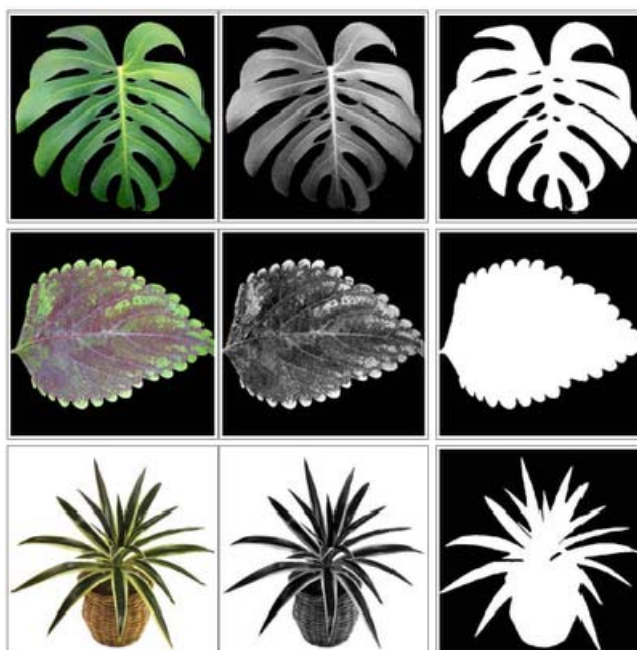




UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA  
DEPARTAMENTO DE SEÑALES Y COMUNICACIONES

## METODOLOGÍA PARA LA IDENTIFICACIÓN DE FORMAS MEDIANTE LA TRANSFORMACIÓN MARKOVIANA DE SU CONTORNO

*Juan Carlos Briceño Lobo*



*Directores: Dr. Carlos Manuel Travieso Gonzalez  
Dr. Jesús Bernardino Alonso Hernández*









**DON PEDRO JOSE QUINTANA MORALES SECRETARIO DEL  
DEPARTAMENTO DE SEÑALES Y COMUNICACIONES DE LA  
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA,**

**CERTIFICA,**

Que la Comisión de Investigación del Departamento (que tiene delegadas las competencias en materia de doctorado), en su sesión de fecha diez de enero de dos mil trece, tomó el acuerdo de dar el consentimiento para su tramitación, a la tesis doctoral titulada "METODOLOGÍA PARA LA IDENTIFICACIÓN DE FORMAS MEDIANTE LA TRANSFORMACIÓN MARKOVIANA DE SU CONTORNO" presentada por el doctorando Don Juan Carlos Briceño Lobo y dirigida por los Doctores Don Carlos Manuel Travieso González y Don Jesús Bernardino Alonso Hernández.

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, diez de enero de dos mil trece.





**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA**  
**DEPARTAMENTO DE SEÑALES Y COMUNICACIONES**

**PROGRAMA DE DOCTORADO**  
**CIBERNÉTICA Y TELECOMUNICACIÓN**



**TESIS DOCTORAL**

**Metodología para la identificación de formas mediante  
la transformación markoviana de su contorno**

**AUTOR:** D. Juan Carlos Briceño Lobo

**DIRECTOR:** Dr. D. Carlos Manuel Travieso González  
Dr. D. Jesús Bernardino Alonso Hernández

**El Director**

**El Codirector**

**El Doctorando,**

Las Palmas de Gran Canaria a 9 de enero de 2013





*A mis Directores de Tesis Doctoral.*



## **AGRADECIMIENTOS**

Quisiera agradecer a los directores de esta Tesis Doctoral, al **Dr. Carlos Manuel Travieso González** y al **Dr. Jesús Bernardino Alonso Hernández** su confianza, su ayuda y su paciencia, gracias a lo cual, se ha posibilitado la realización de este Trabajo.

Me gustaría destacar la ayuda, colaboración y apoyo recibido durante este tiempo, al resto de los miembros de la División de Procesado Digital de Señales (DPDS) del Intituto para el Desarrollo Tecnológico y la Innovación en las Comunicaciones (IDeTIC) de la Universidad de Las Palmas de Gran Canaria: a **Ing. Jaime Roberto Ticay Ricas**, el **Dr. Miguel Ángel Ferrer Ballester**, **Dr. Rafael Pérez Jiménez** y **D. José Cruz Gil**.

