR PARA FRAMES EN SECUENCIAS <i>FOREMAN</i> , <i>FOOTBALL</i> , <i>MOBILE</i> Y <i>SUZIE</i>	121
FIGURA 4.38.- RESULTADOS PROMEDIO PARA VARIOS MÉTODOS (A) PSNR EN DB, (B) SSIM	122
FIGURA 4.39.- RESULTADOS VISUALES PARA LA SECUENCIA <i>FOREMAN</i> CIF (<i>FRAME</i> 70)	123
FIGURA 4.40.- RESULTADOS DE FIGURA DE MÉRITO FM	124
5.- RESULTADOS Y ANÁLISIS: ENTORNO MULTI-CÁMARA	127
FIGURA 5.1.- DIAGRAMA GLOBAL DE APORTACIONES ALGORÍTMICAS RESALTANDO ENTORNO MULTI-CÁMARA	128
FIGURA 5.2.- PROCESO DE CREACIÓN DE LAS SECUENCIAS MC ARTIFICIALES	129
FIGURA 5.3.- <i>TESTBENCH</i> DEL MÉTODO TEMPORAL-ESPACIAL	132
FIGURA 5.4.- PRE-PROCESAMIENTO <i>OVERLAP+BORDERS</i> EN <i>TESTBENCH</i> DEL MÉTODO TEMPORAL-ESPACIAL	133
FIGURA 5.5.- <i>TESTBENCH</i> DEL MÉTODO ESPACIAL-TEMPORAL	134
FIGURA 5.6.- PRE-PROCESAMIENTO <i>OVERLAP+BORDERS</i> EN <i>TESTBENCH</i> DEL MÉTODO ESPACIAL-TEMPORAL	135
FIGURA 5.7.- <i>TESTBENCH</i> DEL MÉTODO MIXTO Y DE LA SR ESPACIAL	136
FIGURA 5.8.- MODIFICACIÓN DEL <i>TESTBENCH</i> DE LA SR TEMPORAL	137
FIGURA 5.9.- <i>ARRAY</i> MC DE DIMENSIONES 3X3 PARA GENERAR LAS SECUENCIAS MC ARTIFICIALES	138
FIGURA 5.10.- <i>FRAMES</i> nº 12 SÚPER-RESUELTOS MRT (A) Y HR (B) PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>FULL-FRAME</i>	142
FIGURA 5.11.- <i>FRAMES</i> nº 12 INTERPOLADOS MRT (A) Y SR+INT (B) PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>FULL-FRAME</i>	142
FIGURA 5.12.- <i>FRAME</i> nº 12 INTERPOLADO INTx4 PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO TEMPORAL- ESPACIAL EN MODO <i>FULL-FRAME</i>	142
FIGURA 5.13.- <i>FRAMES</i> nº 14 SÚPER-RESUELTOS MRT (A) Y HR (B) PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>FULL-FRAME</i>	143
FIGURA 5.14.- <i>FRAMES</i> nº 14 SÚPER-RESUELTOS MRT (A) Y SR+INT (B) PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>FULL-FRAME</i>	143
FIGURA 5.15.- <i>FRAMES</i> nº 14 INTERPOLADO INTx4 PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO TEMPORAL- ESPACIAL EN MODO <i>FULL-FRAME</i>	143
FIGURA 5.16.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>OVERLAP</i>	145
FIGURA 5.17.- <i>FRAMES</i> nº 14 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>OVERLAP</i>	145
FIGURA 5.18.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>OVERLAP+BORDERS</i>	148

FIGURA 5.19.- <i>FRAMES</i> nº 14 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO TEMPORAL-ESPACIAL EN MODO <i>OVERLAP+BORDERS</i>	148
FIGURA 5.20.- <i>FRAMES</i> nº 10 DE <i>WATER COOLER</i> SÚPER-RESUELTO (A) E INTERPOLADO (B) A LA SALIDA DE LA 2ª FASE DE SR TEMPORAL EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>FULL-FRAME</i>	150
FIGURA 5.21.- <i>FRAMES</i> nº 42 DE <i>MOBCAL</i> SÚPER-RESUELTO (A) E INTERPOLADO (B) A LA SALIDA DE LA 2ª FASE DE SR TEMPORAL EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>FULL-FRAME</i>	151
FIGURA 5.22.- <i>FRAMES</i> nº 12 SÚPER-RESUELTOS MRS (A) Y HR (B) E INTERPOLADOS MR (C) Y SR+INT (D) PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>FULL-FRAME</i>	151
FIGURA 5.23.- <i>FRAMES</i> nº 14 SÚPER-RESUELTOS MRS (A) Y HR (B) E INTERPOLADOS MR (C) Y SR+INT (D) PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>FULL-FRAME</i>	152
FIGURA 5.24.- <i>FRAMES</i> nº 10 SÚPER-RESUELTOS (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>WATER COOLER</i> EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>OVERLAP</i>	153
FIGURA 5.25.- <i>FRAMES</i> nº 42 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>MOBCAL</i> EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>OVERLAP</i>	154
FIGURA 5.26.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>OVERLAP</i>	154
FIGURA 5.27.- <i>FRAMES</i> nº 14 SÚPER RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>SHIELDS</i> EN EL MÉTODO ESPACIAL-TEMPORAL EN MODO <i>OVERLAP</i>	154
FIGURA 5.28.- PSNR Y SSIM DE MÉTODO MIXTO EN MODO <i>FULL-FRAME</i> EN LA SECUENCIA <i>WATER COOLER</i>	156
FIGURA 5.29.- PSNR Y SSIM DE MÉTODO MIXTO EN MODO <i>FULL-FRAME</i> EN LA SECUENCIA <i>MOBCAL</i>	156
FIGURA 5.30.- PSNR Y SSIM DE MÉTODO MIXTO EN MODO <i>FULL-FRAME</i> EN LA SECUENCIA <i>STOCKHOLM</i>	156
FIGURA 5.31.- <i>FRAMES</i> nº 10 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>WATER COOLER</i> EN EL MÉTODO MIXTO EN MODO <i>FULL-FRAME</i>	157
FIGURA 5.32.- <i>FRAMES</i> nº 42 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>MOBCAL</i> EN EL MÉTODO MIXTO EN MODO <i>FULL-FRAME</i>	158
FIGURA 5.33.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO MIXTO EN MODO <i>FULL-FRAME</i>	159
FIGURA 5.34.- PSNR Y SSIM DE MÉTODO MIXTO EN MODO <i>OVERLAP</i> EN LA SECUENCIA <i>WATER COOLER</i>	160
FIGURA 5.35.- PSNR Y SSIM DE MÉTODO MIXTO EN MODO <i>OVERLAP</i> EN LA SECUENCIA <i>MOBCAL</i>	160
FIGURA 5.36.- PSNR Y SSIM DE MÉTODO MIXTO EN MODO <i>OVERLAP</i> EN LA SECUENCIA <i>STOCKHOLM</i>	160
FIGURA 5.37.- <i>FRAMES</i> nº 10 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>WATER COOLER</i> EN EL MÉTODO MIXTO EN MODO <i>OVERLAP</i>	161
FIGURA 5.38.- <i>FRAMES</i> nº 42 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>MOBCAL</i> EN EL MÉTODO MIXTO EN MODO <i>OVERLAP</i>	162
FIGURA 5.39.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> EN EL MÉTODO MIXTO EN MODO <i>OVERLAP</i>	162
FIGURA 5.40.- <i>FRAMES</i> SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>FUJIDOG</i> APLICANDO SR ESPACIAL EN MODO <i>FULL-FRAME</i>	164
FIGURA 5.41.- <i>FRAMES</i> SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA EL <i>FRAME</i> nº42 DE LA SECUENCIA <i>PARKRUN</i> APLICANDO SR ESPACIAL EN MODO <i>OVERLAP</i>	165
FIGURA 5.42.- <i>FRAMES</i> nº 10 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>WATER COOLER</i> APLICANDO SR TEMPORAL EN MODO <i>FULL-FRAME</i>	166
FIGURA 5.43.- <i>FRAMES</i> nº 42 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>MOBCAL</i> APLICANDO SR TEMPORAL EN MODO <i>FULL-FRAME</i>	167
FIGURA 5.44.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> APLICANDO SR TEMPORAL EN MODO <i>FULL-FRAME</i>	167
FIGURA 5.45.- <i>FRAMES</i> nº 10 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>WATER COOLER</i> APLICANDO SR TEMPORAL EN MODO <i>OVERLAP</i>	168
FIGURA 5.46.- <i>FRAMES</i> nº 42 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>MOBCAL</i> APLICANDO SR TEMPORAL EN MODO <i>OVERLAP</i>	169
FIGURA 5.47.- <i>FRAMES</i> nº 12 SÚPER-RESUELTO (A) E INTERPOLADO (B) HR PARA LA SECUENCIA <i>STOCKHOLM</i> APLICANDO SR TEMPORAL EN MODO <i>OVERLAP</i>	169
FIGURA 5.48.- MODO TEMPORAL-ESPACIAL PARALELO CON 10 <i>CORES/THREADS</i>	173
FIGURA 5.49.- MÉTRICA SSIM PARA MÉTODOS MC Y BSR EN SECUENCIAS BAJO ESTUDIO	174
FIGURA 5.50.- RESUMEN DE TIEMPOS DE CÓMPUTO (EN MS) DE SECUENCIAS BAJO ESTUDIO	175
FIGURA 5.51.- FIGURA DE MÉRITO <i>FM_MC</i> APLICADA A LAS SECUENCIAS BAJO ESTUDIO	175

6.- ANÁLISIS DEL SISTEMA GLOBAL	177
FIGURA 6.1.- SISTEMA GLOBAL PROPUESTO EN ESTA TESIS DOCTORAL	178
FIGURA 6.2.- PSNR (dB) PARA SECUENCIAS MC	179
FIGURA 6.3.- SSIM (%) PARA SECUENCIAS MC	180
FIGURA 6.4.- DIFERENCIA DE PSNR (dB) CON RESPECTO A BSR PARA SECUENCIAS MC	180
FIGURA 6.5.- DIFERENCIA DE SSIM (%) CON RESPECTO A BSR PARA SECUENCIAS MC	180
FIGURA 6.6.- RESULTADOS PARA SECUENCIA <i>MOBCAL</i>	182
FIGURA 6.7.- RESULTADOS PARA SECUENCIA <i>PARKRUN</i>	183
FIGURA 6.8.- RESULTADOS PARA SECUENCIA <i>SHIELDS</i>	184
FIGURA 6.9.- RESULTADOS DE TIEMPOS DE CÓMPUTO PARA SECUENCIAS MC	185
FIGURA 6.10.- RESULTADOS DE FIGURA DE MÉRITO <i>FM_G</i> PARA <i>MCABMSR</i> Y <i>MCSR</i>	185
FIGURA 6.11.- PÍXELES DE FACTORES DE ESCALA 2 (ROJO), 3 (AZUL) Y 4 (VERDE) PARTIENDO DE 1 PÍXEL DE LR	186
FIGURA 6.12.- RESULTADOS PARA FACTOR DE ESCALA 2, 3 Y 4 PARA INTERPOLACIÓN, BSR Y <i>MCABMSR</i>	187
FIGURA 6.13.- DIFERENCIA DE SR CON RESPECTO A INT PARA DIFERENTES FACTORES DE ESCALA	187
FIGURA 6.14.- RESULTADOS DE PSNR PARA FACTOR DE ESCALA 2, 3 Y 4 <i>FRAME A FRAME</i>	187
FIGURA 6.15.- RESULTADOS DE SSIM PARA FACTOR DE ESCALA 2, 3 Y 4 <i>FRAME A FRAME</i>	187
FIGURA 6.16.- RESULTADOS DE SECUENCIA <i>STOCKHOLM</i> PARA FACTOR DE ESCALA 4	188
FIGURA 6.17.- RESULTADOS DE SECUENCIA <i>STOCKHOLM</i> PARA FACTOR DE ESCALA 3	188
FIGURA 6.18.- RESULTADOS DE SECUENCIA <i>STOCKHOLM</i> PARA FACTOR DE ESCALA 2	188
FIGURA 6.19.- FUNCIONES QUE REALIZAN LA MAYOR PARTE DEL TRABAJO INDIVIDUAL (BSR)	189
FIGURA 6.20.- FUNCIONES QUE REALIZAN LA MAYOR PARTE DEL TRABAJO INDIVIDUAL (<i>MCABMSR</i>)	189
FIGURA 6.21.- RUTA CRÍTICA DE BSR	190
FIGURA 6.22.- RUTA CRÍTICA DE <i>MCABMSR</i>	190
FIGURA 6.23.- MUESTRAS INCLUSIVAS Y EXCLUSIVAS Y LLAMADAS ENTRANTES (%) - FUNCIÓN <i>SR</i> (<i>MCABMSR</i>) .	190
FIGURA 6.24.- PSNR (dB) PARA SECUENCIAS SUBMARINAS	192
FIGURA 6.25.- SSIM (%) PARA SECUENCIAS SUBMARINAS	192
FIGURA 6.26.- DIFERENCIA DE PSNR (dB) PARA SECUENCIAS SUBMARINAS	192
FIGURA 6.27.- DIFERENCIA DE PSNR SSIM (%) PARA SECUENCIAS SUBMARINAS	192
FIGURA 6.28.- RESULTADOS PARA SECUENCIA <i>CORAL</i>	193
FIGURA 6.29.- RESULTADOS PARA SECUENCIA <i>REEF</i>	194
FIGURA 6.30.- RESULTADOS PARA SECUENCIA <i>PEZ_CORAL</i>	195
FIGURA 6.31.- RESULTADOS DE TIEMPOS DE CÓMPUTO PARA SECUENCIAS SUBMARINAS	196
FIGURA 6.32.- RESULTADOS DE FIGURA DE MÉRITO <i>FM_G</i> PARA <i>MCABMSR</i> Y <i>MCSR</i>	196
FIGURA 6.33.- GRÁFICAS DE TAMAÑO DE VENTANA DE TRABAJO DE SR FRENTE A VALOR DE UMBRAL <i>WSFTHR</i> ...	197
FIGURA 6.34.- RESULTADOS PARA LA SECUENCIA <i>ATLANTIS</i>	198
FIGURA 6.35.- APLICACIÓN <i>WEB</i> DE VÍDEOS SUBMARINOS GEORREFERENCIADOS	199
FIGURA 6.36.- CÁLCULO DE LA VELOCIDAD MEDIA (<i>AVERAGE SPEED</i>) DEL VEHÍCULO	200
FIGURA 6.37.- RELACIÓN ENTRE VELOCIDAD MEDIA DE VEHÍCULO Y TAMAÑO DE VENTANA DE TRABAJO DE SR	200
FIGURA 6.38.- RESULTADOS DE EXTRACTO DE TRANSECTO 2N1 (<i>FRAMES 14819-14888</i>)	201
FIGURA 6.39.- RESULTADOS DE EXTRACTO DE TRANSECTO 2N1 (<i>FRAMES 16319-16358</i>)	201
FIGURA 6.40.- DETALLE DE SELECCIÓN DE <i>MCABMSR</i> COMO OPCIÓN PREFERIDA FRENTE A BSR.....	203
FIGURA 6.41.- DETALLE DE MEJORA SIGNIFICATIVA DE <i>MCABMSR</i> FRENTE A <i>FRAME</i> DE LR.....	205

ANEXOS	211
FIGURA A.1.- SELECCIÓN DE CÁMARAS CON QUE SE HA GRABADO LA SECUENCIA <i>WATER COOLER</i>	215
FIGURA A.2.- SELECCIÓN DE CÁMARAS CON QUE SE HA GRABADO LA SECUENCIA <i>FUJIDOG</i>	215
FIGURA A.3.- IMAGEN DE SECUENCIA <i>PEZ_CORAL</i>	216
FIGURA A.4.- IMAGEN DE SECUENCIA <i>GRUTA</i>	216
FIGURA A.5.- IMAGEN DE SECUENCIA <i>TITANIC</i>	216
FIGURA A.6.- IMAGEN DE SECUENCIA <i>CORAL</i>	217
FIGURA A.7.- IMAGEN DE SECUENCIA <i>REEF</i>	217
FIGURA A.8.- IMAGEN DE SECUENCIA <i>ATLANTIS</i>	217
FIGURA A.9.- PENETRACIÓN DE LA LUZ EN AGUA DE MAR, DEPENDIENDO DE LA LONGITUD DE ONDA	219
FIGURA A.10.- IMÁGENES SUBMARINAS CON TONALIDAD AZUL-VERDE (ABSORCIÓN DEL COLOR)	220
FIGURA A.11.- IMÁGENES SUBMARINAS CON TONALIDAD AZUL (ABSORCIÓN DEL COLOR)	220
FIGURA A.12.- DISPERSIÓN DE LA LUZ POR UNA PARTÍCULA	221
FIGURA A.13.- EFECTOS DE DISPERSIÓN DE LA LUZ EN EL MEDIO MARINO	222
FIGURA A.14.- EFECTO DE LA REFRACCIÓN DE LA LUZ SOLAR A CAUSA DE LA SUPERFICIE DEL MAR	222
FIGURA A.15.- EFECTOS DEL PÁRPADO DE LA LUZ EN LA IMAGEN SUBMARINA	223
FIGURA A.16.- EJEMPLOS DEL EFECTO DE LA TURBIDEZ EN LA IMAGEN SUBMARINA	223
FIGURA A.17.- EJEMPLOS DEL EFECTO DE LA ILUMINACIÓN ARTIFICIAL EN LA IMAGEN SUBMARINA	224
FIGURA A.18.- IMÁGENES ORIGINALES (ARRIBA) Y RESTAURADAS (ABAJO) [MIA12]	224
FIGURA A.19.- IMAGEN ORIGINAL (IZQUIERDA) Y RESTAURADA (DERECHA) [SK05]	225
FIGURA A.20.- IMAGEN ORIGINAL (IZQUIERDA) Y RESTAURADA CON 2 MÉTODOS (CENTRO Y DERECHA) [BG12] ...	225
FIGURA A.21.- IMÁGENES ORIGINALES (ARRIBA) Y RESTAURADAS (ABAJO) [TO06]	226
FIGURA A.22.- EJEMPLO DE MEJORA DE COLOR	227
FIGURA A.23.- IMAGEN ORIGINAL (IZQUIERDA) Y MEJORADA (DERECHA) [YCH+11]	227
FIGURA A.24.- IMÁGENES ORIGINALES (A, C) Y MEJORADAS (B, D) [PCC09]	227
FIGURA A.25.- CUESTIONARIO DE <i>TEST</i> DE CALIDAD SUBJETIVA Y EJEMPLO DE PREGUNTA PARA SEC. <i>STOCKHOLM</i> .	229
FIGURA A.26.- <i>FRAMES</i> DE <i>TEST</i> , DE IZQDA. A DCHA: <i>FRAME</i> DE LR, <i>FRAME</i> DE BSR Y <i>FRAME</i> DE MCABMSR	230

Índice de Tablas

2.- ESTADO DEL ARTE	9
TABLA 2.1.- RESUMEN DE REFERENCIAS QUE EMPLEAN UN <i>ARRAY</i> MULTI-CÁMARA CON SÚPER-RESOLUCIÓN	36
TABLA 2.2.- COMPARATIVA ENTRE MÉTRICAS OBJETIVAS DE CALIDAD DE IMAGEN	41
TABLA 2.3.- COMPARATIVA ENTRE MÉTRICAS DE EVALUACIÓN SUBJETIVAS ACR Y DCR	43
3.- CONTRIBUCIONES ALGORÍTMICAS	49
TABLA 3.1.- PÍXELES COMPENSADOS Y VECTORES DE MOVIMIENTO ASOCIADOS	53
TABLA 3.2.- NOMENCLATURA ASOCIADA A FILTROS SELECTIVOS DE VENTANA Y MACRO-BLOQUE	71
TABLA 3.3.- RELACIÓN ENTRE SSIM Y MOS	72
TABLA 3.4.- RELACIONES ENTRE MBS DEPENDIENDO DE LAS CÁMARAS Y LOS INSTANTES DE TIEMPO	79
4.- RESULTADOS Y ANÁLISIS: FILTROS SELECTIVOS Y TOPOLOGÍA DE MACRO-BLOQUES ADAPTABLE	85
TABLA 4.1.- EJEMPLO DE LA COMPATIBILIDAD DE TAMAÑOS	90
TABLA 4.2.- RESOLUCIONES CONSIDERADAS	90
TABLA 4.3.- RESUMEN DE RESULTADOS CON WSFSR	92
TABLA 4.4.- RELACIÓN ENTRE MOVIMIENTO GLOBAL, LOCAL Y VALOR ÓPTIMO DEL UMBRAL DE WSFSR	93
TABLA 4.5.- RESUMEN DE RESULTADOS CON BSFSR	102
TABLA 4.6.- RELACIÓN ENTRE MOVIMIENTO GLOBAL, LOCAL Y VALOR ÓPTIMO DEL UMBRAL DE BSFSR	102
TABLA 4.5.- MEJORA DE PSNR CON UMBRAL ADAPTATIVO	107
TABLA 4.8.- RESUMEN DE RESULTADOS CON VBSSR	110
TABLA 4.9.- RELACIÓN ENTRE MOVIMIENTO GLOBAL, LOCAL Y VALOR ÓPTIMO DE UMBRALES DE VBSSR Y FILTROS	110
TABLA 4.10.- RESULTADOS PROMEDIO PARA LOS 100 PRIMEROS <i>FRAMES</i> DE LA SECUENCIA <i>FOREMAN</i>	118
TABLA 4.11.- TIEMPO DE EJECUCIÓN PROMEDIO (EN SEGUNDOS) PARA APLICAR SR A UN <i>FRAME</i> CON UNA VENTANA DE TRABAJO DE 15 <i>FRAMES</i> (<i>FRAME</i> ACTUAL +/-7)	123
5.- RESULTADOS Y ANÁLISIS: ENTORNO MULTI-CÁMARA	127
TABLA 5.1.- <i>FRAMES</i> SELECCIONADOS DE CADA SECUENCIA PARA REALIZAR LAS SIMULACIONES	138
TABLA 5.2.- CONFIGURACIONES DE LOS PARÁMETROS DEL <i>ARRAY</i> MC EN CADA SECUENCIA	139
TABLA 5.3.- CONFIGURACIÓN ESTABLECIDA EN EL MÉTODO TEMPORAL-ESPACIAL PARA CADA SECUENCIA	139
TABLA 5.4.- VALORES PROMEDIO DE RESULTADOS EN MÉTODO TEMPORAL-ESPACIAL (MODO <i>FULL-FRAME</i>)	141
TABLA 5.5.- VALORES PROMEDIO DE RESULTADOS EN MÉTODO TEMPORAL-ESPACIAL (MODO <i>OVERLAP</i>)	144
TABLA 5.6.- VALORES PROMEDIO DE RESULTADOS EN MÉTODO TEMPORAL-ESPACIAL (MODO <i>OVERLAP+BORDERS</i>)	146
TABLA 5.7.- VALORES PROMEDIO DE LOS BORDES EN EL MÉTODO TEMPORAL-ESPACIAL	146

ÍNDICE DE TABLAS

TABLA 5.8.- RESUMEN DE VALORES PROMEDIO DE LOS RESULTADOS DE SSIM (MÉTODO TEMPORAL-ESPACIAL)	147
TABLA 5.9.- CONFIGURACIÓN ESTABLECIDA DEL MÉTODO ESPACIAL-TEMPORAL PARA CADA SECUENCIA	149
TABLA 5.10.- VALORES PROMEDIO DE RESULTADOS EN MÉTODO ESPACIAL-TEMPORAL (MODO <i>FULL-FRAME</i>)	150
TABLA 5.11.- VALORES PROMEDIO DE RESULTADOS EN MÉTODO ESPACIAL-TEMPORAL (MODO <i>OVERLAP</i>)	153
TABLA 5.12.- CONFIGURACIÓN ESTABLECIDA DEL MÉTODO MIXTO PARA CADA SECUENCIA	155
TABLA 5.13.- VALORES PROMEDIO DE LOS RESULTADOS DEL MÉTODO MIXTO EN MODO <i>FULL-FRAME</i>	157
TABLA 5.14.- VALORES PROMEDIO DE LOS RESULTADOS DEL MÉTODO MIXTO EN MODO <i>OVERLAP</i>	160
TABLA 5.15.- CONFIGURACIÓN ESTABLECIDA EN LA SR ESPACIAL PARA CADA SECUENCIA	163
TABLA 5.16.- VALORES PROMEDIO DE LOS RESULTADOS DE LA SR ESPACIAL EN MODO <i>FULL-FRAME</i>	164
TABLA 5.17.- VALORES PROMEDIO DE LOS RESULTADOS DE LA SR ESPACIAL EN MODO <i>OVERLAP</i>	164
TABLA 5.18.- CONFIGURACIÓN ESTABLECIDA EN EL PROCESO DE SR TEMPORAL PARA CADA SECUENCIA	166
TABLA 5.19.- VALORES PROMEDIO DE LOS RESULTADOS DE LA SR TEMPORAL EN MODO <i>FULL-FRAME</i>	166
TABLA 5.20.- VALORES PROMEDIO DE LOS RESULTADOS DE LA SR TEMPORAL EN MODO <i>OVERLAP</i>	168
TABLA 5.21.- RESULTADOS DE LA PRIMERA FASE DEL MÉTODO TEMPORAL-ESPACIAL (T-E) Y EL MÉTODO ESPACIAL-TEMPORAL (E-T)	170
TABLA 5.22.- RESULTADOS DE LA SEGUNDA FASE DEL MÉTODO TEMPORAL-ESPACIAL (T-E) Y EL MÉTODO ESPACIAL-TEMPORAL (E-T)	170
TABLA 5.23.- COMPARATIVA DE LOS VALORES MEDIOS DE LOS RESULTADOS ENTRE LOS DISTINTOS MÉTODOS	171
TABLA 5.24.- VALORES MÁXIMOS DE MÉTODOS MIXTO, TEMPORAL-ESPACIAL Y ESPACIAL-TEMPORAL	171
TABLA 5.25.- TIEMPO MEDIO DE PROCESAMIENTO DE UN <i>FRAME</i> (MÉTODOS MC, BSR E INTERPOLACIÓN)	172
TABLA 5.26.- TIEMPO MEDIO DE PROCESAMIENTO DE UN <i>MB</i> (MÉTODOS MC, BSR E INTERPOLACIÓN)	173
6.- ANÁLISIS DEL SISTEMA GLOBAL	177
TABLA 6.1.- DIFERENCIA MCABMSR CON RESPECTO A INT PARA SECUENCIAS MC (PSNR)	181
TABLA 6.2.- DIFERENCIA MCABMSR CON RESPECTO A INT PARA SECUENCIAS MC (SSIM)	181
TABLA 6.3.- DIFERENCIA MCABMSR CON RESPECTO A INT EN SECUENCIAS SUBMARINAS (PSNR Y SSIM)	191
TABLA 6.4.- SELECCIÓN DE MCABMSR COMO OPCIÓN PREFERIDA FRENTE A BSR	203
TABLA 6.5.- APRECIACIÓN DE MEJORA SIGNIFICATIVA DE MCABMSR FRENTE A <i>FRAME</i> DE LR	203
7.- CONCLUSIONES Y LÍNEAS FUTURAS	205
TABLA 7.1.- DIFERENCIAS MCABMSR-BSR Y MCABMSR-INT (SECUENCIAS CON MOVIMIENTO LOCAL O MOVIMIENTO GLOBAL ACUSADO)	208
TABLA 7.2.- DIFERENCIAS MCABMSR-BSR Y MCABMSR-INT (SECUENCIAS CON MOVIMIENTO GLOBAL Y MOVIMIENTO LOCAL LIMITADOS)	208
ANEXOS	211
TABLA A.1.- SECUENCIAS UTILIZADAS PARA COMPARACIÓN	213
TABLA A.2.- SECUENCIAS UTILIZADAS PARA COMPARACIÓN APROPIADAS PARA SISTEMAS MULTI-CÁMARA	214
TABLA A.3.- CARACTERÍSTICAS DE SECUENCIA <i>PEZ_CORAL</i>	216
TABLA A.4.- CARACTERÍSTICAS DE SECUENCIA <i>GRUTA</i>	216
TABLA A.5.- CARACTERÍSTICAS DE SECUENCIA <i>TITANIC</i>	216
TABLA A.6.- CARACTERÍSTICAS DE SECUENCIA <i>CORAL</i>	217
TABLA A.7.- CARACTERÍSTICAS DE SECUENCIA <i>REEF</i>	217
TABLA A.8.- CARACTERÍSTICAS DE SECUENCIA <i>ATLANTIS</i>	217

Índice de Acrónimos

<i>Acrónimo</i>	<i>Significado</i>
ACR	Absolut Category Rating
ANSI	American National Standards Institute
AQI	Anysotropic Quality Index
BSF	Block Selective Filter
BSFSR	Block Selective Filter Super-Resolution
BSFTHR	Block Selective Filter THReshold
BSR	Baseline Super-Resolution
CCD	Charged Coupled Device
CDN	Content Distribution Network
CDOM	Colored Dissolved Organic Material
CF	Current Frame
CFT	Continuous Fourier Transform
CIF	Common Intermediate Format
CMOS	Complementary Metal Oxide Semiconductor
CMPQM	Color Moving Pictures Quality Metric
CPU	Central Processing Unit
CR	Crámer-Rao
CT	Computed Tomography
DCR	Degradation Category Rating
DFT	Discrete Fourier Transform
DLT	Direct Lineal Transform
DM	Distortion Measure
DSI	Diseño de Sistemas Integrados
DSIS	Double Stimulus Impairment Scale
EITE	Escuela de Ingeniería de Telecomunicación y Electrónica
E-T	Espacial-Temporal
FBS	Fixed Block Size
FF	First Frame
FM	Figura de Mérito
FPGA	Field Programmable Gate Array
FS	Full Search
FSR	Filtered Super-Resolution
FWW	Frame of Working Window
GIST	Gwanju Institute of Science and Technology
GPIO	General Purpose Input-Output
GPS	Global Position System
GPU	Graphics Processing Unit
GRSR	GeoReferenced Super-Resolution
HD	High Definition
HDTV	High Definition TeleVision
HR	High Resolution

<i>Acrónimo</i>	<i>Significado</i>
HSI	Hue, Saturation, Intensity
HSM	Histogram Stretching Method
HTML5	HyperText Markup Language, version 5
IBP	Iterative Back Projection
IDS	Integrated Dynamic Solutions
INT	INterpolación
ITS	Institute for Telecommunication Science
IUMA	Instituto Universitario de Microelectrónica Aplicada
LF	Last Frame
LHR	Lower High Resolution
LQM	Logarithmic Quality Measure
LR	Low Resolution
LSI	Linearly Spatially Invariant
LSV	Linear Shift Varying
MAD	Mean Absolute Difference
MAGiC	MAtemáticas, Gráficas y Computación
MAP	Maximum A Posteriori
MB	Macro-Bloque / Macro-Block
MBS	Macro-Block Size
MC	Multi-Cámara / Multi-Camera
MCABMSR	Multi-Camera Adaptive Block Matching Super-Resolution
MCD	Máximo Común Divisor
ML	Maximum Likelihood
MOS	Mean Opinion Score
MPD	Maximum Pixel Difference
MPQM	Moving Pictures Quality Metric
MR	Medium Resolution
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
MSE	Mean Square Error
MWW	Mixed Working Window
NMF	Nonnegative Matrix Factorization
NQM	Noise Quality Measure
NTS	New Three Steps
NTSC	National Television System Committee
QCIF	Quarter Common Intermediate Format
PCA	Principal Component Analysis
PCI-X	Peripheral Component Interconnect eXtended
PCO	Pair COmparison
PFC	Proyecto Fin de Carrera
PLOCAN	PLataforma Oceánica de CANarias
POCS	Projection Onto Convex Sets
PSF	Point Spread Function
PSNR	Peak Signal-to-Noise Ratio
PWD	Pseudo-Wigner Distribution

<i>Acrónimo</i>	<i>Significado</i>
QCIF	Quarter Common Intermediate Format
RAM	Random Access Memory
RANSAC	RANdom SAMple Consensus
RGB	Red, Green, Blue
ROI	Region Of Interest
ROV	Remotely Operated Vehicle
RSR	Robust Super-Resolution
SA	Search Area
SC	SCale factor
SF	Selective Filter
SR	Súper-Resolución / Super-Resolution
SS	Single Stimulus
SSCQS	Simple Stimulus Continuous Quality Scale
SSIM	Structural SIMilarity
SWW	Spatial Working Window
T-E	Temporal-Espacial
THR	THReshold
TSS	Three Steps Search
TWW	Temporal Working Window
ULPGC	Universidad de Las Palmas de Gran Canaria
ULR	Upper Low Resolution
USB	Universal Serial Bus
UTM	Universal Transverse Mercator
VBS	Variable Block Size
VBSFSR	Variable Block Selective Filter Super-Resolution
VBSSR	Variable Block Size Super-Resolution
VBSTHR	Variable Block Size THReshold
VBSWSFSR	Variable Block Size & Window Selective Filter Super-Resolution
VHHR	Very High Resolution
VQEG	Video Quality Experts Group
VQM	Video Quality Metric
WBF	Window Back Frames
WBSF	Window & Block Selective Filter
WBSFSR	Window & Block Selective Filter Super-Resolution
WCF	Window Cutting Filter
WCFTHR	Window Cutting Filter THReshold
WCFWW	Window Cutting Filter Working Window
WFF	Window Forward Frames
WSF	Window Selective Filter
WSFSR	Window Selective Filter Super-Resolution
WSFTHR	Window Selective Filter THReshold
WSFWW	Window Selective Filter Working Window
WVBSFSR	Window & Variable Block Selective Filter Super-Resolution
WW	Working Window

Capítulo 1

INTRODUCCIÓN

Índice del capítulo

	<i>Página</i>
1.1.- Planteamiento del problema	2
1.2.- Motivación de la Tesis Doctoral	4
1.3.- Objetivos de la Tesis Doctoral	6
1.4.- Organización de la Tesis Doctoral	7

La creencia no es el principio, sino el fin de todo conocimiento
Johan Wolfgang von Goethe, poeta, novelista, dramaturgo y científico alemán (1749 – 1832)

1.1 PLANTEAMIENTO DEL PROBLEMA

En la mayoría de los sistemas electrónicos de captación de imágenes se requieren imágenes de alta resolución (HR, *High Resolution*) [EHK13]. Las imágenes de alta resolución resultan fundamentales en áreas relacionadas principalmente con dos bloques de aplicación: la mejora de información visual para la interpretación humana, y la ayuda a la representación para sistemas de percepción automática [Mil11]. La resolución de imagen es un término que describe su nivel de detalle, cuanto mayor sea la resolución, más detalles contendrá la imagen. La resolución de una imagen digital puede clasificarse atendiendo a diferentes criterios: resolución de píxel, resolución espacial, resolución espectral, resolución temporal y resolución radiométrica. En este contexto, la presente Tesis Doctoral se centra en mejorar la resolución espacial.

Una imagen digital está formada por elementos básicos denominados píxeles. La resolución espacial se refiere a la densidad de píxeles en una imagen y se mide en píxeles por unidad de área. La Figura 1.1 muestra la imagen de *test (test target) 1951 USAF*, usada para determinar la resolución espacial de sensores y sistemas de imagen.

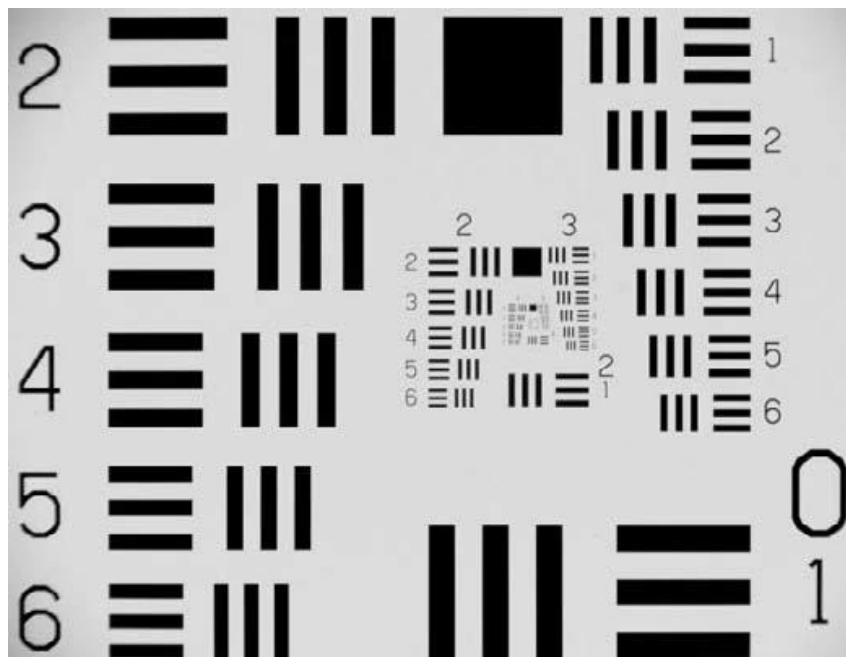


Figura 1.1.- Imagen de test (*test target*) 1951 USAF

Una elevada resolución implica que la densidad de píxel dentro de la imagen es elevada, por lo que una imagen de alta resolución puede ofrecer más detalles que una imagen de baja resolución (LR, *Low Resolution*). En el pasado, los únicos dispositivos de adquisición de imágenes disponibles eran cámaras analógicas de tipologías vidicón u orticón. Desde los años 70 se usan ampliamente los sensores de imagen CCD (*Charge Coupled Device*) y CMOS (*Complementary Metal Oxide Semiconductor*) para capturar imágenes digitales. Aunque estos sensores son adecuados para la mayoría de las aplicaciones de imagen, los niveles actuales de resolución y sus precios asociados no resultan apropiados para sistemas de bajo coste, siendo deseable disponer de niveles de alta resolución a precios tan bajos como sea posible. La demanda de imágenes de alta resolución, ha sido una de las principales motivaciones para encontrar metodologías que incrementen la resolución obtenida usando los dispositivos de adquisición de imagen disponibles actualmente.

La solución directa para incrementar el nivel de resolución podría consistir en reducir el tamaño de píxel en la fabricación del sensor, de forma que el número de píxeles por unidad de área aumente. La desventaja de esta solución es que la cantidad de luz disponible para cada píxel disminuye, lo que conlleva la generación de ruido de disparo o ruido *shot* que degrada significativamente la calidad de la imagen. Desafortunadamente, el tamaño de píxel no puede reducirse por debajo de cierto nivel ($40 \mu\text{m}^2$ para procesos CMOS de $0.35 \mu\text{m}$) para evitar ruido *shot*. Este nivel ya se ha alcanzado en el proceso de fabricación. Otra solución al problema del aumento del nivel de resolución es incrementar el tamaño del sensor manteniendo el tamaño del píxel. Esta aproximación implica un incremento en la capacitancia del sensor. Sin embargo, es bien sabido que una elevada capacitancia limita la velocidad de transferencia de carga, lo cual puede conllevar problemas en el proceso de formación de la imagen. De forma general, todas las soluciones *hardware* a este problema están limitadas por los costes de la óptica de alta precisión y los sensores de imagen requeridos [Mil11].

La manera más factible de afrontar este problema es integrar tanto las capacidades *hardware* como *software* para obtener el nivel de alta resolución requerido. Hacer uso de un nivel de resolución tan alto como sea posible a nivel *hardware*, completa parte de esta tarea. El resto se realiza a nivel *software*. Esta es la tendencia en la mayoría de los dispositivos de captura de imagen actuales. El proceso en el que, a partir de una única imagen, se obtiene otra imagen de mayor resolución, se conoce como interpolación [Ban09]; en este caso, al no proporcionarse información adicional, la calidad de la imagen estará muy limitada y las componentes de alta frecuencia no se podrán recuperar. Por otra parte, el uso de múltiples observaciones de una misma escena, para generar una imagen de mayor resolución se conoce como Súper-Resolución (SR, *Super-Resolution*). Existen otras técnicas de SR tomando otro tipo de información como punto de partida, si bien la planteada en la presente Tesis Doctoral es una de las más comunes, conociéndose generalmente en la literatura como SR por fusión [Mil11]. Las técnicas de SR por fusión incrementan las componentes de alta frecuencia y eliminan las degradaciones causadas en el proceso de captura de imagen de las cámaras. La idea básica consiste en combinar información no redundante contenida en múltiples *frames* de LR para generar una imagen de HR. El aporte de información se basa en los desplazamientos subpíxel que se pueden producir entre las imágenes, debido a movimientos incontrolados entre el sistema de imagen y la escena, movimientos de objetos dentro de la escena, o movimientos controlados, como sistemas de imagen por satélite, con una trayectoria y velocidad predeterminados. Entre otras, las técnicas de SR se aplican en áreas como:

- Vigilancia y seguridad [RRY+13, ZH12, LTB12]: Congelación de la imagen y ampliación de regiones de interés (ROI, *Region of Interest*) por ejemplo para lectura de placas de matrícula, o mejora de la resolución para reconocimiento automático de un objetivo, como en el caso del reconocimiento facial de un potencial criminal.
- Imagen submarina [YU14, CY13, CYX+12]: Mejora de resolución en entornos submarinos en secuencias captadas por vehículos tipo ROV (*Remotely Operated Vehicle*) y en señales acústicas.
- Detección remota [XZZ14, SKC+13, VCB+10]: A partir de varias imágenes en un área geográfica determinada, se busca la mejora de la resolución de la imagen.
- Imagen médica [OII+14, ABG+14, AEH+14, VCT+10]: En aplicaciones como tomografía computarizada (CT, *Computed Tomography*), resonancia magnética (MRI, *Magnetic Resonance Imaging*) o ultrasonidos, normalmente las imágenes presentan una resolución limitada, que puede mejorarse aplicando técnicas de SR.
- Conversión entre estándares de vídeo [Goh13, KSS+12, GSH+11], por ejemplo de señales de vídeo NTSC (*National Television System Committee*) a HDTV (*High Definition TeleVision*).

1.2 MOTIVACIÓN DE LA TESIS DOCTORAL

Las técnicas de SR comenzaron a plantearse hace ya 30 años partiendo del pionero trabajo de Tsai y Huang [TH84], en 1984. En este trabajo inicial se mostró de forma explícita que al menos en teoría era posible mejorar la resolución registrando y fusionando múltiples imágenes. En los años siguientes la potencia de cómputo se incrementó en varios órdenes de magnitud, las cámaras digitales fueron mejorando y los *displays* digitales presentaban cada vez más detalles. De la mano de este progreso, los requerimientos de una imagen de alta calidad se intensificaron de forma natural, quedando estos fuera del alcance de la tecnología del Estado del Arte. De hecho, en los últimos años, la calidad visual de las imágenes y vídeo capturados no ha estado alineada con estas expectativas. La integración de un número de píxeles cada vez mayor en espacios cada vez menores, usando elementos ópticos menos sofisticados, ha hecho que la calidad visual del contenido grabado disminuya. Así, a pesar de lo que se podría ver como una batalla perdida frente a sensores mejores y más baratos, las técnicas de SR se han convertido en un campo de aplicación de gran relevancia. Puesto que prácticamente todo el contenido visual grabado con cámaras digitales se post-procesa, resulta lógico plantear que la SR se terminará por asentar como la *killer application* en lo que a procesamiento de imagen se refiere [Mil11].

Aunque se han propuesto diversas aproximaciones diferentes desde que se introdujo el concepto de SR, la mayoría de ellas opera adecuadamente en imágenes de *test* y no tanto en situaciones reales. Para determinar de forma práctica comparativas adecuadas entre secuencias de vídeo o imágenes, se ha determinado una serie de métricas objetivas y subjetivas; entre éstas se utilizan en la presente Tesis Doctoral el PSNR (*Peak Signal-to-Noise Ratio*) y el índice SSIM (*Structural SIMilarity*), tal y como se justificará posteriormente. Por otra parte, a la hora de concebir un sistema de SR práctico, existen muchos retos para que éste pueda aplicarse en un amplio rango de aplicaciones. Los desafíos clave en el desarrollo y aplicación de las técnicas de SR han propiciado las aportaciones de la presente Tesis Doctoral (centradas en técnicas de SR por fusión), que se resumen en:

- **Eficiencia en el cómputo:** Las técnicas de SR normalmente requieren de una computación intensiva, dado el gran número de operaciones a realizar al operar frecuentemente manipulaciones de matrices. Las aplicaciones reales siempre demandan eficiencia en la reconstrucción por SR para resultar de utilidad práctica. Por ejemplo, en escenarios de vigilancia de vídeo es fundamental que la reconstrucción por SR se produzca en el menor tiempo posible. Los algoritmos de SR requieren además un registro preciso de la imagen, que es intensivo en cómputo. La presente Tesis Doctoral estudia cómo disminuir el cómputo en algoritmos de SR por fusión, partiendo de un algoritmo base BSR (*Baseline Super-Resolution*) y considerando un pre-procesamiento en la información de LR de entrada, determinando los *frames* más adecuados a considerar en el proceso de SR (Filtro Selectivo de Ventana), haciendo uso de la métrica objetiva de similitud estructural SSIM.
- **Registro de la imagen:** Este es un aspecto crítico para el éxito de la reconstrucción de SR de múltiples *frames*, donde se fusionan muestreos espaciales complementarios de la imagen de HR. El registro de la imagen es un problema fundamental en su procesamiento, siendo incluso más difícil en el caso concreto de las técnicas de SR, donde las observaciones son imágenes de LR, con importantes artefactos debidos al *aliasing*. Las prestaciones de los algoritmos estándar para registro de imagen se reducen a medida que la resolución de las observaciones disminuye, dando lugar a un mayor número de errores de registro. Los artefactos, causados por estos errores de

registro, son visualmente más molestos que el efecto de desenfoque resultante de la interpolación de una imagen. En la presente Tesis Doctoral se demuestra cómo el uso de una estructura basada en múltiples cámaras (*array* de cámaras o sistema Multi-Cámara, MC), puede contribuir de forma positiva en el problema planteado por la SR en relación con el registro de la imagen, ya que al disponer de múltiples observaciones de una misma secuencia, y considerando que el desplazamiento presente entre los *frames* capturados sea subpíxel, el resultado de la imagen de HR procesada mejorará las prestaciones, frente a una situación en la que se disponga de una única cámara.

- **Robustez:** Las técnicas tradicionales de SR son vulnerables a la presencia de anomalías debidas a errores de movimiento, modelos de enfoque imprecisos, objetos en movimiento, desenfoque por movimiento, etc. Estos modelos de error imprecisos no se pueden tratar como ruido gaussiano. La robustez del proceso de SR es de interés porque los parámetros de degradación de la imagen no se pueden estimar de forma precisa, y la sensibilidad a las anomalías puede resultar en artefactos visualmente molestos, intolerables en la mayoría de las aplicaciones, por ejemplo en conversión entre estándares de vídeo. Sin embargo, se considera que no se ha dedicado suficiente trabajo a un aspecto tan importante como éste. En esta Tesis Doctoral se introduce un filtro orientado a cada uno de los Macro-Bloques (MB) en los que se divide la secuencia de entrada para realizar el proceso de SR (Filtro Selectivo de MB). Se mostrará que un estudio particular para cada uno de los MBs proporciona una mayor robustez al algoritmo BSR usado como punto de partida, ya que al habilitar un tamaño variable de MB y determinar el grado de movimiento local de cada uno de ellos frente a una aproximación basada únicamente en movimiento global, la calidad global se incrementa. Además, se plantea un caso real orientado a secuencias de vídeo submarino grabadas por vehículos ROV, presentando cómo el cálculo de la velocidad media del vehículo, partiendo de información de georreferenciación, proporciona información clave a la hora de determinar el número de *frames* a seleccionar en el proceso de SR, lo que aporta un mayor grado de robustez al algoritmo.
- **Límites en las prestaciones:** La reconstrucción por SR ha sido una línea de investigación activa desde su planteamiento inicial, sin embargo, no se ha desarrollado mucho trabajo orientado hacia la comprensión fundamental de los límites en las prestaciones de los algoritmos de reconstrucción. La comprensión de este límite de prestaciones es importante. Por ejemplo, esto puede dar luz al diseño de cámaras de SR, ayudando a analizar factores tales como modelos de error, niveles de *zoom* y número de *frames* a combinar. Si bien es difícil perfilar conclusiones consistentes para diferentes técnicas de SR en términos de evaluación de sus prestaciones, es claro que se necesitan bancos de pruebas y conjuntos de secuencias reales y realistas que proporcionen una comparación justa y una comprensión adecuada de los algoritmos desarrollados. En el caso particular de la presente Tesis Doctoral, se han diseñado bancos de prueba específicos, ajustados para proporcionar información fidedigna de los resultados obtenidos. Al mismo tiempo, se han desarrollado sistemas que permiten determinar los máximos teóricos de calidad objetiva que se podrían alcanzar con los algoritmos propuestos.

1.3 OBJETIVOS DE LA TESIS DOCTORAL

La finalidad de esta Tesis Doctoral consiste en aportar soluciones algorítmicas de SR por fusión para mejorar la calidad y reducir el tiempo de cómputo, partiendo de un algoritmo de referencia (BSR, *Baseline Super-Resolution*), desarrollado en el seno de la división de Diseño de Sistemas Integrados (DSI) del Instituto Universitario de Microelectrónica Aplicada (IUMA) de la Universidad de Las Palmas de Gran Canaria (ULPGC). Para conseguir este propósito, esta Tesis Doctoral se centra en alcanzar los siguientes objetivos:

- Estudio del Estado del Arte de SR y en especial, de las posibles soluciones a considerar con respecto a los principales retos que plantea la SR. En este caso, resulta vital determinar las aproximaciones seguidas por otros autores con el fin de desarrollar una estrategia de trabajo adecuada.
- Determinación de las líneas de investigación más adecuadas, partiendo del estudio del Estado del Arte realizado, centrandó el interés en explotar la información espacial y temporal integrada en secuencias de vídeo, tomadas en sistemas que integren múltiples cámaras. En este contexto, los desafíos clave en el desarrollo y aplicación de las técnicas de SR enunciados en la sección anterior, influyen claramente en la toma de decisiones. Así, será fundamental disponer de la información adecuada, para optimizar la calidad subjetiva y objetiva obtenida a partir de los algoritmos desarrollados, estudiando cómo mejorar estos aspectos, desde el punto de vista del registro de la imagen y buscando además la eficiencia en el cómputo.
- Creación de un entorno de análisis que permita investigar en profundidad los procesos de los algoritmos estudiados, dando lugar a que se puedan vislumbrar conclusiones adecuadas a partir de los resultados obtenidos. Este aspecto resulta además básico para ser capaces de comparar la situación previa frente a la propuesta y poder así discernir los límites en las prestaciones del proceso de Súper-Resolución.
- Validación del correcto funcionamiento de las contribuciones propuestas en un entorno de aplicación específico para demostrar su robustez. En este caso se ha tomado como punto de partida, el entorno submarino, ya que como se comentó en la introducción, es un campo que se encuentra actualmente en investigación para la mejora de resolución en entornos submarinos en secuencias captadas por vehículos tipo ROV, así como por otros tipos de vídeos submarinos [YU14, CY13, CYX+12]. Con este objetivo se presentarán los principales aspectos a considerar en este entorno y cómo afectan a los parámetros de los algoritmos de SR.

1.4 ORGANIZACIÓN DE LA TESIS DOCTORAL

El trabajo desarrollado en esta Tesis Doctoral se ha estructurado en siete capítulos, de los cuales el primero de ellos lo constituye la presente introducción. El contenido del resto de los capítulos se describe a continuación:

- **Capítulo 2: Estado del Arte**

Este capítulo introduce el Estado del Arte sobre las técnicas de Súper-Resolución, plantea la relación de la Súper-Resolución con los sistemas Multi-Cámara, analiza las métricas de calidad utilizadas para determinar la aportación de los algoritmos estudiados, y establece las principales limitaciones del proceso de Súper-Resolución, así como las contribuciones aportadas en la presente Tesis Doctoral.

- **Capítulo 3: Contribuciones algorítmicas**

Una vez presentados los conceptos necesarios para comprender el ámbito de aplicación de la presente Tesis Doctoral, en este capítulo se introducen las aportaciones algorítmicas desarrolladas. La exposición del capítulo se desarrolla de forma evolutiva, es decir, de cómo a partir de la situación inicial se pasa a desarrollar nuevos algoritmos con el objetivo de mejorar la calidad de la imagen, disminuir el tiempo de cómputo y en definitiva construir soluciones eficientes.

- **Capítulo 4: Resultados y análisis: filtros selectivos y topología de Macro-Bloques adaptable**

En este capítulo, se presenta el entorno de trabajo utilizado para validar las aportaciones relacionadas con la primera contribución de esta Tesis Doctoral: los filtros selectivos y la topología de Macro-Bloques adaptable. Este entorno se utiliza para analizar los resultados más significativos. Se introducen y se analizan entonces los resultados de simulación, desglosando cada contribución particular.

- **Capítulo 5: Resultados y análisis: Entorno Multi-Cámara**

En este capítulo, se presentan los resultados y el análisis correspondientes a la segunda contribución de la presente Tesis Doctoral: entorno Multi-Cámara, incluyendo los diferentes bancos de pruebas diseñados y los principales resultados alcanzados.

- **Capítulo 6: Análisis del sistema global**

Una vez presentados los resultados de cada una de las contribuciones de la Tesis Doctoral por separado, se integran ambas en este capítulo con el objetivo de evaluar el funcionamiento global del sistema propuesto.

- **Capítulo 7: Conclusiones y Líneas Futuras**

Se presentan en este último capítulo las conclusiones extraídas a partir del trabajo desarrollado en esta Tesis Doctoral, así como las líneas de investigación a considerar.

- **Anexos**

Se incluyen cuatro anexos. El primero de ellos se centra en introducir las diferentes secuencias de vídeo usadas como referencia, el segundo presenta las características de las secuencias submarinas y de las técnicas de mejora de imagen asociadas, el tercero introduce un cuestionario utilizado para la medida de la calidad subjetiva, y finalmente el cuarto muestra el conjunto de publicaciones que se han elaborado en el transcurso del desarrollo de la Tesis Doctoral [C1-C12, R1-R7].

Capítulo 2

ESTADO DEL ARTE

Índice del capítulo

	<i>Página</i>
2.1.- Introducción	10
2.2.- Estado del Arte en técnicas de Súper-Resolución	11
2.3.- Estado del Arte de sistemas Multi-Cámara	22
2.4.- Métricas sobre calidad de imagen	37
2.5.- Desafíos de la Súper-Resolución y contribuciones	44
2.6.- Conclusiones	48

*Cada generación piensa que puede ser más inteligente que la anterior
Aldous Huxley, novelista, ensayista y poeta inglés (1894 – 1963)*

2.1 INTRODUCCIÓN

La aproximación de Súper-Resolución (SR) sobre la que se centran las aportaciones de la presente Tesis Doctoral se basa en la construcción de imágenes de HR a partir de varias observaciones de LR (técnicas de SR por fusión). Cada *frame* de LR es una versión diezmada y con *aliasing* de la escena original. La SR en este caso es posible sólo si existen movimientos subpíxel entre los *frames* de LR. En este capítulo se presentará, tanto esta aproximación, como un Estado del Arte general de las técnicas de Súper-Resolución introducidas a lo largo de los últimos años.

La Figura 2.1 muestra un diagrama simplificado en el que se describe la idea básica de la reconstrucción del proceso de SR usado en la presente Tesis Doctoral. En el proceso de captura de la imagen, la cámara obtiene varios *frames* de LR, lo que equivale a aplicar un submuestreo partiendo de la secuencia HR real con desplazamientos subpíxel entre una y otra imagen. El proceso de SR trata de invertir este proceso, alineando las observaciones con una precisión subpíxel y combinándolas en una rejilla de imagen de HR, superando así las limitaciones del sensor de la cámara.

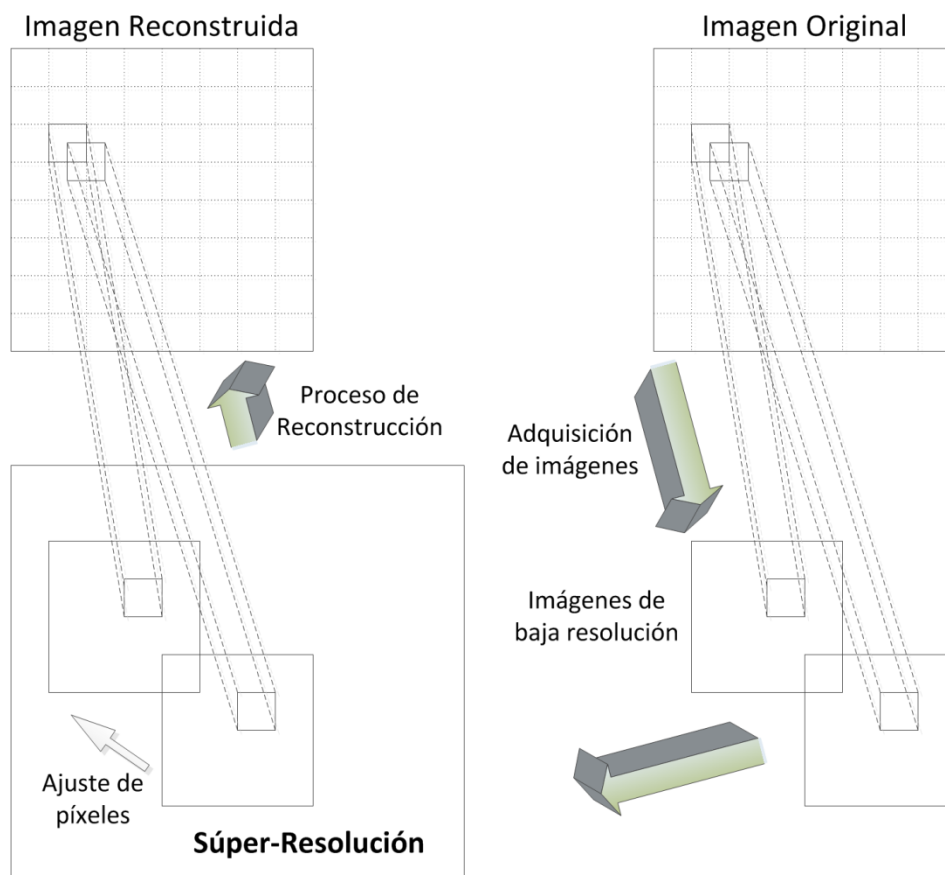


Figura 2.1.- Proceso de SR: reconstrucción de imagen HR a partir de varias imágenes en LR

El presente capítulo introduce un Estado del Arte de las técnicas de SR, plantea la relación de la SR con los sistemas Multi-Cámara (MC), analiza las métricas de calidad utilizadas para determinar las prestaciones de los algoritmos estudiados, y establece las principales limitaciones de la SR.

2.2 ESTADO DEL ARTE EN TÉCNICAS DE SÚPER-RESOLUCIÓN

2.2.1 Introducción

En cuanto a técnicas de procesamiento de imagen se refiere, la SR ha sido uno de los campos más activos desde el trascendental trabajo de Tsai y Huang [TH84] hace ya 30 años, en 1984. A partir de entonces se han propuesto diversas técnicas, presentando aproximaciones desde el dominio frecuencial al dominio espacial, y desde la perspectiva del procesamiento de la señal hasta la perspectiva del aprendizaje automático.

Los primeros trabajos en SR siguieron principalmente la teoría de Tsai y Huang, explorando las propiedades de desplazamiento y *aliasing* de la Transformada de Fourier. Sin embargo, estas aproximaciones en el dominio de la frecuencia estaban muy restringidas en el modelo de observación que podían adoptar, encontrándose que los problemas reales eran mucho más complicados con respecto al modelo inicialmente propuesto. Hoy en día se aborda el problema generalmente en el dominio espacial, debido a su flexibilidad al modelar todos los tipos de degradaciones de la imagen [BS98, PPK03, EFR+04, BK02].

En esta sección se tratan las diferentes técnicas de Súper-Resolución, comenzando por el modelo de observación.

2.2.2 Modelo de observación

El sistema digital de adquisición de la imagen no es perfecto debido fundamentalmente a limitaciones *hardware*, adquiriendo imágenes con varios tipos de degradación. Así, el tamaño de apertura finito causa el efecto de desenfoque óptico (*optical blur*), modelado por la función PSF (*Point Spread Function*). El tiempo de apertura finito resulta en un desenfoque por movimiento (*motion blur*), el cual es muy común en secuencias de vídeo. El tamaño finito del sensor provoca desenfoque debido al sensor (*sensor blur*). La limitada densidad del sensor lleva a efectos de *aliasing*, constriñendo la resolución espacial de la imagen obtenida. Estas degradaciones se modelan parcial o completamente en las diferentes técnicas de SR.

La Figura 2.2 muestra un modelo de observación típico que relaciona la imagen de HR con los *frames* de vídeo de LR [PPK03, EFR+04]. La entrada al sistema de imagen se basa en escenas naturales continuas, aproximadas como señales limitadas en banda. Estas señales pueden estar contaminadas por turbulencia atmosférica antes de alcanzar el sistema de imagen. Al muestrear la señal cumpliendo el criterio de Nyquist, se genera una imagen digital de alta resolución (Figura 2.2.a). En el esquema planteado normalmente existe algún tipo de movimiento entre la cámara y la escena a capturar. Las entradas a la cámara se considera que representan múltiples *frames* de la escena, con movimiento local o global entre ellos, tal y como se presenta en la Figura 2.2.b. A través de la cámara, estos *frames* de alta resolución incurrirán en diferentes tipos de desenfoque. Estas imágenes desenfocadas (Figura 2.2.c) son entonces submuestreadas en los sensores de imagen, por ejemplo detectores CCD (*Charged Coupled Device*), o CMOS (*Complementary Metal Oxide Semiconductor*) en píxeles, mediante la integral de la imagen que presenta el área de cada sensor. A las imágenes submuestreadas les afecta además el ruido del sensor y el ruido del filtro de color. Por tanto, los *frames* capturados por el sistema de imagen de baja resolución se corresponden con versiones desenfocadas, diezmadas y ruidosas de la secuencia original.

Así, si se denomina \mathbf{X} a la imagen de HR deseada, esto es, la imagen digital muestreada cumpliendo el criterio de muestreo de Nyquist desde la secuencia limitada en banda, y sea \mathbf{Y}_k la observación del *frame* k de la cámara, asumiendo que la cámara captura K *frames* de \mathbf{X} , las observaciones de LR están relacionadas con la secuencia HR \mathbf{X} por la ecuación (2.1).

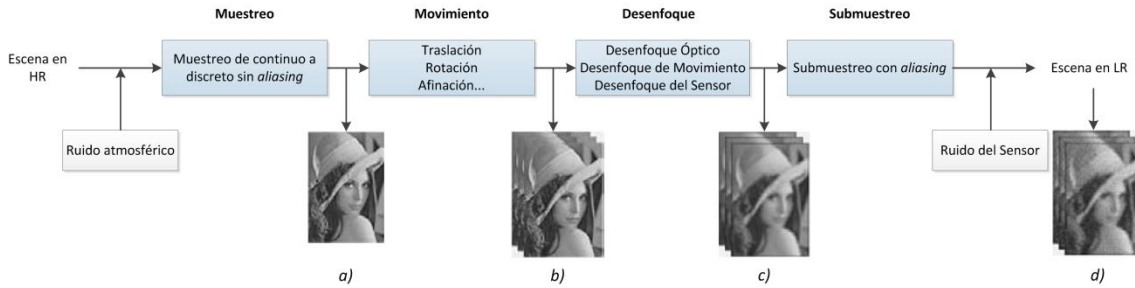


Figura 2.2.- Modelo de observación de un sistema de imagen

$$\mathbf{Y}_k = D_k H_k F_k \mathbf{X} + N_k, k = 1, 2, \dots, K, \quad (2.1)$$

En (2.1) D_k es el operador de submuestreo, H_k modela los efectos de desenfoque, F_k codifica la información de movimiento del *frame* k , y N_k es un término relacionado con el ruido. Estas ecuaciones lineales pueden ser reorganizadas en un sistema de ecuaciones, tal y como se representa en (2.2).

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \cdot \\ \cdot \\ Y_K \end{bmatrix} = \begin{bmatrix} D_1 H_1 F_1 \\ D_2 H_2 F_2 \\ \cdot \\ \cdot \\ D_K H_K F_K \end{bmatrix} \mathbf{X} + \vec{N} \quad (2.2)$$

O de forma equivalente en la expresión (2.3):

$$\vec{Y} = M \mathbf{X} + \vec{N} \quad (2.3)$$

Las matrices implicadas D_k , H_k , F_k o M son muy dispersas, y este sistema lineal está típicamente mal planteado (*ill-posed*), lo cual quiere decir que no cumple alguna de las condiciones de sistemas bien planteados (*well-posed*) propuestas por Hadamard [Had23]:

- Existe una solución.
- La solución es única.
- El comportamiento de la solución varía continuamente con las condiciones iniciales.

Además, en sistemas de imagen reales, estas matrices son desconocidas y necesitan estimarse a partir de las observaciones de LR disponibles, dejando el problema incluso peor condicionado. De esta forma, una regularización previa adecuada para la imagen de HR es siempre deseable y por lo general imprescindible. A continuación se introducirán algunas técnicas de SR básicas propuestas en la literatura y se presentará una visión general de desarrollos recientes.

2.2.3 Súper-Resolución en el dominio de la frecuencia

El trabajo pionero de Súper-Resolución de Tsai y Huang [TH84] relacionó una imagen de HR con varias imágenes de LR en una formulación en el dominio de la frecuencia basada en las propiedades de desplazamiento y *aliasing* de las transformadas de Fourier discreta y continua. Sean t_1 y t_2 las coordenadas horizontal y vertical de una imagen, si se denomina $x(t_1, t_2)$ a una secuencia continua de HR y a las traslaciones globales de K imágenes desplazadas, $x_k(t_1, t_2) = x(t_1 + \Delta_{k1}, t_2 + \Delta_{k2})$, con $k = 1, 2, \dots, K$, donde Δ_{k1} y Δ_{k2} son desplazamientos arbitrarios, pero conocidos, la Transformada Continua de Fourier (CFT, *Continuous Fourier Transform*) de la secuencia vendrá dada por $\chi(u_1, u_2)$ y las transformadas relativas a las secuencias desplazadas por imágenes por $\chi_k(u_1, u_2)$. Entonces, de acuerdo con sus propiedades de desplazamiento, la CFT de las imágenes desplazadas se puede escribir como:

$$\chi_k(u_1, u_2) = e^{j2\pi(\Delta_{k1}u_1 + \Delta_{k2}u_2)} \cdot \chi(u_1, u_2), \quad (2.4)$$

donde u_1 y u_2 son las coordenadas en el dominio frecuencial. Las imágenes desplazadas son muestreadas con períodos de muestreo T_1 y T_2 para obtener la imagen de LR observada $y_k(n_1, n_2) = x_k(n_1T_1 + \Delta_{k1}, n_2T_2 + \Delta_{k2})$, con $n_1 = 1, 2, \dots, N_1 - 1$ y $n_2 = 1, 2, \dots, N_2 - 1$. Las Transformadas de Fourier Discretas (DFT, *Discrete Fourier Transform*) de estas imágenes de LR se denotan por $Y_k[r_1, r_2]$, donde r_1 y r_2 son las coordenadas horizontal y vertical de la DFT, respectivamente. Las CFTs están relacionadas con sus DFTs por el *aliasing*.

$$Y_k(r_1, r_2) = \frac{1}{T_1 T_2} \sum_{m_1=-\infty}^{\infty} \sum_{m_2=-\infty}^{\infty} \chi_k \left(\frac{2\pi}{T_1} \left(\frac{r_1}{N_1} - m_1 \right), \frac{2\pi}{T_2} \left(\frac{r_2}{N_2} - m_2 \right) \right) \quad (2.5)$$

Asumiendo que $\chi_k(u_1, u_2)$ está limitada en banda, $|\chi_k(u_1, u_2)| = 0$ para $|u_1| \geq (N_1 \pi)/T_1$, $|u_2| \geq (N_2 \pi)/T_2$, y combinando las ecuaciones (2.4) y (2.5) es posible relacionar los coeficientes DFT de $Y_k[r_1, r_2]$ con las muestras de la CFT desconocida de $x(t_1, t_2)$ según la expresión (2.6):

$$\vec{Y} = \varphi \vec{\chi}, \quad (2.6)$$

donde \vec{Y} es un vector columna de tamaño $K \times 1$ con el elemento de orden k representando el coeficiente $Y_k[r_1, r_2]$, $\vec{\chi}$ es un vector columna de tamaño $N_1 N_2 \times 1$ que contiene las muestras de los coeficientes CFT desconocidos de $x(t_1, t_2)$, y φ es una matriz $K \times N_1 N_2$ que relaciona \vec{Y} y $\vec{\chi}$. La ecuación 2.6 define un conjunto de ecuaciones lineales para resolver $\vec{\chi}$ y usar entonces la DFT inversa con la finalidad de obtener la imagen reconstruida. Esta formulación de SR asume un modelo de traslación con parámetros conocidos, así como que no existe ruido. El proceso de submuestreo se asume que está basado en la respuesta al impulso, sin modelar efectos de desenfoque del sensor. Basándose en esta línea de trabajo se han propuesto varias extensiones como modelos de observación más complejos. Kim *et al.* [KBV90] extendieron el modelo inicial de Tsai y Huang [TH84] considerando tanto el ruido de observación como el desenfoque espacial. Su trabajo posterior en [KBV93] fue más allá, introduciendo la regularización de Tikohonov [TA97]. Su y Kim [SK94] consideraron un modelo de movimiento local dividiendo las imágenes en bloques solapados y estimando movimientos para cada bloque de forma individual. Tom *et al.* [TKG94] propusieron que la restauración y la estimación de movimiento se realizaran simultáneamente. Sin embargo, la SR en el dominio de la frecuencia no fue más allá. Estas aproximaciones son computacionalmente eficientes, pero limitadas para integrar modelos de degradación más complejos, así como para incluir una regularización de la imagen adecuada y estimar movimiento local. Los trabajos posteriores en SR han continuado de forma casi exclusiva en el dominio espacial [BS98, PPK03, EFR+04, BK02].

2.2.4 Aproximaciones no iterativas. Restauración por interpolación

La imagen de HR y los *frames* de LR están relacionados por un sistema lineal disperso, de forma similar al problema de restauración de una imagen simple, como se describe en el trabajo de Elad y Feuer [EF97]. Por ello, se pueden aplicar múltiples estimadores flexibles a la SR por reconstrucción. Estos incluyen *Maximum Likelihood* (ML), *Maximum a Posteriori* (MAP) [SS96], y *Projection Onto Convex Sets* (POCS) [SO89]. En esta sección se comenzará con el modelo de reconstrucción más simple en el dominio espacial, tal y como se hizo en el dominio de la frecuencia.

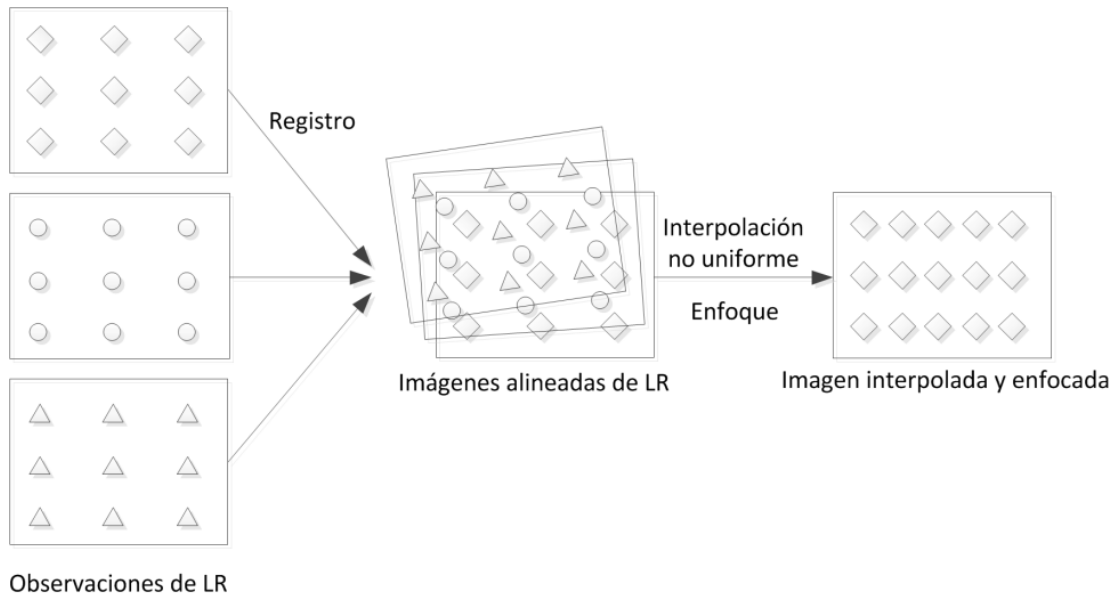


Figura 2.3.- SR basada en alineamiento y pos-procesamiento del desenfoque

Asumiendo que H_k es un sistema lineal espacial invariante (LSI, *Linearly Spatially Invariant*) y que es el mismo para todos los K *frames*, denotándolo como H , si se supone que F_k sólo considera modelos de movimiento simple como la traslación y la rotación, entonces H y F_k conmutan en (2.1) y se obtiene la expresión (2.7):

$$Y_k = D_k F_k H X + N_k = D_k F_k Z, \quad k = 1, 2, \dots, K, \quad (2.7)$$

lo cual lleva a una aproximación no iterativa basada en interpolación y restauración. Existen tres etapas para esta aproximación:

1. Registro de la imagen de LR.
2. Interpolación no uniforme para obtener Z .
3. Eliminación de desenfoque y ruido para obtener X .

La Figura 2.3 muestra el procedimiento de esta aproximación. Los *frames* de LR se alinean en primer lugar haciendo uso de un algoritmo de registro hasta llegar a precisión subpíxel. Estos *frames* alineados de LR se ponen entonces en una rejilla de imagen de HR, donde se usan métodos de interpolación no uniformes para rellenar los píxeles perdidos en la rejilla de HR y obtener Z . Finalmente, Z se enfoca con un algoritmo de deconvolución que incluye eliminación de ruido para obtener X . En la bibliografía se encuentran diferentes propuestas relacionadas con esta aproximación:

- Keren *et al.* [KPB88] propusieron una aproximación de reconstrucción por SR de dos pasos, basado en un modelo de traslación y rotación global.
- Ur *et al.* [UG92] propusieron una interpolación no uniforme de un conjunto de imágenes de LR espacialmente desplazadas utilizando el teorema de muestreo multicanal generalizado por Yen [Yen56], y después Papulis [Pap77], aplicando posteriormente un filtro de enfoque (*deblurring*).
- Nguyen y Milanfar [NM00] propusieron un algoritmo de reconstrucción de SR basado en una interpolación tipo *wavelet* haciendo uso de la estructura de muestreo entrelazada de la información de LR. Destacan en este campo además los trabajos presentados por S.E.El-Khamy *et al.* [EHD+05, EHD+06, EHD+08a, EHD+08b] en una implementación basada en cuatro pasos: alineamiento de la imagen, restauración multicanal de la imagen, fusión basada en *wavelet*, e interpolación de la imagen resultante para obtener una imagen de HR.
- Alam *et al.* [ABH+00] presentaron un esquema de interpolación eficiente basado en el peso de los vecinos más cercanos (*weighted nearest neighbors*), seguido por un filtro de Wiener para enfocar.
- Considerando la reconstrucción por SR, donde la observación está compuesta por traslación pura, desenfoque espacialmente invariante y ruido gaussiano aditivo, Elad y Hel-Or [EH01] presentaron un algoritmo computacionalmente muy eficiente.
- Lertrattanapanich y Bost [LB02] propusieron un método basado en triangulación para interpolar información muestreada irregularmente. Sin embargo, este método no es robusto al ruido presente de forma común en aplicaciones reales.
- Basándose en la convolución normalizada de Knutsson y Westin [KW93], Pham *et al.* [PVS06] propusieron una estructura adaptativa para la fusión de información muestreada irregularmente.
- Takeda *et al.* [TKM07] propusieron un sistema adaptativo basado en regresión para la interpolación de la rejilla de imagen de HR, donde las imágenes de HR se registran y se mapean.

Estas aproximaciones basadas en restauración por interpolación son intuitivas, simples, y computacionalmente eficientes, asumiendo modelos de observación relativamente sencillos. Sin embargo, una aproximación paso a paso no garantiza una estimación óptima. Un error de registro se puede propagar fácilmente al procesamiento posterior. Además, el paso de interpolación queda por debajo del óptimo si no se considera el ruido y los efectos de desenfoque, y sin una regularización adecuada las aproximaciones basadas en interpolación requieren de un tratamiento especial para poder reducir el *aliasing*.

2.2.5 Aproximaciones estadísticas

A diferencia de las aproximaciones basadas en restauración por interpolación, las aproximaciones estadísticas relacionan los pasos de reconstrucción por SR estocásticamente hacia una reconstrucción óptima. La imagen de HR y los movimientos entre las entradas de LR se pueden considerar entonces variables estocásticas. Sea $M(v,h)$ la matriz de degradación definida por el vector de movimiento v y el núcleo de desenfoque h , la reconstrucción por SR puede llevarse a un marco bayesiano de acuerdo con la ecuación (2.8).

$$\begin{aligned}
 \mathbf{X} &= \arg \max_{\mathbf{X}} Pr(\mathbf{X}|\vec{\mathbf{Y}}) = \arg \max_{\mathbf{X}} \int_{v,h} Pr(\mathbf{X}, M(v, h)|\vec{\mathbf{Y}}) dv \\
 &= \arg \max_{\mathbf{X}} \int_{v,h} \frac{Pr(\vec{\mathbf{Y}}|\mathbf{X}, M(v, h)) Pr(\mathbf{X}, M(v, h))}{Pr(\vec{\mathbf{Y}})} dv \\
 &= \arg \max_{\mathbf{X}} \int_{v,h} Pr(\vec{\mathbf{Y}}|\mathbf{X}, M(v, h)) Pr(\mathbf{X}) Pr(M(v, h)) dv
 \end{aligned} \tag{2.8}$$

Se ha de considerar que \mathbf{X} y $M(v, h)$ son estadísticamente independientes, como introdujeron Hardie *et al.* en [HBA97]. En (2.8), $Pr(\vec{\mathbf{Y}}|\mathbf{X}, M(v, h))$ es la probabilidad de la información, $Pr(\mathbf{X})$ es el término previo en la imagen de HR deseada, y $Pr(M(v, h))$ es un término previo en la estimación de movimiento. $\vec{\mathbf{N}}$ en la ecuación (2.3) normalmente se identifica con ruido aditivo, asumiendo que es de media nula y un vector aleatorio de ruido blanco gaussiano. Por tanto,

$$Pr(\vec{\mathbf{Y}}|\mathbf{X}, M(v, h)) \propto e^{-\frac{1}{2\sigma^2} \|\vec{\mathbf{Y}} - M(v, h)\mathbf{X}\|^2}, \tag{2.9}$$

donde σ se identifica con la desviación estándar. Por otra parte $Pr(\mathbf{X})$ se define típicamente usando una distribución de Gibbs en forma exponencial:

$$Pr(\mathbf{X}) = \frac{1}{Z} e^{-\alpha A(\mathbf{X})}, \tag{2.10}$$

donde α balancea la consistencia de información y la fuerza de la imagen de HR, $A(\mathbf{X})$ es una función potencia no negativa, y Z es un factor de normalización. La formulación bayesiana en la ecuación 2.8 es complicada y difícil de evaluar debido a la integración sobre la estimación de movimiento. Si $M(v, h)$ se da o se estima de antemano (redefinido como M) la ecuación 2.8 se puede simplificar como:

$$\begin{aligned}
 \mathbf{X} &= \arg \max_{\mathbf{X}} Pr(\vec{\mathbf{Y}}|\mathbf{X}, M) Pr(\mathbf{X}) \\
 &= \arg \min_{\mathbf{X}} \{ \|\vec{\mathbf{Y}} - M\mathbf{X}\|^2 + \lambda A(\mathbf{X}) \},
 \end{aligned} \tag{2.11}$$

donde λ absorbe la varianza del ruido y α en la ecuación 2.10, balancea la consistencia de información y la fuerza de la imagen HR. La ecuación 2.11 es la formulación de *Maximum a Posteriori* (MAP) para SR, donde se asume que M es conocida. Las diferentes aproximaciones estadísticas varían en las formas de tratar la matriz de degradación $M(v, h)$, el término $Pr(\mathbf{X})$, y los métodos de inferencia estadística de la ecuación 2.8. Estas aproximaciones son:

- *Maximum Likelihood* (ML).
- *Maximum a Posteriori* (MAP).
- Restauración MAP conjunta.
- Tratamientos bayesianos.

2.2.6 Aproximaciones basadas en patrones

Las aproximaciones previas de SR se basan en agregar múltiples *frames* que contienen información espacial complementaria. La regularización en estos escenarios llega a ser especialmente crucial cuando se proporciona un número insuficiente de *frames*, siendo el caso extremo cuando se proporciona un único *frame* de LR [BK02]. Una metodología que ha surgido

recientemente para regularizar la reconstrucción por SR mal planteada es usar patrones para romper el límite de la SR causado por medidas inadecuadas. A diferencia de las aproximaciones previas, donde la información a *priori* se toma de forma paramétrica regularizando la imagen completa, los métodos basados en patrones desarrollan la información a priori muestreando otras imágenes, de manera similar a las desarrolladas por Efron y Leung [EL99] o Hertzmann *et al.* [HJO+01] de forma local. Una rama de aproximaciones basadas en patrones usa directamente los patrones, como en el trabajo propuesto por Freeman *et al.* [FJP02]. Tales aproximaciones normalmente operan manteniendo dos conjuntos de bloques de entrenamiento de n patrones $\{x_i\}_{i=1}^n$ muestreados desde las imágenes de HR, e $\{y_i\}_{i=1}^n$ muestreados desde las imágenes de LR. Cada pareja de bloques (x_i, y_i) está conectada por el modelo de observación $y_i = DHx_i + n$. Este modelo concurrente de HR y LR se aplica a una imagen objetivo para predecir la imagen de HR usando un método basado en bloques (*patches*), con un modelo aleatorio de Markov (MRF, *Markov Random Field*) como se muestra en la Figura 2.4. Los parámetros del modelo de observación han de conocerse a priori, y los conjuntos de entrenamientos deben poderse acoplar con la imagen objetivo. El tamaño del bloque debería elegirse adecuadamente. Si el tamaño del bloque es muy pequeño, la concurrencia será muy débil como para realizar una predicción significativa. Por otra parte, si el tamaño del bloque es muy grande, debería haber un conjunto significativo de entrenamiento como para poder encontrar bloques de proximidad para las observaciones.

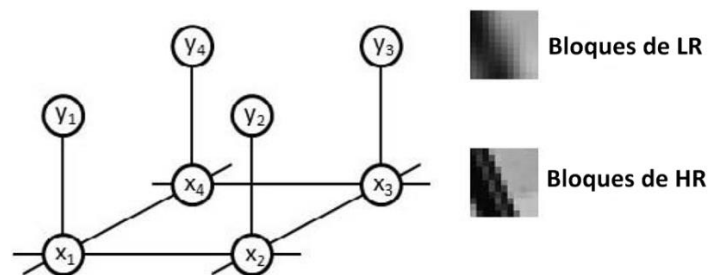


Figura 2.4.- El modelo MRF para SR de un único frame

Una forma inexperta de aplicar SR con este tipo de conjuntos de bloques es, para cada bloque de LR en la imagen de LR, buscar su vecino más cercano \tilde{y} en $\{y_i\}_{i=1}^n$, y entonces ubicar la correspondiente \tilde{x} de $\{x_i\}_{i=1}^n$ en la rejilla de imagen de HR. Desafortunadamente, esta aproximación simple produce artefactos debido al ruido y al mal planteamiento de la naturaleza de la SR, como ocurre en la propuesta de Elad y Datsenko [ED07]. Si la búsqueda del vecino más cercano se extiende, esto puede asegurar que se incluya el camino de proximidad del bloque deseado. Freeman *et al.* [FJP02] propusieron un algoritmo de propagación [YFW01] basado en el modelo MRF para seleccionar el mejor bloque de HR que presente mayor compatibilidad con los bloques adyacentes. Sun *et al.* [SZT+03] extendieron esta idea para usar la información a priori con el fin de mejorar sólo los perfiles de la imagen, y acelerar así el algoritmo. El algoritmo IBP (*Iterative Back Projection*) de Irani y Peleg [IP90] se aplica entonces como un paso de post-procesamiento para asegurar consistencia de información en la imagen completa. Wang *et al.* [WTS05] siguieron esta línea de trabajo y propusieron un modelo estadístico para trabajar con un PSF desconocido.

Los métodos explicados se basan en bloques de imagen, requiriendo grandes conjuntos de entrenamiento para incluir cualquier patrón. Chang *et al.* [CYX04] propusieron otro método simple pero efectivo basado en la inclusión de vecinos (*neighbour embedding*) formulada por Tenenbaum *et al.* [TSL00], con la suposición de correspondencia entre los bloques de imagen de LR y HR. Para cada bloque de LR y_k^t de la imagen de *test* (el superíndice "t" distingue el bloque de *test* del conjunto de los bloques de *test*), el algoritmo encuentra los k vecinos N_i de $\{y_i\}_{i=1}^n$, y calcula los pesos de reconstrucción por inclusión de vecinos:

$$\hat{w}_s = \arg \min_{w_s} \|y_k^t - \sum_{y_s \in N_t} w_s y_s\|^2, \text{ sujeto a } \sum_{y_s \in N_t} w_s = 1 \quad (2.12)$$

Los pesos de reconstrucción se aplican entonces para generar el bloque de HR correspondiente $\hat{x}_k^t = \sum_{y_s \in N_t} \hat{w}_s x_s$. Para tratar el problema de compatibilidad entre bloques adyacentes, se realiza un promediado simple entre las regiones solapadas. El algoritmo funciona incluso con bloques más pequeños [YFW01]. Sin embargo, fijar k para cada bloque de LR puede resultar en un sobreajuste o en un subajuste. Yang *et al.* [YWH+08] propusieron otro método de SR de *frame* simple basado en patrones. El método deriva de la teoría de *compressive sensing*, que asegura que las relaciones lineales entre señales de HR pueden recuperarse de forma precisa a partir de sus proyecciones de baja resolución como propusieron Candes [Can06] y Donoho [Don06a]. El algoritmo modela los conjuntos de entrenamiento como dos diccionarios: $D_h = [x_1, x_2, \dots, x_n]$ y $D_l = [y_1, y_2, \dots, y_n]$. Dado un bloque de imagen de *test* de LR y_k^t , la aproximación busca básicamente los apoyos mediante una minimización l_1 [Don06b].

$$\hat{w} = \arg \min_w \|w\|_{l_1}, \text{ sujeto a } \|y_t - D_l w\|^2 \leq \sigma^2, \quad (2.13)$$

donde σ es la desviación estándar. La ecuación (2.13) puede describirse con un multiplicador de Lagrange como un problema de optimización sin restricciones conocido como *Lasso* en la literatura de la estadística tal y como introdujo Tibshirani en [Tib96]. El correspondiente bloque de HR se recupera como $x_k^t = D_h \hat{w}$. Comparado con el método de inclusión de vecinos con k vecinos, el método de Yang elige adaptativamente los pocos apoyos necesarios para la reconstrucción, evitando de esta manera el sobreajuste. Además, la formulación por minimización l_1 es más robusta al ruido que los métodos previos basados en patrones. En una versión posterior de Huang *et al.* [HYW+10], el caso previo se extendió usando un diccionario de parejas en lugar de bloques simples, permitiendo al algoritmo ser mucho más eficiente.

Una crítica a los mencionados métodos con patrones directos es que operar en los patrones locales no garantiza la optimización global de la estimación. Otro tipo de aproximación basada en patrones busca realizar una estimación MAP con información local a priori. El trabajo pionero de Baker y Kanade [BK02] formuló una regularización explícita que demanda proximidad entre los derivados espaciales de la imagen desconocida y los ejemplos encontrados. Los patrones se forman haciendo uso de una pirámide de conjuntos de características en lugar de usar la información cruda directamente. Pickup *et al.* [PRZ03] propusieron un método similar para la SR de textos. Datsenko y Elad [DE07] presentaron una estimación global MAP donde la regularización basada en patrones se realiza mediante un promedio de pesos binarios en lugar de usar una aproximación al vecino más cercano, evitando valores extremos debidos al ruido. Otra aproximación de SR basada en patrones digna de destacar es la propuesta por Protter *et al.* [PET+09], generalizada a partir de un algoritmo de eliminación de ruido introducido por Buades *et al.* [BCM05]. En lugar de muestrear ejemplos a partir de otras imágenes de entrenamiento, el algoritmo explora autosimilitudes dentro de la imagen (o secuencia) y extrae los patrones de ejemplo de la imagen objetivo (o secuencia).

El uso de patrones puede ser mucho más efectivo cuando se trata con conjuntos reducidos de imágenes, como ocurre con imágenes textuales o faciales. En los últimos años ha surgido un grupo de algoritmos que toman como objetivo la imagen facial debido a su importancia en escenarios de videovigilancia. La SR facial es conocida habitualmente como “alucinación facial”, tomando como referencia el trabajo inicial de Baker y Kanade [BK02]. Capel y Zissersman [CZ01] propusieron un algoritmo donde los modelos del subespacio PCA (*Principal Component Analysis*) introducido por Jolliffe [Jol02] se usan para aprender partes de

las caras. Liu *et al.* [LSF07, LSZ01] propusieron una aproximación en dos pasos para la SR facial, donde el primer paso usa el *Eigenface*, introducido por Turk y Pentland en [TP91], para generar una cara de media resolución, seguida por una aproximación no paramétrica basada en patrones en el segundo paso [FJP02]. Yang *et al.* [YTM+08] propusieron una aproximación similar en dos pasos. En lugar de usar el subespacio PCA, [YTM+08] usa localmente una Factorización de Matriz No-negativa (NMF, *Nonnegative Matrix Factorization*) propuesta por Lee y Seung [LS99] para modelar las caras y el modelo basado en patrones del segundo paso se adopta de [YWH+08]. Jia y Gong [JG05, JG08] propusieron la aproximación de tensor para tratar con más variaciones faciales, tal como iluminaciones y expresiones. Aunque estos algoritmos de alucinación facial trabajan sorprendentemente bien, sólo se aplican en caras frontales, y únicamente algunos trabajos se han basado en evaluar la alucinación facial para reconocimiento, como los propuestos por Gunturk *et al.* [GBA+03] o Hennings-Yeomans *et al.* [HBK08]. La regularización basada en ejemplos resulta efectiva en la SR cuando existen observaciones insuficientes. Existen aún varias cuestiones que han de responderse con respecto a este tipo de aproximaciones:

- Primero, cómo elegir el tamaño de bloque óptimo dada una imagen objetivo. Quizás se requiere un tratamiento de multi-resolución.
- En segundo lugar, cómo elegir la base de datos. Imágenes diferentes presentan estadísticas distintas, y por lo tanto necesitan bases de datos diferentes. Se puede obtener una solución mediante un método eficiente de adaptación a partir de un diccionario objetivo.
- En tercer lugar, cómo usar la información *a priori* basada en ejemplos de una forma más eficiente. El aspecto relativo a la computación puede presentar una dificultad en aplicaciones prácticas. Elad y Datsenko [ED07] realizan un análisis detallado de regularización basada en ejemplos para problemas inversos.

2.2.7 Projection Onto Convex Sets (POCS)

Además de las aproximaciones de optimización derivadas de la visión estocástica discutida en la sección anterior, otro conjunto de métodos se basan en un algoritmo bien conocido como es POCS (*Projection Onto Convex Sets*) que parten normalmente de la versión utilizada por Youla y Webb [YW82]. Los métodos POCS aproximan el problema de SR formulando múltiples conjuntos convexos que contienen la imagen deseada como un punto dentro de los conjuntos. La definición de los conjuntos convexos es flexible y puede incorporar diferentes tipos de restricciones o información *a priori*, incluyendo restricciones no lineales y no paramétricas. Como ejemplo, se introducen varios conjuntos convexos usados de forma habitual en los métodos POCS. La consistencia de la información o las restricciones de la reconstrucción pueden modelarse como K conjuntos convexos:

$$C_k = \{ \mathbf{X} \mid \| D_k H_k F_k \mathbf{X} - \mathbf{Y}_k \|^2 \leq \sigma^2, 1 \leq k \leq K \}, \quad (2.14)$$

donde σ es la desviación estándar. Las restricciones de uniformidad pueden definirse como:

$$C_r = \{ \mathbf{X} \mid \| \Gamma \mathbf{X} \|_p < \sigma^2 \}, \quad (2.15)$$

donde $p = 1, 2, \dots, \infty$ hace referencia a diferentes normas. Las restricciones de amplitud también se pueden modelar de acuerdo con la expresión (2.16).

$$C_A = \{ \mathbf{X} \mid A_1 \leq \mathbf{X}[m, n] \leq A_2 \} \quad (2.16)$$

Con un grupo de M conjuntos convexos, la solución deseada se basa en la intersección de estos conjuntos $\mathbf{X} \in C_s = \bigcap_{i=1}^M C_i$. POCS sugiere el siguiente algoritmo recursivo para encontrar un punto dentro del conjunto de la intersección a partir de la siguiente aproximación:

$$\mathbf{X}_{k+1} = P_M P_{M-1} \dots P_2 P_1 \mathbf{X}_k, \quad (2.17)$$

donde \mathbf{X}_0 es una predicción inicial, y P_i es el operador de proyección que proyecta un punto dentro de un conjunto convexo cerrado C_i .

Las primeras técnicas POCS para reconstrucción por SR fueron propuestas por Stark y Oskoui [SO89]. Patti *et al.* [PST94, PST97a, PST97b] propusieron extensiones para tratar con espacios variantes PSF, desenfoque por movimiento, desenfoque por sensor, y efectos de muestreo por *aliasing*. Muchos trabajos de SR sólo consideran un tamaño de apertura no nulo (el desenfoque de la lente, PSF), pero no un tiempo de apertura finito (desenfoque de movimiento) lo cual es bastante común en vídeos reales de LR. [PST97b] es el primer trabajo que considera desenfoque por movimiento en la reconstrucción por SR en secuencias de vídeo basadas en la técnica POCS. Como el desenfoque causado por un tiempo de apertura finito será en general dependiente del espacio, y en ocasiones del tiempo, no puede independizarse de forma general del problema de la restauración y realizarse como un paso de post-procesamiento independiente. La técnica POCS puede tratar convenientemente tales problemas. Extendiendo este método, Eren *et al.* [EST97] propusieron una aproximación basada en POCS para una SR robusta basada en objetos. El método propuesto emplea un mapa de validación para deshabilitar proyecciones basadas en observaciones con estimación de movimiento imprecisa, y un mapa de segmentación para procesamiento basado en objetos. Elad y Feuer [EF97] analizaron y compararon los métodos ML, MAP y POCS para SR y propusieron una aproximación híbrida. Patti y Altunbasak [PA01] extendieron su trabajo inicial en el modelo de observación de la imagen para permitir una interpolación de alto orden y modificar los conjuntos de restricciones para permitir artefactos de borde.

La ventaja de POCS recae en su simplicidad a la hora de incorporar cualquier tipo de restricción o información *a priori* que pueda presentarse como imposible para estas aproximaciones estocásticas. Sin embargo, POCS es notorio por su elevado nivel de cómputo y su baja convergencia. La solución no es única, dependiendo de la predicción inicial. Los métodos POCS también asumen información *a priori* en los parámetros de movimiento y en los desenfoques del sistema. Estos métodos no pueden estimar los parámetros de registro y la imagen de HR de forma simultánea, como ocurría en las aproximaciones estocásticas. La aproximación híbrida que combina una visión estocástica y la filosofía POCS sugiere una forma prometedora de proseguir.

2.2.8 Topología de Macro-Bloques adaptable

Una de las innovaciones introducidas en la presente Tesis Doctoral se basa en la selección de los Macro-Bloques (MB) más adecuados a utilizar en el proceso de SR. El paso previo a la selección de los MBs se sostiene en una disposición no uniforme de los MBs, esto es lo que se conoce como “topología de Macro-Bloques adaptable”, siendo éste un campo estudiado previamente en la literatura. Su *et al.* [STW+12] dividieron la imagen de salida de HR en bloques de tamaño adaptativo, seleccionando diferentes algoritmos de SR a partir del tamaño de los bloques y aplicando un proceso de “*deblocking*” para reducir artefactos. Este trabajo es una extensión de [STW+08] donde la imagen de salida de HR se dividía en Macro-Bloques de tamaño variable a partir de la textura local y del error de compensación de movimiento para eliminar la influencia de la imprecisión del flujo óptico (*optical flow*).

La clave del trabajo publicado en [STW+12] está en la forma en la que se afronta el problema del error de registro. Existen razones para dividir la imagen de salida de HR en MBs de tamaño adaptable, en lugar de en regiones de formas arbitrarias, como hacen la mayoría de algoritmos de SR basados en regiones. En primer lugar, un análisis de bloques del algoritmo con una estrategia *coarse-to-fine* resulta computacionalmente más eficiente que un algoritmo de segmentación de regiones de píxeles. En segundo lugar, el ruido espacial es más sencillo de eliminar en áreas pequeñas o píxeles que no se adaptan bien al proceso de estimación de movimiento. Además, la SR basada en MBs presenta una facilidad natural en el procesamiento de secuencias de vídeo comprimido, ya que los algoritmos de compresión de vídeo dominantes están basados en MBs de tamaño fijo o MBs de tamaño adaptable [MWS06].

En la propuesta de [STW+12] la imagen de HR se divide en bloques rectangulares $p \times q$ del mismo tamaño m_{init} como el conjunto de división de bloques $\mathcal{B} = \{b_i | i = 1, \dots, p \times q\}$. Entonces, las características de cada bloque b_i se analizan para refinar el conjunto de división de bloques y etiquetar el patrón de bloques. Se usa entonces una matriz para examinar la continuidad alrededor de un píxel, donde la estructura de la matriz para un píxel simple $I(x, y)$ en la imagen I se define como se presenta en la ecuación (2.18).

$$S(x, y) = \nabla I(x, y) \cdot \nabla I(x, y)^T \quad (2.18)$$

De esta forma, en la ecuación (2.19) se define la estructura de la matriz del bloque b_i .

$$S_i = \frac{1}{n_i} \sum_{(x,y) \in b_i} S(x, y) = \frac{1}{n_i} \sum_{(x,y) \in b_i} \nabla \tilde{H}(x, y) \cdot \nabla \tilde{H}(x, y)^T, \quad (2.19)$$

donde n_i representa el número de píxeles del bloque b_i . En la imagen de HR original, H es desconocido; usándose por tanto la predicción \tilde{H} de la H original en lugar de ésta. \tilde{H} se puede generar interpolando directamente la entrada de LR de la imagen, pero esto hace que la ecuación (2.19) sea imprecisa. Esta es una de las razones para usar MBs de tamaño adaptable, lo cual reduce el ruido espacial promediando la estructura de la matriz para cada bloque.

Por otra parte, Anagün y Seke [AS12] implementaron un algoritmo de reconstrucción de tamaño de bloque variable con rotación en el registro de la imagen, consistente en tres pasos:

- Estimación de movimiento: Se aplica un proceso de estimación de movimiento de tamaño de MB variable con rotación. Se presenta un algoritmo que selecciona MBs de tamaño 32×32 , 16×16 u 8×8 píxeles para la estimación de movimiento. Esto reduce la complejidad computacional y proporciona una calidad visual similar o mejor con respecto a un tamaño de MB fijo. Tras determinar el tamaño de MB adecuado, se busca el MB en el *frame* en evaluación mediante una búsqueda angular (que considera rotación del MB).
- Registro: El método registra las imágenes localmente (basándose en los MBs). Se aplica entonces un registro global para obtener la imagen de HR. El registro se realiza con parámetros de movimiento subpíxel, que se obtienen a partir de la descripción de registro de imágenes de Irani y Peleg [IP91].
- Interpolación: Finalmente, a partir del conjunto de imágenes de HR generadas, se crea la imagen de salida de HR haciendo uso de la interpolación.

2.3 ESTADO DEL ARTE DE SISTEMAS MULTI-CÁMARA

2.3.1 Introducción

En esta sección, se realiza una revisión del Estado del Arte de sistemas Multi-Cámara (MC), que es una de las líneas de desarrollo de la presente Tesis Doctoral. Se presentarán diferentes propuestas, así como las principales investigaciones que se están llevando a cabo en este ámbito. Posteriormente, se procederá a analizar aquellas propuestas en las que se usan sistemas MC en combinación con algoritmos de SR. Por último, se expondrán una serie de conclusiones que surgen como resultado de las diferentes fuentes analizadas en el Estado del Arte.

Un sistema o *array* MC está formado por un conjunto de cámaras que se agrupan con un propósito común. Las cámaras se pueden agrupar y posicionar, de forma que capten diferentes ángulos de un mismo objeto, permitiendo mediante un post-procesamiento de esas imágenes obtener un modelo 3D del objeto. Otro propósito puede ser conseguir una secuencia de mejor calidad o mayor resolución a partir de las imágenes capturadas por el *array* MC, como es el caso de los algoritmos de SR.

Existen diversos tipos de sistemas o *arrays* MC, los cuales presentan una amplia variedad de aplicaciones, que van desde la mejora de la calidad de las imágenes, generalmente en términos de resolución, a videovigilancia, estabilización de vídeo, etc. En general, la distribución de las cámaras en el *array* MC está íntimamente ligada con la aplicación a la que este sistema esté destinado.

Como es previsible, estos sistemas requieren de anchos de banda de transmisión elevados, lo que depende directamente del número de cámaras y de la resolución de las imágenes capturadas, así como la tasa de *frame* establecida en la grabación. Por tanto, dentro de las propuestas estudiadas se usan diferentes estándares de transmisión en función del número de cámaras de las que dispone el *array* MC, así como de si el sistema opera en tiempo real o no. Los estándares de transmisión utilizados para la implementación de *arrays* MC se basan en diferentes estándares de comunicación, usando puertos GPIO (*General Purpose Input-Output*), *Ethernet*, *Bluetooth*, USB (*Universal Serial Bus*) 2.0 y USB 3.0. Por otra parte, también existen sistemas a medida, los cuales surgen como solución a los elevados requisitos de ancho de banda. Estos sistemas se encargan de procesar, a partir de la compresión de las imágenes a transmitir o el filtro de aquellas partes de la imagen innecesarias para la aplicación, con el fin de reducir el ancho de banda.

2.3.2 Propuestas de sistemas Multi-Cámara

2.3.2.1 Generación de modelos tridimensionales de objetos

Los sistemas MC se han utilizado para obtener modelos 3D digitalizados de objetos. Así, Weinmann *et al.* [WSR+11] proponen el diseño de un *array* MC fijo en forma de semiesfera, representado en la Figura 2.5, y cuyas condiciones de luz se prepararon para este tipo de procesos, de forma que tanto la disposición de las cámaras como las condiciones de iluminación favorecieran la captura de la información 3D del objeto.

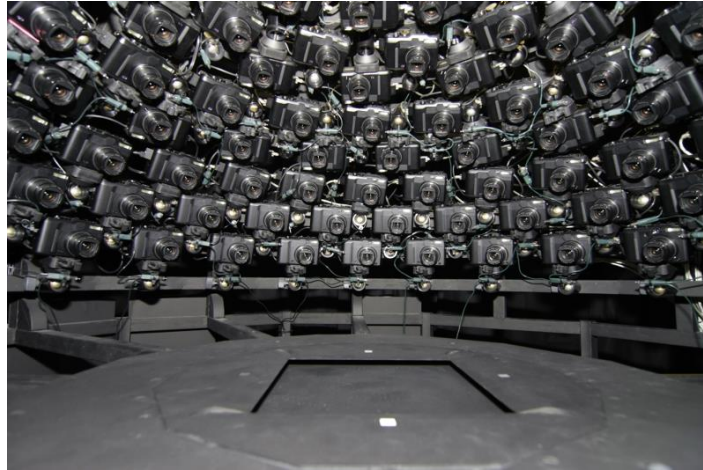


Figura 2.5.- Array MC con geometría semiesférica

Aunque el número de cámaras del *array* es fijo, dependiendo de la complejidad de la figura a modelar se emplea un subconjunto de cámaras, estableciéndose de este modo un punto de equilibrio entre el coste computacional y la resolución del objeto 3D alcanzada con el sistema. A través de cada cámara se obtiene un mallado de la figura desde la posición en la que se encuentra, que se combina con uno o varios mallados adyacentes, creando con esta combinación un mallado de mayor resolución, tal y como se muestra en la Figura 2.6.a, y por tanto obteniendo una mejor calidad del modelo 3D final. En la Figura 2.6.b se presenta un ejemplo de los resultados obtenidos por esta propuesta, haciendo uso de 150 cámaras del *array* MC.

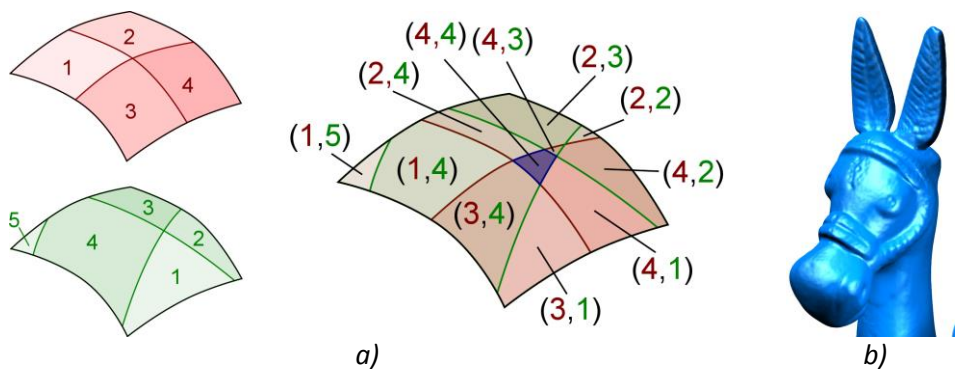


Figura 2.6.- Propuesta de Weinmann et al. [WSR+11] a) Mallado de HR, b) Resultados obtenidos

Otro ejemplo de uso de un *array* MC en modelado 3D es el aportado por Zaharescu et al. [ZCI+08], donde el *array* MC empleado está formado por un conjunto de cámaras situadas alrededor del objeto a modelar, tal y como se muestra en la Figura 2.7.a y en la Figura 2.7.c, las cuales, a la hora de realizar el procesamiento, se agrupan en *clusters* de cámaras. El número de *clusters* en los que se dividirá el *array* MC dependerá de la complejidad de la figura a modelar, dividiendo en un mayor número de *clusters* el *array* MC cuando la figura presente una mayor complejidad. Estos *clusters*, a su vez, pueden estar solapados entre sí, es decir, una o varias cámaras pueden pertenecer a varios *clusters*, siempre y cuando sean *clusters* adyacentes, tal y como se puede ver en las Figuras 2.7.a y 2.7.c, donde se observa que los tres *clusters* comparten cámaras del *array* MC entre sí. Finalmente, las Figuras 2.7.b y 2.7.d, muestran el modelo 3D obtenido tras haber empleado procesamiento a través de este *array* MC y su organización en *clusters*.

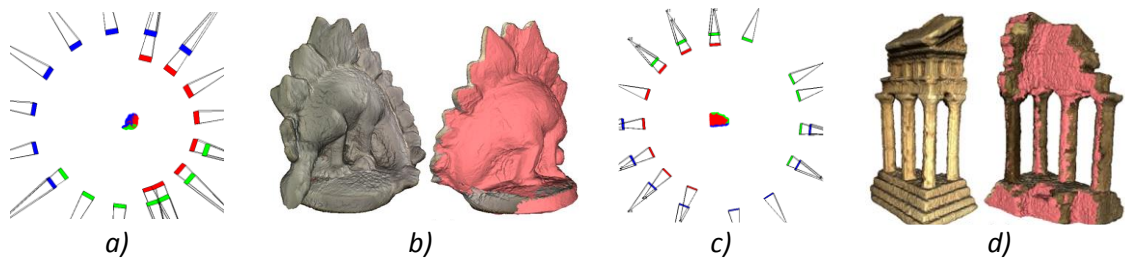


Figura 2.7.- Disposición de cámaras (a y c) y modelo 3D (b y d) de dinosaurio y ruinas [ZCI+08]

2.3.2.2 Generación de espacios tridimensionales

Otra de las aplicaciones en las que se usan *arrays* MC consiste en focalizar una parte de la escena que pueda ser de interés, o incluso generar una nueva cámara virtual a partir de la información capturada por el resto de las cámaras, de forma que el usuario se pueda mover por la escena como si de un espacio 3D se tratase. Ejemplo de ello es el propuesto por Yang *et al.* [YEB+02], cuyo trabajo permite a diferentes usuarios ver en tiempo real partes de una escena captada por el *array* MC. El *array* MC utilizado consta de 64 cámaras distribuidas con geometría cuadrada en una matriz de 8x8, mostrado en la Figura 2.8. Para conseguir que el sistema funcione en tiempo real se procesa la imagen localmente en la cámara, de modo que sólo se envíe la parte de la imagen que es de interés para el resultado final. Para ello existe un sistema denominado *Compositer* que funciona como interfaz entre el usuario y la cámara. Este sistema se encarga de recibir la petición del usuario, traducirla en la información que necesita la cámara para obtener la parte de la imagen deseada, y finalmente componer la imagen final que será enviada posteriormente al usuario.