Quisiera además, tener una nota de agradecimiento a mis compañeros de la Universidad de Costa Rica, por su incondicional ayuda en la elaboración de este Trabajo, en particular **José Luis Vásquez Núñez**, así como a todos los miembros de la Facultad de Biología, como al **Prof. Federico Bolaños**, **Prof. Daniel R. Briceño Lobo**, **Adrián García (Pichi)** y a todos los demás integrantes del PILASIB.

Agradezco a **mi familia** todo su apoyo y sus ánimos, que son sin duda, los verdaderos elementos motrices de este Trabajo. Debido a su implicación, quiero agradecer de manera especial su ayuda, a mi madre **Noni** y mi mujer **Gemma**.

Y por último, quiero agradecer a todos **mis Amigos** su apoyo y aliento en el transcurso de la elaboración de esta Tesis Doctoral.



## **PRÓLOGO**

*“Las ciencias tienen las raíces amargas, pero muy dulces los frutos.”*

**Aristóteles** (384 AC-322 AC) *Filósofo griego.*



## TABLA DE CONTENIDOS

<b>1.</b>	<b>INTRODUCCIÓN.....</b>	<b>1</b>
1.1	MOTIVACIÓN.....	2
1.2	ESTADO DEL ARTE: ANTECEDENTES.....	6
1.2.1	<i>Descriptores del contorno.....</i>	7
1.2.2	<i>Uso de los Modelos Ocultos de Markov.....</i>	10
1.3	LA TESIS: SU OBJETIVO.....	13
1.4	METODOLOGÍA.....	13
1.5	CONTRIBUCIONES Y RESULTADOS.....	14
1.6	ESTRUCTURA DE LA MEMORIA.....	16
<b>2.</b>	<b>DEMOSTRACIÓN DE LA HIPÓTESIS.....</b>	<b>17</b>
2.1	PLANTEAMIENTO DE LA HIPÓTESIS.....	18
2.2	PARAMETRIZACIÓN: CODIFICACIÓN DE LA FORMA.....	18
2.3	DESARROLLO DEL NÚCLEO: ESPACIOS HIPERDIMENSIONALES.....	22
2.3.1	<i>Modelos de Markov.....</i>	22
2.3.2	<i>Extensión a los modelos ocultos de Markov.....</i>	23
2.3.3	<i>Elementos de un HMM.....</i>	23
2.3.4	<i>Los tres problemas básicos del HMM.....</i>	25
2.3.6	<i>Adaptación del evaluador de Fisher.....</i>	29
2.3.7	<i>Sistemas de clasificación por hiperdimensionalidad: uso del SVM.....</i>	36
<b>3.</b>	<b>METODOLOGÍA EXPERIMENTAL.....</b>	<b>39</b>
3.1	BASES DE DATOS.....	42
3.1.1	<i>Base de datos sintética.....</i>	43
3.1.2	<i>Granos de Polen: Variabilidad simple.....</i>	43
3.1.3	<i>Hoja de Árboles: Variabilidad simple.....</i>	45
3.1.4	<i>Labios: Variabilidad moderada.....</i>	47
3.1.5	<i>Manos: Variabilidad moderada.....</i>	47
3.1.6	<i>Palabras: Variabilidad compleja.....</i>	49

3.1.7	<i>Firmas: Variabilidad compleja</i> .....	51
3.1.8	<i>Dactilológico: Variabilidad compleja</i> .....	52
3.2	PRE-PROCESADO .....	53
3.3	PROCESADO Y PARAMETRIZACIÓN, TRANSFORMACIÓN A ESPACIOS HIPERDIMENSIONALES Y CLASIFICACIÓN.....	53
3.4	METODOLOGÍA DE EXPERIMENTACIÓN.....	54
3.4.1	<i>Fase de entrenamiento</i> .....	56
3.4.2	<i>Fase de test</i> .....	57
3.4.3	<i>Validación cruzada:</i> .....	57
<b>4.</b>	<b>EXPERIMENTOS Y RESULTADOS.....</b>	<b>59</b>
4.1	UTILIZACIÓN BASE DE DATOS SINTÉTICA.....	60
4.2	VARIACIÓN SIMPLE .....	61
4.3	VARIACIONES MODERADAS .....	64
4.3.1	<i>Base de datos de labios</i> .....	64
4.3.2	<i>Base de datos manos</i> .....	66
4.4	VARIACIONES COMPLEJAS .....	68
4.4.1	<i>Base de datos de palabras</i> .....	68
4.4.2	<i>Base de datos de firmas</i> .....	69
4.4.3	<i>Base de datos de signos dactilológicos</i> .....	71
<b>5.</b>	<b>ANÁLISIS DE RESULTADOS.....</b>	<b>73</b>
5.1	ANÁLISIS SEGÚN BBDD.....	73
5.1.1	<i>Análisis de la base de datos Sintética</i> .....	74
5.1.2	<i>Análisis de la base de datos Polen</i> .....	74
5.1.3	<i>Análisis de la base de datos Laurisilva de Costa Rica</i> .....	75
5.1.4	<i>Análisis de la base de datos Laurisilva de Canarias</i> .....	75
5.1.5	<i>Análisis de la base de datos Labios</i> .....	76
5.1.6	<i>Análisis de la base de datos Manos</i> .....	76
5.1.7	<i>Análisis de la base de datos Palabras</i> .....	77
5.1.8	<i>Análisis de la base de datos Firmas</i> .....	78
5.1.9	<i>Análisis de la base de datos Dactilológico</i> .....	78