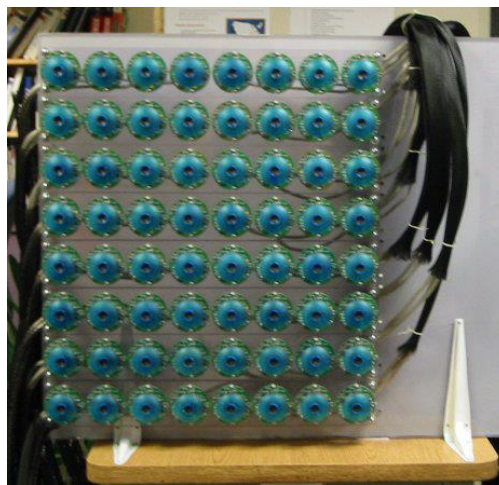


Figura 2.8.- Array MC de 64 cámaras distribuidas en una matriz cuadrada de 8x8

En esta propuesta, se plantean dos configuraciones distintas: una que permite acceder a todos los puntos de vista (*All-Viewpoint*), la cual no ha sido implementada debido al alto ancho de banda necesario, al permitir elegir cualquier punto de vista de la escena a procesar, con lo que sería necesario almacenar los *frames* completos de cada cámara con este propósito, no cumpliendo con los requisitos de tiempo real, y otra que trabaja únicamente con un número finito de puntos (*Finite-Viewpoint*), permitiendo únicamente seleccionar un subconjunto de vistas de la escena con el objetivo de que el sistema pueda operar en tiempo real. Con esta configuración se obtienen los resultados mostrados en la Figura 2.9, donde se puede ver que se enfoca en cada imagen a una persona distinta en planos diferentes.



Figura 2.9.- Tres vistas donde se enfoca a la persona de la izquierda, centro y derecha

Por otro lado, Smith *et al.* [SZJ+09] proponen un algoritmo de estabilización de vídeo cuya función es compensar el movimiento producido por el operador de la cámara cuando se graba una secuencia de vídeo mientras éste se desplaza, es decir, conseguir un algoritmo cuyo resultado sea el mismo que el de una *Steadicam*, que muestra imágenes similares al punto de vista subjetivo del personaje. Para conseguir este efecto se usa un *array* MC comercial, en concreto el dispositivo ProFusion-25C [PFS15], que contiene 25 cámaras fijas, tal y como se muestra en la Figura 2.10.a. De estas 25 cámaras se usan 4 de ellas en sus ejemplos, siendo las cámaras elegidas las centrales situadas en los extremos superior, inferior, izquierda y derecha, formando de este modo una cruz.

A partir de las grabaciones capturadas por estas cámaras se detectan los bordes, estableciendo de este modo puntos de referencia con el fin de conocer la información de movimiento de la secuencia a lo largo del tiempo. Posteriormente, mediante un proceso de minimización del movimiento, se crea una cámara virtual respecto a una cámara de referencia, también conocida como *Free-ViewPoint*, la cual no tiene restricciones en cuanto a su posición, y que pasará finalmente por un proceso de regularización de este movimiento para obtener la secuencia estabilizada. En la Figura 2.10 se puede observar el movimiento de los puntos de referencia a lo largo del tiempo, tanto en la secuencia original de la Figura 2.10.b, como en la procesada, que se representa en la Figura 2.10.c.

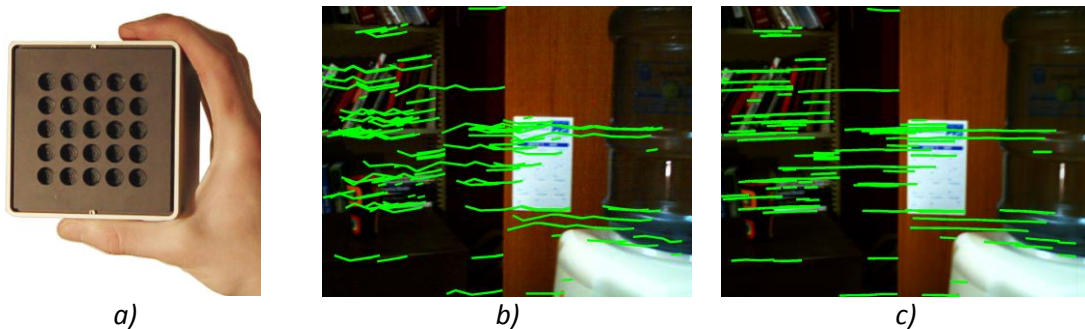


Figura 2.10.- Propuesta de Smith *et al.* [SZJ+09]. a) Cámara ProFusion-25C, b) secuencia original con movimiento y c) secuencia una vez procesada para minimizar el movimiento

Baker y Tanguay [BT06] proponen un sistema modular que puede soportar hasta 24 cámaras. Estas cámaras están conectadas a un banco de concentradores, los cuales usan el bus PCI-X (*Peripheral Component Interconnect eXtended*) para transferir los datos al PC. Este *array* MC, mostrado en la Figura 2.11.a, es configurable, habiéndose realizado montajes correspondientes con *arrays* de dimensiones 2x3, 2x9 cóncavo y 3x6, mostrados en la fila superior de la Figura 2.11.b, y de 22 cámaras y 3x6 convexo en la fila inferior. El propósito de este *array* MC es obtener vídeo de definición variable, mediante la técnica de mosaicos o *Mosaicing*. Para obtener las imágenes de resoluciones variables, el *array* MC se configura de forma que el solape entre las imágenes de las cámaras varíe entre un 20% y 50%, lo que

combinado con los distintos montajes permiten lograr este propósito. En la Figura 2.12 se muestra un ejemplo de esta propuesta, consiguiendo una imagen con gran ángulo de visión, tomada mediante el montaje de cámara cóncava de dimensiones 2x9 mostrado en la Figura 2.11.b.

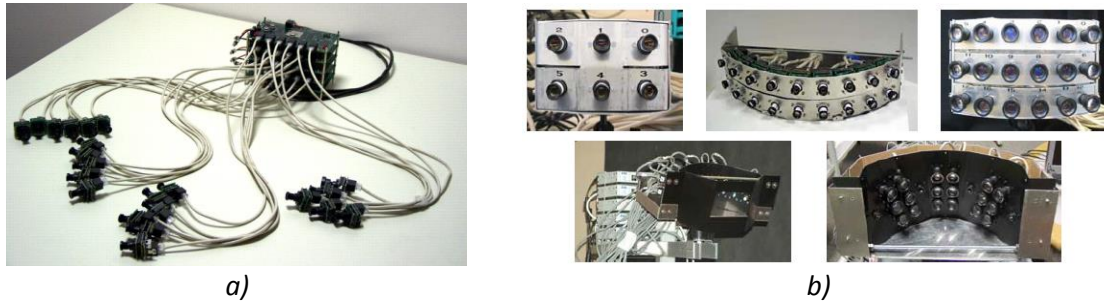


Figura 2.11.- Sistema modular propuesto por Baker y Tanguay [BT06]
 a) Sistema MC configurable, b) Montajes del Sistema MC realizados

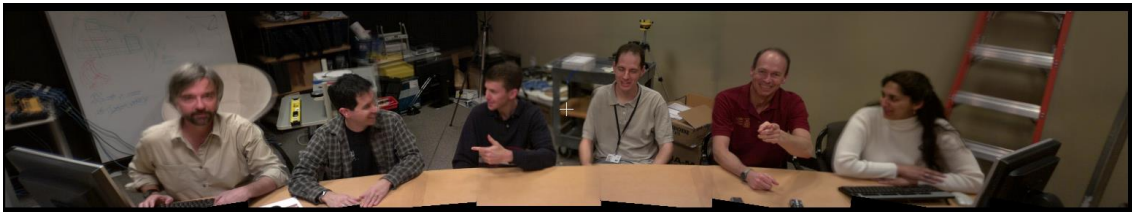


Figura 2.12.- Imagen con gran ángulo de visión tomada mediante un array de dimensiones 2x9

2.3.2.3 Videovigilancia

Otra de las aplicaciones posibles en MC es la de videovigilancia. Así, Bellotto *et al.* [BSB+09] emplean un sistema que consta de cámaras destinadas al seguimiento de personas en un sistema de videovigilancia, donde se puede cambiar el sujeto a seguir de forma dinámica. Para ello se usan 3 cámaras, donde una de ellas es fija y el resto son móviles, es decir, pueden moverse con la intención de seguir al sujeto que se tenga como objetivo. Estas cámaras proporcionan la opción de controlar el zoom, consiguiendo de esta forma agrandar y centrar el seguimiento con mayor precisión. La disposición final de las cámaras se presenta en la Figura 2.13.a, en la cual se puede apreciar la cámara estática en la parte superior, y la cámara activa en la parte inferior. Por otro lado, en la Figura 2.13.b se muestra el momento de detección del sujeto, mientras que en la Figura 2.13.c se observa cómo, una vez detectado, se lleva a cabo un *zoom* sobre el sujeto mientras se realiza el seguimiento.

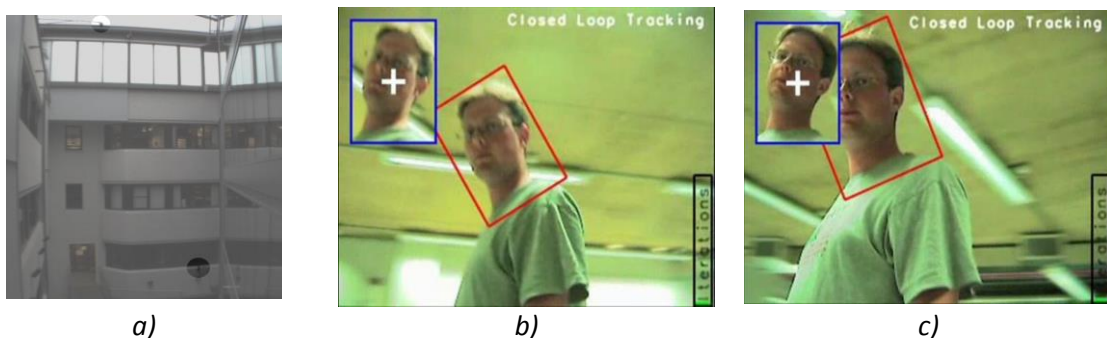


Figura 2.13.- Propuesta de Bellotto *et al.* [BSB+09]. a) Disposición de las cámaras del sistema, b) detección del sujeto y c) seguimiento del sujeto

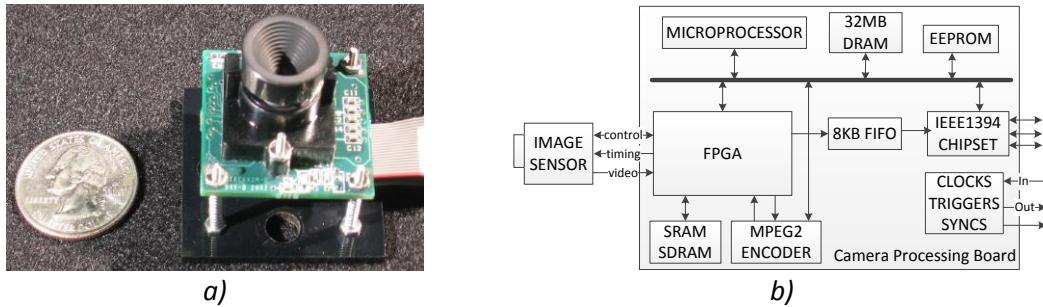


Figura 2.14.- Propuesta de Wilburn et al. [WJV05]. a) Imagen de una cámara a medida y b) Diagrama de bloques de la placa de procesado de cada cámara

Dentro de las propuestas de *arrays* MC más completas que se han estudiado, se encuentra la aportada por Wilburn et al. [WJV05], de la Universidad de Stanford, donde se ha creado un sistema a medida en el que la geometría del *array* es configurable para adaptarlo a las diversas aplicaciones que se presentan en esta propuesta. Este sistema incluye, desde las propias cámaras, hasta la placa de procesamiento para cada cámara. Las cámaras usadas se muestran en la Figura 2.14.a, y el diagrama de bloques de la placa de procesado de las cámaras se puede ver en la Figura 2.14.b. Por otra parte, en la Figura 2.15.a y en la Figura 2.15.b se muestran algunos ejemplos de los montajes de *arrays* MC que se han llevado a cabo en la Universidad de Stanford. Como ejemplo de aplicaciones que se han desarrollado a partir de este sistema MC, se presenta en su propuesta la de vídeo de resolución escalable mediante la técnica de mosaicos o *Mosaicing*, donde estas imágenes se obtienen a partir del solape entre las cámaras del *array* MC. Por otro lado, se consigue vídeo de alta velocidad, es decir, el efecto de tener una cámara con una elevada tasa de *frame*, de modo que se puedan obtener secuencias a cámara lenta, capturando para ello un *frame* por cada cámara del *array* de forma consecutiva. El resultado final produce el efecto de haber aumentado la tasa de *frame*. Como última aplicación a destacar, se ha implementado un algoritmo que crea lo que han denominado *Fotografía de Apertura Sintética No Lineal*, a través del cual se consigue eliminar o poder visualizar lo que hay detrás de objetos a partir de las diferentes perspectivas capturadas de la misma escena, al estar las cámaras del *array* MC en diferentes puntos del espacio, como se observa en las imágenes de la Figura 2.16.



Figura 2.15.- Arrays MC de la Universidad de Stanford. a) Array MC con teleobjetivos de dimensiones 8x12 y b) Array MC donde las cámaras tienen una amplia separación



Figura 2.16.- Secuencia de vídeo procesada con *Fotografía de Apertura Sintética No Lineal*

2.3.2.4 Aplicaciones comerciales

En el mercado existen diversos productos o aplicaciones en los cuales se emplean *arrays* MC con diferentes propósitos, como por ejemplo en el ámbito audiovisual, dentro del cine, videoclips deportivos o musicales. En este caso se genera un efecto en el cual se detiene la escena y se gira alrededor del personaje en cuestión, creando una sensación de rotación 3D. Generalmente, para los vídeos deportivos o promocionales se emplean *arrays* MC contruidos con cámaras GoPro [GPR15]. Algunos ejemplos de estos sistemas MC usados con este propósito se pueden ver en la Figura 2.17.

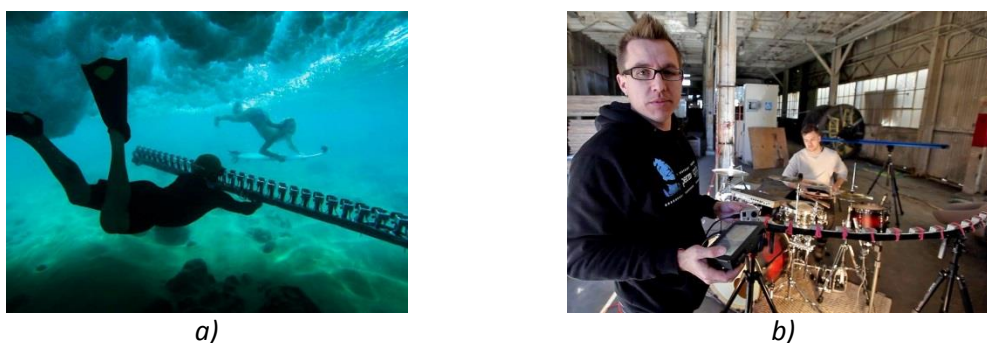


Figura 2.17.- Array MC implementado mediante cámaras GoPro®. a) Disposición submarina, b) Montaje en fábrica

Recientemente, una página *web* de financiación colectiva ha conseguido mediante micromecenazgo implementar y producir la cámara *Panono* [PNN15]. Esta cámara, mostrada en la Figura 2.18, permite tomar fotos en 360°. La geometría de esta cámara consiste en una esfera que contiene 36 cámaras fijas en su interior. El funcionamiento de la cámara *Panono* consiste en lanzar esta esfera hacia arriba, la cual una vez detecta mediante un acelerómetro que ha llegado al punto más alto toma una foto en ese punto con cada una de sus cámaras, cuya información se fusiona mediante la técnica de *Mosaicing* para crear una imagen panorámica esférica de 360°, capturando todo lo que pasa a su alrededor.



Figura 2.18.- Cámara Panono® para obtener panorámicas en 360°

La empresa IDS (*Integrated Dynamic Solutions*) [IDS15] ofrece una serie de cámaras que pueden ser conectadas a un PC mediante comunicación USB 3.0, que combinadas con un software específico de la propia empresa para gestionar la captura y sincronización de las imágenes, permite crear diferentes configuraciones de *arrays* MC, tal y como se muestra en la Figura 2.19. Por otra parte, cabe resaltar que en el cine 3D cobra cierta importancia el uso de *arrays* MC, puesto que para obtener este efecto se necesitan capturar las imágenes desde dos puntos de vista distintos, como si de los ojos del ser humano se tratase, para posteriormente llevar estas imágenes al espectador.



Figura 2.19.- Sistema MC de la empresa IDS® a través de USB 3.0

Finalmente, existen ideas de aplicaciones comerciales basadas en el concepto de MC para SR, como la reciente publicación de solicitud de patente de Apple Inc. [BT14], según la cual plantean un sistema para generar una imagen súper-resuelta a partir de un dispositivo de captura de imágenes que toma muestras ópticas en trayectorias diversas haciendo uso de una lente, tal y como se representa en la Figura 2.20. Así, utilizando un procesador óptico de estabilización de la imagen se ajustan diferentes muestras considerando desplazamientos subpíxel, combinando éstas para obtener una imagen súper-resuelta. Esta característica se integraría en los dispositivos de la empresa, como por ejemplo el Iphone®, tal y como se representa en la Figura 2.21.

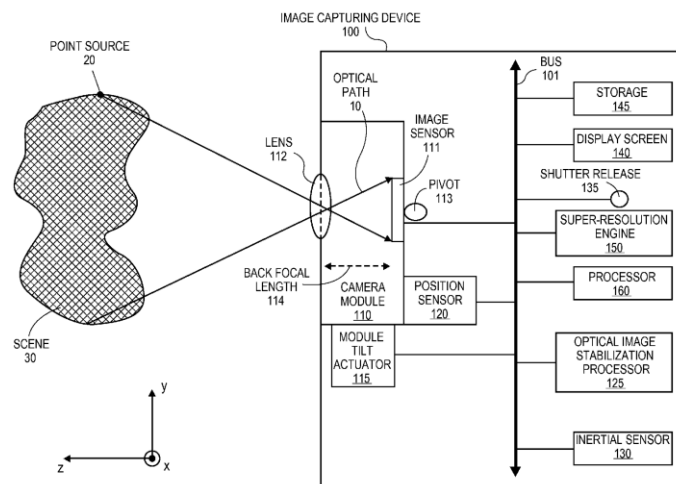


Figura 2.20.- Sistema de SR basado en estabilización óptica de la imagen [BT14]

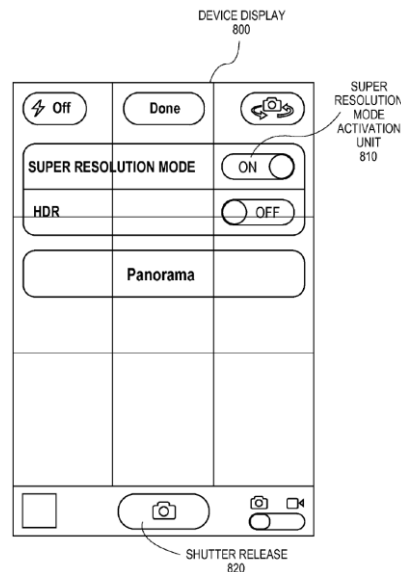


Figura 2.21.- Integración prevista de sistema de SR en cámara de Iphone® [BT14]

2.3.2.5 Otras aplicaciones

En la mayoría de las referencias anteriormente expuestas, los autores, aparte de plantear las aplicaciones a las que han destinado el *array* MC, así como en algunos casos describir la implementación *hardware*, introducen diversos algoritmos para caracterizar los parámetros del *array* MC y corregir el desalineamiento entre las cámaras, regularizar las diferencias de color e iluminación, etc. Sin embargo, existen otras propuestas que se centran única y exclusivamente en el estudio de estos problemas, abstrayéndose o independizándose del *array* MC en sí, y por tanto, planteando soluciones genéricas que se pueden incluir en cualquier sistema MC. Un ejemplo sería la propuesta por Kang *et al.* [KLH08] del Instituto de Ciencia y Tecnología de Gwangju (GIST, *Gwanju Institute of Science and Technology*) en Corea del Sur, donde se parte del uso de un *array* MC plano de una sola dimensión, y cuyo algoritmo se puede extender a más dimensiones de forma iterativa. En este caso se obtiene una línea base de todas las cámaras y se corrigen las diferencias de los desalineamientos verticales y horizontales, así como de las debidas a las rotaciones entre las cámaras respecto a esta línea base. Para el cálculo de esta línea base se realiza un proceso iterativo donde se calculan los puntos medios entre las cámaras, que al unirlos obtienen nuevas rectas, de las cuales se vuelven a obtener sus puntos medios, repitiendo este proceso hasta que sólo quedan 2 puntos que se corresponden con la línea base, tal y como se puede ver en la Figura 2.22.

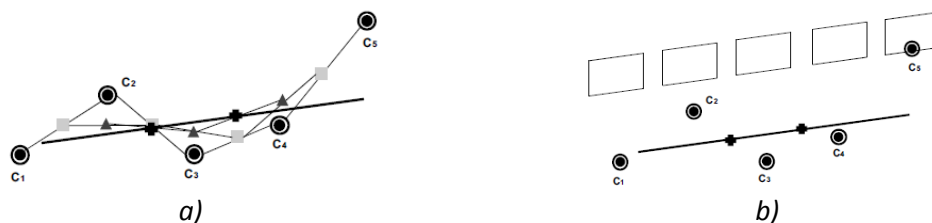


Figura 2.22.- Propuesta genérica de Kang *et al.* [KLH08]

a) Proceso de cálculo de la línea base de corrección de las cámaras, b) Línea base resultante

Por otro lado, en [KH08] Kang y Ho no solo contemplan *arrays* MC planos de una dimensión, sino que también soportan *arrays* de una dimensión en forma de arco. En esta propuesta se añade a la anterior un proceso similar para obtener esta línea base en *arrays* en forma de arco, donde las imágenes tomadas por este *array* MC son compensadas y combinadas posteriormente con Homografía 2D para obtener las imágenes corregidas. Por otra parte, en el trabajo de Weng *et al.* [WAH+92] de la Universidad de Illinois, realizan una propuesta donde se tiene en cuenta un algoritmo de corrección de perspectiva para escenas 3D en las que se producen oclusiones entre las diferentes imágenes captadas por cada una de las cámaras.

2.3.3 Arrays Multi-Cámara en Súper-Resolución

Como se ha presentado en el apartado anterior, existe una amplia variedad de propuestas de sistemas MC diseñados para una gama de múltiples propósitos. Así mismo, existen muchas propuestas en las que basarse para implementar algoritmos que resuelvan el desalineamiento entre las cámaras del *array* MC. En este apartado se pretende estudiar las diferentes propuestas que usan sistemas MC en combinación con algoritmos de SR. La finalidad consiste en aprovechar la información espacial y la información temporal de un sistema de estas características, así como estudiar la problemática que presentan. En este caso, el título de cada apartado coincide con el artículo tomado como referencia.

2.3.3.1 Multi-Camera imaging system using Super-Resolution

En el ámbito de los estudios introducidos, algunos autores plantean diversos sistemas en los que se combina un algoritmo de SR con un *array* MC, de modo que se aprovecha la información espacial que existe entre las cámaras de éste, como por ejemplo en la propuesta de Hyun Park *et al.* [HMG08] del Instituto Universitario de Yonsei en Corea del Sur. En esta propuesta, los autores hacen uso de un *array* MC de dimensiones 3x3, del cual 3 de las cámaras se destinan a capturar la información de color, mientras que las otras 6 obtienen la información de luminancia de la imagen LR. Este *array* MC se muestra en la Figura 2.23. La imagen HR es reconstruida a partir de las imágenes en las que solo se capta la luminancia, para lo cual se aplica un algoritmo basado en Estimación de Imagen Bayesiana Multicanal (*Bayesian Multichannel Image Estimation*), con la característica de considerar el error de movimiento local mediante la regularización de los parámetros. Finalmente, para obtener la imagen HR resultante se aplica un método de fusión que mezcla la información de luminancia y de color. La aplicación concreta de esta propuesta es el renderizado.



Figura 2.23.- Array MC de dimensiones 3x3 del Instituto Universitario de Yonsei [HMG08]

2.3.3.2 Super resolution reconstruction of three view remote sensing images

La propuesta de Zhang *et al.* [ZWM+11] de la Universidad Tianjin en China, dentro del ámbito de renderizado aeroespacial, captura imágenes a través de un *array* de tres líneas (*Three Line-Array*) obteniendo las vistas frontal, *nadir* y hacia atrás tomadas con una cámara de teledetección de barrido asíncrono (*Asynchronous Push Broom Remote Sensing Imaging*), donde la vista *nadir* se corresponde con la intersección entre la vertical de la cámara y la tierra. Estas tres vistas se combinan para obtener la imagen HR súper-resuelta mediante una variante del algoritmo POCS (*Projections onto Convex Sets*). Esta variante consiste en obtener un coeficiente de ponderación global, que se obtiene a través de una media del gradiente, considerando que en las zonas de accidentes geográficos es donde se encuentran mayores variaciones del gradiente sobre la imagen de la superficie lunar, conteniendo por tanto gran cantidad de información. De este modo se propone un algoritmo POCS de Ponderación Global (*Global Weighted POCS*), mediante el cual se obtiene la imagen súper-resuelta.

2.3.3.3 Wireless camera network for image Super-Resolution

Directo *et al.* [DSC04] de La Universidad McMaster en Canadá, proponen un modelo limitado a dos cámaras, relacionadas mediante Homografía Plana 2D (*2D Planar Homography*). El objetivo de este modelo es corregir la perspectiva de una de las cámaras sobre la otra, la cual se establece como cámara de referencia y por tanto, será la cámara cuya secuencia será súper-resuelta. Para obtener la imagen de HR, el proceso planteado consiste en los siguientes pasos: obtener correspondencia entre imágenes (*Image Correspondence*), registrar la imagen (*Image Registration*) y finalmente súper-resolver la imagen (*Image Super-Resolution*).

La fase de correspondencia entre imágenes se encarga de obtener puntos equivalentes entre las imágenes de las cámaras, para lo cual se establecen puntos de control (*Harris feature points*) en ambas imágenes, que pretenden ser un conjunto apropiado de valores aproximados de entrada, y no una correspondencia perfecta entre las imágenes. El proceso de registro de la imagen es el encargado de obtener la relación de homografía plana 2D a partir de los puntos de correspondencia entre imágenes. Para llevar a cabo esta fase se hace uso de la Transformación Lineal Directa (*Direct Linear Transform, DLT*), del algoritmo *RANdom SAmple Consensus* (RANSAC), así como de una fase final de emparejamiento guiado (*Guided Matching*). Todo ello se plantea con el objetivo de obtener una buena estimación de la matriz H de Homografía Plana 2D, que es la matriz de transformación que relaciona ambas imágenes a partir del conjunto de puntos obtenidos en la fase de correspondencia entre imágenes. Para determinar H a través de homografía 2D basta con conocer 4 puntos. Sin embargo, en caso de tener más de 4 puntos, y mediante el uso del algoritmo RANSAC, es posible obtener una estimación de la matriz H de manera más robusta. Esto se debe a que este algoritmo, en caso de tener un conjunto de puntos mayor que 4, toma grupos de 4 puntos seleccionados de manera aleatoria consecutivamente, a partir de los cuales calcula la matriz H mediante DLT. Esta matriz será válida si los puntos de correspondencia están dentro de cierto umbral de distancia. La estimación de la matriz H se dará como buena cuando se hayan tomado suficientes grupos de puntos aleatorios. Una vez estimada esta matriz H se realiza un proceso de correspondencia guiada de modo que se pueda obtener una estimación más refinada. Cuando se consigue una buena estimación de la matriz H , se aplica a la imagen a corregir sobre la establecida como referencia. El último paso consiste en aplicar el algoritmo de SR POCS para súper-resolver las imágenes. Para este caso se ha modelado el sistema de captura considerándolo como un Desplazamiento Lineal Variante (*Linear Shift Varying, LSV*), incluyendo movimiento de cámara y de escena, así como el difuminado, además de incorporar en el algoritmo POCS la PSF, modelada como un núcleo gaussiano (*Gaussian Kernel*).

2.3.3.4 Super-Resolution recovery for Multi-Camera surveillance imaging

La propuesta de Caner *et al.* [CTH03] de la Universidad de Rochester de Nueva York en EEUU, al igual que en [DSC04], propone hacer uso de un modelo de registro entre los *frames* de LR capturados por las distintas cámaras, respecto a un *frame* de referencia que posteriormente será súper-resuelto. En este caso, el modelo genérico propone un sistema de varias cámaras que, no sólo estará desalineado espacialmente, sino que además no tiene que estar sincronizado temporalmente, produciéndose desalineamientos en este sentido entre las secuencias de LR. Este modelo propone como desalineamientos temporales, no solo la falta de sincronización, sino también las posibles diferencias entre la tasa de *frame* de las cámaras. Así mismo, las diferencias de tasa de *frame* se modelan con una transformación unidimensional.

Los desalineamientos espaciales entre cámaras anteriormente comentados, se deben tanto a parámetros internos como externos entre las cámaras. Estos pueden ser modelados, bien mediante Transformación Proyectiva (*Projective Transformation*), o lo que es lo mismo Homografía, o bien mediante una Matriz Fundamental de parámetros. Estos dos modelos contemplan situaciones diferentes, donde la Homografía puede aplicarse cuando las imágenes se puedan considerar como una escena plana. Sin embargo, cuando las imágenes captadas puedan tener características suficientemente significativas debidas a su alineación, o pertenezcan a un escenario 3D, donde por tanto se pueden dar oclusiones debidas a la posición en que se encuentran las cámaras, se requerirá de un modelo más complejo que contenga o extraiga información de la estructura 3D de la escena, apareciendo el uso de una Matriz Fundamental. Sin embargo, en los ejemplos expuestos por estos autores se asume que las tasas de *frame* son iguales, y por tanto, los desalineamientos temporales se modelan mediante una constante. Por otro lado, en los resultados iniciales para estimar el desalineamiento espacial se hace uso de un modelo de movimiento cuadrático 2D. En cuanto al método de SR empleado, se usa POCS considerando: difuminado introducido por el sensor y la óptica, y ruido del sensor, entre otros. Con respecto al *array* de cámaras, se hace uso de dos cámaras situadas muy cerca una de la otra. Este *array* a su vez puede presentar diferentes configuraciones, tal y como se muestra en la Figura 2.24.

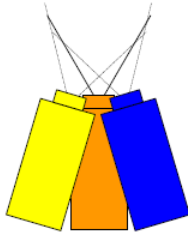


Figura 2.24.- Posibles configuraciones del array MC propuesto por la Universidad de Rochester

2.3.3.5 Regularized Super-Resolution of Multi-View Images

Mientras que de manera general existen diversos algoritmos que operan aprovechando la información temporal para aplicar algún algoritmo de SR, otros aprovechan únicamente la información espacial. En esta Tesis Doctoral desarrollada en la Universidad de Carleton en Canadá se propone un sistema completo de SR donde se contempla la caracterización de una cámara comercial (ProFUSION25) [PFS15], obteniendo su PSF, así como su función de viñeteado (*vignetting*). Además, se propone una nueva técnica de identificación de la PSF. Por otra parte, se usa un esquema para la detección de oclusiones basado en la intensidad, el cual ayuda a eliminar artefactos debidos a estimaciones de movimiento incorrectas. Por último, esta propuesta aporta un identificador de difuminado o *blurring*, así como un sistema cuya función es eliminar este efecto.

2.3.3.6 Super-resolution from multiple views using learnt image models

Capel y Zisserman [CZ01], de la Universidad de Oxford, en Reino Unido, proponen un modelo basado en patrones que ya fue introducido en la sección de “Aproximaciones basadas en patrones” de la presente Tesis Doctoral. La propuesta comprende un sistema compuesto por una única cámara y se aplica un algoritmo de SR Temporal. Este modelo se expresa a través de la ecuación (2.20).

$$g_n = s \downarrow (h * T_n \bar{f}) + \eta \quad (2.20)$$

La ecuación (2.20) representa que una imagen de LR g_n está relacionada con la imagen de HR \bar{f} , mediante una transformación geométrica T_n de tipo proyectiva (Homografía 2D), teniendo en cuenta la función de transferencia de la óptica, modelada en este caso a través de la PSF. Finalmente, se realiza un proceso de diezmado $s \downarrow$, al cual se le añade un término de ruido η , que se asume de distribución gaussiana con media cero.

Lo interesante del modelo aquí presentado es que establece que la relación entre las imágenes de HR y de LR se da a través de una transformación proyectiva. Este concepto puede verse con una perspectiva más amplia, en la cual entre los *frames* de una secuencia de vídeo que han sido grabados en distintos instantes temporales existirá también una relación proyectiva, puesto que puede haber movimiento entre los *frames*, ya sea porque la cámara está en movimiento, o porque sea la escena la que se mueva. De igual forma, existirá la misma relación entre los *frames* capturados en el mismo instante de tiempo por un *array* MC, donde la relación proyectiva entre estos *frames* se debe a una relación espacial basada en la posición de las cámaras en el *array* MC. Sin embargo, se plantea la posibilidad de que, teniendo una serie de imágenes captadas por un *array* MC, de las cuales se ha obtenido la relación de transformación de las cámaras respecto a una de referencia (pudiendo ser esta transformación proyectiva o no), se pueda aplicar un proceso de corrección de estas diferencias de alineamiento, escala, etc. Se considerarían entonces únicamente las traslaciones entre las cámaras, por lo que estas secuencias corregidas podrían ser muy apropiadas para el algoritmo de SR que se ha tomado como base para implementar los métodos propuestos en esta Tesis Doctoral.

2.3.3.7 Resumen de propuestas que incluyen *arrays* Multi-Cámara

Dentro de la variedad de propuestas de *arrays* MC que se han presentado en el apartado 2.3.2, y teniendo en cuenta las características del algoritmo de SR usado como base en esta Tesis Doctoral, que se beneficia de la información subpíxel que pueda encontrar entre las imágenes de la secuencia LR, se consideran como más adecuados aquellos *arrays* MC que presenten o minimicen la transformación proyectiva producida por la propia geometría del *array* MC. Por tanto, aquellos *arrays* MC cuyas cámaras estén lo más alineadas posibles y se encuentren situadas en el mismo plano, se consideran que son los que mejor se adaptan a los algoritmos de SR por fusión, pretendiendo con ello que no presenten cambios de escala o rotaciones en la perspectiva de las cámaras.

Por otro lado, no se establece a priori pauta alguna en cuanto al beneficio que puede presentar la disposición de las cámaras en el *array* MC, es decir, si las cámaras deben estar dispuestas formando una geometría cuadrada, rectangular, circular, en forma de cruz, u otra que pueda ser de especial interés. Sin embargo, tal y como se muestra en la Figura 2.25, se prevé que sea en la parte solapada entre los *frames* capturados por cada cámara (resaltada con líneas oblicuas), donde exista una mayor posibilidad de encontrar la información necesaria para reconstruir la secuencia HR. Desde otro punto de vista, al estar grabando cada cámara la misma escena desde diferentes puntos del espacio, tendrá en la zona solapada un mayor muestreo de la misma parte de la escena, y por tanto, se aumenta la posibilidad de capturar la información necesaria para el proceso de SR del *frame*.

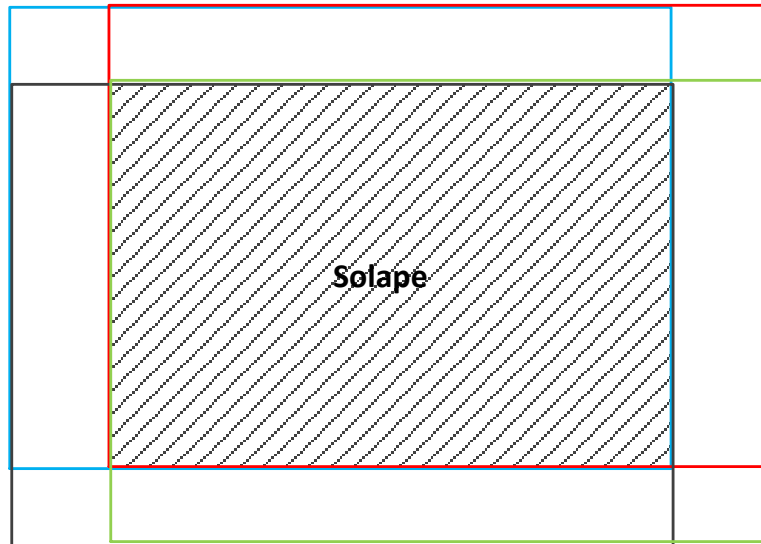


Figura 2.25.- Solape entre las imágenes capturadas por las cámaras

Como consecuencia, aquellas geometrías que aporten un mayor solape serán las que a priori presenten un mejor resultado, siendo las formas cuadrada y rectangular las consideradas como mejores candidatas, tal y como se puede apreciar en la Figura 2.26, aunque pueden resultar igualmente válidas otras geometrías. Con el mismo fin, el de aumentar la dimensión del solape, cuanto más cerca se encuentren las cámaras unas de otras en el *array* MC, mayor será la parte solapada, y por tanto se espera un mejor resultado en la calidad de la secuencia de salida.

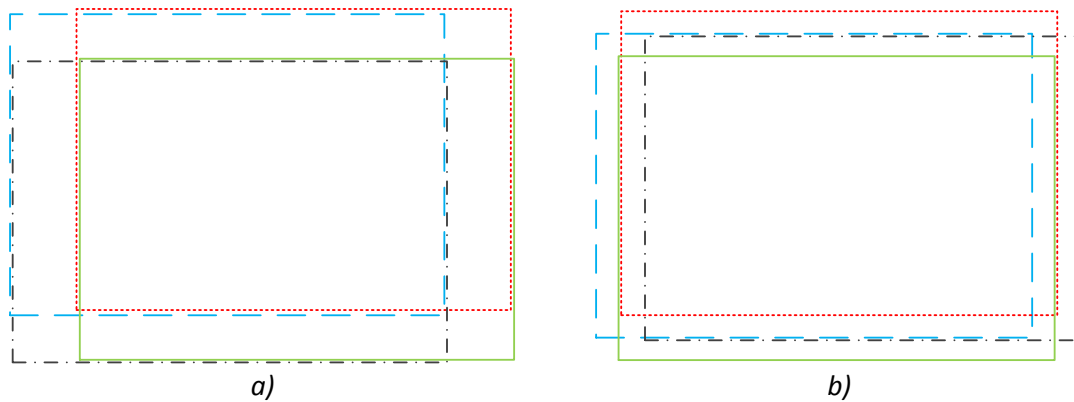


Figura 2.26.- Array MC de dimensiones 2x2 a) Geometría rectangular y b) Geometría en cruz

Como se ha podido ver en el apartado 2.3.3, existen diversos estudios que combinan SR con las imágenes captadas por un *array* MC. Sin embargo, en este estudio del Estado del Arte no se han encontrado propuestas que combinen la información temporal y espacial que proporciona el *array* MC para aportar una mejora en la calidad de las imágenes súper-resueltas. Por ello, no se dispone de una referencia en la que basarse para implementar un algoritmo con el fin de lograr dicho objetivo, así como para comparar adecuadamente los resultados obtenidos en la presente Tesis Doctoral. Finalmente, la Tabla 2.1 presenta un resumen de aquellas propuestas que combinan un *array* MC junto con algoritmos de SR.

Tabla 2.1.- Resumen de referencias que emplean un array Multi-Cámara con Súper-Resolución

Sección	Título	Entidad	Dimensión del array	Nº de cámaras	Registro	Algoritmo de SR	Aplicación
2.3.3.1	Multi-Camera imaging system using Super-Resolution [HMG08]	Yonsei University Institute, Corea del Sur	3x3	9	-	Bayesiano multicanal	Renderizado
2.3.3.2	Super resolution reconstruction of three view remote sensing images based on global weighted POCS algorithm [ZWM+11]	Tianjin University, China	Array de tres líneas	3	-	POCS de Ponderación Global	Renderizado Aeroespacial
2.3.3.3	Wireless camera network for image Super-Resolution [DSC04]	McMaster University, Canadá	2x1	2	Homografía	POCS	Renderizado
2.3.3.4	Super-Resolution recovery for Multi-Camera surveillance imaging [CTH03]	University of Rochester, EEUU	2x1	2	Homografía Matriz fundamental	POCS	Videovigilancia
2.3.3.5	Regularized Super-Resolution of Multi-View Images [Fan09]	Carleton University, Canadá	5x5	25	Warping + Traslaciones + Homografía	POCS	Renderizado

2.4 MÉTRICAS SOBRE CALIDAD DE IMAGEN

Con el objetivo de establecer una medida de calidad objetiva de las secuencias súper-resueltas, tras estudiar cuidadosamente el Estado del Arte, en la presente Tesis Doctoral se ha hecho uso de las métricas de calidad PSNR (*Peak Signal-to-Noise Ratio*) y SSIM (*Structural SIMilarity index*). En esta sección se presentan ambas métricas de calidad y las razones de su utilización, además de otras métricas estudiadas en el proceso de selección, comparadas en un último apartado a modo de conclusión.

2.4.1 Peak Signal-to-Noise Ratio (PSNR)

La métrica PSNR se calcula obteniendo el error cuadrático medio (MSE, *Mean Square Error*) en relación con el valor máximo posible de luminancia (usando 8 bits este valor se corresponde con $2^8 - 1 = 255$), como sigue:

$$MSE = \frac{\sum_{i=1}^M \sum_{j=1}^N [f(i,j) - F(i,j)]^2}{M \times N} \quad (2.21)$$

$$PSNR = 20 \log_{10} \left(\frac{255}{\sqrt{MSE}} \right), \quad (2.22)$$

donde $f(i,j)$ es la señal original en el píxel (i,j) , $F(i,j)$ es la señal reconstruida (súper-resuelta o interpolada), y $M \times N$ es el tamaño de la imagen. El resultado es un valor en decibelios (dB), que va desde 30 a 40 dB para vídeo de calidad medio-alta.

Aunque se han propuesto varios modelos para medir la calidad objetiva, el PSNR continúa siendo uno de los más populares en lo que se refiere a la evaluación de la diferencia de calidad entre imágenes. Recientemente Korhonen y Junyong [KJ12] presentaron que las bajas prestaciones de PSNR se deben frecuentemente a casos en los que existe mucho movimiento local. En escenarios con contenido fijo y con tipos de distorsión relacionados con aplicaciones de comunicación visual, PSNR es una métrica apropiada, y en algunos casos incluso mejor que modelos de calidad objetivos conocidos en la literatura.

2.4.2 Structural SIMilarity index (SSIM)

Wang *et al.* [WBS+04] presentaron una aproximación diferente para la evaluación objetiva de la calidad de vídeo. Este método difiere del PSNR y de la mayoría de los métodos de medida objetiva de la calidad de la imagen en que no está basado en el error, sino en la medida de distorsión estructural. La idea que subyace en este planteamiento es que el sistema de visión humano está altamente especializado en extraer la información estructural del campo de visión y no lo está en la extracción de errores. De esta forma, una medida de la distorsión estructural debería dar una mayor correlación a la impresión subjetiva.

Partiendo de la asunción inicial comentada, se pueden desarrollar diversos métodos de evaluación de calidad. La ventaja de SSIM es que representa un método simple y efectivo que es cada vez más frecuente en muchos trabajos científicos. SSIM se puede expresar de la siguiente manera: Sea $x = \{x_i \mid i = 1, 2, \dots, N\}$ la señal original e $y = \{y_i \mid i = 1, 2, \dots, N\}$ la señal distorsionada, el índice de similitud estructural se puede calcular como:

$$SSIM = \frac{(2xy+C_1)(2\sigma_{xy}+C_2)}{[(\bar{x})^2+(\bar{y})^2+C_1](\sigma_x^2+\sigma_y^2+C_2)}, \quad (2.23)$$

donde:

- \bar{x} es la media de x .
- \bar{y} es la media de y .
- σ_x es la varianza de x .
- σ_y es la varianza de y .
- σ_{xy} es la covarianza entre x e y .
- C_1 y C_2 son constantes.

El valor de SSIM es 1 si $x_i = y_i$ para todos los valores de i . Este índice de calidad se aplica normalmente a cada imagen usando una ventana deslizante de una función de pesos gaussiana simétrica y circular, y calculando el índice total de la imagen como el promedio de todos los índices de calidad de la imagen.

2.4.3 Otras métricas objetivas estudiadas

2.4.3.1 VQM (Video Quality Metric)

VQM (*Video Quality Metric*) propuesta por Pinson y Wolf [PW04] fue desarrollada por ITS (*Institute for Telecommunication Science*) para proporcionar una medida objetiva que permitiera evaluar la calidad de las secuencias de vídeo. Esta métrica considera los efectos de las deficiencias perceptuales en vídeo incluyendo desenfoque, movimiento errático, ruido global, distorsión por bloques y distorsión por color, y los combina en una única métrica. VQM presenta una alta correlación con la evaluación subjetiva de la calidad y fue adoptada por ANSI (*American National Standards Institute*) como una métrica objetiva estándar de calidad de vídeo. VQM toma la secuencia de vídeo original y la secuencia de vídeo procesada como entradas, y las analiza de acuerdo con los siguientes pasos:

- *Calibración*: Este paso calibra el vídeo muestreado. Estima y corrige el desplazamiento espacial y temporal, así como el contraste y el desplazamiento de brillo de la secuencia de vídeo procesada con respecto a la secuencia de vídeo original.
- *Extracción de características de calidad*: Este paso extrae un conjunto de características de calidad que caracteriza cambios perceptuales en las propiedades espaciales, temporales y de crominancia de las subregiones espacio-temporales de secuencias de vídeo usando una función matemática.
- *Cálculo de parámetros de calidad*: Este paso computa un conjunto de parámetros de calidad que describen cambios perceptuales en la calidad de vídeo comparando las características extraídas de las secuencias de vídeo procesada y original.
- *Cálculo VQM*: La métrica VQM se obtiene mediante una combinación lineal de los parámetros calculados en los pasos anteriores.

VQM puede ejecutarse usando varios modelos basados en ciertos criterios de optimización. Estos modelos incluyen televisión, videoconferencia, uso general, modo desarrollador y modo PSNR. El modelo general usa una combinación lineal de siete parámetros: cuatro están basados en características extraídas de gradientes espaciales de la componente de luminancia (Y), dos están basados en características extraídas del vector formado por las dos componentes de crominancia (C_b y C_r), y el último parámetro está basado en el contraste y en características temporales, extraídas de Y .

2.4.3.2 MPQM (*Moving Pictures Quality Metric*)

PSNR no toma en consideración el fenómeno del enmascaramiento visual. En otras palabras, cada error de píxel individual contribuye a una disminución en PSNR, incluso si dicho error no se percibe. Este hecho se afronta incorporando un modelo del sistema de visión humano. En particular, se han estudiado en detalle dos fenómenos de la percepción humana: la sensibilidad al contraste y el enmascaramiento. El primer fenómeno se refiere al hecho de que una señal se detecta por el ojo sólo si el contraste es mayor que un determinado umbral. La sensibilidad al ojo varía en función de la frecuencia espacial, la orientación y la frecuencia temporal. El segundo fenómeno está relacionado con la respuesta del sistema de visión humano a la combinación de varias señales. Un estímulo consiste en dos tipos de señales (*foreground* y *background*). El umbral de detección de *foreground* se modifica en función del contraste de *background*.

MPQM (*Moving Pictures Quality Metric*), propuesta por Labrech y Verscheure [LV96], es una métrica de calidad objetiva de imagen móvil que incorpora las dos características de la visión humana que se han comentado. En primer lugar descompone la secuencia original y su versión distorsionada en dos canales perceptuales. Se calcula entonces una medida de distorsión considerando tanto la sensibilidad al contraste como el enmascaramiento. Finalmente, los datos se analizan sobre todos los canales para calcular la calificación de calidad que se escala entonces de 1 a 5 (de malo a excelente, respectivamente).

Sin embargo, MPQM no toma en consideración el color, y por ello se introdujo el método CMPQM (*Color Moving Pictures Quality Metric*). El primer paso consiste en convertir las componentes de color a valores RGB que son lineales con la luminancia. Entonces los valores RGB se convierten para coordinar valores que corresponden a los canales de luminancia (*Black / White*), Rojo-Verde (*Red / Green*) y Azul-Amarillo (*Blue / Yellow*). Cada componente de las secuencias error y original se analiza con un banco de filtros. El denominado canal "*Black / White*" se procesa como luminancia, pero como "*Red / Green*" y "*Blue / Yellow*" son mucho menos sensibles, sólo se usan 9 filtros espaciales y uno temporal para estas señales. El resto del cálculo es el mismo que para MPQM.

MPQM representa los típicos modelos de evaluación de calidad basados en la sensibilidad al error. La asunción ampliamente adoptada de estos modelos es que la pérdida de la calidad perceptual está directamente relacionada con la visibilidad de la señal error. La mayoría de las aproximaciones de evaluación perceptual propuestas en la literatura intentan pesar diferentes aspectos de la señal error de acuerdo con su visibilidad, como se determina mediante medidas psicofísicas en humanos o fisiológicas en animales. El principio subyacente de la aproximación de la sensibilidad a errores consiste en que la calidad perceptual se estima mejor cuantificando la visibilidad de los errores. Esto se consigue esencialmente estimulando las propiedades funcionales de las primeras etapas del sistema de visión humano, caracterizado tanto mediante experimentos psicosociales como fisiológicos. Aunque esta aproximación *Top-Down* al problema presenta aceptación casi universal, es importante reconocer sus limitaciones. En particular, el sistema de visión humano es un sistema complejo y altamente no lineal, pero la mayoría de los modelos de visión temprana están basados en operadores cuasi-lineales que se han caracterizado usando estímulos simples y restringidos.

2.4.3.3 NQM (Noise Quality Measure)

En esta métrica de calidad, la imagen degradada se modela como la imagen original sujeta a distorsión lineal de frecuencia e inyección de ruido aditivo. Estas dos fuentes de degradación se consideran independientes y están desacopladas en dos medidas de calidad: una medida de distorsión (DM, *Distortion Measure*) y una medida de calidad de ruido (NQM, *Noise Quality Measure*), introducida por Damera-Venkata *et al.* [DGE+00] para medir el efecto del ruido aditivo. NQM toma en consideración los siguientes aspectos:

- Variación de la sensibilidad al contraste con la distancia y con las dimensiones de la imagen.
- Variación en la media local de luminancia.
- Contraste en la interacción entre frecuencias espaciales.
- Efectos de enmascaramiento de contraste.

2.4.3.4 AQI (Anysotropic Quality Index)

La anisotropía es por definición la propiedad de ser dependiente de la dirección y es uno de los aspectos considerados por los investigadores para encontrar la relación entre la estructura operacional del sistema visual y el contenido real. No es posible determinar un patrón de dirección común para todas las escenas reales, ya que la composición de la escena varía, y con ello la dirección de muestra a muestra. Los estudios realizados por Li *et al.* [LPF03] indican que las neuronas alineadas en dirección horizontal son más prevalentes que las alineadas en dirección vertical. La relación existente entre la prevalencia de la escena natural a diferentes orientaciones y la habilidad visual de detectar la dirección de la escena han sido estudiados por Hansen y Essock [HE04]. Kei y Cristóbal [KC00] llevaron a cabo una comparación sistemática entre el contenido horizontal y vertical en función de la frecuencia espacial y encontraron una mayor influencia horizontal a determinadas frecuencias con preponderancia del contenido vertical a otras frecuencias espaciales. Estos experimentos, así como las diferencias que encontraron en la distribución de energía para imágenes naturales en el dominio espacio-frecuencial, sugieren que las medidas entrópicas, tal como la entropía de Rényi, basada en medir el contenido de las frecuencias de las imágenes a través de PWD (*Pseudo-Wigner Distribution*) [WBH91], son apropiadas para medir la anisotropía. La ecuación (2.24) muestra la entropía Rényi $R_3[n]$ asociada a una posición n en función de PWD, que se identifica con una probabilidad de distribución \check{P}_n .

$$R_3[n] = -\frac{1}{2} \log_2 \left(\sum_{k=1}^N \check{P}_n^3[k] \right) \quad (2.24)$$

Para formalizar los cálculos requeridos con el fin de medir la anisotropía de las imágenes, se usa la ecuación (2.24). Esta expresión proporciona un valor de entropía $R_3[n, \theta_s]$ para cada píxel. $\theta_s \in [\theta_1, \theta_2, \dots, \theta_S]$ representa S orientaciones diferentes para medir la entropía. Por tanto, Gabarda y Cristóbal [GC07] introdujeron la ecuación (2.25) con el fin de definir una figura de mérito para la imagen (índice anisotrópico AQI, *Anysotropic Quality Index*).

$$\vec{R}[t, \theta_s] = \sum_n R_3[n, \theta_s] / M, \quad (2.25)$$

donde M representa el tamaño de la imagen y $t \in [1, 2, \dots, T]$ se introduce para considerar las T diferentes imágenes que integran el conjunto de datos.

Así mismo, suponiendo que $\vec{R}(t, \theta_s)$ es el valor esperado de entropía para una imagen $t \in [1, 2, \dots, M]$, medido en direcciones $\theta_s \in [\theta_1, \theta_2, \dots, \theta_S]$, la desviación estándar para el conjunto resultante de valores, referidos a la imagen t , puede definirse como se representa en la ecuación (2.26).

$$A(X_t) = \sigma(t) = \sqrt{\sum_{s=1}^S (\mu_t - \vec{R}(t, \theta_s))^2 / S}, \quad (2.26)$$

identificando $A(X_t)$ como la anisotropía de una imagen X_t , y donde μ_t es la media de los valores $\vec{R}(t, \theta_s)$, como se define en la expresión (2.27):

$$\mu_t = \sum_{s=1}^S \vec{R}(t, \theta_s) / S, \quad (2.27)$$

definiéndose el rango en la imagen t en la ecuación (2.28).

$$rg(t) = \max\{\vec{R}(t, \theta_s)\} - \min\{\vec{R}(t, \theta_s)\} \quad (2.28)$$

Una vez que la relación entre la anisotropía y la calidad de la imagen ha quedado establecida, se puede extender fácilmente a una medida de calidad logarítmica para poder comparar con otras medidas similares basadas en anisotropía. Callicó *et al.* [CLG+09] definieron una medida logarítmica LQM (*Logarithmic Quality Measure*) que estima en dBq la calidad absoluta de las imágenes (q se refiere a calidad) y se ha añadido para identificar su carácter absoluto X_t como se presenta en la ecuación (2.29):

$$LQM(X_t) = B \cdot \log_{10}(A(X_t)), \quad (2.29)$$

donde B es una constante que ha de determinarse con el fin de fijar el rango de valores operativos.

2.4.4 Comparativa entre métricas objetivas

Además de los métodos comentados anteriormente, se han propuesto otros, pero ninguno de ellos se acepta de forma general. La Tabla 2.2 presenta una comparativa entre las métricas previamente presentadas, basándose en el trabajo realizado por Wang [Wan06], así como los análisis presentados por Sheikh *et al.* [SSB06] y Korhonen y Junyong [KJ12].

Tabla 2.2.- Comparativa entre métricas objetivas de calidad de imagen

Métrica	Complejidad Matemática	Correlación con métodos subjetivos
AQI [GC07]	Muy compleja	Buena
PSNR [KJ12]	Simple	Media
SSIM [WBS+04]	Compleja	Buena
VQM [PW04]	Muy compleja	Buena
MPQM [LV96]	Compleja	Variante
NQM [DGE+00]	Compleja	Media

El análisis presentado en la Tabla 2.2, así como la evaluación de una amplia cantidad de artículos científicos, demuestran que la métrica PSNR sigue siendo ampliamente usada, y que SSIM se está adaptando cada vez más. Esto justifica la selección de ambas métricas para la presente Tesis Doctoral, a partir de las cuales se podrá realizar una medida de calidad objetiva adecuada, además de servir como punto de referencia para comparar con las contribuciones propuestas en la literatura.

2.4.5 Métricas subjetivas

El objetivo de la evaluación subjetiva de calidad es conocer, para un conjunto determinado de *frames* (o de secuencias de vídeo), la opinión de un usuario medio. En este sentido, la mejor forma de saberlo es preguntando directamente a los usuarios. Los métodos de calidad objetivos proporcionan guías de cómo preguntar sobre calidad de una forma eficiente.

Existen varios estándares que proporcionan estos métodos de evaluación objetiva, principalmente ITU-R BT.500 [ITU02], ITU-T P.910 [ITU08] e ITU-T P.911 [ITU98]. Todos ellos son bastante similares en la forma que proponen estructurar, realizar y evaluar los *tests*. La mayoría de los *tests* de evaluación subjetiva en la literatura están basados en estos estándares, siendo los *tests* VQEG (*Video Quality Experts Group*) el ejemplo más relevante [Win09].

En las sesiones de *test* se le solicita a un número determinado de sujetos que visualicen un determinado conjunto de *frames* o de secuencias de vídeo y que evalúen su calidad. El número total de evaluadores ha de estar comprendido entre 4 y 40, pudiendo estar distribuidos en diferentes sesiones. De forma general:

- Al menos 15 observadores deberían participar en el experimento.
- No han de ser profesionales en la evaluación de calidad.
- Deberían disponer de una capacidad visual normal o corregida para ser normal (uso de lentillas o gafas).

En el caso de presentar secuencias de vídeo en los *tests*, la localización y los monitores donde se realizan los *tests* deben cumplir con una serie de requisitos en términos de luz, iluminación y contraste de la pantalla, o distancia y ángulo de observación, entre otros. Existen también guías para trabajar, tanto con monitores profesionales, como con televisiones domésticas. Las sesiones realizadas no deberían durar más de media hora. Normalmente al comienzo de la sesión se les presenta a los observadores un conjunto de secuencias de vídeo o *frames* donde pueden ver el tipo de defectos detectables. Las muestras a evaluar pueden estar precedidas de presentaciones de prueba, cuyos resultados no se toman en consideración, para estabilizar la opinión de los observadores. Además, los *frames* o secuencias a evaluar deberían distribuirse aleatoriamente a lo largo de las sesiones.

Aunque existen variaciones en los detalles entre estándares, las estrategias que se usan principalmente, son las siguientes [ITU98]:

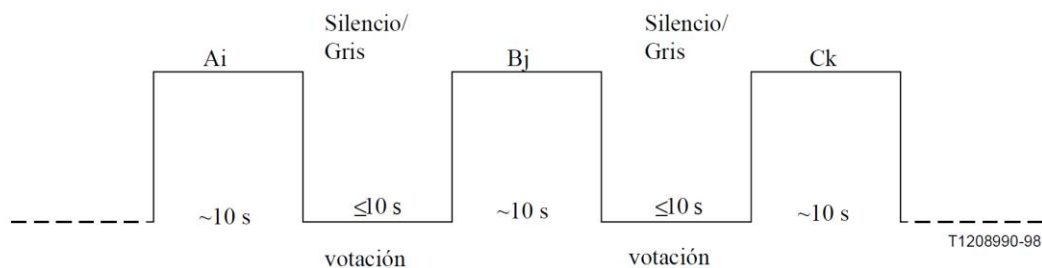
- Evaluación de Categoría Absoluta (ACR, *Absolut Category Rating*), o Método de Estímulo Simple (SS, *Single Stimulus*). Las secuencias de *test* se presentan una única vez y se evalúan independientemente con una escala. Después de cada presentación se le solicita a los sujetos que evalúen la calidad de la secuencia presentada usando una escala absoluta, normalmente con cinco niveles (ver Tabla 2.3). A veces se utilizan escalas de nueve niveles o incluso de once niveles para aumentar la resolución, pero

en la mayoría de las ocasiones esto no da lugar a una variación significativa de los resultados. Se muestra además a modo de ejemplo el modelo de presentación de estímulo para el método ACR según ITU-T P.911 en la Figura 2.27.

- Evaluación de Categoría de Degradación (DCR, *Degradation Category Rating*), o método de Escala de Deficiencias de Doble Estímulo (DSIS, *Double Stimulus Impairment Scale*). En este caso, cada presentación consiste en dos muestras diferentes: el contenido de referencia (sin deficiencias) y la versión procesada o con deficiencias del mismo contenido. Se evalúan las dos muestras de forma consecutiva, y se le solicita al sujeto que evalúe el grado de deficiencia del segundo estímulo con respecto al de referencia. En este caso se usan también escalas de 5 niveles, mostrándose en la Tabla 2.3 la relación existente entre las escalas de evaluación ACR y DCR.
- Método de Comparación por Pares (PCO, *Pair Comparison*). Las secuencias de vídeo o *frames* de *test* se presentan por pares, al igual que en el caso de DCR, pero ahora las secuencias de vídeo o *frames* son dos versiones procesadas de la misma muestra original (por ejemplo con dos niveles diferentes de una misma deficiencia o con tipos diferentes de deficiencias). Tras presentar cada par, el sujeto ha de seleccionar su preferida.
- Evaluación de Calidad Continua de Simple Estímulo (SSCQS, *Simple Stimulus Continuous Quality Scale*). Este método se aplica cuando se plantean como punto de partida secuencias de vídeo de larga duración (de 3 a 30 minutos). Mientras la secuencia se está reproduciendo, se les solicita a los sujetos que evalúen su calidad de forma continua.

Tabla 2.3.- Comparativa entre métricas de evaluación subjetivas ACR y DCR

Evaluación	ACR	DCR
5	Excelente	Imperceptible
4	Bueno	Perceptible pero no molesto
3	Neutro	Ligeramente molesto
2	Pobre	Molesto
1	Malo	Muy molesto



- Ai Secuencia A en condición de prueba i
- Bj Secuencia B en condición de prueba j
- Ck Secuencia C en condición de prueba k

Figura 2.27.- Modelo de presentación de estímulo para el método ACR según ITU-T P.911

Cuando se presentan secuencias de vídeo, la duración propuesta de las secuencias es de 10 segundos, incluyendo otro período de 10 segundos (consistente en una pantalla en gris) para votar cada una de las secuencias, como se puede ver para ACR en la Figura 2.27. Cuando se usan pares de secuencias (DCR y PCO), los pares de secuencias deben separarse haciendo uso de un período de 2 segundos basado en una pantalla en gris.

2.5 DESAFÍOS DE LA SÚPER-RESOLUCIÓN Y CONTRIBUCIONES

En la sección 2.2, se han presentado varias técnicas para la reconstrucción mediante SR y en la sección 2.4 las métricas de calidad de imagen más adecuadas para medir las mejoras introducidas. Si bien se han propuesto diversas aproximaciones diferentes desde que se introdujo el concepto de SR, la mayoría de éstas operan adecuadamente en imágenes de *test* y no tanto en situaciones reales [Mil11]. A la hora de construir un sistema de SR práctico, existen muchos retos para que éste pueda trabajar en un amplio rango de aplicaciones. En esta sección se presentan varios desafíos clave en el desarrollo y aplicación de técnicas de SR y las contribuciones generales de la presente Tesis Doctoral en cada una de las líneas aquí expuestas.

2.5.1 Eficiencia de cómputo

Las técnicas de SR normalmente requieren de una computación intensiva, dado el gran número de operaciones a realizar como consecuencia de requerir con frecuencia manipulaciones de matrices. Las aplicaciones reales siempre demandan eficiencia en la reconstrucción por SR para resultar útiles en la práctica. Por ejemplo, en escenarios de videovigilancia resulta fundamental que la reconstrucción por SR se realice en tiempo real. La eficiencia también es deseable en sistemas de SR que incluyen parámetros ajustables. Muchos algoritmos de SR que tienen como objetivo la eficiencia pertenecen a la aproximación previamente introducida de restauración mediante interpolación, tal como se describe en [ABH+00], [EH01], [Har07] y [NHB+07]. En [Har07], Hardie mostró la superioridad de cómputo de su algoritmo sobre algoritmos previos eficientes propuestos por Alam *et al.* [ABH+00] y Narayanan *et al.* [NHB+07], y expuso que el algoritmo se podía aplicar en tiempo real con un modelo de traslación global. Sin embargo, el cómputo se incrementa significativamente cuando no hay un modelo de traslación, aspecto que puede mejorarse usando computación paralela. Otros autores intentaron examinar escenarios de modelos particulares para acelerar el problema de la optimización. Zomet y Peleg [ZP00] y Farsiu *et al.* [FRE+04] estudiaron la aplicación de D_k , H_k , y F_k directamente como las operaciones de submuestreo, desenfoque y desplazamiento respectivamente, evitando la necesidad de construir explícitamente las matrices y dando lugar a incrementos significativos de velocidad. Bowen y Bouganis [BB08] combinaron una versión ligeramente modificada de [EH01] y [FRE+04] para implementar un sistema de SR en tiempo real usando una FPGA (*Field Programmable Gate Array*).

Sin embargo, estos algoritmos requieren de un registro preciso de la imagen, lo cual resulta intensivo en cómputo. Además, hasta ahora, estos algoritmos sólo pueden gestionar eficientemente modelos de movimiento simple lejos de aplicaciones de vídeo con mucho movimiento. Para vídeos con movimientos arbitrarios, Takeda *et al.* [TMP+09] sugieren alternativas prometedoras para desarrollar algoritmos eficientes. También es interesante ver cómo la computación paralela, por ejemplo usando GPUs (*Graphic Processing Units*), e implementaciones *hardware* afectan las aplicaciones futuras de técnicas de SR.

La presente Tesis Doctoral estudia cómo disminuir el cómputo del algoritmo considerando un pre-procesamiento en la información de entrada de LR con el fin de determinar los *frames* más adecuados a incluir en el proceso de SR (Filtro Selectivo de Ventana) haciendo uso de la métrica objetiva de similitud estructural SSIM, presentada en la sección 2.4.2 del presente capítulo.

2.5.2 Registro de la imagen

El registro de la imagen es crítico para el éxito de la reconstrucción de SR de múltiples *frames*, donde se fusionan muestreos espaciales complementarios de la imagen de HR. El registro de la imagen es un problema fundamental de procesamiento de imagen que es habitualmente conocido como mal planteado (*ill posed*). El problema es incluso más difícil desde el punto de vista del proceso de SR, donde las observaciones son imágenes de LR con artefactos de *aliasing* notables. Las prestaciones de los algoritmos estándar para registro de imagen reducen a medida que la resolución de las observaciones disminuye, dando lugar a más errores de registro. Los artefactos causados por estos errores de registro son visualmente más molestos que el efecto de desenfoque resultante de la interpolación de una imagen. La reconstrucción por SR tradicional trata normalmente el registro de la imagen como un proceso distinto a la estimación de la imagen de HR. Por tanto, la calidad de la imagen de HR recuperada depende en gran medida de la precisión en el registro de la imagen. En la literatura se han propuesto diversas técnicas de registro de imagen derivadas de diferentes principios, recogidas en *surveys*, como los publicados por Brown [Bro92] o Zitová y Flusser [ZF03]. Sin embargo, Robinson y Milanfar [RM04] mostraron que las prestaciones del registro de la imagen dependen de un proceso de traslación global.

El registro de la imagen y la estimación de la imagen de HR son en realidad interdependientes, tal y como expusieron Robinson y Milanfar en [RM06]. Por una parte, una estimación de movimiento subpíxel precisa beneficia la estimación de la imagen de HR. Por otra parte, una imagen de HR de alta calidad puede facilitar una estimación de movimiento precisa. Por tanto, adecuándose al problema de reconstrucción por SR, el registro de imagen de LR puede tratarse conjuntamente con la reconstrucción de imagen de HR, llevando a un marco de estimación simultánea de tipo ML [TKG94] o MAP [HBA97, PRZ03, SZH+07]. Estos algoritmos de estimación capturan la dependencia entre el registro de imagen de LR y la estimación de la imagen de HR, dando lugar a mejoras en las prestaciones. Sin embargo, con observaciones limitadas, la estimación conjunta para parámetros de registro y la imagen de HR pueden resultar en un fenómeno de sobreajuste.

En la presente Tesis Doctoral se demuestra cómo el uso de una estructura basada en múltiples cámaras (*array* de cámaras o sistema Multi-Cámara) influye de forma positiva en el problema planteado por la SR en relación con el registro de la imagen, ya que al disponer de múltiples observaciones de una misma secuencia, y considerando que el desplazamiento presente entre los *frames* capturados es subpíxel, el resultado de la imagen de HR procesada mejorará las prestaciones de calidad objetiva frente a una situación en la que sólo se disponga de una única cámara.

2.5.3 Robustez

Las técnicas tradicionales de SR son vulnerables a la presencia de anomalías debido a errores de movimiento, modelos de enfoque imprecisos, objetos en movimiento, desenfoque por movimiento, etc. Estos modelos de error imprecisos no se pueden tratar como ruido gaussiano. La robustez del proceso de SR es de interés porque los parámetros de degradación de la imagen no se pueden estimar de forma precisa, y la sensibilidad a las anomalías puede resultar en artefactos visualmente molestos, que son intolerables en muchas aplicaciones, como por ejemplo en conversión de vídeo. Sin embargo, no se ha dedicado suficiente trabajo a un aspecto tan importante como éste. Chiang y Boult [CB00] usaron estimación por mediana para combinar las imágenes submuestreadas con el fin de abordar el problema de las anomalías de ruido no estacionario. Zomet *et al.* [ZRP01] se enfrentan a este problema de una

forma diferente, usando un gradiente basado en mediana para evitar la influencia de anomalías. Farsiu *et al.* [EFR+04] utilizaron una estimación robusta similar a [CB00]. Muchos de estos algoritmos mostraron mejoras en las anomalías mostradas en la información de imágenes de prueba, requiriéndose en cualquier caso una evaluación experimental que valide hasta qué punto estos esfuerzos enfocados a la robustez pueden beneficiarse de las prestaciones reales de la SR.

En esta Tesis Doctoral se introduce un filtro adicional al presentado en la sección 2.5.1, en este caso orientado a cada uno de los Macro-Bloques (MB) en los que se divide la secuencia de entrada para realizar el proceso de SR. Se mostrará que un estudio particular para cada uno de los MBs proporciona una mayor robustez a los algoritmos de SR por fusión, ya que al habilitar un tamaño variable de MB y determinar el grado de movimiento local de cada uno de ellos frente a una aproximación basada únicamente en movimiento global, la calidad global se incrementa. Además, se plantea un caso real orientado a secuencias de vídeo submarino grabadas por vehículos tipo ROV (*Remotely Operated Vehicle*) presentando cómo el cálculo de la velocidad media del vehículo partiendo de información de Georreferenciación, proporciona información clave a la hora de determinar el número de *frames* a seleccionar en el proceso de SR, lo que implica un mayor grado de robustez en el algoritmo.

2.5.4 Límites en las prestaciones

La reconstrucción por SR ha sido un tema de investigación muy activo desde que se introdujo hace 30 años, y se han publicado desde entonces miles de artículos relacionados con este tema [SCO15], como se muestra en la Figura 2.28, en la que se presenta una gráfica de los artículos relacionados con SR desde el trabajo de Tsai y Huang en 1984 [TH84] hasta finales del año 2013 (*Super-Resolution OR Superresolution*). Sin embargo, no se ha desarrollado mucho trabajo orientado a la comprensión fundamental de los límites en las prestaciones de los algoritmos de reconstrucción, lo cual podría dar luz al diseño de cámaras de SR, ayudando a analizar factores tales como modelos de error, factores de *zoom* y número de *frames*.

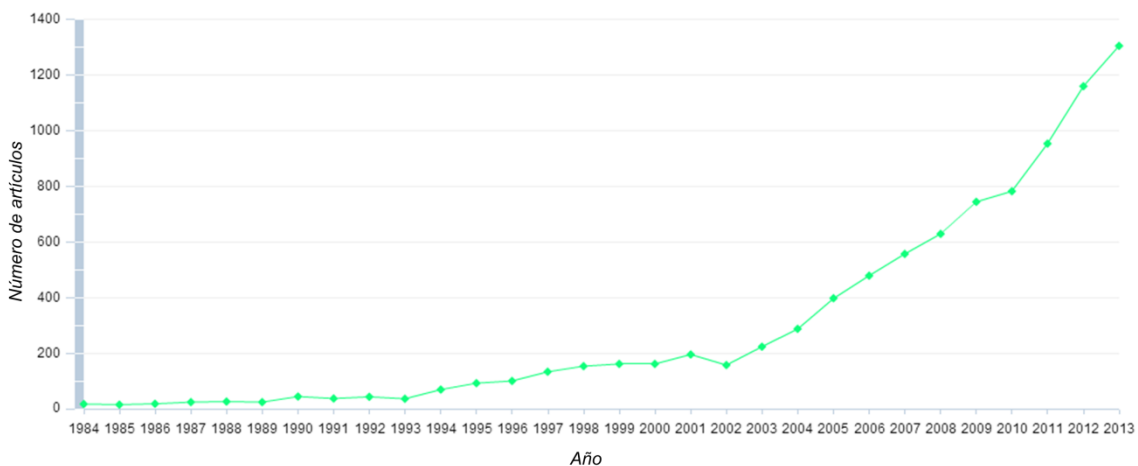


Figura 2.28.- Número de artículos de SR desde el año 1984 hasta el año 2013

En general, un análisis exhaustivo de los límites de las prestaciones de las técnicas de SR puede ser inabordable. En primer lugar, la reconstrucción por SR es una tarea compleja que integra muchas componentes interdependientes. En segundo lugar, se desconoce también cuál es la predicción que proporciona más información para la tarea de SR, sobre todo por las aproximaciones basadas en patrones. Finalmente, se necesita una buena medida de calidad en lugar de una simple MSE para la evaluación de las prestaciones. Se ha comprobado que una

estimación con un MSE más alto no tiene por qué resultar mejor visualmente. Por ejemplo, la interpolación bicúbica normalmente alcanza un MSE más pequeño comparado con aproximaciones de SR basadas en patrones [YWH+08], aunque esta última presente mejor calidad visual.

Se han propuesto varios trabajos en los últimos años intentando acercarse a la aproximación de las prestaciones de las técnicas de SR. Baker y Kanade [BK02] analizaron las condiciones numéricas de los sistemas de SR lineales, y concluyeron que a medida que se incrementa el factor de *zoom*, la predicción realizada ayuda cada vez menos al proceso de SR. Lin y Shum [LS04] se aproximaron a los límites de las prestaciones basados en las perturbaciones de las matrices, pero con la asunción de que el registro de la imagen se considera como predicción. Con un modelo de traslación simple, Robinson y Milanfar [RM04] usaron las restricciones de Crámer-Rao (CR) para analizar los límites de prestaciones en lo que se refiere al registro de la imagen. Este trabajo fue extendido en [RM06] para proporcionar un análisis detallado de las prestaciones de la SR con factores tales como la estimación de movimiento, el factor de diezmado, el número de *frames*, o la información de predicción. El análisis se basó en el uso de MSE como métrica de calidad objetiva y en el modelo de movimiento que se asume una vez como un modelo simple de traslación global. Eekeren *et al.* [ESO+07] evaluaron varios algoritmos de SR con información en tiempo real explorando varios factores influyentes empíricamente. Aunque estos esfuerzos para comprender los límites de las prestaciones están bastante lejos de la SR, pueden sugerir nuevas vías a explorar.

Si bien es difícil perfilar conclusiones consistentes para diferentes técnicas de SR en términos de evaluación de sus prestaciones, es claro que se necesitan bancos de pruebas y conjuntos de secuencias reales y realistas que proporcionen una comparación justa y una comprensión adecuada de los algoritmos. En el caso particular de la presente Tesis Doctoral, se han diseñado bancos de pruebas específicos ajustados para proporcionar información fidedigna de los resultados obtenidos. Al mismo tiempo, se han desarrollado sistemas que permitan determinar los máximos teóricos de calidad objetiva en términos de PSNR y SSIM que se podrían alcanzar con los algoritmos propuestos.

2.6 CONCLUSIONES

En este capítulo, se ha presentado un Estado del Arte de las técnicas de SR, introduciendo las aplicaciones en las que tradicionalmente se han utilizado, y desgranando las características de las aproximaciones a la SR de los últimos 30 años. Con el objetivo de comprender las aportaciones de los diferentes autores considerados, se ha comenzado describiendo un modelo de observación de la imagen mediante un diagrama de bloques completo que integra los diferentes procesos por lo que pasa la imagen cuando es captada por un sensor CMOS o CCD. Se ha presentado además una revisión del Estado del Arte de sistemas Multi-Cámara, introduciendo diferentes propuestas, las principales investigaciones que se están llevando a cabo en este ámbito, y su relación con el campo de estudio de la SR.

Por otra parte, en el análisis de prestaciones de las diferentes técnicas de SR presentadas, resulta fundamental considerar métricas de calidad objetivas adecuadas para evaluar la mejora planteada por los algoritmos introducidos. Así, se ha concluido que las métricas PSNR (*Peak-to-Signal Ratio*) y SSIM (*Structural SIMilarity Index*) son apropiadas para evaluar las aportaciones de la presente Tesis Doctoral, además de servir como punto de referencia para comparar las contribuciones propuestas con la literatura.

Finalmente, se han introducido las principales contribuciones que presenta esta Tesis Doctoral frente a los desafíos que plantea la SR considerando el Estado de Arte actual: eficiencia en el cómputo, registro de la imagen, robustez y límites en las prestaciones.

Capítulo 3

CONTRIBUCIONES ALGORÍTMICAS

Índice del capítulo

	<i>Página</i>
3.1.- Introducción	50
3.2.- Algoritmo Base de Súper-Resolución	51
3.3.- Filtros selectivos y topología de Macro-Bloques adaptable	65
3.4.- Entorno Multi-Cámara	75
3.5.- Conclusiones	84

*Lo bueno necesita aportar pruebas; lo bello, no
Bernard Le Bouvier de Fontenelle, escritor francés (1657 – 1757)*

3.1 INTRODUCCIÓN

Una vez presentados los conceptos necesarios para comprender el ámbito de aplicación de la presente Tesis Doctoral, en este capítulo se introducen las contribuciones algorítmicas desarrolladas. Las aportaciones pasan por dos grandes bloques: el primero de ellos está basado en una etapa de filtro y determinación del tamaño de Macro-Bloque más adecuado en el proceso de Súper-Resolución, mientras que el segundo consiste en describir cómo se pueden integrar las técnicas de Súper-Resolución por fusión en un sistema basado en múltiples cámaras (sistema Multi-Cámara). La Figura 3.1 presenta un esquema general de los principales conceptos que se integran en la Tesis Doctoral.

La etapa de filtro y determinación del tamaño de Macro-Bloque consiste en un filtro inteligente capaz de seleccionar los *frames* más adecuados a usar en el proceso de SR en una primera aproximación con el fin de optimizar el resultado y reducir el tiempo de cómputo (Filtro Selectivo de Ventana, *Window Selective Filter* – WSF) hasta llegar a determinar si se debe o no utilizar un determinado MB en el proceso de SR (Filtro Selectivo de Macro-Bloque, *Block Selective Filter* – BSF). Por último, se plantea que cada Macro-Bloque pueda dividirse generando topologías que sean capaces de maximizar, tanto la calidad objetiva como la calidad subjetiva (topología de Macro-Bloques adaptable basada en un tamaño de MB variable, *Variable Block Size* – VBS).

La integración de las técnicas de Súper-Resolución por fusión en sistemas Multi-Cámara pasa por la presentación de varios métodos que explotan la información temporal (*frames* capturados en distintos instantes de tiempo por una cámara determinada) y espacial (*frames* tomados en un mismo instante de tiempo por cámaras diferentes), conformando los métodos Temporal-Espacial, Espacial-Temporal y Mixto, así como un conjunto de modos de pre-procesamiento para adecuar la información tomada como punto de partida.

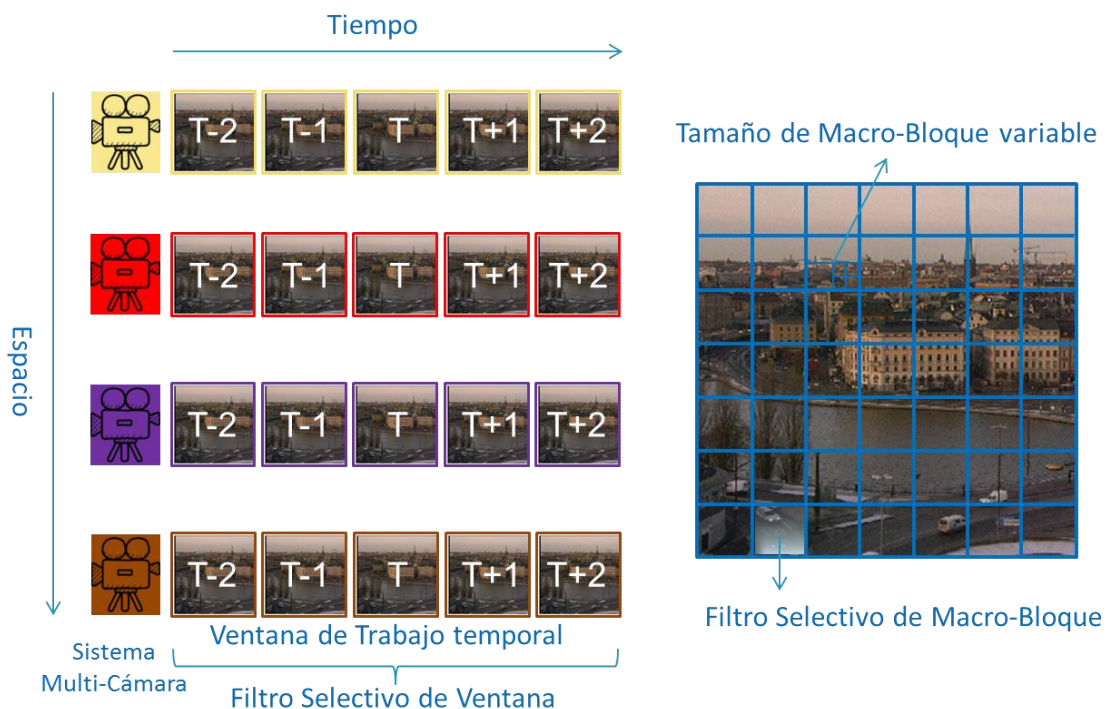


Figura 3.1.- Esquema general de las contribuciones de la presente Tesis Doctoral

3.2 ALGORITMO BASE DE SÚPER-RESOLUCIÓN

El algoritmo base de SR tomado como referencia en la presente Tesis Doctoral (BSR, *Baseline Super-Resolution*) ha sido desarrollado en el seno de la división de Diseño de Sistemas Integrados (DSI) del Instituto Universitario de Microelectrónica Aplicada (IUMA) [Mar03, CLN+02] de la Universidad de Las Palmas de Gran Canaria (ULPGC). El algoritmo se encuadra dentro de la categoría de fusión, más concretamente dentro de la fusión mediante SR dinámica, como se presenta en la Figura 3.2, ofreciendo resultados de alta calidad siempre y cuando la entrada al sistema sea una secuencia correlacionada de imágenes con vectores de movimiento relativamente pequeños en comparación con el tamaño de las imágenes. Lo común es aplicar este tipo de procesos a secuencias filmadas, como por ejemplo una grabación realizada por una cámara de vídeo estándar [BCL+05].



Figura 3.2.- Diagrama de bloques de algoritmo de SR dinámica tomado como referencia

3.2.1 Pseudo-código

En la Figura 3.3 se detalla el pseudo-código del algoritmo base de SR. El algoritmo se ejecuta para cada uno de los *frames* de la secuencia de entrada, tal y como se representa en la Figura 3.2.

El algoritmo BSR genera para cada *frame* una Ventana de Trabajo asociada, donde se almacenan los n *frames* anteriores y posteriores al actual. A continuación, se divide el *frame* actual en Macro-Bloques (MB) de tamaño fijo y se ejecuta el proceso de *Estimación de Movimiento* entre dicho *frame* y cada uno de los *frames* que se encuentran en la Ventana de Trabajo, buscando el mejor vector dentro del Área de Búsqueda (*Search Area*) correspondiente. Una vez hallados todos los vectores de movimiento necesarios, se ejecuta el proceso de *Compensación de Movimiento* generando una imagen compensada de resolución VHR (*Very High Resolution*) de acuerdo con el Factor de Escala (*Scale Factor*) que se haya definido entre LR y VHR. Sobre dicha imagen compensada se realiza el proceso de *Rellenado de Huecos*, cubriendo los huecos para dejar la imagen completamente tratada (imagen rellenada) previo a un paso final de diezmado, que proporciona como resultado la imagen en HR (*High Resolution*) súper-resuelta que se almacenará.

```

for frame actual = primer frame: último frame
  Ventana de Trabajo = Generar Ventana de Trabajo del frame actual
  for frame de referencia = frame inicial de Ventana de Trabajo: frame final de Ventana de Trabajo
    for fila de Macro-Bloque = 1 : filas de Macro-Bloque
      for columna de Macro-Bloque = 1 : columnas de Macro-Bloque
        Macro-Bloque actual = Recortar Macro-Bloque de fila y columna
        vector de movimiento en el frame de referencia con fila y columna =
          Estimación de Movimiento en frame de referencia del Macro-Bloque
          en fila y columna con Área de Búsqueda SA y algoritmo X

  imagen compensada = Compensación de Movimiento de Ventana de Trabajo con vectores de movimiento
  imagen rellenada = Rellenado de Huecos de (1/Factor de Escala) Pixel de imagen compensada
  frame actual de imagen súper-resuelta = Diezmar imagen rellenada
  almacenar frame actual de imagen súper-resuelta

almacenar secuencia
    
```

Figura 3.3.- Proceso de BSR descrito en pseudo-código

3.2.2 Descripción de bloques del algoritmo

3.2.2.1 Estimación de Movimiento

El proceso de *Estimación de Movimiento* se realiza sobre los *frames* de la Ventana de Trabajo. Para ello se divide el *frame* actual en Macro-Bloques (MBs) de tamaño $M \times N$ píxeles y se realiza una *Estimación de Movimiento sub-píxel* sobre cada uno de ellos. Una vez obtenidos los vectores de movimiento con menor distorsión para cada *frame* dentro de la Ventana de Trabajo, se procede a realizar la *Compensación de Movimiento*. En la Figura 3.4 se muestra el proceso de *Estimación de Movimiento* partiendo de la Ventana de Trabajo y obteniendo como resultado los vectores de movimiento, por MB. El algoritmo BSR soporta dos algoritmos de búsqueda: la búsqueda exhaustiva o *Full-Search* y la búsqueda de tres pasos TSS (*Three Steps Search*) [CLT+08].

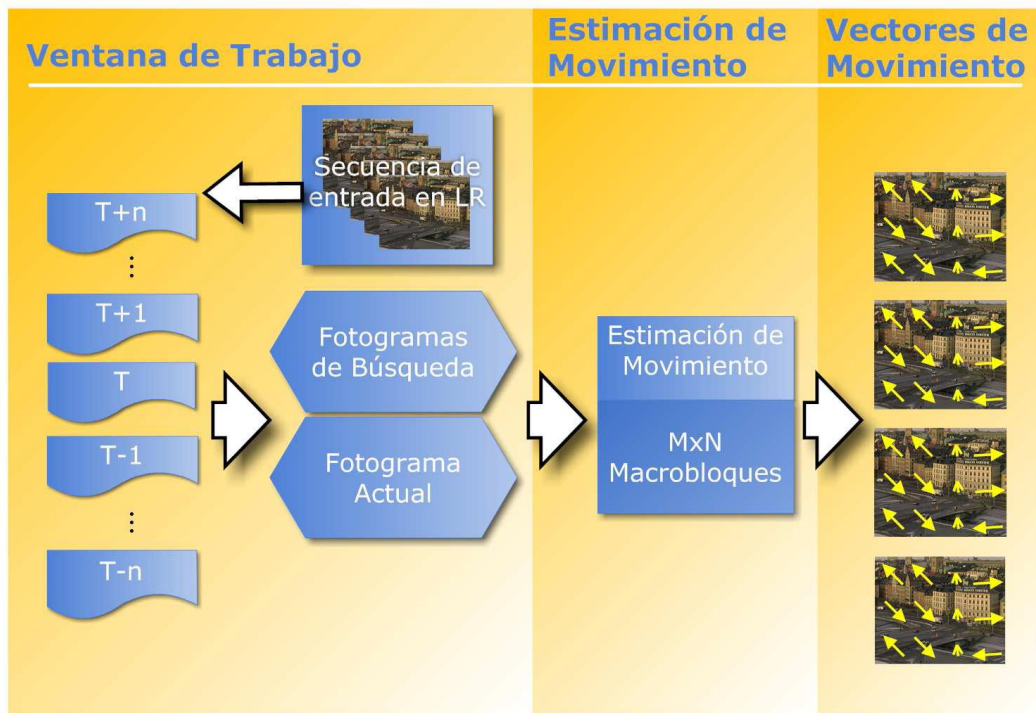


Figura 3.4.- Proceso de Estimación de Movimiento

3.2.2.2 Compensación de Movimiento

El proceso de *Compensación de Movimiento* consiste en generar una malla de resolución VHR que viene determinada por la precisión *sub-píxel* deseada, siendo dicha malla de tamaño igual a $2^{\text{precisión}}$ veces el *frame* de LR. Esto es debido a que la nueva malla de resolución VHR tiene que poder alojar el desplazamiento generado por un vector de movimiento con precisión $1/2^{\text{precisión}}$ de píxel. Dicha malla estará formada en principio por huecos o píxeles vacíos. En el proceso de *Compensación de movimiento* se rellenan esos huecos con píxeles súper-resueltos que provengan de los *frames* de la Ventana de Trabajo procesada.

Como mínimo se podrán cubrir los huecos de los píxeles de LR que provienen del *frame* actual y cuyo vector de movimiento es nulo. En la Figura 3.5 se representa la malla de resolución VHR para un Factor de Escala de 2, donde se aprecian los distintos tipos de posiciones. Las líneas gruesas representan posiciones de precisión entera. Esto quiere decir que para colocar un píxel sobre ellas, el vector de movimiento ha de presentar en alguna de sus coordenadas un valor entero. Los píxeles correspondientes al *frame* actual o de baja resolución (color azul), presentan un vector de movimiento nulo, con lo que en la malla de resolución VHR para una precisión de cuarto de píxel, se colocan cada cuatro píxeles.

La línea más fina y continua representa posiciones de medio píxel de precisión, es decir, para un píxel colocado sobre ella el vector de movimiento asociado ha de presentar en alguna de sus componentes un valor de medio píxel. Por último, la línea discontinua representa las posiciones de un cuarto de píxel. Los píxeles en amarillo representan los píxeles compensados. En su interior se lee un número que representa el *frame* del que proviene, cuya correspondencia se presenta en la Tabla 3.1.

Tabla 3.1.- Píxeles compensados y vectores de movimiento asociados

Nº	Vector de Movimiento
-	(0, 0)
1	(0.5, 0.25)
2	(0, 0.5)
3	(0.25, 0.75)

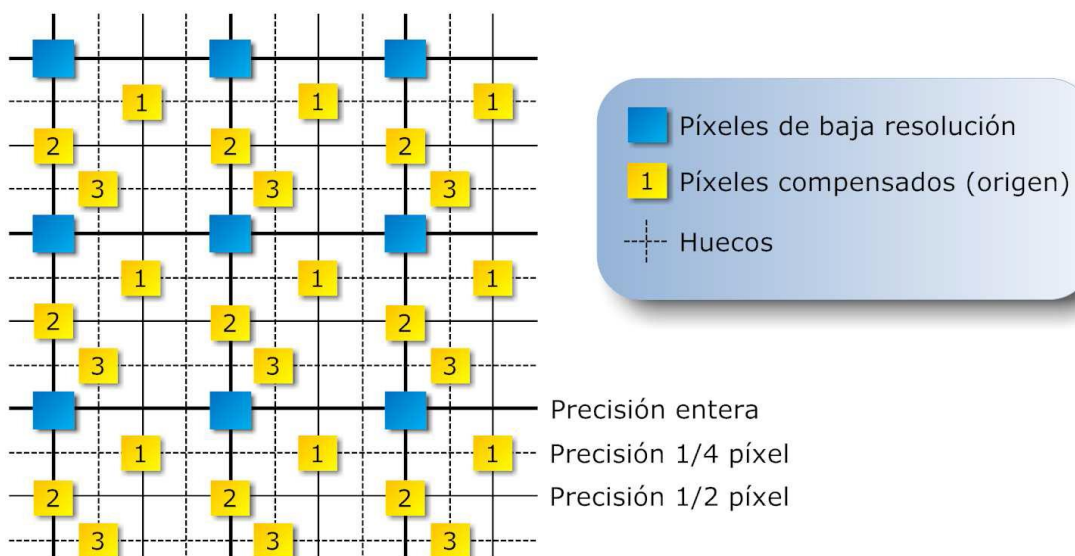


Figura 3.5.- Proceso de Compensación de Movimiento

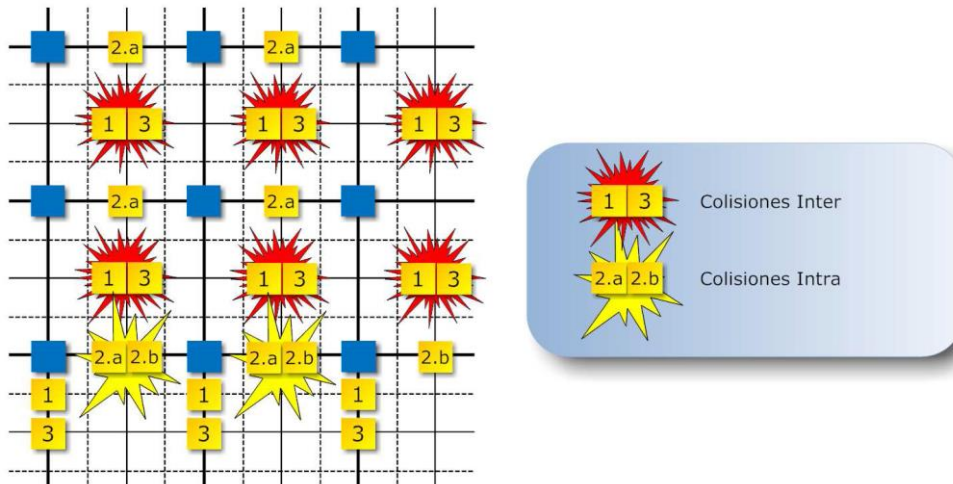


Figura 3.6.- Ejemplo de colisiones en la malla de VHR

En el proceso de *Compensación de Movimiento* puede ocurrir que un MB tenga el mismo vector de movimiento para dos *frames* diferentes, por lo que ambos *frames* intentarían colocar un píxel en la misma posición de la malla de VHR, lo que se denomina colisión entre *frames* o colisión *inter*. Esta situación también puede producirse entre dos MBs diferentes de un mismo *frame*. A este tipo de colisión de MBs se le denomina colisión *intra*. El método más común para solucionar las colisiones es usar la media aritmética de los diferentes valores de píxel. En la Figura 3.6 se muestran ambos casos: marcadas en rojo se representan las colisiones *Inter Frame*, en las que los *frames* 1 y 3 pretenden establecer un valor de píxel en la misma posición. En color amarillo se representan las colisiones *Intra Frame*, donde se observa cómo dos MBs, *a* y *b*, del *frame* 2 pretenden establecer en la misma posición el valor de un mismo píxel.

Por último, dentro de la *Compensación de Movimiento* hay que considerar que los MBs de los diferentes *frames* sólo rellenan los huecos que estén situados dentro del mismo MB en la malla de VHR. En la Figura 3.7 se muestran las posiciones de la malla VHR, en la que, en las posiciones sombreadas en rojo, es posible colocar un píxel súper-resuelto, ya que coincide con el MB actual. Los cuadros azules son los píxeles de LR, y representan el valor del MB al que pertenecen, mientras que en amarillo se representan los píxeles compensados. Los píxeles compensados, tachados con una cruz roja, son los que se encuentran fuera del MB actual y por tanto no se establece un valor sobre la malla de VHR para este MB.

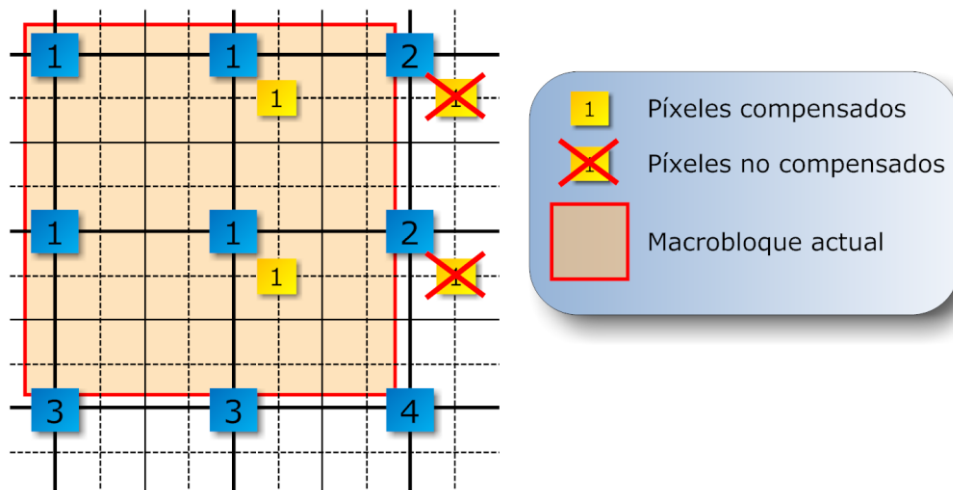


Figura 3.7.- Compensación de Movimiento por MB

3.2.2.3 Rellenado de Huecos

La imagen súper-resuelta se obtiene al diezmar la malla de resolución VHR. Si se diezmas directamente la malla tras realizar el proceso de *Compensación de Movimiento*, existe la posibilidad de que haya píxeles sobre los que no se ha podido colocar ningún píxel súper-resuelto. Estos píxeles son los huecos, y hay que rellenarlos.

Para realizar el *Rellenado de Huecos* se lleva a cabo una interpolación bilineal de los píxeles súper-resueltos próximos al píxel a interpolar. Como referencia, en la Figura 3.8.a se representan los píxeles compensados en color amarillo, mientras que en la Figura 3.8.b se muestran los píxeles compensados en color amarillo y los píxeles interpolados en color verde. Además se representa el área de interpolación para cada uno de ellos enmarcada en rojo, donde se muestran los píxeles que se han tenido en cuenta para la interpolación, sombreados en rojo.

Una vez completado el proceso de *Rellenado de Huecos* se diezma la imagen a HR, constituyendo la imagen súper-resuelta final. Para el *Rellenado de Huecos* la interpolación bilineal utiliza los píxeles que se encuentren en un área de $(2^{\text{precisión}} + 1)$ centrado en el píxel a interpolar, puesto que representa un buen compromiso entre calidad y coste computacional [Ban09].

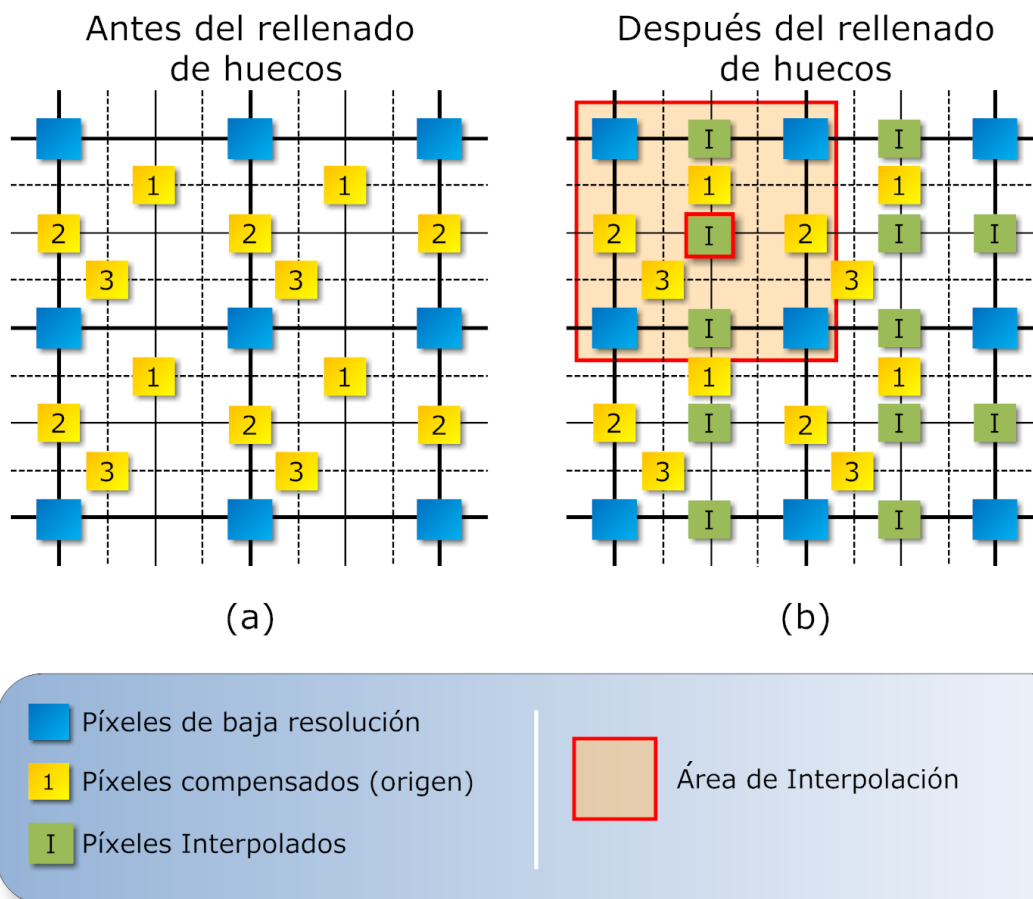


Figura 3.8.- Malla de VHR
 a) Antes del Rellenado de Huecos
 b) Después del Rellenado de Huecos

3.2.3 Parámetros del algoritmo BSR

3.2.3.1 Descripción de parámetros

El algoritmo BSR incluye un conjunto de parámetros de configuración, cuyo valor se selecciona dependiendo de las características de la secuencia de entrada de LR que se esté procesando y de la resolución de la secuencia / imagen de HR que se desee generar. En esta sección se describen estos parámetros, así como los posibles valores que pueden tomar:

- **VIDEO_OUTPUT = 0/1:** Este *flag* determina si la salida súper-resuelta será un único *frame* (0) o una secuencia de vídeo (1) del mismo número de *frames* que la secuencia de entrada.
- **REF_FRAME = #FRAME:** En el caso de tomar como salida un *frame* (**VIDEO_OUTPUT = 0**), se ha de proporcionar el número de *frame* que se usará como referencia. El movimiento de los otros *frames* en la secuencia se ajustará a la posición del número de *frame* proporcionado. El número de *frame* ha de estar incluido en el rango de los *frames* especificados. En el caso de tomar como salida una secuencia de vídeo, el *frame* de referencia será alternativamente cada *frame* de entrada, a los cuales se les denomina *frames* de referencia actuales (*current reference frame*).
- **SCALE_FACTOR = #SCALE:** Factor de Escala que será aplicado a la secuencia de entrada para generar la secuencia súper-resuelta. Se han probado valores entre 2 y 10, si bien en la presente Tesis Doctoral se usarán valores de 2 y 4, como se presentará más adelante.
- **SEARCH_AREA = #SA:** El Área de Búsqueda determina el número de píxeles alrededor de cada Macro-Bloque (MB) que se considera en el proceso de búsqueda del vector de movimiento adecuado. Para escenas con mucho movimiento se requieren altos valores de SA. Para escenas con poco movimiento, elevados valores de SA son contraproducentes, en el sentido de que no incrementarán en gran medida la calidad y sin embargo darán lugar a una gran cantidad de carga computacional ineficiente. Valores entre 8 y 16 son normalmente suficientes para la mayoría de las secuencias.
- **MB_SIZE_LR = #MB:** Tamaño de los MBs de baja resolución en los que se dividirán las imágenes de la secuencia de entrada. Bajos valores de MB permiten disponer de un mayor número de vectores de movimiento independientes, pero esto incrementa significativamente la carga computacional. Valores elevados de MB incrementan la coherencia de los vectores de movimiento pero degradan la calidad si se consideran varios objetos con diferentes movimientos dentro de un mismo MB. Tamaños de MB entre 4 y 32 son normalmente suficientes para la mayoría de las secuencias.
- **MAX_PIX_DIF = #MPD:** Esta es la máxima diferencia de píxel (en valor absoluto) permitida entre un píxel candidato de otro *frame* y el valor interpolado en el *frame* de referencia actual. Valores elevados de MPD implican que se combinarán más píxeles, lo cual incrementará potencialmente la calidad de la SR si existe correlación temporal entre *frames*, pero al mismo tiempo, la contrapartida es que podrían aparecer más artefactos. Bajos valores de MPD evitan en gran medida los artefactos, pero también limitan las mejoras de la SR. Para 8 bits, el máximo valor es 255 (todo SR) y el mínimo valor es siempre cero (todo interpolación).
- **WIN_BACK_FRAMES = #WBF:** Número de *frames* temporalmente localizados antes del *frame* de referencia actual en la secuencia de vídeo de entrada, que se usarán para combinarse con los píxeles del *frame* de referencia actual. Valores altos de WBF incrementan la probabilidad de obtener mejoras con las técnicas de SR, pero al mismo tiempo también incrementan la probabilidad de generar artefactos si los *frames* están poco correlacionados.

- $WIN_FORWARD_FRAMES = \#WFF$: Número de *frames* temporalmente localizados después del *frame* de referencia actual en la secuencia de vídeo, que se usarán para combinarse con los píxeles del *frame* de referencia actual. Valores altos de WFF incrementan la probabilidad de obtener mejoras con las técnicas de SR, pero al mismo tiempo incrementan la probabilidad de generar artefactos si los *frames* están poco correlacionados.
- $KEEP_TMP_FILES = 0/1$: Si este valor se fija a 1, se almacenarán varios archivos temporales con el objetivo de realizar un proceso de depuración.
- $SILENT = 0/1$: Si este valor se fija a 1, no se presentarán mensajes durante la ejecución.
- $KEEP_TRACK = 0/1$: Si este valor se fija a 0, el algoritmo se comporta tal y como se ha descrito previamente. Si se fija a 1, se generará una copia del archivo de configuración con un nombre que incluye una extensión que reflejará los parámetros usados. Las secuencias súper-resuelta e interpolada presentarán el mismo nombre de extensión. Tal extensión tendrá, por ejemplo, el formato que se indica a continuación: FF10_LF25_SC2_SA8_MB16_MPD20_WB2_WF2. El significado de cada *substring* es:
 - FF: Primer *frame* (*First Frame*, 10 en el ejemplo).
 - LF: Último *frame* (*Last Frame*, 25 en el ejemplo).
 - SC: Factor de Escala (*Scale Factor*, 2 en el ejemplo).
 - SA: Área de Búsqueda (*Search Area*, 8 en el ejemplo).
 - MB: Tamaño de macro-bloque (*Macro-Block size*, 16 en el ejemplo).
 - MPD: Máxima Diferencia de Píxel (*Maximum Pixel Difference*, 20 en ejemplo).
 - WBF: Tamaño de ventana hacia atrás (*Window Back Frames*, 2 en el ejemplo).
 - WFF: Tamaño de ventana hacia adelante (*Window Forward Frames*, 2 en el ejemplo).
- $FIRST_FRAME = \#FF$: Si se asigna un valor de -1, se usarán todos los *frames* de la secuencia (sea cual sea el valor de $LAST_FRAME$). En caso de asignar un valor positivo, este valor se usará como primer *frame*.
- $LAST_FRAME = \#LF$: Si se asigna un valor de -1, se usarán todos los *frames* de la secuencia (sea cual sea el valor de $FIRST_FRAME$). En caso de asignar un valor positivo, este valor se usará como último *frame*.

Como referencia, en la Figura 3.9 se representan gráficamente los parámetros $SCALE_FACTOR$, MB_SIZE_LR , $SEARCH_AREA$, WIN_BACK_FRAMES y $WIN_FORWARD_FRAMES$.

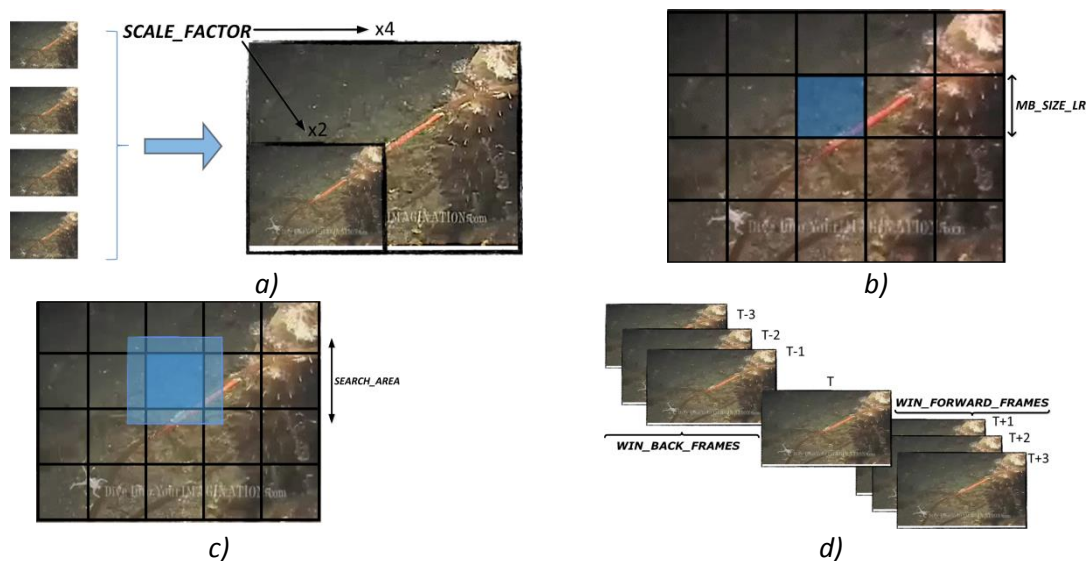


Figura 3.9.- Parámetros de SR

a) $SCALE_FACTOR$, b) MB_SIZE_LR , c) $SEARCH_AREA$ y d) WIN_BACK y $WIN_FORWARD_FRAMES$

3.2.3.2 Parametrización del algoritmo BSR

La calidad del *frame* o de la secuencia de vídeo resultante del proceso de SR depende del conjunto de parámetros utilizado. Los parámetros deben ajustarse a las características de la entrada a súper-resolver. Si los parámetros no son los apropiados, la calidad de la secuencia de salida se podrá degradar significativamente. Los parámetros que afectan principalmente a la calidad del *frame* o de la secuencia súper-resuelta son:

- El **Factor de Escala (*SCALE_FACTOR*)**: A medida que se incrementa su valor se requiere número mayor de *frames* a la entrada. Idealmente, si cada *frame* contiene un desplazamiento sub-píxel diferente, para un Factor de Escala de 2, serán necesarios 4 *frames* (medio píxel a la izquierda y a la derecha, y medio píxel arriba y abajo). Para un Factor de Escala de 3, serán necesarios 9 *frames* (con desplazamientos de precisión 1/3 de píxel) y así sucesivamente, de tal forma que el número de *frames* idealmente necesarios se calcula como la potencia cuadrática del Factor de Escala. En el último caso, usando un Factor de Escala de 10, se necesitarían 100 *frames*. En una situación real, es prácticamente imposible asegurar que todos los desplazamientos sub-píxel estén disponibles (salvo que un sistema mecánico proporcione los desplazamientos) y por tanto, serían necesarios *frames* adicionales para obtener estadísticamente tanta información como fuera posible. El número de *frames* a combinar dependerá de los *frames* disponibles y del movimiento entre *frames*.
- El **Área de Búsqueda (*SEARCH_AREA*)**: Se permite cualquier valor por debajo del tamaño del *frame*. Valores elevados aseguran que los vectores de movimiento se tomarán adecuadamente, a expensas de contar con una carga computacional mucho más elevada. Valores más bajos aceleran la ejecución, pero ponen en compromiso el resultado si existe mucho movimiento entre *frames*. En tal caso, los vectores de movimiento no se calcularán apropiadamente.
- El **tamaño de Macro-Bloque de baja resolución (*MB_SIZE_LR*)**: Se permite cualquier valor por debajo del tamaño del *frame*. Se prefieren valores grandes de tamaño de MB para movimiento global y valores más pequeños para movimiento local. Idealmente, el tamaño de MB debería coincidir con los objetos presentes en la escena.
- El **método *Block-Matching* usado en el proceso de *Estimación de Movimiento***: Se permiten *Full-Search* (FS), *New Three Steps* (NTS) o vectores de movimiento desde un archivo. Los vectores de movimiento desde archivo son sólo para *test*. NTS es más rápido que FS, pero el vector de movimiento no se obtendrá con una distorsión mínima. Por otra parte, FS siempre obtiene el vector de movimiento con menor distorsión, a expensas de un coste computacional más elevado.
- El **número de *frames* hacia atrás o hacia adelante en la Ventana de Trabajo de SR (*WIN_BACK_FRAMES* y *WIN_FORWARD_FRAMES*)**: Estos dos valores están sólo limitados por el número de *frames* disponibles en la secuencia de entrada. Debe considerarse no incluir *frames* no correlacionados dentro de la Ventana de Trabajo.
- La **Máxima Diferencia de Píxel (*MAX_PIX_DIF*)**: Este parámetro constituye un umbral para evitar la inclusión de píxeles no correlacionados en el *frame* súper-resuelto. Un píxel candidato (de otro *frame* diferente al actual) se incluirá en la nueva rejilla de SR si su valor es próximo al correspondiente píxel interpolado. En otras palabras, el píxel candidato se incluirá si la diferencia entre éste y el píxel interpolado está por encima de un umbral denominado MPD (*Maximum Pixel Difference*). Un MPD nulo hará que todos los píxeles sean interpolados, evitando de esta forma artefactos, pero eliminando también las mejoras del proceso de SR. Por el contrario el valor máximo considerando 8 bits (255) permitirá usar todas las contribuciones de la SR, pero con el riesgo de combinar píxeles inadecuados y producir artefactos.

3.2.4 Limitaciones del algoritmo BSR

Una vez explicado el algoritmo BSR y los parámetros que lo caracterizan, se presentan en este apartado las principales limitaciones que presenta el algoritmo BSR y, en general, los algoritmos que se encuadran dentro de la categoría de fusión. Se describirán en este apartado dos conceptos básicos que complementan los parámetros del algoritmo BSR recientemente introducidos, y su caracterización en la presente Tesis Doctoral; estos son la información temporal y la información espacial.

3.2.4.1 Información temporal

La **información temporal** se refiere a los *frames* captados en diferentes instantes de tiempo por una cámara de vídeo. En el algoritmo BSR la información temporal se identifica directamente con la Ventana de Trabajo *WW* (*Working Window*), relacionada con los parámetros *WIN_BACK_FRAMES* y *WIN_FORWARD_FRAMES*, que determinan el tamaño de la Ventana de Trabajo considerando el *Frame Actual* (*CF*, *Current Frame*), según la ecuación 3.1.

$$WW = WIN_BACK_FRAMES + WIN_FORWARD_FRAMES + 1 \quad (3.1)$$

La Figura 3.10 representa cómo se forma la Ventana de Trabajo en el algoritmo BSR. En ésta se presenta una secuencia de 11 frames y se supone que $CF = 6$, así como que la Ventana de Trabajo es de ± 5 ($\#WBF = 5$ y $\#WFF = 5$).

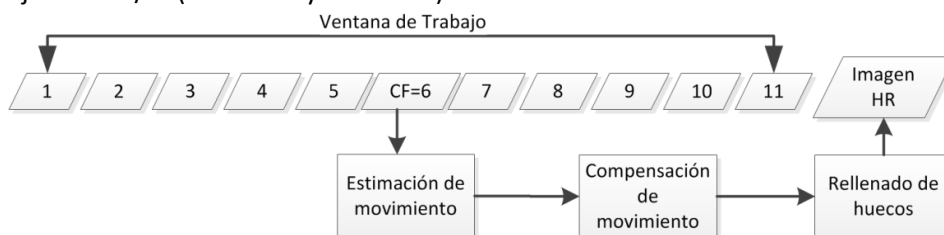


Figura 3.10.- Ejemplo de Ventana de Trabajo de ± 5 ($\#WBF = 5$ y $\#WFF = 5$) con $CF = 6$

La determinación de la Ventana de Trabajo tomada como punto de partida en los algoritmos de fusión, que obtienen su resultado a partir de los *frames* anteriores y posteriores al *frame* actual, es fundamental para obtener resultados adecuados a la salida del algoritmo en forma de una imagen o una secuencia de vídeo de alta resolución. Valores elevados del tamaño de Ventana de Trabajo incrementan la probabilidad de obtener mejoras con las técnicas de SR, pero al mismo tiempo aumentan también la probabilidad de generar artefactos en el caso de que los *frames* estén poco correlacionados. Para comprender mejor este concepto se presentan a continuación tres ejemplos con tres de las secuencias consideradas en la presente Tesis Doctoral, que cubren casos muy generales en lo que se refiere a movimiento global y local, y cuyas características principales se pueden encontrar en el Anexo I, las cuales son *Galdar*, *Mobcal* y *Reef*. Mientras que la secuencia *Galdar* presenta un movimiento global puro, basado en el movimiento de una cámara en un plano fijo, la secuencia *Mobcal* combina tanto movimiento global como movimiento local de un calendario desplazándose hacia arriba y un tren moviéndose hacia la izquierda. Por último, la secuencia *Reef* presenta el caso de una secuencia real grabada en un entorno submarino que presenta movimiento local. La secuencia *Galdar* presenta unas características muy adecuadas para los algoritmos de fusión, proporcionando mejores resultados cuando la secuencia de partida tiene movimiento global, y no tanto con movimiento local. Para contrastar gráficamente este hecho se presentan a continuación varias gráficas obtenidas a partir del algoritmo BSR que relacionan la calidad objetiva de la imagen y el tiempo de cómputo, con el tamaño de la Ventana de Trabajo de SR.

La Figura 3.11 muestra los resultados correspondientes a la secuencia *Galdar*. Las Figuras 3.11.a y 3.11.b presentan cómo el uso de un mayor número de *frames* propicia un incremento de la calidad objetiva (el valor en el eje X muestra el valor de WBF = WFF considerando que la Ventana de Trabajo $WW = WBF + WFF + 1$, de acuerdo con la expresión 3.1). Esto implica como contrapartida incrementar el tiempo de cómputo, tal y como se representa en la Figura 3.11.c. De forma subjetiva, esto se puede apreciar claramente en la formación de las letras de la imagen, que se puede observar cómo mejora en las Figuras 3.11.d, 3.11.e y 3.11.f, que representan la secuencia interpolada y las secuencias súper-resueltas desde $WW = +/-1$ a $WW = +/- 14$, respectivamente.

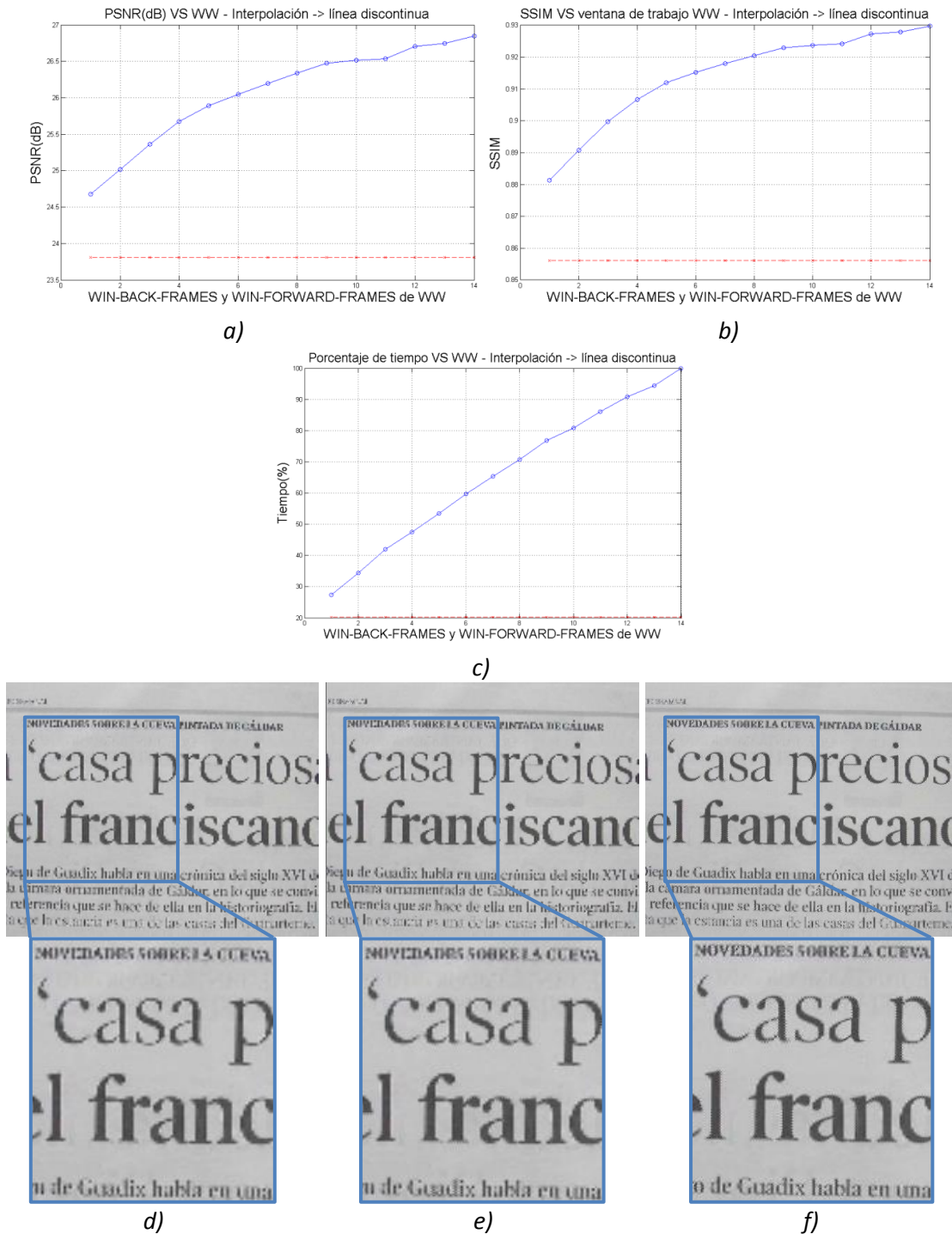


Figura 3.11.- Resultados con secuencia *Galdar*. a) PSNR VS WW, b) SSIM VS WW, c) Tiempo de cómputo vs WW, d) Frame interpolado, e) Frame SR - WW = +/-1, f) Frame SR - WW = +/-14

La Figura 3.12 muestra los resultados correspondientes a la secuencia *Mobcal*. Las Figuras 3.12.a y 3.12.b presentan cómo en este caso el uso de un mayor número de *frames* no propicia un incremento de la calidad objetiva. Se observa que existe un pico de calidad tanto en términos de PSNR como de SSIM cuando se usa $WW = +/-3$, y que en el caso de haber usado una Ventana de Trabajo de mayor tamaño, como $WW = +/-14$, la calidad se habría visto muy deteriorada debido a la aparición de artefactos. De forma subjetiva esto se puede apreciar claramente en el desplazamiento del tren, apareciendo artefactos en las Figuras 3.12.e y 3.12.f, que representan las secuencias súper-resueltas con $WW = +/-3$ y $WW = +/-14$, respectivamente, frente al *frame* interpolado de la Figura 3.12.d, que aunque no presenta artefactos, proporciona una calidad subjetiva global inferior.

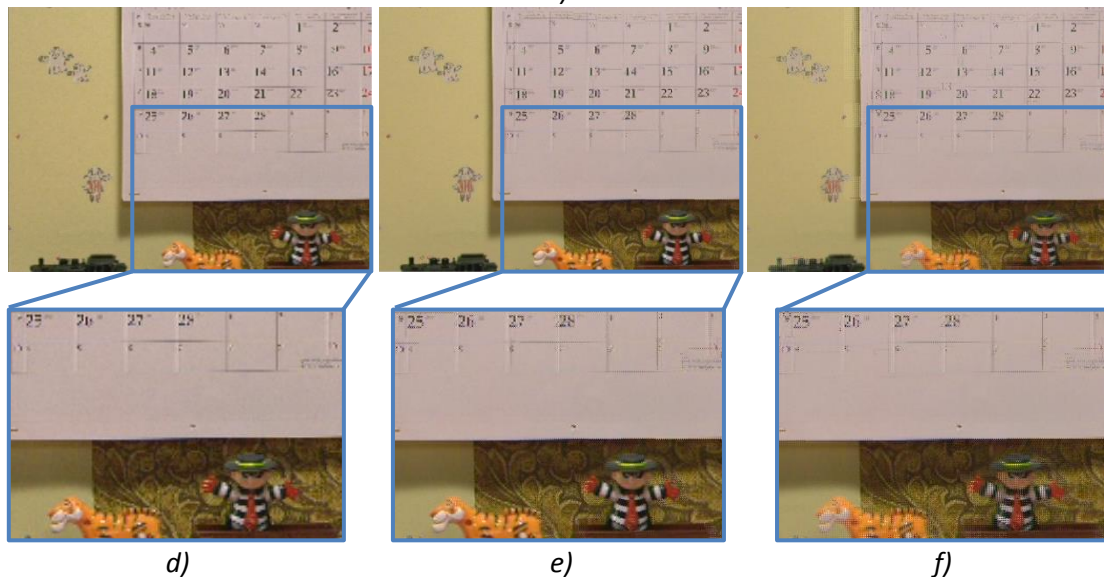
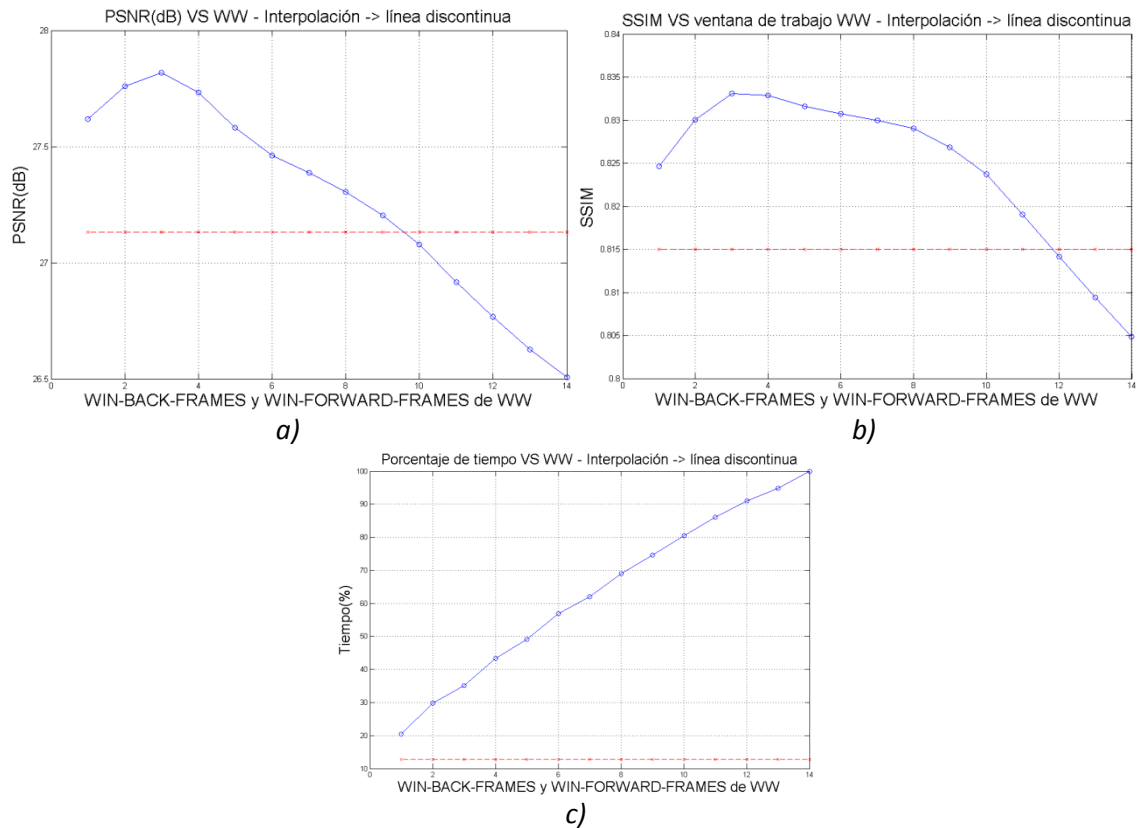


Figura 3.12.- Resultados con secuencia *Mobcal*. a) PSNR VS WW, b) SSIM VS WW, c) Tiempo de cómputo vs WW, d) *Frame* interpolado, e) *Frame* SR - $WW = +/-3$, f) *Frame* SR - $WW = +/-14$

Finalmente, la Figura 3.13 muestra los resultados correspondientes a la secuencia submarina *Reef*, basada en movimiento local. Las Figuras 3.13.a y 3.13.b presentan cómo es necesario usar un número de *frames* adecuado para obtener una calidad objetiva suficiente. Se observa que existe un pico de calidad tanto en términos de PSNR como de SSIM cuando se usa $WW = +/-5$, y que en el caso de haber usado una Ventana de Trabajo mayor, como por ejemplo $WW = +/-14$, la calidad se habría visto muy deteriorada debido a la aparición de artefactos. Así mismo, el uso de una Ventana de Trabajo de pequeño tamaño, como $WW = +/-1$, no sería suficiente ni siquiera para mejorar los resultados en PSNR correspondientes a la interpolación. De forma subjetiva, esto se puede apreciar claramente en el desplazamiento del pez de la secuencia, apareciendo artefactos en las Figuras 3.13.e y 3.13.f que representan la secuencias súper-resueltas con $WW = +/-5$ y $WW = +/-14$, respectivamente, frente al *frame* interpolado representado en la Figura 3.13.d, que aunque no presenta artefactos, proporciona una calidad subjetiva global inferior como en el caso de la secuencia *Mobcal*.

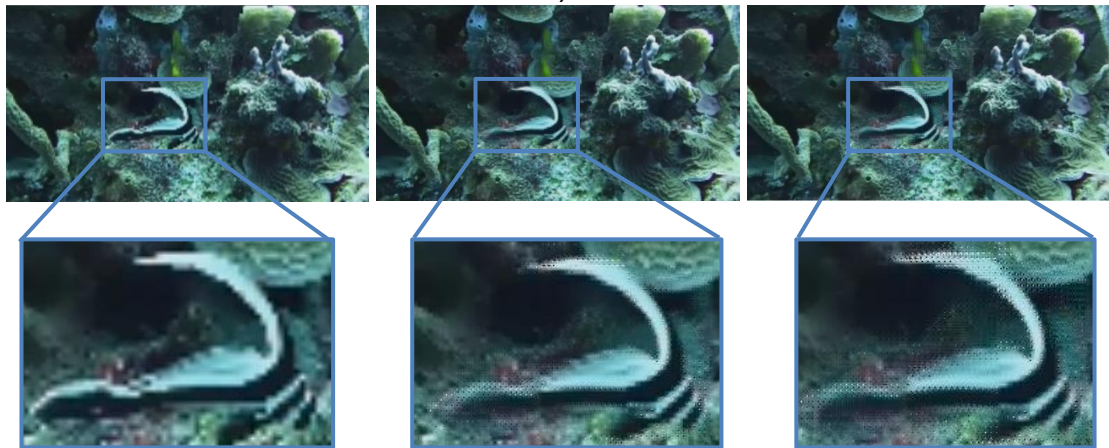
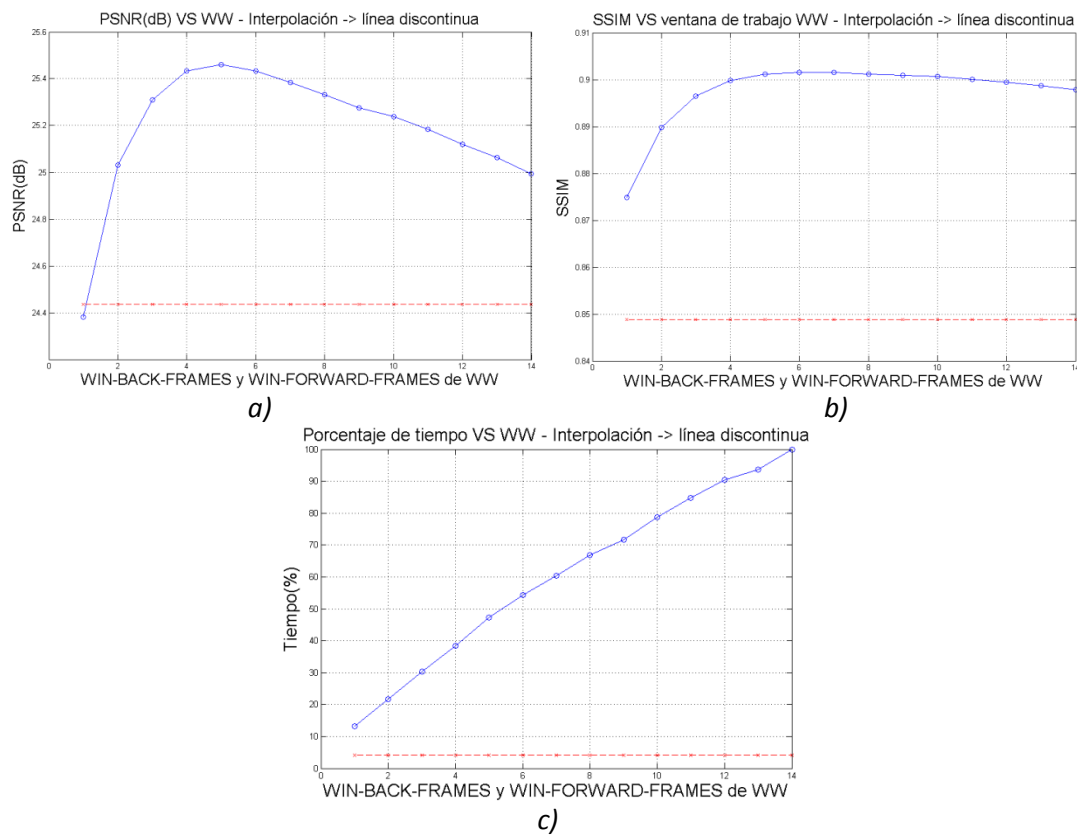


Figura 3.13.- Resultados con secuencia *Reef*. a) PSNR VS WW, b) SSIM VS WW, c) Tiempo de cómputo vs WW, d) *Frame* interpolado, e) *Frame* SR - $WW = +/-5$, f) *Frame* SR - $WW = +/-14$

A partir del análisis realizado sobre un amplio rango de secuencias –del que se han presentado como referencias los resultados mostrados en las Figuras 3.11, 3.12 y 3.13–, se pueden extraer las siguientes conclusiones:

- Las características de la secuencia a procesar son fundamentales a la hora de determinar el tamaño de la WW a utilizar en el proceso de SR. Cada uno de los *frames* considerados en el proceso de SR aportará información que aumentará o reducirá la calidad del *frame* súper-resuelto a la salida. Así, por ejemplo, el uso de WW = +/-9 con respecto al uso de un valor WW = +/-11 no supone un salto de calidad significativo en la secuencia *Galdar*, pero sí se observa una mejora importante al usar WW = +/-12 (ver Figuras 3.11.a y 3.11.b). Así mismo, la calidad objetiva de la secuencia *Mobcal* presenta un pico para WW = +/-3 y el uso de un mayor número de *frames* introduce información que no presenta correlación espacial adecuada, empeorando por tanto los resultados (ver Figuras 3.12.a y 3.12.b). De forma similar ocurre para la secuencia *Reef* a partir de un valor WW = +/-5 (ver Figuras 3.13.a y 3.13.b)
- El tiempo de cómputo aumenta de forma proporcional con el tamaño de la Ventana de Trabajo de SR. Las Figuras 3.11.c, 3.12.c y 3.13.c, que representan el porcentaje del tiempo de cómputo con respecto a un máximo de WW = +/-14 muestran claramente (sobre todo desde WW = +/- 8 a WW = +/- 14) cómo el tiempo de cómputo disminuye en un 10% al reducir el valor de WW en 2 *frames* posteriores y 2 *frames* anteriores (por ejemplo al pasar de WW = +/- 14 a WW = +/-12). Este punto, conjuntamente con el anterior, invitan a pensar que una **selección apropiada de los frames a utilizar en el proceso de SR**, no sólo mantendría o incluso mejoraría la calidad, sino que también disminuiría el tiempo de cómputo del proceso de SR.
- El algoritmo BSR funciona a nivel de MBs, por lo que en el proceso de SR de un *frame* podrá haber MBs en los que el proceso de SR obtenga buenos resultados, mientras que en otros MBs se introduzcan artefactos que disminuyan por tanto la calidad global del *frame* en cuestión. Este aspecto se ha de considerar sobre todo en secuencias que presenten movimiento local. Una vez más, la **selección apropiada de los MBs a utilizar en el proceso de SR** podría implicar efectos beneficiosos tanto en la calidad de la secuencia como en el tiempo de cómputo. En este caso también entra en juego el **tamaño de MB** considerado en el proceso de SR.
- Las métricas PSNR y SSIM presentan un comportamiento muy similar en la determinación de la calidad objetiva de una secuencia. Sin embargo, la aparición de artefactos en un *frame* provoca una caída de calidad mayor en PSNR que en SSIM. Esto se puede apreciar claramente en las Figuras 3.12.a y 3.12.b, donde la media de PSNR aplicando BSR cae por debajo de la interpolación cuando se usa WW = +/-10, mientras que la media de SSIM aplicando BSR cae por debajo de la interpolación cuando se usa WW = +/-12, habiendo introducido 4 *frames* más en el proceso de SR que provocan más de 0.5 dB de pérdida de calidad en términos de PSNR. Así mismo, en las Figuras 3.13.a y 3.13.b se observa cómo los pocos artefactos que se generan con WW = +/-1 en la secuencia *Reef*, hacen que el PSNR para SR se encuentre por debajo de la interpolación, mientras que para SSIM, se obtiene una mejora de aproximadamente un 3%. De aquí se concluye **que la métrica SSIM proporciona una mejor idea de la estructura de una secuencia de vídeo a nivel global**, identificándose de una forma más cercana al sistema visual humano [BSS04].

3.2.4.2 Información espacial

La información espacial se refiere a los *frames* de una misma escena captados desde diferentes puntos del espacio en un mismo instante de tiempo. Es decir, si se dispone de varias cámaras será posible captar una secuencia de vídeo desde diferentes puntos del espacio, obteniendo de esta manera una secuencia temporal por cada una de las cámaras consideradas, tal y como se representa de forma esquemática en la Figura 3.14.

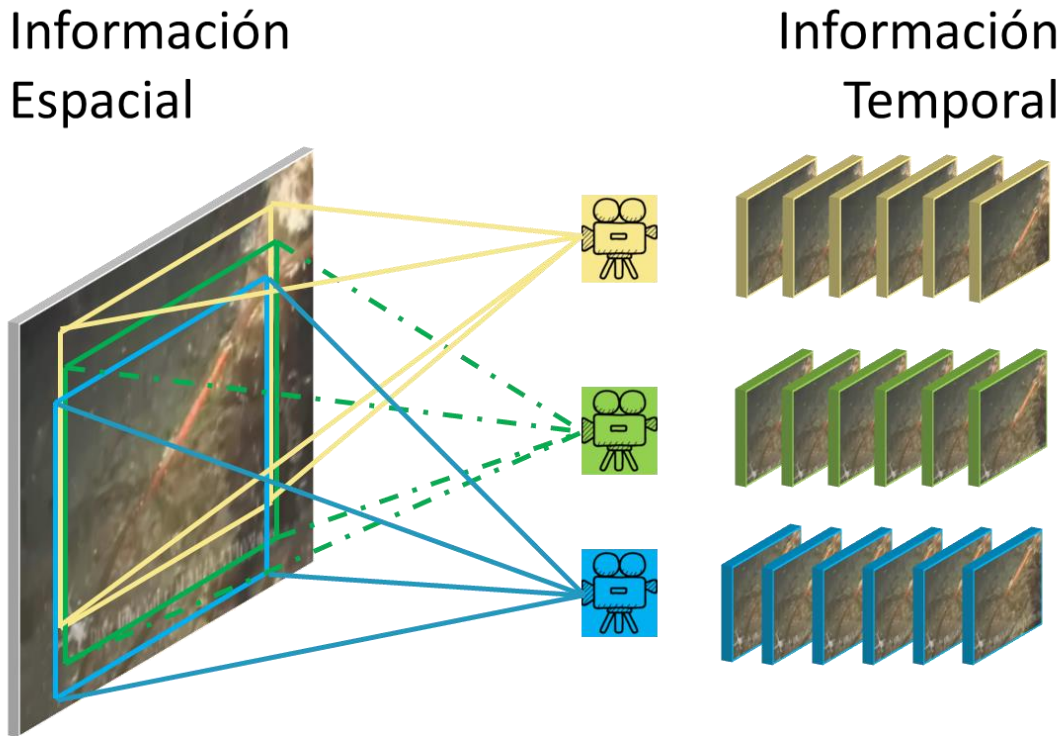


Figura 3.14.- Información espacial e información temporal

Partiendo de los resultados mostrados en la Figura 3.12 y en la Figura 3.13, se observa que en las secuencias de vídeo que presentan un movimiento continuo existe una mayor coherencia espacial entre *frames* cercanos en la línea temporal al *frame* actual a súper-resolver, y que cuando se introducen en la WW *frames* muy alejados del *frame* actual, normalmente se pierde calidad debido a la falta de coherencia espacial, que por lo general implica la aparición de artefactos.

Se deduce entonces que en la medida en que se pueda obtener un mayor número de *frames* cercanos en la línea temporal al *frame* actual a súper-resolver, se podría mejorar la calidad del *frame* de HR a la salida del proceso de SR. Esto se podría lograr gracias a la información espacial siempre y cuando se cumplan una serie de condiciones, como el alineamiento entre cámaras, que permitan obtener información común desde un punto de vista espacial y que exista desplazamiento subpíxel entre las secuencias de vídeo grabadas, de tal forma que se obtengan mejoras aplicando SR. De aquí se infiere que el uso de sistemas Multi-Cámara (MC) podría ser muy adecuado para las técnicas de SR por fusión.

3.3 FILTROS SELECTIVOS Y TOPOLOGÍA DE MACRO-BLOQUES ADAPTABLE

Tal y como se presentó en el capítulo del *Estado del Arte* de la presente Tesis Doctoral, una de las principales limitaciones de las técnicas de SR es la eficiencia de cómputo, ya que la computación es intensiva. Disminuir ese tiempo de cómputo tratando de mantener e incluso mejorar la calidad objetiva es una de las aportaciones de esta Tesis Doctoral, introduciendo para ello un pre-procesamiento en la información de entrada de LR, determinando los *frames* y/o los MBs más adecuados a incluir en el proceso de SR y especificando la topología de MBs (tamaño de los MBs) idónea.

3.3.1 Filtro Selectivo de Ventana y de Macro-Bloque

Partiendo de las conclusiones establecidas en la sección 3.2.4.1 *Información Temporal* se explicará en primer lugar el funcionamiento del Filtro Selectivo de Ventana tomando como punto de partida la Ventana de Trabajo de la Figura 3.10 (#WBF = 5 y #WFF = 5), para a continuación explicar el Filtro Selectivo de MB. La implementación considera umbrales a partir de los cuales se determina si un *frame* o un MB se incluyen o no en el proceso de SR [C2, C4, R1, R4]. Así, se han implementado un total de tres versiones del filtro:

- *Filtro de Ventana Cortante*: La WW se acorta con respecto a la original una vez se encuentra un *frame* que no cumple la condición del umbral.
- *Filtro Selectivo de Ventana*: Se seleccionan los *frames* que cumplen la condición del umbral dentro de la Ventana de Trabajo.
- *Filtro Selectivo de Macro-Bloques*: Se seleccionan los Macro-Bloques que cumplen la condición del umbral dentro de un *frame* determinado. Además, a fin de optimizar la selección de MBs, se considera también la opción del uso de una topología de MBs adaptable (tamaño de MB variable).

3.3.1.1 Filtro de Ventana Cortante

El Filtro de Ventana Cortante (WCF, *Window Cutting Filter*) recibe su nombre debido a que funciona en modo “cortante”, es decir a partir del *frame* actual evalúa temporalmente hacia atrás (*frames* previos) y adelante (*frames* posteriores) si se cumple la condición del umbral establecido (WCFTHR, *Window Cutting Filter THReshold*). Una vez se encuentre un *frame* que no cumpla la condición, se dejan de evaluar *frames* anteriores al *frame* en evaluación si se están considerando *frames* previos al *frame* actual, o adelante en el otro caso. Se presenta como referencia en la Figura 3.15 un caso particular a partir de la WW de la Figura 3.10. Se establecen para este caso los siguientes pasos:

1. Se compara el *frame* actual (6) con los *frames* previos en la línea temporal:
 - a. Al comparar con el *frame* “5”, se cumple la condición del umbral propuesto, por lo que se incluye este *frame* en el proceso.
 - b. Al comparar con el *frame* “4”, se cumple la condición del umbral propuesto, por lo que se incluye este *frame* en el proceso.
 - c. Al comparar con el *frame* “3”, NO se cumple la condición del umbral propuesto, por lo que NO se incluye este *frame* en el proceso y se deja de evaluar hacia atrás, fijando un nuevo valor de #WBF=2.

2. Se compara el *frame* actual (6) con los *frames* posteriores en la línea temporal:
 - a. Al comparar con el *frame* "7", se cumple la condición del umbral propuesto, por lo que se incluye este *frame* en el proceso.
 - b. Al comparar con el *frame* "8", se cumple la condición del umbral propuesto, por lo que se incluye este *frame* en el proceso.
 - c. Al comparar con el *frame* "9", NO se cumple la condición del umbral propuesto, por lo que NO se incluye este *frame* en el proceso y se deja de evaluar adelante, fijando un nuevo valor de #WFF=2.

Por tanto, el efecto global del WCF es una reducción (si procede) de la Ventana de Trabajo de SR. En el caso del ejemplo, la Ventana de Trabajo se ha reducido de +/-5 a +/- 2, tal y como se representa en la Figura 3.15, donde el umbral se representa como WCFTHR. El tamaño de la Ventana de Trabajo del WCF (WCFWW, *Window Cutting Filter Working Window*) queda pues definida por la ecuación 3.2 a partir de los *frames* de la Ventana de Trabajo original (FWW, *Frame of Working Window*) que cumplan la condición del umbral WCFTHR de forma consecutiva.

$$WCFWW = 1(CF) + \sum FWW / SSIM(WW, FA) > WCFTHR \quad (3.2)$$

El Filtro de Ventana Cortante está limitado por el hecho de que una vez se encuentra un *frame* que no cumple la condición del umbral WCFTHR deja de evaluar *frames* anteriores (en caso de que se estén evaluando *frames* temporalmente previos al actual) o *frames* posteriores (en el caso de que se estén evaluando *frames* temporalmente posteriores al actual) que podrían cumplir la condición establecida por el umbral y aportar información al proceso de SR. Por esta razón se decidió implementar una segunda versión basada en un Filtro Selectivo de Ventana.

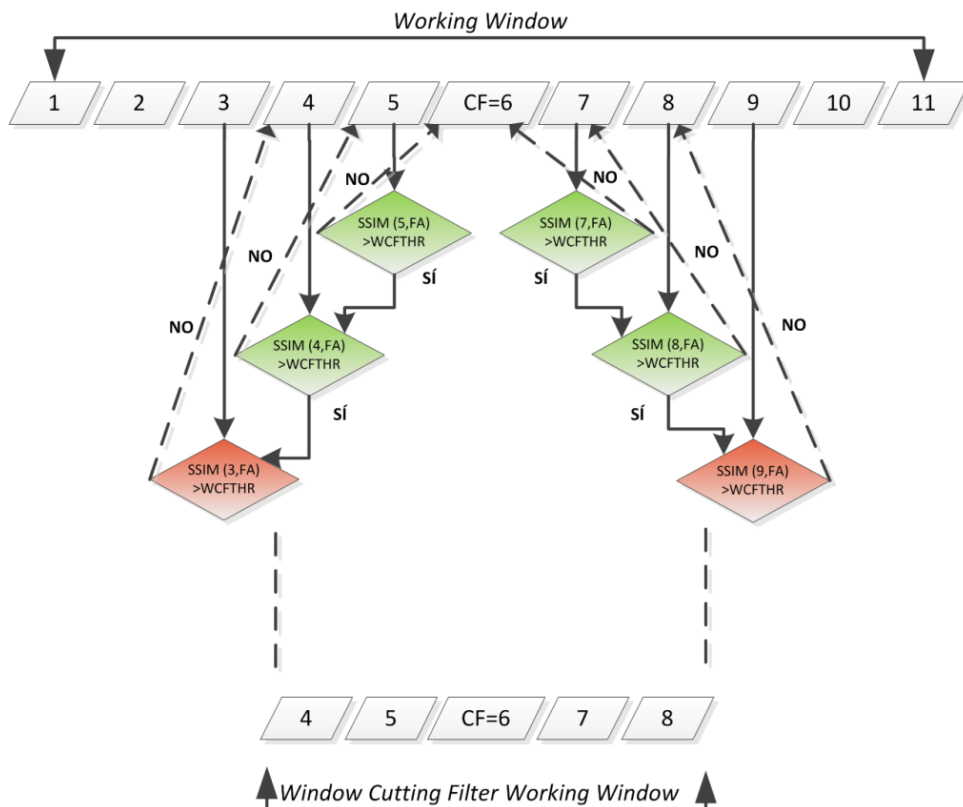


Figura 3.15.- Aplicación del Filtro de Ventana Cortante a Ventana de Trabajo de Figura 3.10

3.3.1.2 Filtro Selectivo de Ventana

El Filtro Selectivo de Ventana (WSF, *Window Selective Filter*) recibe su nombre debido a que funciona en modo “selectivo”. El *frame* actual es comparado con la totalidad de *frames* de la secuencia, y si se cumple la condición del umbral, esto querrá decir que los *frames* son estructuralmente similares uno con respecto a otro y por tanto se considerarán en el proceso de SR. Como referencia se presenta un caso particular en la Figura 3.16 a partir del caso expuesto en la Figura 3.10.

Se observa en el ejemplo que se evalúa la condición de filtrado para todos los *frames* de la secuencia. Esto incrementa ligeramente el coste computacional, al tener que realizar comparaciones para todos los *frames*, pero sin embargo considera *frames* que en el caso del Filtro de Ventana Cortante se podrían perder en el proceso. Este caso particular se da en los *frames* 1, 2 y 10. El tamaño de la Ventana de Trabajo del WSF (WSFWW, *Window Selective Filter Working Window*) queda pues definida por la ecuación 3.3 a partir de los *frames* de la Ventana de Trabajo original (FWW, *Frame of Working Window*) que cumplan la condición del umbral WSFTHR (*Window Selective Filter THReshold*).

$$WSFWW = 1(CF) + \sum FWW / SSIM(WW, FA) > WSFTHR \quad (3.3)$$

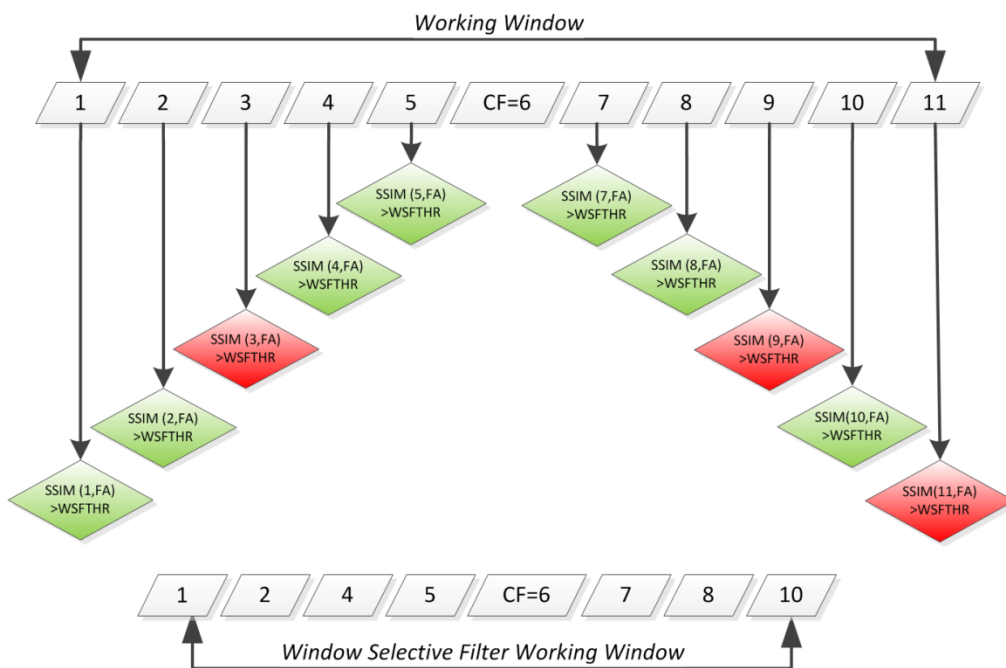


Figura 3.16.- Aplicación de Filtro Selectivo de Ventana a la Ventana de Trabajo de Figura 3.10

3.3.1.3 Filtro Selectivo de Macro-Bloques

Las secciones anteriores han presentado Filtros Selectivos de Ventana, que consideran *frames* con un cierto grado de similitud para realizar el proceso de SR. Sin embargo, es posible que para ciertos *frames* cuyo contenido general sea distinto al *frame* actual, determinadas regiones presenten similitudes a nivel local. Una de las innovaciones introducidas en la presente Tesis Doctoral se basa en la selección de los Macro-Bloques (MB) más adecuados a utilizar en el proceso de SR.

En secuencias con movimiento local, los objetos se pueden desplazar y cambiar de diferentes maneras dentro de un mismo *frame*, pudiendo aparecer oclusiones y vectores de movimiento fuera del Área de Búsqueda. En el mismo *frame* de referencia, es posible encontrar MBs que encajen y contribuyan a mejorar la resolución de la imagen súper-resuelta, mientras que otros introducen artefactos y ruido, y en consecuencia empeoran la calidad de la secuencia de SR. Así, se planteó un filtro a nivel de MB que mejorara la calidad de la imagen basándose en una selección apropiada de MBs a combinar. Para lograr mejores resultados con el algoritmo de SR evitando la aparición de artefactos en la imagen, la combinación de MBs se debe realizar en relación a su semejanza. Se determina si un MB debe utilizarse en el proceso de SR o si debe rechazarse antes de la etapa de *Compensación de Movimiento*, haciendo uso de un umbral a nivel de MB (BSFTHR, *Block Selective Filter THreshold*) basado en la diferencia media absoluta (MAD, *Mean Absolute Difference*) entre MBs. El MAD se calcula entre el bloque actual (*currentBlock*) y el bloque de referencia (*refBlock*) utilizando la expresión (3.4), donde $M \cdot N$ es el tamaño de los bloques, e i y j los índices de píxeles. La condición a cumplir para considerar o no un bloque en el proceso de SR se presenta en la Figura 3.17.

$$MAD = \frac{1}{M \cdot N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |currentBlock_{ij} - refBlock_{ij}| \quad (3.4)$$

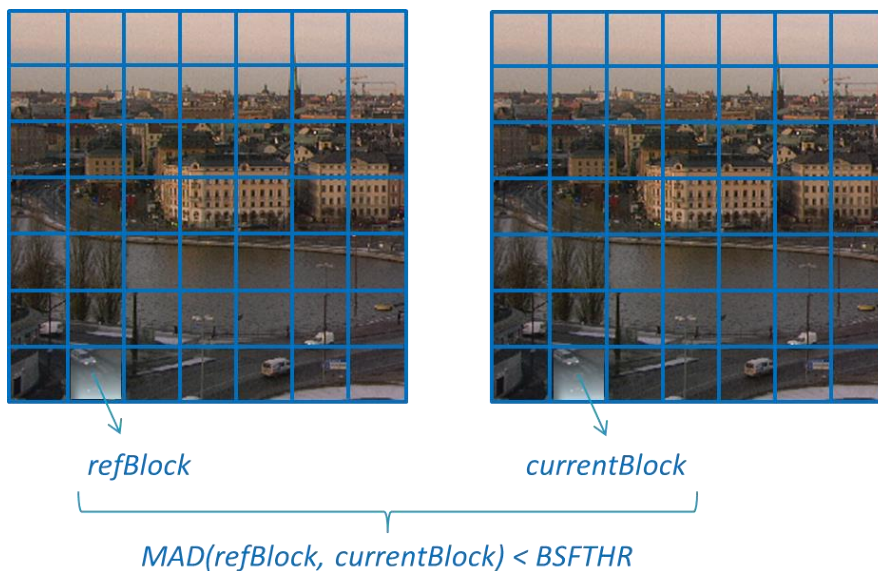


Figura 3.17.- Condición necesaria para considerar un MB en el proceso de SR (Filtro Selectivo de MB)

Una vez presentados el Filtro Selectivo de Ventana (*Window Selective Filter*, WSF) y el Filtro Selectivo de MB (*Block Selective Filter*, BSF) es coherente pensar que una combinación de ambos podría ser muy adecuada para establecer un filtro a dos niveles: un primer nivel *grueso*, a nivel de Ventana de Trabajo, y un segundo nivel *fino*, a nivel de MB; a este filtro conjunto se le ha denominado Filtro Selectivo de Ventana y MB (*Window & Block Selective Filter*) y opera de tal manera que selecciona en primer lugar los *frames* a incluir en el proceso de SR haciendo uso del umbral WSFTHR, y partiendo de los *frames* seleccionados decide los MBs a utilizar comparando con el umbral BSFTHR.

3.3.2 Topología de MBs adaptable – Tamaño de MB variable

En el algoritmo de referencia BSR, el tamaño de MB es un parámetro que se debe ajustar para cada secuencia. Este parámetro tiene una influencia fundamental en la etapa de *Estimación de Movimiento*. En secuencias con movimiento global es preferible por lo general utilizar MBs de mayor tamaño, ya que los vectores de movimiento obtenidos son más precisos, lo que se traduce en una mejora de los resultados de SR. Sin embargo, en secuencias con movimiento local, MBs de tamaño elevado pueden contener objetos con vectores de movimiento diferentes, y la *Estimación de Movimiento* presentará una precisión inferior. En este caso, los MBs deben ser de un tamaño menor, tratando de dividir la imagen en MBs que contienen sólo los objetos con el mismo movimiento.

La consideración de una topología de MBs adaptable se basa pues en un tamaño de MB variable (VBS, *Variable Block Size*) y consiste en una modificación del algoritmo BSR en el que la división de MBs se realiza de forma dinámica utilizando bloques de diferentes tamaños. La división variable de MBs se basa en la consideración de dos umbrales que se identifican directamente con los umbrales WSFTHR y BSFTHR. El primer umbral, VBSTHR1, decide si dividir o no un MB, comparando el SSIM entre bloques. En caso de que se verifique la condición de SSIM se calcula el MAD entre MBs y se utiliza para decidir el tamaño de un MB. Si el MAD de un MB grande es mayor que un umbral VBSTHR2, el MB se divide. El MAD se calcula iterativamente para los nuevos bloques más pequeños y se compara sucesivamente con el umbral VBSTHR2. Este proceso se repite hasta que se alcanza un tamaño de bloque mínimo de 4x4. El MAD se calcula entre el MB actual (*currentBlock*) y el MB de referencia (*refBlock*) utilizando una vez más la expresión (3.4). El proceso queda reflejado en la Figura 3.18.

La primera aproximación que se implementó estaba basada sólo en el contenido del *frame* actual. Se buscó un criterio que fuese capaz de dividir la imagen según la textura del fotograma. Las zonas más homogéneas mantenían MBs de gran tamaño mientras que las zonas con más cambios de textura eran divididas iterativamente en MBs de menor tamaño. Como resultado, se obtenía una única división de MBs del *frame* actual que se aplicaba a toda la Ventana de Trabajo de SR. Para realizar esta división, se utilizó como criterio el *SADintra* por píxel. El *SADintra* proporciona información sobre la diferencia que hay entre los píxeles de un MB y su valor medio. Mientras mayor sea esta diferencia, se observarán mayores contrastes en un MB. El criterio de decisión basado en la textura del *frame* actual no dio los resultados esperados y quedó patente que el criterio de división debía incluir información no sólo del *frame* actual sino que también debía incluir información del *frame* de referencia. Esto llevó a la búsqueda de nuevos criterios de decisión para dividir cada *frame*. Además, utilizar criterios que implicasen la fusión de información proveniente del *frame* de referencia con el *frame* actual conlleva a que se debe realizar una división distinta para cada *frame* de referencia dentro de la Ventana de Trabajo.

Los siguientes criterios que se escogieron para realizar la división de MBs, estaban basados en los gradientes temporales y espaciales del *frame*. Como resultado, se alcanzaron ciertas mejoras en determinadas secuencias con mucho movimiento local utilizando como criterio el gradiente temporal. Estos resultados no eran adecuados en secuencias donde dominaba el movimiento global. El algoritmo dividía innecesariamente los *frames* provocando una pérdida de calidad en comparación con el uso de MBs de tamaño fijo. Analizando los resultados, se desarrolló el criterio definitivo. La combinación idónea para realizar la división de MBs resultó de la combinación de dos criterios: el SSIM entre los *frames* de referencia y el actual, y su gradiente temporal.

Cuando el SSIM entre dos *frames* es alto, no es necesario realizar una división irregular del *frame*, los MBs de mayor tamaño son los que mejor se adaptan para obtener buenos resultados de calidad. En cambio, cuando el SSIM es menor, hay una mayor diferencia entre los *frames* y puede suponer un indicador de la aparición de movimiento local. En este caso, realizar una subdivisión de ciertos MBs aplicando como criterio el gradiente de movimiento mejora la calidad de la secuencia de vídeo.

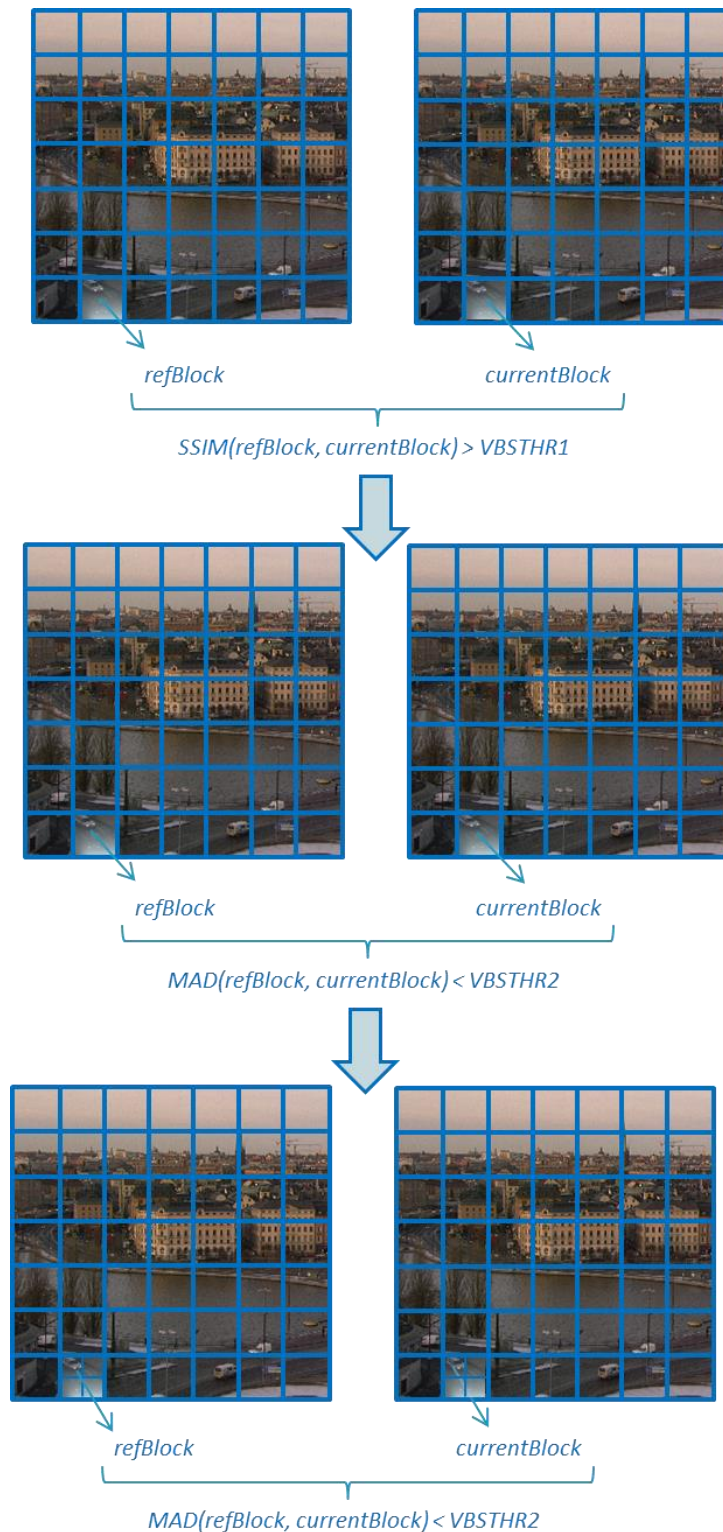


Figura 3.18.- Condiciones para dividir MBs en el proceso de SR usando VBS

El Filtro Selectivo de Ventana, el Filtro Selectivo de MB y la adaptación de la topología de los MBs (tamaño de MB variable) establecen un amplio abanico de variantes dependiendo de cuáles de ellos se activen, o no, con el objetivo de mejorar los resultados de BSR. En la Tabla 3.2 se desglosan las 7 variantes posibles (se considera como “0” la situación inicial presente en la BSR) dependiendo de las siguientes características:

- Selección a nivel de WW (*Working Window*) → WSF (*Window Selective Filter*).
- Selección a nivel de MB (*Macro-Block*) → BSF (*Block Selective Filter*).
- Tamaño de MB variable → VBS (*Variable Block Size*).

Tabla 3.2.- Nomenclatura asociada a Filtros Selectivos de Ventana y Macro-Bloque

ID	Acrónimo	Nomenclatura	WSF	BSF	VBS
0	BSR	Baseline SR	0	0	0
1	VBSSR	Variable Block Size SR	0	0	1
2	BSFSR	Block Selective Filter SR	0	1	0
3	VBSFSR	Variable Block Selective Filter SR	0	1	1
4	WSFSR	Window Selective Filter SR	1	0	0
5	VBSWSFSR	Variable Block Size & Window Selective Filter SR	1	0	1
6	WBSFSR	Window & Block Selective Filter SR	1	1	0
7	WVBSFSR	Window & Variable Block Selective Filter SR	1	1	1

3.3.3 Análisis del valor de los umbrales

Los Filtros Selectivos de Ventana y MB presentados permiten obtener una disminución del tiempo de cómputo y una mejora de la calidad de la imagen, siempre y cuando el valor del umbral establecido sea adecuado. Ocurre lo mismo cuando se plantea mejorar la calidad con los umbrales establecidos para la división de MBs en VBS, aunque en este caso, por contra, el tiempo de cómputo se incrementará. Con el objetivo de modelar un valor de umbral del Filtro Selectivo de Ventana WSFTHR adecuado es fundamental entender cómo funciona éste:

- El valor de WSFTHR está directamente relacionado con la métrica SSIM, que varía entre 0 y 1 (o entre 0 y 100 de forma porcentual) para determinar la calidad global de una imagen. La caracterización del filtro se ha realizado utilizando valores porcentuales para una mayor claridad.
- Cada *frame* analizado se compara con respecto al *frame* actual usando la métrica de SSIM comparándola con el umbral (WSFTHR), de forma que si se cumple la condición de que $SSIM(\text{Frame Analizado}, \text{Frame Actual}) > \text{WSFTHR}$ entonces el *frame* analizado se considerará en el proceso de SR. Esto da lugar a que:
 - Un valor **WSFTHR = 100**, hará que ningún *frame* analizado se incluya en el proceso de SR, ya que $SSIM(\text{Frame Analizado}, \text{Frame Actual})$ siempre será menor o igual a 100 (valor máximo del umbral). Este caso de umbral máximo se corresponde por tanto con la interpolación (sólo se considera el *frame* actual para el tratamiento de la imagen).
 - Un valor **WSFTHR = 0**, hará que todos los *frames* analizados se incluyan en el proceso de SR, ya que $SSIM(\text{Frame Analizado}, \text{Frame Actual})$ siempre será mayor a 0 (valor mínimo del umbral). Este caso de umbral mínimo se corresponde por tanto con la versión inicial del algoritmo de SR (BSR, *Baseline Super-Resolution*).

Atendiendo a las características del umbral utilizado, sería adecuado realizar un cálculo experimental del umbral del filtro a partir de las características de la secuencia. En primer lugar cabe destacar que en secuencias con movimiento global el valor de WSFTHR ha de ser bajo para asegurar de esta forma que los *frames* a descartar sean lo suficientemente diferentes, mientras que en secuencias con movimiento local acusado habrá que incrementar el valor de WSFTHR para de esta forma evitar incluir *frames* que introduzcan artefactos. El análisis de un amplio abanico de secuencias demostrará que valores alrededor de un 20% funcionan apropiadamente en secuencias con movimiento global, mientras que valores alrededor de un 80% presentan buenas prestaciones cuando se combina movimiento global con movimiento local. Un valor por debajo del 50% asegura que se eliminen los *frames* que se parecen muy poco de acuerdo con una evaluación subjetiva basada en MOS (*Mean Opinion Score*) [ZMZ+14] tal y como se presenta en la Tabla 3.3, por tanto, esto se podrá tener en cuenta para la gran mayoría de secuencias que no incluyan un movimiento local acusado.

Tabla 3.3.- Relación entre SSIM y MOS

SSIM	MOS	Calidad	Deficiencia
≥ 0.99	5	Excelente	Imperceptible
[0.95, 0.99)	4	Buena	Perceptible pero no molesta
[0.88, 0.95)	3	Media	Un poco molesta
[0.5, 0.88)	2	Pobre	Molesta
< 0.5	1	Mala	Muy molesta

Así mismo, el criterio utilizado para tomar la decisión de filtrar a nivel de MB se basa en el MAD calculado durante la etapa de *Estimación de Movimiento*, comparándolo con el umbral BSFTHR. De esta manera, se incrementan las métricas de calidad de PSNR y SSIM de la secuencia en alta resolución. La condición sobre BSFTHR está invertida con respecto a WSFTHR con el fin de mantener coherencia con el cálculo establecido con MAD, y que los valores de umbral se mantengan cercanos a 0. El valor de MAD oscila entre 0 y 255, pero se ha normalizado para que varíe entre 0 y 100. Por tanto:

- Un valor **BSFTHR = 0**, hará que ningún MB analizado se incluya en el proceso de SR, ya que nunca se cumplirá que $MAD(MB \text{ Referencia}, MB \text{ Actual})$ sea menor que 0. Este caso de umbral mínimo se corresponde por tanto con la interpolación.
- Un valor **BSFTHR = 100**, hará que todos los MBs analizados se incluyan en el proceso de SR, ya que siempre se cumplirá que $MAD(MB \text{ Referencia}, MB \text{ Actual})$ sea menor que 100 (valor máximo del umbral). Este caso de umbral máximo se corresponde por tanto con la versión inicial del algoritmo de SR (BSR, *Baseline Super-Resolution*).

En la Figura 3.19, se muestran los MBs filtrados para la secuencia *Suzie* en color negro, resaltándose a su izquierda que se trata de un *frame* de 160x128 píxeles formado por 20 MBs de 32x32 píxeles, de los cuales 15 han sido seleccionados para realizar el proceso de SR al superar la condición del umbral. El análisis de un amplio abanico de secuencias demuestra que al trabajar a un nivel mucho más fino a nivel de MB que a nivel de *frame*, un valor normalizado de BSFTHR 20 (MAD = 50) hace que un MB prácticamente no se parezca a otro, por lo que en este caso se verá en el próximo capítulo de la presente Tesis Doctoral que lo más adecuado es mantener un valor de BSFTHR bajo para secuencias con movimiento local y algo más elevado para secuencias con movimiento global.

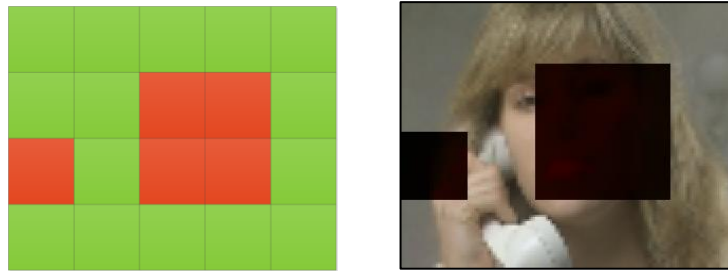


Figura 3.19.- Selección de MBs para aplicar SR en un frame de la secuencia Suzie utilizando FBS

En la Figura 3.20 se muestran dos ejemplos de MB por división adaptativa utilizando VBS para dos *frames* de las secuencias *Mobile* y *Foreman*. La división en cada *frame* se modifica dinámicamente con el contenido de la secuencia. Las regiones de mayor complejidad se dividen en bloques de tamaño inferior para aumentar la precisión durante la etapa de *Estimación de Movimiento*. La definición de los umbrales VBSTHR1 y VBSTHR2 se identifica con la determinación de los umbrales WSFTHR y BSFTHR, respectivamente.



Figura 3.20.- División de MBs en frames de las secuencias *Mobile*(a) y *Foreman*(b) usando VBS

Finalmente, se ha de considerar que en la primera implementación del Filtro Selectivo de MB el umbral de decisión, para descartar o incluir MBs en el proceso de SR era estático y debía fijarse al configurar los parámetros del algoritmo. Con el objetivo de mejorar la respuesta del algoritmo en secuencias con características cambiantes y reducir la dependencia de los resultados con el umbral establecido en los parámetros de configuración del algoritmo, se desarrolló un filtro de MB con un umbral dinámico adaptativo que se adaptase al contenido de la secuencia de vídeo. Para conseguir este objetivo, se desarrolló un umbral de decisión que se modula para cada MB utilizando como criterio los vectores de movimiento resultantes del proceso de *Estimación de Movimiento*.

En la mayoría de las situaciones, se observa que al realizar el proceso de *Estimación de Movimiento* de los MBs del *frame* actual con respecto al *frame* de referencia, aquellos MBs con vectores de movimiento mayores suelen dar como resultado una peor correspondencia entre MBs. Como resultado, en el proceso de SR estos MBs presentan mayor probabilidad de introducir artefactos en la secuencia de HR. Por ello, es coherente introducir un umbral que sea más restrictivo cuando se detecte en la imagen mayor movimiento entre MBs. Así, el criterio de decisión se modificó aplicando la expresión 3.5 como criterio para realizar el filtro de los MBs.

$$MAD < BSFTHR * [1 - (|MV|/SA)], \quad (3.5)$$

donde MV es el vector de movimiento considerado y SA el Área de Búsqueda. Con este criterio de decisión, se obtuvieron mejores resultados de calidad en varias secuencias, especialmente en aquellas en las que existe mayor variación de su contenido, tal y como se presentará en el capítulo de Resultados.

3.3.4 Pseudocódigo del algoritmo

Se realizó un estudio exhaustivo con el fin de determinar la forma óptima de integrar las aportaciones de WSFSR, BSFSR así como la topología de MBs adaptable [R4]. El resultado de este estudio se presenta en la Figura 3.21, que introduce el pseudocódigo asociado. Se observa en el pseudocódigo que se siguen los siguientes pasos:

- Se realiza una selección de los *frames* a considerar en el proceso de SR en caso de que el *flag* WSF esté activado. Para ello se compara el resultado del SSIM del *frame* actual con el *frame* tomado como referencia en el proceso de SR, de forma que si este valor se encuentra por debajo del valor del umbral WSFTHR, se descarta dicho *frame*.
- Partiendo de los *frames* de la Ventana de Trabajo considerados, se realiza el proceso de *Estimación de Movimiento* haciendo uso de un tamaño de MB fijo o variable, dependiendo de si el *flag* VBS está activado o no, respectivamente. Si VBS está activado se considera MB variable y se dividen los MBs considerando el valor de los umbrales VBSTHR1 y VBSTHR2, tal y como se ha explicado en el apartado anterior.
- Finalmente se realiza una selección de los MBs a considerar en el proceso de SR en caso de que el *flag* BSF esté activado. Para ello se compara el resultado del MAD del MB actual (multiplicado por $[1-|MV|/SA]$ si se considera umbral adaptativo) con el MB tomado como referencia en el proceso de SR, de tal forma que si este valor se encuentra por debajo del valor de umbral BSFTHR, se descarta dicho MB.

```

Entradas: Secuencia de frames de LR : primer_frame:último_frame
Flags: WSF, VBS, BSF
Umbrales: WSFTHR, VBSTHR1, VBSTHR2, BSFTHR
Número de frames de la Ventana de Trabajo: tamaño_ventana
Salidas: secuencia_SR
for frame_actual = primer_frame : último_frame
  imagen_SR = Muestrear con huecos(frame_actual)
  Crear Ventana de Trabajo =
  min(frame_actual-tamaño_ventana,primer_frame):max(frame_actual+tamaño_ventana,último_frame)
  for ref_frame = frame en Ventana de Trabajo excepto frame_actual
    if WSF
      if SSIM(frame_actual, ref_frame) >= WSFTHR,
        máscara de Ventana de Trabajo (frame_actual) = 1
      else
        máscara de Ventana de Trabajo (frame_actual) = 0
      end if
    if(máscara de Ventana de Trabajo (frame_actual) = 1)
      if VBS
        if SSIM(frame_actual, ref_frame) <= VBSTHR1
          VBS_frame_division(frame_actual, ref_frame, VBSTHR2)
          for i = 1 : número de Macro-Bloques
            MB_actual[i][j] = seleccionar MB i,j en frame_actual
            ref_MB[i][j] = seleccionar MB i,j en ref_frame
            vector_movimiento[i],MAD = Estimación de Movimiento(MB_actual[i][j],ref_MB[i][j])
          end if
        else
          for i = 1 : filas_MB
            for j = 1 : columnas_MB
              MB_actual[i][j] = seleccionar MB i,j en frame_actual
              ref_MB[i][j] = seleccionar MB i,j en ref_frame
              vector_movimiento[i][j], MAD = Estimación de Movimiento con tamaño de MB fijo
              (MB_actual[i][j], ref_MB[i][j])
            end if
          end if
          BSFTHR = BSFTHR*(1-|vector_movimiento|/SA)
          if BSF & (MAD < BSFTHR)
            imagen_SR = imagen_SR + Sobremuestreo(ref_MB[i+vector_movimiento.x[i][j],
              j+vector_movimiento.y[i][j])
          end if
        imagen_SR = Rellenado de Huecos(imagen_SR)
        almacenar SR_image en secuencia_SR
      almacenar secuencia_SR
    end if
  end for
end for

```

Figura 3.21.- Pseudocódigo incluyendo filtros selectivos y topología de MBs adaptable

3.4 ENTORNO MULTI-CÁMARA

En esta sección, se describen los métodos propuestos para combinar la información temporal y espacial proporcionada por un sistema MC mediante el uso del BSR explicado en la sección 3.2. El objetivo de estos métodos es conseguir combinar la información temporal y espacial de manera que se aumente la calidad de la secuencia de salida de una de las cámaras del *array* MC [C5, R3]. Por otra parte, se han implementado una serie de algoritmos de pre-procesamiento para estudiar qué partes de las secuencias de vídeo aportan más información a la secuencia súper-resuelta. Este estudio tiene una doble finalidad: primero, valorar si es posible descartar partes de los *frames* de las secuencias de vídeo para reducir el coste computacional; y segundo, analizar el efecto que tiene la cantidad de solape entre las cámaras del *array* MC sobre la calidad de la secuencia súper resuelta.

3.4.1 Métodos de SR con MC propuestos

Se presentan aquí, los diferentes métodos de SR usando MC propuestos en la presente Tesis Doctoral. Para ello, se han implementado un total de tres métodos diferentes, donde dos de ellos hacen uso de procesos de SR temporal y SR espacial para obtener la secuencia súper-resuelta de salida. Finalmente un tercer método combina la información temporal y espacial (método Mixto). La SR espacial hace referencia a la posición de las cámaras en el mismo instante de tiempo, mientras que la SR temporal está relacionada con el flujo temporal que sigue cada cámara.

3.4.1.1 Método Temporal-Espacial (*Temporal-Spatial SR*)

El método aquí expuesto, hace uso de la información temporal y espacial aportada por el sistema MC para obtener la secuencia súper-resuelta de salida en dos fases. En una primera fase se combina únicamente la información temporal, y en una segunda fase se combina la información espacial. Es por ello que este método se ha denominado Temporal-Espacial (*Temporal-Spatial SR*), queriendo expresar que la secuencia final obtenida será el resultado de realizar dos procesos de SR consecutivos, así como el orden en que se combina la información temporal y espacialmente.

Tal y como se aprecia en la Figura 3.22, en primer lugar se súper-resuelven las secuencias de las cámaras del *array* MC aplicando un proceso de SR haciendo uso de la información temporal de cada una de ellas. Este proceso de SR se realiza haciendo uso de una Ventana de Trabajo Temporal definida por el usuario (*Temporal Working Window, TWW*), donde se decidirá la cantidad de *frames* a considerar en este proceso, tal y como sucedería en el algoritmo BSR. A la salida de esta primera fase se dispondrá por tanto de K secuencias súper-resueltas, donde el valor K estará determinado por las dimensiones del *array* MC $R \times S$, siendo R las filas y S las columnas que definen el tamaño del *array*, o incluso podrían representar las dimensiones de un subconjunto de cámaras dentro de ese *array* MC.

La secuencia resultante de este primer proceso de SR será considerada como secuencia de Resolución Media_{temporal} (*Medium Resolution_{temporal}*, MR_t), tal y como se indica en la Figura 3.21. La relación entre la resolución de la secuencia de entrada (secuencia de Baja Resolución, LR) y la resolución de la secuencia de salida (MR_t), es un Factor de Escala de tipo entero definido por el usuario, denominado "Escalado_{Temporal}".

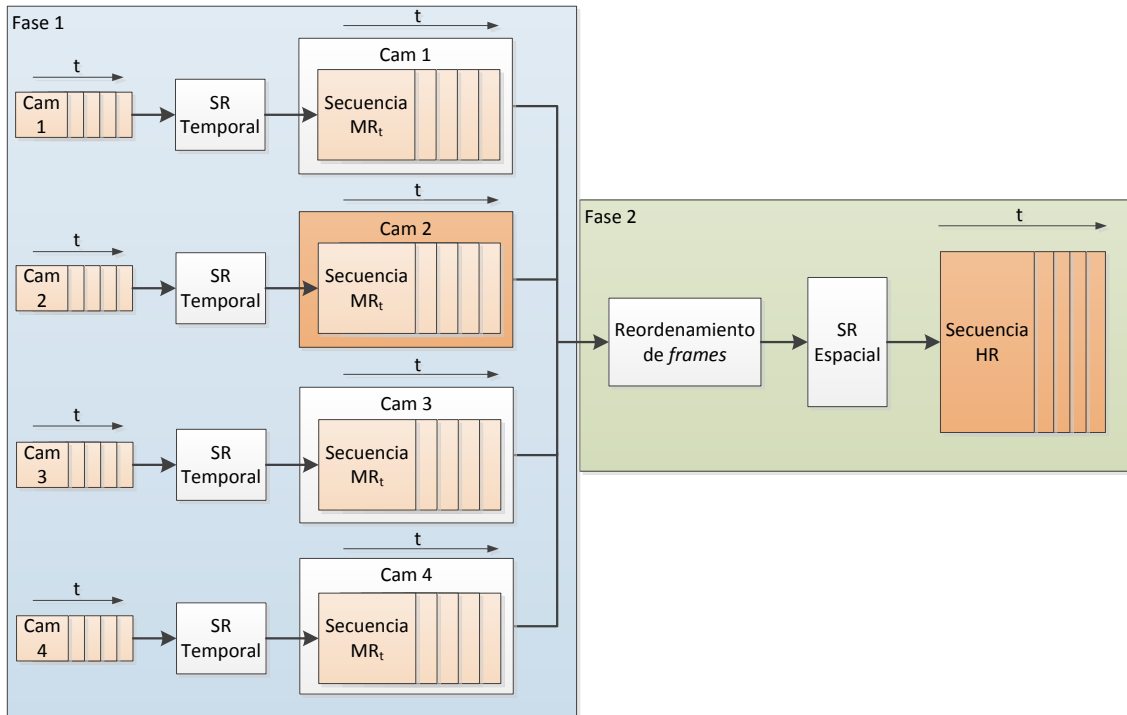


Figura 3.22.- Diagrama del método Temporal-Espacial usando la cámara 2 como referencia

Posteriormente, las secuencias de MR_t pasan por un proceso de reordenamiento de frames, mostrado en la Figura 3.23, el cual consiste en reordenar cada uno de los frames capturados por cada una de las cámaras del array MC en el mismo instante de tiempo en orden lexicográfico, esto es, de izquierda a derecha y de arriba abajo, repitiendo este proceso para los sucesivos instantes de tiempo, convirtiéndose esta nueva secuencia reordenada en la entrada del segundo proceso de SR.

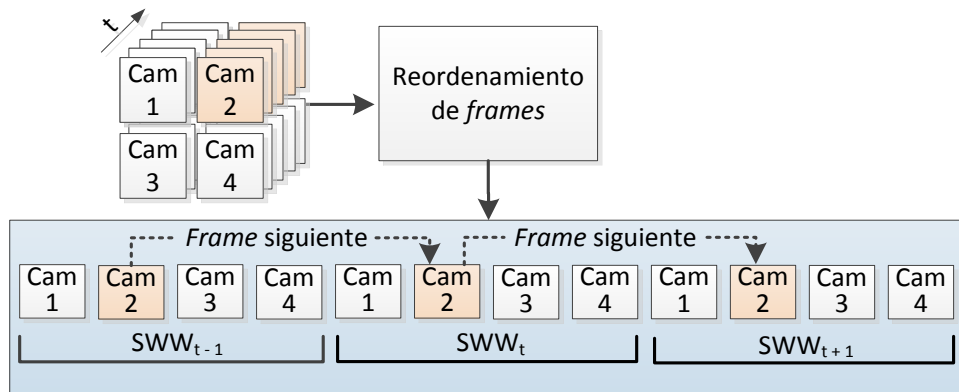


Figura 3.23.- Proceso de reordenamiento de frames

En la segunda fase se hará uso de la información espacial contenida en las secuencias de MR_t para obtener finalmente como resultado una secuencia súper-resuelta de salida de alta resolución (HR), que constituye la salida final del método Temporal-Espacial, tal y como se representa en la Figura 3.22. La secuencia de MR_t y la secuencia HR están relacionadas a través de un Factor de Escala denominado "Escalado_{Espacial}". En este caso, para efectuar la SR Espacial se define una Ventana de Trabajo Espacial (*Spatial Working Window, SWW*) la cual contiene únicamente los frames capturados en un mismo instante de tiempo por el array MC, y que por tanto, serán los únicos frames considerados para realizar el proceso de SR Espacial.

Este método presenta un alto coste computacional, puesto que el número de procesos de SR a realizar en la primera fase crecerá en función de las dimensiones del *array* MC, teniendo que realizar K procesos de SR Temporal, (donde $K = R \times S$); siendo R el número de posiciones horizontales de cámaras en el *array* y S el número de posiciones verticales de cámaras en el *array*, más un proceso de SR Espacial (siendo esta última de mayor cómputo al realizarse sobre imágenes de tamaño $LR \times Escalado_{Temporal}$), resultando por consiguiente en un total de: $(K + 1) = (R \times S + 1)$ procesos de SR.

Finalmente, analizando este método desde un punto de vista global, cabe resaltar que entre la secuencia de entrada de LR y la de HR, que constituye la salida del método propuesto, se define un Factor de Escala global que es el producto entre el Factor de Escala del proceso de SR Temporal y el proceso de SR Espacial, tal y como se define en la ecuación (3.6).

$$Escalado_{Global} = Escalado_{Temporal} \times Escalado_{Espacial} \quad (3.6)$$

Considerando que el valor mínimo posible de cada Factor de Escala es 2, la mínima relación de escala posible entre la secuencia de LR y la secuencia de HR será 4, si bien es posible diezmar la salida en caso de requerirse un Factor de Escala 2.

3.4.1.2 Método Espacial-Temporal

Derivado del método Temporal-Espacial, surge la necesidad de analizar si existe alguna influencia del orden en el que se realiza la SR Temporal o la SR Espacial en la calidad de la secuencia HR resultante, dando pie a un método similar al anterior, pero invirtiendo el orden en el que se realizan los procesos de SR.

Como se puede apreciar en la Figura 3.24, el método Espacial-Temporal se basa en realizar en primer lugar un proceso de SR Espacial, donde la información usada para generar cada uno de los *frames* de la secuencia súper-resuelta será la capturada desde los distintos puntos del espacio en un mismo instante de tiempo por las cámaras del *array* MC. Como resultado se obtendrá una secuencia de Media Resolución_{Espacial} (*Medium Resolution_{Spatial}*, MR_S). Para ello, como paso previo al proceso de SR Espacial, se realiza un reordenamiento de los *frames* provenientes del sistema MC en orden lexicográfico con el fin de simplificar la entrada de los datos al algoritmo de SR usado en esta Tesis Doctoral.

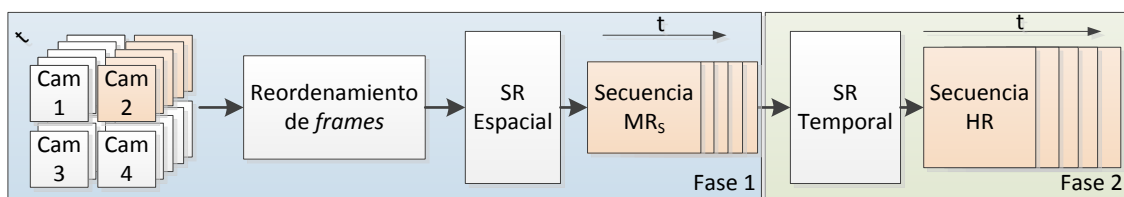


Figura 3.24.- Diagrama del método Espacial-Temporal

Posteriormente se combina la información temporal contenida entre los *frames* de la secuencia de MR_S debido al muestreo realizado por la cámara seleccionada en distintos instantes de tiempo, *frames* que previamente han sido súper-resueltos haciendo uso de la información espacial. Esta segunda fase dará como resultado una secuencia súper-resuelta de HR, que constituye la salida final del método Espacial-Temporal, tal y como se representa en la Figura 3.24. En este caso, el proceso de SR se realiza haciendo uso de una Ventana de Trabajo TWW (*Temporal Working Window*) tal y como sucedería en el algoritmo BSR.

En este método se puede ver que, aparte de obtener un estudio sobre la influencia del orden en el que se realiza la SR temporal o espacial en la calidad de la secuencia HR resultante, se consigue reducir notablemente el coste computacional, puesto que solo se realizan dos procesos de SR para obtener la secuencia de salida de HR. Al igual que en el método Temporal-Espacial, la relación de escala entre la secuencia de LR y HR está determinada por la ecuación (3.6), puesto que el orden de los procesos de SR no afectará al tamaño de la secuencia de salida HR.

3.4.1.3 Método Mixto

Este método surge como una combinación de los métodos anteriormente descritos, los cuales, pese a que hacen uso de la información temporal y la información espacial proporcionadas por el sistema MC para obtener la secuencia HR, aplican la información a los *frames* en dos fases independientes. Es por ello que, aunque se consiguen aprovechar los beneficios que proporciona un sistema MC, sería deseable poder hacer uso de algún modo de la información temporal y espacial en un mismo proceso de SR y de manera combinada, para generar una secuencia de salida de HR. Con este objetivo se propone el método Mixto, que se presenta gráficamente en la Figura 3.25.

A raíz de este análisis surge la idea de que después de realizar el proceso de reordenamiento de los *frames*, en lugar de definir una Ventana de Trabajo que contenga únicamente información espacial, tal y como se realiza en la primera fase del método Espacial-Temporal, se podría ampliar el tamaño de esta WW de forma que contenga *frames* con información de carácter espacial y temporal, definiendo de este modo una nueva Ventana de Trabajo mixta (*Mixed Working Window, MWW*). Todo esto se puede realizar gracias a la forma en la que el proceso de reordenamiento almacena los *frames*. En este caso, al tener dentro de la MWW tanto información temporal como espacial, el algoritmo BSR en la etapa de *Estimación de Movimiento*, realizará una búsqueda en *frames* que estará relacionada con el *frame* a procesar mediante variaciones temporales y espaciales. De esta forma se estarán ampliando las posibilidades de encontrar la información más adecuada para súper-resolver la secuencia de vídeo de la cámara seleccionada.

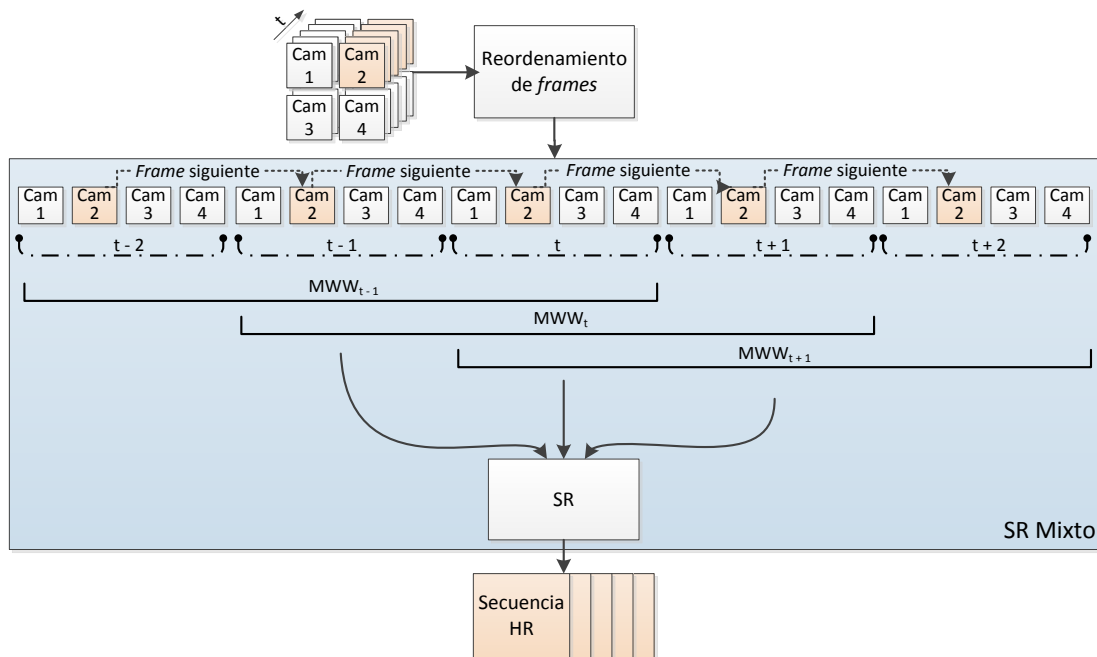


Figura 3.25.- Diagrama del método Mixto

De forma gráfica se presenta en la Figura 3.25 un ejemplo de cómo se define la MWW en el método Mixto. En este ejemplo se ha elegido un *array* MC de tamaño 2x2, en el cual se va a aplicar el proceso de SR la cámara 2, y se ha decidido que la MWW comprenda la información de un instante de tiempo anterior y posterior al *frame* a procesar. Para el instante “*t-1*” se define una Ventana de Trabajo que contiene tanto los *frames* del *array* MC capturados en el instante “*t-1*”, como los *frames* capturados en los instantes “*t-2*” y “*t*”. Una vez procesado el *frame* de la cámara 2 en el instante “*t-1*”, se procede a súper-resolver este mismo *frame* en el instante de tiempo “*t*”. En este caso, se vuelve a generar una MWW similar a la anterior, pero conteniendo los instantes de tiempo “*t*”, “*t-1*” y “*t+1*”, procediendo igualmente en los sucesivos *frames*. Será la etapa de *Estimación de Movimiento* del proceso de SR la encargada de buscar la información más adecuada dentro de la MWW para componer la secuencia de HR de salida del método de SR mixto. Al haber extendido la Ventana de Trabajo de forma que contenga tanto los *frames* capturados en un mismo instante de tiempo, (información espacial), como diferentes instantes tiempo (información temporal), resultará en que en esta etapa se obtienen MBs para reconstruir la secuencia súper-resuelta que estén relacionados de forma temporal, espacial o mixta; siendo esta última una combinación de las dos anteriores.

Para ver lo anteriormente expuesto, considérese la situación en que la MWW que se está procesando es la Ventana de Trabajo que contiene los instantes de tiempo “*t-1*”, “*t*” y “*t+1*”, correspondiente a la MWW central mostrada en la Figura 3.25. Al llevar a cabo la etapa de *Estimación de Movimiento* para la secuencia de la cámara 2, se podría obtener como resultado una de las siguientes relaciones entre MBs, tal y como se resumen en la Tabla 3.4:

- **Relación espacial:** Donde se obtendría como resultado que el MB más adecuado para aplicar SR es un MB que se encuentra en el *frame* correspondiente a la cámara 1, 3 o 4 en el instante de tiempo “*t*”.
- **Relación temporal:** Donde se obtendría como resultado que el MB más adecuado para aplicar SR es un MB que se encuentra en el *frame* correspondiente a la cámara 2 en el instante de tiempo “*t-1*” o “*t+1*”.
- **Relación mixta:** Donde se obtendría como resultado que el MB más adecuado para aplicar SR es un MB que se encuentra en el *frame* correspondiente a la cámara 1, 3 o 4 en el instante de tiempo “*t-1*” o “*t+1*”.

Tabla 3.4.- Relaciones entre MBs dependiendo de las cámaras y los instantes de tiempo

Instante de tiempo	Cámara			
	2	1	3	4
<i>t</i>	Interpolación	Relación Espacial		
<i>t - 1</i>	Relación Temporal	Relación Mixta		
<i>t + 1</i>				

Sin embargo, en cada fase de los métodos Temporal-Espacial y Espacial-Temporal solo se puede dar el caso de que en la etapa de *Estimación de Movimiento* se encuentren MBs que mantengan una relación espacial o temporal, dependiendo del tipo de proceso de SR que se esté aplicando, ya que se procesa la información temporal y espacial independientemente. En el método Mixto se consigue reducir el coste computacional de los anteriores métodos, puesto que solamente se realiza un proceso de SR para obtener la secuencia súper-resuelta, aunque a expensas de aumentar los requerimientos de memoria del algoritmo. Además se dispone de un único Factor de Escala en este método, que se identifica con el factor de “Escalado_{Temporal}” o “Escalado_{Espacial}” de los anteriores métodos, logrando así una mayor flexibilidad en las dimensiones finales de las secuencias súper-resueltas.

3.4.2 Modos de pre-procesamiento para SR con MC

De forma complementaria a los métodos anteriormente expuestos, se han desarrollado una serie de algoritmos de pre-procesamiento de las secuencias de entrada. El objetivo es decidir si descartar alguna parte de los *frames* capturados por el *array* MC de manera que se obtenga una mayor calidad y/o se reduzca el coste computacional. Se establecen restricciones a las características físicas que debe cumplir el *array* MC, para adaptarse lo mejor posible al algoritmo de BSR usado en esta Tesis Doctoral, de tal forma que la disposición de las cámaras en el *array* debe cumplir que:

- El *array* MC debe presentar una geometría rectangular, cuadrada o romboidal.
- Las cámaras deben estar situadas en el mismo plano. Además, no se realiza corrección de perspectiva de la escena.
- La separación entre cámaras debe ser tal que las cámaras estén grabando partes de una misma escena global desde distintos puntos del espacio.

Teniendo en cuenta un *array* MC que cumpla estos requisitos, como el mostrado en la Figura 3.26.a, se puede observar que entre los *frames* grabados por cada una de las cámaras (o al menos un subconjunto de las cámaras del *array*) existirá una parte común o solape, y alrededor de este solape quedará lo que se denominará como *borde*, tal y como se muestra en la Figura 3.26.b. La geometría y la separación entre las cámaras del *array* MC afectarán a la disposición de los bordes y a la forma de obtenerlos, y así mismo, influirá en la dimensión del solape entre las cámaras.

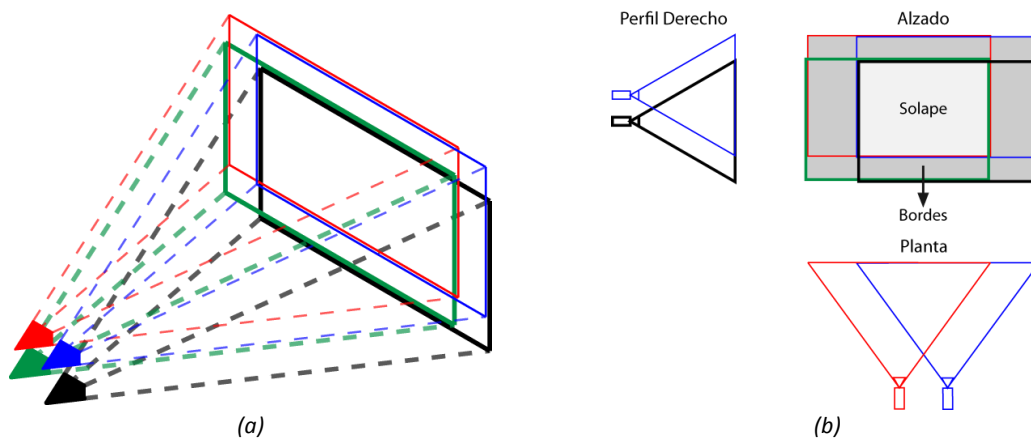


Figura 3.26.- Array MC

(a) Imágenes captadas en perspectiva.

(b) Perfil, alzado y planta de las imágenes captadas por las cámaras

De este análisis, surge la idea de experimentar los posibles beneficios de realizar un proceso de SR considerando toda la información del *frame* (*Full Frame*), solamente el solape entre cámaras (*Overlap*), o si por el contrario sería más conveniente dividir la imagen en la parte solapada y los bordes (*Overlap + Borders*), súper-resolver por separado cada una de estas partes y fusionar los trozos para obtener la imagen súper-resuelta resultante. A continuación se pasará a detallar estos modos de pre-procesamiento.

3.4.2.1 Full-Frame (Frame Completo)

Este es el modo básico de funcionamiento, el cual recoge el *frame* tal cual es capturado por cada cámara, para posteriormente ser procesado por alguno de los métodos de SR anteriormente presentados.

3.4.2.2 Overlap (Solape)

En este caso se pretende obtener la parte común entre todos los *frames* del *array* MC (el solape u *Overlap*), para posteriormente procesarlo mediante los métodos de SR anteriormente descritos. Para obtener el solape, primero se debe hallar el desplazamiento en *píxeles* que existe entre las cámaras (*offset*). Una vez conocido, se procede a recortar todos los *frames* capturados por cada una de las cámaras de manera que solo se conserva el solape. Tras recortar los *frames*, se obtiene una nueva secuencia, la cual será súper-resuelta mediante el método Temporal-Espacial, Espacial-Temporal o Mixto.

3.4.2.3 Overlap + Borders (Solape + Bordes)

Tras haber planteado un método que proporciona el solape, que es donde previsiblemente se espera que haya más información para usar en el proceso de SR, se plantea la posibilidad de estudiar el beneficio que puede aportar aplicar el proceso de SR a los bordes que existen alrededor de este solape, además de estudiar cuánto aporta, en términos de calidad de la secuencia de salida, el solape frente a los bordes a la calidad de la secuencia resultante. De esta forma es posible analizar si es conveniente descartar estos bordes, súper-resolverlos por separado, o realizar la SR con el *frame* completo sin aplicar este pre-procesamiento. Por tanto, se plantea dividir los *frames* capturados por el *array* MC en 9 zonas no solapadas, tal y como se presenta en la Figura 3.27, las cuales no tendrán partes comunes entre ellas.

A continuación, mediante el uso de un *array* MC de 2x2, que es el caso más sencillo, se muestra de forma gráfica lo anteriormente comentado. En la Figura 3.27 se puede observar que las imágenes tomadas por una *array* MC formarán una intersección, en la cual se diferencia una parte de solape y 8 bordes, los cuales pueden categorizarse en laterales (2, 4, 6 y 8) y esquinas (1, 3, 5 y 7). En este caso se observa que para los laterales existe información de todas las cámaras. Sin embargo esto no sucederá así en las esquinas, donde solo se dispone información de una de las cámaras. Por tanto, cuando se creen las secuencias correspondientes para cada parte almacenándolas en orden lexicográfico, en aquellos casos en los que no exista información, se procederá al rellenado con ceros.

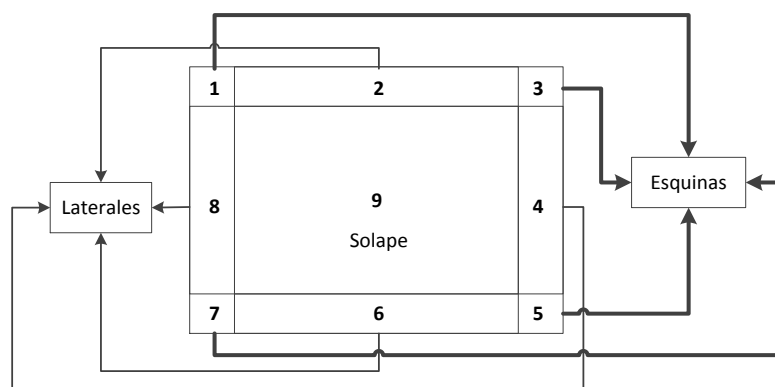


Figura 3.27.- División de los frames del array MC en bordes y solape

3.4.3 Pseudocódigo del algoritmo

Los métodos de SR MC y los modos de pre-procesamiento introducidos se integraron algorítmicamente tal y como se presenta en el pseudocódigo de la Figura 3.28. Del pseudocódigo se desprende que el principal aspecto a tener en cuenta en la integración del sistema MC en el algoritmo de SR consiste en la definición de la Ventana de Trabajo o las Ventanas de Trabajo apropiadas para cada método de SR, considerando adicionalmente el modo de pre-procesamiento seleccionado en cada caso [R3].

Para integrar la información aportada por el *array* MC en el algoritmo de SR es necesario establecer un conjunto de nuevos parámetros de entrada, cuya finalidad es poder activar el modo de pre-procesamiento correspondiente, así como el método de SR deseado. Estos parámetros indican:

- Cámara del *array* MC que se va a súper-resolver (*cámara_referencia*).
- Modo de pre-procesamiento que se va a aplicar (*modo_procesamiento*) incluyendo como opciones *Full_Frame*, *Overlap* u *Overlap+Borders*.
- Método MC que se va a emplear (*Método*), cuyas opciones se corresponden con *Temporal-Espacial*, *Espacial-Temporal*, *Mixto* y *BSR*.
- Número de cámaras (*num_cámaras*) que contiene el *array* MC, que combinándolo con el valor de la cámara de referencia, permite determinar las dimensiones de la Ventana de Trabajo espacial, así como de la Ventana de Trabajo del método Mixto. Para considerar esto último hace falta conocer además el número de instantes de tiempo (*num_instantes_tiempo*), anteriores y posteriores al *frame* actual, que se deseen considerar en el método de SR MC Mixto.

Para la ejecución de los métodos de SR MC, se ha de determinar en primer lugar el tamaño de la Ventana de Trabajo que se va a usar en el proceso de SR Espacial. Para ello se determinan cuáles son el primer *frame* y último *frame* que componen la Ventana de Trabajo a partir de los parámetros de entrada *num_cámaras* y *cámara_referencia*. Posteriormente, para cada *frame* de la cámara a súper-resolver se creará una Ventana de Trabajo que contiene los *frames* correspondientes a cada método (usando la función *Crear_Ventana_Trabajo*). En el caso del método de SR MC Temporal-Espacial y del método de SR MC Espacial-Temporal, se crearán una Ventana de Trabajo para cada proceso de SR.

Posteriormente, si se ha seleccionado el modo de pre-procesamiento *Overlap* u *Overlap+Borders*, se obtendrá el *offset* que existe entre las cámaras. Este *offset* servirá para dividir los *frames* de la Ventana de Trabajo en el solape y los bordes, en función del modo de pre-procesamiento seleccionado. El modo de pre-procesamiento elegido se aplicará directamente en el método de SR MC Mixto y en la primera fase del método de SR MC Espacial-Temporal, mientras que en el método de SR MC Temporal-Espacial se procesará en la segunda fase, puesto que es donde existe una relación espacial entre los *frames*. Una vez pre-procesada la Ventana de Trabajo, según el modo de pre-procesamiento elegido, se procede a ejecutar el proceso de SR sobre el *frame* actual considerando la información que éste contiene. De este modo se obtiene la versión súper-resuelta del *frame* actual, la cual en el método de SR MC Mixto se corresponde con la secuencia HR súper-resuelta. Sin embargo para los métodos de SR MC Temporal-Espacial y de SR MC Espacial-Temporal, el resultado será un *frame* de la secuencia MR, que constituirá la entrada del segundo proceso de SR. En esta segunda fase se creará una nueva Ventana de Trabajo, temporal o espacial, a partir de los *frames* de la secuencia MR para generar finalmente la versión súper-resuelta HR del *frame* actual.

```

Entradas: Secuencia de frames de Baja Resolución: for primer_frame to último_frame
Número de frames para formar la Ventana de Trabajo: tamaño_ventana
Cámara a ser súper-resuelta: cámara_referencia
Número de cámaras del array Multicámara: num_cámaras // cuando BSR -> 1
Método de SR: Método // BSR, Temporal-Espacial, Espacial-Temporal o Mixto
Modo de pre-procesamiento: modo_procesamiento // Full Frame, Overlap, Overlap+Borders
Número de instantes de tiempo a considerar en la SR Mixta: num_instantes_tiempo

Salida: Secuencia_SR

for frame_actual = primer_frame : num_cámaras : último_frame
    imagen_SR = Muestrear con huecos(frame_actual)

    case Método

    when BSR || Temporal-Espacial
        Ventana de Trabajo = Crear_Ventana_Trabajo(SR Temporal)

    when Espacial-Temporal
        Ventana de Trabajo = Crear_Ventana_Trabajo(SR Espacial)

    others // SR Mixto
        Ventana de Trabajo = Crear_Ventana_Trabajo(SR Mixta)

    if Método: not BSR and not Full_Frame
        offset = offset_cámaras (Ventana de Trabajo)
    end if

    if Método: not (Temporal-Espacial or BSR)
        preprocesar(Ventana de Trabajo, offset, modo_procesamiento)
    end if

    if Método: Temporal-Espacial
        for cámara_actual = 1:num_cámaras
            imagen_MR (cámara_actual) = SR(Ventana de Trabajo, frame_actual)
        end for

        Ventana de Trabajo = Crear_Ventana_Trabajo(SR Espacial)

        preprocesar(Ventana de Trabajo, offset, modo_procesamiento)
        imagen_SR = SR(Ventana de Trabajo, frame_actual)

    else if Método: Espacial-Temporal
        imagen_MR = SR(Ventana de Trabajo, frame_actual)

        Ventana de Trabajo = Crear_Ventana_Trabajo(SR Temporal)

        imagen_SR = SR(Ventana de Trabajo, frame_actual)

    else // SR Mixto
        imagen_SR = SR(Ventana de Trabajo, frame_actual)
    end if

    imagen_SR = Rellenado de Huecos(imagen_SR)
    almacenar imagen_SR en Secuencia_SR

almacenar Secuencia_SR

Función Crear_Ventana_Trabajo(Tipo_SR)
    offset_primera_cámara = cámara_referencia + 1
    offset_última_cámara = num_cámaras - cámara_referencia

    case tipo_SR
    when SR Temporal
        Ventana de Trabajo = max(frame_actual - tamaño_ventana, primer_frame) to
            min(frame_actual + tamaño_ventana, último_frame)

    when SR Espacial
        Ventana de Trabajo = max(frame_actual - offset_primera_cámara, primer_frame) to
            min(frame_actual + offset_última_cámara, último_frame)

    others // SR Mixto
        primera_cámara_mixto = offset_primera_cámara + (num_cámaras x num_instantes_tiempo)
        última_cámara_mixto = offset_última_cámara + (num_cámaras x num_instantes_tiempo)

        Ventana de Trabajo = max(frame_actual - primera_cámara_mixto, primer_frame) to
            min(frame_actual + última_cámara_mixto, último_frame)

    return Ventana de Trabajo
end Función

```

Figura 3.28.- Pseudocódigo del algoritmo de SR incluyendo métodos Multi-Cámara

3.5 CONCLUSIONES

En este capítulo, se han introducido las aportaciones algorítmicas desarrolladas en la presente Tesis Doctoral. Aunque la exposición desarrollada en el capítulo se ha presentado de forma evolutiva con respecto al esquema temporal de la presente Tesis Doctoral, como conclusión al capítulo se presenta en la Figura 3.29 un esquema global de las aportaciones, tomando como referencia una secuencia de entrada de baja resolución LR hasta llegar a una secuencia de vídeo de salida de alta resolución HR.

Se observa en la Figura 3.29 que tras un reordenamiento de los *frames* de las cámaras de entrada de un *array* MC (por ejemplo aquí 2x2), se pasa a realizar el pre-procesamiento MC seleccionando la opción *Full-Frame* (se considera toda la información), el solape (*Overlap*), o la información combinada de solape y bordes (*Overlap+Borders*). Se aplica entonces uno de los tres métodos de SR con MC desarrollados: Temporal-Espacial, Espacial-Temporal o Mixto. Finalmente, se aplica el Filtro Selectivo de Ventana o de MB, además del uso de la topología Macro-Bloques adaptable (si se selecciona esta opción), obteniendo una secuencia de HR a la salida.

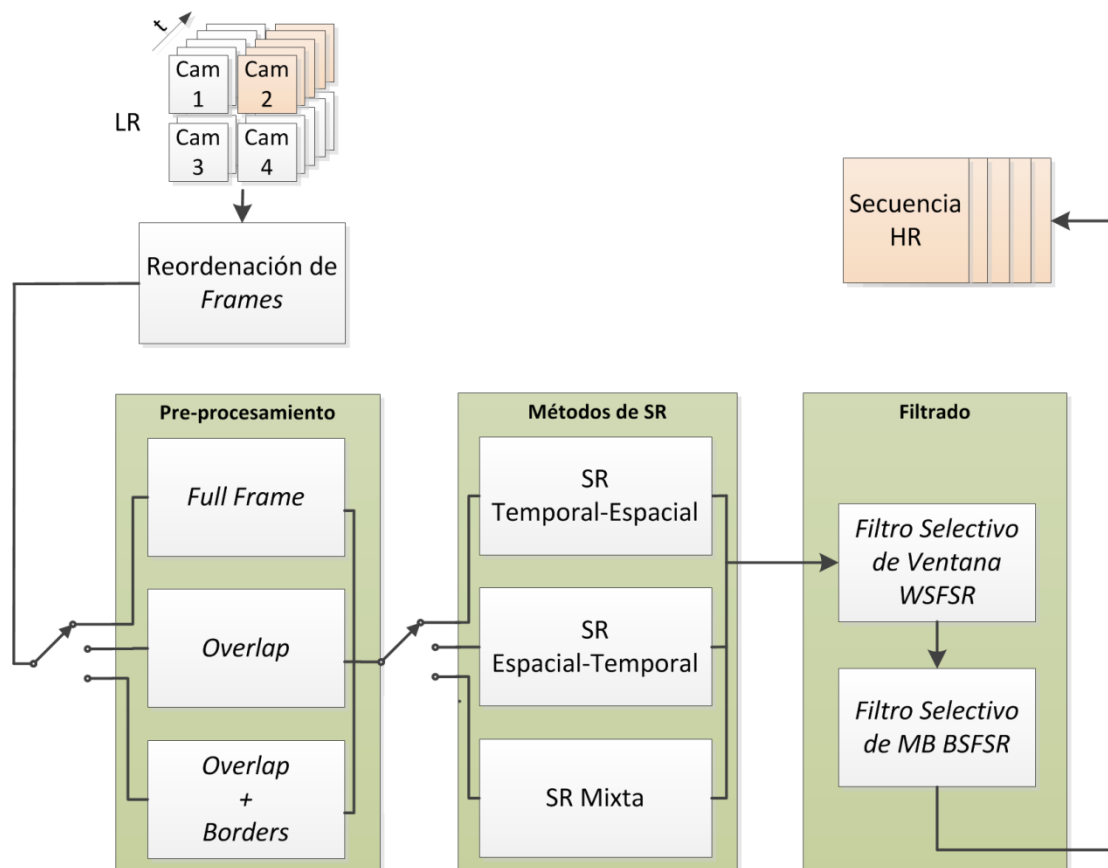


Figura 3.29.- Diagrama global de las aportaciones algorítmicas desarrolladas

Capítulo 4

RESULTADOS Y ANÁLISIS: FILTROS SELECTIVOS Y TOPOLOGÍA DE MACRO-BLOQUES ADAPTABLE

Índice del capítulo

	<i>Página</i>
4.1.- Introducción	86
4.2.- Entorno de simulación y análisis	87
4.3.- Simulaciones	92
4.4.- Conclusiones	125

Si buscas resultados distintos, no hagas siempre lo mismo
Albert Einstein, científico alemán nacionalizado estadounidense (1879 – 1955)

4.1 INTRODUCCIÓN

Una vez presentadas las contribuciones algorítmicas aportadas en la presente Tesis Doctoral, se dedica este capítulo a introducir el entorno de simulación específico desarrollado para validar las aportaciones relativas a filtros selectivos y topología de Macro-Bloques adaptable, así como para analizar los resultados más significativos obtenidos a partir del proceso de caracterización de prestaciones.

El entorno de simulación, que se introduce en la sección 4.2, está basado en un banco de pruebas o *testbench* que integra un modelo de verificación completo para el conjunto de contribuciones realizadas, introduciendo el pre-procesamiento y el post-procesamiento necesarios para ejecutar la validación de los algoritmos desarrollados.

Los resultados de simulación siguen el flujo de desarrollo de la Tesis Doctoral, tal y como se presentaron en el capítulo anterior, partiendo del Filtro Selectivo de Ventana, pasando por el Filtro Selectivo de Macro-Bloques hasta llegar a la topología de Macro-Bloques adaptable. Por otra parte, si bien el filtro constituye la salida del sistema completo presentado en las conclusiones del capítulo anterior, tal y como se presenta en la Figura 4.1, en este capítulo se presentará con independencia de dicho sistema, aplicándolo a secuencias tomadas por cámaras simples.