5.2	ANÁLISIS SEGÚN LA VARIABILIDAD DEL CONTORNO DE LAS BASE DE DATOS. ....	79
5.2.1	<i>Análisis de la bases de datos de tipo simple. ....</i>	79
5.2.2	<i>Análisis de la bases de datos de tipo moderada. ....</i>	80
5.2.3	<i>Análisis de la bases de datos de tipo compleja. ....</i>	80
5.2.4	<i>Análisis de todas las bases de datos. ....</i>	81
5.3	EXTENSIÓN DE LA USABILIDAD DEL MODELO Y OTRAS CONSIDERACIONES. ....	81
5.3.1	<i>Coste computacional del sistema. ....</i>	81
5.3.2	<i>Robustez frente a ruido. ....</i>	82
5.3.3	<i>Generalización del sistema para verificación en manos. ....</i>	83
5.3.4	<i>Reducción de la resolución en la base de datos de labios. ....</i>	85
5.3.5	<i>Evaluación del sistema de parametrización de contornos con otros clasificadores aplicado a firmas manuscritas. ....</i>	88
5.3.6	<i>Uso dos de bases de datos en la identificación de usuarios a partir de su firma manuscrita. ....</i>	89
5.4	COMPARATIVA CON EL ESTADO DEL ARTE. ....	90
<b>6.</b>	<b>CONCLUSIONES Y LÍNEAS FUTURAS. ....</b>	<b>93</b>
6.1	CONCLUSIONES. ....	93
6.2	LÍNEAS FUTURAS. ....	95
<b>7.</b>	<b>BIBLIOGRAFÍA. ....</b>	<b>97</b>



## ÍNDICE DE FIGURAS

Figura 1.1: a) Imagen ecográfica del corazón. b) Diferentes imágenes de melanomas. ...	4
Figura 1.2: Detección del contorno de manzanas en una cadena de producción. ....	4
Figura 1.3: Sistema automático de detección de nudos en tablones de madera. ....	4
Figura 1.4: a) Imágenes de granos de polen. b) Imágenes de hojas de plantas. ....	5
Figura 2.1: Ejemplo de un fragmento de un contorno de imagen descompuesto en $G_j$ grafos. ....	20
Figura 2.2: Ejemplo de estructuración por grafos $G_i$ en una hoja fuertemente depredada. La presentación de los grafos se hace siguiendo la dirección contraria de las manecillas del reloj, y en orden de secuencia de colores: rojo, verde, azul. ....	20
Figura 2.3: Ejemplo general de una codificación angular. ....	21
Figura 2.4: Modelo de estados de Markov. ....	22
Figura 2.5: Estados del modelo <i>left-right</i> . ....	28
Figura 2.6: Hiperplano de separación lineal para el caso de separación lineal. ....	37
Figura 2.7: Diagrama de bloques del clasificador SVM multiclase en modo entrenamiento para un sistema de N clases. ....	38
Figura 3.1: Esquema general de los bloques utilizados en este trabajo. ....	41
Figura 3.2: Algunas muestras de la base de datos sintética. ....	43
Figura 3.3: Ejemplos de la base de datos de polen. ....	45
Figura 3.4: Imágenes de las 16 variedades de Laurisilva canariensis. ....	46
Figura 3.5: Géneros de Laurisilva de Costa Rica. ....	47
Figura 3.6 Imágenes de la base de datos de palabras. ....	50
Figura 3.7: Signos que conforman la base de datos dactilológica. ....	52
Figura 3.8: Pre-procesado aplicado a la base de datos de granos de polen. ....	54
Figura 4.1: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento. ....	65

Figura 5.1: (a, b) Ejemplos de muestras de hojas con ruido: muestras de comidas por insectos. (c, d) Mutación o deformaciones relacionadas con enfermedades. (e) Una muestra joven. (f) deformación circunstancial. .... 83

Figura 5.2: Curva ROC utilizando las bases de datos GPDS144 y UST usando dos muestras de entrenamiento. .... 84

Figura 5.3: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 3:1 del tamaño original..... 85

Figura 5.4: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 4:1 del tamaño original..... 86

Figura 5.5: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 5:1 del tamaño original..... 87

## ÍNDICE DE TABLAS

Tabla 3.1: Familia, género y especie de la base de datos de granos de polen .....	44
Tabla 3.2: Variedades de la base de datos de hojas de árboles de Laurisilva Canaria ...	45
Tabla 3.3: Características más importantes de la base de datos caras para el reconocimiento de labios .....	47
Tabla 3.4: Características más importantes de la base de datos de manos GPDS144....	48
Tabla 3.5: Características más importantes de la base de datos de manos GPDS100hand3Band.....	49
Tabla 3.6: Características más importantes de la base de datos de manos UST .....	49
Tabla 3.7: Descripción del contenido de la base de datos de palabras.....	50
Tabla 3.8: Distribución del corpus de GPDS-800 .....	51
Tabla 4.1: Resultados con HMM usando 40, 50 y 60 estados .....	60
Tabla 4.2: Resultados del clasificador SVM con los datos transformados con el kernel HMM con 50 estados.....	61
Tabla 4.3: Tasa de éxito usando el clasificador HMM.....	61
Tabla 4.4: Tasa de acierto de los datos transformados utilizando SVM como clasificador .....	62
Tabla 4.5: Tasa de acierto utilizando el kernel HMM y SVM variando el número de estados .....	62
Tabla 4.6: Resultados usando 40 estados definidos en HMM.....	63
Tabla 4.7: Resultados usando 50 estados definidos en HMM.....	63
Tabla 4.8: Resultados usando 60 estados en el HMM.....	63
Tabla 4.9: Resultados con SVM usando un HMM con 40 estados .....	64
Tabla 4.10: Resultados con SVM usando un HMM con 50 estados .....	64
Tabla 4.11: Resultados con SVM usando un HMM con 60 estados .....	64

Tabla 4.12: Tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con el tamaño original .....	65
Tabla 4.13: Tasa de acierto usando SVM con un kernel RBF para diferentes números de muestras de entrenamiento con imágenes de tamaño original .....	66
Tabla 4.14: Tasa de acierto para un clasificador HMM, usando información del contorno.....	67
Tabla 4.15: Tasa de acierto usando el clasificador SVM con los datos transformados por el kernel HMM .....	67
Tabla 4.16: Tasa de acierto usando clasificadores HMM y SVM para datos sin transformar y transformados, usando un 20% de las muestras para el entrenamiento .....	68
Tabla 4.17: Tasa de acierto usando clasificadores HMM y SVM para datos sin transformar y transformados, usando 200 puntos de control.....	69
Tabla 4.18: Tasa de acierto usando clasificadores HMM para firmas manuscritas .....	70
Tabla 4.19: Tasa de acierto usando clasificadores SVM para firmas manuscritas de la base de datos MYCT-75 .....	70
Tabla 4.20: Tasa de acierto usando clasificadores SVM para firmas manuscritas de la base de datos GPDS100.....	71
Tabla 4.21: Tasa de acierto usando clasificadores HMM para el lenguaje dactilológico. ....	71
Tabla 4.22: Tasa de acierto usando SVM con kernel lineal y RBF.....	72
Tabla 5.1: Descripción de los valores más significativos de cada base de datos .....	79
Tabla 5.2: Tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con las imágenes a una escala 3:1.....	86
Tabla 5.3: Tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con las imágenes a una escala 4:1.....	86
Tabla 5.4: Tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con las imágenes a una escala 5:1.....	87

Tabla 5.5: Tasa de acierto usando SVM con un kernel RBF para diferentes números de muestras de entrenamiento con una escala 4:1 del tamaño original.....	88
Tabla 5.6: Tasas de éxito en la base de datos de firmas usando diversos clasificadores	89
Tabla 5.7: Tasa de éxito para el mismo clasificador variando la parametrización de las firmas.....	89
Tabla 5.8: Comparativa del trabajo propuesto para la biometría de la mano frente al estado del arte .....	91
Tabla 5.9: Comparativa del trabajo propuesto para la biometría de la firma frente al estado del arte .....	91