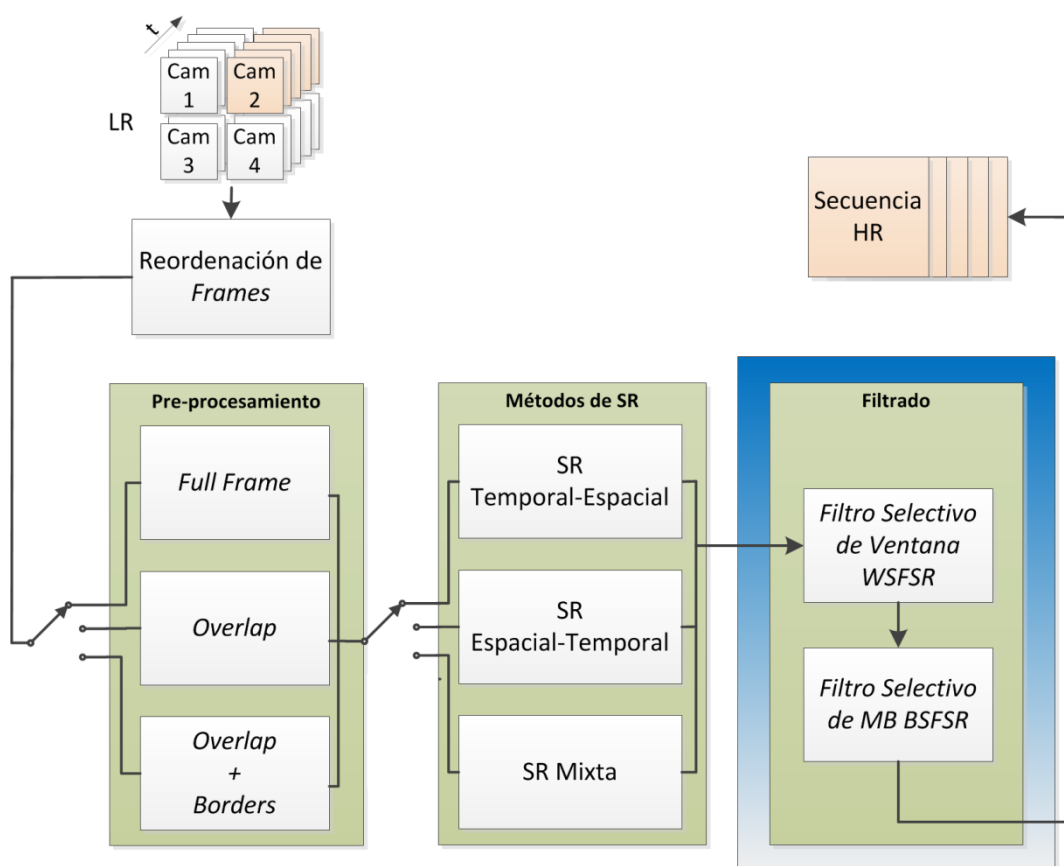


Figura 4.1.- Diagrama global de las aportaciones algorítmicas desarrolladas destacando filtros

4.2 ENTORNO DE SIMULACIÓN Y ANÁLISIS

Es necesario validar el procedimiento empleado en esta Tesis Doctoral y verificar el flujo de diseño en lo que se refiere a las aportaciones relacionadas con filtros selectivos y topología de Macro-Bloques adaptable. Para ello se ha implementado un banco de pruebas o *testbench*, que no constituye un sistema aislado, sino que se integra en el código de partida. Por este motivo, la implementación de este banco de pruebas representa la modificación más compleja a realizar. Además, se debe tener especial cuidado a la hora de tratar con tamaños de imágenes variados en cada fase del proceso. Todo ello es fundamental para poder establecer comparaciones con otras aportaciones desarrolladas en el Estado del Arte.

En la Figura 4.2 se muestra el diseño del banco de pruebas desarrollado. En éste se puede observar la evolución de las imágenes y de sus tamaños, aspecto que será tratado posteriormente.

En primer lugar, se recibe una secuencia de imágenes de alta resolución como entrada y se recortan para adecuar las imágenes de forma previa al diezmado. En el siguiente paso se diezma la secuencia de imágenes desplazadas según el factor indicado por el parámetro *Scale Factor* asociado al proceso de SR, para así crear la secuencia de baja resolución que simula la entrada original al proceso. Esta secuencia es almacenada en un archivo, cuyo nombre puede indicar el usuario a través del fichero de configuración.

A partir de este punto se aplican, tanto el proceso de SR como el proceso de interpolación. El resultado obtenido es de alta resolución, si bien no es necesariamente el mismo tamaño que la secuencia original, pues recuérdese la posibilidad de tener que recortar las imágenes originales para realizar el diezmado. El proceso de SR se puede ejecutar en las diferentes aproximaciones introducidas en la presente Tesis Doctoral. Los posibles modos de trabajo del algoritmo son:

- *Baseline Super-Resolution (BSR)*: Es el algoritmo que se toma como punto de partida para la presente Tesis Doctoral, adaptado para cumplir con los requisitos exigidos que permitan obtener a la salida, tanto *frames*, como secuencias de vídeo con el formato adecuado.
- *Window Selective Filter (WSF)*: Modo Filtro Selectivo de Ventana activado.
- *Block Selective Filter (BSF)*: Filtro Selectivo de MB activado.
- *Variable Block Size (VBS)*: Tamaño de MB Variable activado.

Para poder aplicar la relación señal a ruido de pico, PSNR (*Peak Signal to Noise Ratio*), y el índice de similitud estructural, SSIM (*Structural SIMilarity index*) con el fin de obtener una medida objetiva de la calidad entre la imagen original y las obtenidas como resultado de aplicar los procesos de SR e interpolación, es necesario recortar también la imagen original. Así, las imágenes tienen el mismo tamaño y pueden ser comparadas adecuadamente. Para el estudio de secuencias MC se han desarrollado *testbenches* específicos para cada método, que se explicarán posteriormente.

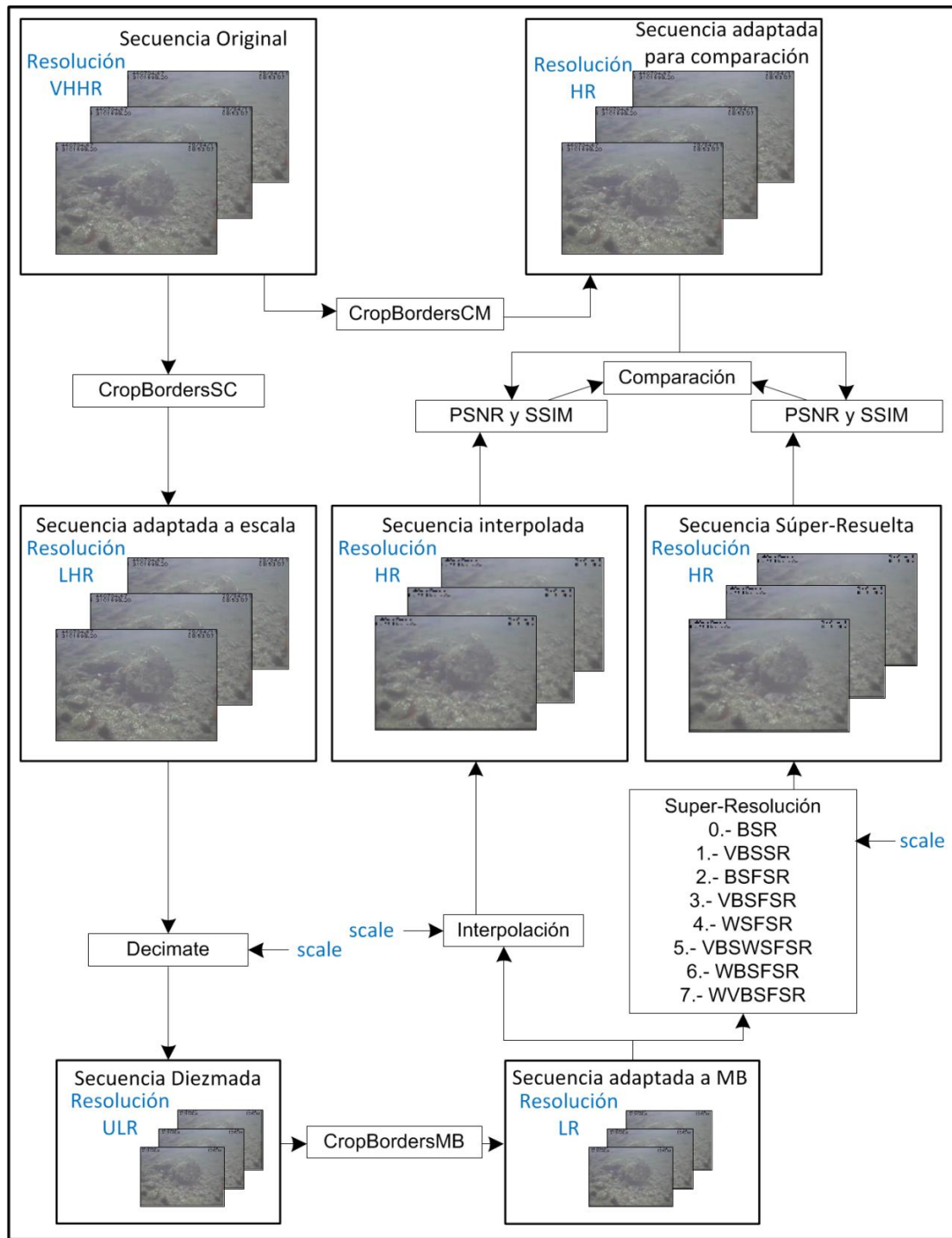


Figura 4.2.- Diseño del Banco de Pruebas (Testbench)

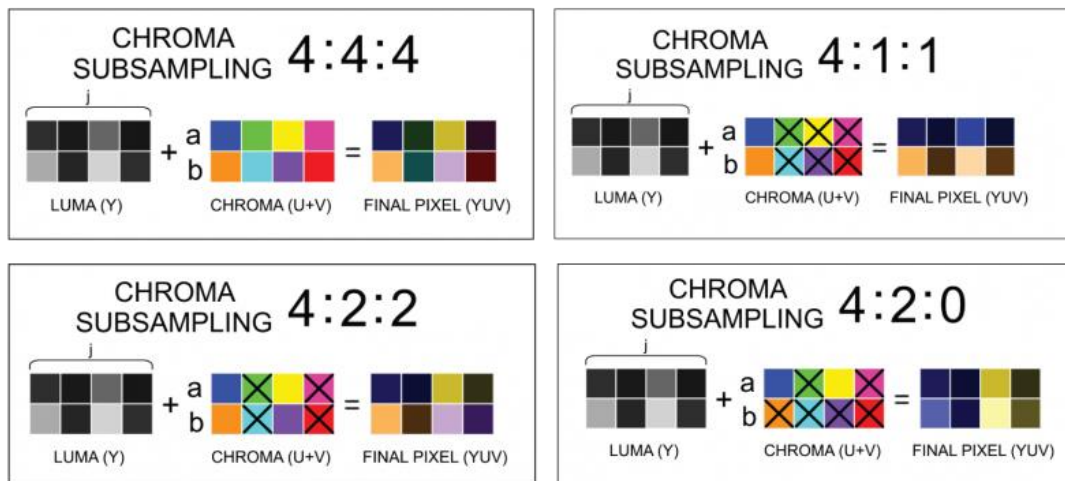
4.2.1 Análisis de la adecuación del tamaño de la imagen

Supóngase que el tamaño de la imagen original es de $m \times n$ píxeles. Este tamaño puede no ser compatible con el factor de diezmado, por lo que se utiliza la función *cropborders()*, que elimina las filas y columnas necesarias de modo que se pueda diezmarse la imagen resultante. Suponiendo que el número de filas a recortar es *cropRowsSC*, y que *cropColsSC* representa el número de columnas, tras *cropborders()* la secuencia tendrá un tamaño de $(m - \text{cropRowsSC}) \times (n - \text{cropColsSC})$ píxeles.

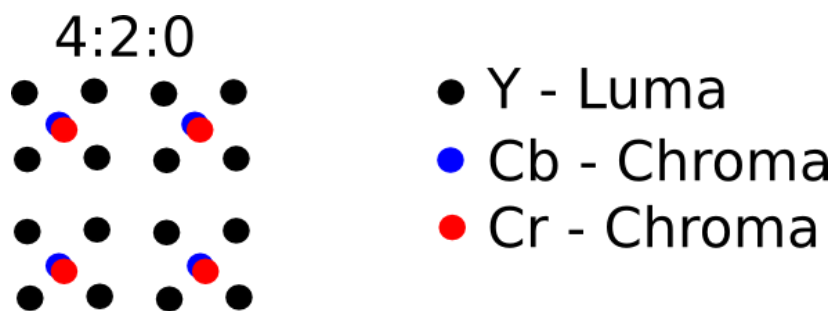
El formato de muestreo que se utiliza en la presente Tesis Doctoral es el 4:2:0, que es uno de los formatos de muestreo de crominancias (*chroma*) más utilizados (ver en Figura 4.3.a diferentes formatos). En el formato 4:2:0 se muestrea toda la información de luminancia y las crominancias se muestrean a la mitad de frecuencia, tanto en horizontal como en vertical, tal y como se representa en la Figura 4.3. Se ha de considerar que para que la secuencia recortada tras *cropBorders()*, cuya resolución se denominará LHR (*Lower High Resolution*), pueda ser diezmada, ha de cumplir las siguientes condiciones:

1. Las columnas y filas de resolución LHR han de ser divisibles por el Factor de Escala (*Scale Factor*).
2. Tras dividir por el Factor de Escala, las columnas y filas resultantes han de ser divisibles por 2 para poder mantener de forma adecuada la relación entre crominancias y luminancias.

Por tanto, para que la resolución LHR pueda ser diezmada, las columnas y filas de resolución LHR han de ser divisibles por un valor doble al Factor de Escala.



a)



b)

Figura 4.3.- Formatos de muestreo (a) y representación de píxeles en formato 4:2:0 (b)

En el código de partida se comprueba la compatibilidad del tamaño de la imagen de entrada con el tamaño seleccionado para el MB. Para lograr que esto se cumpla tras haber diezclado se emplea de nuevo la función *cropBorders()*, que recorta las columnas necesarias de la imagen de modo que el tamaño resultante no presente problemas de compatibilidad. Si las filas que es necesario recortar se denominan *cropRowsMB*, y las columnas *cropColsMB*, la secuencia de LR será de resolución $[((m - \text{cropRowsSC}) / \text{Scale Factor}) - \text{cropRowsMB}] \times [((n - \text{cropColsSC}) / \text{Scale Factor}) - \text{cropColsMB}]$ píxeles.

Como se puede observar, en función de los parámetros seleccionados, el tamaño final de la imagen variará. Además, es necesario emplear diferentes identificadores para los tamaños que aparecen a lo largo del proceso. Hacer uso en dos ocasiones de la función *cropBorders()* requiere la creación de una nueva resolución, denominada en este caso ULR (*Upper Lower Resolution*). La Tabla 4.1 muestra la solución planteada para el siguiente caso de referencia:

- Secuencia: *galdar_O_352x288.yuv*.
- Factor de Escala: 3.
- Tamaño de MB: 16.

Tabla 4.1.- Ejemplo de la compatibilidad de tamaños

Situación	Filas (píxeles)	Columnas (píxeles)	Resolución
Imagen original	288	352	VHHR
Secuencia tras <i>cropBorders_{SC}</i>	288	348	LHR
Secuencia diezmada	96	116	ULR
Secuencia tras <i>cropBorders_{MB}</i>	96	112	LR
Imagen resultante	288	336	HR

La Tabla 4.2 resume de manera genérica las resoluciones consideradas en el dimensionamiento de las imágenes descrito.

Tabla 4.2.- Resoluciones consideradas

Resolución	Filas (píxeles)	Columnas (píxeles)
VHHR	m	n
LHR	$(m - cropRowsSC)$	$(n - cropColsSC)$
ULR	$[(m - cropRowsSC) / Scale Factor]$	$[(n - cropColsSC) / Scale Factor]$
LR	$[((m - cropRowsSC) / Scale Factor) - cropRowsMB]$	$[((n - cropColsSC) / Scale Factor) - cropColsMB]$
HR	$[m - cropRowsSC - (cropRowsMB \times Scale Factor)]$	$[n - cropColsSC - (cropColsMB \times Scale Factor)]$

Finalmente, con el fin de generar la imagen necesaria para llevar a cabo el cálculo del PSNR y del SSIM en relación con la imagen súper-resuelta y la imagen interpolada, es necesario partir del resultado obtenido por la función *cropBorders()* para la compatibilidad del tamaño con el factor de diezmado, empleando de nuevo la misma función con el fin de recortar las filas y columnas necesarias para compensar el recorte que se realizará posteriormente, asegurando así la compatibilidad del tamaño de MB seleccionado.

4.2.2 Estudio sobre la unificación de la función *cropBorders()*

Durante el desarrollo de la presente Tesis Doctoral, se intentaron unificar los dos bloques en los que se emplea la función *cropBorders()*. De acuerdo al procedimiento descrito, el primero se emplea para la compatibilidad con el Factor de Escala, y el segundo para la compatibilidad con el tamaño de MB seleccionado.

A partir de la Figura 4.2, se ha desarrollado un estudio, con el fin de obtener las condiciones que se tendrían que cumplir para poder implementar un solo bloque que unifique los criterios requeridos por ambos. La primera secuencia, de tamaño VHHR (*Very High Resolution*), tiene unas dimensiones de $m \times n$ píxeles. De este modo:

$$\text{Si } [m / (2 * \text{Scale Factor})] \neq \text{entero} \rightarrow m' = m - \text{resto}[m / (2 * \text{Scale Factor})] \quad (4.1)$$

$$\text{Si } [n / (2 * \text{Scale Factor})] \neq \text{entero} \rightarrow n' = n - \text{resto}[n / (2 * \text{Scale Factor})] \quad (4.2)$$

Es decir, si estas dimensiones no son compatibles con el doble del Factor de Escala indicado, será necesario aplicar un recorte en los bordes de la imagen, un número de filas y columnas (*cropRowsSC* y *cropColsSC*) tal que el tamaño resultante obtenido, m' y n' , respectivamente sea divisible por el factor de diezmado. Esta nueva resolución de imagen se denomina LHR (*Lower High Resolution*). Resulta obvio que *cropRowsSC* y *cropColsSC* se obtienen a partir de las ecuaciones anteriores del siguiente modo:

$$\text{cropRowsSC} = \text{resto}[m / (2 * \text{Scale Factor})] \quad (4.3)$$

$$\text{cropColsSC} = \text{resto}[n / (2 * \text{Scale Factor})] \quad (4.4)$$

Posteriormente se aplica el diezmado y se obtiene la secuencia de tamaño ULR (*Upper Lower Resolution*), de $P \times Q$ píxeles, donde:

$$P = m' / (2 * \text{Scale Factor}) \quad (4.5)$$

$$Q = n' / (2 * \text{Scale Factor}) \quad (4.6)$$

Por otro lado, si este tamaño no es compatible con el correspondiente al tamaño de MB establecido, es necesario también recortar tantas filas y columnas (*cropRowsMB* y *cropColsMB*) como sea necesario para poder procesar la información:

$$\text{Si } P / \text{MB_SIZE} \neq \text{entero} \rightarrow P' = P - \text{resto}(P / \text{MB_SIZE}) \quad (4.7)$$

$$\text{Si } Q / \text{MB_SIZE} \neq \text{entero} \rightarrow Q' = Q - \text{resto}(Q / \text{MB_SIZE}) \quad (4.8)$$

De estas expresiones se deduce que el modo de obtener el número filas y columnas a recortar es:

$$\text{cropRowsMB} = \text{resto}(P / \text{MB_SIZE}) \quad (4.9)$$

$$\text{cropColsMB} = \text{resto}(Q / \text{MB_SIZE}) \quad (4.10)$$

Para poder integrar las dos funcionalidades en un solo bloque es imprescindible cumplir las siguientes condiciones, o conseguir que se unifiquen en una sola operación:

$$[m / (2 * \text{Scale Factor})] \equiv \text{entero} \quad , \quad [n / (2 * \text{Scale Factor})] \equiv \text{entero} \quad (4.11)$$

$$P / \text{MB_SIZE} \equiv \text{entero} \quad , \quad Q / \text{MB_SIZE} \equiv \text{entero} \quad (4.12)$$

Para ello, habría que obtener el número de filas y columnas que es necesario recortar de forma que, a partir de este punto, los siguientes tamaños de imagen que se vayan obteniendo sean compatibles con cada uno de los procesos a ejecutar posteriormente. Ante la problemática que esto conlleva, y siendo posible que no exista una solución válida para este planteamiento, se decidió finalmente mantener dos bloques *cropBorders()* independientes.

4.3 ANÁLISIS DE SIMULACIONES

Las simulaciones de este capítulo se desglosan a partir de las siguientes aportaciones algorítmicas:

- Simulaciones con Filtro Selectivo de Ventana (*Window Selective Filter*, WSF).
- Simulaciones con Filtro Selectivo de MB (*Block Selective Filter*, BSF).
- Simulaciones con topología de MBs adaptable (*Variable Block Size*, VBS).

4.3.1 Simulaciones con Filtro Selectivo de Ventana

Se presentan en este apartado las simulaciones realizadas con Filtro Selectivo de Ventana (*Window Selective Filter Super-Resolution*, WSFSR) que han considerado secuencias típicamente utilizadas para la comparación entre diferentes propuestas. Estas secuencias se detallan en el Anexo A.1 de la presente Tesis Doctoral. En la Tabla 4.3 se presenta un resumen de los resultados obtenidos. Esta tabla presenta las siguientes columnas:

- **ID:** Identificador de secuencia. Se presentan un total de 9 secuencias.
- **Secuencia:** Nombre de la secuencia.
- **SA:** Área de Búsqueda (*Search Area*), en píxeles, con la que se obtienen los mejores resultados.
- **WSFTHR:** Umbral (*Window Selective Filter THReshold*) óptimo para la secuencia en cuestión. En el siguiente capítulo se detalla el proceso de su obtención.
- **MBS:** Tamaño de MB (*Macro-Block Size*), en píxeles, con la que se obtienen los mejores resultados.
- **Diferencia de PSNR:** Diferencias del valor de PSNR entre BSR e interpolación bilineal (INT) y entre BSR y WSFSR considerando valor de umbral WSFTHR óptimo.
- **Diferencia de SSIM:** Diferencias del valor de SSIM entre BSR e INT en porcentaje, y entre BSR y WSFSR, considerando el valor de umbral WSFTHR óptimo.
- **T:** Porcentaje de reducción de tiempo de cómputo de WSFSR frente a BSR.

Tabla 4.3.- Resumen de resultados con WSFSR

ID	Secuencia	SA	MBS	WSFTHR	PSNR (dB)		SSIM		T
					BSR-INT	WSFSR-INT	BSR-INT	WSFSR-INT	
1	<i>Deadline</i>	2	8	80%	-0.3	0.01	-2.32%	0.19%	64%
2	<i>Flower</i>	4	32	30%	-1.76	0.4	-13.66%	2.23%	55%
3	<i>Football</i>	16	4	80%	-7.49	0.05	-31.30%	0.2%	90%
4	<i>Foreman</i>	4	16	80%	-2.09	0.58	-5.33%	0.8%	56%
5	<i>Galdar</i>	2	32	20%	2.98	3.18	7.18%	0.76%	5%
6	<i>Mobcal</i>	4	32	50%	-1.04	0.63	-2.68%	1.53%	58%
7	<i>Mobile</i>	4	16	20%	0.56	1.22	5.92%	7.29%	28%
8	<i>Stockholm</i>	8	16	40%	-3.64	0.55	-3.09%	0.71%	77%
9	<i>Suzie</i>	4	4	70%	0.26	0.51	0.29%	0.71%	30%

En la Tabla 4.3 se observa que el umbral óptimo del Filtro Selectivo de Ventana (WSFTHR) oscila entre valores que van desde un 20% hasta un 80%. Dichos valores guardan una estrecha relación con la cantidad de movimiento local y /o global de la secuencia en cuestión. La Tabla 4.4 presenta una relación para cada secuencia entre la cantidad de movimiento global y local medida subjetivamente, y el umbral óptimo del Filtro Selectivo de Ventana.

Tabla 4.4.- Relación entre movimiento global, local y valor óptimo del umbral de WSFSR

ID	Secuencia	SA	MBS	Tamaño (píxeles)	Movimiento Global	Movimiento Local	WSFTHR
1	<i>Deadline</i>	2	8	352x288	Nulo	Medio	80%
2	<i>Flower</i>	4	32	352x288	Medio	Bajo	30%
3	<i>Football</i>	2	4	352x288	Bajo	Muy Alto	80%
4	<i>Foreman</i>	4	16	176x144	Medio	Alto	80%
5	<i>Galdar</i>	2	32	352x288	Bajo	Nulo	20%
6	<i>Mobcal</i>	4	32	480x360	Medio	Medio	50%
7	<i>Mobile</i>	4	16	176x144	Medio	Bajo	30%
8	<i>Stockholm</i>	8	16	704x576	Medio	Bajo	40%
9	<i>Suzie</i>	4	4	176x144	Nulo	Alto	70%

En la Tabla 4.4 se observa cómo las secuencias con movimiento global medio-bajo y movimiento local medio-bajo se ajustan a valores del umbral WSFTHR en torno al 30%. Esto ocurre para las secuencias *Flower*, *Galdar*, *Mobile* o *Stockholm*. Por otra parte, secuencias con movimiento local elevado y movimiento global bajo o nulo, presentan valores de umbral elevados, como ocurre con las secuencias *Deadline*, *Football*, *Foreman* o *Suzie*. Esto es coherente considerando que la métrica que se usa para tomar decisiones con respecto al número de *frames* a utilizar es el SSIM, que mide cuánto se parecen estructuralmente dos *frames* determinados. Cuando existe un movimiento local acusado los *frames* no se ajustarán según esta métrica y, por tanto, el valor del umbral tendrá que elevarse para buscar una similitud adecuada entre *frames*. En las simulaciones realizadas se ha comprobado que el Filtro Selectivo de Ventana opera de tal manera que selecciona los *frames* adecuados para maximizar la calidad, de tal forma que en tanto en cuanto la información que aporte un *frame* sea significativa para el proceso de SR se considerará, mientras que se prescindirá de éste en caso contrario. Con el objetivo de evaluar adecuadamente los resultados obtenidos, se obtuvieron gráficas que justificasen el correcto funcionamiento del filtro y que determinaran en qué medida la calidad se veía aumentada, así como la disminución del tiempo de cómputo cuando se utiliza un valor de umbral WSFTHR adecuado. Las gráficas utilizadas fueron:

- Diferencias de métricas promedio frente al valor del umbral WSFTHR.
- Tiempo de ejecución frente al valor del umbral WSFTHR.
- Métricas para el valor óptimo del umbral WSFTHR para cada *frame*.
- Tamaño de Ventana de Trabajo, en *frames*, para cada *frame* analizado.
- Medida subjetiva de la calidad.

A continuación, se presentarán a modo de ejemplo cada una de estas gráficas y la información que aportan para los 30 primeros *frames* de la secuencia *Foreman*, así como algunas particularizaciones del resto de secuencias analizadas.

4.3.1.1 Diferencia de métricas promedio frente al valor del umbral WSFTHR

En primer lugar, se busca una forma de caracterizar el umbral WSFTHR del Filtro Selectivo de Ventana. Para ello, se muestran gráficas que comparan la diferencia promedio de una determinada métrica (PSNR o SSIM) del número de *frames* analizados en el proceso de SR frente al valor del umbral del Filtro Selectivo de Ventana utilizado, de acuerdo con las ecuaciones (4.13) y (4.14).

$$DIF_PSNR_{WSFTHR} = \frac{\sum_{i=1}^n (PSNR_WSFSR_{WSFTHRi} - PSNR_BSR_i)}{n} \quad (4.13)$$

$$DIF_SSIM_{WSFTHR} = \frac{\sum_{i=1}^n (SSIM_WSFSR_{WSFTHRi} - SSIM_BSR_i)}{n}, \quad (4.14)$$

donde:

- $PSNR_WSFSR_{WSFTHRi}$: Es el valor de PSNR del *frame i* súper-resuelto considerando el umbral WSFTHR, frente al *frame i* original.
- $SSIM_WSFSR_{WSFTHRi}$: Es el valor de SSIM del *frame i* súper-resuelto considerando el umbral WSFTHR, frente al *frame i* original.
- $PSNR_BSR_i$: Es el valor de PSNR del *frame i* súper-resuelto con el algoritmo BSR, frente al *frame i* original.
- $SSIM_BSR_i$: Es el valor de SSIM del *frame i* súper-resuelto con el algoritmo BSR, frente al *frame i* original.
- n : Es el número de *frames* totales de la secuencia analizada. En el caso del presente ejemplo, $n = 30$.

En la Figura 4.4 se presentan las gráficas de diferencia de métricas promedio en términos de PSNR (izquierda) y SSIM (derecha) para la secuencia *Foreman*. En éstas se puede observar cómo se maximiza la calidad para un valor de WSFTHR = 80% tanto para PSNR como para SSIM. Recuérdese que un valor de WSFTHR = 0% se corresponde con BSR, de ahí que la diferencia ($PSNR_WSFR_{0i} - PSNR_BSR_{0i}$) sea nula para todo i . Por otra parte, WSFTHR = 100% se identifica con la interpolación bilineal INT.

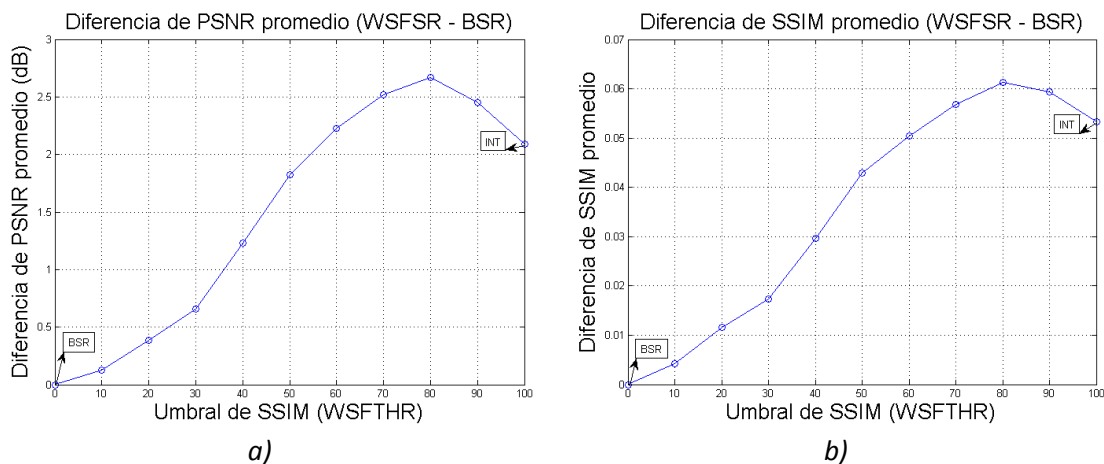


Figura 4.4.- Diferencia entre WSFSR y BSR respecto a WSFTHR – Secuencia *Foreman*.
a) PSNR, b) SSIM

Las gráficas presentadas en la Figura 4.4 son muy intuitivas, observándose con claridad que en el punto máximo (WSFTHR = 80%) el resultado correspondiente a WSFSR queda prácticamente 2.6 dB por encima del de BSR en PSNR, y 0.06 (un 6%) en SSIM. Este umbral óptimo de un 80% se obtiene para la gran mayoría de las secuencias con movimiento global medio y movimiento local alto. Así, en la Figura 4.5 se presenta el contraste de la diferencia entre WSFSR y BSR para la secuencia *Suzie*, muy similar a la de la secuencia *Foreman*. Por otra parte, secuencias adecuadas para aplicar BSR, típicamente con movimiento local nulo y movimiento global bajo, presentan una situación de partida en BSR compleja de superar con WSFSR, lo que se traduce en un ligero máximo para el umbral óptimo, en este caso, por ejemplo, del 20% en la secuencia *Galdar*, como se puede observar en la Figura 4.6. Finalmente, secuencias con movimiento global medio y movimiento local medio o bajo presentarán el pico

de WSFSR en un umbral similar al caso de secuencias con movimiento global, como ocurre con la secuencia *Mobile*, donde el pico se encuentra también en el 20%, aunque se consigue una mayor mejora con respecto a BSR, como se representa en la Figura 4.7.

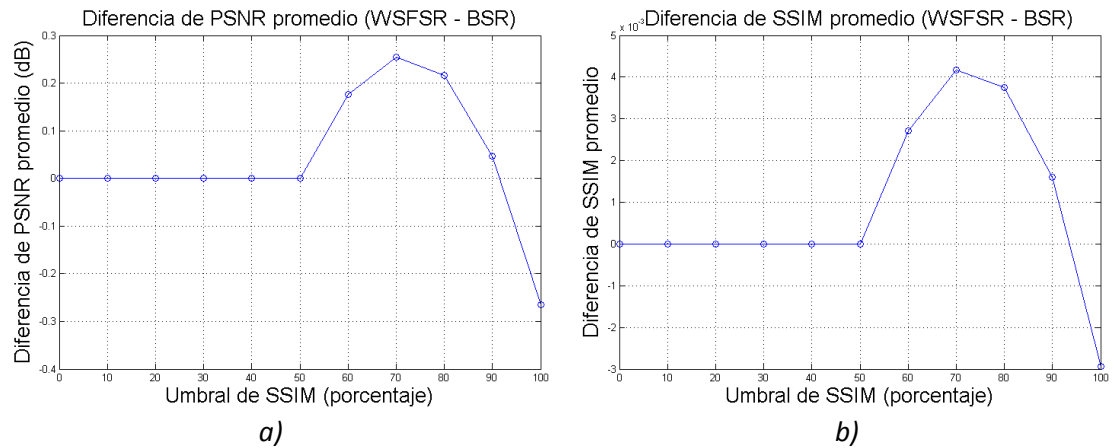


Figura 4.5.- Diferencia entre WSFSR y BSR respecto a WSFTHR – Secuencia Suzie.

a) PSNR, b) SSIM

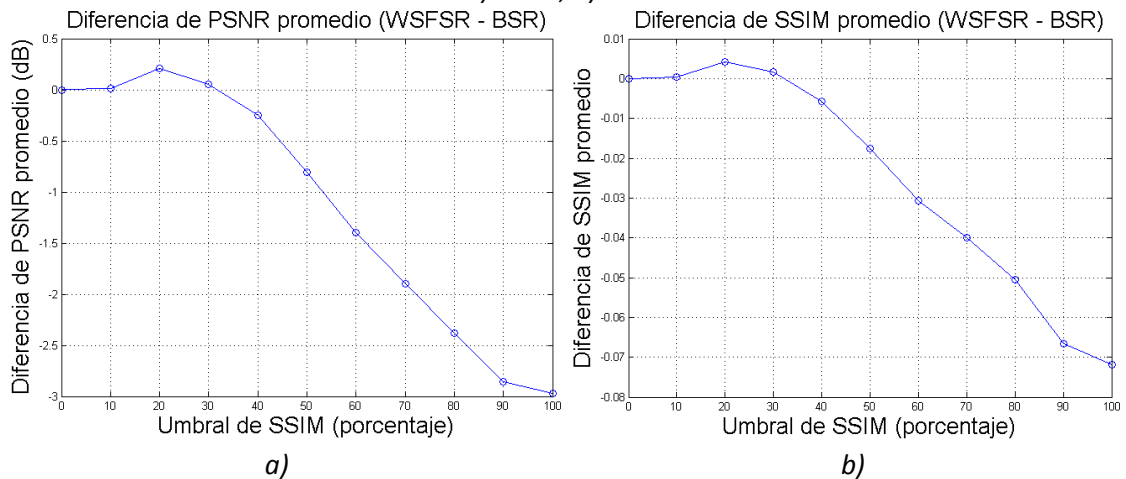


Figura 4.6.- Diferencia entre WSFSR y BSR respecto al umbral de SSIM – Secuencia Galdar.

a) PSNR, b) SSIM

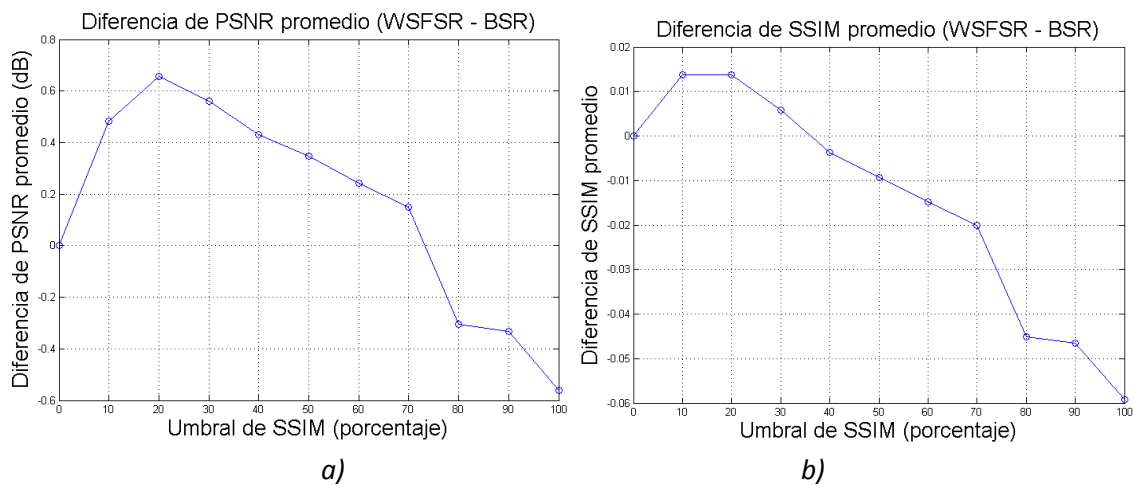


Figura 4.7.- Diferencia entre WSFSR y BSR respecto al umbral de SSIM – Secuencia Mobile.

a) PSNR, b) SSIM

4.3.1.2 Tiempo de ejecución frente al valor del umbral WSFTHR

Una vez caracterizada la calidad objetiva frente al umbral del filtro, resulta fundamental determinar cómo varía el tiempo de ejecución. Para ello se utilizan gráficas que comparan el tiempo de ejecución (en porcentaje con respecto al máximo) del proceso de SR frente al valor del umbral del Filtro Selectivo de Ventana utilizado. El máximo tiempo de ejecución se corresponderá para un valor de WSFTHR = 0% (BSR), mientras que será mínimo para WSFTHR = 100% (interpolación bilineal INT). En la Figura 4.8 se presenta la gráfica de tiempo de ejecución en porcentaje con respecto al máximo para la secuencia *Foreman*. Se observa cómo para un valor de umbral de un 80%, el tiempo de ejecución se reduce en aproximadamente un 56% para $WSFSR_{WSFTHR=80\%}$. Se interpreta además que el tiempo necesario para la interpolación es de un 36% del total.

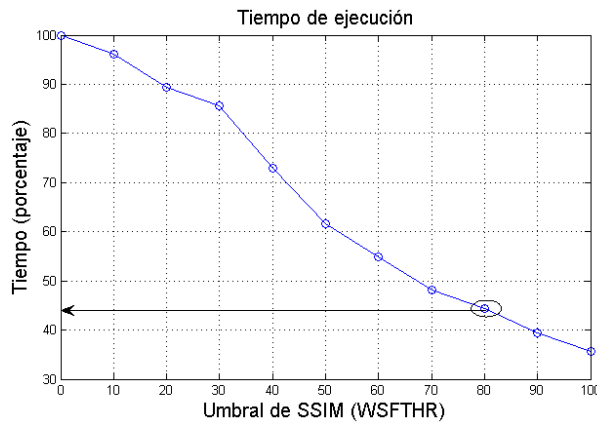


Figura 4.8.- Tiempo de ejecución respecto a WSFTHR – secuencia *Foreman*. Para WSFTHR = 80% el tiempo es un 44% sobre BSR

La disminución del tiempo de ejecución, está directamente ligada con el descarte de *frames* que se realiza de acuerdo con el decisor que analiza el parecido de estos *frames* frente al umbral WSFTHR. De esta forma, en secuencias como *Galdar*, con movimiento local nulo y movimiento global bajo, la reducción del tiempo de ejecución será gradual desde valores bajos, encontrándose el valor óptimo cerca del 100% de tiempo (ver Figura 4.9.a), ya que en este caso es adecuado utilizar la gran mayoría de los *frames* en el proceso de SR, debido a la alta correlación existente entre estos (umbral óptimo 20%). Por otra parte, en secuencias sin movimiento global y con movimiento local medio o alto, el descarte de *frames* se producirá a valores mayores del umbral, ya que la estructura de la secuencia será similar para valores pequeños del umbral, como se representa en la Figura 4.9.b para la secuencia *Suzie*.

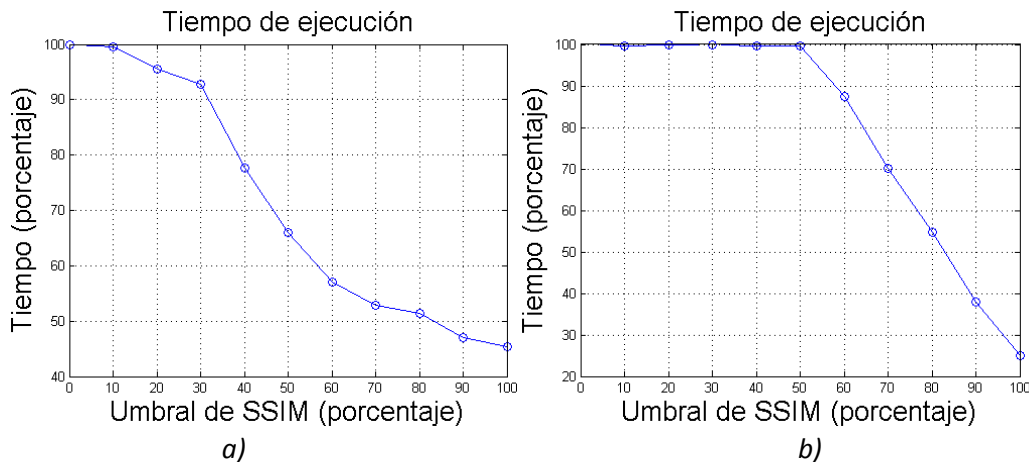


Figura 4.9.- Variación porcentual de tiempo respecto al umbral. a) *Galdar*, b) *Suzie*

4.3.1.3 Métricas para el valor óptimo del umbral WSFTHR para cada *frame*

Habiendo caracterizado la calidad objetiva y el tiempo de ejecución con respecto al umbral WSFTHR, se busca representar el valor de una determinada métrica (PSNR o SSIM) para cada uno de los *frames* analizados en el proceso de SR, considerando el valor óptimo del umbral, y analizando tanto WSFSR como BSR e interpolación (INT). A modo de ejemplo, la Figura 4.10 representa los valores de PSNR (izquierda) y SSIM (derecha) correspondientes a cada uno de los 30 primeros *frames* de la secuencia *Foreman*.

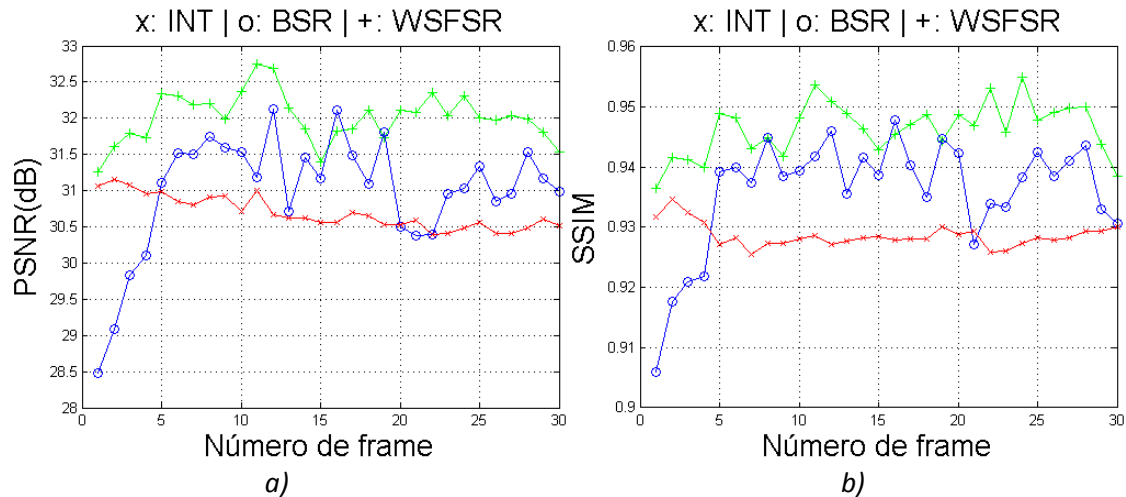


Figura 4.10.- Secuencia *Foreman* para umbral óptimo = 80%. a) PSNR, b) SSIM

Se observa en la Figura 4.10 cómo la calidad objetiva de la secuencia *Foreman* para WSFSR se mantiene siempre por encima de BSR y de INT, tanto en términos de PSNR como de SSIM, igualando los valores de BSR en algunos casos donde no es posible mejorar la calidad. Es también de destacar que la variabilidad entre valores de PSNR y SSIM es más estable, existiendo aproximadamente 1.5 dB de diferencia en PSNR para WSFSR entre máximo y mínimo, y de 4 dB de diferencia en PSNR para BSR entre máximo y mínimo. La situación para SSIM es similar, existiendo aproximadamente un 2% de diferencia para WSFSR entre máximo y mínimo y un 7% de diferencia para BSR entre máximo y mínimo. Esto es importante, ya que no sólo se mejora la calidad, sino que se estabiliza en la secuencia. El objetivo será siempre que WSFSR esté por encima de las curvas de interpolación y de BSR, aunque la situación de partida puede variar. Así, por ejemplo la secuencia *Galdar* presenta unos resultados para BSR muy superiores a los de la interpolación, superada ligeramente por WSFSR cuando se usan los *frames* más adecuados, como se representa en la Figura 4.11. Por otra parte, en secuencias en las que el uso de un elevado número de *frames* produce artefactos al aplicar SR, como es el caso de *Flower*, la situación será inversa (interpolación por encima de BSR), como se muestra en la Figura 4.12.

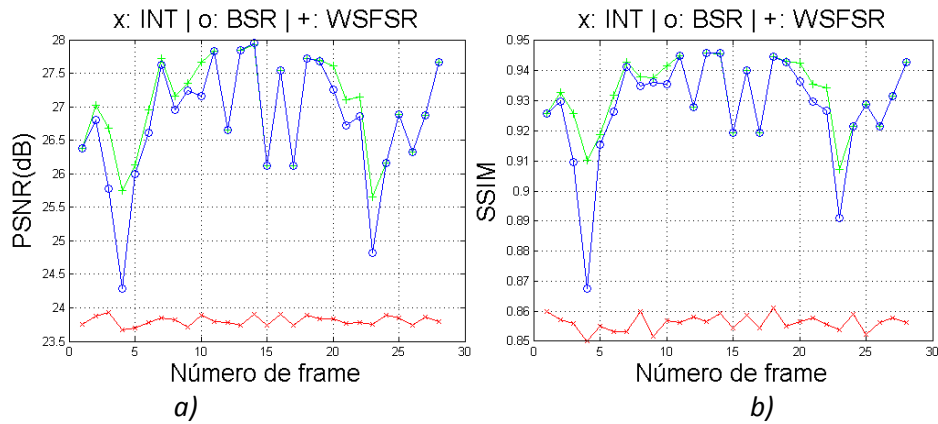


Figura 4.11.- Secuencia Galdar para umbral óptimo = 20%. a) PSNR, b) SSIM

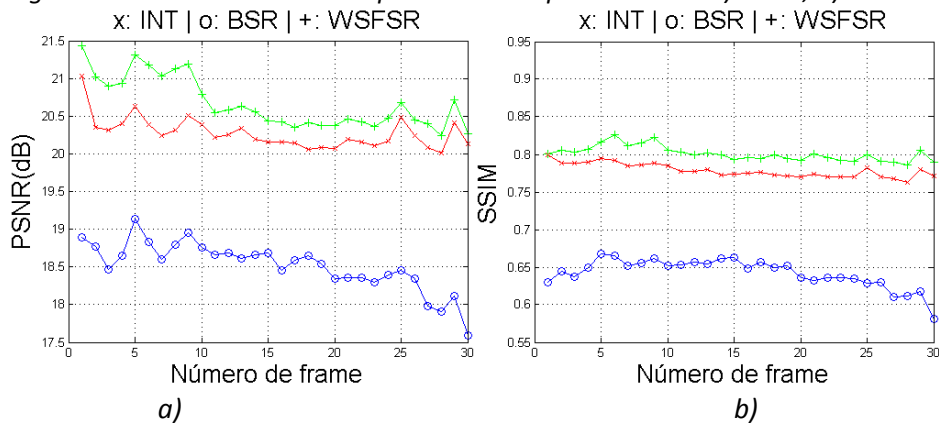


Figura 4.12.- Secuencia Flower para umbral óptimo = 30%. a) PSNR, b) SSIM

4.3.1.4 Tamaño de Ventana de Trabajo en frames para cada frame analizado

Finalmente, sería interesante valorar cómo varía el tamaño de la Ventana de Trabajo (en frames) dependiendo del frame analizado en la secuencia. Para ello se presentan gráficas que especifican el valor de la Ventana de Trabajo de SR utilizada para cada uno de los frames analizados en el proceso de SR considerando el valor óptimo del Filtro Selectivo de Ventana. Para el ejemplo de referencia se utilizará un número de frames totales de 30. Por otra parte, la Ventana de Trabajo de SR utilizada por BSR es de +/-20. Esto significa que, como se representa en la Figura 4.13, los frames utilizados en el proceso de SR para BSR irán desde 21 para el frame nº1 (el actual más los 20 posteriores) y el frame nº 30 (el actual y los 20 anteriores) hasta un máximo de 30 (caso particular de los frames 10 hasta el 21). Por otra parte, los frames utilizados para WSFSR se seleccionarán dependiendo de la decisión tomada por la métrica SSIM con respecto al umbral de decisión establecido. Así, en la Figura 4.13 se pueden ver los siguientes casos:

- Para el frame nº1, WSFSR utiliza tan solo 2 frames (el actual y 1 posterior) frente a los 21 frames utilizados en BSR. Esto se traduce en un incremento de calidad de 0.2 dB en PSNR (0.05% en SSIM) frente a INT y de casi 3 dB (3% en SSIM) frente a BSR.
- Para el frame nº5, WSFSR utiliza 4 frames frente a los 25 frames utilizados en BSR. Esto se traduce en un incremento de calidad de 1.4 dB (2% en SSIM) frente a INT y de 1.3 dB (1% en SSIM) frente a BSR.
- Para el frame nº16, WSFSR utiliza 8 frames, frente a los 30 frames utilizados en BSR. Es el único caso de los 30 frames analizados en el que BSR se encuentra por encima de WSFSR en PSNR (0.2 dB) y SSIM (0.2%). En cualquier caso, la diferencia es prácticamente nula.

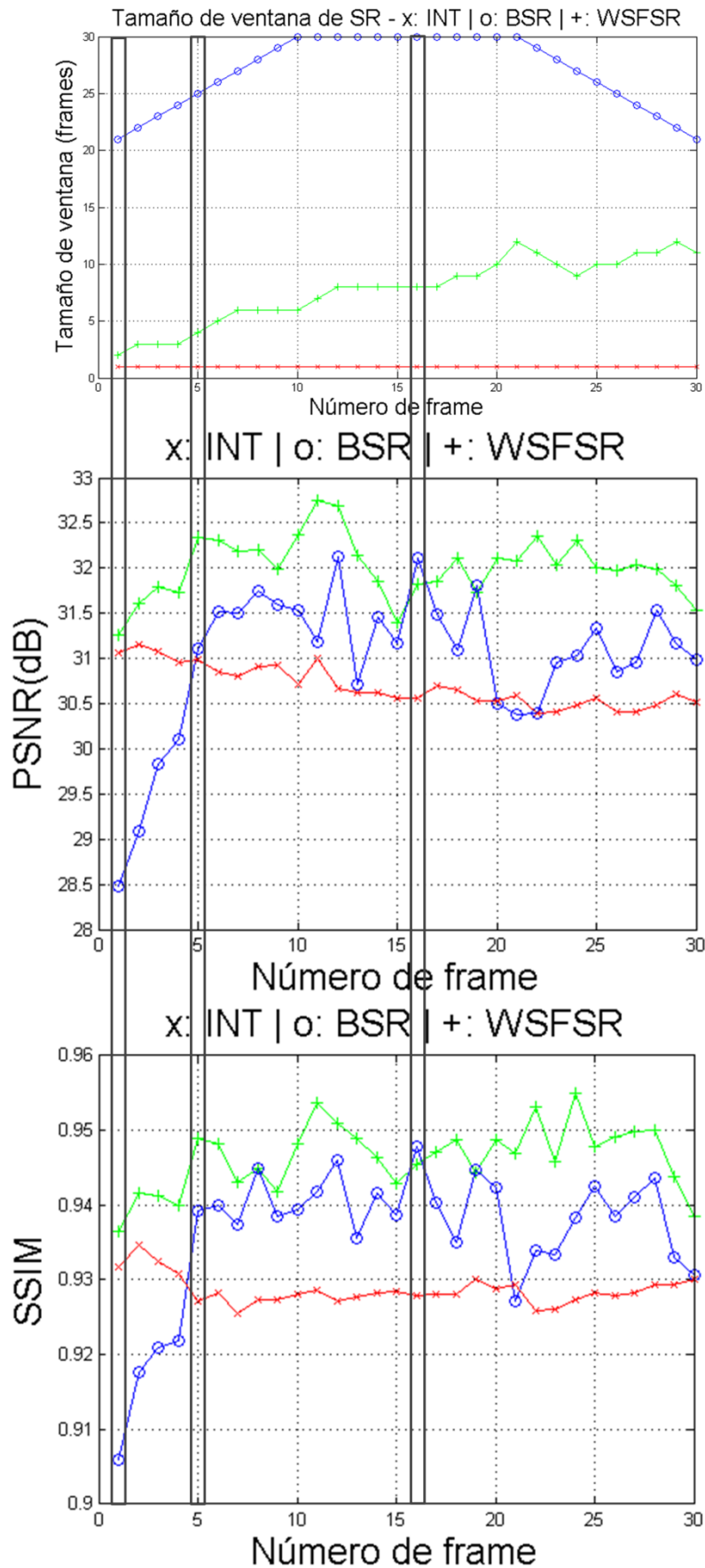


Figura 4.13.- Ventana de SR, PSNR y SSIM con respecto a número de frame – secuencia Foreman. Se analizan diferentes valores de umbral (de arriba abajo: 0, 80 y 100)

Tal y como se comentó en la sección 4.3.1.2, la disminución del tiempo de cómputo a medida que aumenta el umbral se debe al descarte del número de *frames*. Se presentan aquí los casos particulares de las secuencias *Galdar* y *Suzie*, en los que se observa cómo para *Galdar*, cuyas curvas de tamaño de Ventana de Trabajo están representadas en la Figura 4.14, se presenta una disminución progresiva a medida que aumenta el umbral, mientras que para *Suzie*, la situación, representada en la Figura 4.15, es más abrupta, descartando *frames* a valores altos del umbral.

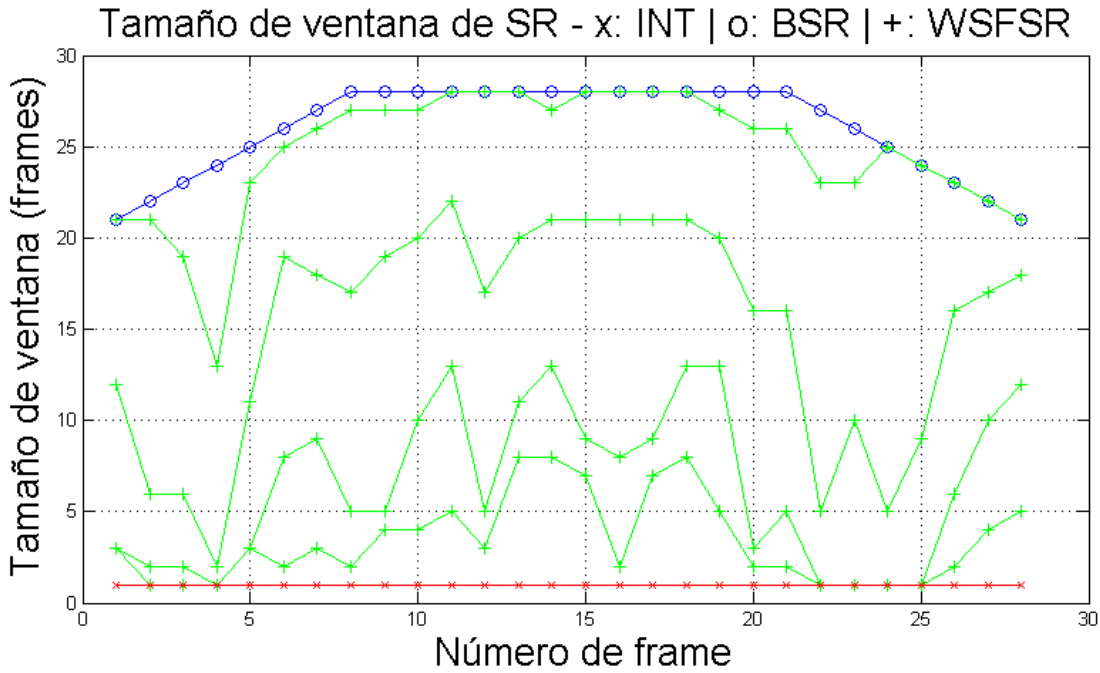


Figura 4.14.- Ventana de Trabajo de SR con respecto a número de frame – secuencia *Galdar*. Se usan diferentes valores de umbral (de arriba abajo: 0, 20, 40, 60, 80 y 100)

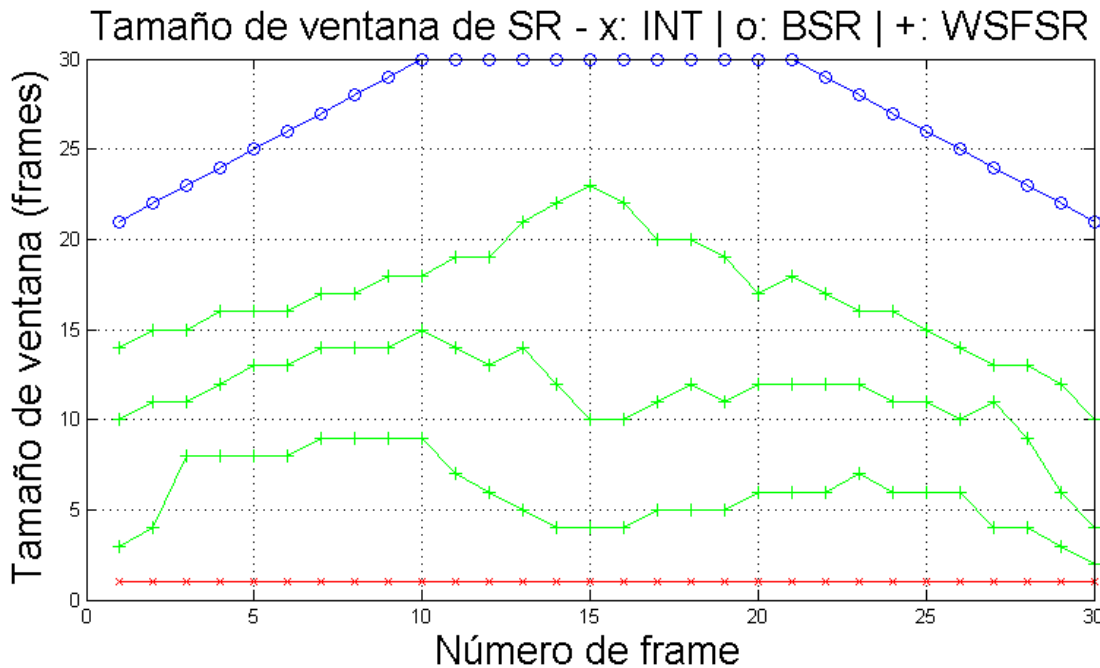


Figura 4.15.- Ventana de Trabajo de SR con respecto a número de frame – secuencia *Suzie*. Se usan diferentes valores de umbral (de arriba abajo: 0, 70, 80, 90 y 100)

4.3.1.5 Medida subjetiva de calidad

La medida subjetiva de calidad se basa en la comparación directa de los *frames* de salida. Para ello se comparan WSFSR, BSR, e INT con la secuencia original. Aun siendo una medición que depende del sujeto que la realice, es adecuada para tomar decisiones en la mejora continua de los algoritmos propuestos.

En la Figura 4.16 se presenta el *frame* 20 de la secuencia *Foreman*, observándose que en este caso, una selección adecuada de *frames* para el cálculo del *frame* súper-resuelto a la salida (WSFSR selecciona para este *frame* un total de 24 *frames* en lugar de los 30 *frames* utilizados por BSR), evita la aparición de artefactos a la vez que reduce el tiempo de cómputo.

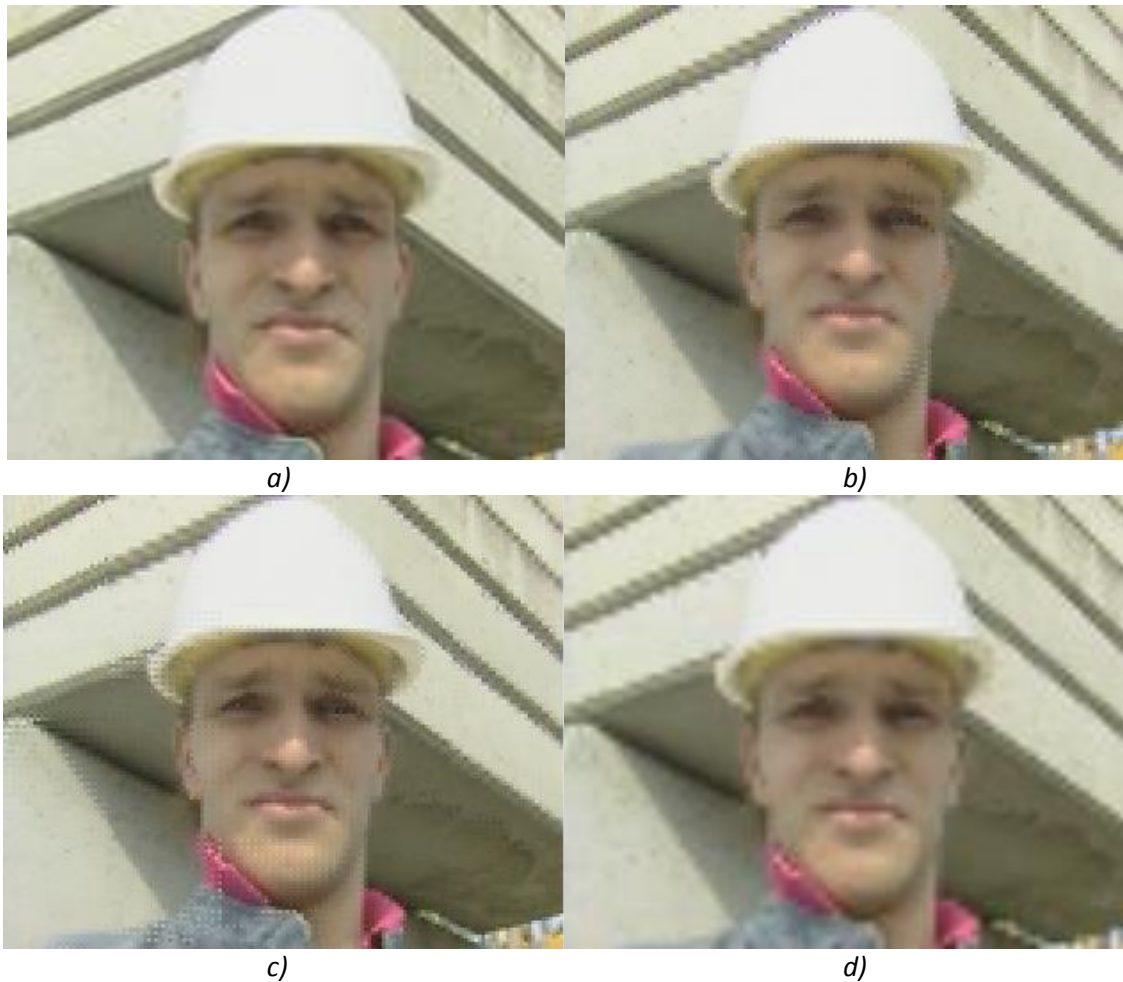


Figura 4.16.- Análisis del *frame* 20 de la secuencia *Foreman* QCIF para WSFSR.
a) Original, b) WSFSR, c) BSR, d) Interpolación

La mejora de calidad se da en la mayoría de los casos en la disminución de artefactos y el suavizado de bordes. En cualquier caso, cabría esperar que la aplicación de un filtro a nivel más fino (Macro-Bloque) mejorara aún más tanto la calidad objetiva como subjetiva a nivel de artefactos.

4.3.2 Simulaciones con Filtro Selectivo de Macro-Bloque

En este apartado se presentan las simulaciones realizadas con Filtro Selectivo de Macro-Bloque (*Block Selective Filter SR*, BSFSR) considerando las secuencias de referencia utilizadas en el apartado anterior. En la Tabla 4.5 se presenta un resumen de los resultados obtenidos. Esta tabla presenta las siguientes columnas:

- **ID:** Identificador de secuencia. Se presentan un total de 9 secuencias.
- **Secuencia:** Nombre de la secuencia.
- **SA:** Área de Búsqueda (*Search Area*), en píxeles, con la que se obtienen los mejores resultados.
- **BSFTHR:** Umbral (*Block Selective Filter THReshold*) óptimo para la secuencia en cuestión. En el siguiente capítulo se detalla el proceso de su obtención.
- **MBS:** Tamaño de MB (*Macro-Block Size*), en píxeles, con el que se obtienen los mejores resultados.
- **Diferencia de PSNR:** Diferencias del valor de PSNR entre BSR e interpolación bilineal (INT) y entre BSR y BSFSR, considerando el valor de umbral BSFTHR óptimo.
- **Diferencia de SSIM:** Diferencias del valor de SSIM entre BSR e INT en porcentaje y entre BSR y BSFSR, considerando el valor de umbral BSFTHR óptimo.
- **T:** Porcentaje de reducción de tiempo de BSFSR frente a BSR.

Tabla 4.5.- Resumen de resultados con BSFSR

ID	Secuencia	SA	MBS	BSFTHR	PSNR (dB)		SSIM		T
					BSR-INT	BSFSR-INT	BSR-INT	BSFSR-INT	
1	<i>Deadline</i>	2	8	2%	-0.3	0.01	-2.32%	0.1%	2%
2	<i>Flower</i>	4	32	6%	-1.76	0.1	-13.66%	1.08%	33%
3	<i>Football</i>	16	4	2%	-7.49	0.03	-31.30%	0.2%	4%
4	<i>Foreman</i>	4	16	4%	-2.09	0.59	-5.33%	0.3%	26%
5	<i>Galdar</i>	2	32	8%	2.98	3.18	7.18%	0.76%	4%
6	<i>Mobcal</i>	4	32	5%	-1.04	0.83	-2.68%	1.61%	3%
7	<i>Mobile</i>	4	16	6%	0.56	1.3	5.92%	7.49%	3%
8	<i>Stockholm</i>	8	16	8%	-3.64	0.52	-3.09%	0.94%	2%
9	<i>Suzie</i>	4	4	1%	0.26	0.36	0.29%	1.81%	13%

En la Tabla 4.5 se observa que el umbral óptimo del Filtro Selectivo de MB (BSFTHR) oscila entre valores comprendidos entre un 2% y un 8%. Dichos valores guardan una estrecha relación con la cantidad de movimiento local y /o global de la secuencia en cuestión. La Tabla 4.6 presenta una relación para cada secuencia entre la cantidad movimiento global y local medida subjetivamente, y el umbral BSFTHR óptimo del Filtro Selectivo de MB.

Tabla 4.6.- Relación entre movimiento global, local y valor óptimo del umbral de BSFSR

ID	Secuencia	SA	MBS	Tamaño (píxeles)	Movimiento Global	Movimiento Local	BSFTHR
1	<i>Deadline</i>	2	8	352x288	Nulo	Medio	2%
2	<i>Flower</i>	4	32	352x288	Medio	Bajo	6%
3	<i>Football</i>	16	4	352x288	Bajo	Muy Alto	2%
4	<i>Foreman</i>	4	16	176x144	Medio	Medio	4%
5	<i>Galdar</i>	2	32	352x288	Bajo	Nulo	8%
6	<i>Mobcal</i>	4	32	480x360	Medio	Medio	5%
7	<i>Mobile</i>	4	16	176x144	Medio	Bajo	6%
8	<i>Stockholm</i>	8	16	704x576	Medio	Bajo	8%
9	<i>Suzie</i>	4	4	176x144	Nulo	Alto	1%

En la Tabla 4.6 se observa cómo secuencias con movimiento global medio-bajo y movimiento local medio-bajo, se ajustan a valores del umbral de Filtro Selectivo de MB en torno al 6%. Esto ocurre para las secuencias *Flower*, *Galdar*, *Mobile* o *Stockholm*. Por otra parte, secuencias con movimiento local elevado y movimiento global reducido o nulo, presentan valores de umbral más bajos, como ocurre con las secuencias *Deadline*, *Football*, *Foreman* o *Suzie*. Esto es coherente considerando que la métrica que se usa para tomar decisiones con respecto al número de *frames* a utilizar es el MAD, que mide la diferencia de píxeles entre los *frames*. Cuando existe un movimiento local acusado los *frames* no se ajustarán según esta métrica y, por tanto, el valor del umbral tendrá que bajar para buscar un parecido adecuado entre *frames*. Por tanto, considerando la forma en la que se establece la comparación, la situación es inversa a la que se establecía en el Filtro Selectivo de Ventana. Así mismo, tal y como ocurría con los *frames* en el Filtro Selectivo de Ventana, en este caso se seleccionan los MBs adecuados para maximizar la calidad, de tal forma que en tanto en cuanto la información que aporte un MB resulte significativa para el proceso de SR se considerará, mientras que se prescindirá de éste en caso contrario. Con el objetivo de presentar adecuadamente los resultados obtenidos se obtuvieron gráficas que justificasen el correcto funcionamiento del filtro y que determinaran en qué medida la calidad se veía aumentada y el tiempo de cómputo disminuido cuando se utiliza un valor de umbral de Filtro Selectivo de MB adecuado. Las gráficas utilizadas para determinar los resultados fueron:

- Diferencias de métricas promedio frente al valor del umbral BSFTHR.
- Tiempo de ejecución frente al valor del umbral BSFTHR.
- Métricas para el valor óptimo del umbral BSFTHR para cada *frame*.
- Número de MBs utilizados para cada *frame* analizado.
- Medida subjetiva de la calidad.

4.3.2.1 Diferencia de métricas promedio frente al valor del umbral BSFTHR

Tal y como se procedió en el caso del análisis del Filtro Selectivo de Ventana, se compara en este apartado la diferencia promedio de una determinada métrica (PSNR o SSIM) del número de *frames* analizados en el proceso de SR frente al valor del umbral del Filtro Selectivo de MB utilizado, de acuerdo con las ecuaciones (4.15) y (4.16).

$$DIF_PSNR_{BSFTHR} = \frac{\sum_{i=1}^n (PSNR_BSFSR_{BSFTHRi} - PSNR_BSR_i)}{n} \quad (4.15)$$

$$DIF_SSIM_{BSFTHR} = \frac{\sum_{i=1}^n (SSIM_BSFSR_{BSFTHRi} - SSIM_BSR_i)}{n}, \quad (4.16)$$

donde:

- $PSNR_BSFSR_{BSFTHRi}$: Es el valor de PSNR del *frame* i súper-resuelto, considerando el umbral BSFTHR frente al *frame* i original.
- $SSIM_BSFSR_{BSFTHRi}$: Es el valor de SSIM del *frame* i súper-resuelto, considerando el umbral BSFTHR frente al *frame* i original.
- $PSNR_BSR_i$: Es el valor de PSNR del *frame* i súper-resuelto con el algoritmo BSR, frente al *frame* i original.
- $SSIM_BSR_i$: Es el valor de SSIM del *frame* i súper-resuelto con el algoritmo BSR, frente al *frame* i original.
- n : Es el número de *frames* totales de la secuencia analizada. En el caso del presente ejemplo, $n = 30$.

Se presentan en la Figura 4.17 las gráficas de diferencia de métricas promedio en términos de PSNR (izquierda) y SSIM (derecha) para la secuencia *Foreman*. En éstas se puede observar cómo se maximiza la calidad para un valor de BSFTHR = 4%, tanto para PSNR como para SSIM. Recuérdese que un valor de BSFTHR = 100% se corresponde con BSR, de ahí que la diferencia ($PSNR_{WSFR_{0i}} - PSNR_{BSR_{0i}}$) sea nula para todo i , aunque normalmente, como en el caso de esta secuencia, a partir de valores cercanos a BSFTHR = 20% se produce una saturación del umbral. Por otra parte, BSFTHR = 0% se identifica con la interpolación bilineal INT.

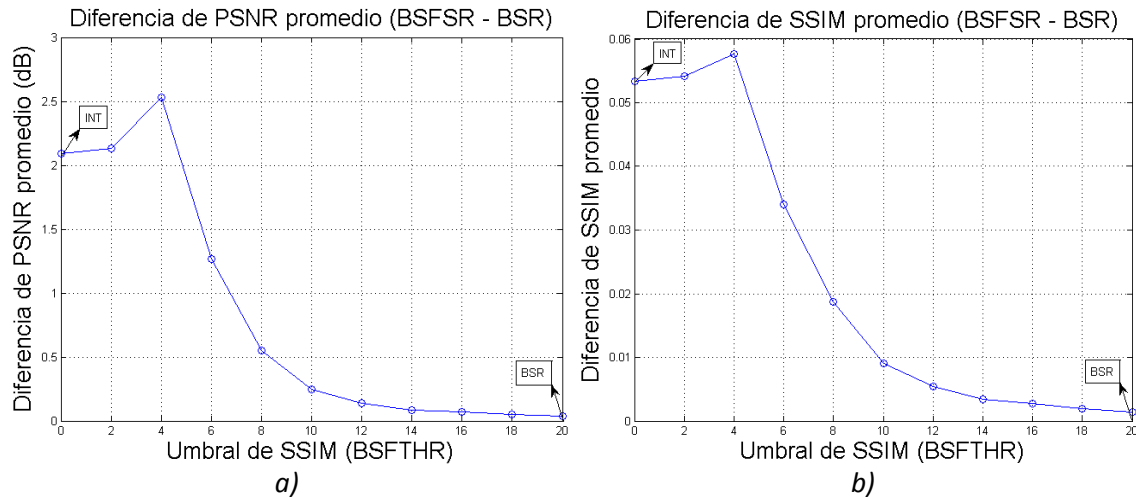


Figura 4.17.- Diferencia entre BSFSR y BSR respecto a BSFTHR – Secuencia *Foreman*.
a) PSNR, b) SSIM

4.3.2.2 Tiempo de ejecución frente al valor del umbral BSFTHR

De forma similar a como se procedió con el Filtro Selectivo de Ventana, se compara en este apartado el tiempo de ejecución (en porcentaje con respecto al tiempo que utiliza BSR) del proceso de SR frente al valor del umbral del Filtro Selectivo de MB utilizado. El máximo tiempo de cómputo se corresponderá para un valor de BSFTHR = 100%, mientras que será mínimo para BSFTHR = 0% (para este caso se toma como referencia el tiempo asociado a la interpolación). Se presenta en la Figura 4.18 la gráfica de tiempo de ejecución en porcentaje con respecto al máximo para la secuencia *Foreman*. Se observa cómo para un valor de umbral de un 4% el tiempo de ejecución se reduce en aproximadamente un 29% para $BSFSR_{BSFTHR=4\%}$.

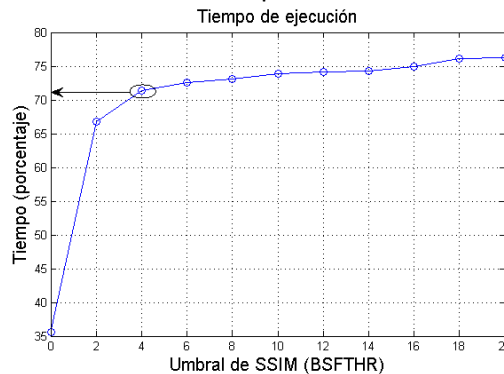


Figura 4.18.- Tiempo de ejecución respecto a BSFTHR – Secuencia *Foreman*. Para BSFTHR = 4% el tiempo es un 71% respecto a BSR

Tal y como se representa en la Figura 4.18, el tiempo de cómputo que optimiza la calidad para BSFSR es un 71% con respecto a BSR, mientras que el tiempo de ejecución que optimiza la calidad para WSFSR es un 44% con respecto a BSR.

4.3.2.3 Métricas para el valor óptimo del umbral BSFTHR para cada *frame*

Habiendo caracterizado la calidad objetiva y el tiempo de ejecución con respecto al umbral BSFTHR, se busca representar el valor de una determinada métrica (PSNR o SSIM) para cada uno de los *frames* analizados en el proceso de SR, considerando el valor óptimo del umbral BSFTHR, y analizando tanto BSFSR como BSR e interpolación (INT). A modo de ejemplo, la Figura 4.19 representa los valores de PSNR (izquierda) y SSIM (derecha) de cada uno de los 30 primeros *frames* de la secuencia *Foreman*.

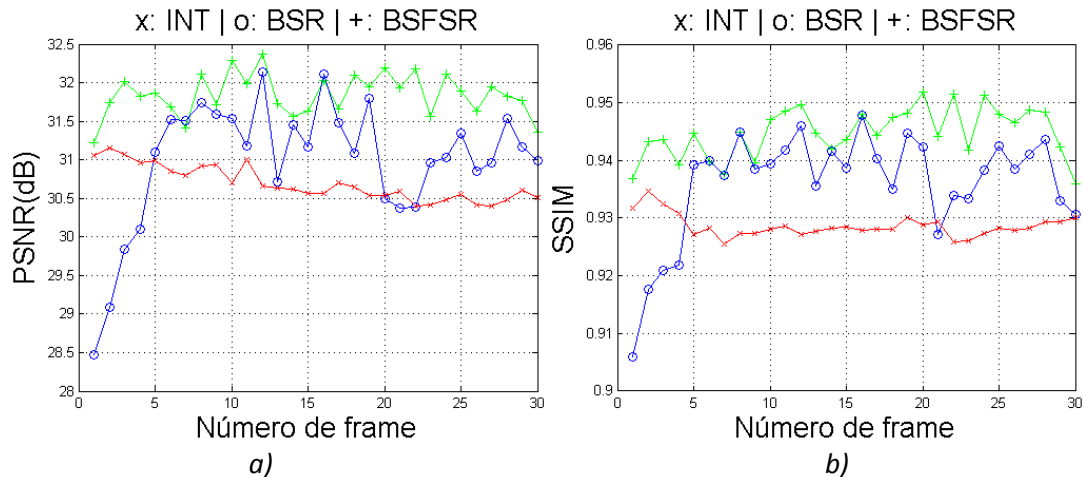


Figura 4.19.- Secuencia *Foreman* para umbral óptimo = 4%. a) PSNR, b) SSIM

4.3.2.4 Número de MBS utilizados para cada *frame* analizado

En este apartado, se presenta el procedimiento utilizado para visualizar el número de MBs utilizados para cada *frame* analizado, de forma análoga a como se representó el número de *frames* en la Ventana de Trabajo del Filtro Selectivo de Ventana para WSFSR. Se representa por tanto el valor del número de MBs utilizados con respecto al *frame* actual para cada uno de los *frames* analizados en el proceso de SR considerando el valor óptimo del Filtro Selectivo de MB. Para el ejemplo de referencia se utilizará, al igual que para el caso de WSFSR, un número de *frames* total de 30 y una Ventana de Trabajo de SR utilizada por BSR de +/-20. Esto significa que, como se representa en la Figura 4.28, los MBs utilizados en el proceso de SR para BSR irán desde 400 para el *frame* nº1 (20 MBs por *frame*, para un total de 20 *frames* posteriores) y el *frame* nº 30 (20 MBs por *frame*, para un total de 20 *frames* anteriores) hasta un máximo de 580 (caso particular de los *frames* 10 hasta el 21). Por otra parte, los MBs utilizados para BSFSR se seleccionarán dependiendo de la decisión tomada por la métrica SSIM con respecto al umbral establecido. Así, por ejemplo, considerando también la Figura 4.20 se pueden ver los siguientes casos:

- Para el *frame* nº1 BSFSR utiliza tan solo 295 MBs frente a los 400 MBs utilizados en BSR. Esto se traduce en un incremento de calidad de 0.2 dB en PSNR (0.05% en SSIM) frente a INT, y de casi 3 dB (3% en SSIM) frente a BSR, al igual que ocurría con WSFSR.
- Para el *frame* nº5 BSFSR utiliza 455 MBs frente a los 480 MBs utilizados en BSR. Esto se traduce en un incremento de calidad de 0.8 dB (1.5% en SSIM) frente a INT y de 0.6 dB (0.05% en SSIM) frente a BSR, que son valores ligeramente inferiores a los resultados de BSFSR.
- Para el *frame* nº16 BSFSR utiliza 560 MBs, frente a los 580 MBs utilizados en BSR. Se obtienen prácticamente los mismos resultados que para BSR, mejorando ligeramente los resultados de BSFSR.

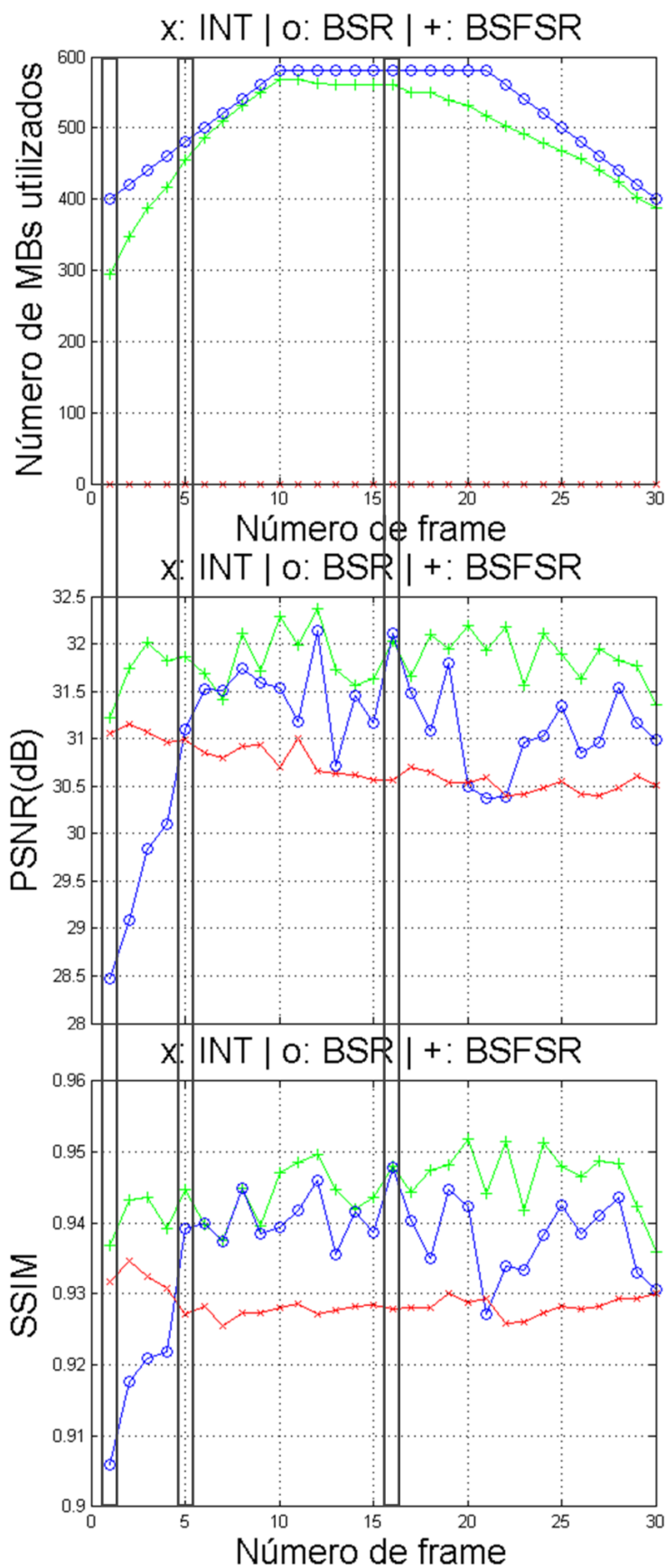


Figura 4.20.- Número de MBs con respecto a número de frame, PSNR y SSIM – Secuencia Foreman. Se analizan diferentes valores de umbral (de abajo a arriba: 0, 4 y 20)

Por otra parte, tal y como se comentó en el capítulo *Contribuciones Algorítmicas*, si se considera un umbral adaptativo los resultados pueden mejorar aún más. Particularmente para el caso de la secuencia *Foreman* (para los 100 primeros *frames*) se presenta a modo de referencia en la Tabla 4.7 y en la Figura 4.21 la mejora de PSNR, tanto en media como en varianza. Así, la calidad al comparar el filtro de MB con umbral adaptativo y con umbral estático mejora en 0.45 dB de promedio con máximos de 1.95 dB para determinados *frames*, reduciéndose la varianza en un 54%. La mejora con respecto al algoritmo BSR es mucho mayor, obteniéndose una mejora total de 2.78 dB de media con máximos de 5.98 dB. Así, en los resultados obtenidos al aplicar un Filtro Selectivo de MB con un umbral adaptativo se consigue una mejora de calidad y una mayor estabilidad con respecto al Filtro Selectivo de MB con umbral estático. Se han realizado pruebas con varias secuencias, obteniendo resultados apropiados en secuencias que incluyen movimiento local.

Tabla 4.7.- Mejora de PSNR con umbral adaptativo

	BSR	BSF	BSF con umbral adaptativo
Media PSNR (dB)	28.57	30.90	31.35
Varianza (dB)	3.03	0.51	0.28

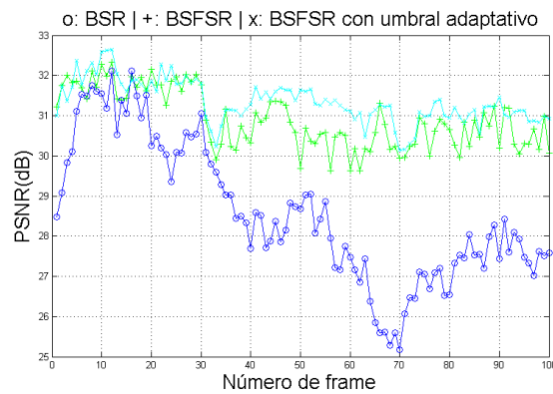


Figura 4.21.- Mejora de PSNR con umbral adaptativo frame a frame

Considerando las mejoras introducidas por WSFSR y por BSFSR de forma independiente, tanto en tiempo de cómputo como en calidad, surge el planteamiento de usar los dos filtros en cascada de forma que se use WSFSR como un filtro *grueso* y BSFSR como un filtro *fino*, conformando un filtro dual, que como se comentó en el capítulo *Contribuciones Algorítmicas* se denomina WBSFSR (*Window & Block Selective Filter SR*). Partiendo de esta idea se obtienen gráficas de calidad y tiempo de cómputo que comparan el funcionamiento de WSFSR con el de WBSFSR para diferentes valores del umbral WSFTHR considerando el valor de BSFTHR que optimiza la curva de WBSFSR. Los resultados, mostrados en la Figura 4.22, representan la comparativa con respecto a calidad y tiempo de cómputo, respectivamente.

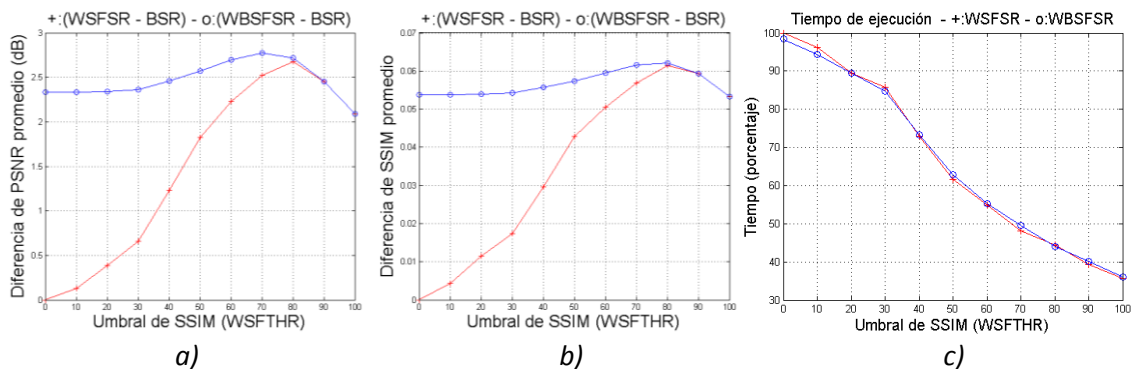


Figura 4.22.- Diferencias WSFSR-BSR y WBSFSR-BSR respecto a WSFTHR – Secuencia Foreman.
a) PSNR, b) SSIM c) Tiempo de ejecución

Tal y como se representa en la Figura 4.28, el aporte del valor óptimo del umbral del Filtro Selectivo de MB BSFSR (8%) al Filtro Selectivo de Ventana WSFSR es tal que el filtro conjunto WBSFSR se encuentra por encima de BSR y de la interpolación para cualquier valor de WSFTHR, obteniendo un máximo global superior a 2.7 dB en PSNR y a 6% en SSIM con un umbral WSFTHR = 70%, mejorando el máximo de WSFSR. Por otra parte, el tiempo de cómputo de WBSFSR es muy similar al de WSFSR, como se puede observar en la Figura 4.22, que representa los porcentajes de tiempo de cómputo utilizados con respecto al máximo (establecido para WSFSR trabajando en BSR que implica que WSFTHR = 0).

Tal y como se realizó en el Filtro Selectivo de Ventana, se representa a continuación el valor de una determinada métrica (PSNR o SSIM) para cada uno de los *frames* analizados en el proceso de SR considerando los valores óptimos de los filtros, analizando WBSFSR, BSFSR, WSFSR, BSR e INT. A modo de referencia, la Figura 4.23 representa los valores de PSNR (izquierda) y SSIM (derecha) de cada uno de los 30 primeros *frames* de la secuencia *Foreman*. Por otra parte, la Figura 4.24 particulariza los resultados para cada filtro. Se observa claramente el funcionamiento mejorado de WBSFSR con respecto al resto, obteniendo valores mejores que WSFSR en prácticamente todos los *frames* analizados.

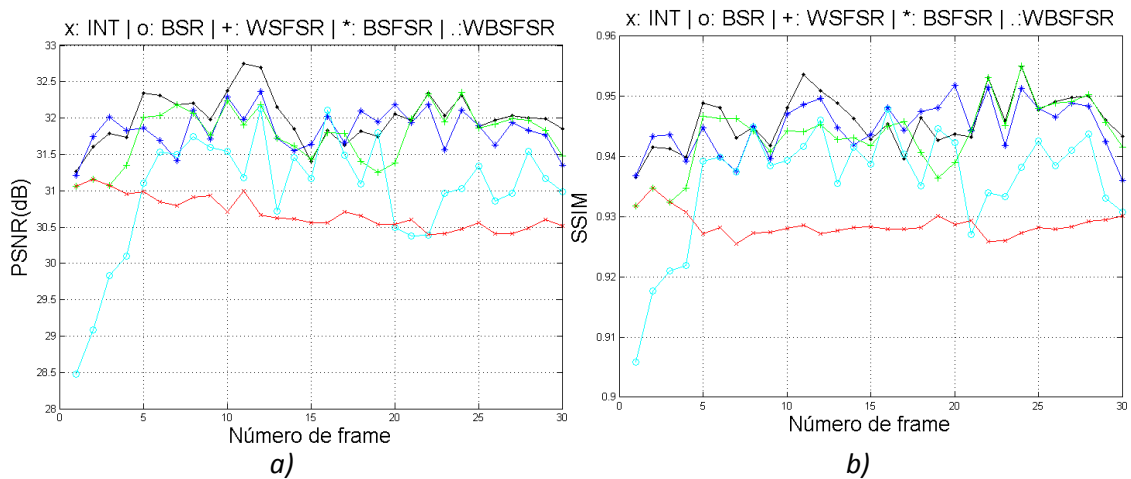


Figura 4.23.- Secuencia *Foreman* para INT, BSR y umbrales óptimos de WSFSR, BSFSR y WBSFSR
a) PSNR, b) SSIM

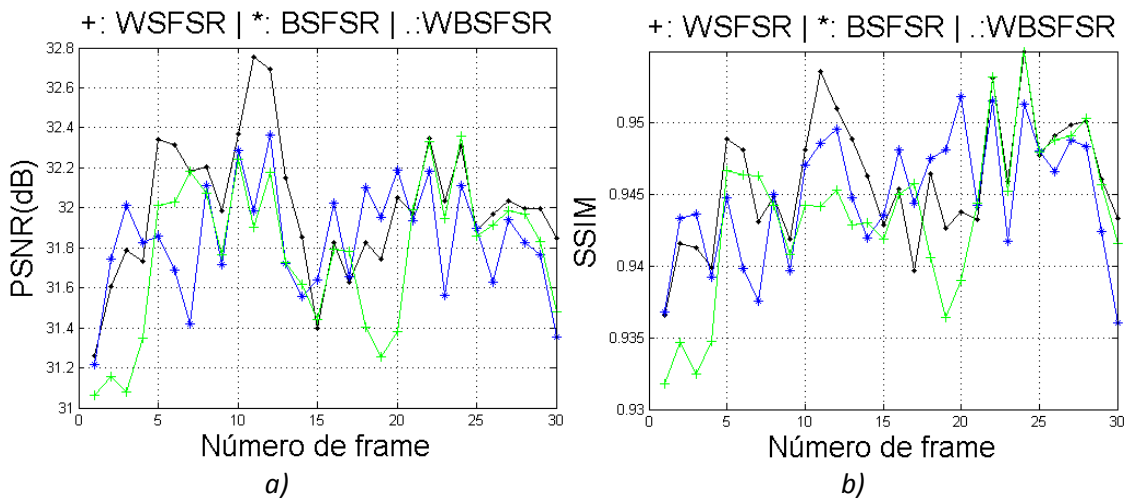


Figura 4.24.- Secuencia *Foreman* para umbrales óptimos de WSFSR, BSFSR y WBSFSR
a) PSNR, b) SSIM

4.3.2.5 Medida subjetiva de calidad

La medida subjetiva de calidad se basa en la comparación directa de los *frames* de salida. Para ello se comparan BSFSR, WSFSR y WBSFSR. En la Figura 4.25 se presenta el *frame* 20 de la secuencia *Foreman*. Se observa que en este caso una selección adecuada de *frames* y MBs para el cálculo del *frame* súper-resuelto a la salida (WBSFSR selecciona para este *frame* un total 580 MBs en lugar de los 580 MBs usados por BSR), evita la aparición de artefactos a la vez que reduce el tiempo de cómputo.

Se observa cómo la aplicación de un filtro a nivel más fino (Macro-Bloque) mejora aún más, tanto la calidad objetiva como subjetiva a nivel de artefactos.



Figura 4.25.- Análisis del *frame* 20 de la secuencia *Foreman* QCIF para BSFSR.
a) BSR, b) BSFSR, c) WSFSR, d) WBSFSR

4.3.3 Simulaciones con topología de Macro-Bloques adaptable

En este apartado, se presentan las simulaciones realizadas con topología de MBs adaptable, considerando un tamaño de MB variable (*Variable Block Size SR*, VBSSR). En la Tabla 4.8 se presenta un resumen de los resultados obtenidos. Los valores de SA y MBS utilizados son los mismos que para WSFSR y BSFSR. Se relacionará en este caso VBSSR con respecto a BSR. Esta tabla presenta las siguientes columnas:

- **ID:** Identificador de secuencia. Se presentan un total de 9 secuencias.
- **Secuencia:** Nombre de la secuencia.
- **VBSTHR1:** Umbral óptimo relacionado con SSIM para la secuencia en cuestión. Se fija normalmente a un valor similar al umbral WSFTHR.
- **VBSTHR2:** Umbral óptimo relacionado con MAD para la secuencia en cuestión. Se fija normalmente a un valor similar al umbral BSFTHR.
- **Diferencia de PSNR:** Diferencias de PSNR entre BSR e interpolación bilineal (INT) y entre BSR y VBSSR, considerando los valores de umbral VBSTHR1 y VBSTHR2 (*Variable Block Size THResholds*) óptimos.
- **Diferencia de SSIM:** Diferencias de PSNR entre BSR e INT y entre BSR y VBSSR, considerando los valores de umbral VBSTHR1 y VBSTHR2 óptimos.

Tabla 4.8.- Resumen de resultados con VBSSR

ID	Secuencia	VBSTHR1	VBSTHR2	VBSSR-BSR (PSNR en dB)	VBSSR-BSR (SSIM)
1	<i>Deadline</i>	90%	1%	0.07	0.4
2	<i>Flower</i>	40%	5%	1.57	2.07%
3	<i>Football</i>	80%	2%	0.02	0.1%
4	<i>Foreman</i>	80%	4%	1.25	2.8%
5	<i>Galdar</i>	10%	6%	0	0%
6	<i>Mobcal</i>	50%	3%	0.01	0.38%
7	<i>Mobile</i>	20%	6%	0.05	0.1
8	<i>Stockholm</i>	30%	6%	1.54	1.45%
9	<i>Suzie</i>	80%	3%	0.59	3.6%

En la Tabla 4.8 se observa que los umbrales óptimos de VBSSR oscilan entre valores comprendidos entre un 20% hasta un 90% para VBSTHR1 y entre un 0% hasta un 20% para VBSTHR2. Dichos valores guardan una estrecha relación con la cantidad de movimiento local y/o global de la secuencia en cuestión, y con los umbrales de los Filtros Selectivos de Ventana y MB (el primero de ellos se identifica con WSFTHR y el segundo de ellos con BSFTHR). La Tabla 4.9 presenta un resumen de resultados y la cantidad movimiento global y local medida subjetivamente. En esta tabla se puede apreciar claramente la relación entre VBSTHR1 y WSFSR, y entre VBSTHR2 y BSFSR

Tabla 4.9.- Relación entre movimiento global, local y valores óptimos de umbrales de VBSSR y filtros

ID	Secuencia	Movimiento Global	Movimiento Local	VBSTHR1	WSFTHR	VBSTHR2	BSFTHR
1	<i>Deadline</i>	Nulo	Medio	90%	80%	1%	2%
2	<i>Flower</i>	Medio	Bajo	40%	30%	5%	6%
3	<i>Football</i>	Bajo	Muy Alto	80%	80%	2%	2%
4	<i>Foreman</i>	Medio	Alto	80%	80%	4%	4%
5	<i>Galdar</i>	Bajo	Nulo	10%	20%	6%	8%
6	<i>Mobcal</i>	Medio	Medio	50%	50%	3%	5%
7	<i>Mobile</i>	Medio	Bajo	20%	20%	6%	6%
8	<i>Stockholm</i>	Medio	Bajo	30%	40%	6%	8%
9	<i>Suzie</i>	Nulo	Alto	80%	70%	3%	1%

Las gráficas utilizadas para analizar VBSSR fueron:

- Diferencias de métricas promedio frente a valores de umbrales VBSTHR.
- Métricas de evaluación de la secuencia para cada *frame*.
- Porcentaje de píxeles clasificados en un tamaño de MB.
- Tiempo de ejecución frente a valores de umbrales VBSTHR.
- Medida subjetiva de la calidad.

A continuación, se presentará a modo de referencia, cada una de estas gráficas y la información que aportan para los 30 primeros *frames* de la secuencia *Foreman*.

4.3.3.1 Diferencia de métricas promedio frente a valores de umbrales VBSTHR

En estas gráficas, al igual que en casos anteriores, se compara la diferencia promedio de una determinada métrica (PSNR o SSIM) del número de *frames* analizados en el proceso de SR frente a los valores de los umbrales VBSTHR1 y VBSTHR2 utilizados, fijando previamente a un valor similar o igual a los umbrales del Filtro Selectivo de Ventana WSFTHR y del Filtro Selectivo de MB BSFTHR, de acuerdo con las ecuaciones (4.17) y (4.18).

$$DIF_PSNR_{VBSTHR} = \frac{\sum_{i=1}^n (PSNR_VBSSR_{VBSTHRi} - PSNR_BSR_i)}{n} \quad (4.17)$$

$$DIF_SSIM_{VBSTHR} = \frac{\sum_{i=1}^n (SSIM_VBSSR_{VBSTHRi} - SSIM_BSR_i)}{n}, \quad (4.18)$$

donde:

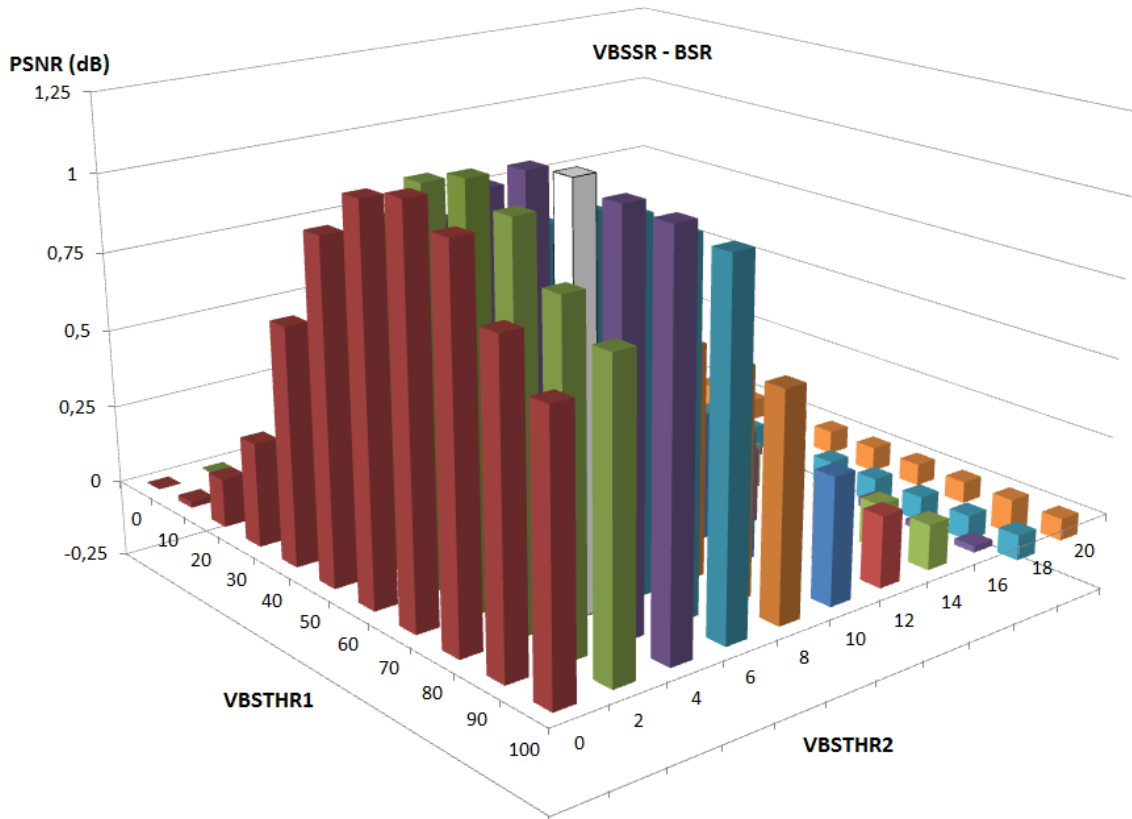
- PSNR_VBSSR_{VBSTHRi}: Es el valor de PSNR del *frame i* súper-resuelto, considerando los umbrales VBSTHR1 y VBSTHR2 frente al *frame i* original.
- SSIM_VBSSR_{VBSTHRi}: Es el valor de SSIM del *frame i* súper-resuelto, considerando los umbrales VBSTHR1 y VBSTHR2 frente al *frame i* original.
- PSNR_BSR_i: Es el valor de PSNR del *frame i* súper-resuelto con el algoritmo BSR, frente al *frame i* original.
- SSIM_BSR_i: Es el valor de SSIM del *frame i* súper-resuelto con el algoritmo BSR, frente al *frame i* original.
- *n*: Es el número de *frames* totales de la secuencia analizada. En el caso del presente ejemplo, *n* = 30.

Se presentan en la Figura 4.26 las gráficas de diferencia de métricas promedio para PSNR (izquierda) y SSIM (derecha) para la secuencia *Foreman*. En éstas se puede observar cómo se maximiza la calidad para valores de VBSTHR1 = 80% y VBSTHR2 = 4%, tanto para PSNR (1.25 dB de mejora) como para SSIM (2.8% de mejora), casos que quedan representados por las barras de color blanco y gris.

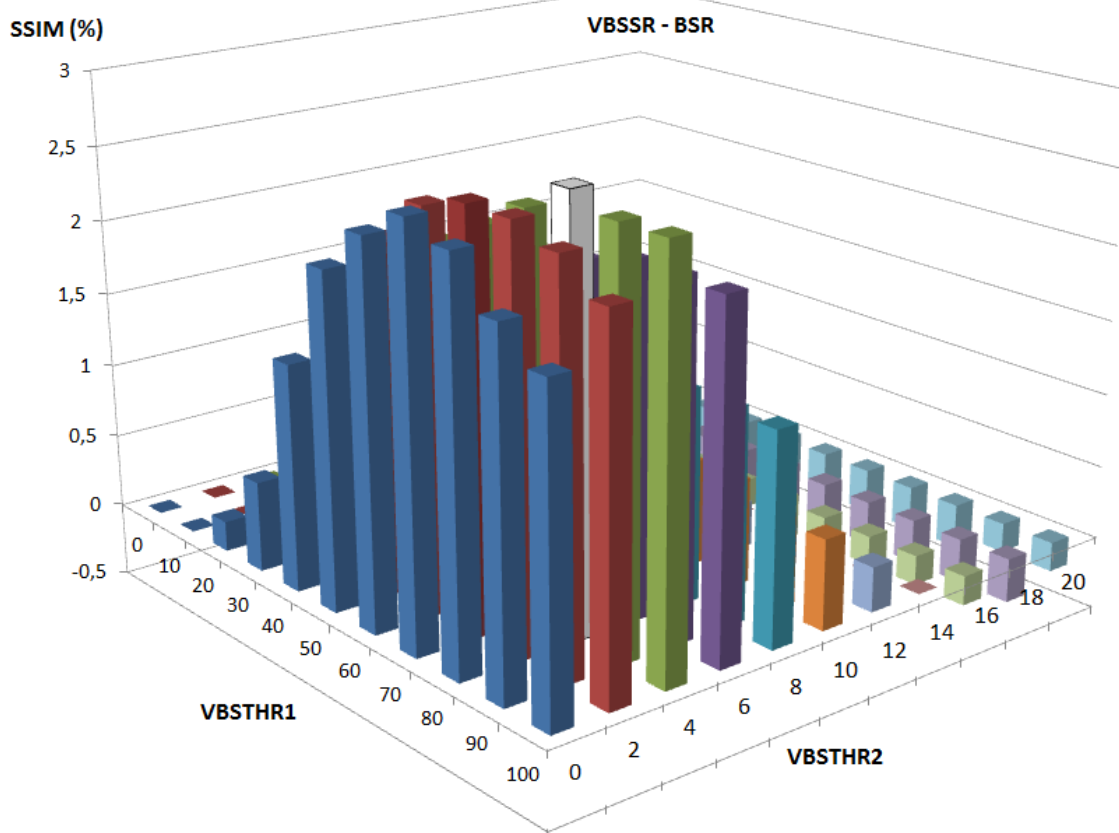
A partir de aquí, se deduce la interrelación existente entre los umbrales WSFTHR y VBSTHR1, y entre los umbrales BSFTHR y VBSTHR2, cuyos valores óptimos son idénticos para la secuencia *Foreman*. Por tanto, una definición adecuada de los umbrales WSFTHR y BSFTHR puede ayudar a la descripción posterior de los valores de los umbrales VSBTHR1 y VSBTHR2.