## ACRÓNIMOS

<b>ATVS</b>	<i>Biometric Recognition Group</i> de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid
<b>CINAT</b>	Centro de Investigación Nacional de Apicultura Tropical
<b>DHMM</b>	del inglés <i>Discrete Hidden Markov Model</i> , Modelo oculto de Markov discreto
<b>EER</b>	del inglés <i>Equal Error Rate</i> , Ratio de igual error
<b>FAR</b>	del inglés <i>False Acceptance Rate</i> , Ratio de falsa aceptación
<b>FRR</b>	del inglés <i>False Rejection Rate</i> , Ratio de falso rechazo
<b>g</b>	gamma
<b>GLDS</b>	del inglés, <i>Generalized Linear Discriminant Sequence</i> , Secuencia discriminante lineal generalizada
<b>GMM</b>	del inglés, <i>Gaussian Mixture Model</i> , Model de Mezcla de Gausianas
<b>HMM</b>	del inglés <i>Hidden Markov Model</i> , Modelo oculto de Markov
<b>HSV</b>	del inglés <i>Hue, Saturation and Value</i> , Matiz, saturación y valor
<b>LDA</b>	del inglés <i>Linear Discriminant Analysis</i> , Análisis discriminante lineal
<b>ppp</b>	píxeles por pulgada
<b>RBF</b>	del inglés <i>Radial Basis Function</i> , Función de base radial
<b>ROC</b>	del inglés <i>Receiver operating characteristic</i> , Característica operativa del receptor
<b>SRM</b>	del inglés <i>Structural Risk Minimization</i> , Minimización del riesgo estructural
<b>SVM</b>	del inglés <i>Support Vector Machine</i> , Máquina de soporte vectorial



## **CAPÍTULO 1**

### **Introducción**

---

En este primer capítulo se van a determinar aquellos aspectos que justifican y motivan la realización de esta tesis. Por otra parte, también se explicarán aquellos conocimientos previos, de los cuales se ha partido para el desarrollo del proyecto.

## **1.1 Motivación**

La mayor parte de la información que nos rodea esta formada en su esencia por patrones complejos. En psicología, uno de los principales temas de estudio, son los mecanismos que convierten los estímulos externos en experiencias perceptuales y significativas. A pesar de haber realizado un gran avance en la comprensión de estos procesos, aún quedan muchas incógnitas por resolver, para que se pueda construir un modelo definitivo de cómo nuestro cerebro lleva a cabo este reconocimiento de patrones. No obstante, es aceptado por la comunidad científica, que este proceso se debe llevar a cabo siguiendo un esquema general. Antes de la fase de reconocimiento, una característica debe ser percibida por el sistema sensorial. Previamente a esta fase, una característica debió ser percibida y recordada, para que pueda realizarse una correspondencia entre la característica percibida en el instante actual y la característica memorizada.

Hoy en día, el desarrollo tecnológico de equipos de procesamiento computacional y de los algoritmos de aprendizaje en los campos de visión computarizada e inteligencia artificial, están permitiendo resolver los problemas de una forma real y eficiente. Además de permitir enfrentarnos a los problemas actuales desde una perspectiva innovadora, nos permite enfrentarnos a problemas antes intratables. Esta tesis se centra en el reconocimiento de formas, siendo el reconocimiento de patrones en dos dimensiones, un campo cada vez más extendido debido a su gran importancia y aplicabilidad.

La forma es una importante característica visual y es una de las características básicas que se usan para describir el contenido de una imagen. En la actualidad, existe una demanda de herramientas efectivas que facilitan la búsqueda de imágenes. Sin embargo, la descripción correcta y su reconocimiento automático es una tarea complicada. Esto es debido a que, cuando un objeto del mundo real en 3D es proyectado a un plano de imagen en 2D, se pierde información de una dimensión del objeto. Como resultado, la forma extraída de una imagen es una representación parcial del objeto proyectado. Para hacer el problema aún más complejo, la forma del objeto es normalmente degenerada con ruido, distorsiones, defectos y/u oclusiones.

La importancia del reconocimiento de formas se ve reflejada en diversas aplicaciones reales en la Biología, la Medicina, la Industria e incluso en el estándar MPEG-7 [MPEG, 2001]. En el estándar MPEG-4 se consigue más libertad de

interacción con el contenido audiovisual que en MPEG-1 y MPEG-2, donde quedaban cubiertas las necesidades de obtener información audiovisual en cualquier sitio. Con MPEG-7, se busca la forma de enlazar los elementos del contenido audiovisual, encontrar y seleccionar la información que el usuario necesita e identificar y proteger los derechos del contenido. MPEG-7 surge a partir del momento en que aparece la necesidad de describir los contenidos audiovisuales debido a la creciente cantidad de información. El hecho de gestionar los contenidos es una tarea compleja (encontrar, seleccionar, filtrar, organizar... el material audiovisual) [MPEG, 2001].