4.3.3.2 Métricas de evaluación de la secuencia para cada *frame*

En la Figura 4.27, se presenta la mejora que presenta VBSSR con respecto a BSR *frame a frame*, considerando los valores óptimos de VBSSR y tomando como ejemplo los 30 primeros *frames* de la secuencia *Foreman*. Se observa cómo la tendencia de la curva de VBSSR es a situarse por encima de la curva de BSR tanto en PSNR como en SSIM a costa de un tiempo de cómputo mayor (un 33% para este ejemplo concreto).



a)



b)

Figura 4.26.- Diferencia entre VBSSR y BSR respecto a VBSTHR1 y VBSTHR2 – Secuencia Foreman. a) PSNR, b) SSIM

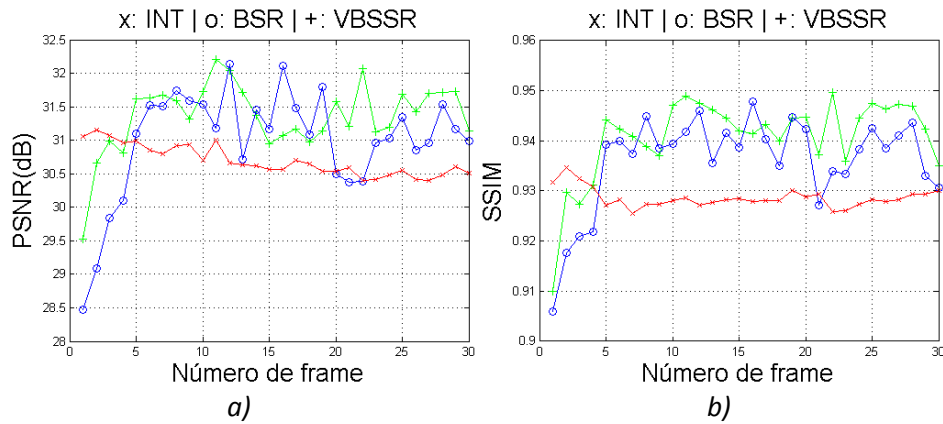


Figura 4.27.- Secuencia Foreman considerando VBSSR frente a BSR e INT. a) PSNR, b) SSIM

4.3.3.3 Porcentaje de píxeles clasificados en un tamaño de MB

Siguiendo una estrategia similar a las presentadas para WSFSR y BSFSR, se representa en este caso el porcentaje de píxeles que se clasifican en un determinado tamaño de MB cuando se realiza el proceso de división de MBs (topología variable). Así, dependiendo de las características de la secuencia, los píxeles quedarán clasificados en MBs de mayor o menor tamaño, habiéndose observado que en secuencias con movimiento global puro la clasificación se realiza mayoritariamente en MBs de mayor tamaño, mientras que cuando hay mucho movimiento local los píxeles se clasifican en MBs de menor tamaño, aglutinándose la mayoría de estos en MBs del mínimo tamaño permitido. Además, de forma general, cuando hay más cantidad de movimiento tiende a haber un mayor porcentaje de píxeles de menor tamaño.

Tal y como se presentó en el capítulo 3 de la presente Tesis Doctoral, se permiten 4 tamaños posibles de MB: 32, 16, 8 y 4. Así, las gráficas de porcentaje obtenidas incluyen 4 curvas, cada una de ellas para un tamaño de MB, y sumando siempre para cada *frame* de la secuencia un 100%. Se presenta en la Figura 4.28 un ejemplo de este tipo de gráfica para los 30 primeros *frames* de la secuencia *Foreman*. Se observa cómo va variando el porcentaje de los píxeles dependiendo de cada situación particular, y ya que en los primeros *frames* hay mayor movimiento, el porcentaje de píxeles incluidos en MBs de menor tamaño (4) es siempre mayoritario (ver 5 primeros *frames*). Por otra parte, puesto que en los últimos *frames* el movimiento es reducido, el porcentaje de píxeles incluidos en MBs de mayor tamaño (32) es siempre mayoritario (ver 5 últimos *frames*)

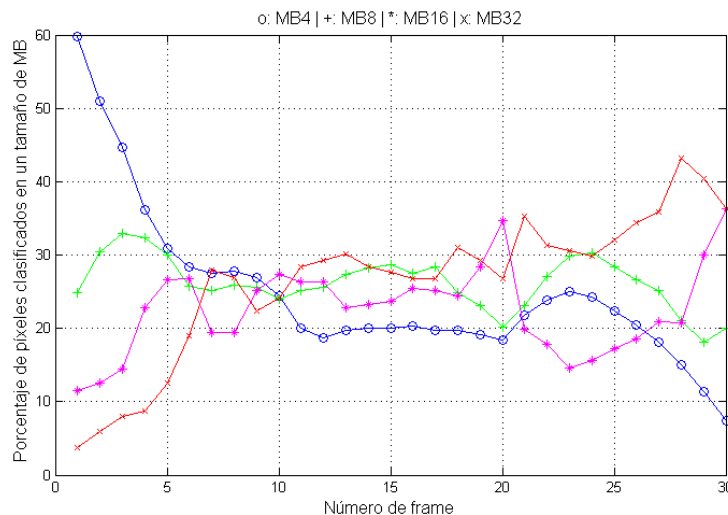


Figura 4.28.- Porcentaje de píxeles clasificados en un tamaño de MB – Secuencia Foreman

4.3.3.4 Medida subjetiva de calidad

La medida subjetiva de calidad se basa en la comparación directa de los *frames* de salida. Para ello se compara INT con BSR y VBSSR. Se presenta en la Figura 4.29 el *frame* 20 de la secuencia *Foreman*. Se observa que en este caso una selección adecuada del tamaño de MB para el cálculo del *frame* súper-resuelto mejora la reconstrucción del *frame* en cuestión. Esto se presenta claramente en la definición del edificio que queda detrás del obrero.

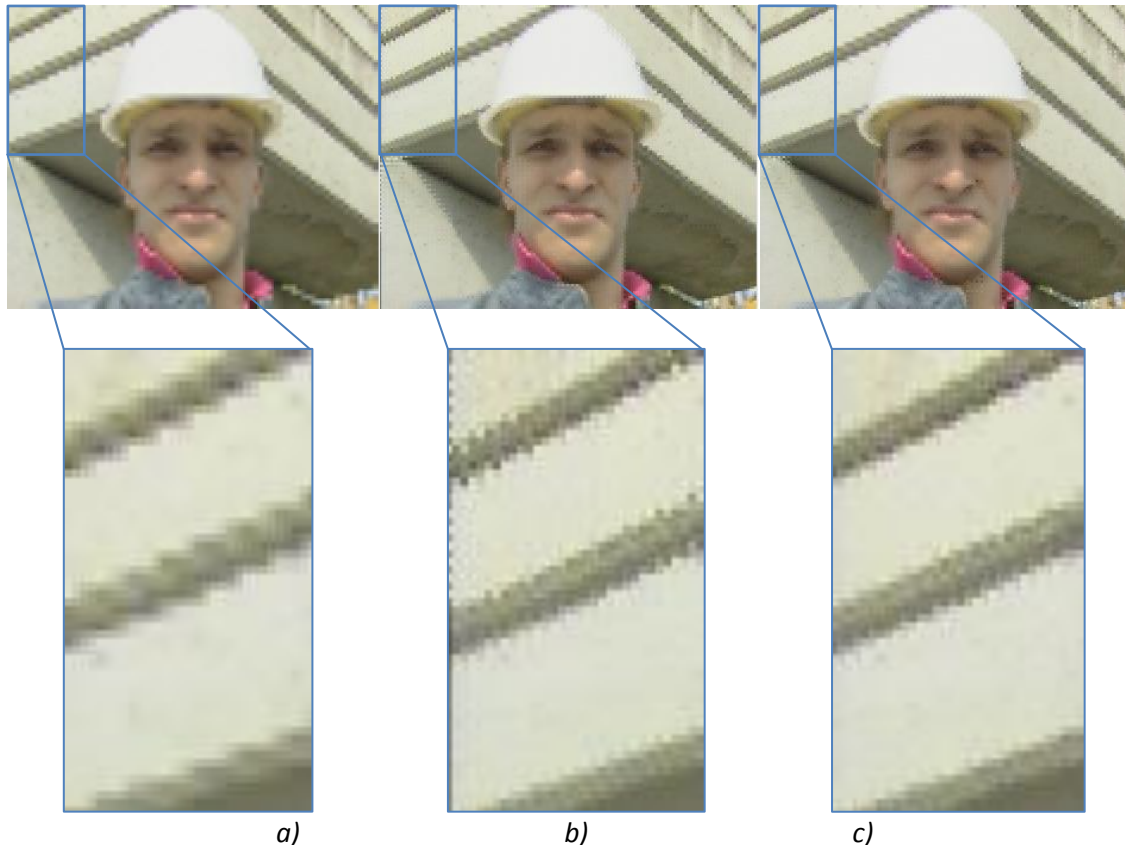


Figura 4.29.- Análisis del *frame* 20 de la secuencia *Foreman* QCIF para VBSSR.
a) INT, b) BSR, c) VBSSR

4.3.3.5 Variable Block Selective Filter Super-Resolution (VBSFSR)

Si bien los resultados de VBSSR mejoran claramente a BSR, estos son aún mejorables. Una solución adecuada para mejorar aún más la calidad es la incorporación a VBSSR de un Filtro Selectivo de MB, dando lugar al filtro VBSFSR (*Variable Block Size Filter SR*), representándose en la Figura 4.30 la diferencia de PSNR y de SSIM con respecto a BSR usando los filtros BSFSR y VBSFSR, en la que se puede observar claramente la mejora que supone VBSFSR cuando se considera un valor óptimo de BSFTHR = 4%, alcanzando una mejora en PSNR con respecto a BSR de 2.7 dB, y de un 6% en SSIM, tal y como se presenta en las Figuras 4.30.a y 4.30.b, respectivamente. Sin embargo, como contrapartida, el tiempo de cómputo pasa a ser de más del doble, como se puede ver en la Figura 4.30.c.

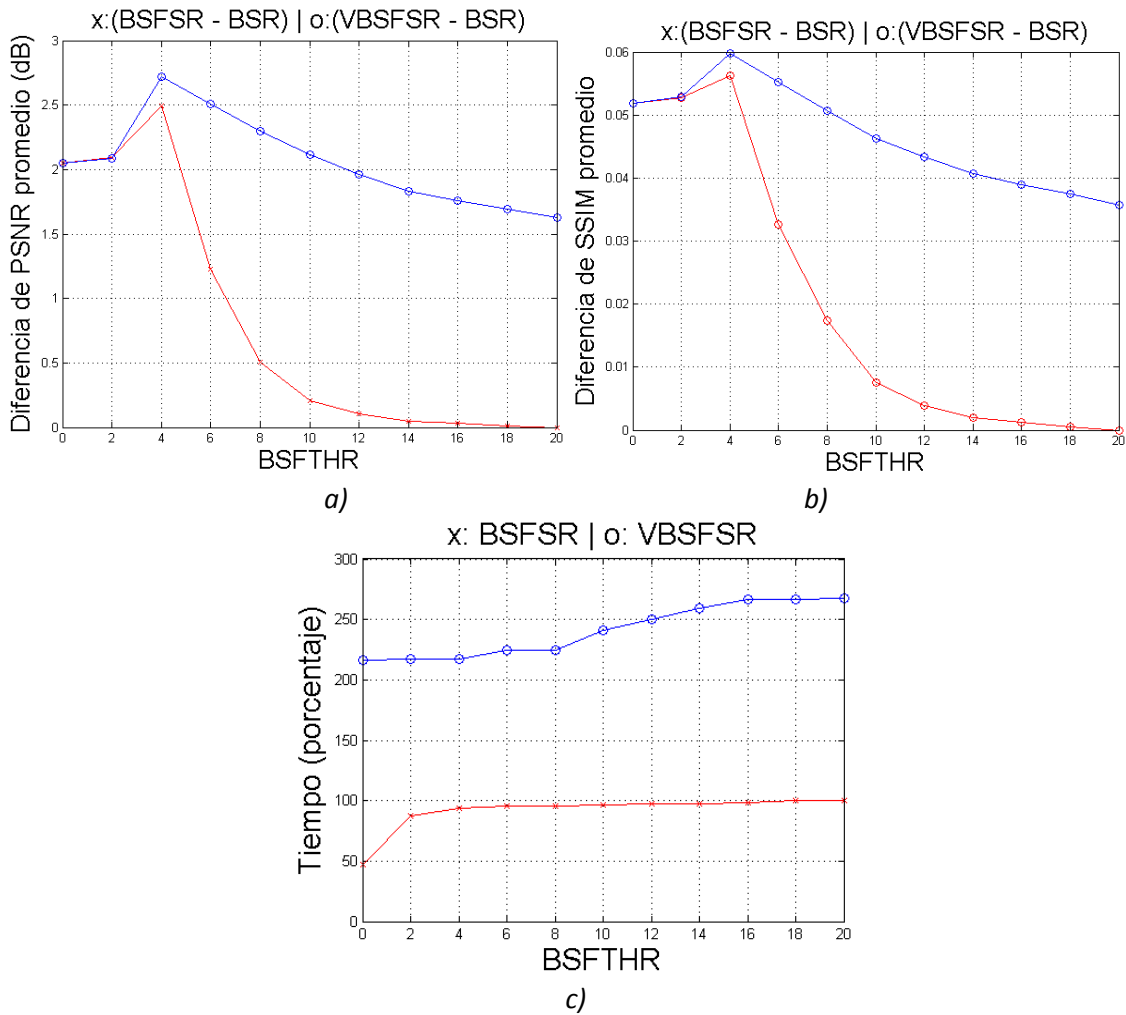


Figura 4.30.- Prestaciones del filtro VBSFSR con respecto a BSFSR. a) Diferencia de PSNR promedio frente a BSR, b) Diferencia de SSIM promedio frente a BSR, c) Tiempo de ejecución

4.3.3.6 Variable Block Size Window Selective Filter Super-Resolution (VBSWSFSR)

Ya que en la versión del filtro VBSFSR el tiempo de cómputo aparece como el principal obstáculo a considerar, se puede valorar como alternativa aplicar un Filtro Selectivo de Ventana a VBSSR, dando lugar al filtro VBSWSFSR (*Variable Block Size Window Selective Filter SR*). En la Figura 4.31 se presenta la diferencia de PSNR y de SSIM con respecto a BSR usando los filtros WSFSR y VBSWSFSR, donde se puede observar claramente la mejora que representa VBSFSR cuando se considera un valor óptimo de WSFTHR = 70%, alcanzando una mejora en PSNR con respecto a BSR de 2.75 dB, y de más de un 6% en SSIM, tal y como se presenta en la Figura 4.31.a y en la Figura 4.31.b, respectivamente. Además, el tiempo de cómputo para el filtro VBSWSFSR para el umbral óptimo es de menos de un 60% con respecto a BSR, tal y como se presenta en la Figura 4.31.c, mejorando mucho las prestaciones con respecto a BSFSR, que requería un tiempo de cómputo mayor que el doble del tiempo de cómputo de BSR.

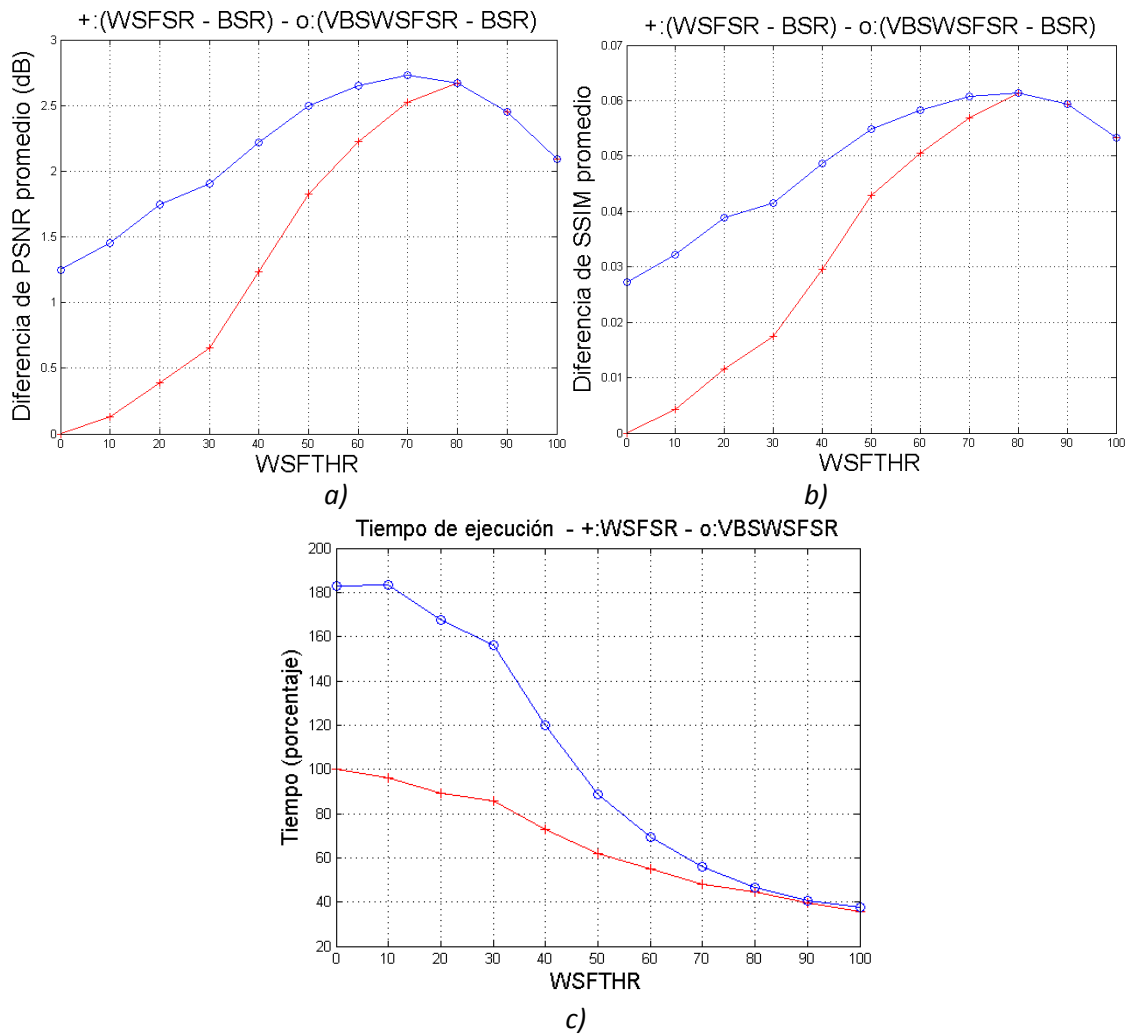


Figura 4.31.- Prestaciones del filtro VBSWSFSR con respecto a WSFSR. a) Diferencia de PSNR promedio frente a BSR, b) Diferencia de SSIM promedio frente a BSR, c) Tiempo de ejecución

4.3.3.7 Window & Variable Block Selective Filter Super-Resolution (WVBSFSR)

Considerando las mejoras introducidas, tanto en calidad objetiva como en tiempo de cómputo por parte de WBSFSR, se añade finalmente la capacidad de que el tamaño de MB sea variable, conformando el filtro WVBSFSR (*Window & Variable Block Selective Filter SR*), que se introdujo en el capítulo *Contribuciones Algorítmicas*. Partiendo de esta propuesta se obtienen gráficas comparativas de calidad y tiempo que presenten el funcionamiento de WSFSR, WBSFR y WVBSFSR con respecto a BSR para diferentes valores del umbral WSFTHR considerando umbrales óptimos para BSFTHR, VBSTHR1 y VBSTHR2. Los resultados se presentan en la Figura 4.32 y en la Figura 4.33, que representan la comparativa con respecto a calidad y tiempo de cómputo, respectivamente.

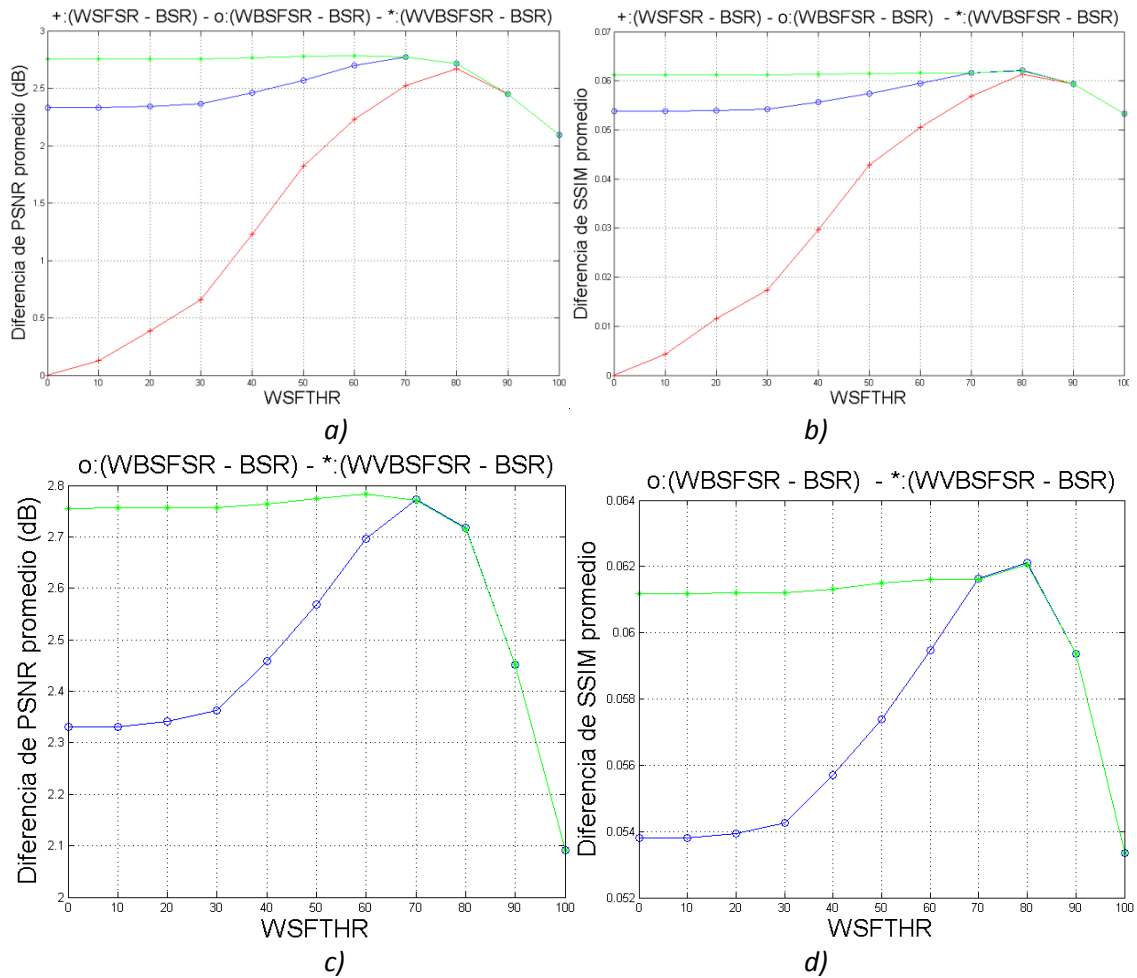


Figura 4.32.- Prestaciones de filtros respecto a WSFTHR – Secuencia Foreman. a) PSNR, b) SSIM, c) Zoom de PSNR, d) Zoom de SSIM

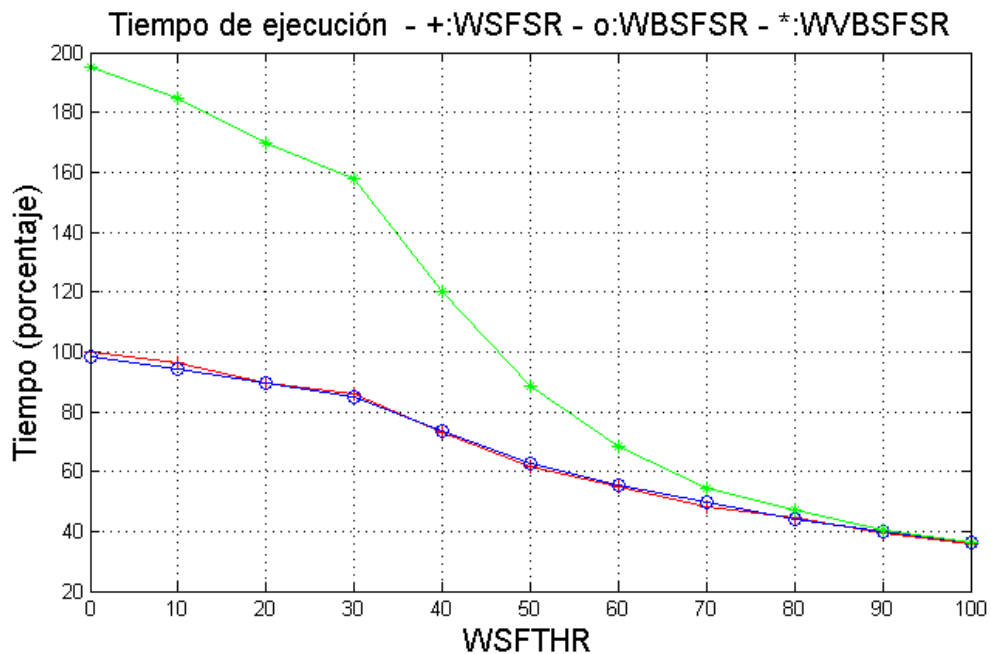


Figura 4.33.- Tiempos de ejecución respecto a WSFTHR – secuencia Foreman. Para WSFTHR = 60% el tiempo es un 70% sobre BSR

Tal y como se representa en la Figura 4.32, el aporte de los valores óptimos de umbral de VBSSR al Filtro Selectivo de Ventana y MB WBSFSR supone que el funcionamiento conjunto WVBSFSR se encuentre por encima de BSR y de la interpolación para cualquier valor de WSFTHR, obteniendo un máximo global superior a 2.8 dB en PSNR, y estando por encima de 6% en SSIM con independencia del valor del umbral WSFTHR. De aquí se deduce que si se optimizan los valores de los umbrales BSFTHR, VBSTHR1 y VBSTHR2, el umbral WSFTHR podría no ser necesario. Sin embargo esto implica como contrapartida los siguientes aspectos:

- Un mayor coste computacional; como se presenta en la Figura 4.33, el tiempo de ejecución es prácticamente el doble con respecto a BSR cuando WSFTHR = 0, siendo un 70% mayor que el correspondiente a BSR para un valor de WSFTHR que optimiza la calidad (60%).
- Una mayor complejidad en la selección de umbrales. Para WVBSFSR se han de seleccionar 4 valores de umbral con respecto a los 2 umbrales de WBSFSR.

Considerando estos aspectos, que se plantean como contraproducentes, y añadiendo el hecho de que la mejora de calidad que obtiene con WVBSFSR no es por lo general significativa con respecto a WBSFR, se concluye que WBSFSR representa la opción más recomendable. Esta afirmación se extenderá en el apartado 4.3.5, en la que se realiza una comparativa con el Estado del Arte, determinando que WBSFSR presenta un compromiso óptimo entre calidad objetiva y tiempo de cómputo.

4.3.4 Comparativa global de resultados

A modo de resumen se presentan los resultados de la secuencia *Foreman* para los 100 primeros *frames* de ésta. La Tabla 4.10 muestra los resultados promedio, donde destacan WBSFR y WVBSFSR, que se presentan *frame a frame* conjuntamente con INT y BSR en la Figura 4.34.

Tabla 4.10.- Resultados promedio para los 100 primeros frames de la secuencia *Foreman*

	PSNR(dB)	SSIM(%)	WSFTHR(%)	BSFTHR(%)	VBSTHR1(%)	VBSTHR2(%)
INT	30.66	92.73	NA	NA	NA	NA
BSR	28.57	87.40	0	0	0	0
VBSSR	29.82	90.12	0	0	80	4
BSFSR	31.10	93.16	0	4	0	0
VBSFSR	31.33	93.51	0	4	70	4
WSFSR	31.24	93.53	80	0	0	0
VBSWSFSR	31.30	93.47	70	0	80	4
WBSFSR	31.34	93.52	70	4	0	0
WVBSFSR	31.35	93.56	60	4	80	4

Tal y como se presenta en la Figura 4.34, las mejoras introducidas por WBSFSR y WVBSFSR destacan en la secuencia *Foreman* en la que zona en la que ésta alcanza un máximo de movimiento local (*frames* entre 60 y 80), que es donde normalmente la SR funciona peor debido a la aparición de artefactos. Como referencia, en la Figura 4.35 se presenta el *frame* nº 70 donde se aprecia claramente la eliminación de artefactos, incrementándose la calidad en casi 6 dB en PSNR (13% en SSIM) con respecto a BSR, y 0.5 dB en PSNR (1% en SSIM) con respecto a INT. Por otra parte, en la Figura 4.36 se muestra el *frame* 20, que presenta menor movimiento, y para el que se obtiene una mejora de alrededor de 2 dB (2% en SSIM) tanto con frente a BSR e INT, que presentan resultados similares.

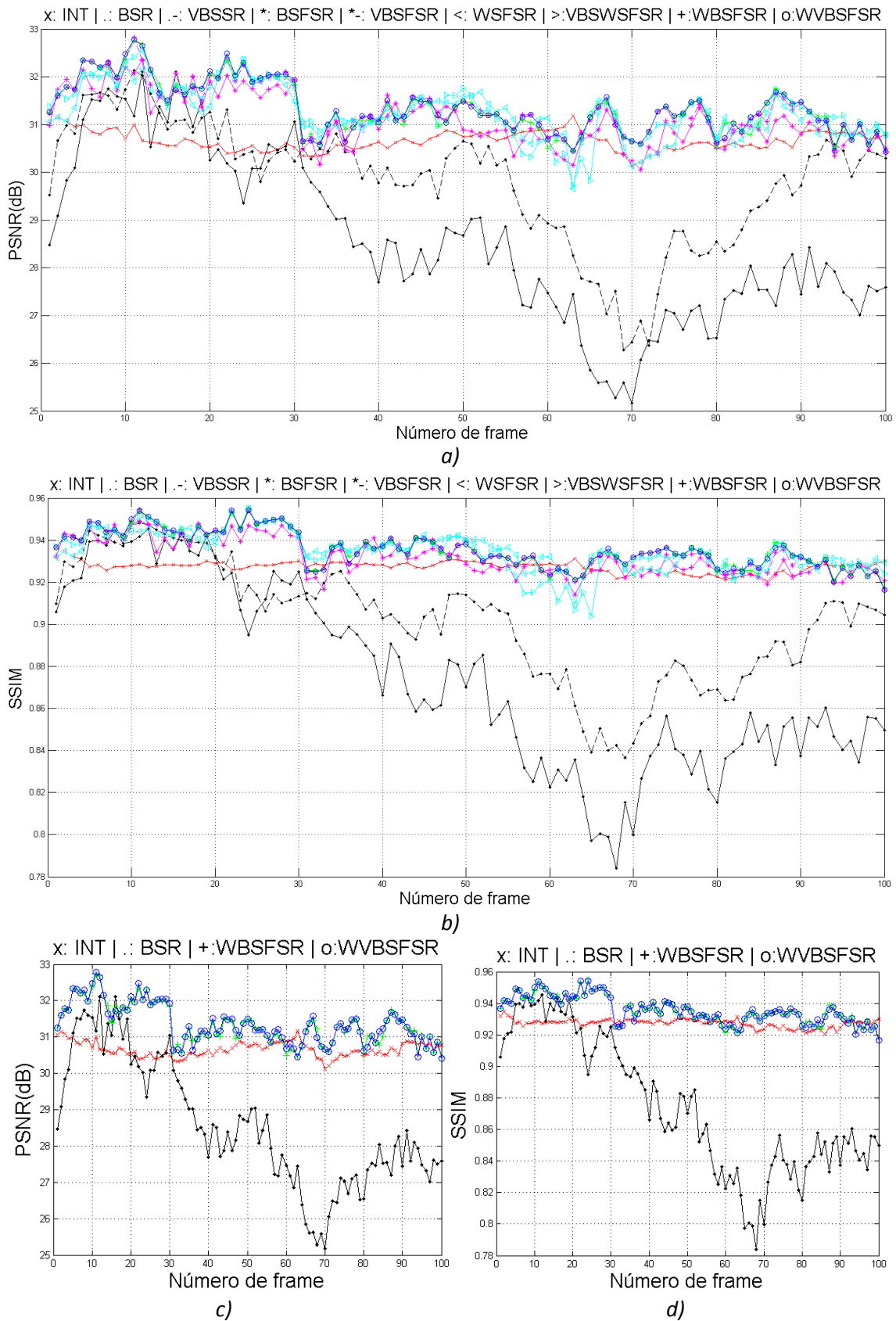


Figura 4.34.- Resumen de aportación de WSFSR, BSFSR y VBSSR. a) PSNR, b) SSIM, c) PSNR de INT, BSR, WBSFSR y WVBSFSR, d) SSIM de INT, BSR, WBSFSR y WVBSFSR



Figura 4.35.- Frame 70 de la secuencia Foreman QCIF.
a) INT, b) BSR, c) WVBSFSR



Figura 4.36.- Frame 20 de la secuencia Foreman QCIF.
a) INT, b) BSR, c) WVBSFSR

Por último, comentar que para disponer de una referencia del máximo de calidad que el algoritmo puede alcanzar en todo momento, se ha hecho uso de máximos teóricos partiendo de estimaciones a priori.

4.3.5 Comparativa con el Estado del Arte

Una vez caracterizadas las siete variantes que proporcionan el Filtro Selectivo de Ventana, el Filtro Selectivo de MB y la topología de tamaño de MBs adaptable, se realizó una comparativa con el Estado del Arte [R4]. En concreto se compararon las propuestas introducidas en la presente Tesis Doctoral con los siguientes trabajos:

- *Real-time image super resolution using an FPGA*. Se usa técnica *Iterative Back Projection* (IBP), propuesta por Bowen *et al* [BB08].
- *Robust Super-Resolution* (RSR), propuesto por Zomet *et al* [ZRP01].
- *Super-Resolution using variable size block-matching motion estimation with rotation*, propuesto por Anagün *et al* [AS12].

La comparativa se realizó haciendo uso de las versiones CIF (352x288 píxeles) de las secuencias *Flower*, *Football*, *Foreman*, *Mobcal* y *Mobile* y la versión QCIF (176x144 píxeles) de la secuencia *Suzie*. La consideración de estas secuencias concretas se debe a que para una de las referencias utilizadas, la de Anagün *et al* [AS12], tan sólo se disponían datos de PSNR de algunos *frames* representativos de cuatro las secuencias introducidas (*Foreman*, *Football*, *Mobile* y *Suzie*). Así, se presenta en primer lugar en la Figura 4.37 la comparativa de resultados global de PSNR considerando los *frames* propuestos en [AS12]. Las simulaciones consideran un Factor de Escala (*Scale Factor*) y una Ventana de Trabajo inicial de 15 (*frame* actual +/-7).

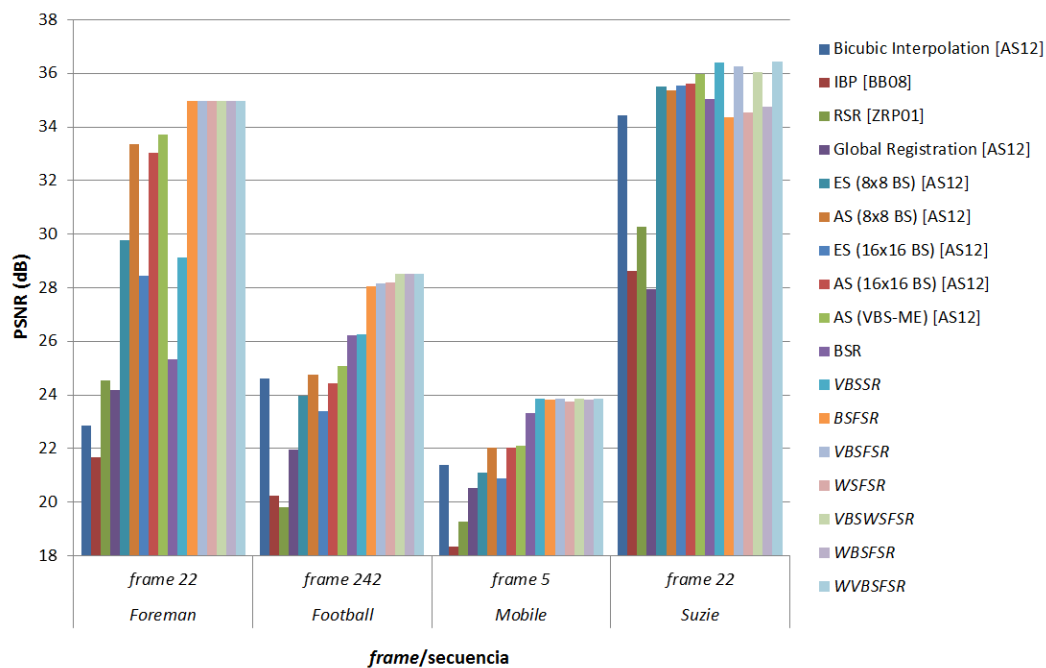
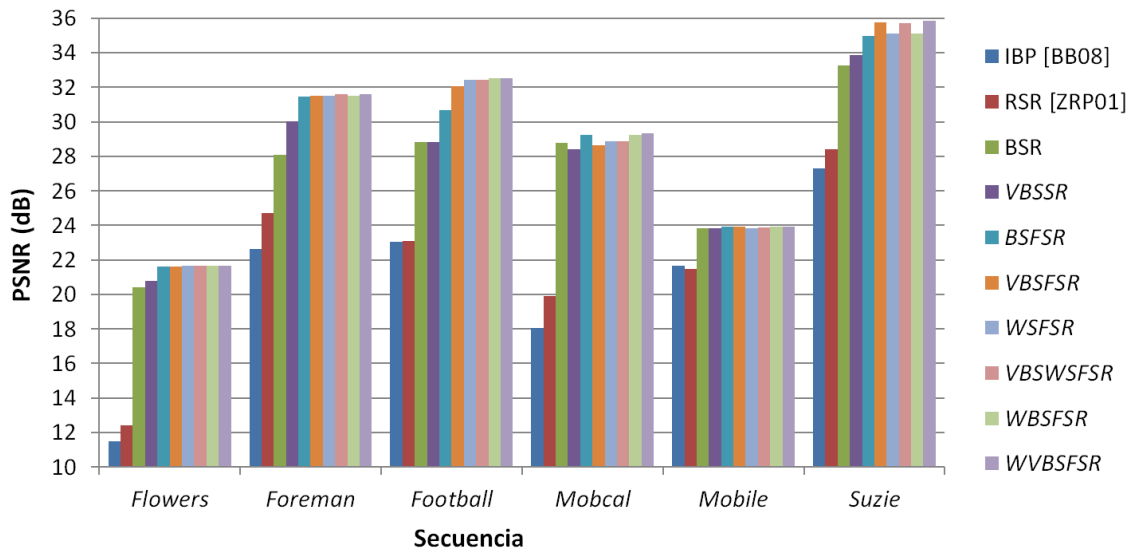
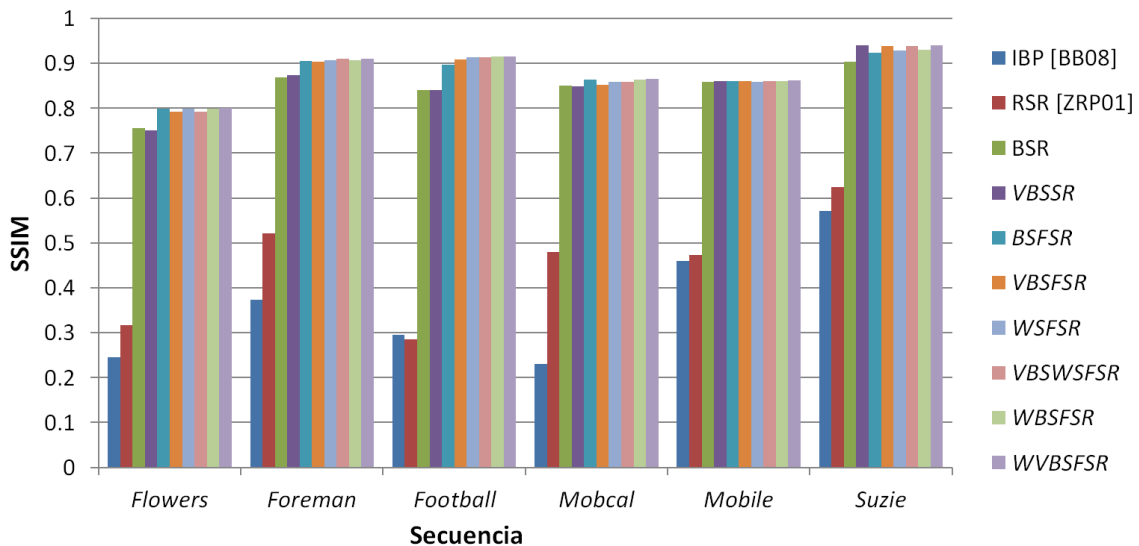


Figura 4.37.- Resultados de PSNR para frames concretos en secuencias *Foreman*, *Football*, *Mobile* y *Suzie*

Por otra parte, la Figura 4.38 muestra los resultados promedio de PSNR y SSIM para los métodos descritos en [BB08] y [ZRP01], BSR y las propuestas introducidas en esta Tesis Doctoral. Se observa cómo las métricas objetivas calculadas para las secuencias consideradas mejoran el Estado del Arte actual. La mejora en secuencias complejas, como *Football* o *Suzie*, es destacable, obteniendo 32.54 dB en PSNR (0.915 en SSIM) y 35.86 dB en PSNR (0.941 en SSIM), respectivamente.



a)



b)

Figura 4.38.- Resultados promedio para varios métodos (a) PSNR en dB, (b) SSIM

Con el objetivo de minimizar el tiempo de ejecución del algoritmo BSR, para la comparativa de tiempos de cómputo sólo se considerarán las propuestas basadas en WSF (*Window Selective Filter*). Como se ha presentado en la Figura 4.38 los métodos WSFSR, WBSFSR y WVBSFSR mejoran el Estado del Arte. Esto se puede comprobar también con una inspección subjetiva a la Figura 4.39, donde se muestra el *frame* 70 de la secuencia *Foreman* (*frame* en una parte de la secuencia con elevado movimiento local) para IBP [BB08] (a), RSR [ZRP01] (b), BSR (c) y WVBSFSR (d).

Finalmente, la Tabla 4.11 presenta el tiempo de cómputo medio (en segundos) para aplicar el proceso de SR a un *frame* con una Ventana de Trabajo de 15 (*frame* actual +/-7) considerando todas las secuencias estudiadas. Los *tests* se han ejecutado en un *core* 2 Quad (3 GHz, 4 GB RAM). Los resultados muestran que el tiempo de ejecución es mínimo para el filtro propuesto que usa un tamaño de MB fijo, y que las propuestas WSFSR, WBSFSR y WVBSFSR introducen una mejora notoria en términos de tiempo de cómputo frente al Estado del Arte.



Figura 4.39.- Resultados visuales para la secuencia Foreman CIF (frame 70)
a) IBP [BB08], b) RSR [ZRP01], c) BSR y d) WVBSFSR

Tabla 4.11.- Tiempo de ejecución promedio (en segundos) para aplicar SR a un frame con una Ventana de Trabajo de 15 frames (frame actual +/-7)

Método / Secuencia	Flower	Foreman	Football	Mobcal	Mobile	Suzie
IBP [BBP08]	12.19	25.29	19.04	35.83	29.26	6.75
RSR [ZRP01]	23.99	26.10	34.99	38.12	27.26	10.52
BSR	2.35	6.95	12.42	3.93	8.63	3.25
WSFSR	1.81	2.40	5.22	3.72	4.01	1.45
WBSFSR	1.66	2.38	5.19	3.42	3.94	1.41
WVBSFSR	2.32	3.02	5.41	8.08	7.78	1.52

De acuerdo con los resultados de la Tabla 4.11, y considerando que los resultados de calidad objetivos del método WBSFSR son bastante similares al método WVBSFSR se puede concluir que el método WBSFSR es la opción más apropiada, ya que presenta el mejor compromiso en términos de calidad y tiempo de ejecución. Esto también se puede ver si se relaciona la calidad objetiva (SSIM) y el tiempo de ejecución de acuerdo con la Figura de Mérito (FM) presentada en la expresión (4.19), que se introduce como un porcentaje con un valor máximo igual a 100 y un valor mínimo igual a 0.

$$FM(\%) = 100 * \frac{SSIM(Sequencia) + \{1 - [Tiempo(Sequencia)/Tiempo_{max}]\}}{2}, \quad (4.19)$$

donde $Tiempo_{max}$ es el tiempo máximo calculado para todos los métodos estudiados. Los resultados de FM se muestran en la Figura 4.40. Estos resultados se han calculado para las secuencias *Flower*, *Foreman*, *Football*, *Mobcal*, *Mobile* y *Suzie*, basándose en los resultados presentados en la Figura 4.38.b y en la Tabla 4.11.

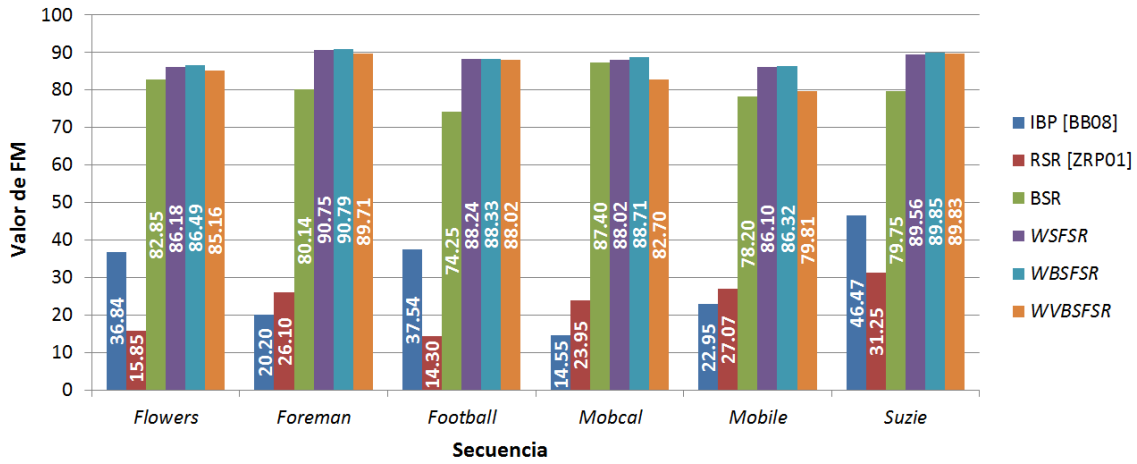


Figura 4.40.- Resultados de Figura de Mérito FM

4.4 CONCLUSIONES

En este capítulo, se ha presentado el entorno de simulación y los resultados más significativos de las contribuciones de filtros selectivos y topología de Macro-Bloques adaptable de la presente Tesis Doctoral.

Los algoritmos propuestos, además de mejorar la calidad objetiva en términos de PSNR y SSIM, aceleran significativamente el tiempo de cómputo con respecto al Estado del Arte, lo cual implica una aproximación clara hacia implementaciones en tiempo real, concepto que depende enormemente del tiempo de ejecución global. Usando plataformas dedicadas (GPUs, FPGAs, plataformas de multitud de núcleos, etc.) se podría alcanzar tiempo real para aplicaciones de vídeo de electrónica de consumo. Por otra parte, tras haber realizado un estudio de las 7 variantes que proporcionan el Filtro Selectivo de Ventana, el Filtro Selectivo de Macro-Bloque y el la topología de Macro-Bloques adaptable, se ha concluido que una aproximación que combine un filtro conjunto a nivel de Ventana de Trabajo y de Macro-Bloque fijo es muy adecuada para trabajar sobre Súper-Resolución.

Capítulo 5

RESULTADOS Y ANÁLISIS: ENTORNO MULTI-CÁMARA

Índice del capítulo

	<i>Página</i>
5.1.- Introducción	128
5.2.- Entorno de simulación y análisis	129
5.3.- Análisis de simulaciones	138
5.4.- Conclusiones	176

*Es la búsqueda lo que da auténtico sentido al hallazgo
José Saramago, escritor portugués (1922 – 2010)*

5.1 INTRODUCCIÓN

Las técnicas Súper-Resolución por fusión requieren que la coherencia espacial entre los *frames* analizados sea adecuada con el fin de poder estimar y compensar el movimiento existente entre estos. La segunda contribución de la presente Tesis Doctoral se centra en la integración de estas técnicas en un entorno Multi-Cámara. De esta forma se consigue incrementar la coherencia espacial entre *frames* y maximizar así tanto la calidad objetiva, como la calidad subjetiva.

Partiendo del sistema global propuesto en la Tesis Doctoral, reproducido en la Figura 5.1, en el que se consideran los métodos Multi-Cámara introducidos: Temporal-Espacial, Espacial-Temporal y Mixto, así como los modos de pre-procesamiento que sirven como etapa previa a estos: *Full-Frame*, *Overlap*, *Overlap+Borders*, se han desarrollado bancos de pruebas o *testbenches* que integran modelos de verificación completos para el conjunto de aportaciones introducidas.

A partir de los bancos de pruebas desarrollados se analizarán los resultados de calidad objetiva en términos de PSNR y SSIM, así como la calidad subjetiva para cada uno de los métodos Multi-Cámara propuestos, combinándolos con los diferentes modos de pre-procesamiento.

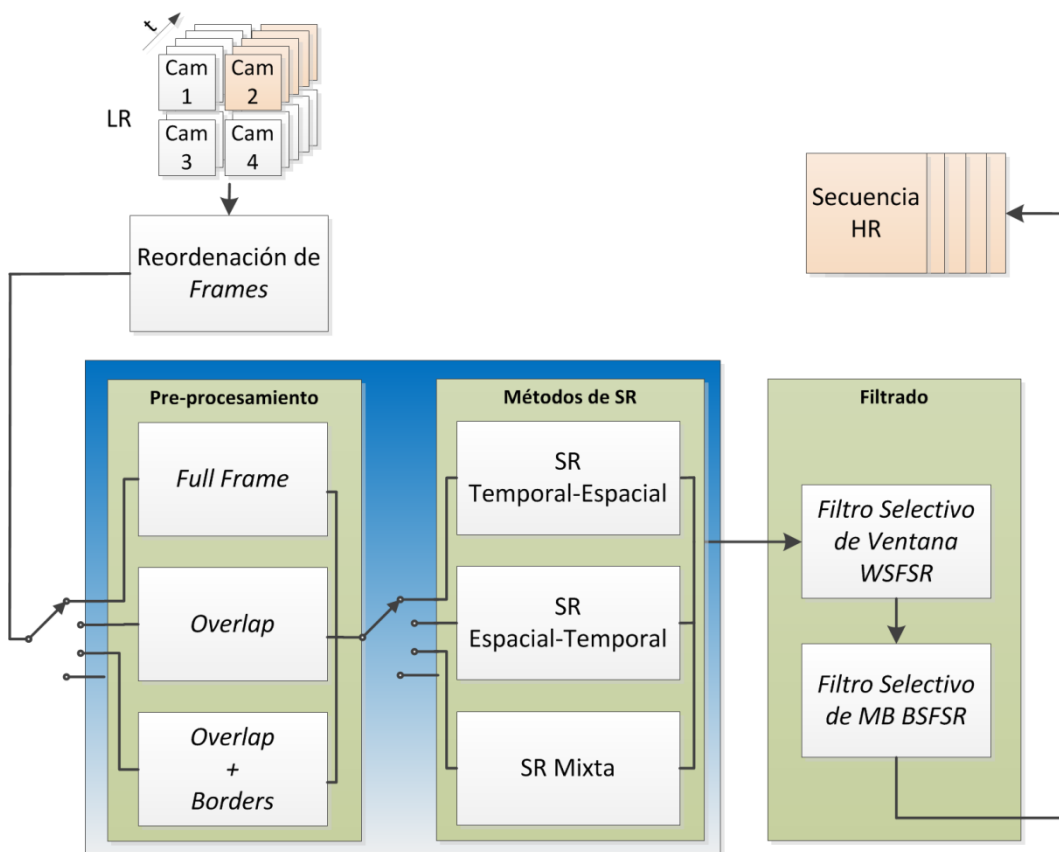


Figura 5.1.- Diagrama global de aportaciones algorítmicas resaltando entorno Multi-Cámara

5.2 ENTORNO DE SIMULACIÓN Y ANÁLISIS

El entorno de simulación desarrollado requiere, en primera instancia, disponer de secuencias de *test* suficientes para realizar las pruebas requeridas. El principal motivo por el cual surge la necesidad de crear secuencias MC artificiales es la escasez de secuencias de este tipo de libre disposición en la red. Pese a realizar una búsqueda intensiva e intentar contactar con distintas entidades, entre ellas la Universidad de Wisconsin [SZJ+09] y la Universidad de Nagoya [OFY+09] no se consiguieron más que un número reducido de secuencias, algunas de las cuales no se adaptaban a las necesidades de los métodos propuestos en esta Tesis Doctoral, al no cumplir con los requisitos tomados como punto de partida. A partir de ello surge la necesidad de obtener secuencias para realizar los *tests* de los métodos MC implementados simulando las características de un *array* MC. Por otro lado, la creación de secuencias artificiales presenta la ventaja de poder generar un banco de pruebas controlado, en el que por ejemplo se pueden analizar parámetros como la separación entre cámaras o las dimensiones del *array* MC.

Para crear estas secuencias MC de manera artificial, se parte de una secuencia de vídeo de alta resolución, como puede ser una secuencia HD (*High Definition*) o *Full HD*, aunque como se indicará a continuación, esto no tiene que ser estrictamente así. La idea es crear una serie de secuencias de menor resolución que estén solapadas entre sí, y que emulen a un *array* MC a partir de esta secuencia de alta resolución. Para ello, entre otros datos, se deben conocer las dimensiones de la secuencia de entrada y el Factor de Escala del algoritmo de SR seleccionado por el usuario.

A modo de referencia, se considerará que se desea utilizar una imagen de entrada de tamaño QCIF (*Quarter Common Intermediate Format*, 176x144 píxeles) con un Factor de Escala de 2, con la finalidad de obtener un tamaño de salida CIF (*Common Intermediate Format*, 352x288 píxeles). Para obtener dos secuencias de estos tamaños de una de las cámaras del *array* MC, a partir de una secuencia de alta resolución, habría que recortar de cada *frame* de la secuencia HD un nuevo *frame* de tamaño CIF y posteriormente diezmarlo, obteniendo una secuencia de entrada del tamaño deseado, como si de una nueva cámara se tratase.

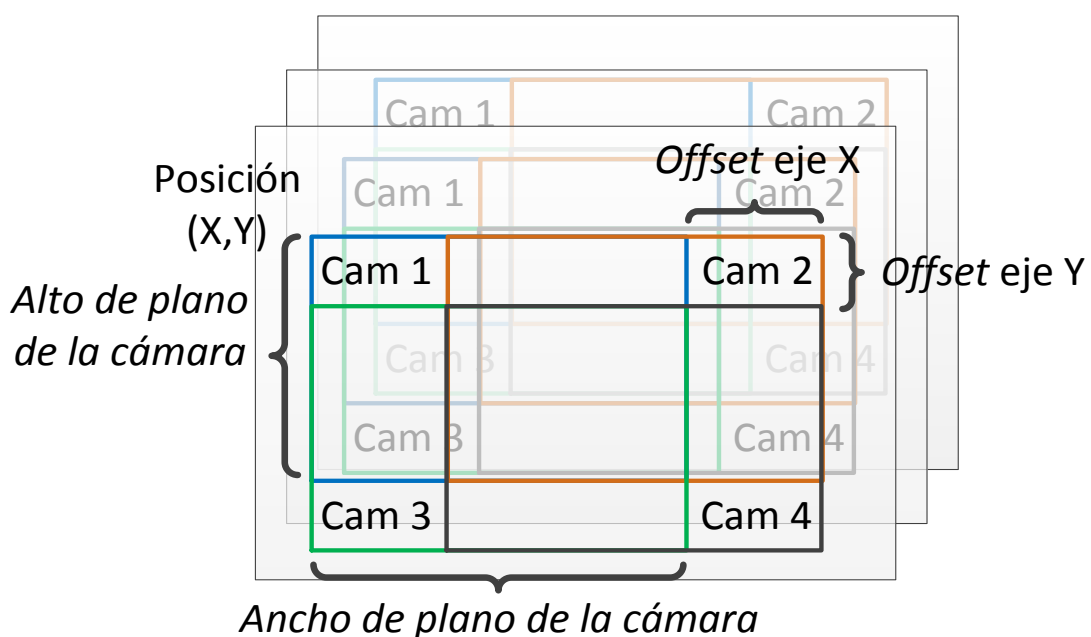


Figura 5.2.- Proceso de creación de las secuencias MC artificiales

Por tanto, si se conocen las dimensiones de la matriz que representa el *array* MC ($R \times S$), y a su vez se define el *offset* para los ejes X e Y , en píxeles, que van a tener las cámaras del *array* entre sí, basta con repetir el proceso anteriormente descrito desplazándose en las coordenadas X e Y el número de píxeles indicados por el valor del *offset* para obtener un *frame* de tamaño CIF de cada cámara, generado artificialmente emulando un *array* MC. A partir de estos *frames* en tamaño CIF se diezmará, reordenará o se realizará cualquier otro proceso necesario para implementar los diferentes bancos de prueba que han sido necesarios llevar a cabo a fin de evaluar la calidad de los procesos de SR propuestos haciendo uso de un sistema MC. De forma gráfica se puede ver representado este planteamiento en la Figura 5.2 para un *array* de dimensiones 2×2 .

Como parámetro adicional es necesario indicar la posición, definida por las coordenadas X e Y , dentro de la secuencia HD donde se desean obtener las secuencias artificiales del *array* MC. Esto permitirá además captar diferentes partes de la secuencia HD, aportando una mayor flexibilidad a las simulaciones. Por último, se ha de resaltar que las dimensiones de la secuencia no tienen por qué ser necesariamente HD o *Full* HD, bastaría con que se cumpliera que tanto el ancho como el alto de la secuencias tuviesen el tamaño de la secuencia MC a obtener, más el número de cámaras por el tamaño del *offset* en X e Y , según corresponda. Es importante tener en cuenta que el valor del *offset* también se verá afectado por el factor de diezmo, cuando se realice el proceso de diezmo. Por tanto, en función del valor del *offset*, estas secuencias aportarán información subpíxel o no, de la que hace uso el algoritmo BSR para reconstruir la secuencia. Así, un valor de *offset* 8 con un Factor de Escala 2 entre las cámaras en la secuencia HR, generará una secuencia LR con un *offset* 4, por lo que no tendrá información subpíxel, mientras que un *offset* de valor 19 generará una secuencia LR con un *offset* entre cámaras de 8.5, conteniendo esta secuencia información subpíxel. Es por tanto clave seleccionar valores de *offset* que, combinados con el Factor de Escala, generen valores subpíxel.

5.2.1 Modelos de testbench desarrollados

Una vez explicado de forma global el proceso seguido para crear las secuencias MC artificiales, se procede a exponer los distintos modelos de *testbench* que se han llevado a cabo para los métodos de SR propuestos en esta Tesis Doctoral, con el fin de hacer uso de las ventajas que aporta un sistema MC, como son los métodos de SR Temporal-Espacial, Espacial-Temporal y Mixto. Así mismo, a través del proceso de SR Espacial que se realiza en los métodos Temporal-Espacial y Espacial-Temporal, es posible generar una secuencia HR comparable con las generadas por los métodos MC, por lo que se ha desarrollado un modelo de *testbench* para el proceso de SR Espacial, de modo que se puedan comparar los resultados de este proceso con los de los métodos de SR MC. Así mismo, se ha establecido la posibilidad de crear secuencias artificiales para el proceso BSR, con el fin de poder establecer una comparativa de los resultados obtenidos a partir de los diferentes métodos anteriormente mencionados.

Por otra parte, aunque en el apartado anterior se ha explicado la manera en la que se obtienen, a partir de una secuencia HR, las secuencias artificiales MC, el proceso se ha descrito de forma muy general, mientras que a continuación, en los bancos de prueba desarrollados, se podrán apreciar las secuencias necesarias para cada método concreto, tanto para las secuencias de entrada como para las secuencias de referencia, y poder calcular así las métricas definidas.

5.2.1.1 *Testbench* del método Temporal-Espacial

A continuación, se expondrá el primero de los *testbenches* desarrollados, asociado al método Temporal-Espacial. En primer lugar se resaltarán una serie de aspectos a modo de recordatorio sobre este método. En este método se realizan n procesos de SR temporales, uno por cada cámara del *array*, y posteriormente se realiza un proceso de reordenación de los *frames* con el fin de agruparlos en una sola secuencia para, a continuación, llevar a cabo un proceso de SR Espacial. Puesto que este método se implementa en dos fases, es interesante tomar medida de la calidad, no solo en la secuencia final resultante del proceso, sino que además se considera de interés medir la mejora introducida a la salida de la primera fase. Por otro lado, se ha de recordar también que este método se basa en dos Factores de Escala, uno para el proceso de SR temporal y otro para el proceso de SR espacial.

El esquema mostrado en la Figura 5.3, presenta tres ramas principales distribuidas de forma horizontal. La estructura de este esquema, es la misma que la que se empleará en el resto de diagramas de los bancos de prueba, con alguna pequeña variación. En esta estructura se encontrará:

- En la parte superior, el proceso de SR.
- En la parte central, el proceso de creación de las secuencias de referencia.
- En la parte inferior, las distintas interpolaciones que se realizarán para comparar los resultados obtenidos.

Se comenzará por el proceso de creación de las secuencias de referencia, a través de las cuales se podrán calcular las métricas PSNR y SSIM, tanto en las secuencias resultantes de los distintos procesos de SR, como en las interpolaciones que se usarán a modo de referencia para contrastar la mejora obtenida en el proceso de SR. Como se puede observar en la rama central, (procesamiento de izquierda a derecha) una vez obtenidos los *frames* a partir de la secuencia HR, una de estas secuencias será la secuencia de referencia utilizada para obtener las métricas de la calidad de la secuencia resultante del método Temporal-Espacial, la cual ha sido denominada como HR (cámara seleccionada por el usuario para SR, en la Figura 5.3 es la cámara #2). En este caso, no se realiza ningún proceso de reordenamiento de *frames* en las secuencias, es decir, se almacenaría una secuencia de vídeo por cada cámara, tal y como se puede apreciar en la Figura 5.3. Posteriormente, a la secuencia interpolada correspondiente con la cámara seleccionada como referencia para realizar el proceso de SR, se le volverá a realizar otro proceso de interpolación. En este caso, el factor de interpolación será igual al escalado de la SR espacial correspondiente (*Scale Factor* se denomina aquí Esc_{Esp} para distinguirlo claramente del Factor de Escala temporal Esc_{Temp} .) con el segundo proceso de SR de este método, dando como resultado una secuencia HR equivalente a una interpolación por el escalado total del método de SR Temporal-Espacial. Por otro lado, al mismo tiempo que se realiza esta segunda interpolación, a la secuencia súper-resuelta MR_t de la cámara seleccionada a súper-resolver, es decir, la secuencia MR_t de la cámara #2 en la Figura 5.3, se le realiza un proceso de interpolación con el objetivo de evaluar si compensa realizar este segundo proceso de SR, o si por el contrario, es mejor realizar una interpolación.

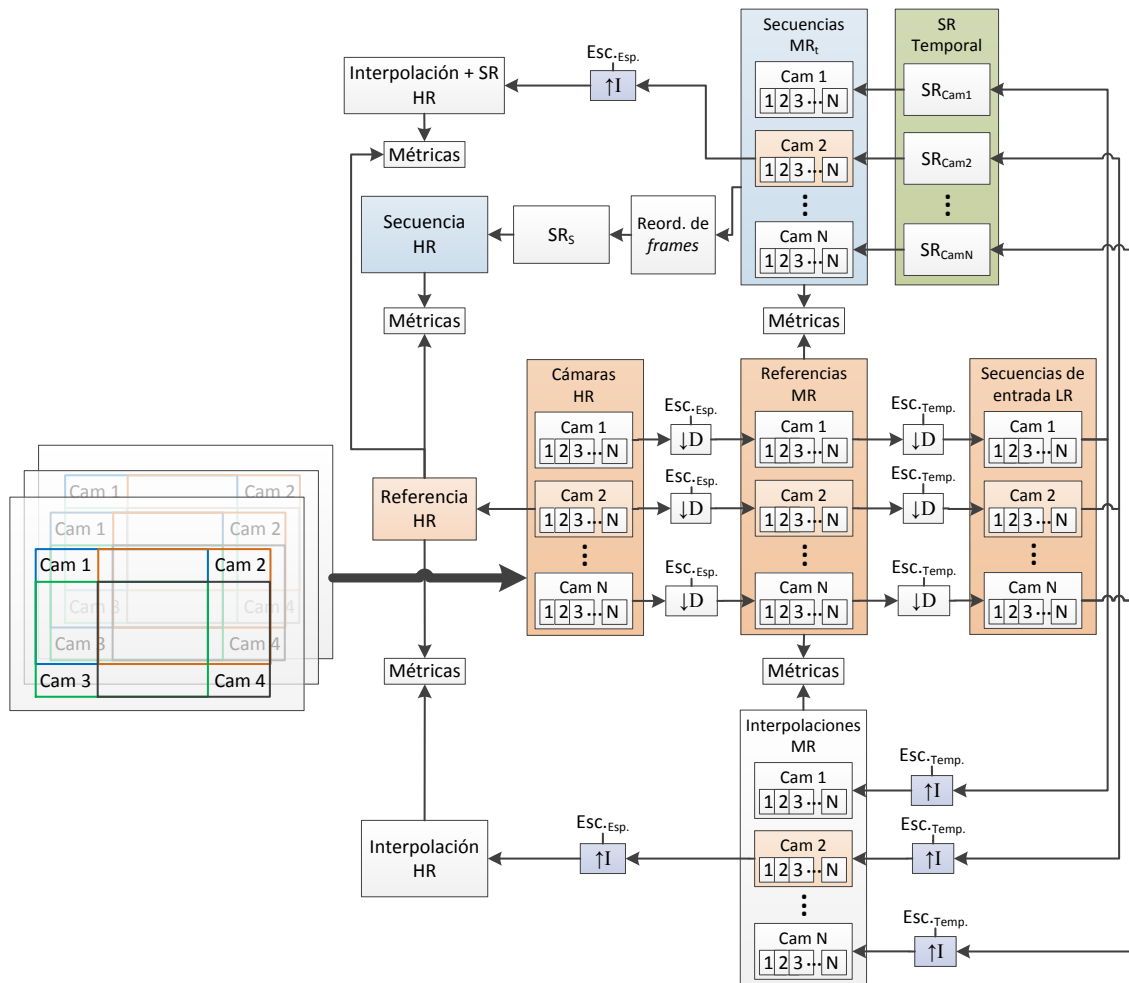


Figura 5.3.- Testbench del método Temporal-Espacial

Los motivos por los cuales se realiza el proceso de interpolación en dos fases son los siguientes:

- Poder obtener una referencia en las medidas de calidad de las secuencias generadas en cada una de las fases del método de SR Temporal-Espacial.
- Como se ha explicado previamente, se da una situación especial en el algoritmo BSR, el cual realiza un recorte de los bordes en el caso de que el tamaño de las secuencias de entrada no sean múltiplos enteros de MB.

Realizando las interpolaciones en dos fases, se asegura que el tamaño de la “Interpolación HR” coincide con la “Secuencia HR” resultante del método de SR Temporal-Espacial.

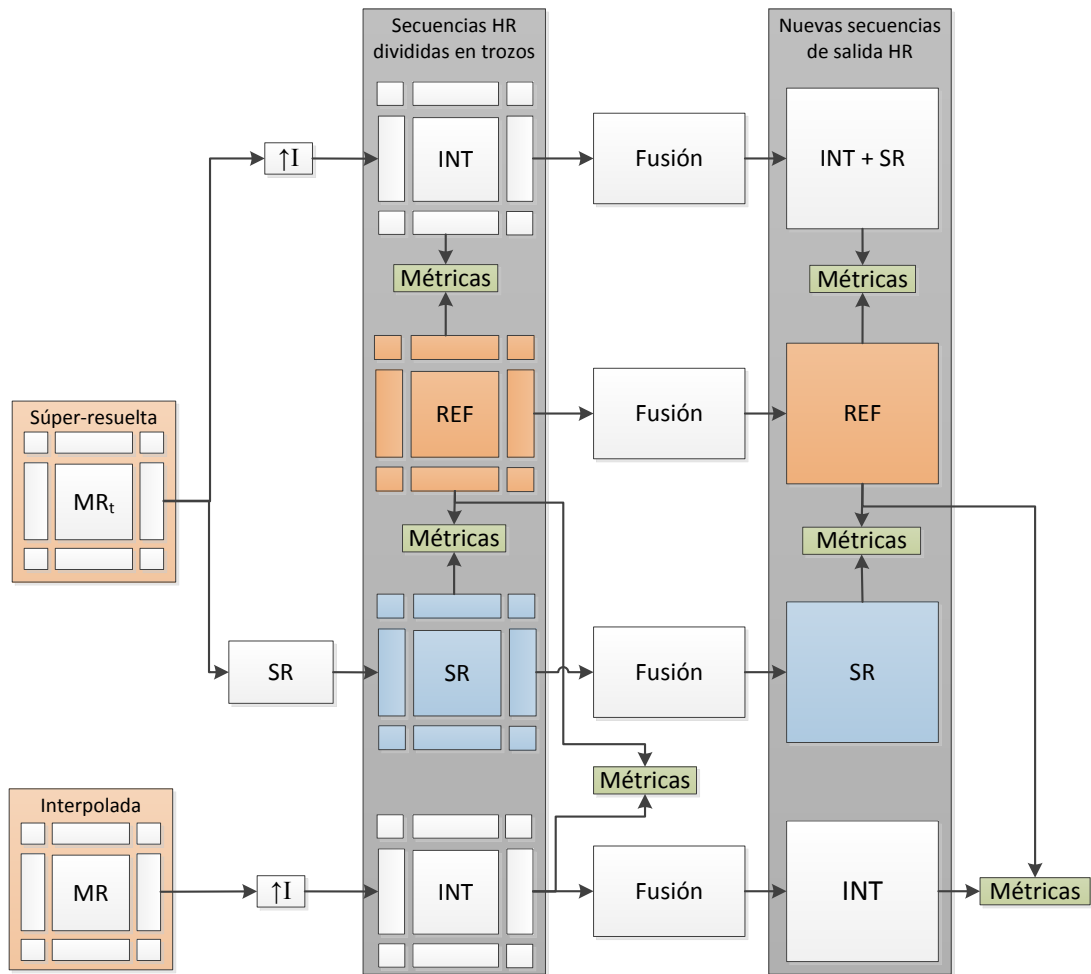


Figura 5.4.- Pre-procesamiento Overlap+Borders en testbench del método Temporal-Espacial

Finalmente, se procederá a comentar las modificaciones necesarias en este *testbench* para que soporte el pre-procesamiento *Overlap+Borders*. Antes de comenzar, se ha de recalcar que este modo de pre-procesamiento sólo es aplicable en la SR Espacial de este método, puesto que es donde únicamente hay una relación entre los *frames* en el proceso de SR que presenta un solape y unos bordes. El primero de los pasos a realizar, tal y como se presenta en la Figura 5.4, consiste en que una vez se han reordenado los *frames* súper-resueltos, se dividen todos los *frames* de la secuencia MR_t súper-resuelta en nueve secuencias, que son: la parte central solapada, cuatro laterales y cuatro esquinas. En aquellas partes donde no haya información, se introducirán ceros para rellenar y no perder la correspondencia entre los *frames*. Del mismo modo, se dividirá la secuencia de referencia HR usada para obtener las métricas, así como la secuencia MR interpolada. Posteriormente, todas y cada una de estas nueve secuencias MR_t se súper-resolverán y se interpolarán, midiendo todas y cada una de ellas con respecto a la de referencia (que previamente se ha dividido) con el fin de obtener los resultados de PSNR y SSIM, y por otro lado se interpolarán las secuencias MR interpoladas para obtener la interpolación HR, tal y como se muestra en la Figura 5.4. Por tanto, en este punto se tienen medidas del proceso realizado por separado, pudiendo tener un mayor control sobre lo que está sucediendo. Finalmente, se realiza un proceso de fusión de todos los trozos para volver a calcular las métricas de las secuencia fusionadas y valorar la calidad del resultado final.

5.2.1.2 Testbench del método Espacial-Temporal

A continuación se presentará el modelo de *testbench* usado con este método, el cual es en esencia similar al del método Temporal-Espacial. Así, su diagrama de bloques presenta las tres mismas ramas, estando en este caso la diferencia tanto en la secuencia de entrada, como en las secuencias de salida y las diferentes secuencias de referencia. Como se indicó en el Capítulo *Contribuciones Algorítmicas*, en este caso la salida generada por cada uno de los procesos de SR de este método es una única secuencia, la cual se corresponde con la secuencia súper-resuelta de la cámara de referencia seleccionada por el usuario, que en este caso está asociada a la cámara #2 del *array* MC. Es por este motivo que una vez adquiridos los *frames* de cada cámara, se toma como referencia HR la secuencia de la cámara #2, siendo ésta la única secuencia que se diezma por el Factor de Escala del segundo proceso de SR (SR Temporal, cuyo Factor de Escala es $Esc_{Temp.}$), obteniendo por tanto una secuencia MR que servirá de referencia para realizar el cómputo de las métricas. Por otro lado, los *frames* tomados de la secuencia HR son diezmos por el Factor de Escala total del método ($Esc_{Temp.} \times Esc_{Esp.}$), a la vez que van siendo reordenados en orden lexicográfico, de forma que se genere la secuencia LR de entrada al algoritmo de SR, al mismo tiempo que la secuencia MR de referencia es diezmada por el Factor de Escala de la SR Espacial ($Esc_{Esp.}$), para obtener los *frames* de la secuencia LR correspondientes de la cámara #2 del *array* MC, pudiéndose apreciar este proceso en la Figura 5.5.

El resto del proceso es el mismo que en el *testbench* del método Temporal-Espacial. En la rama inferior el objetivo es obtener dos interpolaciones, una de MR y otra de HR para comparar los resultados obtenidos al realizar el proceso de SR espacial y posteriormente el temporal, respectivamente. En paralelo al segundo proceso de SR, su secuencia de entrada, la cual ha sido previamente súper-resuelta, se interpolará por un Factor de Escala igual al escalado del proceso de SR temporal ($Esc_{Temp.}$), lo cual permitirá evaluar, al igual que anteriormente, si es apropiado o no realizar este segundo proceso de SR, o si por el contrario es mejor realizar una interpolación. Por último, resaltar que el primer proceso de interpolación se aplicará únicamente a los *frames* correspondientes a la segunda cámara del *array* MC, que en este caso es la seleccionada por el usuario para aplicar el proceso de SR.

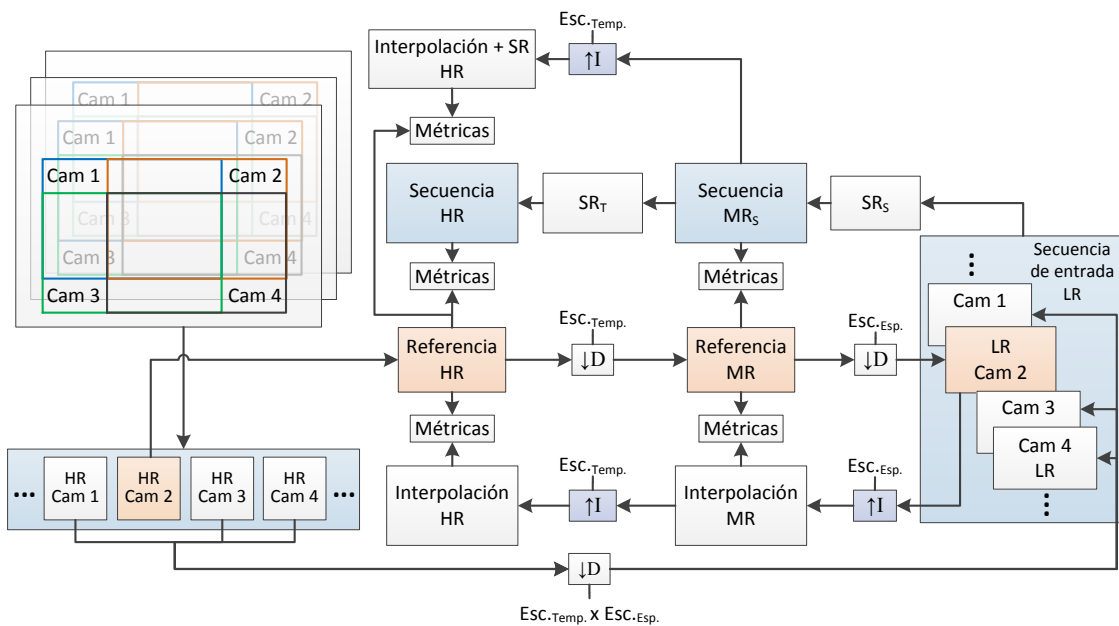


Figura 5.5.- Testbench del método Espacial-Temporal

Finalmente, se presentarán las modificaciones necesarias para contemplar en este *testbench* el modo de pre-procesamiento *Overlap+Borders*. Como se puede observar en la Figura 5.6, este caso es similar al explicado en el método Temporal-Espacial, exceptuando que, al realizar el pre-procesamiento en la primera fase del método, se simplifican las modificaciones en el *testbench*. Al igual que en el caso anterior, hay que dividir la secuencia de entrada en esta fase del método Espacial-Temporal, que en este caso se corresponde con la secuencia LR. Una vez dividida la secuencia LR, se procesan las secuencias interpolándolas y súper-resolviéndolas. De este modo, una vez obtenidas las nueve secuencias interpoladas, así como las nueve secuencias súper-resueltas, es posible realizar el cálculo de las métricas de cada borde y del solape, así como de la secuencia resultante del proceso de fusión.

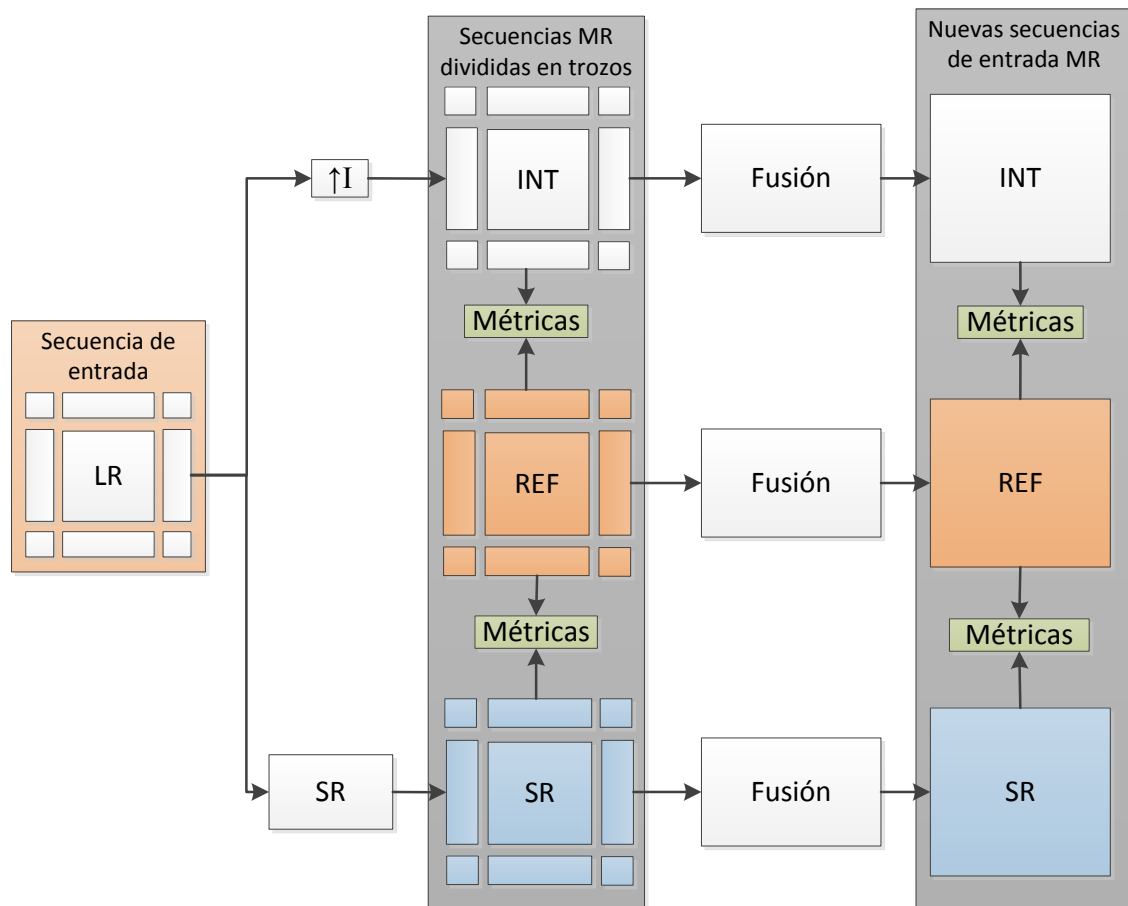


Figura 5.6.- Pre-procesamiento *Overlap+Borders* en *testbench* del método Espacial-Temporal

Además, tal y como se presenta en la Figura 5.6, no es necesario dividir la secuencia interpolada MR, puesto que se está en un paso previo a este proceso en el cual se obtiene la primera secuencia interpolada, es decir, la interpolación MR, simplificando el *testbench*. Así mismo, se ha de resaltar que estas nuevas secuencias fusionadas constituirán la entrada del segundo proceso de SR, donde se aprovechará la información temporal contenida en esta secuencia, para generar la secuencia de salida HR de este método.

5.2.1.3 Testbench del método Mixto y de la SR Espacial

Dado el interés de estudiar las posibilidades que puede aportar el realizar únicamente un proceso de SR Espacial, y considerando que las modificaciones realizadas sobre el algoritmo BSR permiten realizar un proceso de SR Espacial de forma aislada, se ha posibilitado la realización de simulaciones aplicando únicamente SR Espacial. Por otro lado, comentar que el *testbench* diseñado para el método Mixto es el mismo que el implementado en una SR Espacial, y se explicará a continuación.

En el método Mixto, se realiza un único proceso de SR para obtener la secuencia súper-resuelta HR. Así, el *testbench* se simplifica de forma significativa con respecto a los métodos anteriores, aunque el esquema mantiene la misma estructura. Por lo tanto, tal y como se muestra en la Figura 5.7, se realiza el proceso de reordenamiento de los *frames* extraídos de la secuencia HR para generar la secuencia LR de entrada al método Mixto o Espacial, al mismo tiempo que los *frames* correspondientes a la cámara seleccionada como referencia para súper-resolver conformarán la secuencia de referencia para medir las métricas a la salida, y posteriormente ser diezmada por el Factor de Escala del método Mixto (Esc_{Mixto}) o Espacial (Esc_{Esp}) con el fin de generar el *frame* LR y mezclarlo con el resto de *frames*, para generar la secuencia de entrada al método.

Posteriormente, los *frames* correspondientes a la cámara #2 en la secuencia LR serán, por un lado súper-resueltos, y por otro, interpolados por el Factor de Escala seleccionado (Esc_{Mixto} o Esc_{Esp}) en el método de SR empleado. Así, se obtendrá una versión de la secuencia LR interpolada y otra súper-resuelta, que en conjunto con la secuencia HR de referencia permitirá obtener las métricas, y por consiguiente poder realizar una valoración de la calidad resultante de estos métodos.

Finalmente, las modificaciones en este *testbench* para incluir el modo de pre-procesamiento *Overlap+Borders*, son las mismas que se han explicado en el apartado del *testbench* Espacial-Temporal, representado en la Figura 5.6, con la excepción de que en este caso las secuencias fusionadas MR, tanto de la interpolación como del proceso de SR, constituyen la nueva secuencia de salida de este método.

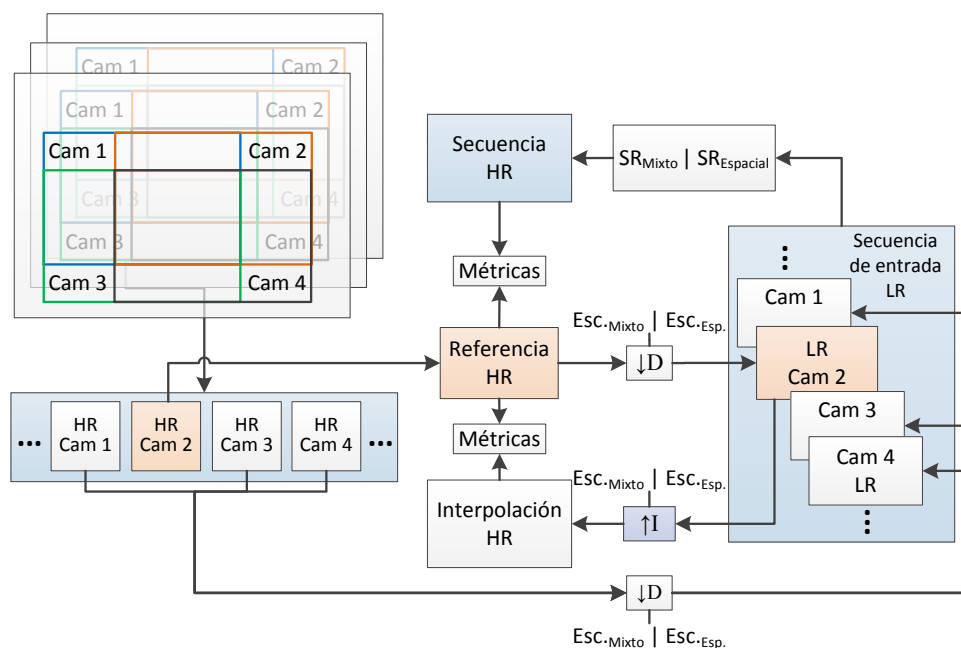


Figura 5.7.- Testbench del método Mixto y de la SR Espacial

5.2.1.4 Modificación del testbench de la SR Temporal

La SR Temporal es el tipo de SR que hasta ahora soportaba el algoritmo BSR, añadiéndosele la posibilidad de usar, bien secuencias reales como se ha venido haciendo hasta ahora, o bien secuencias artificiales creadas del mismo modo que para los métodos MC desarrollados en esta Tesis Doctoral, tal y como se muestra en la Figura 5.8. Añadiendo esta nueva opción se posibilita la comparación de los resultados obtenidos en los métodos MC con los del algoritmo BSR sobre una misma base, y por otro lado, se permite seguir realizando pruebas con nuevas modificaciones que vayan surgiendo para la mejora del algoritmo BSR.

En este caso, la secuencia de referencia se obtendría recortando la cámara a súper-resolver de la secuencia HR para tomarla como punto de partida para el cálculo de las métricas, y posteriormente aplicarle un factor de diezmo igual al Factor de Escala de la SR Temporal ($Esc_{Temp.}$), obteniendo por tanto la secuencia LR de entrada al algoritmo. Al igual que se ha comentado en el resto de métodos, sobre esta secuencia LR se aplica el proceso de SR y se interpola para posteriormente evaluar las métricas respecto de la secuencia de referencia, pudiendo obtener una comparativa que aportará una valoración de la calidad objetiva obtenida en el proceso de SR.

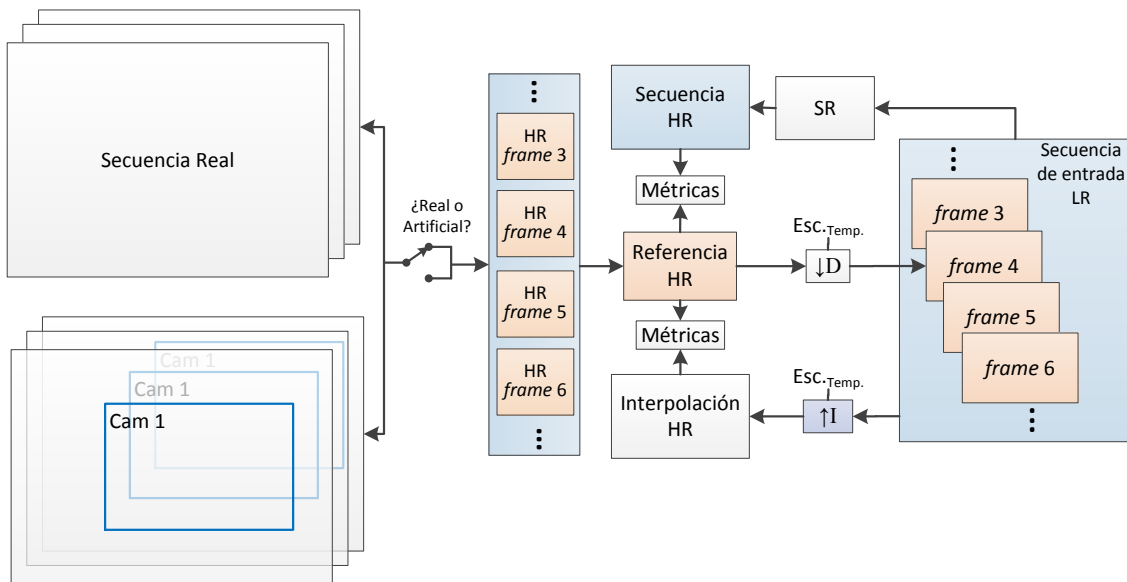


Figura 5.8.- Modificación del testbench de la SR Temporal

5.3 ANÁLISIS DE SIMULACIONES

Para evaluar la mejora aportada por cada uno de los métodos de SR propuestos, y los diferentes modos de pre-procesamiento implementados en esta Tesis Doctoral, se han usado un total de seis secuencias de referencia: *Mobcal*, *Stockholm*, *Shields*, *Parkrun*, *Water Cooler* y *Fujidog*, cuyas características se presentan en el anexo A.1 de la presente Tesis Doctoral. De estas secuencias, *Water Cooler* y *Fujidog* son secuencias MC reales, mientras que el resto de secuencias se han generado artificialmente.

El *array* MC que se ha usado para realizar las simulaciones de las secuencias artificiales se corresponde con una matriz de dimensiones 3x3. La disposición de las cámaras en el *array* MC tiene forma cuadrada, siendo la cámara elegida a súper-resolver la cámara central de este *array*, correspondiéndose con la cámara número 5 siguiendo un orden lexicográfico, tal y como se muestra en la Figura 5.9. De este modo se consigue que tanto el modo de pre-procesamiento *Full-Frame* como *Overlap+Borders* sean comparables entre sí, puesto que tienen el mismo número de cámaras, siendo esta cámara la seleccionada para aplicar el proceso de SR.

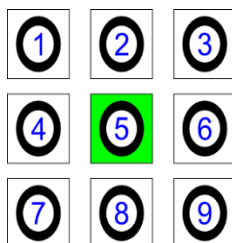


Figura 5.9.- Array MC de dimensiones 3x3 para generar las secuencias MC artificiales

En cuanto a las dimensiones utilizadas para generar las secuencias de entrada, y por consiguiente caracterizar los diferentes métodos, se ha seleccionado en las secuencias reales MC unas dimensiones de 112x80 píxeles para la secuencia *Water Cooler*, puesto que las dimensiones de la secuencia original son de 480x360 píxeles, mientras que para *Fujidog* se establecen las dimensiones de la secuencia de entrada en 160x96 píxeles, puesto que las dimensiones de la secuencia original son de 640x480 píxeles. Por otra parte, las dimensiones que tienen las secuencias artificiales de entrada a los diferentes métodos serán de 192x160 píxeles, con el objetivo de que sean divisibles por los tamaños de MB más usados, siendo estos 4, 8, 16 y 32. Al ser las dimensiones de las secuencias divisibles por los tamaños de MB mencionados, no será necesario adaptar el tamaño de éstas en cualquiera de los procesos de SR a realizar, y por consiguiente todas las secuencias de salida tendrán las mismas dimensiones. De este modo, se asegura que los resultados obtenidos para diferentes tamaños de MB puedan ser comparados entre sí, permitiendo por tanto presentar comparativas entre los resultados obtenidos a partir de los diferentes métodos. Los *frames* seleccionados de cada secuencia, para generar los resultados presentados en este capítulo se corresponden con los indicados en la Tabla 5.1.

Tabla 5.1.- Frames seleccionados de cada secuencia para realizar las simulaciones

Secuencia	<i>Water Cooler</i>	<i>Fujidog</i>	<i>Mobcal</i>	<i>Stockholm</i>	<i>Shields</i>	<i>Parkrun</i>
Frames	1-60	1	190-249	5-64	3-62	5-64

Por otra parte, la configuración asignada al *array* MC para generar las secuencias artificiales, se corresponde con la mostrada en la Tabla 5.2, en la que se representan las coordenadas X e Y correspondiente a la posición de cada cámara, así como el *offset* vertical y horizontal entre cada una de las cámaras, que es igual para todas las simulaciones (se han probado varios valores y los resultados no varían notoriamente, siempre que se considere que exista desplazamiento subpíxel y que el desplazamiento quede dentro del área de búsqueda seleccionado). Todos los valores de esta tabla están expresados en píxeles. Para mostrar los resultados obtenidos, a partir de las simulaciones realizadas, se presentarán los valores medios de un total de 60 *frames* a través de diferentes tablas.

Tabla 5.2.- Configuraciones de los parámetros del *array* MC en cada secuencia

Secuencia	Coordenadas (píxeles)		Offset (píxeles)	
	X	Y	Vertical	Horizontal
<i>Mobcal</i>	200	18	21	21
<i>Stockholm</i>	400	15		
<i>Shields</i>	200	15		
<i>Parkrun</i>	200	15		

5.3.1 Método Temporal-Espacial

En este apartado, se presentan los resultados obtenidos para cada una de las secuencias consideradas, con el fin de caracterizar el método de SR MC Temporal-Espacial. La configuración empleada para cada secuencia se muestra en la Tabla 5.3. Cada uno de los procesos de SR Temporal que se lleva a cabo en este método, sobre cada cámara, tienen la misma configuración de parámetros, en los cuales se ha establecido una Ventana de Trabajo de SR de ± 7 *frames*, ya que tras varias simulaciones se ha determinado que es un valor adecuado para maximizar la calidad objetiva en promedio para la mayoría de las secuencias.

Por otra parte, los valores del parámetro SA (*Search Area*) elegido en la fase de SR Temporal han sido determinados experimentalmente para cada secuencia, calculándose su valor en un punto de compromiso, entre la mejora resultante y el tiempo de procesamiento de cada *frame*. Sin embargo, los valores del parámetro SA de la SR Espacial están determinados directamente por la separación que existe en píxeles. Al ser este valor conocido en las secuencias artificiales, basta con saber que la separación entre cámaras se corresponde con el *offset* entre cámaras HR, el cual dividido por el Factor de Escala que se establece entre la secuencia HR y la secuencia MR o LR, proporciona el valor mínimo de SA. En este caso, el valor del desplazamiento entre cámaras HR es de 21 píxeles, y el escalado entre la secuencia HR y MR_t se corresponde con un Factor de Escala 2, implicando que existe una separación entre los *frames* de las cámaras en la fase de SR Espacial de $21/2 = 10.5$ píxeles, por lo que cualquier valor de SA superior a 11 píxeles sería válido para súper-resolver la secuencia. Por otra parte, el tamaño de MB en la Tabla 5.3 es el que maximiza la calidad objetiva de la secuencia resultante.

Tabla 5.3.- Configuración establecida en el método Temporal-Espacial para cada secuencia

Secuencia	1ª Fase (SR Temporal) - Secuencia MR_t		2ª Fase (SR Espacial) - Secuencia HR	
	SA (píxeles)	MBS (píxeles)	SA (píxeles)	MBS (píxeles)
<i>Water Cooler</i>	16	16	32	8
<i>Mobcal</i>	20	8	12	32
<i>Stockholm</i>	20	16	12	32
<i>Shields</i>	20	16	12	16
<i>Parkrun</i>	20	16	12	32

A continuación se mostrarán las gráficas con los resultados más representativos del método de SR MC Temporal-Espacial, incluyendo los valores medios de PSNR y SSIM de las simulaciones, así como *frames* de las secuencias resultantes con el fin de realizar una medida subjetiva de la calidad obtenida, para cada uno de los modos de pre-procesamiento propuestos, además de los resultados obtenidos al realizar una interpolación equivalente a la misma secuencia de entrada aplicando el mismo modo de pre-procesamiento.

5.3.1.1 Modo de pre-procesamiento *Full-Frame*

En la Tabla 5.4, se muestran los valores medios de PSNR y SSIM obtenidos a partir del método Temporal-Espacial, para el modo de pre-procesamiento *Full-Frame* en el caso de las secuencias de *test* consideradas, junto con la versión correspondiente a la secuencia interpolada (INTx4). En ellas, se puede observar cómo en cada cámara se consigue en la primera fase que los valores medios de PSNR y SSIM en las secuencias súper-resueltas estén en términos generales, por encima de los correspondientes a la secuencia interpolada. Por otra parte, en la segunda fase se puede ver cómo en la gran mayoría de los casos se supera tanto en PSNR como en SSIM a la secuencia interpolada HR (INTx4) en media.

Así, a modo de referencia, como se deduce de los valores medios de la segunda fase de la secuencia *Water Cooler*, existen casos en que la secuencia súper-resuelta se encuentra por debajo de los resultados obtenidos tras haber interpolado la secuencia MR_t (SR+INT) resultante de la primera fase del método de SR. Por tanto, se considera esta segunda fase crítica, pudiendo haber ciertos casos en los que resulte más conveniente llevar a cabo una interpolación sobre la secuencia MR_t , dependiendo sobre todo de los resultados obtenidos a partir de la primera fase.

Finalmente, se resaltan algunos valores significativos, como por ejemplo los resultados obtenidos para las secuencias *Stockholm* o *Shields*, en los que, por ejemplo, en la cámara 5 de *Stockholm* se supera en media en la primera fase a la secuencia interpolada MR en 1.25 dB en PSNR y 0.056 en SSIM, y en la cámara 1 de la secuencia *Shields* se supera en media 0.92 dB en PSNR y de 0.044 en SSIM. Por otra parte, en la segunda fase de este método y por tanto en la secuencia final, para el caso de *Stockholm* se supera a la secuencia interpolada en media 1.63 dB en PSNR y 0.068 en SSIM, y en la secuencia *Shields* se obtiene una mejora de 1.5 dB en PSNR y de 0.059 en SSIM. Así mismo, en la Tabla 5.4 se han destacado los valores más representativos de cada secuencia, para cada fase.

En la secuencia *Stockholm*, se puede apreciar la mejora aportada por el método Temporal-Espacial entre las secuencias MR_t y HR, mostradas en la Figura 5.10.a y en la Figura 5.10.b respectivamente, en los tejados de las casas, así como en los coches situados en la esquina inferior izquierda. Así mismo, si se compara la secuencia súper-resuelta HR con respecto a la interpolada HR, representadas en la Figura 5.11.b y en la Figura 5.12, respectivamente, se puede observar cómo además de los tejados y los coches anteriormente mencionados, las fachadas de las casas muestran una mejor calidad de imagen.

Por otra parte, en la secuencia *Shields*, se puede apreciar la mejora aportada por el método Temporal-Espacial entre las secuencias MR_t y HR, mostradas en la Figura 5.12.a y en la Figura 5.12.b, respectivamente, en las letras y en la definición de los escudos. Así mismo si se compara la secuencia súper-resuelta HR con respecto a la interpolada HR, representadas en la Figura 5.13.b y en la Figura 5.14, respectivamente, se puede observar una mejor definición en el suavizado de los bordes.

Tabla 5.4.- Valores promedio de resultados en método Temporal-Espacial (modo Full-Frame)

Secuencia	C A M	1ª Fase (SR Temporal) Secuencia MR _t				2ª Fase (SR Espacial) Secuencia HR				INTx4	
		PSNR (dB)		SSIM		PSNR (dB)		SSIM		PSNR (dB)	SSIM
		SR	INT	SR	INT	SR	SR+INT	SR	SR+INT		
Water Cooler	1	27.43	27.36	0.897	0.893	27.45	27.60	0.861	0.869	27.10	0.856
	2	28.04	27.98	0.902	0.901						
	3	27.87	27.76	0.899	0.897						
	4	27.51	27.49	0.903	0.900						
Mobcal	1	26.97	26.87	0.815	0.812	26.89	26.81	0.887	0.872	26.63	0.866
	2	26.90	26.79	0.812	0.809						
	3	26.80	26.70	0.807	0.806						
	4	27.10	27.02	0.818	0.816						
	5	27.00	26.93	0.814	0.813						
	6	26.91	26.87	0.810	0.810						
	7	27.20	27.17	0.820	0.819						
	8	27.13	27.09	0.817	0.816						
	9	27.04	27.03	0.813	0.813						
Stockholm	1	27.03	25.75	0.796	0.741	26.96	26.43	0.904	0.880	25.33	0.836
	2	27.01	25.69	0.796	0.740						
	3	26.95	25.65	0.797	0.741						
	4	26.90	25.69	0.791	0.737						
	5	26.90	25.64	0.792	0.736						
	6	26.85	25.59	0.792	0.737						
	7	26.84	25.60	0.789	0.732						
	8	26.80	25.54	0.789	0.731						
	9	26.76	25.50	0.790	0.732						
Shields	1	26.63	25.71	0.821	0.777	26.88	26.22	0.928	0.900	25.38	0.869
	2	26.65	25.67	0.819	0.776						
	3	26.66	25.73	0.821	0.778						
	4	26.53	25.69	0.816	0.777						
	5	26.51	25.65	0.815	0.776						
	6	26.54	25.71	0.817	0.778						
	7	26.44	25.68	0.812	0.776						
	8	26.40	25.65	0.811	0.774						
	9	26.45	25.70	0.813	0.777						
Parkrun	1	19.34	18.56	0.653	0.587	19.83	19.33	0.828	0.784	18.64	0.738
	2	19.30	18.53	0.653	0.588						
	3	19.28	18.53	0.655	0.590						
	4	19.70	18.91	0.661	0.598						
	5	19.64	18.88	0.660	0.599						
	6	19.63	18.87	0.662	0.600						
	7	20.13	19.26	0.674	0.608						
	8	20.10	19.23	0.674	0.609						
	9	20.05	19.22	0.675	0.610						



a)

b)

Figura 5.10.- Frames nº 12 súper-resueltos MR_t (a) y HR (b) para la secuencia Stockholm en el método Temporal-Espacial en modo Full-Frame



a)

b)

Figura 5.11.- Frames nº 12 interpolados MR_t (a) y SR+INT (b) para la secuencia Stockholm en el método Temporal-Espacial en modo Full-Frame



Figura 5.12.- Frame nº 12 interpolado INTx4 para la secuencia Stockholm en el método Temporal-Espacial en modo Full-Frame



a) b)
Figura 5.13.- Frames nº 14 súper-resueltos MRt (a) y HR (b) para la secuencia Shields en el método Temporal-Espacial en modo Full-Frame



a) b)
Figura 5.14.- Frames nº 14 interpolados MRt (a) y SR+INT (b) para la secuencia Shields en el método Temporal-Espacial en modo Full-Frame



Figura 5.15.- Frame nº 14 interpolado INTx4 para la secuencia Shields en el método Temporal-Espacial en modo Full-Frame

5.3.1.2 Modo de pre-procesamiento *Overlap*

Debido a que los resultados de la primera fase del método de SR MC Temporal-Espacial coinciden en todos los modos de pre-procesamiento, al aplicarse estos sobre el proceso de SR Espacial, en este apartado se muestran únicamente los resultados de esta segunda fase. A continuación, se representan los valores medios de PSNR y SSIM obtenidos a partir del método Temporal-Espacial, con el modo de pre-procesamiento *Overlap* para las secuencias de *test* consideradas. Los resultados del proceso de SR Temporal coinciden con los mostrados en la Tabla 5.4, puesto que el modo de pre-procesamiento *Overlap* se realiza en la fase de SR Espacial. Junto a ellos se muestra el resultado de haber interpolado la secuencia del modo de pre-procesamiento *Overlap* (INTx4).

Tal y como se puede observar en la Tabla 5.5, los valores medios que se obtienen en las secuencias de *test* súper-resueltas mediante este modo de pre-procesamiento, superan los valores de PSNR y SSIM de la secuencia interpolada (INTx4), así como la secuencia MR_t interpolada (SR+INT). Sin embargo, en la secuencia *Water Cooler* sus valores se encuentran por encima o muy próximos a los de la secuencia súper-resuelta en la mayoría de los *frames*, situación que también se produce en la secuencia *Mobcal*, aunque en mucha menor medida. Al igual que se comentó en el modo de pre-procesamiento *Full-Frame*, se debería considerar en esas situaciones realizar una interpolación de esos *frames*.

Tabla 5.5.- Valores promedio de resultados en método Temporal-Espacial (modo *Overlap*)

Secuencia	1ª Fase (SR Temporal) Secuencia MR _t				2ª Fase (SR Espacial) Secuencia HR				INTx4	
	PSNR (dB)		SSIM		PSNR (dB)		SSIM		PSNR (dB)	SSIM
	SR	INT	SR	INT	SR	SR+INT	SR	SR+INT		
<i>Water Cooler</i>	Ver Tabla 5.4				26.77	26.80	0.852	0.857	25.94	0.836
<i>Mobcal</i>					26.84	26.81	0.857	0.847	26.60	0.840
<i>Stockholm</i>					26.62	26.00	0.846	0.815	24.89	0.761
<i>Shields</i>					27.04	26.23	0.884	0.846	25.37	0.806
<i>Parkrun</i>					19.47	18.86	0.718	0.657	18.22	0.612

Así mismo, como valores más representativos de este modo de pre-procesamiento se pueden observar los obtenidos en la secuencia *Stockholm* y en la secuencia *Shields*, consiguiendo con respecto a la interpolación una mejora de 1.73 dB en PSNR y de 0.085 en SSIM en *Stockholm*, mientras que en la secuencia *Shields* se consigue superar en 1.67 dB en PSNR y 0.078 en SSIM.

Al igual que sucedía en el modo de pre-procesamiento *Full-Frame*, se observa cómo para la secuencia *Stockholm*, mostrada en la Figura 5.16, se ha conseguido reconstruir con una mejor definición los tejados de las casas, así como las fachadas de los edificios en la secuencia súper-resuelta (Figura 5.16.a) con respecto a la secuencia interpolada (Figura 5.16.b).

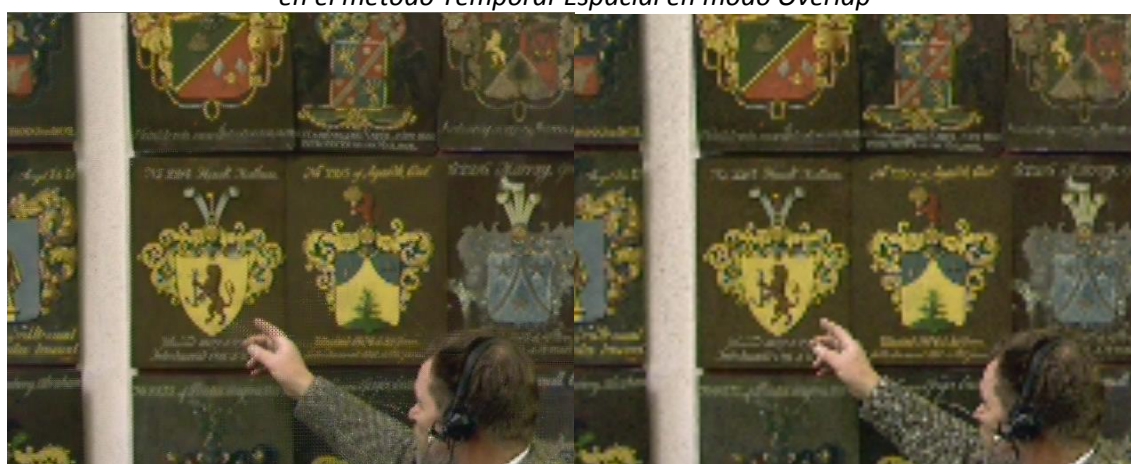
Así mismo para la secuencia *Shields*, mostrada en la Figura 5.17, se ha conseguido reconstruir con mejor definición los escudos y las palabras que aparecen encima y debajo de estos en la secuencia súper-resuelta (Figura 5.17.a) con respecto a la secuencia interpolada (Figura 5.17.b).



a)

b)

Figura 5.16.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia Stockholm en el método Temporal-Espacial en modo Overlap



a)

b)

Figura 5.17.- Frames nº 14 súper-resuelto (a) e interpolado (b) HR para la secuencia Shields en el método Temporal-Espacial en modo Overlap

5.3.1.3 Modo de pre-procesamiento *Overlap+Borders*

En este modo de pre-procesamiento, al igual que en la sección anterior, sólo se mostrarán los resultados de la segunda fase del proceso de SR. Las curvas que se muestran a continuación se corresponden con los resultados de la secuencia una vez reconstruidos los *frames*, tras haber sido fusionados los bordes con la parte solapada en la secuencias. A continuación, se muestran los valores medios de PSNR y SSIM obtenidos en el método Temporal-Espacial para las secuencias de *test* consideradas con el modo de pre-procesamiento *Overlap+Borders*. Los resultados del proceso de SR Temporal coinciden con los mostrados en la Tabla 5.15, puesto que el modo de pre-procesamiento *Overlap+Borders* se realiza en la fase de SR Espacial. Así mismo, se muestran los resultados obtenidos tras realizar el mismo pre-procesamiento en el proceso de interpolación (INTx4). En este caso se puede observar en la Tabla 5.6 cómo los valores medios de los resultados de la secuencia súper-resuelta se encuentran por encima de los correspondientes a la secuencia interpolada (INTx4), en la cual se ha procesado cada borde por separado al igual que la secuencia súper-resuelta. Como valores más representativos de este modo de pre-procesamiento, se observa que en *Stockholm* se supera en 1.55 dB de PSNR y 0.079 de SSIM a los valores de la secuencia interpolada, mientras que en *Shields* se produce una mejora de 1.55 dB en PSNR y de 0.073 en SSIM, con respecto de la secuencia interpolada.

Tabla 5.6.- Valores promedio de resultados en método Temporal-Espacial (modo Overlap+Borders)

Secuencia	1ª Fase (SR Temporal) Secuencia MR _t				2ª Fase (SR Espacial) Secuencia HR				INTx4	
	PSNR (dB)		SSIM		PSNR (dB)		SSIM		PSNR (dB)	SSIM
	SR	INT	SR	INT	SR	SR+INT	SR	SR+INT		
Mobcal	Ver Tabla 5.4				26.86	26.80	0.767	0.766	26.63	0.762
Stockholm					27.00	26.42	0.705	0.677	25.35	0.626
Shields					26.92	26.21	0.748	0.719	25.37	0.675
Parkrun					19.86	19.33	0.560	0.508	18.63	0.449

Tabla 5.7.- Valores promedio de los bordes en el método Temporal-Espacial

Secuencia	Borde	Método Temporal-Espacial – 2ª Fase					
		PSNR (dB)		SSIM		INTx4	
		SR	SR+INT	SR	SR+INT	PSNR (dB)	SSIM
Mobcal	Izquierdo	26.92	26.99	0.755	0.759	26.98	0.758
	Izq. Inferior	32.62	32.55	0.864	0.867	32.94	0.865
	Izq. Superior	38.19	38.10	0.901	0.900	37.83	0.896
	Inferior	27.31	27.16	0.741	0.745	26.94	0.743
	Solape	26.95	26.87	0.860	0.847	26.71	0.841
	Derecho	31.11	31.18	0.871	0.873	31.00	0.868
	Der. Inferior	24.81	24.62	0.621	0.634	25.41	0.688
	Der. Superior	24.47	24.50	0.583	0.593	24.26	0.592
Stockholm	Inferior	25.25	25.27	0.692	0.698	24.93	0.693
	Izquierdo	25.85	25.52	0.724	0.716	24.00	0.658
	Izq. Inferior	28.58	28.45	0.678	0.679	28.00	0.664
	Izq. Superior	38.07	37.99	0.870	0.869	37.34	0.861
	Inferior	29.32	29.25	0.731	0.732	28.19	0.696
	Solape	26.74	26.12	0.844	0.812	25.00	0.761
	Derecho	27.16	27.03	0.709	0.707	26.24	0.673
	Der. Inferior	29.97	29.89	0.736	0.743	28.32	0.700
Shields	Der. Superior	37.48	37.37	0.868	0.867	38.54	0.857
	Inferior	34.53	34.44	0.864	0.863	33.33	0.849
	Izquierdo	28.50	28.10	0.796	0.791	26.90	0.751
	Izq. Inferior	32.52	32.37	0.840	0.839	31.62	0.815
	Izq. Superior	29.95	29.66	0.814	0.800	28.58	0.750
	Inferior	25.40	25.32	0.695	0.700	25.10	0.684
	Solape	26.96	26.15	0.881	0.843	25.34	0.804
	Derecho	27.85	27.61	0.774	0.770	26.84	0.733
Parkrun	Der. Inferior	30.24	29.89	0.797	0.791	28.70	0.737
	Der. Superior	27.51	27.25	0.739	0.728	26.00	0.664
	Inferior	26.61	26.23	0.773	0.753	24.57	0.665
	Izquierdo	19.92	18.82	0.534	0.533	19.10	0.484
	Izq. Inferior	27.13	26.84	0.714	0.704	25.93	0.672
	Izq. Superior	14.84	14.87	0.326	0.345	14.45	0.311
	Inferior	27.37	27.00	0.692	0.679	26.38	0.662
	Solape	20.10	19.44	0.750	0.687	18.71	0.633
Derecho	21.43	21.32	0.546	0.546	20.51	0.492	
Der. Inferior	28.00	27.75	0.680	0.671	26.77	0.645	
Der. Superior	16.46	16.36	0.374	0.385	15.66	0.352	
Inferior	15.16	15.20	0.326	0.345	14.73	0.322	

Para este modo de pre-procesamiento se presentan también, en la Tabla 5.8, los valores medios de PSNR y SSIM intermedios, obtenidos en cada uno de los bordes en los que se han dividido los *frames* de la secuencia en el proceso de SR Espacial. A través de este desglose de valores medios de la calidad de cada borde y el solape, se puede observar cómo en la mayoría de los casos es en el solape donde se produce una mayor diferencia de PSNR y SSIM entre la secuencia súper-resuelta (SR) e interpolada (INTx4), concluyéndose que es en el solape donde se encuentra concentrada una mayor cantidad de información.

Si se comparan los resultados de la secuencia súper-resuelta para este modo de pre-procesamiento con los obtenidos en la secuencia interpolada en el modo de pre-procesamiento *Full-Frame*, se puede observar que la calidad de la secuencia súper-resuelta se encuentra por debajo de la interpolada. Así, al haber sido procesados los bordes y el solape por separado, se ha producido una degradación en la imagen súper-resuelta, reflejándose este hecho sobre todo en los valores de SSIM. La Tabla 5.9 presenta una comparativa global de los modos de pre-procesamiento *Full-Frame*, *Overlap* y *Overlap+Borders* para el método Temporal-Espacial a la salida de la segunda fase, donde se hace evidente el efecto comentado.

Tabla 5.8.- Resumen de valores promedio de los resultados de SSIM (método Temporal-Espacial)

Modo	Full-Frame			Overlap			Overlap+Borders		
	SR	SR+INT	INTx4	SR	SR+INT	INTx4	SR	SR+INT	INTx4
<i>Mobcal</i>	0.887	0.872	0.866	0.852	0.857	0.836	0.767	0.766	0.762
<i>Stockholm</i>	0.904	0.880	0.836	0.857	0.847	0.840	0.705	0.677	0.626
<i>Shields</i>	0.928	0.900	0.869	0.846	0.815	0.761	0.748	0.719	0.675
<i>Parkrun</i>	0.828	0.784	0.738	0.884	0.846	0.806	0.560	0.508	0.449
<i>PROMEDIO</i>	0.887	0.859	0.827	0.860	0.841	0.811	0.695	0.667	0.628
<i>POSICIÓN</i>	1	3	5	2	4	6	7	8	9

Este efecto se produce de forma similar con los métodos Espacial-Temporal y Temporal-Espacial, por lo que en el desarrollo de la presente Tesis Doctoral se decidió descartarlo; por esta razón en los apartados posteriores relativos a los métodos Espacial-Temporal y Temporal-Espacial tan sólo se presentarán resultados para los modos de pre-procesamiento *Full-Frame* y *Overlap*.

Como se puede apreciar en las imágenes mostradas en la Figura 5.18.a y en la Figura 5.19.a, la calidad de las imágenes en la parte central o solape es similar a la obtenida en el resto de modos de pre-procesamiento. Sin embargo, como se comentó anteriormente, los resultados que se obtienen tras fusionar los bordes con el solape se encuentran por debajo de los valores obtenidos a partir del modo de pre procesamiento *Full-Frame*, lo cual se puede observar sobre todo en la esquina inferior izquierda de la secuencia *Stockholm*, representada en la Figura 5.18.a, y en la esquina inferior derecha en la secuencia *Shields*, mostrada en la Figura 5.19.a, ya que el resto de bordes son muy homogéneos.

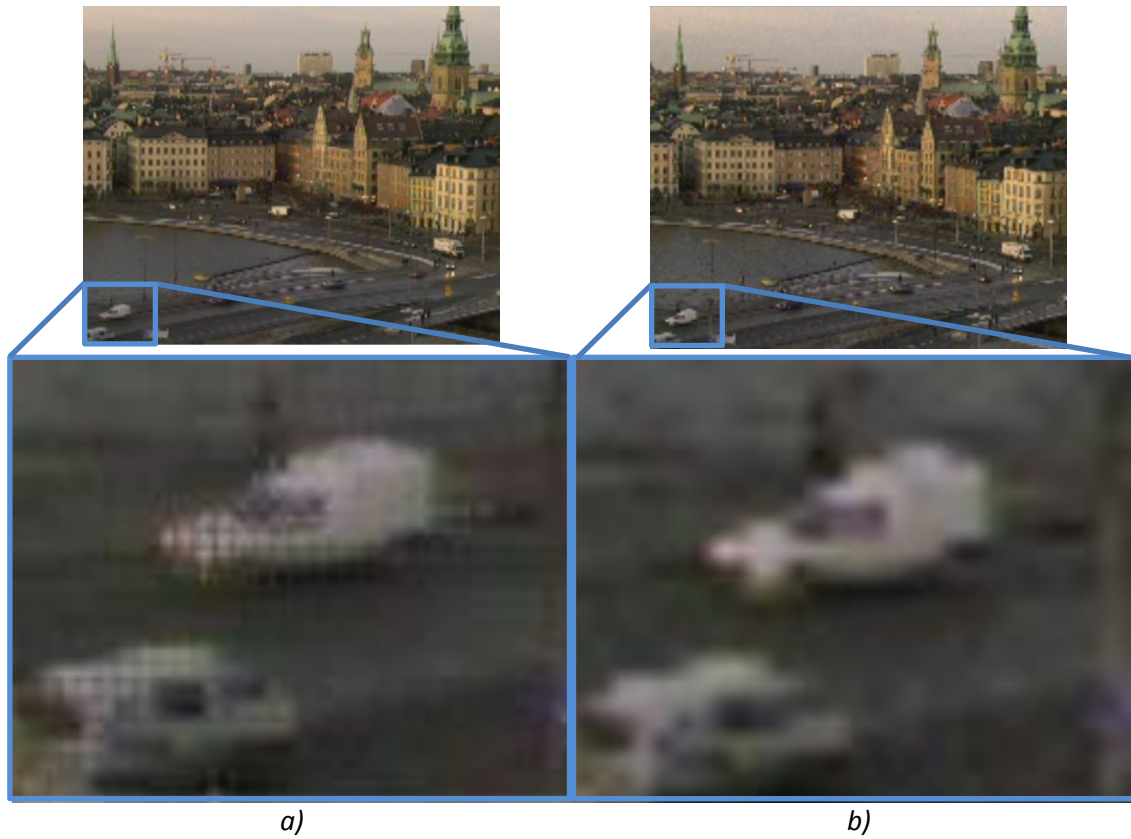


Figura 5.18.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia Stockholm en el método Temporal Espacial en modo Overlap+Borders



Figura 5.19.- Frames nº 14 súper-resuelto (a) e interpolado (b) HR para la secuencia Shields en el método Temporal Espacial en modo Overlap+Borders

5.3.2 Método Espacial-Temporal

En este apartado, se presentan los resultados obtenidos para cada una de las secuencias propuestas en la caracterización del método Espacial-Temporal. La configuración empleada para cada secuencia, en cada uno de los procesos de SR que se llevan a cabo en este método, se muestra en la Tabla 5.9, maximizando la calidad objetiva en cada caso. Al igual que en el método Temporal-Espacial, los procesos de SR Temporal realizados en este método para cada secuencia constan de una Ventana de Trabajo de SR de ± 7 frames.

Del mismo modo que sucedía en el método Temporal-Espacial, a la hora de elegir el valor de SA en la fase de SR Espacial, este valor debe seleccionarse teniendo en cuenta el desplazamiento entre cámaras. En este caso, al igual que en anteriores casos, la separación entre las cámaras HR se corresponde con 21 píxeles, por lo que al haber realizado un proceso de diezmado con un Factor de Escala 4, la separación entre cámaras en la secuencia LR se corresponde con $21/4 = 5.25$. Este valor condicionará que el valor mínimo de SA sea de 6, optándose en este caso por asignarle un valor estándar de SA de 8 píxeles, ya que este incremento en SA no va a afectar significativamente al rendimiento.

Tabla 5.9.- Configuración establecida del método Espacial-Temporal para cada secuencia

Secuencia	1ª Fase (SR Espacial) - Secuencia MR _s		2ª Fase (SR Temporal) - Secuencia HR	
	SA (píxeles)	MBS (píxeles)	SA (píxeles)	MBS (píxeles)
Water Cooler	16	8	32	32
Mobcal	8	16	32	16
Stockholm	8	16	32	32
Shields	8	16	32	16
Parkrun	8	16	32	32

A continuación, se mostrarán los valores medios de PSNR y SSIM resultantes de las simulaciones para el método Espacial-Temporal, así como algunos frames de las secuencias obtenidas, con el fin de realizar una medida subjetiva de calidad para los modos de pre-procesamiento *Full-Frame* y *Overlap*.

5.3.2.1 Modo de pre-procesamiento Full-Frame

Se observa que en la primera fase del método Espacial-Temporal se consiguen unos resultados mejores que en el método Temporal-Espacial. Esto se debe a que el efecto de escalado que existe entre los frames de la Ventana de Trabajo de SR temporal, en la primera fase del método Temporal-Espacial, no se produce en el proceso de SR Espacial, puesto que todas las cámaras capturan el mismo instante de tiempo, lo cual hace que la Ventana de Trabajo de SR espacial esté compuesta por frames en los que sólo habrá escalado debido a la posición de las cámaras en el array MC, que es menor que el que hay entre los frames de la Ventana de Trabajo temporal. De esta forma, el método Espacial-Temporal se beneficia de los buenos resultados proporcionados por la SR Espacial de la primera fase, compensando la degradación introducida en el proceso de SR Temporal, lo que conlleva que los resultados de la secuencia HR se mantengan mucho más estables, estando por encima o muy próximos a la interpolación bilineal.

En la Tabla 5.10 se muestran los valores medios de PSNR y SSIM obtenidos a partir del método Espacial-Temporal con el modo de pre-procesamiento *Full-Frame* para las secuencias de test propuestas, junto con la versión interpolada correspondiente (INTx4). Estos valores muestran cómo, tanto en la primera fase, en la que se lleva a cabo un proceso de SR Espacial, como en la segunda fase, en la que se realiza un proceso de SR Temporal, las secuencias súper-

resueltas MR_s y HR superan los resultados de la secuencia interpolada MR e INTx4, respectivamente. De entre estos resultados se pueden destacar los correspondientes a la secuencia *Stockholm*, para la que en la fase de SR Espacial se supera en 1.09 dB en PSNR y en 0.043 en SSIM a la secuencia interpolada, mientras que en su segunda fase se observa una diferencia de 1.2 dB en PSNR y 0.061 en SSIM. Así mismo, para la secuencia *Shields*, en la fase de SR Espacial se supera en 1.4 dB en PSNR y en 0.053 en SSIM respecto a la secuencia interpolada, mientras que en su segunda fase se observa una diferencia de 1.13 dB en PSNR y de 0.053 en SSIM. Por otra parte, se observa en los resultados de la segunda fase de este método, mostrados en la Tabla 5.10, cómo la calidad de la secuencia SR+INT se encuentra por encima de la secuencia súper-resuelta, lo cual se da en las secuencias *Water Cooler* y *Mobcal*. A partir de estos valores, se puede determinar que ésta es una fase crítica en la que puede ser más conveniente interpolar la secuencia MR_s , al igual que sucedía en el método Temporal-Espacial, apreciándose más este hecho en la segunda fase en el método Espacial-Temporal.

Tabla 5.10.- Valores promedio de resultados en método Espacial-Temporal (modo Full-Frame)

Secuencia	1ª Fase (SR Espacial) Secuencia MR_s				2ª Fase (SR Temporal) Secuencia HR				INTx4	
	PSNR (dB)		SSIM		PSNR (dB)		SSIM		PSNR (dB)	SSIM
	SR	INT	SR	INT	SR	SR+INT	SR	SR+INT		
<i>Water Cooler</i>	27.74	27.36	0.903	0.893	27.52	27.55	0.861	0.868	27.10	0.856
<i>Mobcal</i>	27.70	26.93	0.834	0.813	26.77	27.42	0.886	0.895	26.63	0.866
<i>Stockholm</i>	26.73	25.64	0.779	0.736	26.53	26.38	0.897	0.883	25.33	0.836
<i>Shields</i>	27.05	25.65	0.829	0.776	26.51	26.63	0.922	0.916	25.38	0.869
<i>Parkrun</i>	20.00	18.88	0.675	0.599	19.57	19.63	0.822	0.815	18.64	0.738

Finalmente, en la Figura 5.20 se muestra el *frame* 10 correspondiente a la secuencia *Water Cooler*, en la Figura 5.21 el *frame* 42 de la secuencia *Mobcal*, en la Figura 5.22 el *frame* 12 de la secuencia *Stockholm*, y en la Figura 5.23 el *frame* 14 de la secuencia *Shields*, de modo que se pueda realizar una valoración subjetiva de la calidad. A modo de referencia, se presentan los *frames* interpolados, y el *frame* resultante de haber interpolado la secuencia súper-resuelta MR_t , el cual se ha designado como SR+INT (para las secuencias *Stockholm* y *Shields*), además de los *frames* súper-resueltos e interpolados (en todas las secuencias).

En las imágenes mostradas en la Figura 5.20.a y en la Figura 5.20.b, se puede observar cómo en la secuencia *Water Cooler* existe una mejoría de la secuencia de salida de la segunda fase SR+INT frente a la secuencia SR, sobre todo en los papeles encima de la mesa.



Figura 5.20.- Frames nº 10 de *Water Cooler* súper-resuelto (a) e interpolado (b) a la salida de la 2ª fase de SR temporal en el método Espacial-Temporal en modo Full-Frame

El mismo efecto se produce en la secuencia *Mobcal*, tal y como se representa en la Figura 5.21.a y en la Figura 5.21.b, donde se observa principalmente una mejora en la reconstrucción del muñeco con camisa a rayas.

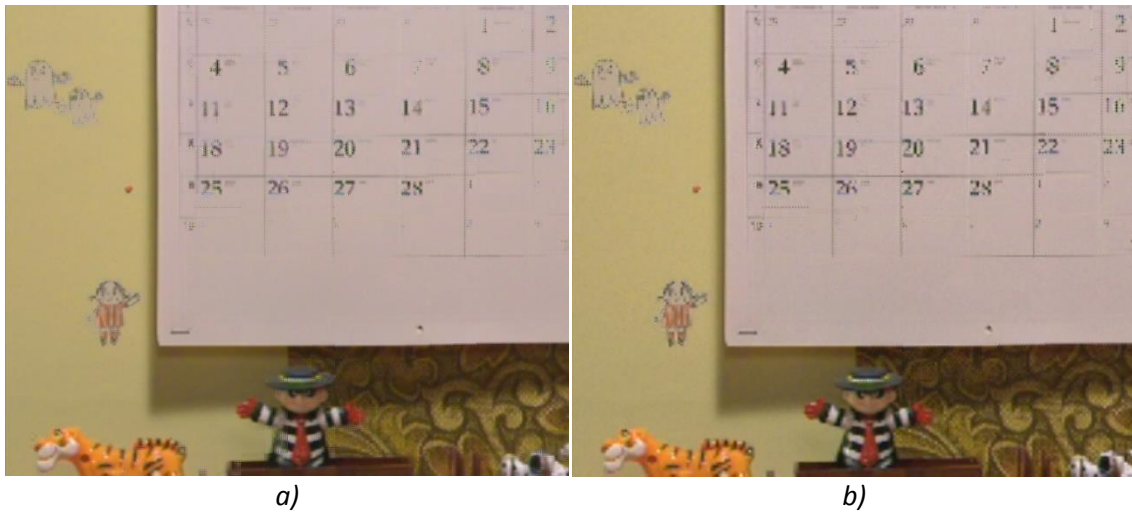


Figura 5.21.- Frames nº 42 de *Mobcal* súper-resuelto (a) e interpolado (b) a la salida de la 2ª fase de SR temporal en el método Espacial-Temporal en modo Full-Frame



Figura 5.22.- Frames nº 12 súper-resueltos MR₅ (a) y HR (b) e interpolados MR (c) y SR+INT (d) para la secuencia *Stockholm* en el método Espacial-Temporal en modo Full-Frame

En la secuencia *Stockholm*, se puede ver cómo las fachadas de las casas, así como los tejados de éstas, mostradas en la Figura 5.22.a y en la Figura 5.22.b, presentan una mejor definición que en la secuencia interpolada, representada en la Figura 5.22.c. Del mismo modo, si se compara la secuencia súper-resuelta HR frente a la secuencia SR+INT, mostrada en la Figura 5.22.d, se observa un menor pixelado de los tejados en la secuencia súper-resuelta, si bien los coches situados en la esquina inferior izquierda se encuentran mejor definidos en la secuencia SR+INT. Así mismo, en la secuencia *Shields* se puede ver cómo los escudos y las palabras que aparecen encima y debajo de estos, mostradas en la Figura 5.23.a y en la Figura 5.23.b, respectivamente presentan una mejor definición que en la secuencia interpolada, representada en la Figura 5.23.c. Del mismo modo, si se compara la secuencia súper-resuelta HR frente a la secuencia SR+INT, mostrada en la Figura 5.23.d, se observa un menor pixelado de los bordes en la secuencia súper-resuelta, pero sin embargo se evitan artefactos en el dedo del guía en la secuencia SR+INT.



Figura 5.23.- Frames nº 14 súper-resueltos MR_s (a) y HR (b) e interpolados MR (c) y SR+INT (d) para la secuencia *Shields* en el método Espacial-Temporal en modo Full-Frame

5.3.2.2 Modo de pre-procesamiento *Overlap*

En la Tabla 5.11 se muestran los valores medios de PSNR y SSIM obtenidos a partir del método de SR MC Espacial-Temporal aplicando el modo de pre-procesamiento *Overlap* para las secuencias de *test* propuestas. Junto a estos valores se muestra el resultado

correspondiente a haber interpolado la secuencia de entrada a este modo de pre-procesamiento (INTx4). A través de estos valores medios, se puede observar cómo la secuencia súper-resuelta MR_s supera generalmente la calidad de la secuencia interpolada MR, del mismo modo que sucede en la secuencia súper-resuelta HR. Sin embargo, se puede observar cómo en este modo de pre-procesamiento la segunda fase de SR se encuentra en más casos superada por los resultados de la secuencia SR+INT, considerándose una fase delicada en este método.

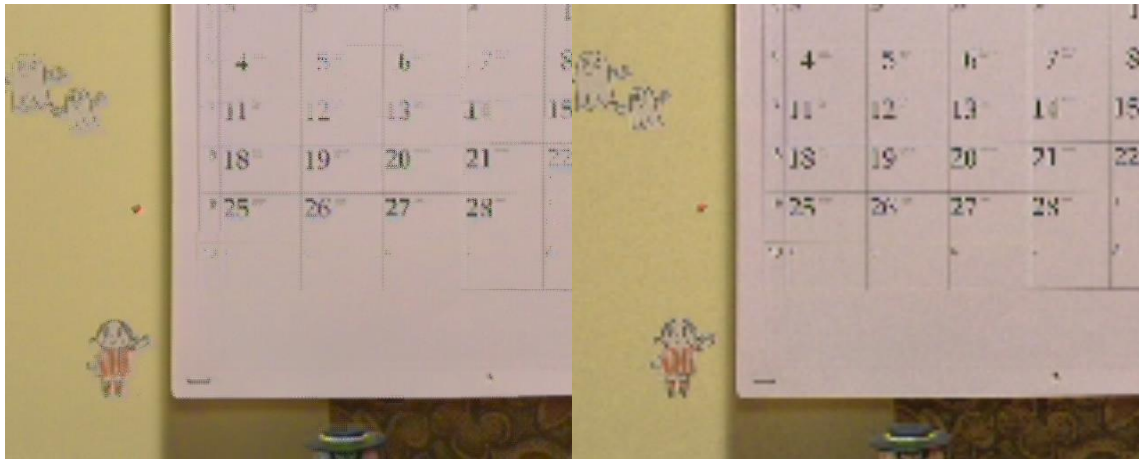
Tabla 5.11.- Valores promedio de resultados en método Espacial-Temporal (modo Overlap)

Secuencia	1ª Fase (SR Espacial) Secuencia MR _s				2ª Fase (SR Temporal) Secuencia HR				INTx4	
	PSNR (dB)		SSIM		PSNR (dB)		SSIM		PSNR (dB)	SSIM
	SR	INT	SR	INT	SR	SR+INT	SR	SR+INT		
<i>Water Cooler</i>	26.79	26.12	0.892	0.878	26.68	26.60	0.848	0.853	25.94	0.836
<i>Mobcal</i>	27.51	26.89	0.832	0.816	26.72	27.26	0.858	0.864	26.60	0.840
<i>Stockholm</i>	26.38	25.20	0.776	0.726	26.17	26.00	0.833	0.818	24.89	0.761
<i>Shields</i>	27.11	25.65	0.830	0.775	26.52	26.66	0.869	0.862	25.37	0.806
<i>Parkrun</i>	19.59	18.44	0.651	0.574	19.12	19.21	0.691	0.693	18.22	0.612

Finalmente, a continuación se muestra el *frame* 10 correspondiente a la secuencia *Water Cooler*, el *frame* 42 de la secuencia *Mobcal*, el *frame* 12 de la secuencia *Stockholm*, y el *frame* 14 de la secuencia *Shields*, de modo que se puede realizar una valoración subjetiva de la calidad. En las imágenes mostradas a continuación, se puede observar la secuencia de salida HR, así como sus correspondientes *frames* interpolados. En las imágenes mostradas en la Figura 5.24, correspondientes a la secuencia *Water Cooler*, se puede observar una mejoría de la secuencia súper-resuelta (Figura 5.24.a) frente a la secuencia interpolada (Figura 5.24.b), especialmente sobre los papeles de la mesa, así como en el borde de ésta, si bien también se aprecia una mejoría en los objetos sobre la mesa situada al fondo de esta habitación. Por otra parte, en la Figura 5.25 se presenta la secuencia *Mobcal*, en la que se muestra en la secuencia súper-resuelta (Figura 5.25.a) cómo la niña sobre el fondo amarillo presenta una mejor definición frente a la secuencia interpolada, representada en la Figura 5.25.b. Respecto a los resultados obtenidos para la secuencia *Stockholm*, mostrados en la Figura 5.26.a y en la Figura 5.26.b, se observa cómo especialmente en los tejados de la secuencia súper-resuelta presenta una mejor calidad en comparación con la secuencia interpolada. Para concluir, en los resultados correspondientes a la secuencia *Shields*, mostrados en la Figura 5.27.a y en la Figura 5.27.b, se observa cómo especialmente en los bordes de los escudos, la secuencia súper-resuelta presenta una mejor calidad en comparación con la secuencia interpolada.



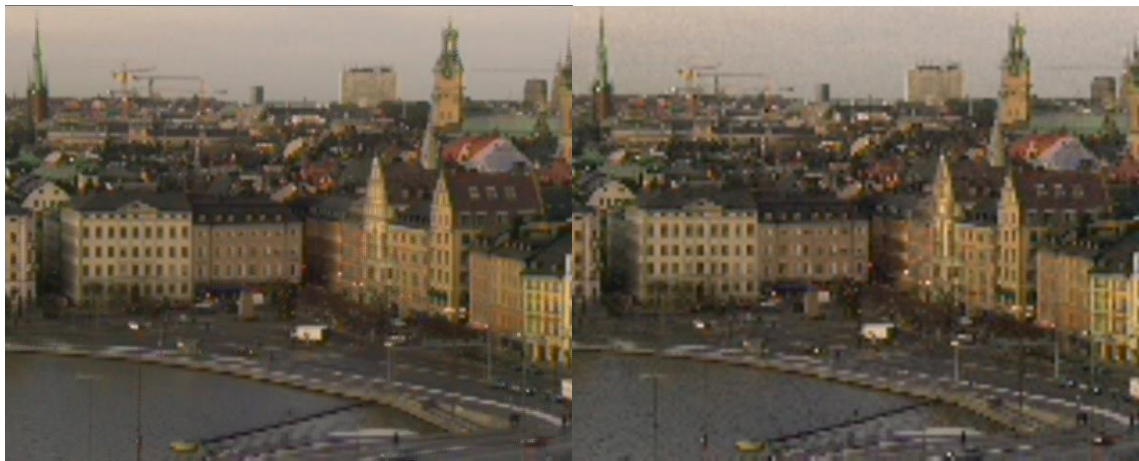
Figura 5.24.- Frames nº 10 súper-resuelto (a) e interpolado (b) HR para la secuencia *Water Cooler* en el método Espacial-Temporal en modo Overlap



a)

b)

Figura 5.25.- Frames nº 42 súper-resuelto (a) e interpolado (b) HR para la secuencia Mobcal en el método Espacial-Temporal en modo Overlap



a)

b)

Figura 5.26.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia Stockholm en el método Espacial-Temporal en modo Overlap



a)

b)

Figura 5.27.- Frames nº 14 súper-resuelto (a) e interpolado (b) HR para la secuencia Shields en el método Espacial-Temporal en modo Overlap

5.3.3 Método Mixto

En este apartado, se presentan los resultados obtenidos para cada una de las secuencias propuestas en la caracterización del método de SR MC Mixto. La configuración empleada para cada secuencia, en cada uno de los procesos de SR que se llevan a cabo en este método, se muestra en la Tabla 5.12, que presenta también las dimensiones de la Ventana de Trabajo de SR empleada en cada secuencia con el objetivo de maximizar la calidad objetiva.

Tabla 5.12.- Configuración establecida del método Mixto para cada secuencia

Secuencia	Método Mixto			
	SA (píxeles)	MBS (píxeles)	Winback (frames)	Winforward (frames)
<i>Water Cooler</i>	8	16	12	15
<i>Mobcal</i>	16	32	31	31
<i>Stockholm</i>	16	32	58	58
<i>Shields</i>	16	32	31	31
<i>Parkrun</i>	16	32	31	31

Tal y como se mostrará en las secciones posteriores, el método Mixto es el que mejor resultados proporciona en la mayoría de las secuencias, por lo que se presentará un estudio completo de éste con los modos de pre-procesamiento *Full-Frame* y *Overlap*. Se mostrarán algunas gráficas con los resultados más representativos, los valores medios de PSNR y SSIM resultantes de las simulaciones para el método Mixto, así como *frames* de las secuencias obtenidas con ampliación de detalles con el fin de realizar una medida subjetiva de la calidad para cada uno de los modos de pre-procesamiento.

5.3.3.1 Modo de pre-procesamiento *Full-Frame*

El método Mixto, a diferencia de los anteriores, consta de un único proceso de SR, por lo que los resultados mostrados a continuación presentan únicamente las curvas de PSNR y SSIM para la secuencia súper-resuelta (SR) y su versión interpolada equivalente (INTx4) como referencia para valorar los resultados.

Como se puede observar en la Figura 5.28.a y en la Figura 5.28.b, los resultados obtenidos con el método Mixto para la secuencia *Water Cooler* muestran, al igual que en el método Temporal-Espacial, cómo en los primeros 30 *frames* se supera la calidad de la secuencia interpolada, mientras que a partir del *frame* 30 los resultados se encuentran por debajo de los correspondientes a la secuencia interpolada. Si bien en este método se combina tanto la información espacial como la información temporal en su Ventana de Trabajo, con lo que se consigue aumentar las posibilidades de encontrar un MB más adecuado para reconstruir la secuencia, al mismo tiempo aumentan las posibilidades de que se produzca una mala decisión en la etapa de estimación de movimiento. Esta situación se produce a partir del *frame* 30, resultando en una reducción de la calidad en la secuencia súper-resuelta.

Como casos de secuencias cuya calidad mejora significativamente con respecto a la interpolación, en la Figura 5.29 y en la Figura 5.30 se presentan las gráficas de las secuencias *Mobcal* y *Stockholm*, respectivamente, donde se observa además cómo la mejora de la calidad sobre la secuencia interpolada es mayor si se compara con los resultados obtenidos en los métodos Temporal-Espacial y Espacial-Temporal. Sobre todo se puede resaltar el valor de SSIM que se obtiene en la secuencia *Stockholm*, el cual se encuentra alrededor de 0.94, tal y como se muestra en la Figura 5.30.b.

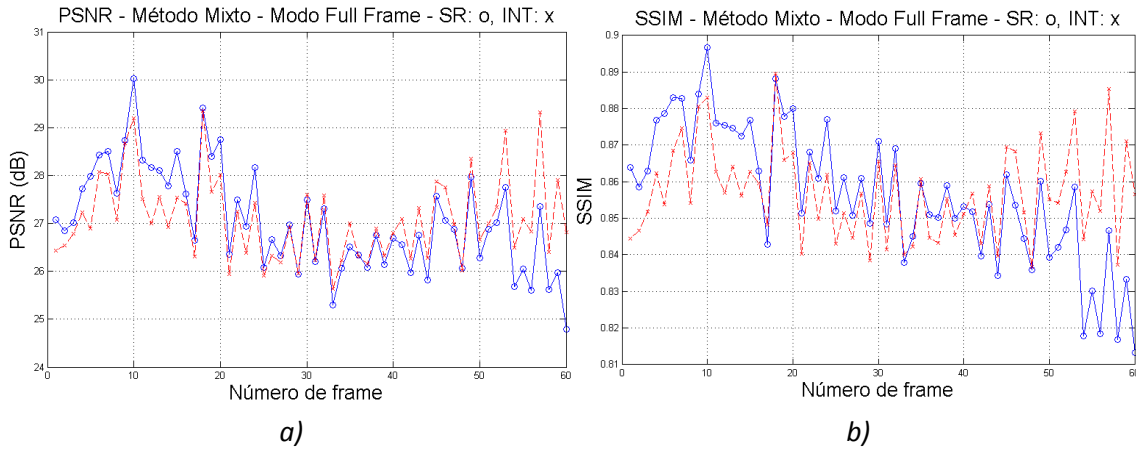


Figura 5.28.- PSNR y SSIM de método Mixto en modo Full-Frame en la secuencia Water Cooler

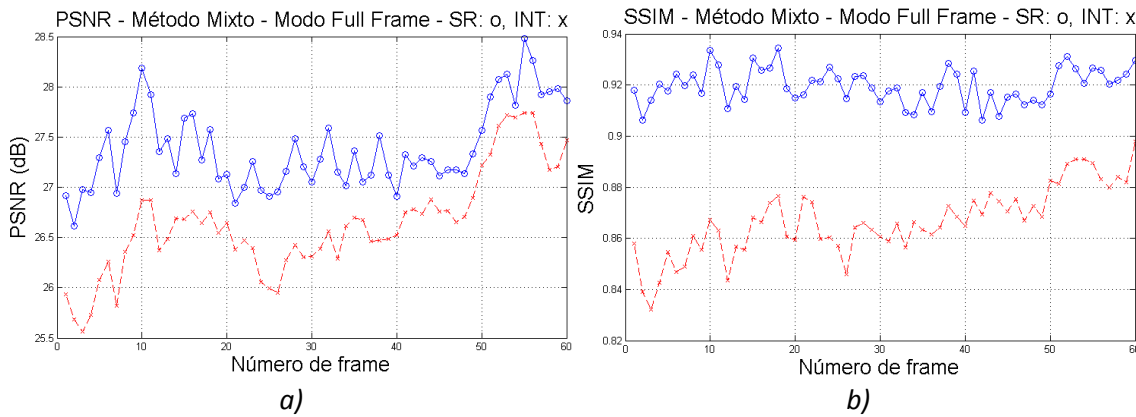


Figura 5.29.- PSNR y SSIM de método Mixto en modo Full-Frame en la secuencia Mobcal

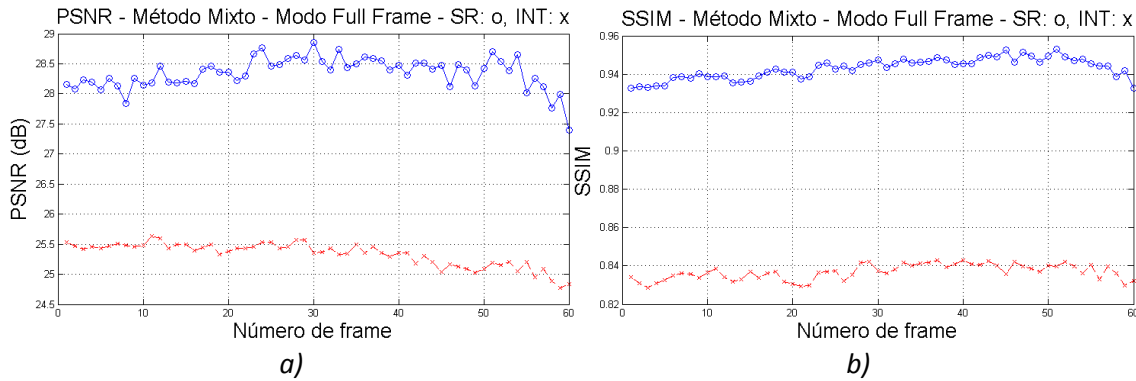


Figura 5.30.- PSNR y SSIM de método Mixto en modo Full-Frame en la secuencia Stockholm

A continuación se muestran los valores medios de PSNR y SSIM obtenidos a partir del método Mixto con el modo de pre-procesamiento *Full-Frame* para las secuencias de *test* consideradas. Los valores mostrados se corresponden con los de la secuencia súper-resuelta y su correspondiente versión interpolada (INTx4). Los valores medios representados en la Tabla 5.13 muestran cómo este método proporciona una mejora superior a los métodos Temporal-Espacial y Espacial-Temporal, en los que se puede destacar que se logra una mejora de 3.01 dB en PSNR y 0.106 en SSIM en la secuencia *Stockholm*, una mejora de 2.2 dB en PSNR y 0.082 en SSIM en *Shields*, y por último una mejora de 1.84 dB en PSNR y 0.164 en SSIM en *Parkrun*, respecto a la secuencia interpolada. Con respecto a las gráficas de la Figura 5.30 destaca además el hecho de que en el caso de la secuencia *Water Cooler*, aun siendo una secuencia muy compleja para MC, se produce mejora de calidad objetiva en la mayoría de los *frames*. Por otra parte, en una secuencia como *Mobcal*, en la que se combina movimiento global y

movimiento local, se consiguen mejoras para todos los *frames*. Finalmente, los resultados se optimizan en la secuencia *Stockholm*, al tratarse de una secuencia basada principalmente en movimiento global, que resulta muy adecuado para maximizar la coherencia espacial entre *frames*. Finalmente, en la Figura 5.31 se muestra el *frame* 10 correspondiente a la secuencia *Water Cooler*, en la Figura 5.32 el *frame* 42 de la secuencia *Mobcal*, y en la Figura 5.33 el *frame* 12 de la secuencia *Stockholm*, de modo que se pueda realizar una valoración subjetiva de la calidad. En las imágenes mostradas a continuación, se puede observar la secuencia de salida HR, así como sus correspondientes *frames* interpolados.

Tabla 5.13.- Valores promedio de los resultados del método Mixto en modo Full-Frame

Secuencia	Método Mixto					
	PSNR (dB)			SSIM		
	SR	INT	Diferencia	SR	INT	Diferencia
<i>Water Cooler</i>	27.05	27.1	-0.05	0.856	0.856	0
<i>Mobcal</i>	27.38	26.63	0.75	0.919	0.866	0.053
<i>Stockholm</i>	28.34	25.33	3.01	0.942	0.836	0.106
<i>Shields</i>	27.58	25.38	2.2	0.951	0.869	0.082
<i>Parkrun</i>	20.48	18.64	1.84	0.902	0.738	0.164

Tal y como se observa en la Figura 5.31, en la imagen de la secuencia súper-resuelta (Figura 5.31.a), se puede ver cómo los papeles sobre la mesa presentan una mejor definición que en la secuencia interpolada (Figura 5.31.b). Así mismo, en este caso se observa también cómo los objetos sobre la mesa situada debajo del cuadro al final de la habitación se ven mejor definidos en la secuencia súper-resuelta que en la secuencia interpolada, aunque se generen artefactos.

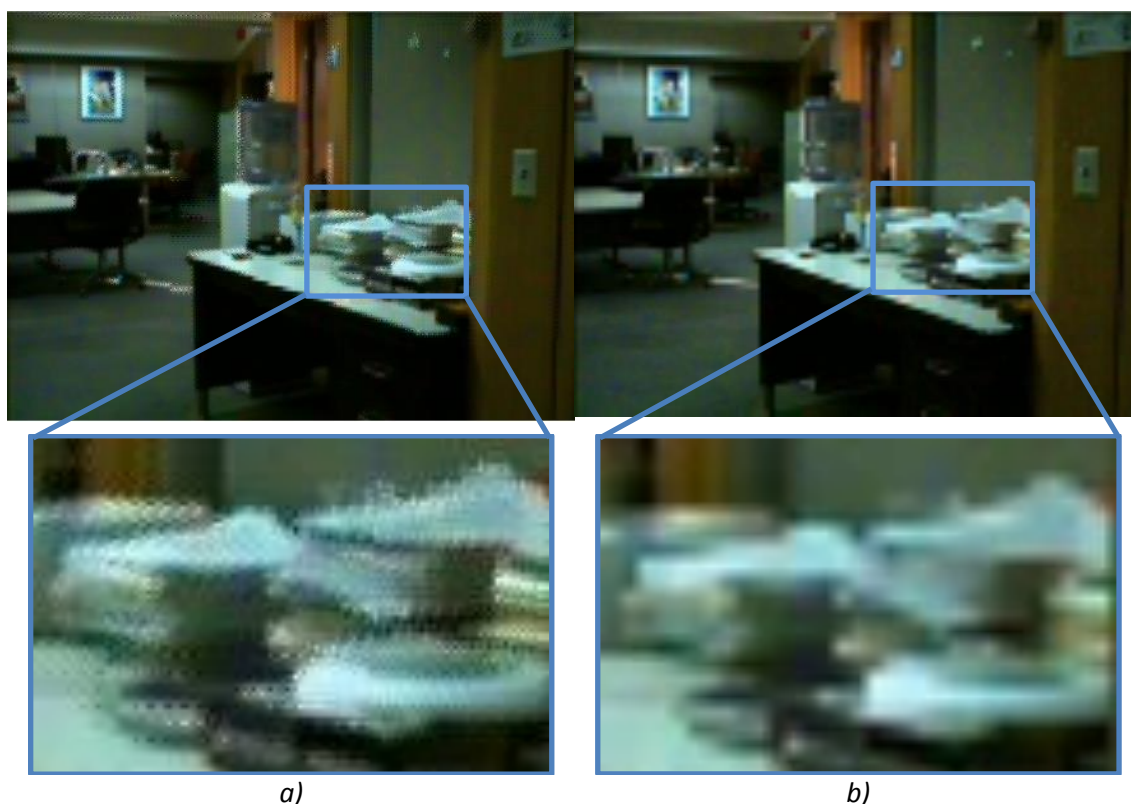


Figura 5.31.- Frames nº 10 súper-resuelto (a) e interpolado (b) HR para la secuencia *Water Cooler* en el método Mixto en modo Full-Frame

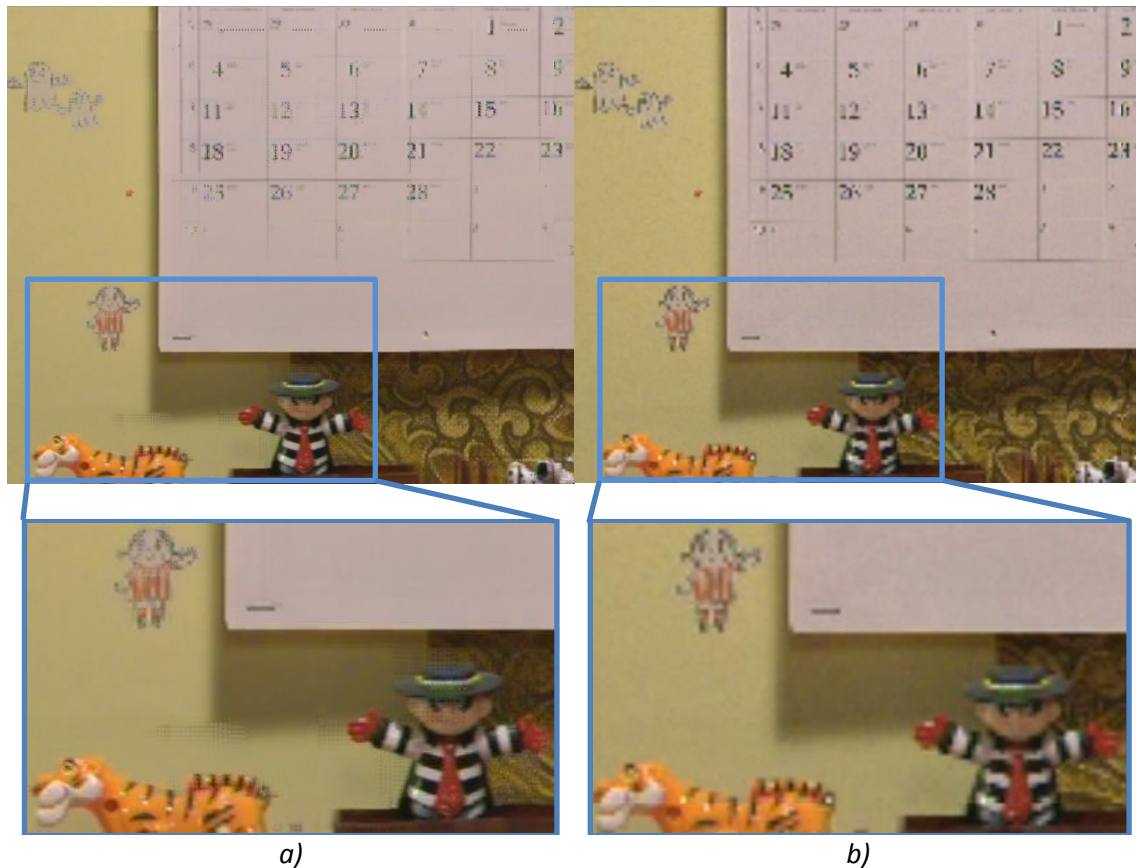


Figura 5.32.- Frames nº 42 súper-resuelto (a) e interpolado (b) HR para la secuencia *Mobcal* en el método Mixto en modo Full-Frame

En la Figura 5.32 se muestra una imagen de la secuencia *Mobcal*, en la que se puede observar cómo los números del calendario se encuentran mejor definidos en la secuencia súper-resuelta (Figura 5.32.a) que en la secuencia interpolada (Figura 5.32.b). Así mismo, se puede destacar cómo el muñeco de camisa de rayas y el tigre se encuentran mejor definidos en la secuencia súper-resuelta (Figura 5.32.a) que en la secuencia interpolada (Figura 5.32.b).

En cuanto a la calidad subjetiva de la secuencia *Stockholm*, se puede observar en la Figura 5.33.a una mayor definición en los tejados, así como las fachadas de todos los edificios respecto a la interpolada representada en la Figura 5.33.b, así como respecto a los métodos Temporal-Espacial y Espacial-Temporal. Así mismo, se observa cómo los coches de la esquina inferior izquierda presentan una mejor calidad. En esta secuencia, se muestra además una ampliación de los tejados donde se pueden apreciar mejor los buenos resultados obtenidos.

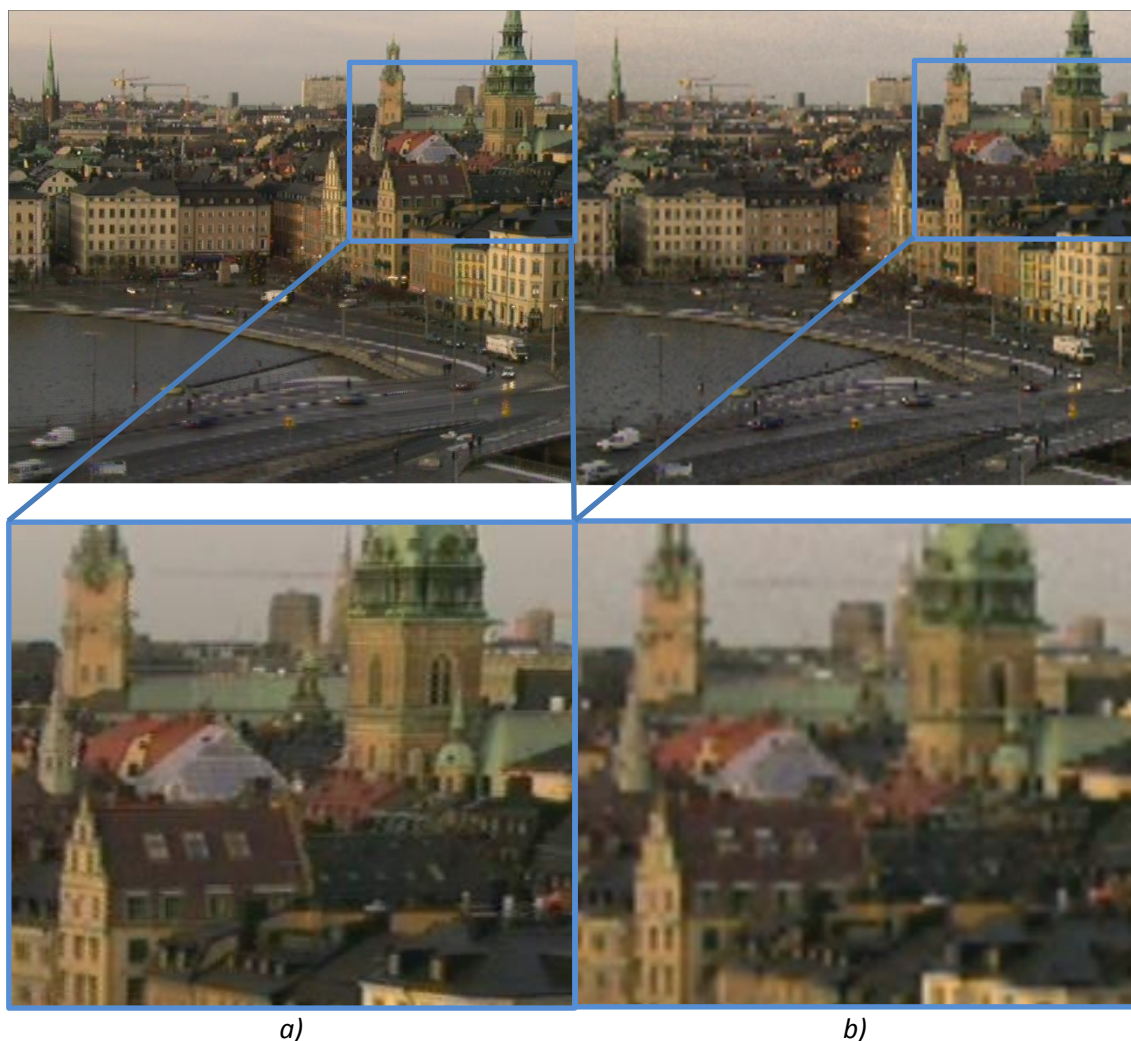


Figura 5.33.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia *Stockholm* en el método Mixto en modo *Full-Frame*

5.3.3.2 Modo de pre-procesamiento *Overlap*

A continuación se muestran las gráficas de PSNR y SSIM del método Mixto con el modo de pre-procesamiento *Overlap* para las secuencias *Water Cooler*, *Mobcal*, y *Stockholm*. Los resultados representados en la Figura 5.34 para la secuencia *Water Cooler* muestran cómo tras haberse descartado los bordes en esta secuencia, se logra que los resultados de la secuencia súper-resuelta se mantengan por encima de los correspondientes a la secuencia interpolada en la mayoría de casos, consiguiendo una mejora en la calidad respecto al modo *Full-Frame*. Por otro lado, en la Figura 5.35 y en la Figura 5.36 se observa cómo los resultados para las secuencias *Mobcal* y *Stockholm* se mantienen por encima de la secuencia interpolada, al igual que sucedía en el modo *Full-Frame*. Del mismo modo, ambas secuencias mantienen una amplia diferencia frente a la secuencia interpolada, especialmente en el caso de la secuencia *Stockholm*, ya que tal y como se ha comentado, la coherencia espacial es tal que a medida que se incrementa el número de *frames* disponibles en el proceso de SR se consigue maximizar la calidad objetiva sin introducir un número apreciables de artefactos, al ser el movimiento local en esta secuencia bajo.

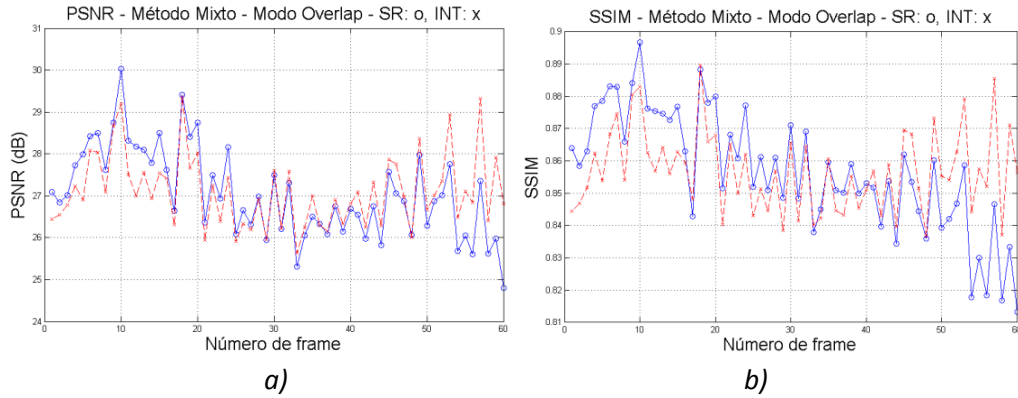


Figura 5.34.- PSNR y SSIM de método Mixto en modo *Overlap* en la secuencia *Water Cooler*

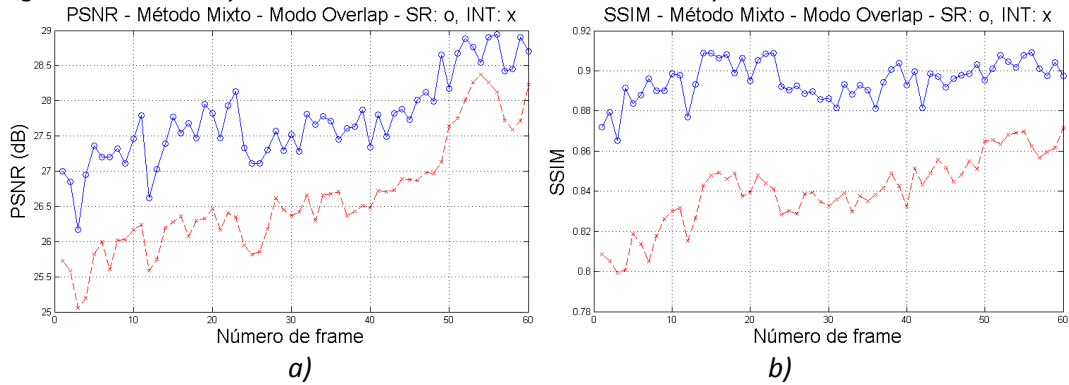


Figura 5.35.- PSNR y SSIM de método Mixto en modo *Overlap* en la secuencia *Mobcal*

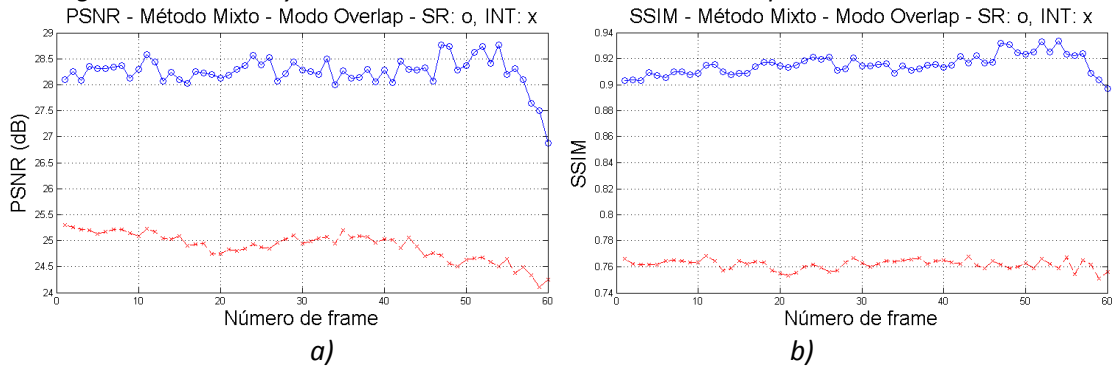


Figura 5.36.- PSNR y SSIM de método Mixto en modo *Overlap* en la secuencia *Stockholm*

En la Tabla 5.14, se muestran los valores medios de PSNR y SSIM obtenidos con el método Mixto, aplicando el modo de pre-procesamiento *Overlap* para las secuencias de *test* consideradas. En ella se puede observar, cómo los resultados para la secuencia súper-resuelta se encuentran por encima de los correspondientes a la secuencia interpolada en la mayoría de casos. Los resultados más significativos se corresponden con los de la secuencia *Stockholm*, en la que se obtiene una mejora de 3.35 dB en PSNR y 0.154 en SSIM, así como para la secuencia *Shields*, en la que en la que se obtiene una mejora de 1.86 dB en PSNR y 0.096 en SSIM.

Tabla 5.14.- Valores promedio de los resultados del método Mixto en modo *Overlap*

Secuencia	Método Mixto			
	PSNR (dB)		SSIM	
	SR	INT	SR	INT
<i>Water Cooler</i>	26.65	25.94	0.853	0.837
<i>Mobcal</i>	27.72	26.60	0.895	0.840
<i>Stockholm</i>	28.24	24.89	0.915	0.761
<i>Shields</i>	27.23	25.37	0.902	0.806
<i>Parkrun</i>	19.20	18.22	0.746	0.612

Finalmente, en la Figura 5.37 se muestra el *frame* 10 correspondiente a la secuencia *Water Cooler*, en la Figura 5.38 el *frame* 42 de la secuencia *Mobcal*, y en la Figura 5.39 el *frame* 12 de la secuencia *Stockholm*, de modo que se pueda realizar una valoración subjetiva de la calidad. En las imágenes mostradas a continuación, se puede observar la secuencia de salida HR, así como sus correspondientes *frames* interpolados.

En la Figura 5.37, en la que se muestra el *frame* 10 de la secuencia *Water Cooler*, se puede observar cómo los papeles de la mesa, así como el borde de ésta muestran una mayor definición en la secuencia súper-resuelta (Figura 5.37.a) sobre la interpolada (Figura 5.37.b).

Por otro lado, en la secuencia *Mobcal*, tal y como se muestra en la Figura 5.38.a, los números del calendario, así como la niña sobre el fondo amarillo, presentan una mejor calidad con respecto a la secuencia interpolada (Figura 5.38.b). En este caso no se ha podido analizar el muñeco al quedar fuera de la zona de solapamiento.

Finalmente, en la secuencia *Stockholm* se puede observar cómo la secuencia súper-resuelta, representada en la Figura 5.39.a, no sólo muestra una mejoría significativa en los tejados y fachadas de las casas, sino que también se observa una mejor calidad en el agua y en el puente que cruza el río.

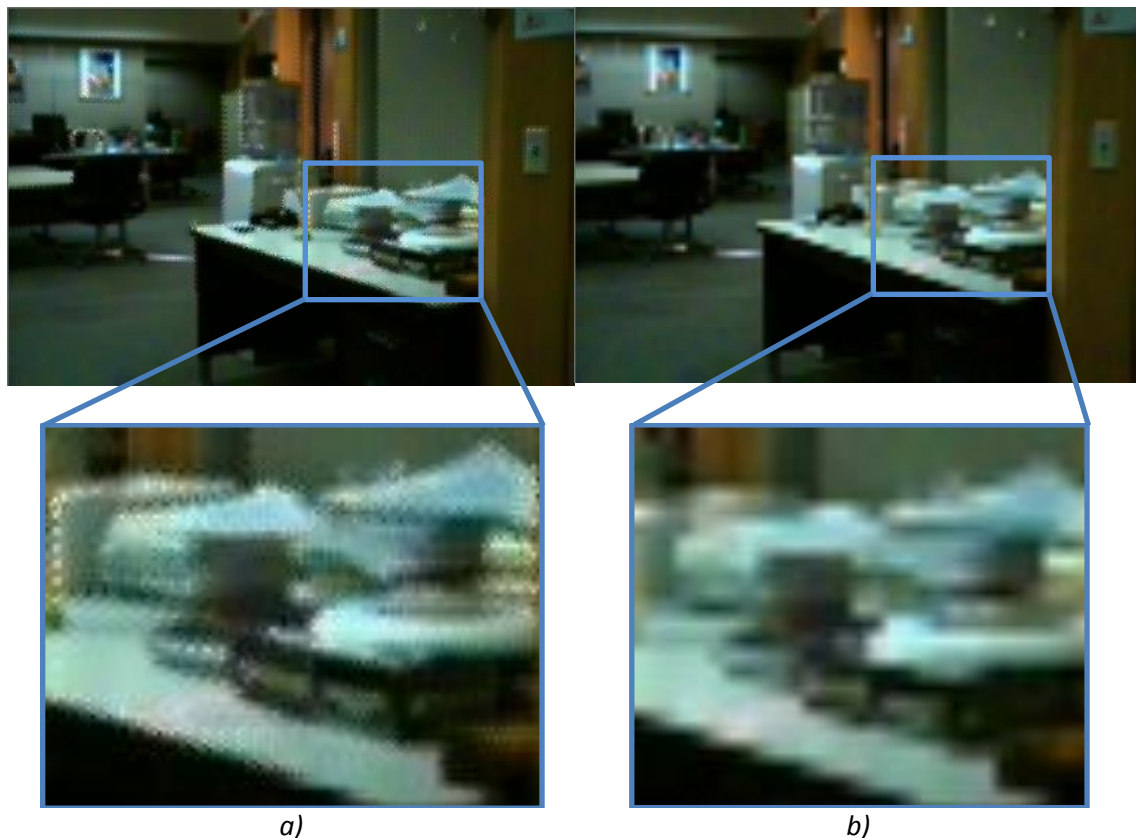


Figura 5.37.- Frames nº 10 súper-resuelto (a) e interpolado (b) HR para la secuencia *Water Cooler* en el método Mixto en modo *Overlap*

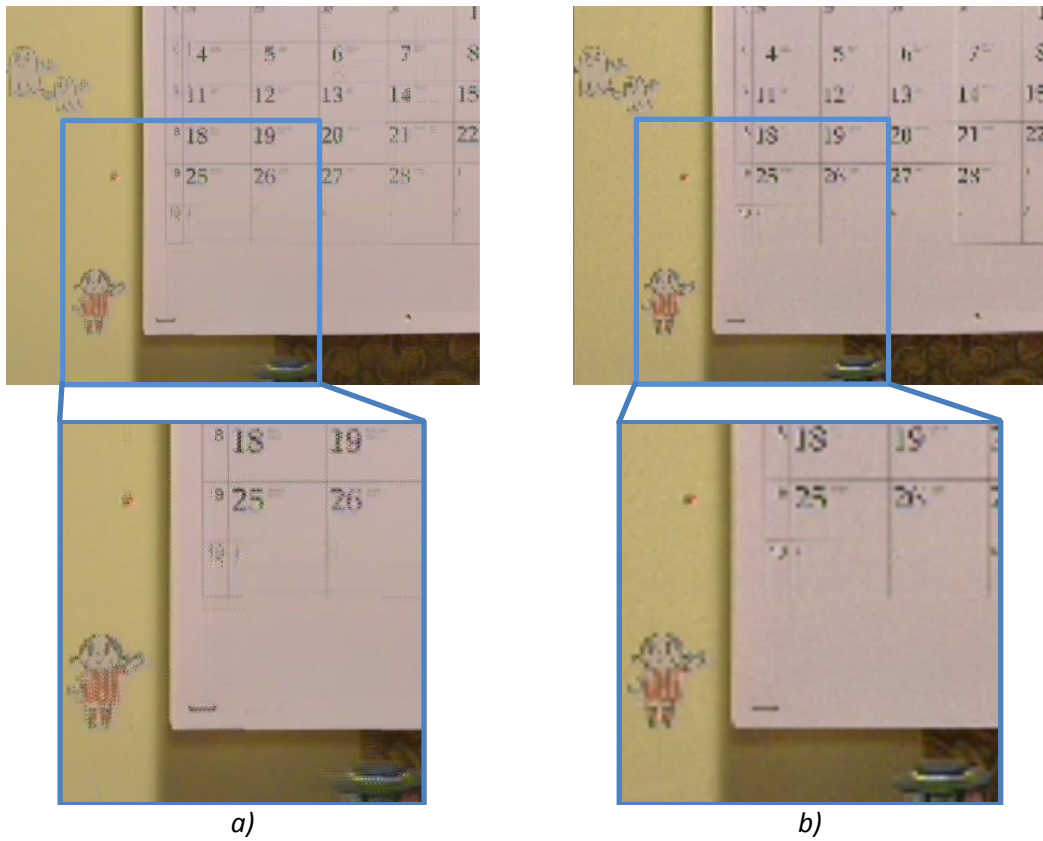


Figura 5.38.- Frames nº 42 súper-resuelto (a) e interpolado (b) HR para la secuencia Mobcal en el método Mixto en modo Overlap

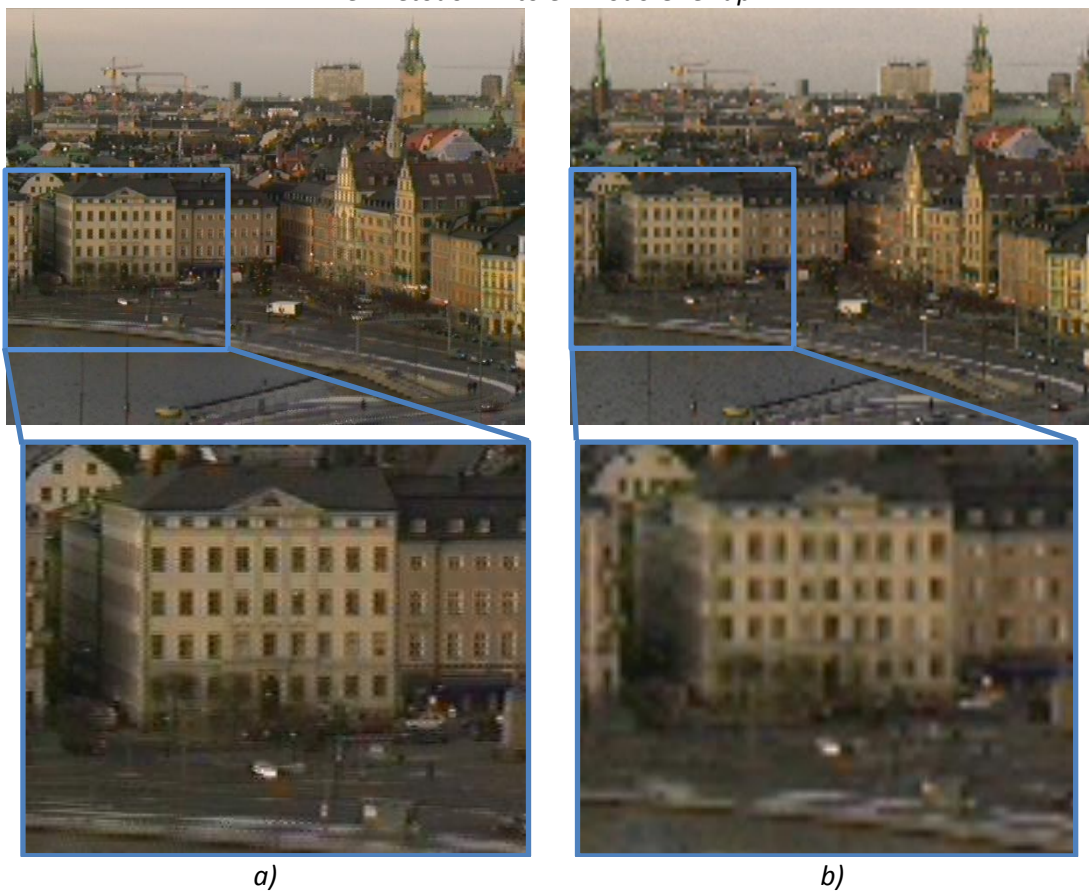


Figura 5.39.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia Stockholm en el método Mixto en modo Overlap

5.3.4 Súper-Resolución Espacial

En este apartado, se presentan los resultados obtenidos para cada una de las secuencias propuestas en la caracterización del proceso de SR Espacial. El interés se centra en representar un caso particular del método Mixto de SR MC en el cual no se tenga en cuenta la variable temporal; es decir un *array* MC que capta una misma escena en un mismo instante de tiempo. Con ello se podrá evaluar hasta qué punto la coherencia espacial entre *frames* aporta al proceso de SR sin considerar la variable temporal. En este caso se introduce además la secuencia *Fujidog*, ya que es una secuencia “espacial” (contiene *frames* tomados en un mismo instante de tiempo por diferentes cámaras de un *array* MC). La configuración empleada en cada secuencia, en aras de maximizar la calidad objetiva, se muestra en la Tabla 5.15.

Tabla 5.15.- Configuración establecida en la SR Espacial para cada secuencia

Secuencia	Método Espacial	
	SA (píxeles)	MBS (píxeles)
<i>Water Cooler</i>	16	16
<i>Fujidog</i>	8	32
<i>Mobcal</i>	8	32
<i>Stockholm</i>	8	32
<i>Shields</i>	8	32
<i>Parkrun</i>	8	32

A continuación, se mostrarán los valores medios de PSNR y SSIM resultantes de las simulaciones realizadas aplicando SR Espacial, así como *frames* de las secuencias resultantes con el fin de realizar una medida subjetiva de la calidad para cada los modos de pre-procesamiento *Full-Frame* y *Overlap*.

5.3.4.1 Modo de pre-procesamiento *Full-Frame*

El proceso de SR Espacial, al igual que el método Mixto, consta de un único proceso de SR, por lo que se presentan únicamente las curvas de PSNR y SSIM para las secuencias súper-resuelta (SR) y su versión interpolada equivalente (INTx4) como referencia para valorar los resultados. En la Tabla 5.16 se muestran los valores medios de PSNR y SSIM obtenidos aplicando únicamente SR Espacial con el modo de pre procesamiento *Full-Frame* para las secuencias de *test* consideradas.

En términos generales, se puede observar cómo la calidad de la SR Espacial supera a la secuencia interpolada, lográndose con este método unos buenos resultados. Destaca cómo se consigue 1.52 dB en PSNR y 0.079 en SSIM de mejora en la secuencia *Stockholm*, mientras en *Parkrun* se obtiene 1.66 dB en PSNR y 0.138 en SSIM de mejora sobre la secuencia interpolada, y en *Fujidog* 1.6 dB en PSNR y 0.06 en SSIM. Como se comentó al principio de esta sección, la secuencia *Fujidog* es realmente una foto tomada al mismo tiempo por cada una de las cámaras del *array* MC. Se presenta por tanto en la Figura 5.40 los resultados para el *frame* súper-resuelto de *Fujidog*, donde se aprecia cómo se han reconstruido mucho mejor los textos y símbolos de los libros apilados debajo del muñeco de peluche, así como los bordes de los marcos de los cuadros de colores, en el *frame* súper-resuelto (Figura 5.40.a) en comparación con el interpolado (Figura 5.40.b). En este caso, se ha realizado una magnificación de la parte de los libros, de forma que se puedan apreciar mejor los buenos resultados en la secuencia *Fujidog*.

Tabla 5.16.- Valores promedio de los resultados de la SR Espacial en modo Full-Frame

Secuencia	SR Espacial			
	PSNR (dB)		SSIM	
	SR	INT	SR	INT
<i>Water Cooler</i>	26.39	27.10	0.855	0.856
<i>Fujidog</i>	26.67	25.07	0.922	0.862
<i>Mobcal</i>	27.98	26.63	0.927	0.866
<i>Stockholm</i>	26.84	25.32	0.915	0.836
<i>Shields</i>	27.23	25.38	0.945	0.869
<i>Parkrun</i>	20.30	18.64	0.876	0.738

5.3.4.2 Modo de pre-procesamiento *Overlap*

En la Tabla 5.17 se muestran los valores medios de PSNR y SSIM obtenidos a partir del método de SR Espacial aplicando el modo de pre-procesamiento *Overlap* para las secuencias de *test* consideradas. Al igual que sucedía en el modo de pre-procesamiento *Full-Frame*, la calidad de la secuencia súper-resuelta se encuentra en valores medios por encima de la secuencia interpolada. Entre estos valores se pueden destacar los correspondientes a la secuencia *Stockholm*, obteniéndose una mejora de 1.25 dB en PSNR y 0.083 en SSIM, a *Parkrun* en la que se mejora 1.25 dB en PSNR y 0.122 en SSIM, cuyo *frame* 42 se representa en la Figura 5.41.

Tabla 5.17.- Valores promedio de los resultados de la SR Espacial en modo *Overlap*

Secuencia	SR Espacial			
	PSNR (dB)		SSIM	
	SR	INT	SR	INT
<i>Water Cooler</i>	25.51	25.94	0.847	0.837
<i>Mobcal</i>	27.11	26.59	0.875	0.840
<i>Stockholm</i>	26.29	25.04	0.845	0.762
<i>Shields</i>	26.84	25.33	0.886	0.806
<i>Parkrun</i>	19.47	18.22	0.734	0.612



Figura 5.40.- Frames súper-resuelto (a) e interpolado (b) HR para la secuencia *Fujidog* aplicando SR Espacial en modo *Full-Frame*

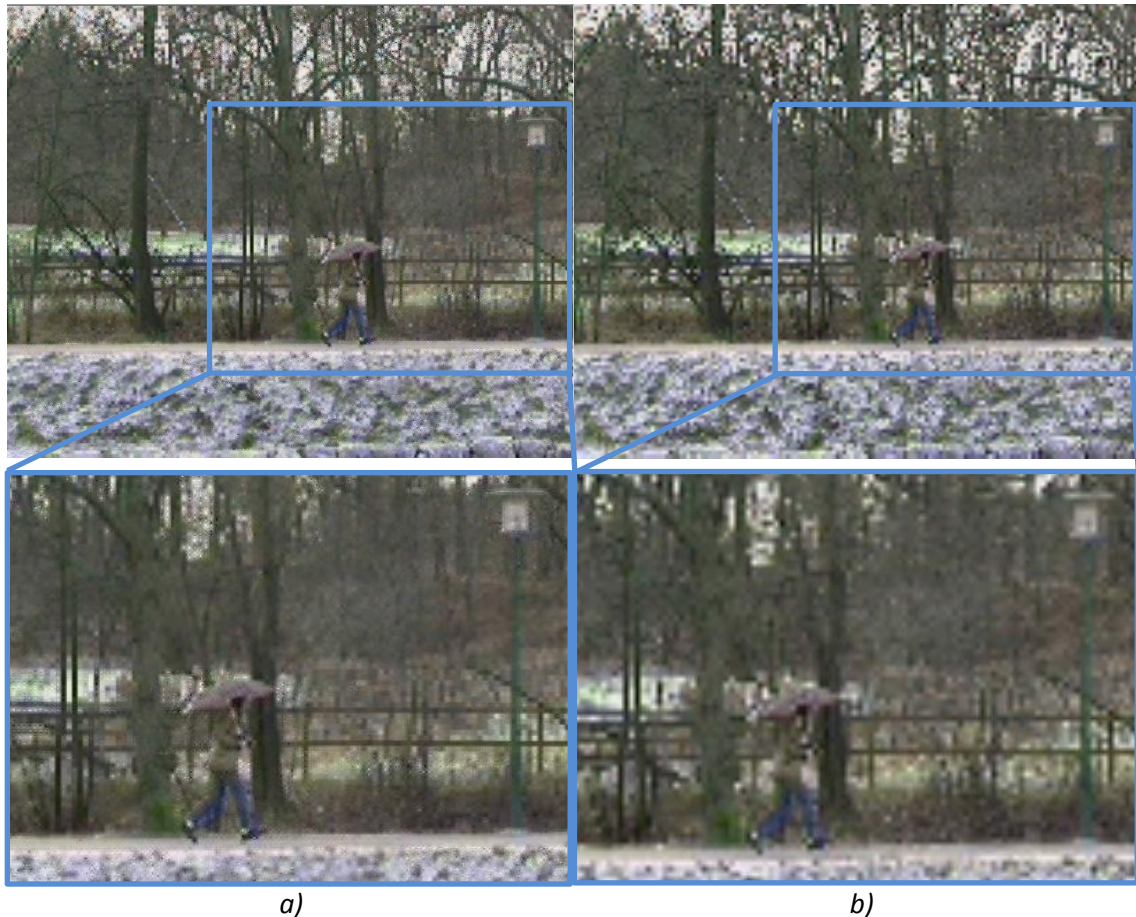


Figura 5.41.- Frames súper-resuelto (a) e interpolado (b) HR para el frame nº42 de la secuencia Parkrun aplicando SR Espacial en modo Overlap

5.3.5 Súper-Resolución Temporal - BSR

En este apartado, se analizan los resultados del algoritmo de referencia BSR, con el objetivo de compararlos posteriormente con los obtenidos a partir de los métodos propuestos, y de este modo poder estudiar la mejora que estos introducen sobre el algoritmo inicial. En la Tabla 5.18 se presenta la configuración establecida en el proceso de SR para cada secuencia maximizando la calidad objetiva, teniendo en cuenta una Ventana de Trabajo igual a la utilizada para el método Mixto. Se generó por tanto una simulación en modo *Full-Frame* para cada secuencia, la cual sirvió para comparar los resultados obtenidos en el resto de métodos con los modos de pre-procesamiento *Full-Frame* y *Overlap+Borders*, así como una simulación para cada secuencia en modo de pre-procesamiento *Overlap*, con el fin de contrastar con este mismo modo en el resto de métodos. Estos resultados, así como los mostrados en el proceso de SR Espacial, proporcionan una visión de los resultados esperables en los métodos de SR MC. Es decir, si los resultados de la SR Temporal o Espacial logran una buena mejora sobre la secuencia interpolada, existe una buena correlación temporal y espacial entre los *frames* de la secuencia (como es el caso de *Stockholm*), y por tanto esto va a aportar unas buenas condiciones a los métodos MC, resultando en una buena calidad en la secuencia HR. Por el contrario, si la secuencia no presenta una buena correlación temporal o espacial, como es el caso de *Mobcal*, la cual da lugar a unos buenos resultados en la SR Espacial, pero no tan buenos en la SR Temporal, habrá implicaciones en la calidad de los resultados en los métodos de SR MC.

Tabla 5.18.- Configuración establecida en el proceso de SR Temporal para cada secuencia

Secuencia	BSR – SR Temporal	
	SA (píxeles)	MBS (píxeles)
<i>Water Cooler</i>	16	16
<i>Mobcal</i>	20	32
<i>Stockholm</i>	20	32
<i>Shields</i>	20	16
<i>Parkrun</i>	20	32

5.3.5.1 Modo de pre-procesamiento *Full-Frame*

En la Tabla 5.19 se muestran los valores medios de PSNR y SSIM obtenidos en el método de SR Temporal para el modo de pre-procesamiento *Full-Frame* en las secuencias de *test* consideradas. Los valores mostrados se corresponden con la secuencia súper-resuelta y su correspondiente versión interpolada (INTx4). De entre estos valores medios cabe destacar el obtenido en la secuencia *Stockholm*, donde se logra una mejora de 0.9 dB en PSNR y 0.058 en SSIM sobre la secuencia interpolada.

Tabla 5.19.- Valores promedio de los resultados de la SR Temporal en modo *Full-Frame*

Secuencia	BSR – SR Temporal			
	PSNR (dB)		SSIM	
	SR	INT	SR	INT
<i>Water Cooler</i>	26.71	27.10	0.847	0.856
<i>Mobcal</i>	26.54	26.63	0.890	0.866
<i>Stockholm</i>	26.23	25.33	0.894	0.836
<i>Shields</i>	25.83	25.38	0.903	0.869
<i>Parkrun</i>	19.33	18.64	0.813	0.738

Se muestra en la Figura 5.42 el *frame* 10 correspondiente a la secuencia *Water Cooler*, en la Figura 5.43 el *frame* 42 de la secuencia *Mobcal*, y en la Figura 5.44 el *frame* 12 de la secuencia *Stockholm*, para realizar una valoración subjetiva de la calidad comparando la secuencia de salida HR, así como sus correspondientes *frames* interpolados.

El *frame* súper-resuelto de la secuencia *Water Cooler*, mostrado en la Figura 5.42.a, muestra una mejor definición en los papeles, así como en los objetos situados en la mesa que se encuentra debajo del cuadro, en comparación con el *frame* interpolado, presentado en la Figura 5.42.b.



Figura 5.42.- Frames nº 10 súper-resuelto (a) e interpolado (b) HR para la secuencia *Water Cooler* aplicando SR Temporal en modo *Full-Frame*

Por otra parte, en la Figura 5.43 se muestra el *frame* súper-resuelto (Figura 5.43.a) de la secuencia *Mobcal*, el cual obtiene una mejor reconstrucción de los números del calendario, donde se pueden destacar los números 11, 12, 14 y 21, en comparación con el *frame* de la secuencia interpolada (Figura 5.43.b). Sin embargo, a la izquierda del muñeco de rayas se puede observar una franja con muchos artefactos en la secuencia súper-resuelta.

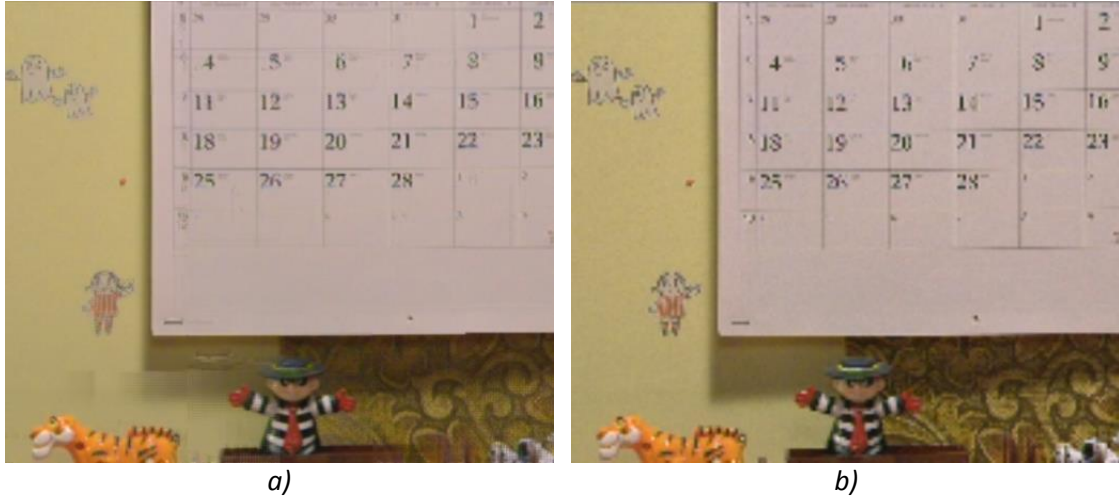


Figura 5.43.- Frames nº 42 súper-resuelto (a) e interpolado (b) HR para la secuencia *Mobcal* aplicando SR Temporal en modo Full-Frame

Finalmente, se puede apreciar una mejora en la secuencia *Stockholm*, mostrada en la Figura 5.44, sobre todo en las fachadas de las casas en la secuencia súper-resuelta (Figura 5.44.a) frente a la secuencia interpolada (Figura 5.44.b).

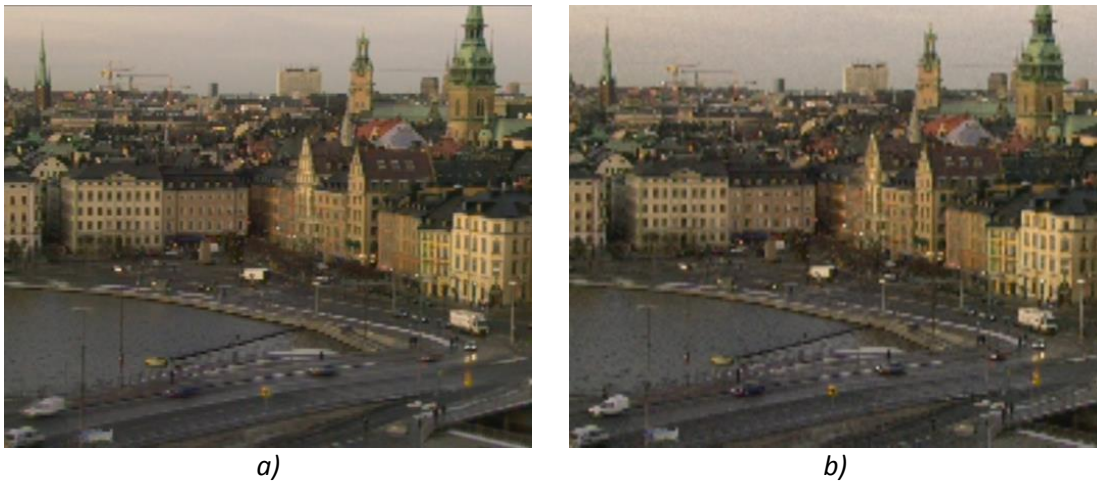


Figura 5.44.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia *Stockholm* aplicando SR Temporal en modo Full-Frame

5.3.5.2 Modo de procesamiento *Overlap*

En la Tabla 5.20 se muestran los valores medios de PSNR y SSIM obtenidos a partir del algoritmo BSR para las secuencias de *test*, aplicándolo a la parte correspondiente al solape entre los *frames* de las cámaras del *array* MC. Como se puede observar, los valores medios de la secuencia súper-resuelta se encuentran en la mayoría de los casos por encima de los correspondientes a la secuencia interpolada. Sin embargo, si se comparan los resultados del método BSR frente a los resultados obtenidos en el resto de métodos, no se logran unas diferencias tan significativas.

Tabla 5.20.- Valores promedio de los resultados de la SR Temporal en modo Overlap

Secuencia	BSR – SR Temporal			
	PSNR (dB)		SSIM	
	SR	INT	SR	INT
<i>Water Cooler</i>	26.00	25.94	0.836	0.837
<i>Mobcal</i>	26.45	26.60	0.851	0.840
<i>Stockholm</i>	25.71	24.89	0.827	0.761
<i>Shields</i>	25.70	25.37	0.847	0.806
<i>Parkrun</i>	18.61	18.22	0.658	0.612
<i>Reef</i>	25.00	24.55	0.817	0.785

En la Figura 5.45 se muestra el *frame* 10 correspondiente a la secuencia *Water Cooler*, en la Figura 5.46 el *frame* 42 de la secuencia *Mobcal*, y en la Figura 5.47 el *frame* 12 de la secuencia *Stockholm*, de modo que se pueda realizar una valoración subjetiva de la calidad. En las imágenes mostradas a continuación se puede observar la secuencia de salida HR, así como sus correspondientes *frames* interpolados.

En el *frame* súper-resuelto de la secuencia *Water Cooler*, mostrado en la Figura 5.45, hay partes de la mesa que no presentan una buena calidad en comparación con la secuencia interpolada, como la caja situada en la parte más próxima a la cámara o el borde de la mesa, pese a que los papeles se aprecian sobre la mesa con una mejor calidad.

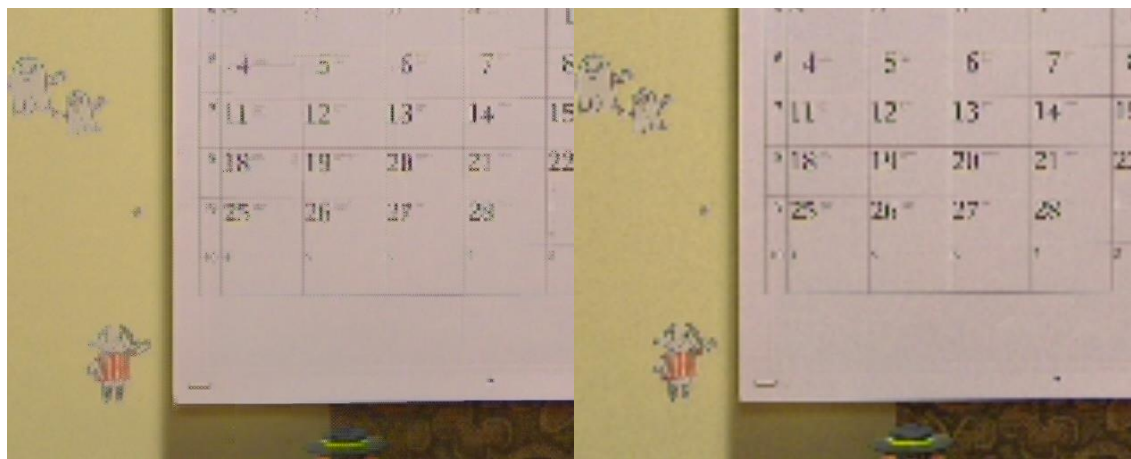


(a)

(b)

Figura 5.45.- Frames nº 10 súper-resuelto (a) e interpolado (b) HR para la secuencia *Water Cooler* aplicando SR Temporal en modo Overlap

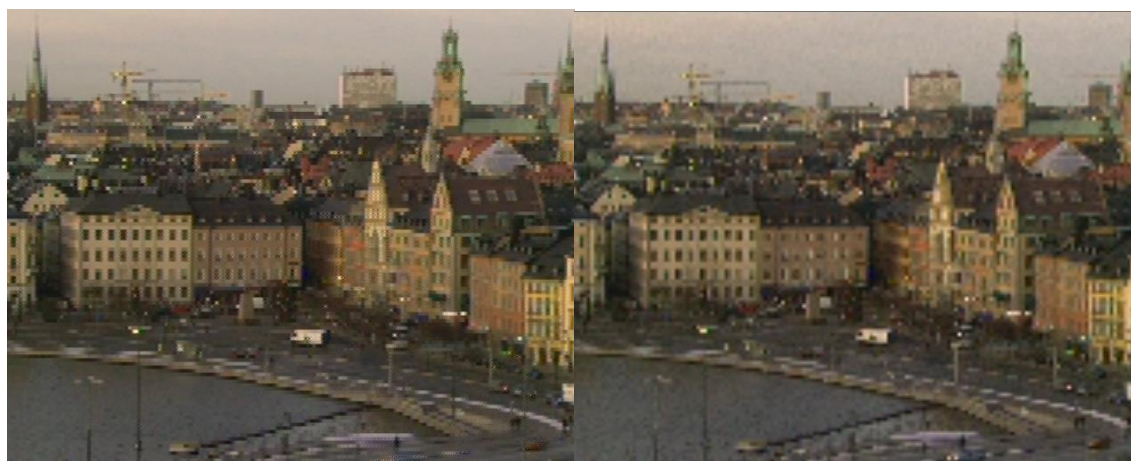
Por otro lado, en la Figura 5.46, correspondiente a la secuencia *Mobcal*, se aprecia cómo en la secuencia súper-resuelta (Figura 5.46.a) los números no se encuentran tan pixelados, del mismo modo que la niña se ve mejor definida, en comparación con la secuencia interpolada (Figura 5.46.b). Finalmente, el *frame* de la secuencia *Stockholm*, representado en la Figura 5.47, muestra una mejora en las fachadas de las casas, así como el agua del río, en la secuencia súper-resuelta (Figura 5.47.a) en comparación con la secuencia interpolada (Figura 5.47.b).



a)

b)

Figura 5.46.- Frames nº 42 súper-resuelto (a) e interpolado (b) HR para la secuencia Mobcal aplicando SR Temporal en modo Overlap



a)

b)

Figura 5.47.- Frames nº 12 súper-resuelto (a) e interpolado (b) HR para la secuencia Stockholm aplicando SR Temporal en modo Overlap

5.3.6 Comparativa de calidad de los métodos Multi-Cámara

En este apartado, se va a proceder a comparar los resultados obtenidos a partir de la aplicación de los distintos métodos presentados. Para ello, se reproducirán a modo de resumen algunas de las tablas de valores medios, como base para justificar las conclusiones que se establecen en este apartado. Así, se comenzará por incluir alguna comparativa que pueda resultar de interés en el análisis de los resultados intermedios y finales obtenidos a partir de los métodos Temporal-Espacial y Espacial-Temporal, prosiguiendo con una comparativa de todos los métodos presentados, donde para cada uno de ellos se estudiarán y se justificarán ciertos casos especiales que se puedan dar. Finalmente, se mostrarán y compararán determinados *frames* con el fin de valorar la calidad subjetiva de estos.

A modo de síntesis, teniendo en cuenta los comentarios que se han hecho a lo largo del capítulo en lo que respecta a los métodos Temporal-Espacial y Espacial-Temporal, se procede a analizar en mayor profundidad los resultados de ambos métodos, de modo que se obtenga una visión más concreta de lo que sucede en la segunda fase de estos métodos. Como se puede ver en la Tabla 5.21, los valores de PSNR y SSIM de la primera fase del método

Temporal-Espacial son inferiores a los que presenta el método Espacial-Temporal en su primera fase. Sin embargo, tal y como se observa en los resultados de la Tabla 5.22, sucede justamente lo contrario cuando se ejecuta el segundo proceso de SR, siendo en este caso el método de SR MC Temporal-Espacial el que proporciona unos mejores resultados frente al método Espacial-Temporal.

Tabla 5.21.- Resultados de la primera fase del método Temporal-Espacial (T-E) y el método Espacial-Temporal (E-T)

Secuencia	1ª Fase de SR			
	PSNR (dB)		SSIM	
	T-E	E-T	T-E	E-T
Water Cooler	27.43	27.74	0.897	0.903
Mobcal	27.00	27.70	0.814	0.834
Stockholm	26.90	26.73	0.792	0.779
Shields	26.51	27.05	0.815	0.829
Parkrun	19.64	20.00	0.660	0.675

Tabla 5.22.- Resultados de la segunda fase del método Temporal-Espacial (T-E) y el método Espacial-Temporal (E-T)

Secuencia	2ª Fase de SR			
	PSNR (dB)		SSIM	
	T-E	E-T	T-E	E-T
Water Cooler	27.45	27.52	0.861	0.861
Mobcal	26.89	26.77	0.887	0.886
Stockholm	26.96	26.53	0.904	0.897
Shields	26.88	26.51	0.928	0.922
Parkrun	19.83	19.57	0.828	0.822

Los motivos que propician esta situación en ambas fases, se deben a que en la primera fase ambos métodos parten de una secuencia sin artefactos, no introduciendo errores en el proceso de *Estimación de Movimiento*, mientras que en la segunda fase normalmente se parte de una secuencia con artefactos. Por ello, cuando se realiza la fase de *Estimación de Movimiento* en el núcleo del algoritmo base de SR, si el MB para el que se realiza esta estimación contiene artefactos, al buscar un MB similar a él, se estará buscando un MB que no tiene por qué parecerse al MB deseado para reconstruir la secuencia original, por lo que puede seleccionarse un MB erróneo debido a estos artefactos. Por otro lado, aunque seleccione el MB más adecuado para reconstruir la secuencia, el MB que se va a incorporar en la imagen HR puede contener artefactos, y por tanto, estar introduciendo artefactos generados en el primer proceso de SR, en la segunda fase, combinados con los que ya contiene el *frame* a súper-resolver. En este punto cabe pensar que sería adecuado aplicar el Filtro Selectivo de MB desarrollado en esta Tesis Doctoral. En resumen, si se generan artefactos en la primera fase, estos van a provocar errores en la *Estimación de Movimiento*, afectando directamente a la calidad de la secuencia HR. Así mismo, un MB seleccionado como el mejor candidato para reconstruir la secuencia deseada, puede contener artefactos que va a introducir en el resultado de la segunda fase.

Por tanto, se considera que el proceso de SR Espacial es más robusto que el proceso de SR Temporal, puesto que éste, partiendo de una secuencia de MR_t cuya calidad es inferior a la calidad de la secuencia de MR_s , es capaz de proporcionar una secuencia de salida con mayor calidad. El proceso de SR Temporal del método Espacial-Temporal, pese a tener una entrada con mayor calidad que el método Temporal-Espacial, degrada la calidad de la secuencia HR, quedando por debajo de la del método Temporal-Espacial, donde además hay que añadir que ambas secuencias contienen artefactos derivados del primer proceso de SR. Este análisis explica también lo que produce que el proceso de SR Temporal proporcione resultados próximos, o incluso inferiores a los que se obtienen en la secuencia SR+INT para el método Espacial-Temporal.

A partir de estos resultados, se puede ver que existe una dependencia del orden en el que se procese la información aportada por el *array* MC, dependiendo de las características de la secuencia que se va a procesar, pudiendo ser más apropiado usar el método Temporal-Espacial o el Espacial-Temporal para obtener mejores resultados.

Una vez estudiados los detalles relativos a los métodos Temporal-Espacial y Espacial-Temporal, se procede a realizar una comparativa de todos los métodos entre sí, atendiendo únicamente a la calidad objetiva de la secuencia final HR que proporcionan, que es el objetivo principal de estos métodos.

La Tabla 5.23 muestra una comparativa de los valores medios de PSNR y SSIM de todos los métodos de SR propuestos en esta Tesis Doctoral, para cada una de las secuencias de *test* utilizadas en relación con los obtenidos para BSR e INT. A través de estos valores, se puede establecer que el método que mejores resultados aporta de forma general es el método Mixto, estando en la mayoría de casos por encima de los métodos Temporal-Espacial y Espacial-Temporal, así como del proceso de SR Espacial, del algoritmo BSR y de la secuencia interpolada, mejorando en muchos casos aproximadamente entre 1 y 2 dB en PSNR y 0.5 en SSIM de media sobre el resto de métodos de SR MC. Además se presentan en la Tabla 5.24 los valores máximos de los métodos Mixto, Temporal-Espacial y Espacial-Temporal, que justifican aún más el buen funcionamiento del método de SR MC Mixto. En segundo lugar, se podría considerar como una buena opción aplicar SR Espacial, puesto que es la que después del método Mixto proporciona una secuencia de mayor calidad, si bien este método, pese a ser buen candidato a considerar, no combina la información espacial con la información temporal que aporta un *array* MC.

Tabla 5.23.- Comparativa de los valores medios de los resultados entre los distintos métodos

Secuencia	Métrica	Temporal-Espacial	Espacial-Temporal	Mixto	SR Espacial	BSR	INT
Water Cooler	PSNR (dB)	27.45	27.52	27.05	26.39	26.71	27.10
	SSIM	0.861	0.861	0.856	0.855	0.847	0.856
Mobcal	PSNR (dB)	26.89	26.77	27.38	27.98	26.54	26.63
	SSIM	0.887	0.886	0.919	0.927	0.890	0.866
Stockholm	PSNR (dB)	26.96	26.53	28.34	26.97	26.23	25.33
	SSIM	0.904	0.897	0.942	0.914	0.894	0.836
Shields	PSNR (dB)	26.88	26.51	27.58	27.23	25.83	25.38
	SSIM	0.928	0.922	0.951	0.945	0.903	0.869
Parkrun	PSNR (dB)	19.83	19.57	20.48	20.30	19.33	18.64
	SSIM	0.828	0.822	0.902	0.876	0.813	0.738

Tabla 5.24.- Valores máximos de métodos Mixto, Temporal-Espacial y Espacial-Temporal

Secuencia	Métrica	Temporal-Espacial	Espacial-Temporal	Mixto	SR Espacial	BSR	INT
Water Cooler	PSNR (dB)	29.95	29.97	30.03	28.41	29.8	29.35
	SSIM	0.893	0.895	0.897	0.886	0.894	0.889
Mobcal	PSNR (dB)	28.18	27.83	28.48	29.37	27.47	27.74
	SSIM	0.907	0.905	0.935	0.946	0.911	0.897
Stockholm	PSNR (dB)	27.24	26.82	28.86	27.53	26.56	25.64
	SSIM	0.911	0.904	0.953	0.921	0.902	0.843
Shields	PSNR (dB)	27.79	26.95	28.39	27.94	26.3	25.82
	SSIM	0.944	0.928	0.958	0.951	0.915	0.88
Parkrun	PSNR (dB)	20.38	20	21.03	21.01	19.74	19.16
	SSIM	0.845	0.835	0.917	0.899	0.831	0.76

5.3.7 Comparativa de tiempos de cómputo de los métodos Multi-Cámara

Se presentan en este apartado los tiempos de cómputo de los métodos de SR MC introducidos en la presente Tesis Doctoral, obtenidos a partir de un PC de sobremesa con un procesador *core 2 quad* a 3,0 GHz con sus recursos al completo. Así, en la Tabla 5.25 se presenta el tiempo medio en procesar un *frame* en cada fase de cada uno de los métodos de

SR MC. En estos valores se puede apreciar cierta uniformidad en los resultados en las secuencias *Mobcal*, *Stockholm*, *Shields* y *Parkrun*, hecho que se debe a que comparten iguales dimensiones en sus *frames*, existiendo aun así pequeñas diferencias de tiempo en cada uno de los métodos. Estas diferencias se deben a que la etapa de *Rellenado de Huecos* del proceso de SR requiere un mayor tiempo de procesamiento en función de la cantidad de píxeles que se hayan rellenado en la fase de *Compensación de Movimiento*, lo cual también se puede observar en los tiempos de cómputo a nivel de MB, mostrados en la Tabla 5.26. Tomando como ejemplo la secuencia *Shields*, como se puede ver en la Tabla 5.25, en el método Temporal-Espacial el tiempo medio en procesar cada cámara en la SR Temporal es de 19.25 s, por lo que si se realiza la SR Temporal de cada cámara de forma secuencial o iterativa considerando las 9 cámaras del *array* MC, se obtiene que el tiempo total para este método se corresponde con 173.25 s, a lo cual hay que añadir el tiempo de cómputo de la segunda fase de este método, que es de 20.75 s para la realización de la fase de SR Espacial, implicando un tiempo total de 194 s, lo que supera con creces el tiempo de cómputo estimado del resto de los métodos.

Tabla 5.25.- Tiempo medio de procesamiento de un frame (métodos MC, BSR e Interpolación)

Método	Fase	Tiempo por secuencia y método (s)				
		<i>Water Cooler</i>	<i>Mobcal</i>	<i>Stockholm</i>	<i>Shields</i>	<i>Parkrun</i>
<i>Temporal-Espacial</i> <i>WW = ±7</i>	Temporal	3.8	23.61	19.1	19.25	19.2
	Espacial	23.9	26.76	20.91	20.75	20.9
	Paralelización	27.7	50.37	40.01	40	40.1
	Total	58.1	239.25	192.81	194	193.7
<i>Espacial-Temporal</i> <i>WW = ±7</i>	Espacial	2.31	4.8	3.66	3.71	3.85
	Temporal	48.89	194.1	153	157	162.15
	Total	51.2	198.9	156.66	160.71	166
<i>Mixto - WW = ±31</i>	Total	5.71	88.33	73.05	69.65	74.44
<i>BSR - WW = ±31</i>	Total	4.65	26.56	21.23	22.03	22.01
<i>Interpolación</i>	Total	1.8	1.08	2.76	3.04	2.86

Sin embargo, si se paralelizaran los nueve procesos de SR, uno por cada cámara del *array* que se ha usado en esta simulación, el tiempo total de cómputo sería de $19.25 + 20.75 = 40$ s, disminuyendo en casi 5 veces el tiempo de cómputo y llegando a ser el método de SR MC más rápido con diferencia, procesando cada *frame* prácticamente en la mitad de tiempo que el método Mixto en la mayoría de casos, aunque hay que resaltar que el método Mixto está trabajando con una Ventana de Trabajo de ± 31 *frames*, que es el tamaño que mejores resultados ha proporcionado en la mayoría de casos. Sin embargo, esta solución conlleva que para conseguir que el método Temporal-Espacial se ejecute en este tiempo, es necesario incrementar considerablemente los requerimientos *hardware* del sistema, teniendo que disponer de 9 *cores/threads* para la SR Temporal, y uno más para la SR Espacial, en total 10 *cores/threads*, tal y como se presenta en la Figura 5.48, para conseguir reducir el tiempo de cómputo lo máximo posible. Finalmente, este tiempo de cómputo está sujeto al tamaño de la Ventana de Trabajo usada, donde el tiempo de cómputo se incrementará a medida que se aumente el tamaño de la Ventana de Trabajo del proceso de SR Temporal, pudiendo no ser éste el método más rápido.

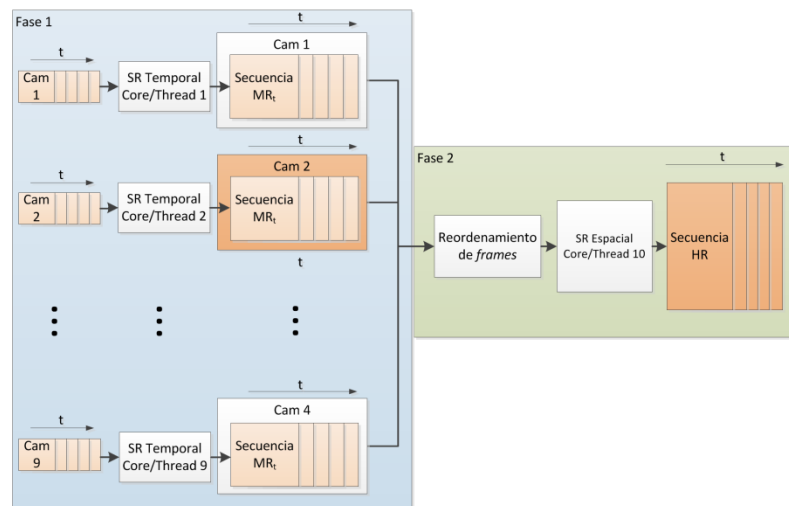


Figura 5.48.- Modo Temporal-Espacial paralelo con 10 cores/threads

Por otra parte, el método Espacial-Temporal requiere para su implementación de dos procesos de SR, por lo que el tiempo total será la suma de los tiempos de cómputo de cada fase. En este caso, como se muestra en la Tabla 5.25, el tiempo necesario para llevar a cabo el proceso de SR Espacial es de 3.71 s, y tarda unos 157 s en procesar el *frame* de salida HR en la segunda fase, donde se realiza un proceso de SR Temporal. Por ello, el tiempo total para procesar un *frame* con este método es de 160.71 s, siendo este tiempo más del cuádruple que el tiempo de cómputo del método Temporal-Espacial en caso de paralelizar los nueve procesos de SR Temporal de la primera fase, y siendo menor que si no se paralelizan estos métodos. Para la implementación de este método se requiere de al menos 2 *cores/threads*, para poder ejecutar ambos procesos a la vez y conseguir obtener los *frames* de salida de forma óptima. Hay que destacar que tanto en el método Temporal-Espacial, como en el Espacial-Temporal, el tiempo de cómputo dependerá del tamaño de la Ventana de Trabajo, el cual en este caso es de ± 7 *frames*. Finalmente, el método Mixto para una Ventana de Trabajo de ± 31 *frames* consigue procesar un *frame* de la secuencia en 69.65 s, como se puede observar en la Tabla 5.25. Por un lado, considerando que las Ventanas de Trabajo se han incrementado considerablemente respecto al resto de métodos, y teniendo en cuenta los resultados de calidad obtenidos en las secuencias artificiales, así como que para el método Mixto solo es necesario el uso de un *core/thread* para obtener estos resultados, hace que se considere como el método que mejores prestaciones presenta, ya que la relación entre requerimientos *hardware* y tiempos de cómputo, es más equilibrada respecto a la mejora de calidad que este método proporciona.

Tabla 5.26.- Tiempo medio de procesamiento de un MB (métodos MC, BSR e Interpolación)

Método	Fase	Tiempo por secuencia y método (ms)				
		Water Cooler	Mobcal	Stockholm	Shields	Parkrun
Temporal-Espacial	Temporal	108	196	159	160.4	160
	Espacial	42	223	174	172	174
	Paralelización	150	419	333	332.4	334
	Total	1014	1987	1605	1615.6	1614
Espacial-Temporal	Espacial	16	40	30.5	30.9	32
	Temporal	1396	1617	1275	1308	1351
	Total	1412	1657	1305.5	1338.9	1383
Mixto	Total	163	736	608	580	620
BSR	Total	132	221	176	183.5	183.4
Interpolación	Total	15.86	9.06	23.57	25.33	23.83

En cuanto a los resultados de tiempo de cómputo obtenidos para la secuencia *Water Cooler*, mostrados en la Tabla 5.25, el correspondiente del método Mixto es mucho menor debido al número de cámaras del *array* MC, el cual influye directamente en el tamaño de la Ventana de Trabajo. En este caso, el método Temporal-Espacial es más rápido que el Espacial-Temporal, aunque es cinco veces más lento que el método Mixto.

En la Tabla 5.26 se muestra el tiempo medio de cómputo que se ha empleado en procesar cada MB en un *frame* de las secuencias de *test*. En esta tabla, se observa cómo el tiempo en procesar un MB en el método Temporal-Espacial es menor que en el resto de métodos cuando se realiza paralelización, presentando el método Mixto tiempos de cómputo inferiores al método Espacial-Temporal. El hecho de que los tiempos de cómputo del método Mixto sean superiores a los del método Temporal-Espacial se debe a las dimensiones que se han establecido en la Ventana de Trabajo Mixta, siendo ésta de ± 31 *frames*, frente a los ± 7 *frames* que tiene la Ventana de Trabajo del proceso de SR Temporal en el método Temporal-Espacial.

5.3.8 Figura de Mérito de los métodos Multi-Cámara (calidad VS tiempo)

Los métodos de SR MC propuestos conllevan un mayor tiempo de cómputo que el algoritmo BSR, lo cual es normal puesto que añaden más información a procesar para obtener la secuencia súper-resuelta. En este apartado, se valorará hasta qué punto este incremento de cómputo se compensa con la mejora en la calidad resultante que se obtiene en estos.