MPEG-7 ofrece un mecanismo para describir información audiovisual, de manera que sea posible desarrollar sistemas capaces de indexar grandes bases de material multimedia (este puede incluir: gráficos, imágenes estáticas, audio, modelos 3D, vídeo y escenarios de cómo estos elementos se combinan) y buscar en estas bases de materiales manual o automáticamente. Dentro de los objetivos del MPEG-7, está el informar de cómo los objetos están combinados dentro de una escena. Queda de manifiesto la importancia de los objetos, y en particular, se pueden mostrar como formas bidimensionales, como la base de datos MPEG7 CE-Shape-1, científicamente muy usada para la identificación de formas [Bai et al., 2010].

En aplicaciones reales, se puede ver la importancia del contorno en diversas disciplinas como la Medicina, la Industria y la Biología, entre muchos ámbitos. En particular, en Medicina se pueden encontrar contornos para definir el tamaño de las Aurículas y los Ventrículos del corazón, en imágenes ecográficas, comprobando así su funcionamiento (véase la Figura 1.1a); o bien en la definición de un lunar o un melanoma (véase la Figura 1.1b).

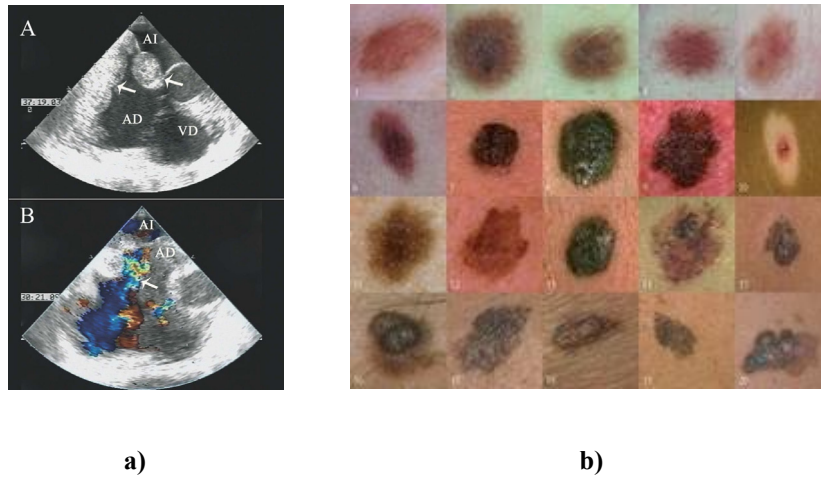


Figura 1.1: a) Imagen ecográfica del corazón. b) Diferentes imágenes de melanomas.

En el ámbito industrial, también se pueden encontrar diversos ejemplos, como la identificación de frutas, para su conteo o identificación (véase la Figura 1.2), o bien en la calidad de materiales, para evitar grietas (véase la Figura 1.3).

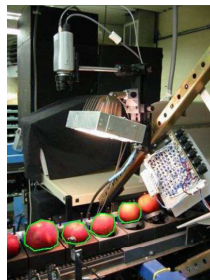
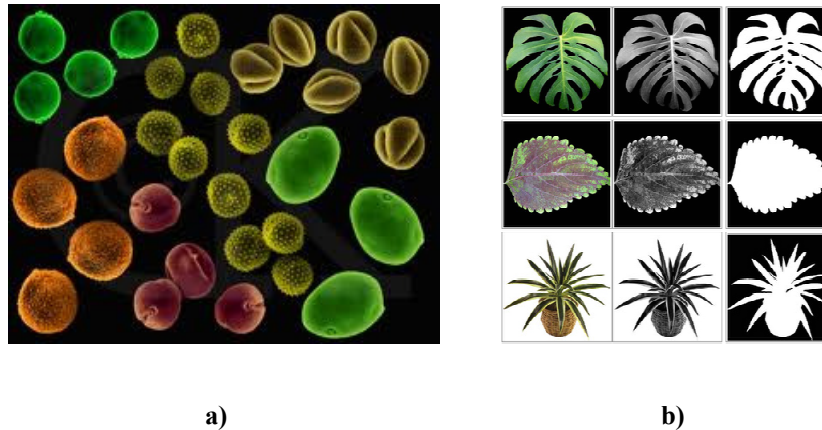


Figura 1.2: Detección del contorno de manzanas en una cadena de producción.