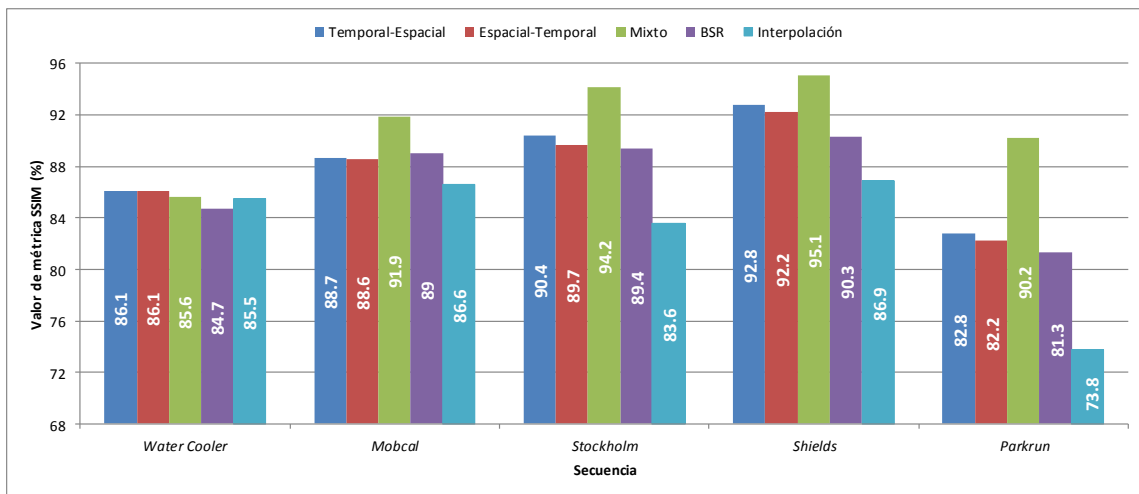


Figura 5.49.- Métrica SSIM para métodos MC y BSR en secuencias bajo estudio

A fin de poder comparar los diferentes métodos desde un punto de vista de la calidad objetiva y el tiempo de cómputo, se establecen figuras de mérito adecuadas partiendo de los resultados obtenidos, que se presentan en la Figura 5.49 a partir de la Tabla 5.24 y de la Tabla 5.26, que muestra la métrica de calidad objetiva SSIM (que se ha tomado como referencia al ser una métrica definida entre 0 y 1) y en la Figura 5.50, que presenta un resumen de los tiempos de cómputo. Se propone por tanto una figura de mérito FM_MC que relaciona directamente la calidad objetiva de la imagen con el tiempo de cómputo del método utilizado en *ms*, considerando además el número de *cores/threads* utilizados. Esta métrica, se aplica para comparar los métodos Temporal-Espacial (simple y paralelo), Espacial-Temporal y Mixto, tal y como se presenta en la Figura 5.51. La expresión aplicable a la métrica se introduce en la expresión (5.1).

$$FM_{MC} = \frac{SSIM(Sequencia) + \{1 - [Tiempo(Sequencia)/Tiempo_{max}]\}}{2 \cdot \sqrt{N_{ct}}}, \quad (5.1)$$

donde $Tiempo_{max}$ es el tiempo máximo utilizado para procesar una secuencia determinada y N_{ct} es el número de cores/threads utilizados (normalmente corresponde al método Temporal-Espacial).

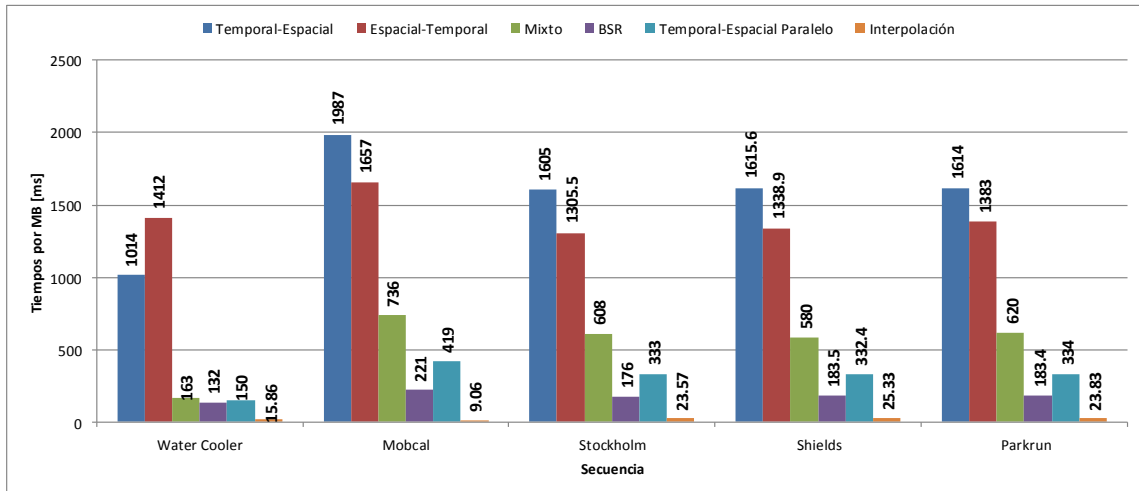


Figura 5.50.- Resumen de tiempos de cómputo (en ms) de secuencias bajo estudio

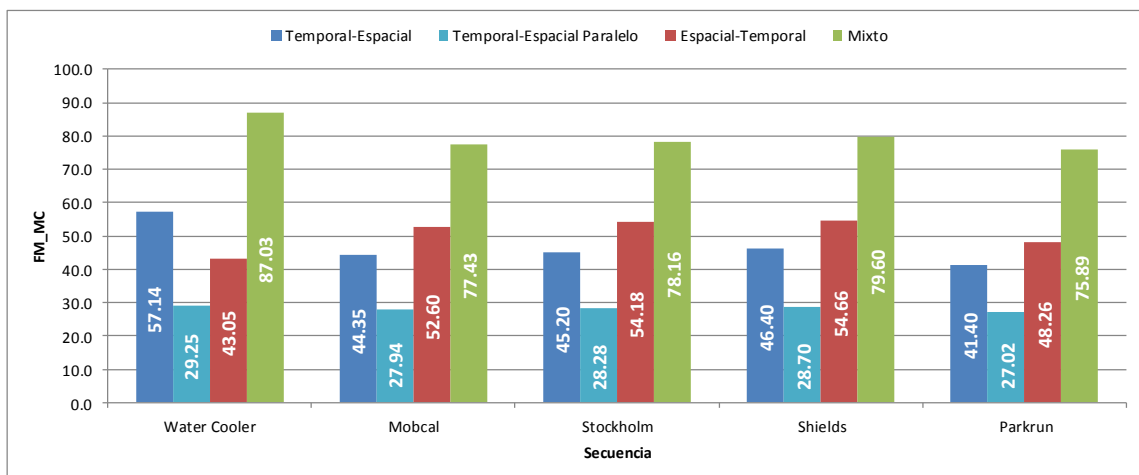


Figura 5.51.- Figura de Mérito FM_{MC} aplicada a las secuencias bajo estudio

De la Figura 5.51 se concluye que tras haber analizado todo los métodos MC en detalle, el método Mixto es el más adecuado para establecer un compromiso apropiado entre calidad objetiva, tiempo de cómputo y recursos *hardware*.

5.4 CONCLUSIONES

En este capítulo, se ha presentado el entorno de caracterización y los resultados más significativos de las contribuciones de la presente Tesis Doctoral con respecto a la aplicación de Súper-Resolución en entornos Multi-Cámara. A partir del estudio llevado a cabo en este capítulo, se concluye que la coherencia espaciotemporal que proporcionan los sistemas Multi-Cámara a la Súper-Resolución es muy adecuada.

Habiendo investigado los tres métodos propuestos: Temporal-Espacial, Espacial-Temporal y Mixto (además de los casos particulares de SR Espacial y SR Temporal), se ha concluido que es el método de SR MC Mixto el que proporciona mejores resultados en términos de calidad objetiva, lo cual se ve reforzado por el buen compromiso que establece este método con respecto al tiempo de cómputo. Los otros dos métodos: Temporal-Espacial y Espacial-Temporal, presentan un peor compromiso entre calidad objetiva y tiempo de cómputo, si bien el método Temporal-Espacial proporciona buenos resultados cuando se paraleliza haciendo uso de varios *cores/threads*, a costa de un mayor uso de recursos *hardware*.

Con respecto a los modos de pre-procesamiento, los que mejores resultados proporcionan son *Full-Frame* y *Overlap*, siendo en la parte solapada donde se encuentra la mayor cantidad de información para reconstruir la secuencia.

Capítulo 6

ANÁLISIS DEL SISTEMA GLOBAL

Índice del capítulo

	<i>Página</i>
6.1.- Introducción	178
6.2.- Resultados de Sistema Global	179
6.3.- Aplicación en entorno submarino	191
6.4.- <i>Test</i> de calidad subjetiva	202
6.5.- Conclusiones	204

*Cualquier poder, si no se basa en la unión, es débil
Jean de la Fontaine, fabulista francés (1621 – 1695)*

6.1 INTRODUCCIÓN

Partiendo de las contribuciones de la presente Tesis Doctoral: Filtro Selectivo de Ventana, Filtro Selectivo de Macro-Bloque, topología de Macro-Bloques adaptable y aplicación en entorno Multi-Cámara para Súper-Resolución, y considerando la gran cantidad de información que se maneja en un sistema Multi-Cámara, se planteó un sistema completo que filtrara la información que fuera menos apropiada para incluir en el proceso de Súper-Resolución, generando un sistema completo. En este capítulo se introducen los resultados que se obtienen al caracterizar el sistema global propuesto en esta Tesis Doctoral.

A partir de las conclusiones extraídas de los capítulos anteriores, el sistema se constituye a partir de la combinación del método de Súper-Resolución Multi-Cámara Mixto con modo de pre-procesamiento *Full Frame* y una estructura de filtro de Súper-Resolución a dos niveles: un primer nivel que selecciona los *frames* que constituyen la Ventana de Trabajo de Súper-Resolución (*Working Window*) y un segundo nivel que determina los Macro-Bloques a los que se aplica el proceso de Súper-Resolución dentro de la Ventana de Trabajo. Este filtro conjunto se denomina WBSFSR (*Window & Block Selective Filter Super-Resolution*). El sistema global propuesto en esta Tesis Doctoral, considerando estas particularidades, se presenta en la Figura 6.1.

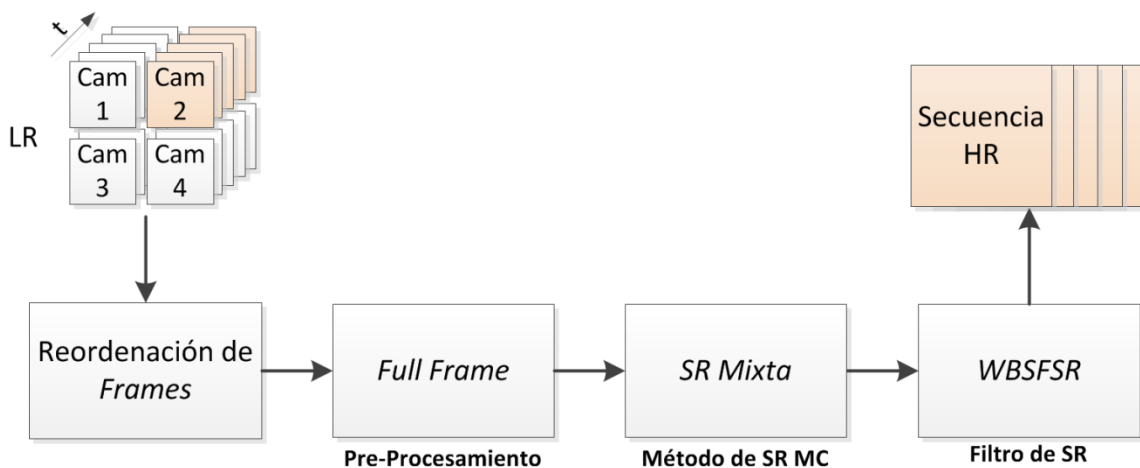


Figura 6.1.- Sistema global propuesto en esta Tesis Doctoral

6.2 RESULTADOS DE SISTEMA GLOBAL

6.2.1 Resultados de calidad y tiempo de cómputo del sistema global

En esta sección, se muestran los resultados en términos de calidad y tiempo de cómputo del sistema global propuesto para las secuencias de *test* consideradas en el Anexo A.1 de la presente Tesis Doctoral.

Las Figuras 6.2, 6.3, 6.4 y 6.5 muestran las mejoras introducidas por un sistema que combina las aportaciones de los filtros selectivos, identificándose esta aportación como FSR (*Filtered Super-Resolution*), conjuntamente con la aportación MC utilizando el método Mixto con modo de pre-procesamiento *Full Frame*, ya que es el que en general presenta mejor relación calidad / tiempo de cómputo, como se ha demostrado en el capítulo anterior, y que se identifica como MCSR (*Multi-Camera Super-Resolution*). Por último, la combinación de ambas aportaciones se representará como MCABMSR (*Multi-Camera Adaptive Block-Matching Super-Resolution*) [C8], que como se verá mantendrá la calidad de MCSR o incluso la superará, además de reducir el tiempo de cómputo. Para los análisis realizados se ha hecho uso de una Ventana de Trabajo de SR para BSR de ± 10 , y una Ventana de Trabajo temporal de MC de ± 3 , que son valores adecuados para obtener en promedio una calidad objetiva óptima para las secuencias estudiadas.

Las Figuras 6.2 y 6.3 muestran cómo el sistema global planteado presenta los mejores resultados de calidad objetiva, tanto en términos de PSNR como de SSIM, incluso en la secuencia *Water Cooler*, que representa la situación más exigente: una secuencia MC real con movimiento global acusado. Con respecto al resto de secuencias destaca la mejora de calidad de MCABMSR con respecto a MCSR en secuencias que incluyen movimiento local, como *Mobcal*, *Shields* o *Parkrun*, y no tanto en secuencias con movimiento local muy limitado como *Stockholm*.

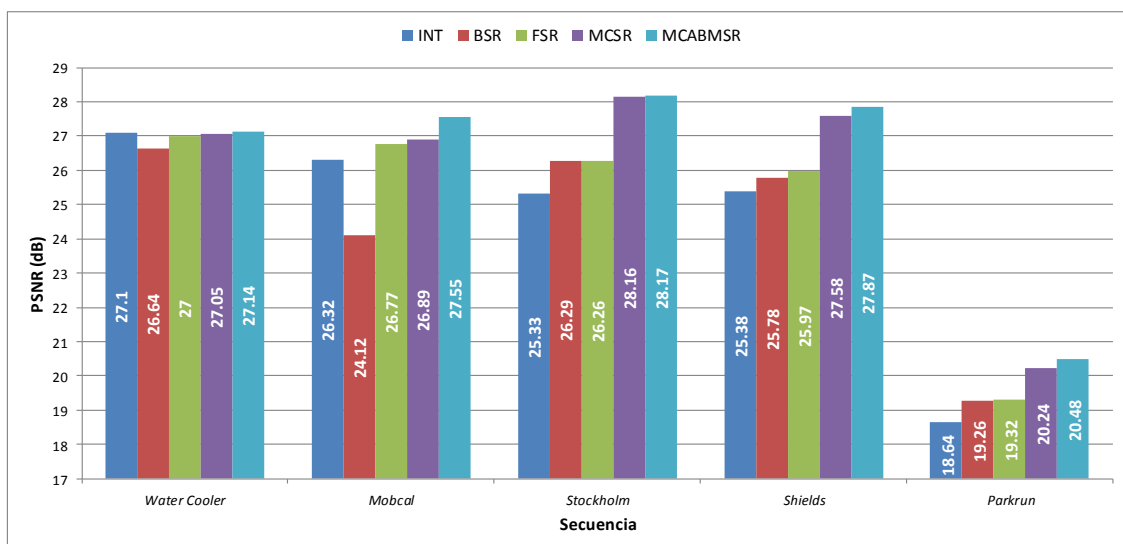


Figura 6.2.- PSNR (dB) para secuencias MC

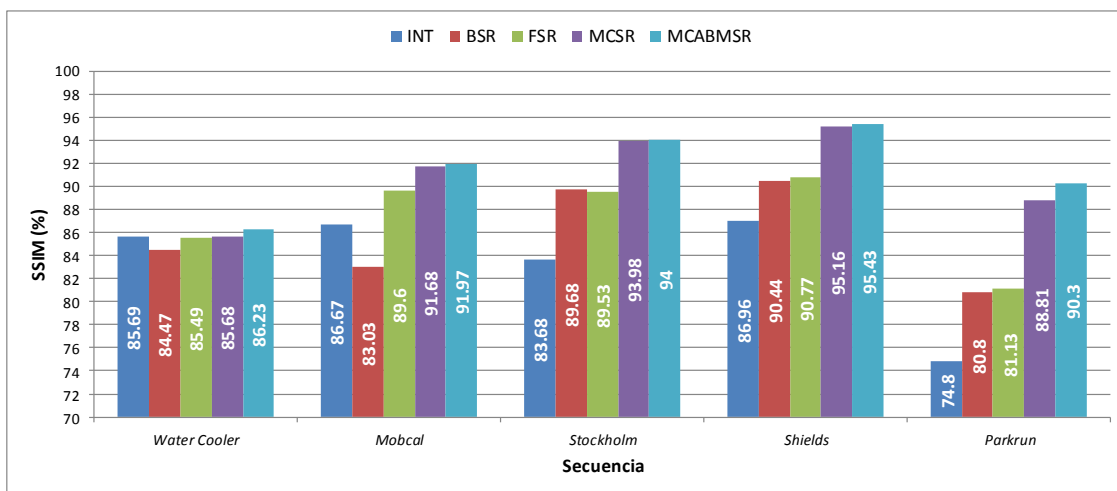


Figura 6.3.- SSIM (%) para secuencias MC

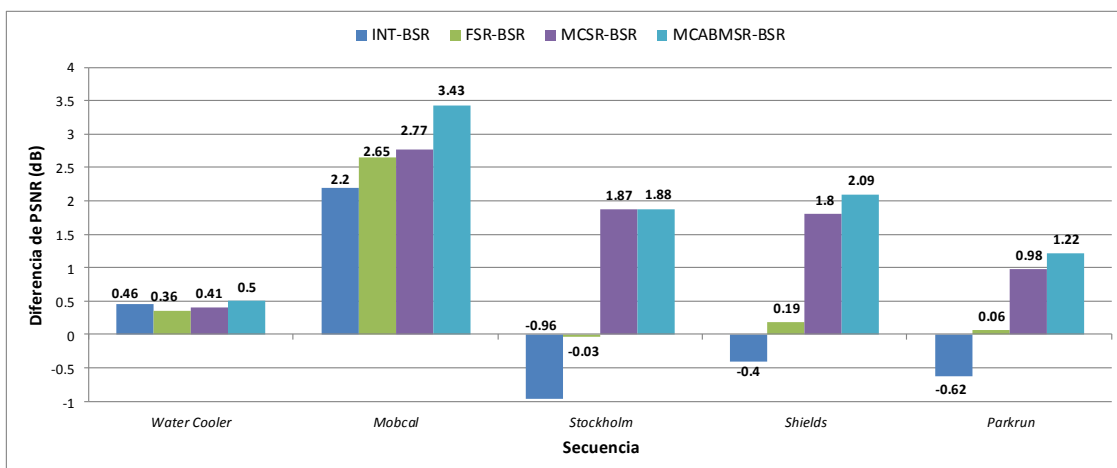


Figura 6.4.- Diferencia de PSNR (dB) con respecto a BSR para secuencias MC

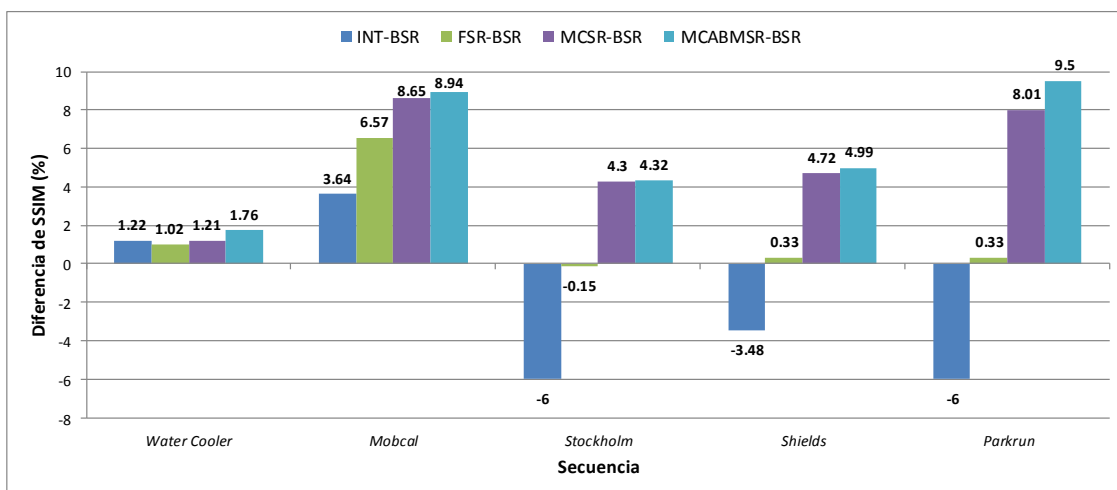


Figura 6.5.- Diferencia de SSIM (%) con respecto a BSR para secuencias MC

La Figura 6.4 y la Figura 6.5 proporcionan una clara representación de la mejora que conlleva la aplicación de cada uno de los métodos propuestos con respecto a BSR, tanto en términos de PSNR como de SSIM.

Por una parte, destacan las prestaciones del sistema global MCABMSR en secuencias con movimiento local o movimiento global acusado, como *Water Cooler* o *Mobcal*, con mejoras que van desde 0.5 dB en PSNR y un 1.76 % en SSIM para *Water Cooler*, hasta 3.43 dB en PSNR y un 8.94 % en SSIM para *Mobcal*, como se presenta en la Figura 6.4 y en la Figura 6.5, respectivamente. En estos casos, la conjunción entre los filtros selectivos desarrollados y el entorno MC (método MCABMSR) resulta muy adecuada para optimizar la calidad de la secuencia resultante, tal y como se representa en la Figura 6.6, correspondiente a la secuencia *Mobcal*, o si la mejora de calidad con respecto a MCSR no es muy significativa, como ocurre en la secuencia *Water Cooler*, se obtiene una reducción del tiempo de cómputo, como se presenta en la Figura 6.9, mejorando por tanto las prestaciones globales del sistema.

Por otra parte, en secuencias con movimiento global medio, en las que los resultados correspondientes a BSR se encuentran bastante por encima de INT, como *Parkrun* o *Stockholm*, la coherencia espaciotemporal entre *frames* con la aplicación de MC es tal, que la mejora de MCABMSR con respecto a MCSR se verá prácticamente limitada al tiempo de cómputo, con mejoras limitadas en calidad objetiva, tal y como se representa en la Figura 6.7, correspondiente a la secuencia *Parkrun*. Esto es debido a que la gran mayoría de *frames* y MBs se utilizan en el proceso de SR, alcanzando una calidad objetiva frente a BSR que va desde 1.22 dB en PSNR para *Parkrun* y un 4.32 % en SSIM para la secuencia *Stockholm*, hasta 1.88 dB en PSNR para la secuencia *Stockholm* y 9.5 % en SSIM para *Parkrun*, como se presenta en la Figura 6.4 y en la Figura 6.5, respectivamente. En estas secuencias, es además muy destacable la diferencia existente entre INT y MCABMSR, que alcanza una variación que va desde 1.83 dB en PSNR para la secuencia *Parkrun* o de 10.11 % de SSIM para la secuencia *Stockholm*, hasta 2.73 dB para la secuencia *Stockholm* y 16.42 % de SSIM para la secuencia *Parkrun*, como se presenta en la Tabla 6.1 y en la Tabla 6.2. Se observa que los máximos de calidad de PSNR no coinciden con los de SSIM, lo cual es debido a que SSIM destaca en secuencias para las que la calidad de la imagen en términos generales mejora (tal y como ocurre en la secuencia *Parkrun*), mientras que PSNR obtiene resultados significativos siempre y cuando no aparezcan artefactos (tal y como ocurre en la secuencia *Shields*).

Finalmente, en secuencias que comparten movimiento global y movimiento local medios, como es el caso de la secuencia *Shields*, se obtienen resultados muy adecuados, tal y como se presenta en la Figura 6.8, alcanzando MCABMSR un valor promedio de 2.49 dB en PSNR con respecto a INT y un 8.47% promedio de mejora en SSIM.

Tabla 6.1.- Diferencia MCABMSR con respecto a INT para secuencias MC (PSNR)

Secuencia	Diferencia de PSNR (dB)
<i>Water Cooler</i>	0.04
<i>Mobcal</i>	1.23
<i>Stockholm</i>	2.73
<i>Shields</i>	2.49
<i>Parkrun</i>	1.83

Tabla 6.2.- Diferencia MCABMSR con respecto a INT para secuencias MC (SSIM)

Secuencia	Diferencia de SSIM (%)
<i>Water Cooler</i>	0.54
<i>Mobcal</i>	5.3
<i>Stockholm</i>	10.11
<i>Shields</i>	8.47
<i>Parkrun</i>	16.42

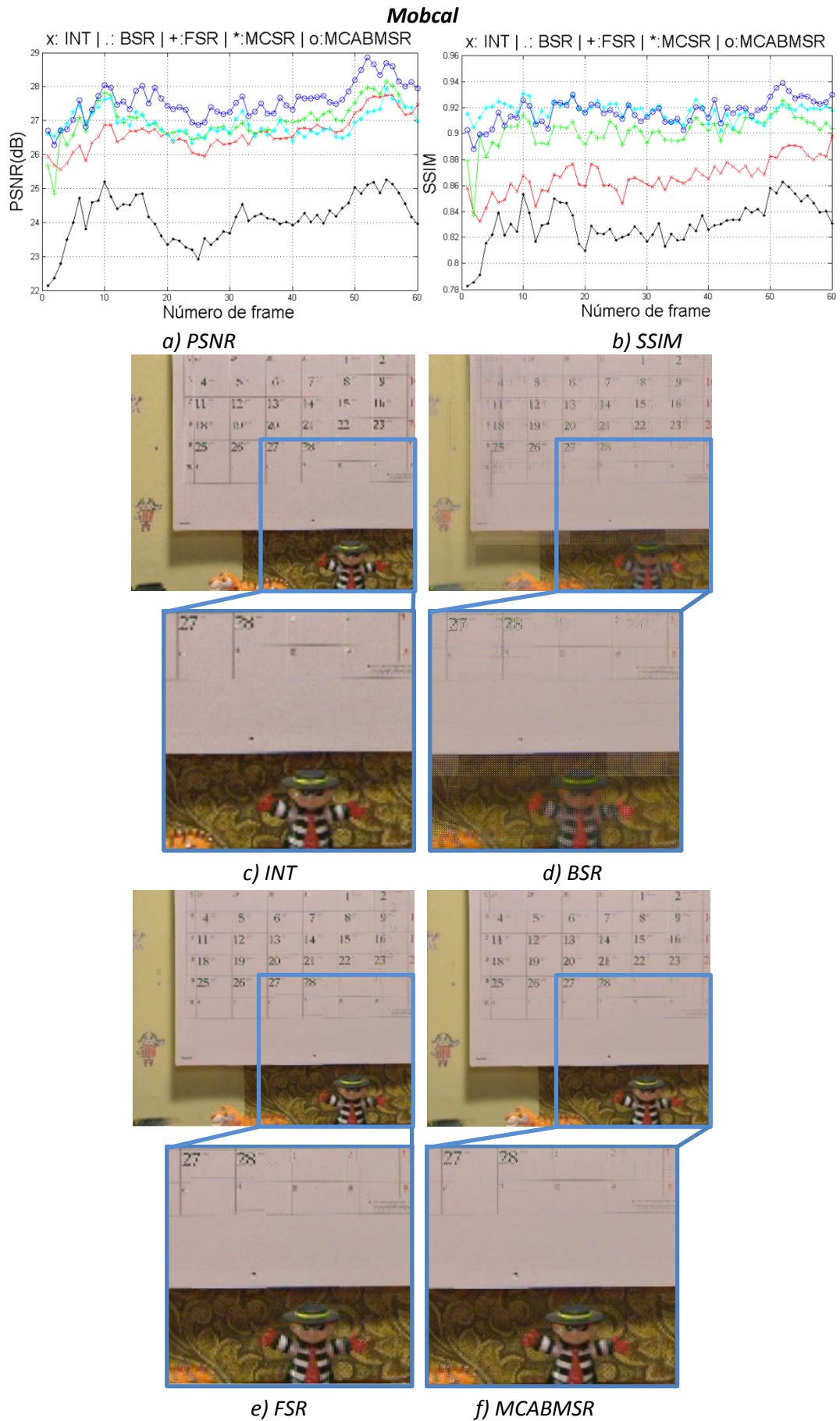


Figura 6.6.- Resultados para secuencia Mobcal

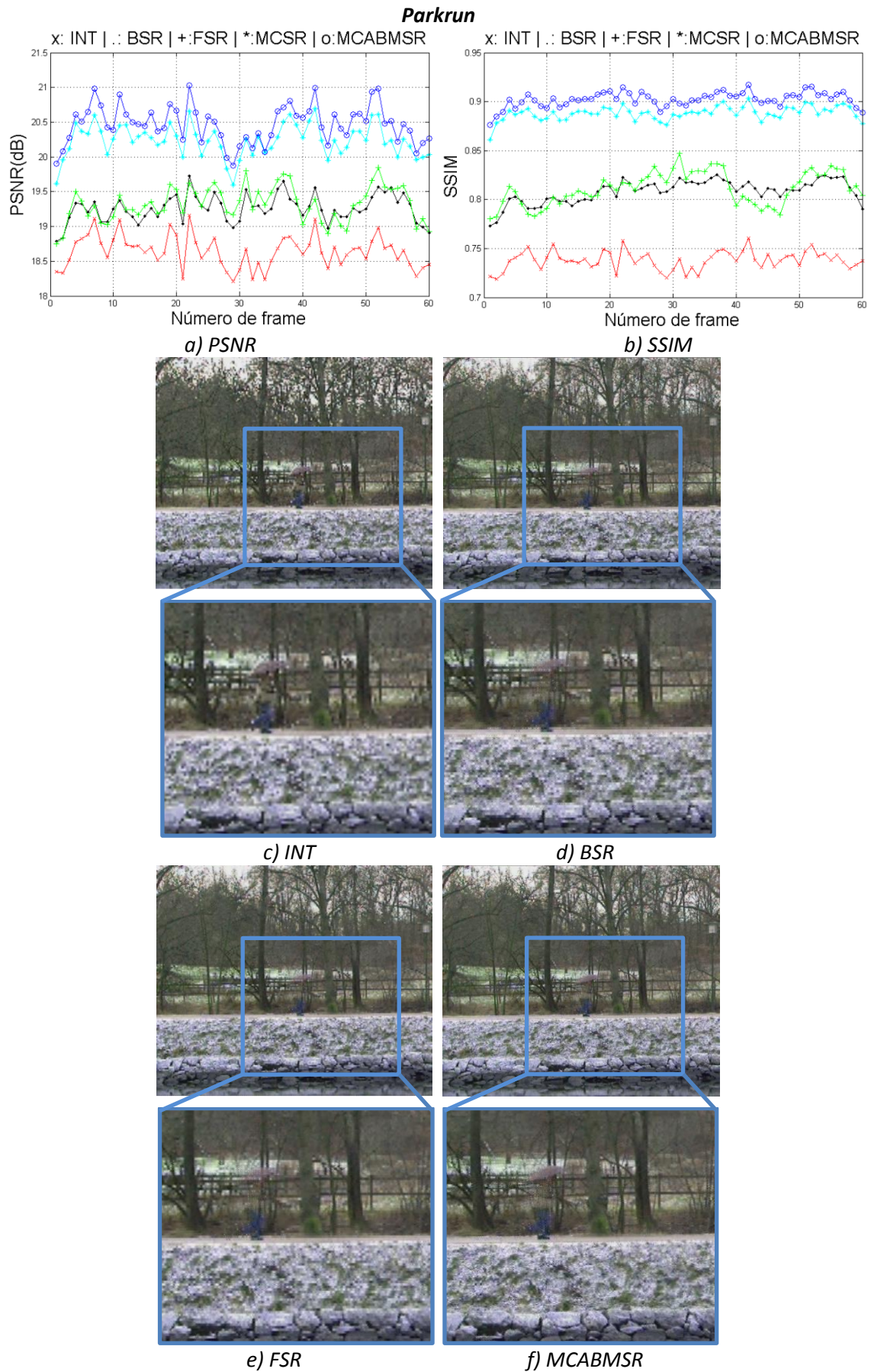


Figura 6.7.- Resultados para secuencia Parkrun

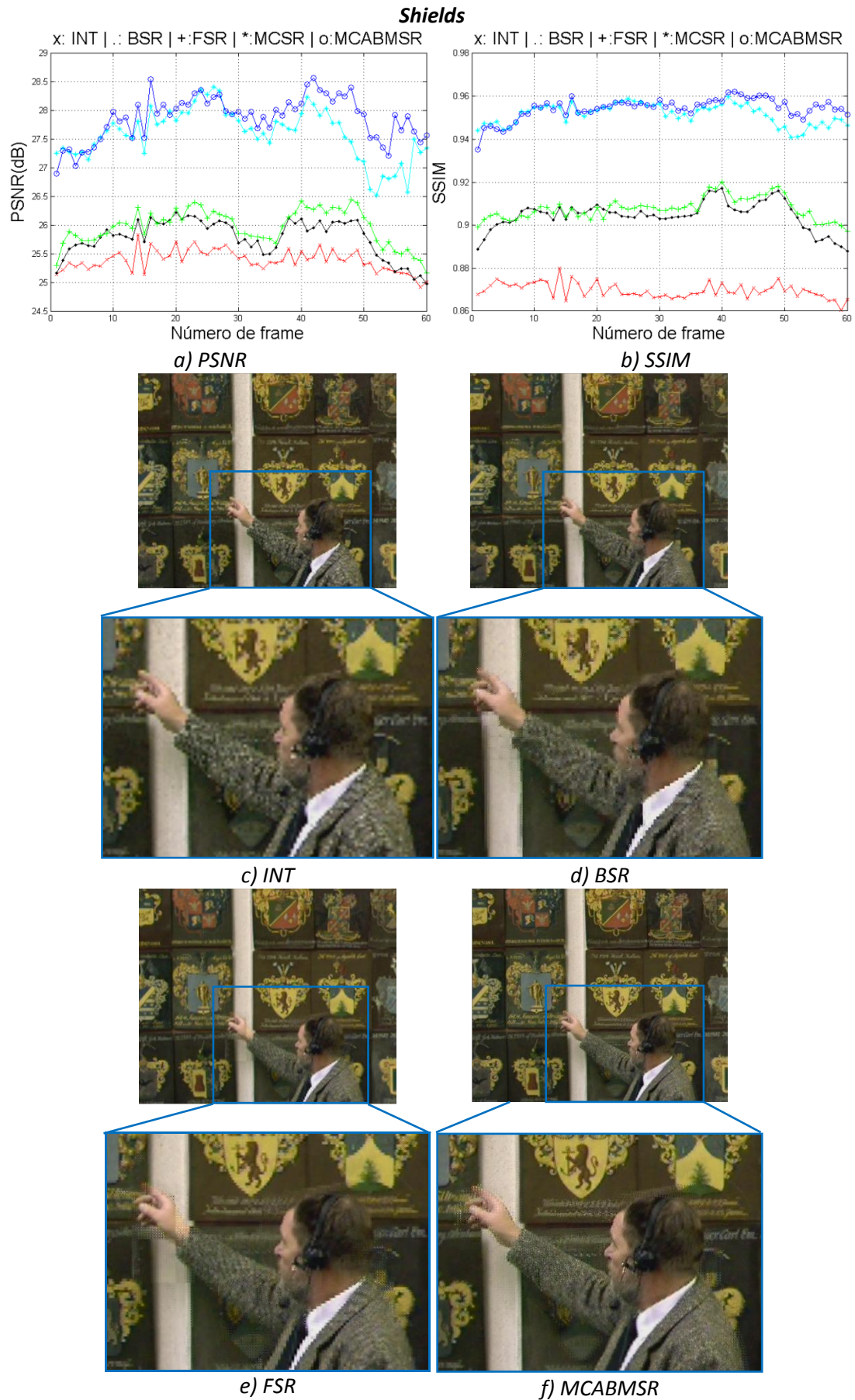


Figura 6.8.- Resultados para secuencia Shields

Finalmente, haciendo uso de una Figura de Mérito global FM_G, de forma similar al capítulo anterior, se relacionan los métodos MCABMSR y MCSR en términos de calidad objetiva (SSIM) y tiempo de cómputo de acuerdo con la expresión 6.1:

$$FM_G = \frac{SSIM(Sequencia) + \{1 - [Tiempo(Sequencia)/Tiempo_{max_met}]\}}{2}, \quad (6.1)$$

donde $Tiempo_{max_met}$ es el tiempo máximo para una determinada secuencia considerando los diferentes métodos propuestos. Así, en la Figura 6.9 se presenta el tiempo de cómputo a nivel de MB de los diferentes métodos, y en la Figura 6.10 la relación entre MCABMSR y MCSR para las secuencias estudiadas considerando la Figura de Mérito FM_G, donde se aprecia claramente la mejora de MCABMSR frente a MSR, sobre todo en la secuencia *Mobcal*, ya que es la que incluye mayor cantidad de movimiento local.

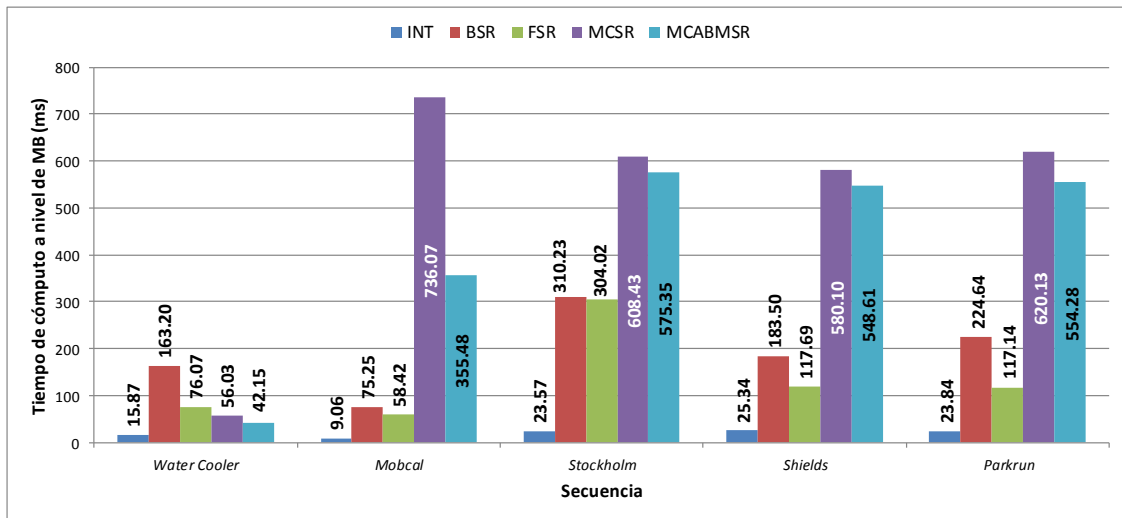


Figura 6.9.- Resultados de tiempos de cómputo para secuencias MC

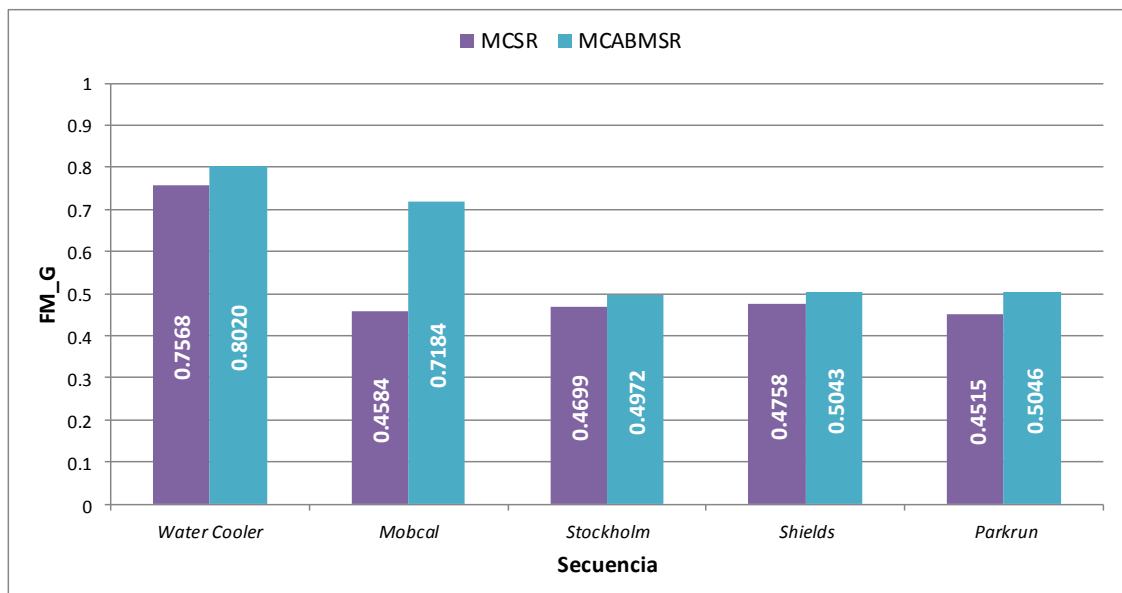


Figura 6.10.- Resultados de Figura de Mérito FM_G para MCABMSR y MCSR

6.2.2 Influencia del Factor de Escala

En el capítulo anterior de la presente Tesis Doctoral, se introdujo que con el fin de comparar los resultados del método de SR MC Mixto, con los métodos Temporal-Espacial y Espacial-Temporal, el Factor de Escala mínimo para el método Mixto debía ser 4, asignando por tanto un Factor de Escala 2 para cada una de las fases de los métodos Temporal-Espacial y Espacial-Temporal. Esto quiere decir que la situación de partida, que se ha considerado para la secuencia de entrada es bastante exigente, en el sentido de que partiendo de cada píxel de LR, se reconstruirán 16 píxeles en HR, ya que un Factor de Escala 4 implica el cuádruple de píxeles en el eje X y en el eje Y en HR con respecto a LR, como se representa en la Figura 6.11, con un píxel de LR en la esquina inferior izquierda y los saltos a Factor de Escala 2, 3 y 4 con línea roja (4 píxeles), con línea azul (9 píxeles) y con línea verde (16 píxeles), respectivamente.

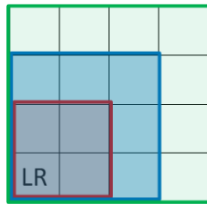


Figura 6.11.- Píxeles de factores de escala 2 (rojo), 3 (azul) y 4 (verde) partiendo de 1 píxel de LR

Se plantea por tanto en este apartado evaluar la influencia del Factor de Escala sobre la calidad objetiva de la secuencia de salida tras aplicar el proceso de SR. Con el objetivo de generar una secuencia de salida de idéntico tamaño, con independencia del Factor de Escala seleccionado, se ha de considerar que la división del número de píxeles tanto en el eje X como en el eje Y con el Factor de Escala resulten en un número entero, y que, al mismo tiempo, dicho valor sea divisible por el tamaño de MB seleccionado. A modo de referencia se muestran los resultados de la secuencia *Stockholm* considerando 3 factores de escala diferentes (2, 3 y 4) y un máximo tamaño de MB de 32 (considerando que los tamaños utilizados son múltiplos de 2, es decir 2, 4, 8, 16 y 32). El tamaño mínimo de la secuencia de salida, tanto en el eje X como en el eje Y, habrá de ser pues el producto del MCD (Máximo Común Divisor) de 2, 3 y 4, y del tamaño máximo de MB, tal y como se presenta en la expresión 6.2:

$$\text{Tamaño_salida}_{\min}(\text{píxeles}) = \text{MCD}(2,3,4) * \text{Tamaño_MB}_{\max} = 12 * 32 = 384 \quad (6.2)$$

En consecuencia se considera por tanto un recorte de la secuencia *Stockholm* de 384x384 píxeles utilizando un sistema MC 3x3 y variando el Factor de Escala para los valores 2, 3 y 4. Los resultados para la secuencia interpolada de salida y la secuencia de salida del método MCABMSR (mejor caso) y BSR (misma Ventana de Trabajo temporal que MCABMSR) se presentan en la Figura 6.12 (promedio) y en la Figura 6.13 (*frame a frame*). En la Figura 6.12 se observa cómo para un Factor de Escala 4 los resultados de PSNR y SSIM son muy similares a los calculados para la secuencia completa, así como la mejora que se obtiene para un Factor de Escala 3, y sobre todo para un Factor de Escala 2 (promedio de 41.31 dB en PSNR y 99.35% en SSIM), con máximos cercanos a 45 dB en PSNR (44.44 dB) tal y como se puede observar en la Figura 6.14, y al 100% en SSIM (99.52%) tal y como se presenta en la Figura 6.15. Por otra parte se muestra en la Figura 6.13 cómo una vez más la métrica SSIM representa mejor el sistema visual humano, ya que el mayor incremento porcentual de calidad se produce para un Factor de Escala 4 (13.24% de MCABMSR con respecto a INT) frente a los valores para un Factor de Escala 3 (5.21%) y un Factor de Escala 2 (4.55%), hecho que se puede constatar realizando una inspección subjetiva a las Figuras 6.16, 6.17 y 6.18, donde el salto de calidad entre interpolación y MCABMSR es mayor para un Factor de Escala 4 (Figura 6.16), que para un Factor de Escala 3 (Figura 6.17) y un Factor de Escala 2 (Figura 6.18).

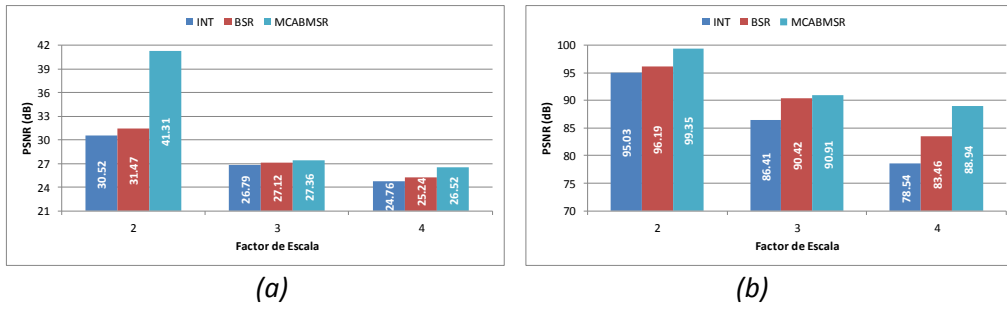


Figura 6.12.- Resultados para Factor de Escala 2, 3 y 4 para interpolación, BSR y MCABMSR. Secuencia Stockholm. (a) PSNR, (b) SSIM

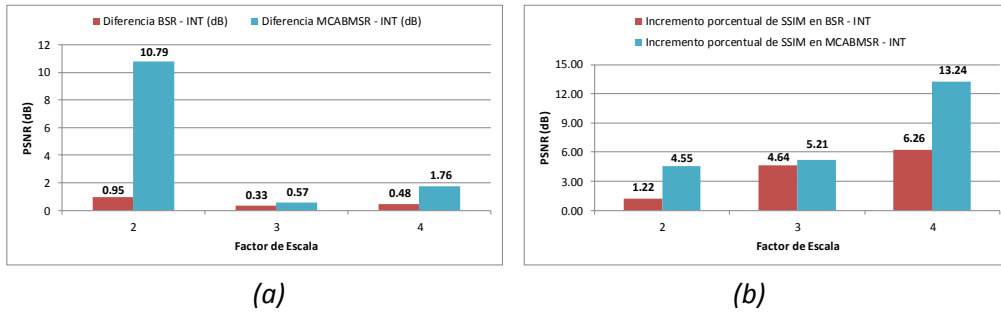


Figura 6.13.- Diferencia de SR con respecto a INT para diferentes factores de escala. Secuencia Stockholm. (a) PSNR, (b) SSIM

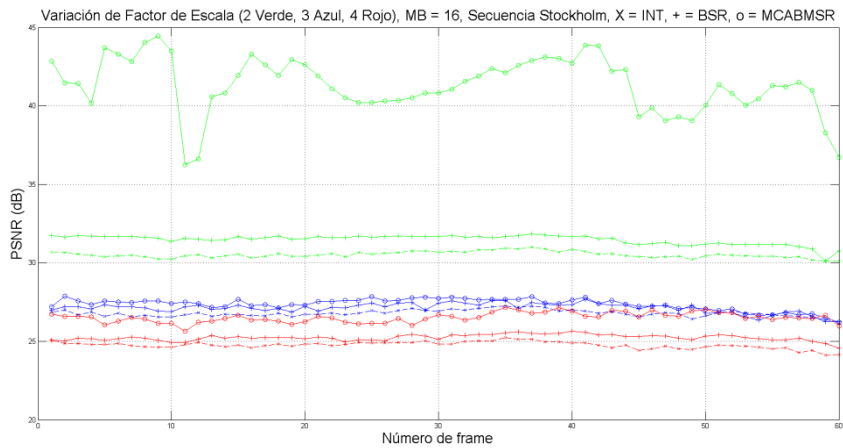


Figura 6.14.- Resultados de PSNR para Factor de Escala 2, 3 y 4 frame a frame

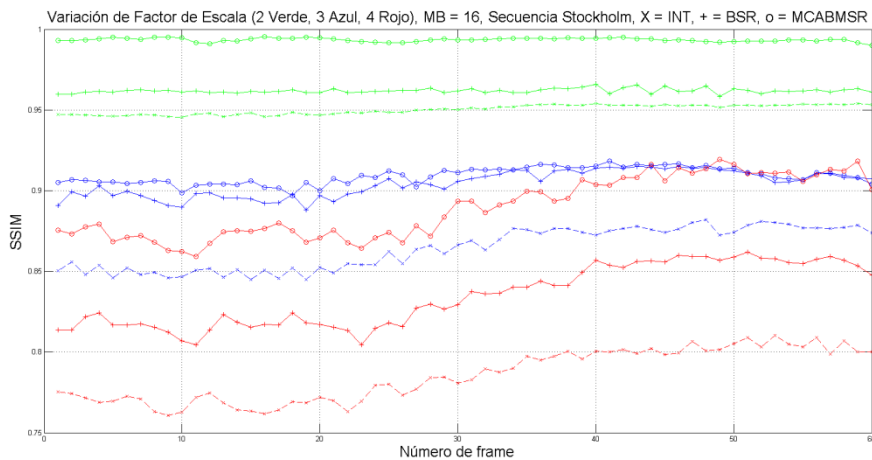


Figura 6.15.- Resultados de SSIM para Factor de Escala 2, 3 y 4 frame a frame

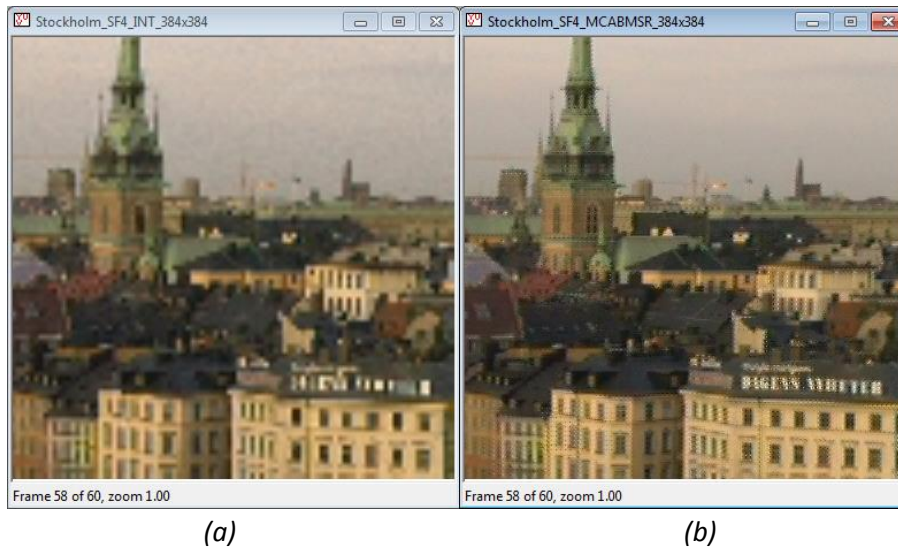


Figura 6.16.- Resultados de secuencia Stockholm para Factor de Escala 4. (a) INT, (b) MCABMSR

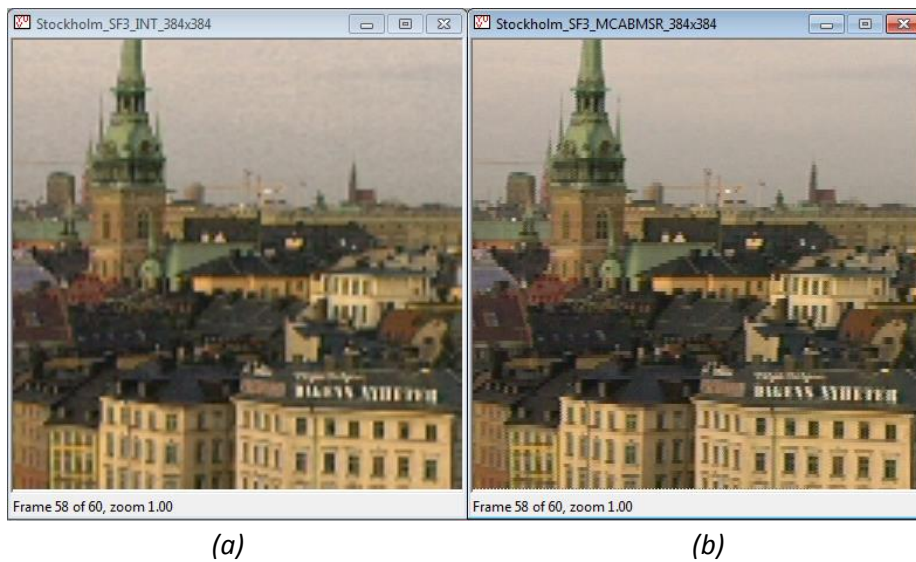


Figura 6.17.- Resultados de secuencia Stockholm para Factor de Escala 3. (a) INT, (b) MCABMSR

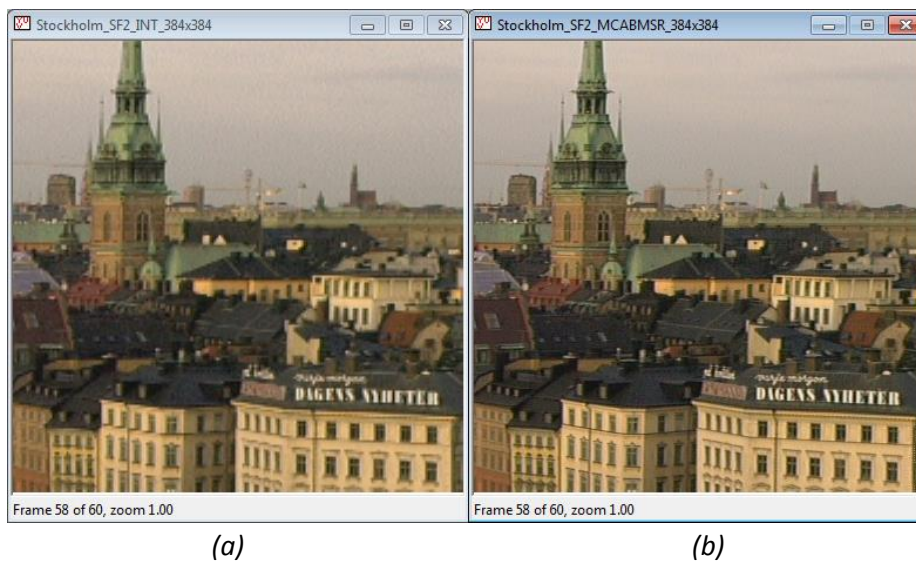


Figura 6.18.- Resultados de secuencia Stockholm para Factor de Escala 2. (a) INT, (b) MCABMSR

6.2.3 Caracterización del sistema global

En esta sección, se realiza una caracterización completa del sistema global MCABMSR propuesto en la presente Tesis Doctoral, con el objetivo de determinar aquellas fases que requieren un mayor cómputo, y cómo se modifica la situación inicial que presentaba el algoritmo BSR, ejecutando ambos algoritmos en igualdad de condiciones y varias veces para considerar la carga del Sistema Operativo. Se ha realizado por ello un *profile* haciendo uso de la herramienta de *Visual Studio Performance Wizard*. Antes de comenzar con el análisis de los resultados obtenidos a partir de este *profile*, se explicarán una serie de conceptos con el fin de poder interpretar adecuadamente los datos que aporta. Estos conceptos se corresponden con muestras o llamadas exclusivas, llamadas entrantes y las muestras inclusivas.

- *Muestra exclusiva*: La cantidad de la métrica de rendimiento que se produjo en la ejecución del cuerpo de la función.
- *Llamadas entrantes*: La cantidad de la métrica de rendimiento que se produjo en las funciones a las que llamó la función seleccionada.
- *Total inclusivo*: El total de los valores de muestra exclusiva y llamadas entrantes.

La *Estimación de Movimiento (blockMatching)* es la etapa del algoritmo que presenta un mayor cómputo en el cuerpo de su función, sin tener en cuenta el cómputo realizado por funciones que son llamadas desde ésta, seguido del SAD tanto para BSR como para MCABMSR, como se muestra en la Figura 6.19 y en la Figura 6.20, respectivamente. La principal diferencia se encuentra en el tercer lugar, en el cual para MCABMSR destaca la función *iqa_convolve* con un 17.96%, interna del cálculo del SSIM, que se utiliza para realizar comparativas entre *frames* cuando se utiliza el filtro de WSFSR. Esta es la principal diferencia existente a este nivel entre ambos algoritmos. Por otra parte, la existencia de la función que realiza el cálculo de valor absoluto *abs*, en cuarto lugar para BSR y en quinto lugar para MCABMSR se debe a que el valor mínimo de SAD es el criterio que se usa en la etapa *blockMatching* para estimar el movimiento que se ha desplazado un MB entre los *frames* de la Ventana de Trabajo respecto al *frame* de referencia o *frame* actual, y el SAD hace uso a su vez de la función *abs* para calcular su valor.



Figura 6.19.- Funciones que realizan la mayor parte del trabajo individual (BSR)



Figura 6.20.- Funciones que realizan la mayor parte del trabajo individual (MCABMSR)

Por otro lado, el *profile* de BSR, presentado en la Figura 6.21, indica que la ruta crítica de este método está determinada principalmente por las fases de *Estimación de Movimiento (blockMatching)*, y de *Compensación de Movimiento (shiftAdd)*, siendo la etapa de *blockMatching* la que mayor tiempo de proceso requiere en la ejecución completa de este método, tal y como se muestra en el valor de las muestras inclusivas, el cual indica un total de 93.18%. Sin embargo, la ruta crítica de MCABMSR, mostrada en la Figura 6.22, se encuentra a

un nivel superior, debido a la entrada de la función SSIMT (17.4% de muestra inclusivas), quedando la ruta crítica definida principalmente por la función *sr*.

Function Name	Inclusive Samples %	Exclusive Samples %
↳ __tmainCRTStartup	99,38	0,00
↳ _main	99,38	0,00
↳ _sr	99,25	0,01
🔥 _blockMatching	93,18	32,82
🔥 _shiftAdd	4,79	4,54

Figura 6.21.- Ruta crítica de BSR

Function Name	Inclusive Samples %	Exclusive Samples %
↳ _mainCRTStartup	100,00	0,00
↳ __tmainCRTStartup	99,47	0,00
↳ _main	99,47	0,00
🔥 _sr	81,79	0,05
🔥 _SSIMT	17,40	0,06

Figura 6.22.- Ruta crítica de MCABMSR

En la Figura 6.23 se muestra un desglose del porcentaje de tiempo de proceso empleado en el cuerpo de la función *sr* para MCABMSR y el porcentaje de tiempo de ésta en cada uno de los procesos que utiliza para llevar a cabo la etapa *Estimación de Movimiento*. Tal y como se observa en esta figura, la función *blockMatching* es la función que mayor tiempo se utiliza dentro de la función *sr*, debido a la gran cantidad de veces que ésta se emplea.

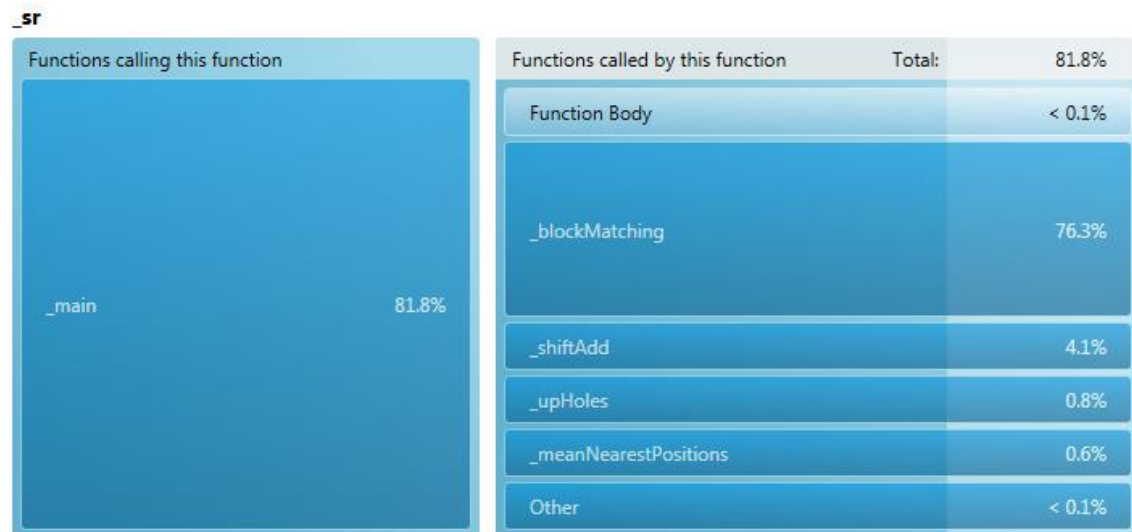


Figura 6.23.- Muestras inclusivas y exclusivas y llamadas entrantes (%) - función *sr* (MCABMSR)

A través de este *profile*, se determina que, tal y como ocurría en la situación original de BSR, es principalmente en la etapa de *Estimación de Movimiento* en la que se debe hacer un mayor esfuerzo para reducir el coste computacional de este algoritmo. Para obtener unas mejores prestaciones sería recomendable llevar a cabo una paralelización de este código, como podría ser a través de una FPGA, una unidad de múltiples núcleos o una GPU (*Graphics Processing Unit*), de forma que se mejoren los tiempos de cómputo.

6.3 APLICACIÓN EN ENTORNO SUBMARINO

6.3.1 Resultados en secuencias en entorno submarino

Partiendo de los resultados obtenidos en la sección anterior, se planteó validar el sistema global desarrollado en un entorno de aplicación real. En este sentido, tal y como se ha comentado en la introducción de la presente Tesis Doctoral, el entorno submarino es un medio adecuado para aplicar Súper-Resolución, si bien presenta una serie de características propias que se han de considerar a la hora de aplicar técnicas de restauración o mejora de imagen. En el Anexo A.2 se detallan las características propias del entorno submarino, así como algunas técnicas de restauración y mejora de imagen del Estado del Arte.

En este apartado, se presentará cómo influyen las características del entorno submarino al proceso de SR por fusión, con el objetivo de evaluar la forma en la que el algoritmo de SR introducido en la presente Tesis Doctoral, puede mejorar secuencias de vídeo submarino. Se realizaron pruebas con una amplia variedad de secuencias submarinas (ver Anexo A.1 para más información de las secuencias consideradas). El análisis planteado es similar al presentado al inicio de este capítulo, pero utilizando como referencia 6 secuencias submarinas: *Reef*, *Atlantis*, *Gruta*, *Titanic*, *Pez_Coral* y *Coral*.

La Figura 6.24 y la Figura 6.25 muestran, respectivamente, los resultados de PSNR y SSIM para las secuencias submarinas analizadas. Se observa claramente la mejora de calidad conseguida con respecto a BSR, que a modo de resumen se introduce en la Figura 6.26 (PSNR) y en la Figura 6.27 (SSIM). Por una parte destacan las prestaciones del sistema global MCABMSR en secuencias con movimiento local o movimiento global acusado. En estos casos, la conjunción entre los filtros selectivos desarrollados y el entorno MC (método MCABMSR) resulta fundamental para optimizar la calidad de la secuencia resultante, tal y como se representa en la Figura 6.28 correspondiente a la secuencia *Coral*, o bien, aun representando una ligera mejora de calidad con respecto a MCSR como ocurre en la secuencia *Reef*, presentada en la Figura 6.29, disminuye el tiempo de cómputo, mejorando por tanto las prestaciones del sistema global. Por otra parte, en secuencias con movimiento global medio, en las que BSR se encuentra bastante por encima de INT, como *Atlantis*, *Pez_Coral*, o *Titanic*, la coherencia espaciotemporal entre *frames* con la aplicación de MC es mayor, alcanzando una calidad objetiva frente a BSR que va desde 1.72 dB en PSNR para *Atlantis* y un 4.53 % en SSIM para *Titanic*, hasta 2.27 dB en PSNR para *Pez_Coral* y 9.42 % en SSIM para *Parkrun*. Así, tal y como se representa en la Figura 6.30, correspondiente a la secuencia *Pez_Coral*, la mejora de calidad objetiva es tal que las curvas de MCABMSR se estabilizan en un promedio elevado tanto en PSNR como en SSIM. Es además muy destacable la diferencia existente entre INT y MCABMSR, alcanzando una variación desde 1.04 dB para la secuencia *Coral* y un 3.11% de SSIM para la secuencia *Gruta*, hasta 3.52 dB en PSNR para la secuencia *Titanic*, o de un 18.62% de SSIM para la secuencia *Atlantis*, tal y como se presenta en la Tabla 6.3.

Tabla 6.3.- Diferencia MCABMSR con respecto a INT en secuencias submarinas (PSNR y SSIM)

Secuencia	<i>Reef</i>	<i>Atlantis</i>	<i>Gruta</i>	<i>Titanic</i>	<i>Pez_Coral</i>	<i>Coral</i>
Diferencia de PSNR (dB)	2.19	3.43	2.19	3.52	3.09	1.04
Diferencia de SSIM (%)	9.59	18.62	3.11	8.66	12.01	1.22

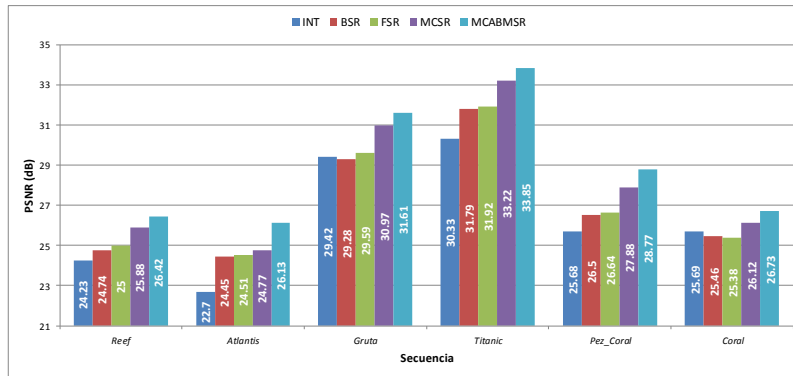


Figura 6.24.- PSNR (dB) para secuencias submarinas

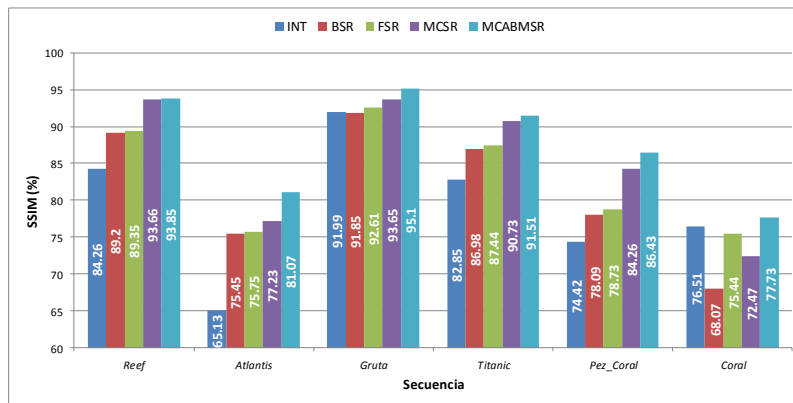


Figura 6.25.- SSIM (%) para secuencias submarinas

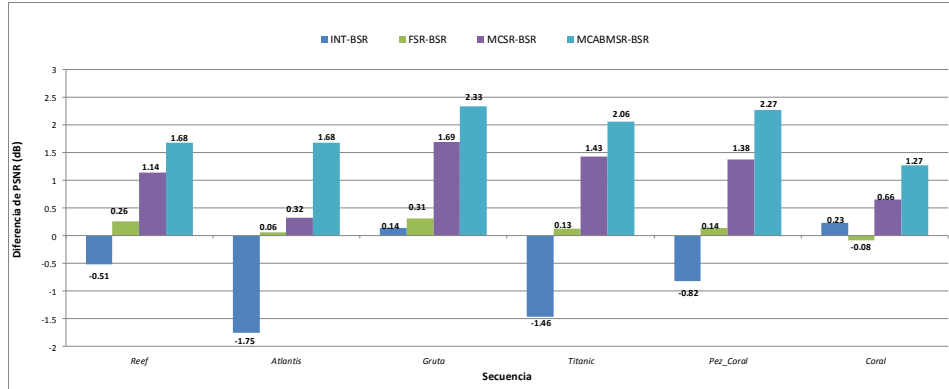


Figura 6.26.- Diferencia de PSNR (dB) para secuencias submarinas

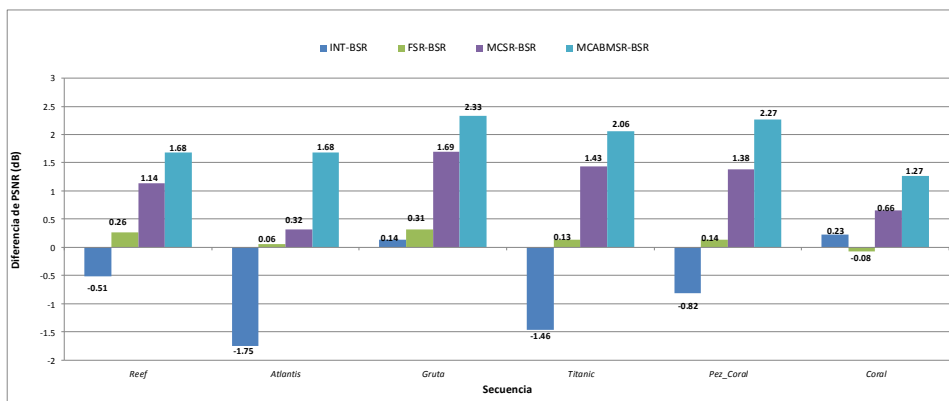
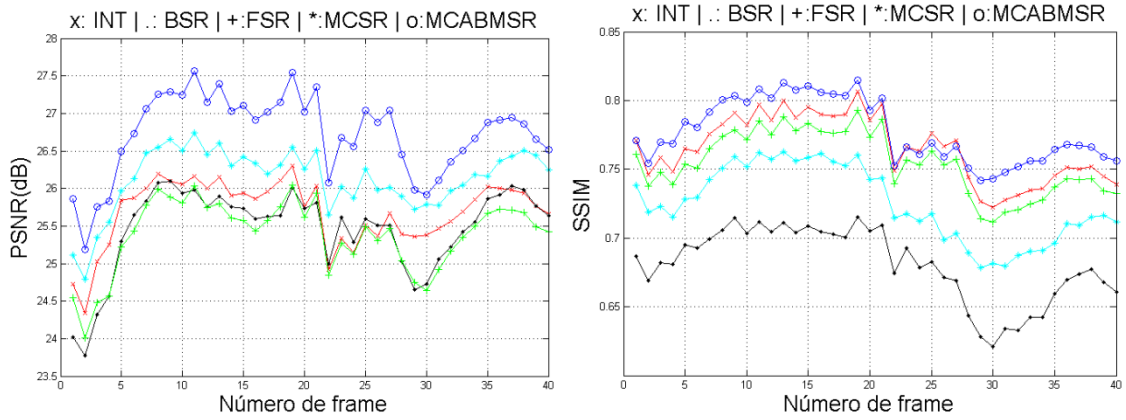


Figura 6.27.- Diferencia de SSIM (%) para secuencias submarinas

Coral



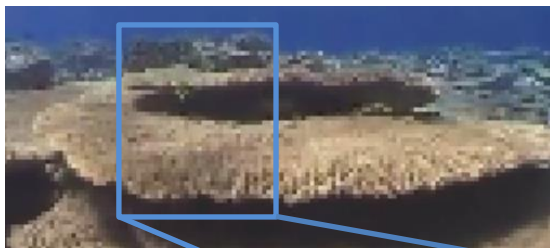
a) PSNR

b) SSIM



c) INT

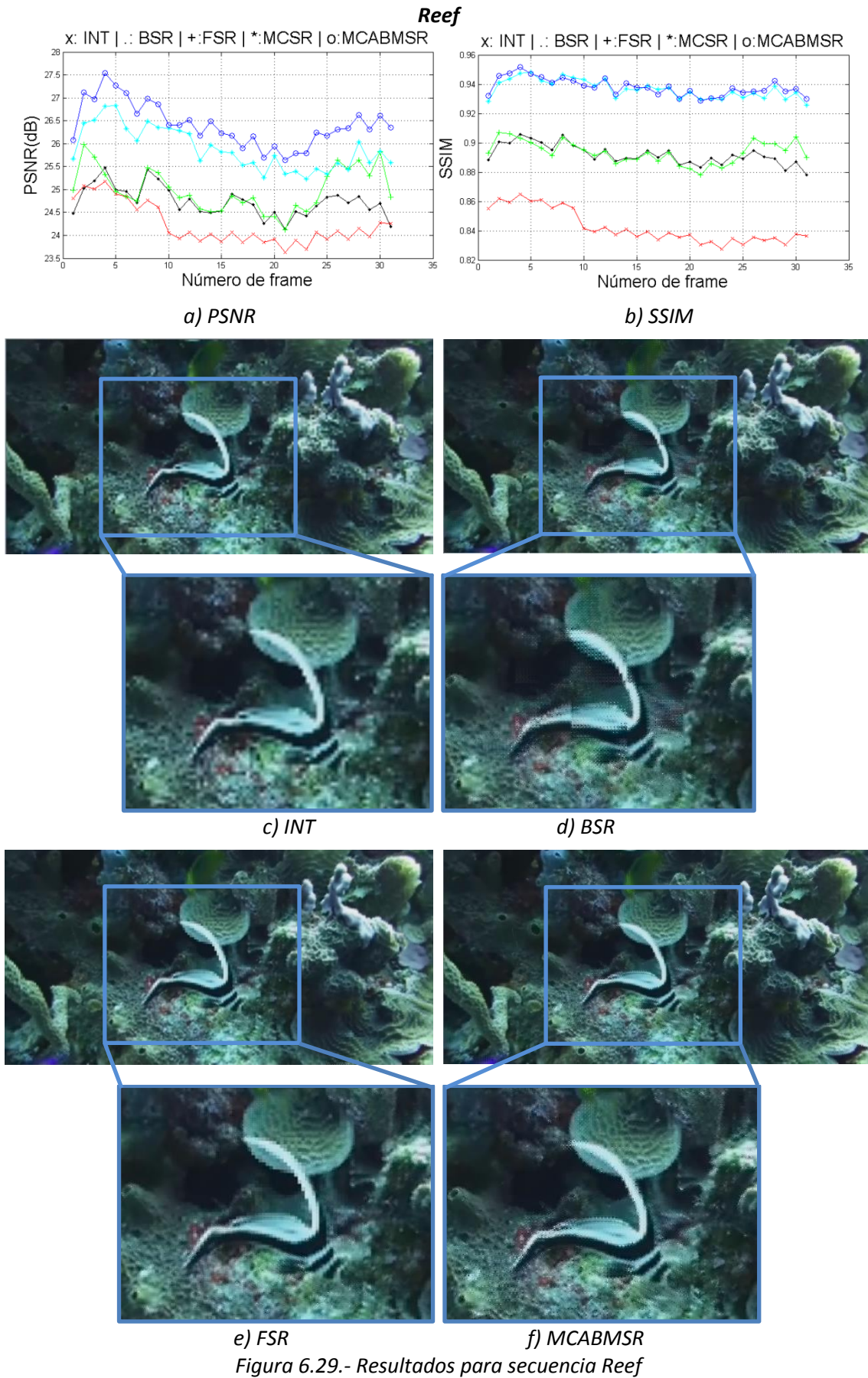
d) BSR



e) FSR

f) MCABMSR

Figura 6.28.- Resultados para secuencia Coral



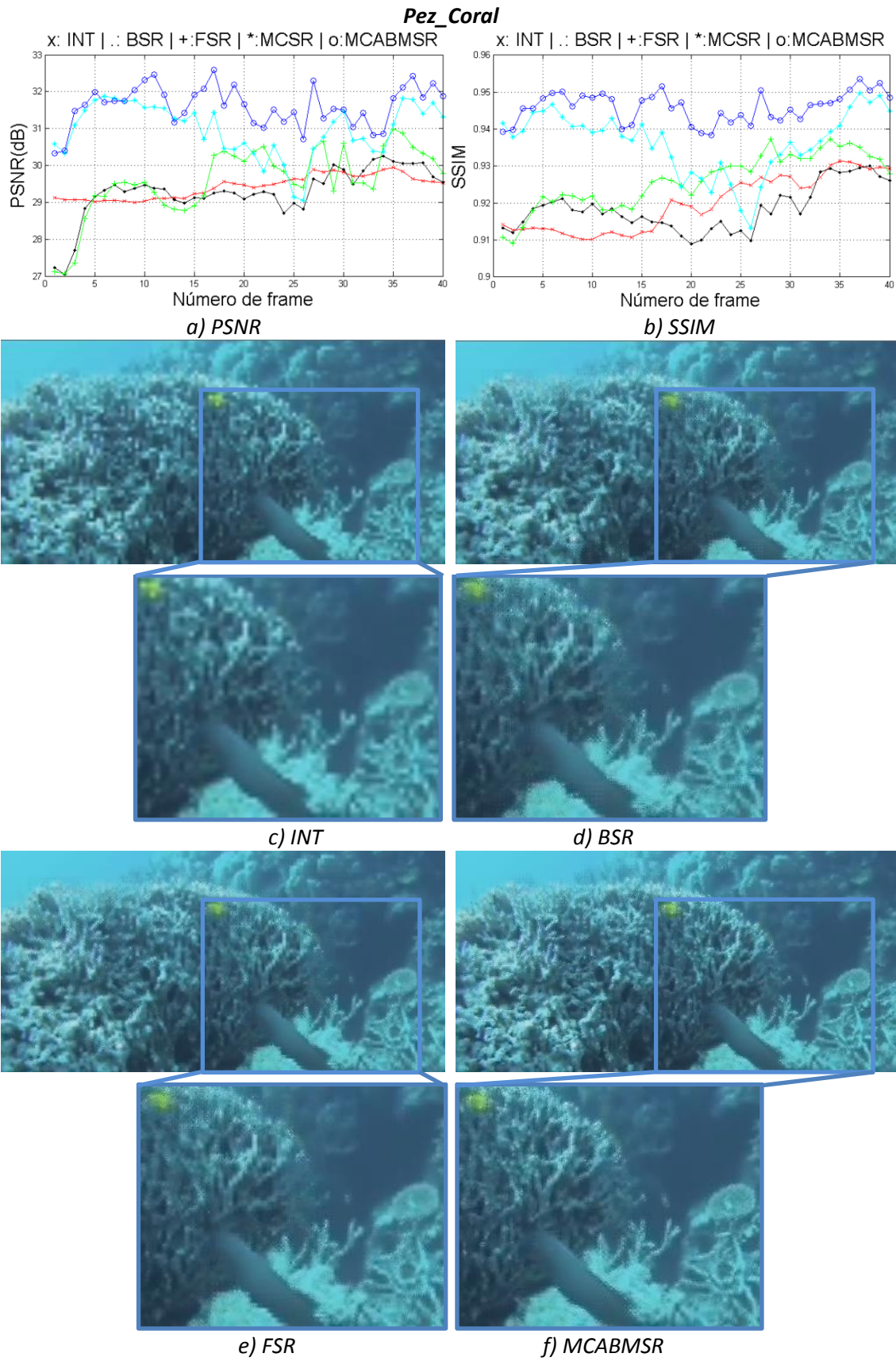


Figura 6.30.- Resultados para secuencia Pez_Coral

Finalmente, en la Figura 6.31 se presenta el tiempo de cómputo a nivel de MB de los diferentes métodos considerados y en la Figura 6.32 la relación entre MCABMSR y MCSR para las secuencias estudiadas, tomando como referencia la Figura de Mérito FM_G introducida en la expresión 6.1. En este caso, se muestra en la Figura 6.31 cómo la integración del Filtro Selectivo de Ventana y el Filtro Selectivo de MB propician una mejora sustancial en lo que respecta a los tiempos de cómputo, situándose en la mayoría de los casos en valores que quedan por debajo del 50% de los tiempos de cómputo de MCSR (6.2% en *Coral*, 38.17% en *Reef*, 48.24% en *Titanic*) o que quedan ligeramente por encima (52.95% en *Titanic*, 54.15% en *Atlantis*, y 55.05% en *Pez_Coral*). Esta reducción considerable del tiempo de cómputo con respecto a MCSR, unida a la mejora de calidad, hace que la Figura de Mérito FM_G de MCABMSR, representada en la Figura 6.32, mejore de forma notoria con respecto a la FM_G de MCSR. Como contrapartida, los tiempos de cómputo invertidos son mayores que en el resto de métodos considerados, por lo que a la hora de plantear un sistema MC habrá que tomar en consideración hasta qué punto el sistema puede asumir el coste computacional que éste conlleva.

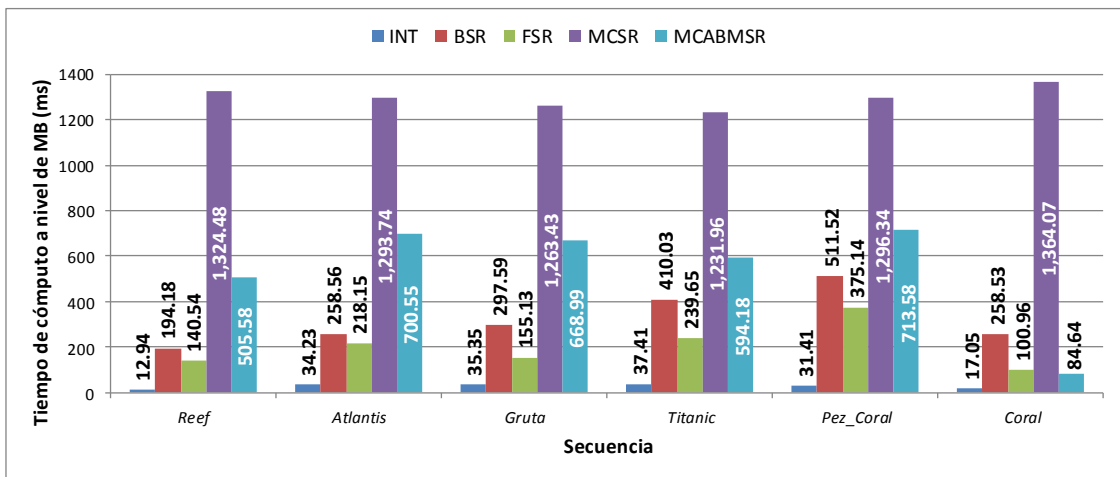


Figura 6.31.- Resultados de tiempos de cómputo para secuencias submarinas

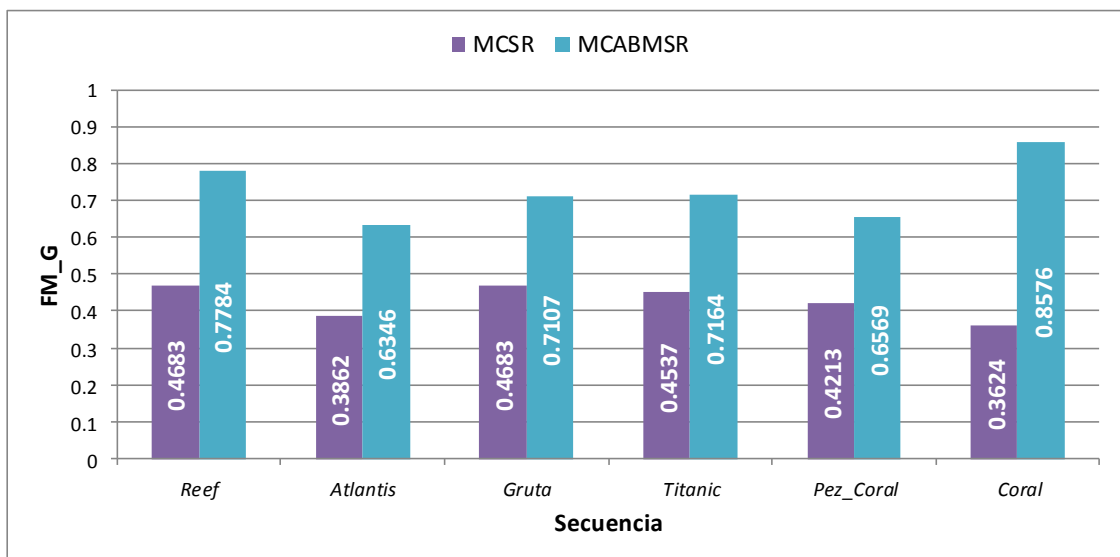


Figura 6.32.- Resultados de Figura de Mérito FM_G para MCABMSR y MCSR

Partiendo del estudio realizado en secuencias submarinas, se alcanzan las siguientes conclusiones:

- El fenómeno de la turbidez, principalmente caracterizado en secuencias submarinas por nieve marina (*marine snow*) y el fenómeno de parpadeo (*light flickering*) producen un efecto de movimiento local sobre la secuencia, que hace más compleja la reconstrucción mediante SR. Esto da lugar, además, a que el número de *frames* a utilizar en el proceso de SR sea menor y que por lo tanto el valor de umbral de Filtro Selectivo de Ventana a utilizar sea menor. La Figura 6.33.a presenta el número de *frames* utilizados para la secuencia *Titanic*, dependiendo del valor del umbral del Filtro Selectivo de Ventana WSFTHR para una reconstrucción con Factor de Escala 2, caracterizada por contener el fenómeno de nieve submarina, mientras que la Figura 6.33.b presenta una gráfica similar para la secuencia *Reef*. En ambos casos, se puede observar que el uso del número de *frames* se limita para evitar la aparición de artefactos, usando una Ventana de Trabajo promedio de unos 6 *frames* en ambas secuencias.

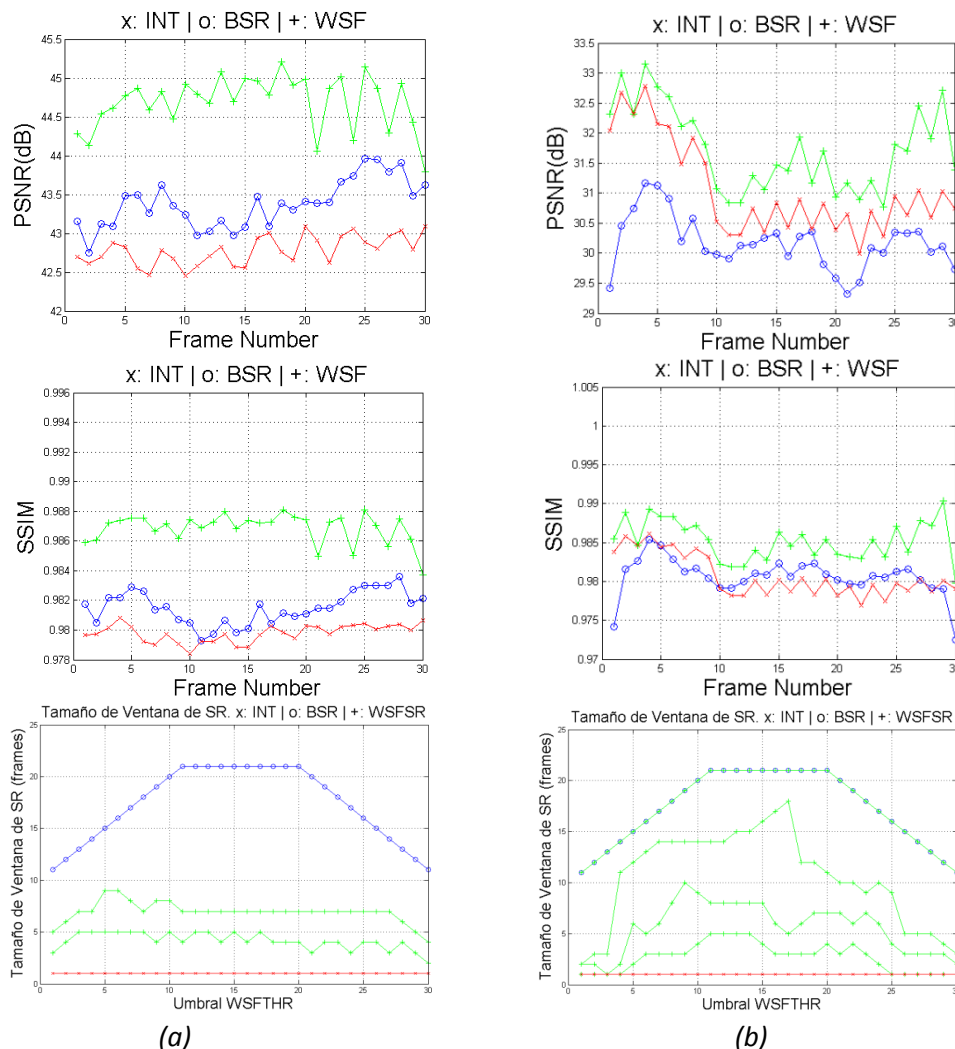


Figura 6.33.- PSNR, SSIM y tamaño de Ventana de Trabajo de SR frente a valor de umbral WSFTHR (a) Secuencia *Titanic*, PSNR y SSIM para WSFTHR = 60 y Ventana de Trabajo con WSFTHR = 0, 60, 80 y 100 de arriba abajo, (b) Secuencia *Reef*, PSNR y SSIM para WSFTHR = 40 WSFTHR = 0, 20, 40, 60, 80 y 100 de arriba abajo

- La iluminación artificial en secuencias submarinas normalmente va ligada a la cámara que realiza la grabación, por lo que se puede decir que la iluminación “se desplaza” con el movimiento de la cámara. Esto hace que la iluminación artificial se pueda considerar movimiento global. Por tanto, en secuencias con poco nivel de turbidez y parpadeo, pero con luz artificial, los resultados del proceso de SR serán mucho mejores y los valores de Filtro Selectivo de Ventana serán menores, usando por tanto un mayor número de *frames* en el proceso de SR. Esto ocurre en la secuencia *Atlantis*, cuyos resultados de PSNR y SSIM se muestran en la Figura 6.34.a y la Figura 6.34.b, respectivamente, y los resultados en lo referente al número de *frames* a utilizar y el porcentaje de tiempo de cómputo con respecto al umbral se muestran en la Figura 6.34.c y en la Figura 6.34.d, respectivamente.

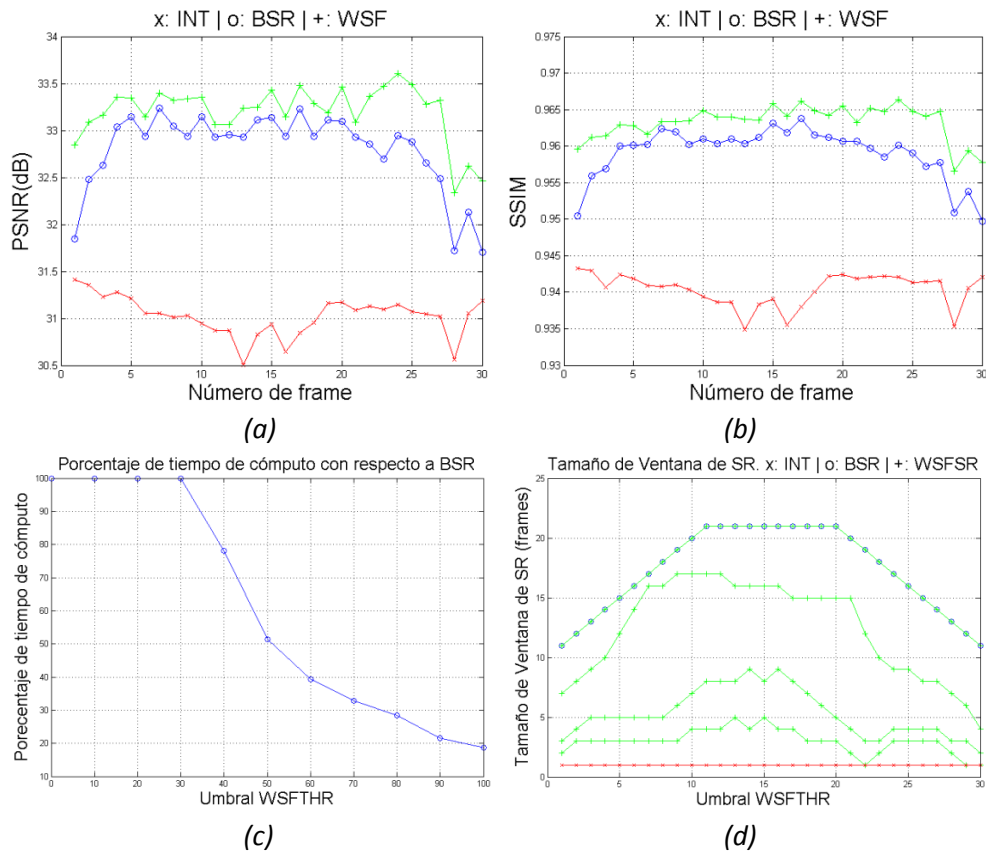


Figura 6.34.- Resultados para la secuencia Atlantis (a) PSNR, (b) SSIM, (c) Porcentaje de tiempo de cómputo, (d) Tamaño de Ventana de Trabajo frente a WSFTHR = 0, 20, 40, 60, 80 y 100 de arriba abajo

- En secuencias en las que la grabación se realiza desde una cámara móvil, normalmente operada por un buceador o un vehículo submarino, el desplazamiento de la cámara es un aspecto fundamental a considerar a la hora de aplicar el proceso de SR. En el caso particular de vehículos submarinos de exploración operados remotamente (ROV, *Remotely Operated Vehicle*), es posible determinar la velocidad del vehículo partiendo de la información de georreferenciación asociada al vehículo y del número de *frames* utilizados en el proceso de grabación. Se plantea por tanto en la presente Tesis Doctoral una aplicación de georreferenciación submarina para determinar el número de *frames* a utilizar en el proceso de SR a partir de dicha información, siguiendo una metodología similar a la desarrollada en los filtros selectivos anteriormente planteados, pero tomando como punto de partida la información espaciotemporal del vehículo. Esta aplicación se introduce en el siguiente apartado.

6.3.2 Georreferenciación submarina

Partiendo del conocimiento generado a partir del proceso de filtros selectivos y del entorno MC, se plantea una aplicación combinada de la técnica de SR que considera la posición y el movimiento de la cámara en un flujo de tiempo determinado. Tomando como campo de aplicación el entorno submarino y un vehículo operado remotamente tipo ROV (*Remotely Operated Vehicle*) se determina cómo mediante Georreferenciación es posible especificar el número de *frames* a utilizar en el proceso de SR [C3, C6, R2].

Con el objetivo de procesar adecuadamente los vídeos submarinos tomados como referencia, se hizo uso de una aplicación *web* que integra un visor con tecnología HTML5. El usuario selecciona un transecto (representación gráfica del recorrido del ROV en un mapa) y el vídeo empieza entonces en el instante de tiempo asociado al punto seleccionado. Para disminuir la latencia y hacer que el visor fuera lo más rápido posible, los vídeos se subieron a una red distribuida de contenidos CDN (*Content Distribution Network*) que sirve a los usuarios directamente a partir de las solicitudes realizadas. Esta decisión, mejoró drásticamente los tiempos de latencia y la experiencia del usuario, ya que se produce la descarga de fragmentos de vídeo, lo que soluciona el hecho de que los archivos sean de varios GB de tamaño. La Figura 6.35 presenta la aplicación web desarrollada. En la parte superior izquierda se presenta el mapa que incluye los transectos recorridos por el ROV, en este caso se trata de la evaluación del fondo marino realizada en el ámbito del estudio de impacto ambiental de la Plataforma Oceánica de Canarias (PLOCAN). En la parte superior derecha se presenta el visor de vídeos georreferenciados, así como el *frame* actual. Finalmente, en la parte inferior se muestra información adicional del vídeo como:

- Coordenadas GPS (latitud y longitud) en grados minutos y segundos.
- Coordenadas UTM (E – *Easting*, N – *Northing*) en metros. Se presenta además la zona de las coordenadas UTM (*Zone*).
- Ancho (*Width*) y Alto (*Height*) en píxeles de la secuencia de vídeo analizada.
- Tasa de *frame* (*Framerate*) de la secuencia de vídeo analizada.

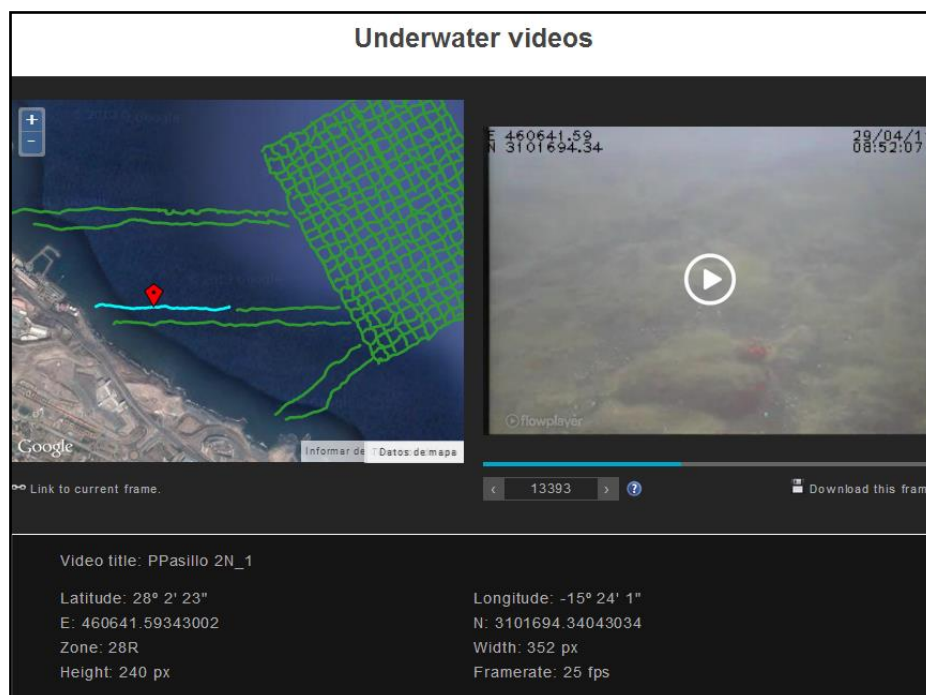


Figura 6.35.- Aplicación web de vídeos submarinos georreferenciados

El número de *frames* seleccionados para un vídeo determinado depende directamente de la velocidad media del vehículo entre un *frame* de control (*key frame*) y el siguiente *frame* de control (un *frame* de control es un aquel que contiene información georreferenciada en coordenadas UTM de *Easting* y *Northing*). La velocidad media se calcula entonces como el cociente entre la distancia recorrida y el tiempo transcurrido como se muestra en la Figura 6.36, teniendo en cuenta una curvatura de la tierra nula (las pequeñas distancias garantizan esta suposición). Como resultado del estudio realizado, se obtuvo que la calidad de la secuencia de vídeo resultante mejoraba cuando el tamaño de la Ventana de Trabajo de SR disminuía a medida que se incrementaba la velocidad media del vehículo, y viceversa, tal y como se presenta en la Figura 6.37.

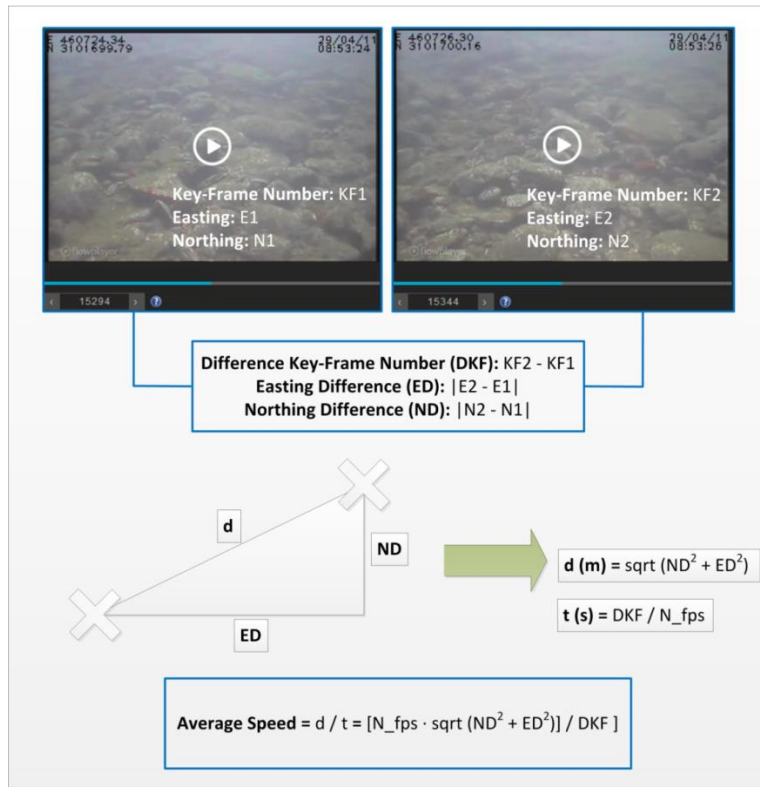


Figura 6.36.- Cálculo de la velocidad media (average speed) del vehículo

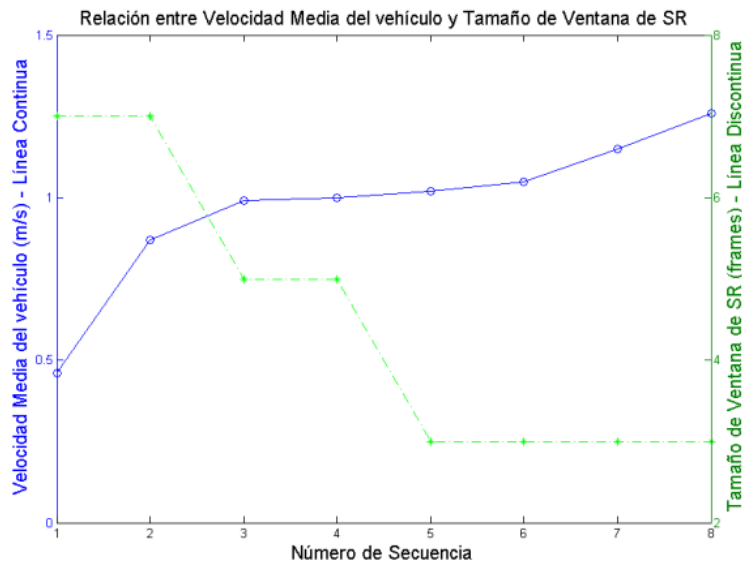


Figura 6.37.- Relación entre velocidad media de vehículo y tamaño de Ventana de Trabajo

Toda vez que la relación entre la velocidad del vehículo y el número de *frames* seleccionados se ensaya convenientemente, la selección de los *frames* más adecuados para el proceso de SR, cuyo método asociado se ha denominado GRSR (*GeoReferenced Super-Resolution*) proporciona muy buenos resultados, igualando los obtenidos con WSFSR, como se representa en la Figura 6.38 para un extracto de los vídeos submarinos georreferenciados presentados al principio de esta sección; o incluso superándolos ampliamente, como se muestra en la Figura 6.39, correspondiente a otro extracto de los vídeos submarinos georreferenciados. Se parte de una situación inicial de BSR con una Ventana de Trabajo de SR de +/- 5 (se establece un valor no muy elevado ya que los vídeos en cuestión incluyen mucho movimiento global y local).

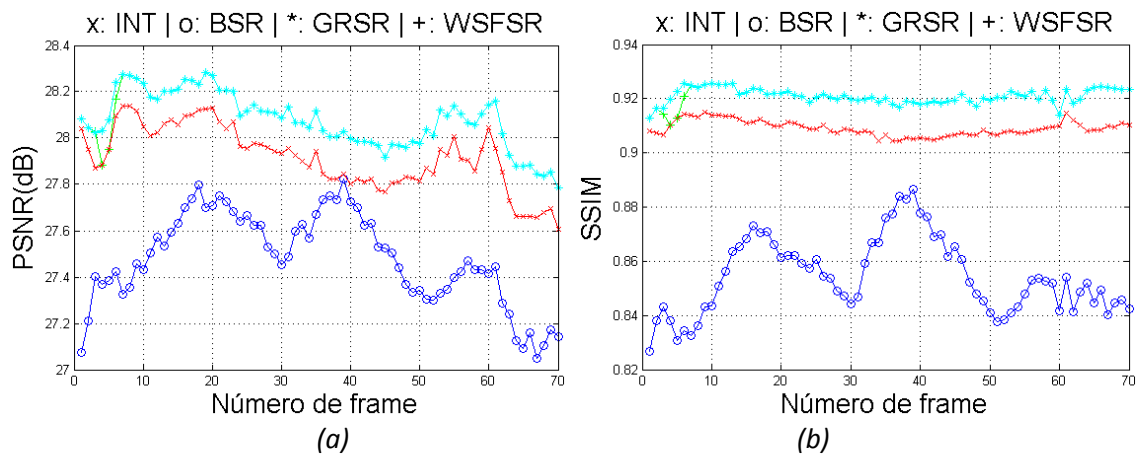


Figura 6.38.- Resultados de extracto de transecto 2N1 (frames 14819-14888) (a) PSNR, (b) SSIM

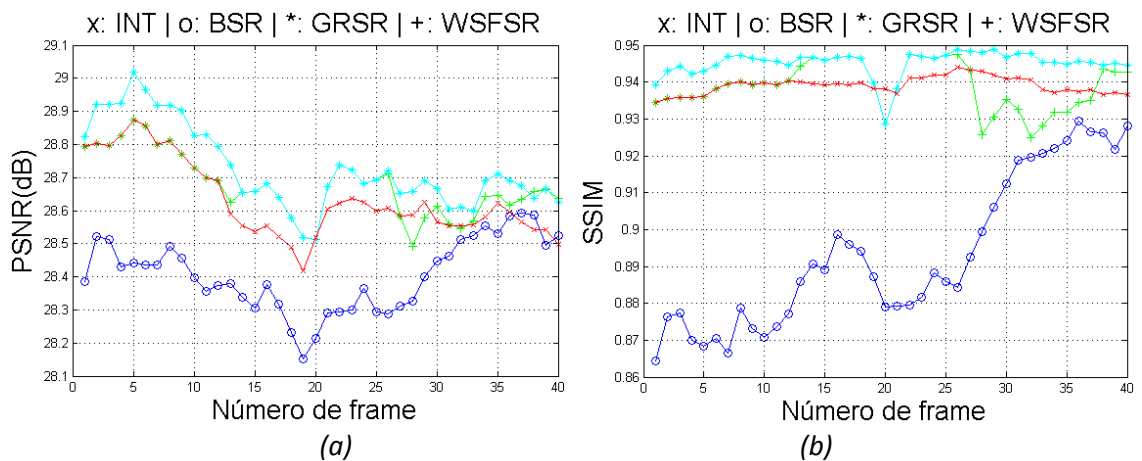


Figura 6.39.- Resultados de extracto de transecto 2N1 (frames 16319-16358) (a) PSNR, (b) SSIM

6.4 TEST DE CALIDAD SUBJETIVA

Una vez que se han presentado los resultados de las contribuciones de la presente Tesis Doctoral en términos de calidad objetiva y tiempos de cómputo, se realiza finalmente un *test* de calidad subjetiva. Tal y como se presentó en el capítulo 2 *Estado del Arte*, el objetivo de la evaluación subjetiva de calidad es conocer, para un conjunto determinado de *frames*, la opinión de un usuario medio. En este sentido, la mejor forma de saberlo es preguntando directamente a los usuarios. Para realizar el *test* se utilizaron las recomendaciones más habituales [ITU98, ITU02, ITU08]:

- Al menos 15 observadores deberían participar en el experimento. En este caso se ha tomado una muestra de 20 observadores.
- No han de ser profesionales en la evaluación de calidad. Se ha cumplido con este criterio, preguntando a profesionales de varios sectores.
- Deberían disponer de una capacidad visual normal o corregida para ser normal (uso de lentillas o gafas).

La estrategia de evaluación subjetiva seguida ha sido el método de comparación por Pares (PCO, *Pair COmparison*). En el cuestionario se presentaron los *frames* de *test* se por pares, mostrando versiones procesadas (BSR y MCABMSR) de la misma muestra original (*frame* de LR de entrada al algoritmo de SR). Tras presentar cada par, el sujeto seleccionaba su preferida. Además se preguntó si se detectaba una mejora significativa. Para el *test* se utilizaron las secuencias *Stockholm*, *Shields*, *Parkrun*, *Mobcal*, *Reef* y *Coral*, con el objetivo de obtener una muestra representativa de secuencias con características variadas. Tanto el cuestionario como los *frames* de *test* introducidos se pueden consultar en el Anexo A.3 de la presente Tesis Doctoral. Una vez recogidos, los resultados se presentan en las tablas y figuras que se muestran a continuación de la siguiente forma:

- Tabla 6.4: Porcentaje de veces en las que se selecciona MCABMSR como opción preferida frente a BSR.
- Figura 6.40: Detalle de veces en las que se selecciona MCABMSR como opción preferida frente a BSR. Cada color representa una secuencia.
- Tabla 6.5: Porcentaje de veces en las que se aprecia una mejora significativa de MCABMSR frente al *frame* de LR.
- Figura 6.41: Detalle de veces en las que se aprecia una mejora significativa de MCABMSR frente al *frame* de LR. Cada color representa una secuencia.