Figura 1.3: Sistema automático de detección de nudos en tablones de madera.

También se pueden encontrar ejemplos de contornos en el ámbito de la Biología, como se puede observar en la figura 1.4, para la identificación mediante imágenes, de granos polen y de hojas de plantas.



**Figura 1.4: a) Imágenes de granos de polen. b) Imágenes de hojas de plantas.**

Debido a la complejidad del problema de reconocimiento de formas, las aplicaciones que se encuentran en el mercado son específicas respecto al problema que se desea resolver y por lo tanto, existen restricciones derivadas de la naturaleza de las características, por ejemplo, escritura manuscrita, imágenes médicas, imágenes de naturaleza biológica, espectogramas, rostros, piezas industriales, reconocimiento de fallos en materiales, etc. También, las especificaciones del sistema suponen otra limitación ya que se pueden encontrar en el estado del arte sistemas que ofrecen respuestas con un alto grado de fiabilidad, sin embargo los tiempos de cómputo pueden hacer inviables, para un sistema en tiempo real o comercial. Finalmente, los aspectos económicos juegan un papel muy importante, cuando se desea desarrollar un sistema que pueda ser accesible a los usuarios en general. Sensores de alta fiabilidad y dispositivos de computos más potente, implica un coste mayor del sistema final.

De esta forma se deduce que, un sistema diseñado para resolver un problema concreto, no será el más eficaz para resolver otro, con distintas restricciones. Esta visión del uso de formas o contornos en la vida diaria, así como la necesidad de investigar y desarrollar nuevas técnicas que permitan diseñar sistemas que resuelvan el mayor número de problemas, independientemente de los factores antes mencionados, es la que da motivos para desarrollar la actual tesis doctoral y la propuesta de su hipótesis.

## **1.2 Estado del arte: antecedentes**

La representación de formas busca características efectivas y perceptualmente importantes basadas en la información de los bordes de la forma del objeto, como el del borde más la información del interior. El descriptor de la forma debe también ser capaz de describir formas afectadas por el ruido, formas distorsionadas de diversas formas y defectuosas, efectos que son tolerados por los seres humanos cuando se comparan formas. Esto se conoce como el requisito de robustez. Se desean características compactas ya que son ideales para la indexación y recuperación en línea. Si un descriptor de forma tiene una característica jerárquica de general a fina, se puede lograr un alto nivel de eficiencia de emparejamiento de formas. Esto se debe, a que las formas pueden ser comparadas a nivel general, para eliminar primero grandes cantidades de formas disímiles, y a nivel fino, las formas pueden ser comparadas en detalles. Un descriptor de forma deseable debería ser independiente de la aplicación y no sólo funcionar bien para cierto tipo de formas. Una característica importante de un descriptor de formas es que tenga una baja complejidad de cómputo. Para un descriptor de forma, complejidad baja de computo significa minimizar cualquier incierto o factores especiales, que están involucrados en los procesos derivados. Cuantos menos factores de incertidumbre están involucrados en los procesos de cómputo, más robusto se vuelve el descriptor. En resumen, baja complejidad de cálculo significa claridad y estabilidad.

En el espacio de imágenes en dos dimensiones, la forma de un objeto se define como el área rodeada por una curva de contorno cerrada. Existen dos tipos de descriptores de la forma: descriptores basados en el contorno de la forma y descriptores basados en la región de la forma. Cuando el contorno de un objeto es descrito con precisión, el objeto puede ser identificado.

Las técnicas basadas en contorno explotan solamente la información contenida en los bordes del objeto y se usan mediante dos enfoques: enfoque continuo (información global) y enfoque discreto (información estructural). El enfoque continuo no divide al forma en sub-partes, normalmente se obtiene un vector de características del borde que describe la forma del objeto. La medida de similitud se lleva a cabo normalmente usando la distancia métrica. Por otro lado, el enfoque discreto divide el contorno en segmentos llamados primitivas. La representación final de estos segmentos son una cadena, un árbol o un gráfico.



El proceso de obtener el conjunto de características que puedan ser usadas en el proceso de clasificación no es una tarea fácil. Cuando las características elementales están explícitamente presentes en los datos adquiridos y son lo suficientemente discriminantes, éstas pueden ser pasadas directamente a la etapa de clasificación. Sin embargo, cuando la información discriminante no es tan evidente, es necesario llevarlas a un espacio superior donde sea posible separarlas. Esta operación se lleva a cabo normalmente a través de transformaciones de los datos, a espacios de orden superior mediante el uso de un núcleo o kernel.