Se observa en la Tabla 6.4 cómo prácticamente para el 100% de los casos se ha seleccionado MCABMSR frente a BSR, salvo el caso extremo (*outlier*) del sujeto 15, tal y como se presenta en la Figura 6.40. Por otra parte, en la Tabla 6.4 a la hora de decidir si existe una mejora significativa, se aprecia claramente una distribución diferente dependiendo de las características de cada secuencia:

- Movimiento global puro (*Stockholm*, *ID=1*): Esta secuencia, muy apropiada para algoritmos de SR por fusión, se selecciona para el 100% de los casos.
- Secuencias que combinan movimiento global y ligero movimiento local (*Shields*, *Parkun* o *Reef*, *IDs = 2, 3 y 5*). Se observa una mejora significativa en el 85% de los casos.
- Secuencias con movimiento local acusado y/o mucho movimiento global (*Mobile* o *Coral*, *IDs = 4 y 5*). Se observa una mejora significativa en el 50% de los casos (*Coral*) y en el 65% de los casos (*Mobile*).

Tabla 6.4.- Selección de MCABMSR como opción preferida frente a BSR

ID	1	2	3	4	5	6
Secuencia	Stockholm	Shields	Parkrun	Mobile	Reef	Coral
Porcentaje	100	100	95	100	95	100

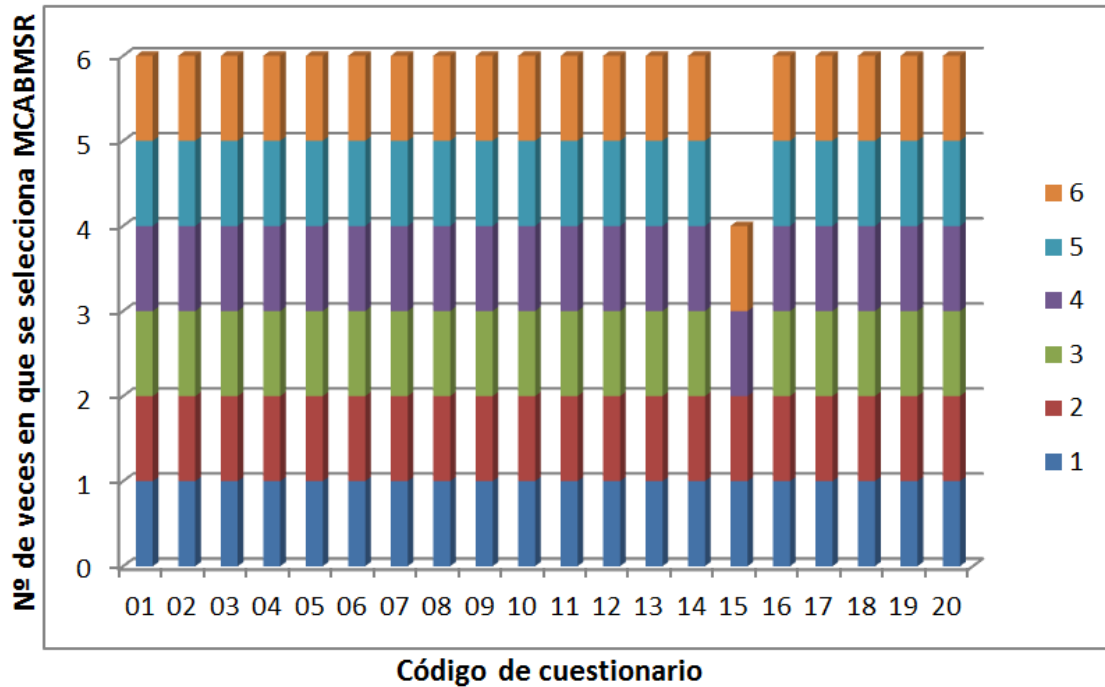


Figura 6.40.- Detalle de selección de MCABMSR como opción preferida frente a BSR

Tabla 6.5.- Apreciación de mejora significativa de MCABMSR frente a frame de LR

ID	1	2	3	4	5	6
Secuencia	Stockholm	Shields	Parkrun	Mobile	Reef	Coral
Porcentaje	100	85	85	65	85	50

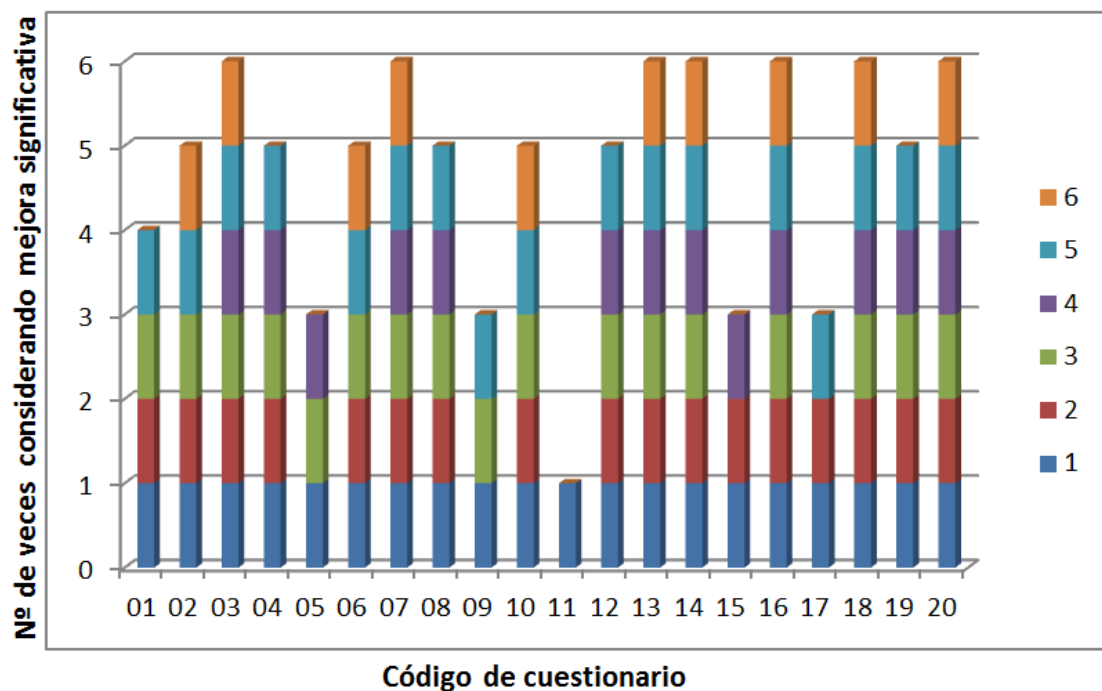


Figura 6.41.- Detalle de mejora significativa de MCABMSR frente a frame de LR

6.5 CONCLUSIONES

En este capítulo, se ha presentado el análisis del sistema global propuesto en la presente Tesis Doctoral. En primer lugar, se ha realizado un estudio de la integración entre los Filtros Selectivos de Ventana y de MB y la topología de MBs adaptable con el entorno MC, partiendo de las conclusiones extraídas a partir de los capítulos 4 y 5, y estableciendo así un sistema conjunto que contemplase un *array* MC haciendo uso de un método de SR MC Mixto y un modo de pre-procesamiento *Full Frame* con Filtros Selectivos de Ventana y de MB a su salida. A partir de los resultados obtenidos, se concluye la mejora que conlleva la integración de los filtros sobre el método de SR MC Mixto, disminuyendo el tiempo de cómputo e incrementando la calidad objetiva en la totalidad de los casos estudiados. Como contrapartida, el tiempo de cómputo supera ligeramente la situación del algoritmo original BSR.

Se ha llevado a cabo entonces la combinación de las contribuciones de la presente Tesis Doctoral en un caso concreto de aplicación, como es el entorno submarino, validando las prestaciones del sistema global en una amplia variedad de secuencias, considerando las particularidades de este tipo de secuencias como son las variaciones de color, brillo y contraste, el parpadeo de la luz solar, la nieve marina, o la iluminación artificial. Además, se ha estudiado una innovadora idea, basada en la combinación de la Súper-Resolución con la Georreferenciación Submarina. De todo ello, se deduce la importancia de la coherencia espaciotemporal entre *frames* para que los algoritmos de SR por fusión, como el estudiado en la presente Tesis Doctoral, operen correctamente, siendo fundamental para ello que exista un desplazamiento subpíxel entre *frames*.

Finalmente, se ha realizado un *test* de calidad subjetiva, concluyendo que, el algoritmo final resultante de esta Tesis Doctoral, MCABMSR, mejora claramente la calidad del algoritmo BSR que se tomó como punto de partida, y se aprecia una mejora significativa con respecto a la entrada de baja resolución en prácticamente la totalidad de las secuencias, exceptuando casos muy complejos para la Súper-Resolución, que incluyen movimiento local acusado y/o excesivo movimiento global.

Capítulo 7

CONCLUSIONES Y LÍNEAS FUTURAS

Índice del capítulo

	<i>Página</i>
7.1.- Introducción	206
7.2.- Conclusiones	207
7.3.- Líneas Futuras	209

La verdadera generosidad hacia el futuro consiste en entregarlo todo al presente
Albert Camus, autor, periodista y filósofo francés (1913-1960)

7.1 INTRODUCCIÓN

En esta Tesis Doctoral, se han presentado un conjunto de contribuciones en el ámbito de las técnicas de Súper-Resolución por fusión, cuyo objetivo ha sido mejorar la calidad objetiva y subjetiva de las secuencias de vídeo resultantes, así como disminuir el tiempo de cómputo. Para ello, por una parte, se han aportado dos filtros que permiten seleccionar los *frames* o los Macro-Bloques más adecuados dentro de la Ventana de Trabajo del proceso de Súper-Resolución, integrando además una topología de Macro-Bloques adaptable. Por otra parte, se han implementado tres métodos de SR, que permiten combinar la información temporal y espacial proporcionada por un sistema Multi-Cámara: Temporal-Espacial, Espacial-Temporal y Mixto. Finalmente, se han integrado las aportaciones realizadas en un sistema completo, demostrando su validez en un entorno real, como es el medio submarino. Además, el sistema se ha probado también con éxito en secuencias del ámbito de la electrónica de consumo.

Este último capítulo, proporciona una visión general de los hitos más destacables alcanzados en el desarrollo de la presente Tesis Doctoral. Se analizarán entonces las principales conclusiones del trabajo realizado, para entonces exponer las líneas de investigación futuras que se abren a partir de éste.

7.2 CONCLUSIONES

La presente Tesis Doctoral ha pretendido abordar los desafíos clave que se plantean en el desarrollo y la aplicación de las técnicas de SR: eficiencia en el cómputo, registro de la imagen, robustez y límites en las prestaciones. Para ello se ha tomado como punto de partida un algoritmo base BSR (*Baseline Super-Resolution*) y se ha implementado un conjunto de sistemas, con el objetivo de mejorar el Estado del Arte actual. El análisis de los resultados obtenidos a partir de las contribuciones realizadas, ha demostrado cómo las técnicas de Súper-Resolución por fusión pueden alcanzar un mayor grado de robustez, cuando se utiliza una topología variable en el análisis de los MBs de un *frame* específico, considerando los límites de prestaciones de la SR, expresados con máximos teóricos. Por otra parte, la aplicación de filtros selectivos, tanto a nivel de *frame* como de MB, ha resultado ser una vía apropiada para mantener o incluso mejorar la calidad de la secuencia de vídeo resultante disminuyendo el tiempo de cómputo. Además, la implementación de los métodos de SR aplicados a MC: Temporal-Espacial, Espacial-Temporal y Mixto, ha dado lugar a una mejora significativa de la calidad del algoritmo de partida, así como de la interpolación bilineal. Finalmente, la integración de las contribuciones de la presente Tesis Doctoral en un sistema completo, se ha ajustado a una situación real en entorno submarino.

El análisis conjunto del Filtro Selectivo de Ventana y de MB con la implementación de una topología MBs adaptable, ha proporcionado muy buenos resultados en términos de calidad tanto objetiva como subjetiva. Este hecho, ha propiciado además la independencia actual del algoritmo con los parámetros de entrada del algoritmo BSR tomado como referencia. Como contrapartida, el tiempo de cómputo aumenta debido a la adaptación de los MBs a un tamaño óptimo. Por ello, la implementación conjunta de un filtro *grueso* a nivel de *frames* con otro a nivel *fino* a nivel de MBs, proporciona resultados similares, se toma esta opción, como la más favorable para minimizar el tiempo de cómputo y mantener o mejorar la calidad. Se infiere una implementación factible del algoritmo partiendo de casos de éxito previos [CLN+02, CLL+06, CLS+08, LCT+09, SCN11].

Con respecto a los métodos de SR con MC, a partir de los resultados obtenidos se ha concluido que, de entre los métodos propuestos, el que presenta la mejor relación calidad / cómputo / recursos es el método Mixto. En términos generales, este método proporciona los mejores resultados y requiere de un único proceso de SR. Esto, unido a que para este método tan solo es necesario un *core/thread*, hace que se considere el que mejor prestaciones presenta, de entre los propuestos. Los métodos Temporal-Espacial y Espacial-Temporal presentan unos resultados, que dependen del orden en el que se aplican los procesos de SR, y requieren de varios procesos de SR para su implementación, siendo el método Temporal-Espacial el que presenta mejores tiempos de procesamiento, a expensas de requerir un alto incremento de los recursos *hardware*. Por otra parte, los modos de pre-procesamiento que mejores resultados proporcionan son *Full Frame* y *Overlap*, siendo en la parte solapada, donde se encuentra la mayor cantidad de información para reconstruir la secuencia. En cuanto al modo *Overlap+Borders*, al procesar por separado los bordes del solape, no se han obtenido buenos resultados. Por tanto, puede ser interesante aplicar alguna modificación a este modo, de forma que se puedan obtener mejores prestaciones.

Una vez evaluadas las contribuciones realizadas, se decidió conformar un sistema completo formado por el método de SR MC Mixto y usando el modo de pre-procesamiento *Full Frame* conjuntamente con un Filtro Selectivo de Ventana de Trabajo y un Filtro Selectivo de MB (*Window & Block Selective Filter*, WBSF) conformando la contribución global

denominada MCABMSR (*Multi-Camera Adaptive Block Matching Super-Resolution*), que se aplica tanto a secuencias MC, como a secuencias reales tomadas en un entorno submarino.

De la contribución global MCABMSR, se destaca por una parte que, partiendo de una secuencia que presenta unas características inicialmente desfavorables, como es el caso de *Water Cooler*, cuya grabación está destinada a probar algoritmos de estabilización de vídeo, y que además presenta unas malas características espaciales entre cámaras, y temporales entre *frames*, como se recoge en el Anexo A.1 de la presente Tesis Doctoral, se han logrado obtener mejoras sobre la secuencia generada con el algoritmo BSR y la secuencia interpolada. Por otra parte, también se han de resaltar los buenos resultados del sistema conjunto en secuencias normalmente complejas para técnicas de SR por fusión, debido al movimiento local que incluyen, como son la secuencia artificial *Mobcal* o las secuencias submarinas *Coral*, *Gruta* y *Reef*. Así, los valores representados en la Tabla 7.1 son los ejemplos más representativos de la mejora aportada por la contribución MCABMSR sobre el algoritmo BSR y la secuencia interpolada para secuencias con movimiento local o movimiento global acusado, destacando los resultados obtenidos para la secuencia *Reef*.

Tabla 7.1.- Diferencias MCABMSR-BSR y MCABMSR-INT (secuencias con movimiento local o movimiento global acusado)

Métrica	PSNR (dB)		SSIM (%)		
	Secuencia / Dif.	MCABMSR - BSR	MCABMSR - INT	MCABMSR - BSR	MCABMSR - INT
<i>Reef</i>		1.68	2.19	4.65	9.59
<i>Gruta</i>		2.33	2.19	3.25	3.11
<i>Coral</i>		1.27	1.04	9.66	1.22
<i>Mobcal</i>		3.43	1.23	8.94	5.3
<i>Water Cooler</i>		0.5	0.04	1.76	0.54

Por otra parte, en casos en los que la situación es tal que las técnicas de SR por fusión funcionan normalmente bien, como aquellas en las que el movimiento es principalmente global y el movimiento local está muy limitado, la coherencia espacial y temporal cuando se integra un sistema MC es muy elevada, por lo que la calidad de MCABMSR mejorará significativamente con respecto a BSR, tal y como lo hace BSR con respecto a INT. Esto ocurre en secuencias artificiales como *Stockholm*, *Shields* o *Parkrun*, y en secuencias submarinas como *Atlantis*, *Pez_Coral* o *Titanic*, destacando los resultados obtenidos para la secuencia *Atlantis*, tal y como se presenta en la Tabla 7.2.

Tabla 7.2.- Diferencias MCABMSR-BSR y MCABMSR-INT (secuencias con movimiento global y movimiento local limitados)

Métrica	PSNR (dB)		SSIM (%)		
	Secuencia / Dif.	MCABMSR - BSR	MCABMSR - INT	MCABMSR - BSR	MCABMSR - INT
<i>Atlantis</i>		1.72	3.47	8.3	18.62
<i>Parkrun</i>		1.22	1.83	9.42	16.42
<i>Pez_coral</i>		2.27	3.09	8.34	12.01
<i>Stockholm</i>		1.88	2.73	4.32	10.11
<i>Titanic</i>		2.06	3.52	4.53	8.66
<i>Shields</i>		2.09	2.49	4.99	8.47

Por tanto, la combinación de un sistema MC para SR conjuntamente con filtros adecuados que discriminen los *frames* y MBs más apropiados a considerar en el proceso de SR resulta muy apropiada para técnicas de SR por fusión, si bien se ha de considerar el tiempo de cómputo adicional que representa con respecto a los algoritmos de SR por fusión tradicionales.

7.3 LÍNEAS FUTURAS

A partir del análisis de los resultados de la presente Tesis Doctoral, se desprenden una serie de aspectos que se podrían optimizar para mejorar las prestaciones de las contribuciones realizadas. Se proponen, por tanto, una serie de posibles líneas futuras de investigación.

Con respecto al desarrollo de los filtros selectivos de Ventana de Trabajo y de MBs, así como de la topología variable de MBs, es fundamental la definición de decisores adecuados relacionados con los umbrales definidos. Se plantea por ello, que podría resultar apropiado incluir un mayor número de variables en la toma de decisiones. Este es un proceso complejo, ya que cada variable que se integra en la toma de decisiones, ha de estar muy bien seleccionada para funcionar de forma adecuada con las secuencias de vídeo de baja resolución. Para apoyar esta línea de investigación, sería además interesante la utilización de clasificadores, que permitieran caracterizar las secuencias de entrada.

Por otra parte, en relación con los sistemas MC, una de las posibles mejoras es la creación de algoritmos de pre-procesamiento que corrijan la diferencia de perspectiva, los desalineamientos, los cambios de escala, o la variación de la luminosidad y de color, que pueden existir entre las cámaras del *array* MC. Como se ha podido ver, especialmente a través de los resultados obtenidos en la secuencia *Fujidog*, esta mejora puede resultar muy ventajosa. Una posible solución a la hora de afrontar los inconvenientes mencionados puede consistir en introducir una fase de calibración del *array* MC, justo antes de realizar una sesión de grabación con el sistema MC, en la cual se extraigan una serie de parámetros que se usen para posteriormente aplicar una corrección directa sobre los *frames* de cada cámara. Todo ello partiendo de la base de que durante una sesión de grabación con el *array* MC, estos parámetros no deberían variar. Éste es un paso que se realiza habitualmente cuando se trabaja con este tipo de sistemas y que no sólo reduce el coste computacional del pre-procesamiento, ya que no es necesario obtener para cada *frame* grabado en cada instante de tiempo estos parámetros, sino que además, el cálculo de los parámetros comentados también puede ser de interés, ya que las oclusiones, cambios de escala, etc., pueden deberse al movimiento local de objetos o personas que se encuentren en la escena. En consecuencia, puede resultar conveniente obtener estos parámetros en cada instante en que el *array* MC capture un *frame*. Es más, se podrían combinar ambas propuestas, considerando la calibración del *array* como una extracción de parámetros *grosso modo*, los cuales se afinarán mediante el cálculo de los parámetros *frame a frame*.

Por otra parte, otra línea de investigación posible estaría relacionada con la automatización del parámetro del Área de Búsqueda (*Search Area*, SA). En la caracterización de este parámetro podrían entrar en juego dos de las contribuciones de la presente Tesis Doctoral. Por un lado, la determinación de una topología variable del tamaño de MB podría relacionarse directamente con el tamaño del Área de Búsqueda, tratando de establecer una correspondencia biunívoca entre estos parámetros. Por otro lado, en un sistema MC sería interesante obtener la distancia en píxeles entre las cámaras, puesto que si ésta es muy grande, implicaría tener que aumentar el valor del Área de Búsqueda del algoritmo de SR, incrementando el coste computacional de éste. Sin embargo, si se conoce esta distancia, se podría proporcionar este dato como valor de entrada al algoritmo de estimación de movimiento, sobre el cual se establecería el Área de Búsqueda, reduciendo por tanto el coste computacional cuando fuera posible.

La combinación planteada en la presente Tesis Doctoral, entre sistemas MC y filtros selectivos de Ventana de Trabajo y de MB, ha proporcionado buenos resultados para la disminución de los típicos artefactos que se reproducen en los algoritmos de SR por fusión. Como mejora adicional, se podrían plantear mapas de profundidad o de oclusión, de modo que ayuden a reducir la cantidad de artefactos introducidos en los procesos de SR con MC Espacial o Mixto, al conocer qué objetos quedan ocultos debido a la posición de las cámaras del *array* MC. Así mismo, esto ayudaría a optimizar los recursos de memoria y de procesamiento, puesto que con ello se pueden descartar ciertas partes de los *frames* de la Ventana de Trabajo, en los que no conviene aplicar SR por la falta de información debida a las oclusiones.

En cuanto al modo de pre-procesamiento *Overlap+Borders*, podría modificarse para intentar obtener mejores resultados, súper-resolviendo todos los bordes y el solape sin llevar a cabo la fase de *Rellenado de Huecos*, y una vez súper-resuelto cada borde y el solape, unirlos para posteriormente aplicar la fase de *Rellenado de Huecos* a todos ellos de forma conjunta.

Finalmente, en el ámbito de aplicación en entorno submarino, sería adecuado tratar de realizar pruebas reales integrando un sistema MC en un vehículo tipo ROV y aplicar post-procesamiento en las secuencias de vídeo captadas aplicando SR.

Anexos

Índice del capítulo

	<i>Página</i>
A.1.- Secuencias de vídeo utilizadas	213
A.2.- Entorno submarino: Características y mejora de imagen	219
A.3.- <i>Test</i> de Calidad Subjetiva: Cuestionario	229
A.4.- Publicaciones	231

Hay que unirse, no para estar juntos, sino para hacer algo juntos
Juan Donoso Cortés, filósofo y político español (1809 – 1853)

A.1.- SECUENCIAS DE VÍDEO UTILIZADAS

En este anexo, se resumen las principales características espaciales y temporales de las secuencias de vídeo utilizadas en esta Tesis Doctoral. Por una parte, se utilizan estas secuencias con el objetivo de establecer una comparación posterior con otros algoritmos de mejora de imagen existentes en el Estado del Arte. Estas secuencias se representan en la Tabla A.1. En la Tabla A.2 se presentan además secuencias adecuadas para comparación en sistemas Multi-Cámara.

Tabla A.1.- Secuencias utilizadas para comparación

<i>Nombre y Tamaño (píxeles)</i>	<i>Imagen de Secuencia</i>	<i>Características principales</i>
<i>Deadline</i> 352x288		Secuencia con el fondo estático y con una gran cantidad de altas frecuencias, mientras que en primer plano se encuentra un individuo que realiza movimientos locales y pequeños.
<i>Flower</i> 352x288		En esta secuencia se aprecia un jardín con una gran cantidad de texturas y unas personas caminando lentamente. Existe mucha textura de fondo y la cámara se desplaza con movimiento horizontal.
<i>Football</i> 352x288		Secuencia con elevado movimiento local y global. Se trata de un partido de fútbol americano con los jugadores pasándose el balón entre ellos.
<i>Foreman</i> 176x144		Esta secuencia presenta un doble movimiento; por una parte un movimiento local de la persona y por otra un movimiento global de la cámara.
<i>Galdar</i> 352x288		Secuencia con movimiento global de cámara. Se presentan varias palabras tipografiadas en un periódico.
<i>Mobile</i> 176x144		Secuencia en la que se aprecia un calendario desplazándose hacia arriba y un tren moviéndose hacia la izquierda con un desplazamiento suave de la cámara. El fondo presenta una textura homogénea, al igual que el calendario.
<i>Suzie</i> 176x144		Secuencia con elevado movimiento local. Se trata de una mujer hablando por teléfono y moviendo su cabeza continuamente. El movimiento local es por tanto elevado.

Tabla A.2.- Secuencias utilizadas para comparación apropiadas para sistemas Multi-Cámara

Nombre	Imagen de Secuencia	Características principales
<i>Mobcal</i> 480x360		Es una secuencia muy parecida a <i>Mobile</i> : un calendario que sube y baja y un tren que se mueve. El movimiento local es sin embargo mayor que en <i>Mobile</i> .
<i>Stockholm</i> 704x576		Panorámica de la ciudad de Estocolmo en la que la cámara realiza un suave desplazamiento horizontal. La secuencia dispone una gran cantidad de textura. Existe un poco de movimiento local de los vehículos.
<i>Shields</i> 768x640		Grabación de un hombre con barba y chaqueta caminando en frente de una pared llena de escudos muy detallados. Al final de la secuencia se realiza una ampliación. Secuencia caracterizada por tener movimiento global medio y mucho detalle.
<i>Parkrun</i> 768x640		Grabación de un hombre corriendo por un parque, con un paraguas. La escena tiene árboles, nieve y agua de fondo. Esta secuencia contiene muchos detalles, una alta textura y movimiento global alto. Así mismo contiene un bajo nivel de movimiento local, producido tan solo por el hombre corriendo por el parque.
<i>Water Cooler</i> 448x320		Secuencia de vídeo grabada por un sistema MC real. En esta secuencia se avanza por el pasillo de una oficina, girando a la altura del dispensador de agua para tomar un primer plano de éste. Posteriormente, se continúa avanzando por una puerta situada al lado de este dispensador hacia un pasillo exterior, que tiene muy poca iluminación. Esta secuencia está grabada cámara en mano, produciéndose un elevado movimiento en la secuencia.
<i>Fujidog</i> 640x480		Imagen tomada mediante un sistema MC real. En la imagen se puede ver en un primer plano una pila de libros, con un perro en la parte superior. De fondo se puede ver una imagen de unas olas (realizada a través de fractales), un tablero de ajedrez y un panel de colores en perspectiva. Con esta distribución de objetos se consigue disponer de imágenes de cada cámara con oclusiones y cambios de perspectiva.

La secuencia *Water Cooler* [SZJ+09] fue grabada por la Universidad de Wisconsin utilizando una cámara ProFusion25, que contiene un total de 25 cámaras, de las cuales para grabar la secuencia *Water Cooler* se han usado cuatro cámaras, que se corresponden con las cámaras superior, inferior, izquierda y derecha situadas en el centro de los bordes del *array* MC, tal y como se muestra en la Figura A.1. En consecuencia, el *array* presenta forma de romboide y dimensiones 2x2. Como características físicas aportadas por los autores se sabe que existe una distancia total de 23.7 mm entre cada cámara y el centro del *array* MC.

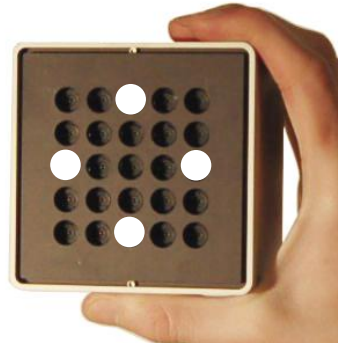


Figura A.1.- Selección de cámaras con que se ha grabado la secuencia Water Cooler

Por otro lado, en el método de SR Espacial se incluye el resultado obtenido a partir de la secuencia *Fujidog*, la cual se corresponde realmente con una foto tomada por cada una de las cámaras en el mismo instante de tiempo, y en consecuencia sólo se puede aplicar un proceso de SR Espacial, el cual generará un *frame* súper-resuelto como resultado. Estas imágenes han sido capturadas con un *array* MC de 81x81 cámaras, las cuales posteriormente han pasado por un proceso de alineación de las cámaras, tras el cual solo quedan las traslaciones entre cámaras. Con el objetivo de simular un *array* MC de tamaño 3x3 se han tomado dos subconjuntos de cámaras de este *array* MC para, posteriormente, usarlo con el fin de realizar simulaciones mediante el proceso de SR Espacial. Un subconjunto de cámaras está situado en la esquina superior izquierda del *array* MC, mientras que el otro se encuentra situado en la parte superior central de este *array* MC, marcado con color azul y rojo en la Figura A.2, respectivamente. El subconjunto resaltado en rojo, permite llevar a cabo una simulación de esta secuencia con una mayor separación entre cámaras, y por tanto un solape menor.

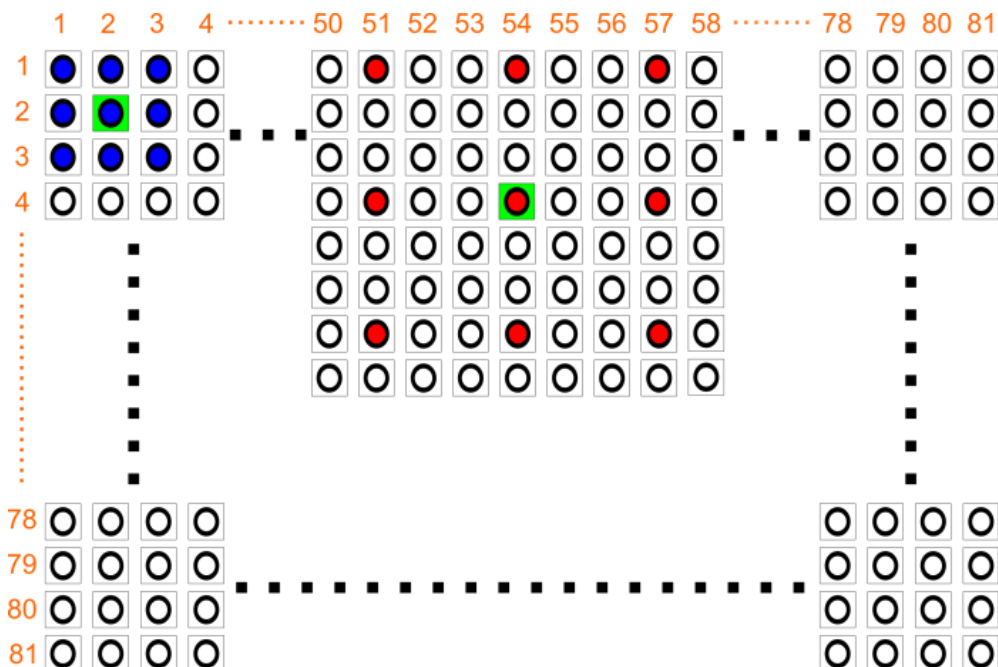


Figura A.2.- Selección de cámaras con que se ha grabado la secuencia Fujidog

Finalmente se presentan las secuencias submarinas usadas en la presente Tesis Doctoral.

1.- *Secuencia Pez_Coral* – 640x360 píxeles – 25 fps

Característica	Evaluación subjetiva
<i>Movimiento global</i>	3
<i>Movimiento local</i>	3
<i>Color</i>	3
<i>Brillo y Contraste</i>	4
<i>Parpadeo de luz solar</i>	1
<i>Turbidez</i>	2
<i>Luz artificial</i>	1

Figura A.3.- Imagen de secuencia *Pez_Coral* Tabla A.3.- Características de secuencia *Pez_Coral*

Se trata de una secuencia con movimiento global y ligero movimiento local. Al estar tomada en fondo marino hay una considerable pérdida de color. No existe parpadeo de la luz solar. La secuencia presenta una ligera turbidez y un brillo y contraste considerables.

2.- *Secuencia Gruta* – 480x360 píxeles – 25 fps

Característica	Evaluación subjetiva
<i>Movimiento global</i>	2
<i>Movimiento local</i>	2
<i>Color</i>	4
<i>Brillo y Contraste</i>	3
<i>Parpadeo de luz solar</i>	4
<i>Turbidez</i>	3
<i>Luz artificial</i>	1

Figura A.4.- Imagen de secuencia *Gruta*

Tabla A.4.- Características de secuencia *Gruta*

Se trata de una secuencia con movimiento global y ligero movimiento local. Debido a que está tomada cerca de la superficie del mar, no hay pérdida de color. El parpadeo de la luz solar es importante, si bien está condensado en una zona determinada, integrándose en el movimiento global de la cámara. La secuencia presenta poca turbidez y un brillo y contraste moderados.

3.- *Secuencia Titanic* – 480x360 píxeles – 25 fps

Característica	Evaluación subjetiva
<i>Movimiento global</i>	4
<i>Movimiento local</i>	2
<i>Color</i>	5
<i>Brillo y Contraste</i>	3
<i>Parpadeo de luz solar</i>	1
<i>Turbidez</i>	4
<i>Luz artificial</i>	5

Figura A.5.- Imagen de secuencia *Titanic*

Tabla A.5.- Características de secuencia *Titanic*

Se trata de una secuencia con movimiento global. Debido a que está tomada en el fondo marino, hay una pérdida de color considerable. La falta de iluminación propicia además el uso

de luz artificial, siendo en este caso el foco fijo, por lo que se integra en el movimiento global de la cámara. La secuencia presenta, además de bastante turbidez en forma de nieve marina, y un brillo y contraste moderados.

4.- *Secuencia Coral – 480x360 píxeles – 25 fps*



Figura A.6.- Imagen de secuencia Coral

Característica	Evaluación subjetiva
<i>Movimiento global</i>	4
<i>Movimiento local</i>	4
<i>Color</i>	2
<i>Brillo y Contraste</i>	2
<i>Parpadeo de luz solar</i>	5
<i>Turbidez</i>	1
<i>Luz artificial</i>	1

Tabla A.6.- Características de secuencia Coral

Se trata de una secuencia con alto movimiento global y movimiento local en forma de parpadeo solar. El resto de características presentan una valoración moderada o baja.

5.- *Secuencia Reef – 480x360 píxeles – 25 fps*



Figura A.7.- Imagen de secuencia Reef

Característica	Evaluación subjetiva
<i>Movimiento global</i>	3
<i>Movimiento local</i>	4
<i>Color</i>	2
<i>Brillo y Contraste</i>	2
<i>Parpadeo de luz solar</i>	4
<i>Turbidez</i>	2
<i>Luz artificial</i>	1

Tabla A.7.- Características de secuencia Reef

Se trata de una secuencia con movimiento global moderado y movimiento local alto en forma de parpadeo solar, y de un pez que se desplaza en la secuencia.

6.- *Secuencia Atlantis – 640x360 píxeles – 30 fps*



Figura A.8.- Imagen de secuencia Atlantis

Característica	Evaluación subjetiva
<i>Movimiento global</i>	3
<i>Movimiento local</i>	1
<i>Color</i>	3
<i>Brillo y Contraste</i>	3
<i>Parpadeo de luz solar</i>	1
<i>Turbidez</i>	1
<i>Luz artificial</i>	5

Tabla A.8.- Características de secuencia Atlantis

Se trata de una secuencia con movimiento global medio y movimiento local nulo. El resto de características presentan una valoración subjetiva moderada o baja salvo la luz artificial.

A.2.- ENTORNO SUBMARINO: CARACTERÍSTICAS Y MEJORA DE IMAGEN

En este anexo, se describen el tipo de características de la imagen que se ven afectadas por el entorno acuático para pasar entonces a estudiar diferentes trabajos de investigación orientados hacia la mejora de calidad de secuencias de vídeo submarinas, proporcionando un pre-procesamiento que podría ser adecuado para maximizar los resultados de la aplicación de algoritmos de SR.

A.2.1.- Características del entorno submarino

A la hora de evaluar un entorno submarino es fundamental considerar las diferencias existentes entre agua dulce y agua salada, ya que es el segundo escenario el que se ha considerado en el ámbito de esta Tesis Doctoral, siendo además el que presenta características más exigentes. El agua de mar es una solución de sales, por lo que sus propiedades físicas son muy diferentes de las del agua dulce y varían de acuerdo con la cantidad de sales que contenga. Por la gran complejidad que presenta el agua de mar en su composición, y debido a su riqueza en seres vivos, sustancias orgánicas en suspensión, y gases disueltos, algunos autores la describen como “una sopa turbia de seres vivos”. A continuación se introducen las principales características del entorno submarino en medio marino.

La transmisión de luz en el medio acuático se ve afectada por varios aspectos, dando lugar a la variación de características de la imagen como el color, el brillo o el contraste. La Figura A.9 presenta la penetración de la luz en agua de mar dependiendo de la longitud de onda.

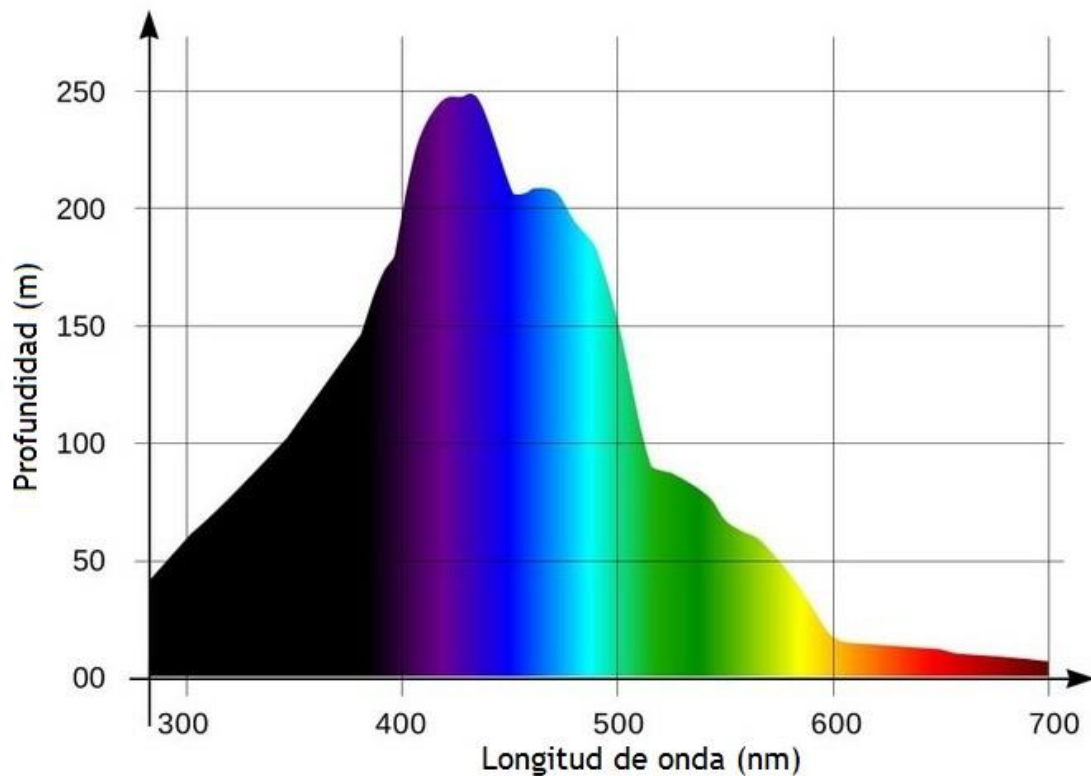


Figura A.9.- Penetración de la luz en agua de mar, dependiendo de la longitud de onda

A.2.1.1 El color

Las radiaciones que forman la luz son absorbidas por el agua de mar y le transmiten calor. Esta absorción es selectiva y depende de la longitud de onda de la radiación. Dentro del espectro visible, la absorción es máxima para el rojo y mínima para los colores verde, y sobre todo azul, como se puede observar en la Figura A.9 [KG12]. En el infrarrojo se produce el transporte de la mayor parte de la energía calorífica, por lo que la absorción se genera prácticamente en el primer metro de agua. La pérdida de luz es además no sólo cuantitativa, sino también cualitativa, ya que no todos los colores que componen la luz visible pierden intensidad de igual forma. Así, el rojo y el violeta son los que se pierden más cerca de la superficie, seguidos del naranja y el amarillo, siendo el verde y el azul los colores que más profundizan, como se puede observar en la Figura A.10 y en la Figura A.11.

Por otra parte, la calidad del agua también influye en la absorción. En aguas claras la longitud de onda larga o luz roja se pierde en primer lugar; el naranja se filtra a continuación, seguido por el amarillo, verde, y luego azul. Otras aguas, aguas costeras en particular, contienen limo, descomposición de materia vegetal y animal, plancton y posibilidad de variedad de contaminantes, que suman su coeficiente de absorción al del agua. El plancton, por ejemplo, absorbe violetas y azules. La absorción de la luz dependerá de la presencia de diversos componentes disueltos en el agua del entorno a estudiar. Algunos de estos componentes son:

- Moléculas de agua.
- Fitoplancton.
- Detritos orgánicos.
- Partículas minerales.
- Materia orgánica disuelta (CDOM, *Colored Dissolved Organic Material*).

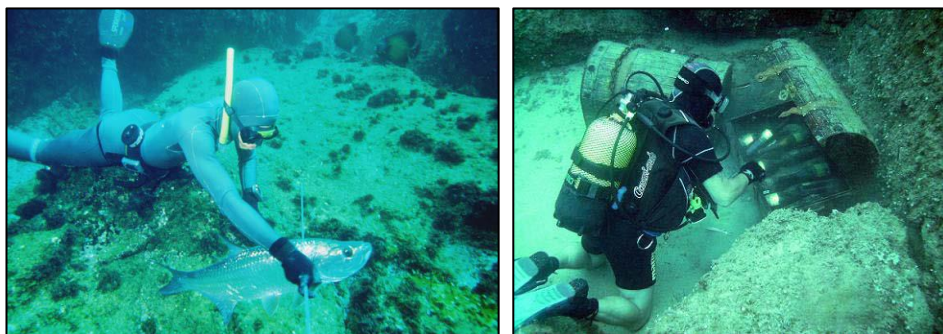


Figura A.10.- Imágenes submarinas con tonalidad azul-verde (absorción del color)



Figura A.11.- Imágenes submarinas con tonalidad azul (absorción del color)

A.2.1.2 El brillo y el contraste

La cantidad de energía lumínica que llega a la superficie del mar depende de la latitud (máxima en el ecuador y mínima en los polos), del ciclo de estaciones (máxima en primavera y verano), del ciclo día-noche y de una serie de factores locales, como nubosidad, contaminación, turbidez del aire, etc. De esta energía incidente, una parte va a ser reflejada por la superficie del mar como consecuencia del fuerte gradiente de densidad existente entre el aire y el agua de mar.

A su vez, la fracción de luz que penetra en el mar sufre una pérdida de intensidad muy marcada a medida que gana en profundidad. Esta pérdida es ocasionada por la absorción y dispersión de los rayos luminosos por las sustancias orgánicas disueltas y las partículas en suspensión. Su magnitud es tal, que en aguas oceánicas a los 100 m de profundidad ya sólo llega el 1% de la luz incidente, mientras que en aguas costeras ese porcentaje se alcanza a una profundidad comprendida entre 10 y 30 m [KG12]. En la Figura A.12 se representa el fenómeno de dispersión de la luz debido a una partícula.

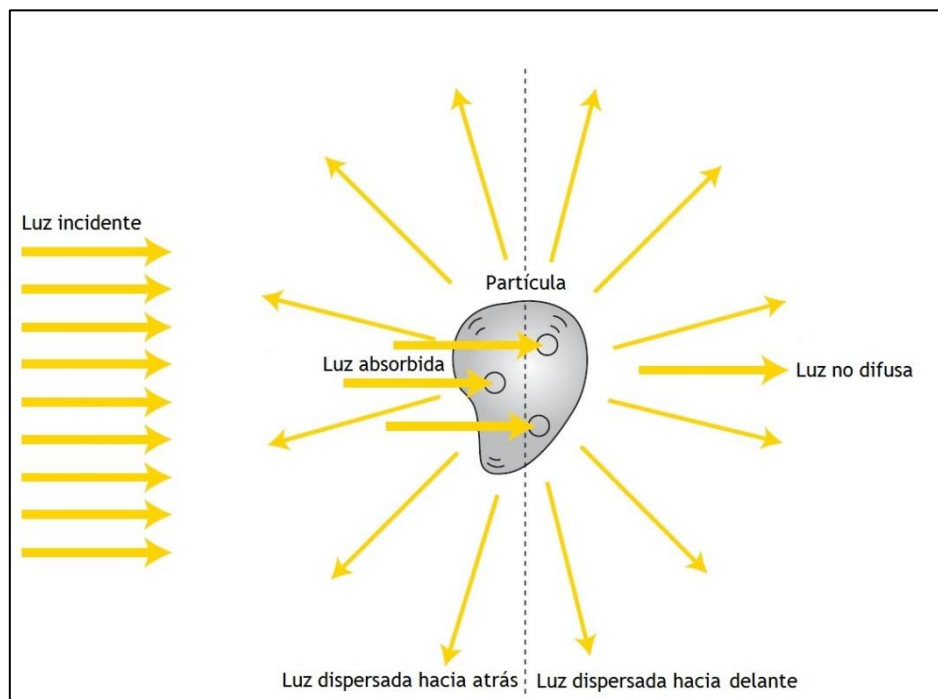


Figura A.12.- Dispersión de la luz por una partícula

Finalmente, las características como el brillo y el contraste de la imagen se verán afectadas por los efectos de refracción y difracción de la luz en el agua:

- El efecto de refracción se produce cuando la onda lumínica pasa de un medio a otro y cambia de dirección, en el caso de la imagen submarina la luz pasa por los siguientes medios: aire-agua-cristal-aire-cristal hasta la lente de la cámara.
- La difracción se produce cuando el haz de luz cambia de dirección al encontrarse con un obstáculo (partículas en suspensión en el agua de tamaño comparable a la longitud de onda de la luz).

Los efectos de la dispersión de la luz en la imagen submarina son el desenfoque y la creación de un velo que esconde la escena, tal y como se representa en la Figura A.13.



Figura A.13.- Efectos de dispersión de la luz en el medio marino

A.2.1.3 Parpadeo de la luz solar

En los sistemas acuáticos hay varios factores que causan variaciones irregulares debido al movimiento de las olas, a la presencia de nubes y al movimiento vertical del fitoplancton.

El fenómeno de fluctuación de la luz, conocido como parpadeo de luz (*light flickering*), es producido por el efecto lente del movimiento del agua en superficie, que simultáneamente enfoca y difumina la luz solar a poca profundidad [SS13], tal y como se representa en la Figura A.14. Los objetos sumergidos y el volumen de agua en sí se iluminan con un patrón aleatorio natural, que es espacial y temporalmente variable. Este fenómeno se ha considerado principalmente como una perturbación significativa en la visión. Una superficie de agua ondulada refracta la luz natural de forma espacialmente variante. Bajo el agua, los rayos refractados crean un patrón espacial de iluminación. Este patrón varía temporalmente, debido al movimiento de las olas en la superficie [SSH+10], tal y como se presenta en la Figura A.15.

Dependiendo del objetivo del estudio, este efecto puede ser considerado perjudicial o no. Hay estudios que utilizan este efecto para crear imágenes más estables y con una iluminación uniforme, mientras que otros utilizan estas variaciones de luz en diferentes puntos para detectar texturas y construir imágenes 3D [SS13], [SSN11].

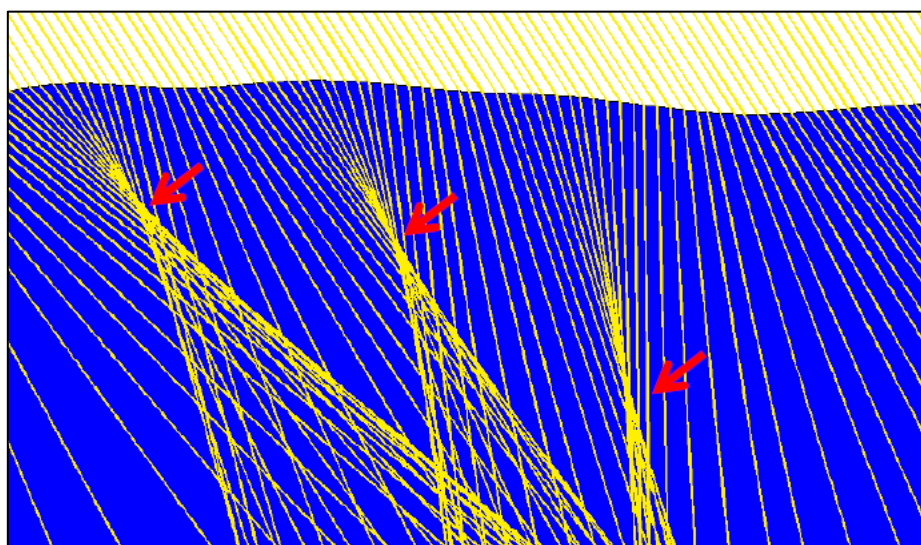


Figura A.14.- Efecto de la refracción de la luz solar a causa de la superficie del mar. Causa del parpadeo (*light flickering*)



Figura A.15.- Efectos del parpadeo de la luz en la imagen submarina

A.2.1.4 Nieve marina

La turbidez del agua depende de la cantidad de partículas o seres vivos de pequeñas dimensiones en suspensión, a este efecto se le conoce como nieve marina “*marine snow*”. La concentración de partículas dependerá de la zona, época del año, temperatura del agua, corrientes oceánicas, etc.

La nieve marina está compuesta principalmente por materia orgánica (incluyendo plantas o animales muertos, bacterias y células), restos fecales y partículas de arena muy fina, entre otros. Esta nieve marina es un proceso por el cual materia orgánica que se genera en las capas altas e iluminadas de la superficie del mar llegan a las capas intermedias e incluso profundas para servir de alimento a otros organismos. El promedio de caída de estas partículas es de 20 metros por día [SSH+10]. La visión submarina se ve afectada aumentando la absorción y dispersión de la luz en el agua, es decir afectando a los colores, brillo y contraste de la imagen. Además introduce en la imagen puntos blanquecinos que distorsionan la visión hasta llegar a tal punto que ocultan el objeto. La Figura A.16 presenta imágenes que presentan el efecto de nieve marina.

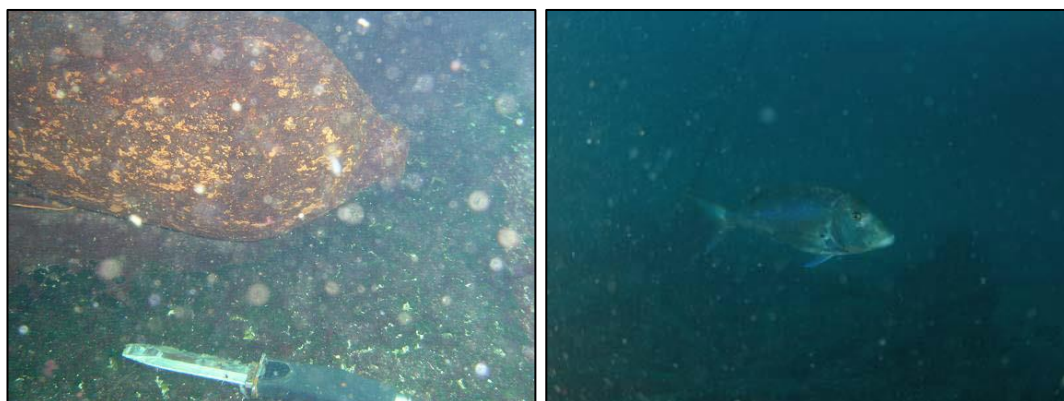


Figura A.16.- Ejemplos del efecto de la turbidez en la imagen submarina

A.2.1.5 Iluminación artificial

La falta de luz, a medida que se avanza en profundidad, hace necesaria la instalación de focos para poder visualizar el entorno. Dicha iluminación artificial provoca una iluminación no uniforme generando una zona muy iluminada en el centro con un área poco iluminada a su alrededor. Además, dependiendo del tipo de foco variará la coloración de la imagen obtenida, por lo que en ocasiones, dependiendo del objetivo de la misión, es interesante utilizar en vez de luz blanca, otro tipo de iluminación. La Figura A.17 muestra imágenes submarinas con iluminación artificial, que presentan una iluminación no uniforme en la imagen.

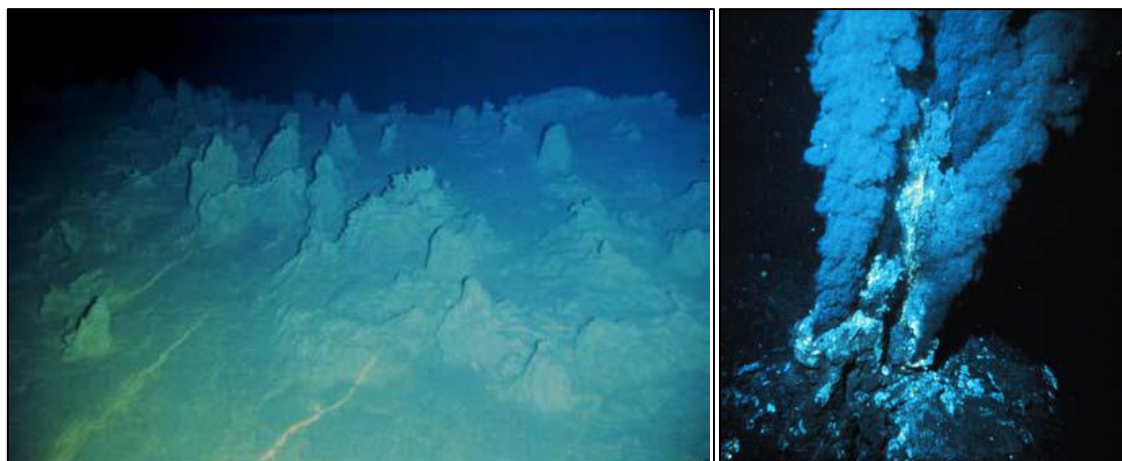


Figura A.17.- Ejemplos del efecto de la iluminación artificial en la imagen submarina

A.2.2.- Mejora de calidad en imagen submarina

El procesamiento de la imagen submarina se puede dividir principalmente en dos técnicas: técnicas de restauración de imagen o métodos de mejora de imagen.

A.2.2.1 Técnicas de restauración de imagen submarina

Estas técnicas tienen por objeto recuperar la imagen original de la imagen observada, usando (si está disponible) el conocimiento explícito acerca de la función de degradación y las características de ruido. Se presentan a continuación algunos estudios previos que trabajan en este sentido.

Miao [Mia12] planteó un proceso de restauración de la imagen subacuática utilizando un algoritmo basado en el modelo de turbulencia atmosférica. Este algoritmo es automático y no requiere ajuste de parámetros y ningún conocimiento a priori de las condiciones de adquisición. Esto se debe a que las funciones evalúan el modelado de parámetros usando métrica de calidad de la imagen. Este método de restauración mejora en gran medida el contraste y la definición de imágenes bajo el agua y también aumenta la calidad visual de la imagen, como se representa en la Figura A.18.

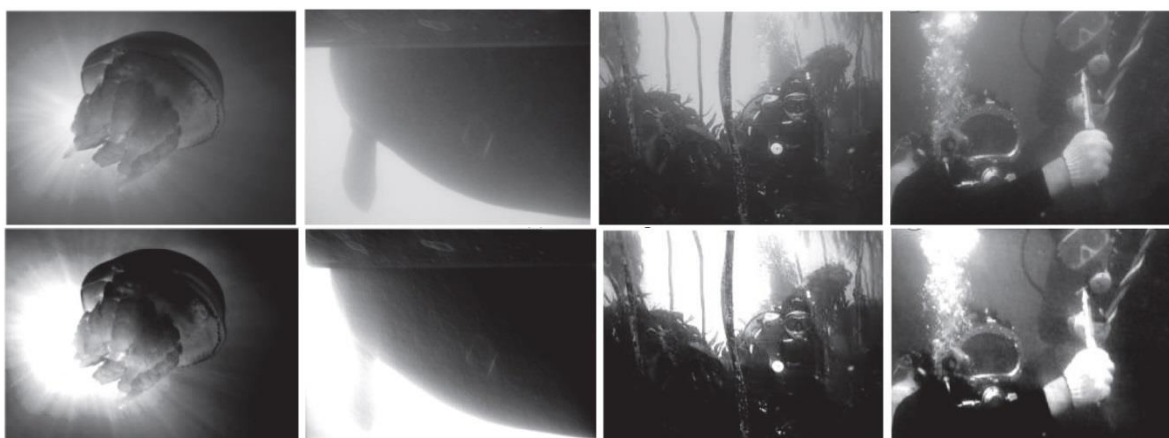


Figura A.18.- Imágenes originales (arriba) y restauradas (abajo) [Mia12]

Schechner y Karpel [SK05] introdujeron un método que presenta el análisis de un par de imágenes adquiridas a través de un filtro de polarización. Dado que se basa en la física de la imagen, el método también recupera información acerca de la estructura de la escena (distancias). El algoritmo no requiere ninguna calibración de los parámetros ambientales y explota la iluminación natural. En este estudio se lograron grandes mejoras de contraste de la escena y la corrección de color, casi duplicando el rango de visibilidad bajo el agua, como se presenta en la Figura A.19. El estudio explota los efectos de polarización en la dispersión bajo el agua para compensar la degradación de la visibilidad. Los autores afirman que la falta de definición de la imagen no es la causa principal de la degradación del contraste de la imagen y establecen la polarización como la principal perturbación en la visibilidad bajo el agua, por lo que es el objetivo a eliminar (efecto denominado luz de retrodispersión).



Figura A.19.- Imagen original (izquierda) y restaurada (derecha) [SK05]

El trabajo de Boffety y Galland [BG12] demuestra que la turbidez o nieve marina es una perturbación importante que puede disminuir fuertemente el rendimiento de los algoritmos de restauración del color. Por ello, se desarrolló un modelo que permite la simulación de imágenes bajo el agua en presencia de nieve marina. A continuación se analizó el comportamiento de la técnica de restauración del color, usando el método del histograma expandido (HSM, *Histogram Stretching Method*) estándar en la presencia de nieve marina, permitiendo aumentar su robustez frente a la presencia de esta perturbación en las aplicaciones con datos reales, como se puede observar en la Figura A.20. La perspectiva inmediata de este trabajo es generalizar este enfoque con el fin de determinar automáticamente cuál es el mejor valor cuantil a utilizar en la versión robusta del algoritmo de restauración del color. Por otra parte, el interés de tal simulador no se limita a la restauración del color y puede ser utilizado para caracterizar la robustez de otros algoritmos de procesamiento de imágenes frente a la presencia de nieve marina.

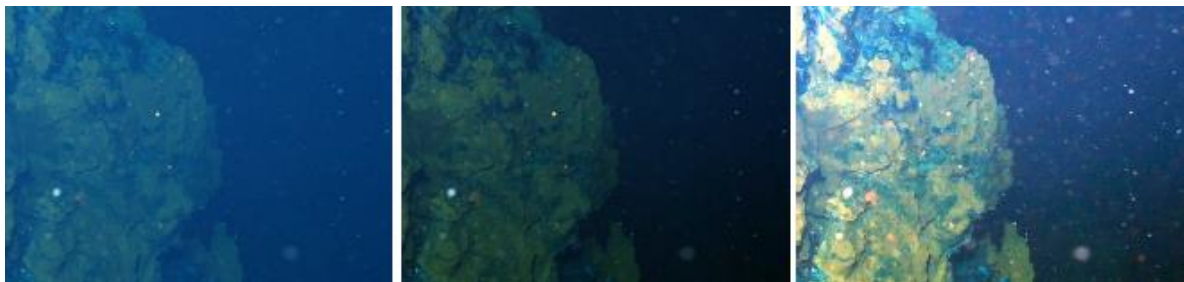


Figura A.20.- Imagen original (izquierda) y restaurada con 2 métodos (centro y derecha) [BG12]

Trucco y Olmos-Antillón [TO06] presentaron un filtro de restauración auto-ajustable basado en una simplificación del modelo Jaffe-McGlamery [McG79] de formación de una imagen submarina. El modelo simplificado es ideal para aguas poco profundas (superficiales), y condiciones de luz difusa con retrodispersión limitada, pero resultados experimentales sugieren que dicha mejoría también se consigue en una variedad de las condiciones de formación de imágenes, como se representa en la Figura A.21. El algoritmo presentado es de auto-sintonización, en el sentido de que los valores óptimos de los parámetros se estiman para cada imagen.

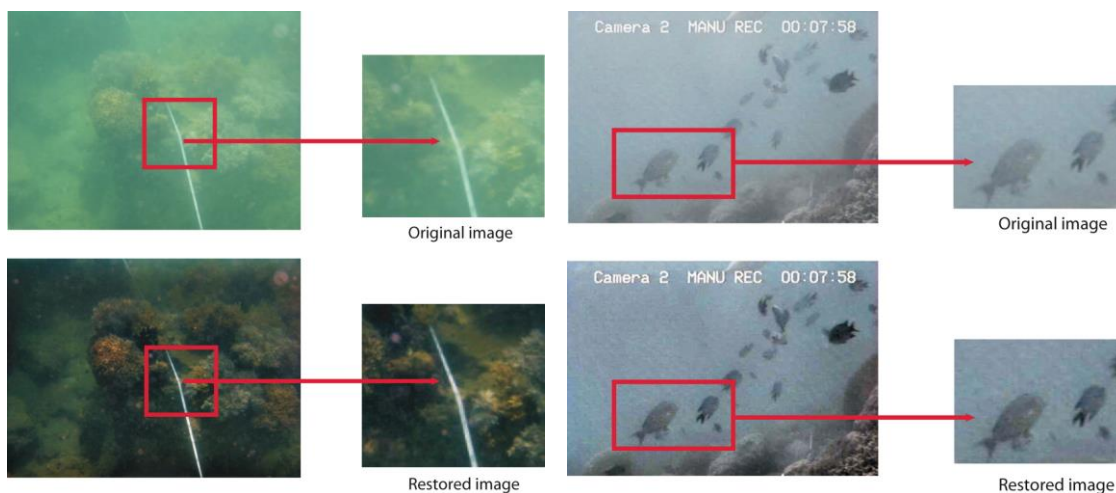


Figura A.21.- Imágenes originales (arriba) y restauradas (abajo) [TO06]

A.2.2.2 Técnicas de mejora de imagen submarina

Estos métodos hacen abstracción total del proceso de formación de la imagen, y sin conocimiento a priori del entorno (no utiliza coeficientes de atenuación y dispersión por ejemplo). Son generalmente más simples y rápidos que las técnicas de restauración de imágenes. Se presentan a continuación algunos de estos trabajos.

Iqbal *et al.* [ISO+07] utilizan un algoritmo de deslizamiento en expansión, tanto en los modelos de color RGB (*Red, Green, Blue*) como HSI (*Hue, Saturation, Intensity*) para mejorar las imágenes submarinas. En primer lugar, se realiza un ajuste del contraste en el modelo de color RGB, para ecualizar el contraste de color de la imagen. En segundo lugar, se realiza la saturación y la expansión de la intensidad en el modelo de color HSI. La ventaja de la aplicación de los dos modelos de expansión es que ayuda a equilibrar el contraste de color en las imágenes y también aborda el problema de la iluminación. La Figura A.22 muestra un ejemplo de mejora del color tras la aplicación de este método a nivel de imagen e histograma.

Yang *et al.* [YCH+11] proponen un método de mejora de la imagen de baja complejidad. El enfoque propuesto contiene principalmente dos procedimientos: un filtro de mediana que se utiliza para estimar la transmisión de la imagen de entrada, y un cálculo de luz atmosférica que se obtiene mediante el uso de un canal pro-oscuridad. Para mejorar aún más la calidad visual, se emplea un método de corrección de color que actúa sobre el contraste de color del objeto bajo el agua. Los resultados experimentales muestran que el método propuesto puede mejorar la eficacia de la imagen bajo el agua, tal y como se presenta en la Figura A.23. Además, este método requiere menos recursos y es muy adecuado para la implementación en vigilancia y navegación subacuática en tiempo real.

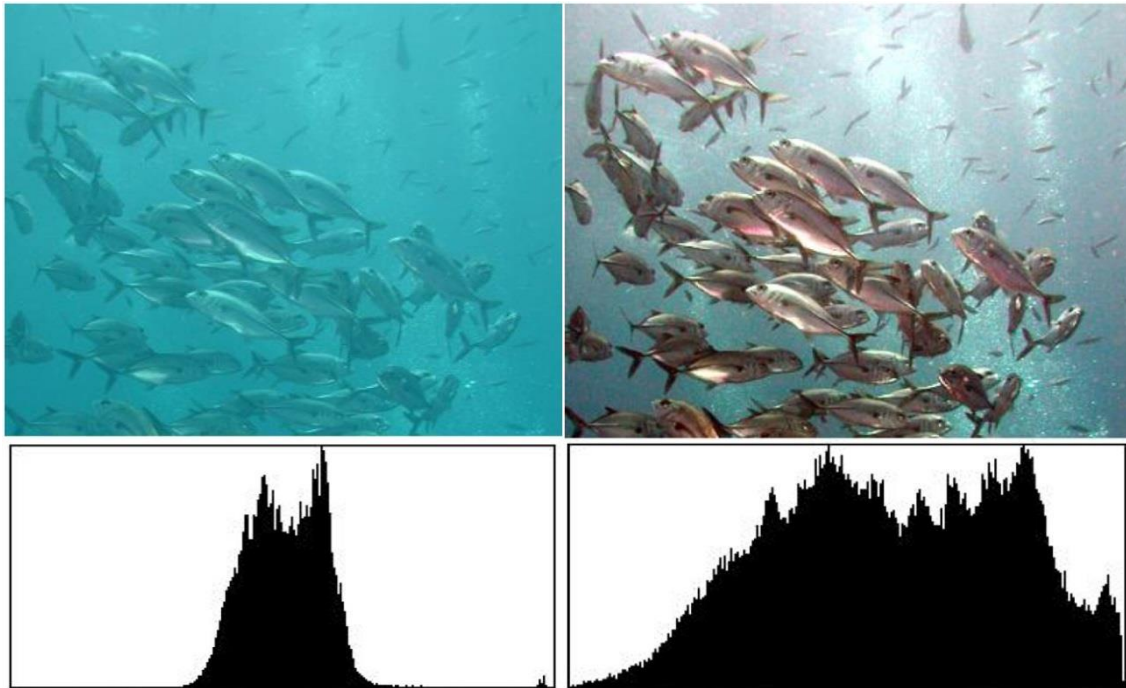


Figura A.22.- Ejemplo de mejora de color. Izquierda: imagen original e histograma, derecha: imagen mejorada e histograma [ISO+07]



Figura A.23.- Imagen original (izquierda) y mejorada (derecha) [YCH+11]


Finalmente, el método presentado por Petit *et al.* [PCC09] mejora el contraste y el color dinámico en las imágenes submarinas cambiando los píxeles de agua a gris o colores de baja saturación mientras que los objetos permanecen totalmente coloreados, como se puede observar en la Figura A.24. Se ha utilizado una transformación geométrica con cuaternios en torno a un eje de matiz, representando el color del agua. El contraste de los objetos mejora con bastante calidad, y los matices azulados de la imagen debido a la atenuación de la luz, se eliminan.



Figura A.24.- Imágenes originales (a, c) y mejoradas (b, d) [PCC09]

A.3.- TEST DE CALIDAD SUBJETIVA: CUESTIONARIO

En este anexo, se presenta el cuestionario de *test* de calidad subjetiva realizado, que se presenta en la Figura A.25, así como los *frames* de *test* utilizados, que se muestran en la Figura A.26.



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
DIVISIÓN DE DISEÑO DE SISTEMAS INTEGRADOS

TEST DE CALIDAD SUBJETIVA

Código: _____

Fecha: _____ Hora: _____ Lugar: _____

Datos para el estudio estadístico

Edad: _____
Sexo: _____
Estudios: _____
Ocupación: _____

Marca la respuesta correspondiente a la imagen que consideres de mayor calidad:

- 1.- *Stockholm* A B
- 2.- *Shields* A B
- 3.- *Parkrun* A B
- 4.- *Mobcal* A B
- 5.- *Reef* A B
- 6.- *Coral* A B

¿Consideras que hay una mejora significativa con respecto a la imagen original? Si es así, ¿en qué imágenes? (escribe nº de imagen)

Test de imagen *Stockholm*





Imagen de entrada al algoritmo



A

¿Cuál es la imagen de mayor calidad?, ¿A o B?
¿Hay una mejora significativa con respecto a la imagen de entrada?


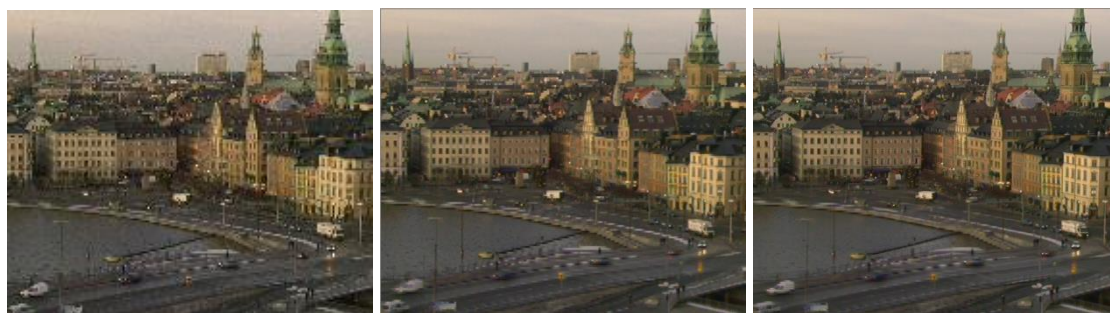


Figura A.25.- Cuestionario de test de calidad subjetiva y ejemplo de pregunta para secuencia *Stockholm*



1.- Stockholm



2.- Shields



3.- Parkrun



4.- Mobcal



5.- Reef



6.- Coral

Figura A.26.- Frames de test, de izquierda a derecha: frame de LR, frame de BSR y frame de MCABMSR

A.4.- PUBLICACIONES

En este anexo, se resumen las publicaciones realizadas en el ámbito de la presente Tesis Doctoral en orden cronológico, además de otras publicaciones desarrolladas como resultado del trabajo realizado en el Programa de Doctorado de Ingeniería de Telecomunicación Avanzada.

A.4.1.- Congresos

[C1] Eduardo Quevedo, Olivia Rodríguez, Gustavo M. Callicó, Félix Tobajas, and Valentín de Armas, "Image Resolution Enhancement in Underwater applications", *27th Conference on Design of Circuits and Integrated Systems (DCIS)*, Avignon, France, Nov. 2012.

[C2] Eduardo Quevedo, David Horat, Gustavo M. Callicó, and Félix Tobajas, "Computation Time Optimization in Super-Resolution applications", *14th International Conference on Computer Aided Systems Theory*, Las Palmas de Gran Canaria, Spain, pp. 110-111, Feb. 2013.

[C3] Jorge Rodríguez, Alexis Quesada, David Horat, and Eduardo Quevedo "Web Georeferenced Video Player with Super-Resolution Screenshot Feature", *14th International Conference on Computer Aided Systems Theory*, Las Palmas de Gran Canaria, Spain, pp. 104-105, Feb. 2013.

[C4] Eduardo Quevedo, Luis Sánchez, Gustavo M. Callicó, Félix Tobajas, Jesús de la Cruz, and Roberto Sarmiento, "Selective filter with adaptive size macro-block for Super-Resolution applications", *The 17th IEEE International Symposium on Consumer Electronics*, Hsinchu, Taiwan, pp. 101-102, June 2013.

[C5] Eduardo Quevedo, Jesús de la Cruz, Gustavo M. Callicó, Félix Tobajas, and Roberto Sarmiento, "Spatial-Temporal Video Enhancement using Super-Resolution from a Multi-Camera System", *Institute of Electrical and Electronic Engineers (IEEE) International Conference on Consumer Electronics*, Las Vegas, USA, pp. 536-537, Jan. 2014.

[C6] E. Quevedo, J. Rodríguez, D. Horat. A. Quesada-Arencia, F. Tobajas, G.M. Callicó, and R. Sarmiento, "Improving underwater video navigation systems using Georeferencing and Super-Resolution techniques", *Institute of Electrical and Electronic Engineers (IEEE) Marine Technology Society (MTS) OCEANS conference*, Taipei, Taiwan, pp. 1-7, Apr. 2014.

[C7] T. Szydzik, E. Quevedo, G. M. Callicó, A. Nunez, F. Tobajas, and R. Sarmiento, "Optimization of non-uniform grid projection image super-resolution algorithms by reduced granularity and modified addressing", *29th Conference on Design of Circuits and Integrated Systems (DCIS)*, Madrid, Spain, Nov. 2014.

[C8] Eduardo Quevedo, Jesús de la Cruz, Luis Sánchez, Gustavo M. Callicó, and Félix Tobajas, "Variable Size Block-Matching Super-Resolution Applied to a Multi-Camera System", *Institute of Electrical and Electronic Engineers (IEEE) International Conference on Consumer Electronics*, Las Vegas, USA, pp. 662-663, Jan. 2015.

A.4.2.- Revistas

[R1] Eduardo Quevedo, David Horat, Gustavo M. Callicó, and Félix Tobajas, "Computation Time Optimization in Super-Resolution applications", *Lecture Notes in Computer Science*, vol. 8112, pp. 101-108, Dec. 2013.

[R2] Jorge Rodríguez, Alexis Quesada Arencibia, David Horat, and Eduardo Quevedo "Web Georeferenced Video Player with Super-Resolution Screenshot Feature", *Lecture Notes in Computer Science*, vol. 8112, pp. 87-92, Dec. 2013.

[R3] E. Quevedo, J. de la Cruz, G. M. Callicó, F. Tobajas and R. Sarmiento, "Video Enhancement using Spatial and Temporal Super-Resolution from a Multi-Camera System", *IEEE Transactions on Consumer Electronics*, JCR = 1.157 , vol. 60, issue 3, pp. 420-428, Aug. 2014.

[R4] E. Quevedo, L. Sánchez, G. M. Callicó, F. Tobajas, J. de la Cruz, V. de Armas, and R. Sarmiento, "Super-Resolution with selective filter based on adaptive window and variable macro-block size", *Journal on Real-Time Image Processing*, JCR = 1.111 , 10.1007/s11554-015-0489-3, Feb. 2015.

[R5] E. Quevedo, E. Delory, G. M. Callicó, F. Tobajas, and R. Sarmiento, "Underwater Video Enhancement using Multi-Camera Super-Resolution", *IEEE Journal of Oceanic Engineering*, JCR = 1.325, submitted, Feb. 2015.

[R6] E. Quevedo, J. de la Cruz, L. Sánchez, G. M. Callicó, and F. Tobajas, "Super Resolution with Adaptive Macro-Block Topology Applied to a Multi Camera System", *IEEE Transactions on Consumer Electronics*, JCR = 1.157, submitted after minor review, Mar. 2015.

A.4.3.- Otras publicaciones

Por ultimo, se presentan otras publicaciones desarrolladas como resultado del trabajo realizado en el Programa de Doctorado de Ingeniería de Telecomunicación Avanzada. Estos trabajos se han desarrollado en colaboración con la división MAGiC (MATemáticas Gráficas y Computación) del IUMA.

[C9] F. Perdomo, E. Quevedo, J.P Suárez, and A. Plaza, "Results on the convergence of the longest-edge trisection method for triangles", *Meeting on Applied Scientific Computing and Tools (MASCOT)*, Las Palmas de Gran Canaria, Spain, Oct. 2010.

[C10] F. Perdomo, A. Plaza, E. Quevedo, and J.P Suárez, "A lower bound on the angles of triangles constructed by LE-trisection", XIV Spanish Meeting on Computational Geometry, Madrid, Spain, June 2011.

[C11] J.P Suárez , T. Moreno, E. Quevedo, F. Perdomo, A. Plaza, M.A. Padrón, and P. Abad, "Results on the Convergence of the Longest-Edge Trisection Method for Tetrahedral Meshes", *Meeting on Applied Scientific Computing and Tools - International Society for Grid Generation (MASCOT-ISGG)*, Las Palmas de Gran Canaria, Spain, Oct. 2012.

[C12] E. Quevedo, J.P. Suárez, F. Perdomo, and A. Plaza, "Quality Assessment in Longest-Edge Refinement Schemes for Real-Time Terrain Triangulations", (*Meeting on Applied Scientific Computing and Tools - International Society for Grid Generation (MASCOT-ISGG)*), Las Palmas de Gran Canaria, Spain, Oct. 2012.

[R7] Francisco Perdomo, Ángel Plaza, Eduardo Quevedo, and José P. Suárez, “A mathematical proof of how fast the diameters of a triangle mesh tend to zero after repeated trisection”, *Mathematics and Computers in Simulation*, vol. 106, pp. 95-108, JCR = 0.856, Aug. 2014.

Bibliografía

Mientras haya libros, no existe el pasado
Edward George Bulwer-Lytton (1803-1873), escritor inglés.

[ABG+14] W. van Aarle, K.J. Batenburg, G. van Golpel, E. van de Castele, and J. Sijbers. Super-Resolution for Computed Tomography Based on Discrete Tomography. *IEEE Transactions on Image Processing*, 23(3):1181-1193, 2014.

[ABH+00] M. S. Alam, J. G. Bognar, R. C. Hardie, and B. J. Yasuda. Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames. *IEEE Transactions on Instrumentation and Measurement*, 49(5):915–923, 2000.

[AEH+14] H. Ashikaga, H.L. Estner, D. A. Herzka, E. R. Mcveigh, and H. R. Halperin. Quantitative Assessment of Single-Image Super-Resolution in Myocardial Scar Imaging. *IEEE Journal of Transnational Engineering in Health and Medicine*, 2:1-12, 2014.

[AS12] Yıldıray Anagün and Erol Seke. Super Resolution Using Variable Size Block-Matching Motion Estimation With Rotation. Super resolution using variable size block-matching motion estimation with rotation. In *Proceedings of the International Symposium on Innovations in Intelligent Systems and Applications*, pp.1-5, 2012.

[Ban09] Vivek Bannore, “*Iterative-Interpolation Super-Resolution Image Reconstruction. A Computationally Efficient Technique*”, Springer, 2009.

[BB08] O. Bowen and C. S. Bouganis. Real-time image super resolution using an FPGA. In *Proceedings of International Conference on Field Programmable Logic and Applications*, pp. 89–94, 2008.

[BCL+05] D. Barreto, G. Callicó, S. López, L. García, and A. Núñez, “Real-time super-resolution over raw and compressed video sequences”, *International Society for Optics and Photonics (SPIE) Microtechnologies for the New Millenium conference*, Seville, Spain, vol. 5837, pp. 628–637, May 2005.

[BCM05] A. Buades, B. Coll, and J. M. More. A non-local algorithm for image denoising. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 60–65, 2005.

- [BG12] M. Boffety and F. Galland. Phenomenological marine snow model for optical underwater image simulation: Applications to color restoration. In *Proceedings of IEEE OCEANS International Conference*, pp. 1-6, 2012.
- [BK02] S. Baker and T. Kanade. Limits on super-resolution and how to break them. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1167–1183, 2002.
- [Bro92] L. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, 1992.
- [BS98] Sean Borman and Robert L. Stevenson. Super-resolution from image sequences - A review. In *Proceedings of Midwest Symposium on Circuits and Systems*, pp 374–378, 1998.
- [BSB+09] N. Bellotto, E. Sommerlade, B. Benfold, C. Bibby, I. Reid, D. Roth, C. Fernandez, L. Van Gool, and J. Gonzalez. A distributed camera system for multi-resolution surveillance. In *Proceedings of Third ACM/IEEE International Conference on Distributed Smart Cameras pp. 1-8*, 2009.
- [BT06] H.H. Baker and D. Tanguay. A Multi-imager Camera for Variable-Definition Video. In *Proceedings of the 2006 International Workshop on Classification and Security in Multimedia Content Representation*, pp. 594-601, 2006.
- [BT14] Richard L. Baer and Damien J. Thivent. Super-Resolution based on optical image stabilization, *Patent Application Publication US 2014/0125825 A1*, 2014.
- [Can06] E. Candes. Compressive sensing. In *Proceedings of International Congress of Mathematicians*, vol. 3, pp. 1433–1452, 2006.
- [CB00] M. C. Chiang and T. E. Boult. Efficient super-resolution via image warping. *Image and Vision Computing*, 18(10):761–771, 2000.
- [CLG+09] G. M. Callicó, Sebastian Lopez, Salvador Gabarda, Elena Gil, Gabriel Cristóbal, José López, and Roberto Sarmiento. Anisotropic Quality Measurement applied to H.264 Video Compression. In *Proceedings of the International Society for Optics and Photonics – VLSI Circuits and Systems*, vol. 7363, 2009.
- [CLL+06] G. M. Callicó, R. L. Llopis, S. López Suárez, J. F. López. A. Núñez, R. Sethuraman, and R. Sarmiento. Low-Cost.super-Resolution Algorithms Implementation over a HW/SW Video Compression Platform. *EURASIP Journal on Applied Signal Processing. Special Issue on Super-Resolution*, 2006(10):1-29, 2006.
- [CLN+02] G.M. Callicó, R. Peset Llopis, A. Núñez, R. Sethuraman, and Marc Op de Beek “A Low-Cost Implementation of Super-Resolution based on a Video Encoder”, *28th Annual Conference of the IEEE Industrial Electronics Society (IECON)*, Seville, Spain, vol. 2, pp. 1439-1444, Nov. 2002.
- [CLS+08] G. M. Callicó, S. López Suárez, O. Sosa, J. López, and R. Sarmiento. Analysis of Fast Block Matching Motion Estimation Algorithms for Video Super-Resolution Systems. *IEEE Transactions on Consumer Electronics*, 54(3): 1430-1438, 2008.

- [CLT+08] G.M. Callicó, S. López, K. Tarajano, J. López, and R. Sarmiento. "Impact of Fast Motion Estimation Algorithms on Super-Resolved Video Sequences", *International Conference on Digest of Technical Papers, Consumer Electronics*, Las Vegas, USA, pp. 1-2, Jan. 2008.
- [CTH03] G. Caner, A.M. Tekalp, and W. Heinzelman. Super resolution recovery for multi-camera surveillance imaging. In *Proceedings of the 2003 International Conference on Multimedia and Exposition*, pp. 109-12, 2003.
- [CY13] Yuzhang Chen and Kecheng Yang. MAP-regularized robust reconstruction for underwater image detection. *International Journal for Light and Electron Optics*, 124(20):4514-4518.
- [CYX04] H. Chang, D. Y. Yeung, and Y. Xiong. Super-resolution through neighbor embedding. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 275–282, 2004.
- [CYX+12] Yuzhang Chen, Bofei Yang, Min Xia, Wei Li, Kecheng Yang, and Xiaohui Zhang. Model-based super-resolution techniques for underwater imaging. In *Proceedings of Photonics and Optoelectronics Meetings: Optoelectronic Sensing and Imaging*, vol.8332, 2012.
- [CZ01] D. Capel and A. Zisserman. Super-resolution from multiple views using learnt image models. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 627–634, 2001.
- [DE07] D. Datsenko and M. Elad. Example-based single document image superresolution: a global MAP approach with outlier rejection. *Multidimensional System and Signal Processing*, 18(2-3):103–121, 2007.
- [DGE+00] N. Damera-Venkata, T.D. Kite, W.S. Geisler, B.L. Evans, and A.C. Bovik. Image quality assessment based on degradation model. *IEEE Transactions on Image Processing*, 9(4):636-650, 2000.
- [Don06a] D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [Don06b] D. L. Donoho. For most large underdetermined systems of linear equations, the minimal 1-norm near-solution approximates the sparsest near-solution. *Communications on Pure and Applied Mathematics*, 59(7):907–934, 2006.
- [DSC04] M. Directo, S. Shirani, and D. Capson. Wireless camera network for image superresolution. In *Proceedings of 17th Canadian Conference on Electrical and Computer Engineering*, pp.681-684, 2004.
- [ED07] M. Elad and D. Datsenko. Example-based regularization deployed to super-resolution reconstruction of a single image. *The Computer Journal*, 52(1):15–30, 2007.
- [EF97] M. Elad and A. Feuer. Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images. *IEEE Transactions on Image Processing*, 6(12): 1646–1658, 1997.
- [EFR+04] M. Elad S. Farsiu, D. Robinson, and P. Milanfar. Advances and challenges in super-resolution. *International Journal of Imaging Systems and Technology*, 14(2):47–57, 2004.

- [EH01] M. Elad and Y. Hel-Or. A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur. *IEEE Transactions on Image Processing*, 10(8):1187–1193, 2001.
- [EHD+05] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouki, B. M. Salam, and F. E. Abd El-Samie. Regularized Super-Resolution Reconstruction of Images Using Wavelet Fusion. *Journal of Optical Engineering*, 44(9):097001, 2005.
- [EHD+06] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouki, B. M. Salam, and F. E. Abd El-Samie. Wavelet Fusion: A Tool to break the limits on LMMSE Image Super-Resolution. *International Journal of Wavelets, Multiresolution and Information Processing*, 04(105), 2006.
- [EHD+08a] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouki, B. M. Salam, and F. E. Abd El-Samie. New Techniques to Conquer the Image Resolution Enhancement Problem. *Progress In Electromagnetics Research B*, 7:13-51, 2008.
- [EHD+08b] S. E. El-Khamy, M. M. Hadhoud, M. I. Dessouki, B. M. Salam, and F. E. Abd El-Samie. A Wavelet-Based Entropic Approach to High-Resolution Reconstruction of Images. *Journal on Machine Graphics & Vision International*, 17(4):389-402, 2008.
- [EHK13] Fathi E. abd El-Samie, Mohiy M. Hadhoud, and Said E. El-Khamy, “*Image Super-Resolution and Applications*”, CRC Press, 2013.
- [EL99] A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *Proceedings of IEEE International Conference on Computer Vision*, pp. 1033–1038, 1999.
- [ESO+07] A. W. M. van Eekeren, K. Schutte, O. R. Oudegeest, and L. J. van Vilet. Performance evaluation of super-resolution reconstruction methods on real-world data. *EURASIP Journal on Advances in Signal Processing*, 2007.
- [EST97] P. E. Eren, M. I. Sezan, and A. M. Tekalp. Robust, object-based high resolution image reconstruction from low-resolution video. *IEEE Transactions on Image Processing*, 6(10):1446–1451, 1997.
- [Fan09] M.H. Fanaswala. *Regularized Super-Resolution of Multi-View Images*, M.S. thesis. Department of Systems and Computers Engineering, Carleton Institute for Electrical and Computers Engineering, Ottawa, Ontario, 2009.
- [FJP02] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based superresolution. *IEEE Computer Graphics and Applications*, 22(2):56–65, 2002.
- [FRE+04] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar. Fast and robust multi-frame super-resolution. *IEEE Transactions on Image Processing*, 13(10):1327–1344, 2004.
- [GBA+03] B. K. Gunturk, A. U. Batur, Y. Altunbasak, M. H. Hayes, and R. M. Mersereau. Eigenface-domain super-resolution for face recognition. *IEEE Transactions on Image Processing*, 12(5):597–606, 2003.
- [GC07] S. Gabarda and G. Cristóbal. Blind image quality assessment through anisotropy. *J. Opt. Soc. Am. A*, 24(12):B42–B51, 2007.

[Goh13] S. Gohshi. Limitation of Super Resolution Image Reconstruction for Video. In *Proceedings of International Conference on Computational Intelligence, Communication Systems and Networks*, pp. 217-221, 2013.

[GPR15] GoPro [Online] Available: <http://es.gopro.com/>. Last visit: March 2015.

[GSH+11] T. GoTo, S. Suzuki, S. Hirano, M. Sakurai and T. Q. Nguyen. Fast and high quality learning-based super-resolution utilizing TV regularization method. In *Proceedings of IEEE International Conference on Image Processing*, pp. 1185-1188, 2011.

[Had23] J. Hadamard, *Lectures on Cauchy's Problem in Linear Partial Differential Equations*, Yale University Press, New Haven CT, 1923.

[Har07] R. Hardie. A fast image super-resolution algorithm using an adaptive Wiener filter. *IEEE Transactions on Image Processing*, 16(12):2953–2964, 2007.

[HBA97] R. C. Hardie, K. J. Barnard, and E. E. Armstrong. Joint MAP registration and high resolution image estimation using a sequence of undersampled images. *IEEE Transactions on Image Processing*, 6(12):1621–1633, 1997.

[HBK08] P. H. Hennings-Yeomans, S. Baker, and B. V. K. V. Kumar. Simultaneous super-resolution and feature extraction for recognition of low resolution faces. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.

[HE04] B. C. Hansen and E. A. Essock. A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes, *J. Vision* 4, 1044–1060, 2004.

[HJO+01] A. Hertzmann, C. E. Jacobs, N. Oliver, B. Curless, and D. H. Salesin. Image analogies. In *Proceedings of the 28th annual conference Computer Graphics and Interactive Techniques*, pp. 327–340, 2001.

[HMG08] J. Hyun Park, H. Mook Oh, and M. Gi Kang. Multi-Camera imaging system using Super-Resolution. In *Proceedings of the International Technical Conference on Circuits/Systems, Computers and Communications*, 2008.

[HYW+10] T. S. Huang J. Yang, J. Wright and Y. Ma. Image Super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 19(11):2861–2873, 2010.

[IDS15] IDS [Online] Available: <https://en.ids-imaging.com/>. Last Visit: March 2015.

[ISO+07] Kashif Iqbal, Rosalina Abdul Salam, Azam Osman, and Abdullah Zawawi Talib. Underwater Image Enhancement Using An Integrated Colour Model. *IAENG International Journal of Computer Science*, 32(2):239-244, 2007.

[IP90] M. Irani and S. Peleg. Super resolution from image sequences. In *Proceedings of 10th International Conference on Pattern Recognition*, vol. 2, pp. 115–120, 1990.

[IP91] M. Irani and S. Peleg. Improving resolution by image registration. *Graphical Models and Image Processing*, vol. 53, pp. 231–239, 1991.

- [ITU98] ITU-T Tech. Rec. P.911. *Subjective audiovisual quality assessment methods for multimedia applications*, 1998.
- [ITU02] ITU-R Tech. Rec. BT 500.11. *Methodology for the subjective assessment of the quality of television pictures*, 2002.
- [ITU08] ITU-T Tech. Rec. P.910. *Subjective video quality assessment methods for multimedia applications*, 2008.
- [JG05] K. Jia and S. Gong. Multi-model tensor face for simultaneous superresolution and recognition. In *Proceedings of IEEE International Conference on Computer Vision*, vol. 2, pp. 1683–1690, 2005.
- [JG08] K. Jia and S. Gong. Generalized face super-resolution. *IEEE Transactions on Image Processing*, 17(6):873–886, 2008.
- [Jol02] I. T. Jolliffe. *Principal Component Analysis*. Series: Springer Series in Statistics, 2002.
- [KBV90] S. P. Kim, N. K. Bose, and H. M. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframe. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(6):1013–1027, 1990.
- [KBV93] N. K. Bose, H. C. Kim, and H. M. Valenzuela. Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframe. In *Proceedings of the IEEE Conference on Acoustics, Speech and Signal Processing*, volume 5, pages 269–272, 1993.
- [KC00] M. S. Keil and G. Cristóbal. Separating the chaff from the wheat: possible origins of the oblique effect, *J. Opt. Soc. Am. A* 17, 697–710, 2000.
- [KG12] Dmitry Karpenko, and Vahe Ganapetyan. Light in the Reef Aquaria
[Online] Available: <http://www.advancedaquarist.com/2012/10/aafeature>
Last visit: March 2015.
- [KH08] Yun-Suk Kang and Yo-Sung Ho. Geometrical compensation for multi-view video in multiple camera array. In *Proceedings of the 50th International Symposium ELMAR*, pp.83-86, 2008.
- [KJ12] J. Korhonen and Y. Junyong, Peak Signal-To-Noise Ratio revisited: is simple beautiful?. *Fourth International Workshop on Quality of Multimedia Experience (QoMEX)*, 2012.
- [KLH08] Yun-Suk Kang, Cheon Lee and Yo-Sung Ho. An Efficient Rectification Algorithm for Multi-View Images in Parallel Camera Array. In *Proceedings of 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video*, pp.61-64, 2008.
- [KPB88] D. Keren, S. Peleg, and R. Brada. Image sequence enhancement using subpixel displacements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 742–746, 1988.
- [KSS+12] Y. Kawamoto, S. Suzuki, Y. Sakuta, T. Goto, and M. Sakurai. A study on fast learning-based super-resolution utilizing TV regularization for HDTV. In *Proceedings of IEEE International Conference on Consumer Electronics*, pp. 725-726, 2012.

- [KW93] H. Knutsson and C.-F. Westin. Normalized and differential convolution. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 515–523, 1993.
- [LB02] S. Lertrattanapanich and N. K. Bost. High resolution image formation from low resolution frames using delaunay triangulation. *IEEE Transaction on Image Processing*, 11(12):1427–1441, 2002.
- [LCT+09] S. López, G. M. Callicó, F. Tobajas, J. F. López, and R. Sarmiento. A Novel Real-Time DSP-Based video Super-Resolution System. *IEEE Transactions on Consumer Electronics*, 55(4): 2264-2270, 2009.
- [LPF03] B. Li, M. R. Peterson, and R. D. Freeman, Oblique effect: a neural bias in the visual cortex, *J. Neurophysiol.* 90, 204–217, 2003.
- [LS99] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.
- [LS04] Z. Lin and H.-Y. Shum. Fundamental limits on reconstruction-based superresolution algorithms under local translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):83–97, 2004.
- [LSF07] C. Liu, H. Y. Shum, and W. T. Freeman. Face hallucination: theory and practice. *International Journal of Computer Vision*, 75(1):115–134, 2007.
- [LSZ01] C. Liu, H. Y. Shum, and C. S. Zhang. Two-step approach to hallucinating faces: global parametric model and local nonparametric model”. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 192–198, 2001.
- [LTB12] A. Le, N. Thakoor, and B. Bhanu. Vehicle logo super-resolution by canonical correlation analysis. In *Proceedings of 19th International Conference on Image Processing*, pp. 2229-2232, 2012.
- [LV96] C. J. Branden Lambrecht and O. Verscheure. Perceptual Quality Measure using a Spatio-Temporal Model of the Human Visual System. In *Proceedings of SPIE (International Society for Optics and Photonics)*, vol. 2668, pp. 450-461, 1996.
- [Mar03] G. Marrero Callicó, *Mejora de la calidad en imagen estática y video basada en Súper-Resolución con prestaciones de tiempo real y bajo coste mediante codificación híbrida (Tesis Doctoral)*, Universidad de Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, España, 2003.
- [Mcg79] B. J. McGlamery. A computer model for underwater camera systems. *Oceans Optics VI*, vol. 28, 1979.
- [Mia12] Yang Miao. Underwater image adaptive restoration and analysis by turbulence model. In *Proceedings of World Congress on Information and Communication Technologies*, pp. 1182-1187, 2012.
- [Mil11] P. Milanfar, *Super-Resolution Imaging*. Ed. CRC Press, 2011.

- [MWS06] D. Marpe, T. Wiegand, and G. Sullivan. The H.264/MPEG4 advanced video coding standard and its applications. *IEEE Communications Magazine*, vol. 44, nº 8, pp. 134-143, 2006.
- [NHB+07] B. Narayanan, R. C. Hardie, K. E. Barner, and M. Shao. A computationally efficient super-resolution algorithm for video processing using partition filters. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(5):621–634, 2007.
- [NM00] N. Nguyen and P. Milanfar. An efficient wavelet-based algorithm for image super-resolution. In *Proceedings of International Conference on Image Processing*, vol. 2, pp. 351–354, 2000.
- [OFY+09] M. Ota, N. Fukushima, T. Yendo, M. Tanimoto, and T. Fujii. Rectification of Pure Translation 2D Camera Array. In *Proceedings of the International Workshop on Advanced Image Technology*, pp.245-249, 2009.
- [OII+14] H. Okuhata, R. Imai, M. Ise, R . Y. Omaki, H. Nakamura, S. Hara, and I. Shirakawa. Implementation of dynamic-range enhancement and super-resolution algorithms for medical image processing. In *Proceedings of IEEE International Conference on Consumer Electronics*, pp. 181-184, 2014.
- [PA01] A. J. Patti and Y. Altunbasak. Artifact reduction for set theoretic super resolution image reconstruction with edge adaptive constraints and higher-order interpolants. *IEEE Transactions on Image Processing*, 10(1):179–186, 2001.
- [Pap77] A. Papulis. Generalized sampling expansion. *IEEE Transactions on Circuits and Systems*, 24(11):652–654, 1977.
- [PCC09] F. Petit, A. Capelle-Laize, and P. Carré. Underwater image enhancement by attenuation inversion with quaternions. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1177-1180, 2009.
- [PET+09] M. Protter, M. Elad, H. Takeda, and P. Milanfar. Generalizing the nonlocal-means to super-resolution reconstruction. *IEEE Transactions on Image Processing*, 18(1):36–51, 2009.
- [PFS15] Point Grey: Innovation in imaging. ProFusion25-C [Online]
Available: <http://www.ptgrey.com/KB/10134>
Last Visit: March 2015.
- [PNN15] Panoramic Ball Camera: Panono [Online] Available: <http://www.panono.com/>
Last visit: March 2015.
- [PPK03] Sung C. Park, Min K. Park, and Moon G. Kang. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20(3):21–36, 2003.
- [PRZ03] L. C. Pickup, S. J. Robert, and A. Zisserman. A sampled texture prior for image super-resolution. In *Proceedings of Advances in Neural Information and Processing System*, pp. 1587–1594, 2003.
- [PST94] A. J. Patti, M. I. Sezan, and A. M. Tekalp. High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur. In *Proceedings of the IEEE International Conference on Image Processing*, vol. 1, pp. 343–347, 1994.

- [PST97a] A. J. Patti, M. Sezan, and A. M. Tekalp. Robust methods for high quality stills from interlaced video in the presence of dominant motion. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(2):328–342, 1997.
- [PST97b] A. J. Patti, M. I. Sezan, and A. M. Tekalp. Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time. *IEEE Transactions on Image Processing*, 6(8):1064–1076, 1997.
- [PVS06] T. Q. Pham, L. J. Vliet, and K. Schutte. Robust fusion of irregularly sampled data using adaptive normalized convolution. *EURASIP Journal on Applied Signal Processing*, 2006.
- [PW04] M. Pinson and S. Wolf. A New Standardized Method for Objectively Measuring Video Quality. *IEEE Transactions on Broadcasting*, 50(3):312-322, 2004.
- [RM04] D. Robinson and P. Milanfar. Fundamental performance limits in image registration. *IEEE Transactions on Image Processing*, 13(9):1185–1199, 2004.
- [RM06] D. Robinson and P. Milanfar. Statistical performance analysis of superresolution. *IEEE Transactions on Image Processing*, 15(6):1413–1428, 2006.
- [RRY+13] R. Raghavendra, K. B. Raja, B. Yang, and C. Bush. Comparative evaluation of super-resolution techniques for multi-face recognition using light-field camera. In *Proceedings of 18th International Conference on Digital Signal Processing (DSP)*, pp. 1-6, 2013.
- [SCN11] T. Szydzik, G.M. Callicó, and A. Núñez. Efficient FPGA Implementation of a High-Quality Super-Resolution Algorithm with Real-Time Performance. *IEEE Transactions on Consumer Electronics*. 57(2):664-672, 2011.
- [SCO15] Scopus Digital Library, 2015.
- [SK94] W. Su and S. P. Kim. High-resolution restoration of dynamic image sequences. *International Journal of Imaging Systems and Technology*, 5(4):330–339, 1994.
- [SK05] Y.Y. Schechner and N. Karpel. Recovery of underwater visibility and structure by polarization analysis. *IEEE Journal of Oceanic Engineering*, 30(3):570-587, 2005.
- [SKC+13] S. Sriwilai, T. Kasetkasem, T. Chanwimaluang, T. Srinark, and T. Isshiki. A super-resolution mapping algorithm based on the level set method. In *Proceedings of International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, pp 1-6, 2013.
- [SO89] H. Stark and P. Oskoui. High-resolution image recovery from image plane arrays, using convex projections. *Journal of Optical Society of America A*, 6(11):1715–1726, 1989.
- [SS96] R. R. Schultz and R. L. Stevenson. Extraction of high-resolution frames from video sequences. *IEEE Transactions on Image Processing*, 5(6):996–1011, 1996.
- [SS13] Y. Swirski and Y. Y. Schechner. 3DFlicker from motion. In *Proceedings of IEEE International Conference on Computational Photography*, 2013.

- [SSB06] H.R. Sheikh, M.F. Sabir, and A.C. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 15(11): 3440-3451, 2006.
- [SSH+10] Y. Swirski, Y. Y. Schechner, B. Herzberg, and S. Negahdaripou. Underwater stereo using natural flickering illumination. In *Proceedings of IEEE OCEANS International Conference*, pp. 1-7, 2010.
- [SSN11] Y. Swirski, Y. Y. Schechner, and T. Nir. Variational stereo in dynamic illumination. In *Proceedings of IEEE International Conference on Computer Vision*, pp. 1124-1131, 2011.
- [STW+08] H.Su, L. Tang, Y. Wu, D. Tretter and J. Zhou. A practical and adaptive framework for Super-Resolution. In *Proceedings of 15th IEEE International Conference on Image Processing*, pp 1236-1239, 2008.
- [STW+12] H. Su, L. Tang, Y. Wu, D. Tretter, and J. Zhou. Spatially Adaptive Block-Based Super-Resolution. *IEEE Transactions on Image Processing*, 21(3): 1031-1045, 2012.
- [SZH+07] H. Shen, L. Zhang, B. Huang, and P. Li. A MAP approach for joint motion estimation, segmentation & superresolution. *IEEE Transactions on Image Processing*, 16(2):479-490, 2007.
- [SZJ+09] B.M. Smith, Li Zhang, H. Jin, and A. Agarwala. Light Field Video Stabilization. In *Proceedings of the IEEE International Conference on Computer Vision*, pp.341-348, 2009.
- [SZT+03] J. Sun, N. N. Zheng, H. Tao, and H. Shum. Image hallucination with primal sketch priors. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 729-736, 2003.
- [TA97] A. N. Tikhonov and V. A. Arsenin. *Solution of ill-posed problems*. Winston & Sons, Washington, 1997.
- [TH84] R. Y. Tsai and T. S. Huang. Multiple frame image restoration and registration. *Advances in Computer Vision and Image Processing*, pp 317-339. Greenwich, CT: JAI Press Inc., 1984.
- [Tib96] R. Tibshirani. Regression shrinkage and selection via the Lasso. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, 59(1):267-288, 1996.
- [TKG94] B. C. Tom, A. K. Katsaggelos, and N. P. Galatsanos. Reconstruction of a high resolution image from registration and restoration of low resolution images. In *Proceedings of IEEE International Conference on Image Processing*, pages 553-557, 1994.
- [TKM07] H. Takeda, S. Farsiu, and P. Milanfar. Kernel regression for image processing and reconstruction. *IEEE Transactions on Image Processing*, 16(2):349-366, 2007.
- [TMP+09] H. Takeda, P. Milanfar, M. Protter, and M. Elad. Super-resolution without explicit subpixel motion estimation. *IEEE Transaction on Image Processing*, 18(9):1958-1975, 2009.
- [TO06] E. Trucco and A. T. Olmos-Antillon. Self-Tuning Underwater Image Restoration. *IEEE Journal of Oceanic Engineering*, 31:511-519, 2006.
- [TP91] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586-591, 1991.

- [TSL00] J. B. Tenenbaum, V. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000.
- [UG92] H. Ur and D. Gross. Improved resolution from subpixel shifted pictures. *CVGIP: Graphical Models and Image Processing*, 54(2):181–186, 1992.
- [VCB+10] A. Villa, J. Chanussot, J.A. Benediktsson, M. Ulfarsson, and C. Jutten. Super-resolution: an efficient method to improve spatial resolution of hyperspectral images. In *Proceedings of IEEE International Symposium on Geoscience and Remote Sensing*, pp. 2003–2006, 2010.
- [VCT+10] L. G. Villanueva, G. M. Callicó, F. Tobajas, S. López, and V. de Armas, J. F. López and R. Sarmiento. Medical Diagnosis Improvement through Image Quality Enhancement Based on Super-Resolution. In *Proceedings of Euromicro Conference on Digital System Design: Architectures, Methods and Tools*, pp. 259–262, 2010.
- [WAH+92] Juyang Weng, N. Ahuja, Huang, and S. Thomas. Matching two perspective views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8): 806–825, 1992.
- [Wan06] Y. Wang. *Survey of objective video quality measurements*. Worcester Polytechnic University, Technical Report, 2006.
- [WBH91] W. J. Williams, M. L. Brown, and A. O. Hero. Uncertainty, information and time-frequency distributions, . In *Proceedings of SPIE (International Society for Optics and Photonics)*, vol. 1566, pp. 144–156, 1991.
- [WBS+04] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Imaging Processing*, 2004.
- [Win09] S. Winkler. Video quality measurement standards - current status and trends. In *International Conference on Information, Communications and Signal Processing, ICICS 2009*, Macau, China, 2009.
- [WJV05] B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Transactions on Graphics*, 24(3): 765–776, 2005.
- [WSR+11] M. Weinmann, C. Schwartz, R. Ruiters, and R. Klein. A Multi-camera, Multi-projector Super-Resolution Framework for Structured Light. In *Proceedings of the Int. Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission*, pp.397–404, 2011.
- [WTS05] Q. Wang, X. Tang, and H. Shum. Patch based blind image superresolution. In *Proceedings of IEEE International Conference on Computer Vision*, vol. 1, pp. 709–716, 2005.
- [XZZ14] X. Xu, Y. Zong, and L. Zang. Adaptive Subpixel Mapping Based on a Multiagent System for Remote-Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 52(2):787–804, 2014.
- [YCH+11] Hung-Yu Yang, Pei-Yin Chen, Chien-Chuan Huang, Ya-Zhu Zhuang, and Yeu-Horng Shiau. Low Complexity Underwater Image Enhancement Based on Dark Channel Prior. In *Proceedings of Second International Conference on Innovations in Bio-inspired Computing and Applications*, pp. 17–20, 2011.

- [YEB+02] J. C. Yang, M. Everett, C. Buehler, and L. McMillan. A Real-Time Distributed Light Field Camera. In *Proceedings of the 13th Eurographics Workshop on Rendering*, pp.1-10, 2002.
- [Yen56] L. J. Yen. On non-uniform sampling of bandwidth limited signals. *IRE Transactions on Circuits Theory*, 3(4):251–257, 1956.
- [YFW01] J. S. Yedidia, W. T. Freeman, and Y. Weiss. Generalized belief propagation. In *Proceedings of Advances in Neural Information Processing Systems*, pages 689–695, 2001.
- [YTM+08] Jianchao Yang, Hao Tang, Yi Ma, and Thomas Huang. Face hallucination via sparse coding. In *Proceedings of IEEE International Conference on Image Processing*, pp. 1264–1267, 2008.
- [YU14] Zihan Yu and Kiichi Urahama. Iterative Method for Inverse Nonlinear Image Processing. *Transactions on Fundamentals of Electronics, Communications and Computer Sciences*. E97-A(2):719-721, 2014.
- [YW82] D. C. Youla and H. Webb. Image registration by the method of convex projections: Part 1-theory. *IEEE Transactions on Medical Imaging*, 1(2):81–94, 1982.
- [YWH+08] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma. Image superresolution as sparse representation of raw image patches. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2008.
- [ZCI+08] A. Zaharescu, C. Cagniard, S. Ilic, E. Boyer, and R. Horaud. Camera-Clustering for Multi-Resolution 3-D Surface Reconstruction. In *Proceedings of Workshop on Multi-Camera and Multi-Modal Sensor Fusion*, 2008.
- [ZF03] B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11):977–1000, 2003.
- [ZH12] X. Zeng and H. Huang. Super-Resolution Method for Multiview Face Recognition from a Single Image per Person using Nonlinear Mappings on Coherent Features. *IEEE Signal Processing Letters*, 19(4):195-198, 2012.
- [ZMZ+14] M. Zanforlin, D. Munaretto, A. Zanella, and M. Zorzi. SSIM-based video admission control and resource allocation algorithms. *Proceedings of the WiOpt workshop “Wireless Video Performance”*, pp. 1-6, 2014.
- [ZP00] A. Zomet and S. Peleg. Efficient super-resolution and applications to mosaics. In *Proceedings of International Conference on Pattern Recognition*, pp. 579–583, 2000.
- [ZRP01] A. Zomet, A. Rav-Acha, and S. Peleg. Robust super-resolution. In *Proceedings of the IEEE Workshop on Applications of Computer Vision*, pp. 645–650, 2001.
- [ZWM+11] Z. Zang, X. Wang, J. Ma, and G. Jia. Super resolution reconstruction of three view remote sensing images based on global weighted POCS algorithm. In *Proceedings of International Conference on Remote Sensing, Environment and Transportation Engineering*, 2011.