De la misma forma, el sistema de clasificación es un elemento crítico en el binomio eficacia-eficiencia cuando se trata de grandes volúmenes de datos y cuando éstos, presentan un alto grado de complejidad en cuanto a las características que lo componen. A continuación, se describirán los trabajos las técnicas más utilizadas para parametrizar los contornos de las formas de los objetos y posteriormente, se describirán los trabajos sobre los HMMs, tanto como clasificador como herramienta de transformación.

### ***1.2.1 Descriptores del contorno***

Entre las técnicas de representación de la forma basadas en información global, se encuentran los descriptores de forma simple: el área, circularidad, excentricidad, orientación del eje mayor y la energía de curvatura, es decir, el número de transformaciones necesarias para alinear dos formas [Young et al., 1974]. Peura y Livarinen propusieron otros descriptores simples en [Peura and Livarinen, 1997] en donde incluyen la convexidad, el radio de los ejes principales, varianza circular y varianza elíptica.

Dentro de estas técnicas, se encuentran la denominada coincidencia de formas basadas en correspondencia, en donde se mide la similitud de formas midiendo la coincidencia punto a punto, en otras palabras, cada punto se trata como una característica. El método más utilizado en esta técnica es la distancia Hausdorff, el cual fue extendido por Rucklidge [Rucklidge, 1997] a la coincidencia afin invariante y cuyo alto coste computacional fue mejorado por Chetverikov and Khenokh [Chetverikov y Khenokh, 1999].

Otra técnica utilizada, es la denominada firma de la forma, la cual representa una forma mediante una función unidimensional derivada de los puntos del borde del objeto.

Esta firma comprende el perfil del centroide, coordenadas complejas, distancia al centroide, ángulo tangencial, curvatura, área y longitud de la cuerda [Davies, 1997], [Van Otterloo, 1991], [Zhang y Lu, 2002]. Estas características son normalizadas para ser invariantes a la traslación y al escalado, pero presentan el problema del alto coste computacional y además son sensibles al ruido, y ligeros cambios en los bordes, pueden causar errores significativos en el proceso de coincidencia de las formas. La técnica del cálculo de los momentos de los bordes, se puede utilizar para reducir la dimensionalidad de la representación del contorno [Sonka et al., 1993]. La ventaja que presentan es su fácil implementación, sin embargo es difícil asociar momentos de orden alto a interpretaciones físicas.

Los métodos estocásticos también han sido utilizados para calcular descriptores de la formas, basados en el contorno. Se utilizan modelos de serie temporales y especialmente modelos auto-regresivos [Chellappa y Bagdazian, 1984], [Das y Paulink, 1990], [Dubois y Glanz, 1986], [Eom y Park, 1990], [Kashyap y Chellappa, 1981], [He y Kundu, 1981] y [Sekita et al., 1992]. La desventaja de utilizar métodos auto-regresivos es en el caso de bordes complejos, un pequeño número de parámetros auto-regresivos no son suficientes para una descripción adecuada.

El problema de la sensibilidad al ruido y a las variaciones en la mayoría de los dominios espaciales, empuja al uso de un análisis de escalado espacial. La representación de la forma usando escalado espacial, se crea a través del seguimiento de los puntos de inflexión en los bordes de una forma, filtrados por filtros paso bajo Gaussianos con ancho de banda variable [Asada y Brady, 1984], [Asada y Brady, 1986]. Mokhtarian y Mackworth adoptaron el método de interpretación de Asada y Brady, para extenderlo como registro del contorno. Ellos definieron a la firma de escalado espacial, como una curvatura de escalado espacial (CSS Curvature scale space) [Mokhtarian y Mackworth, 1986].

Los descriptores espectrales resuelven el problema de la sensibilidad al ruido y a las variaciones del contorno, a través del análisis del contorno en el dominio espectral. Uno de los métodos más usados en el dominio espectral son los descriptores de Fourier [Brill, 1968], [Arbter, 1989], [Arbter et al., 1990], [Marine, 2000] y [Zhang y Lu, 2001].

Otros descriptores que analizan el contorno en el dominio espectral son los descriptores wavelet [Tieng y Boles, 1997], [Yang et al., 1998], [Ohm et al., 2000].

Éstos últimos tienen la ventaja sobre los descriptores de Fourier, que realizan un análisis de multi-resolución tanto en el dominio espacial como en el espectral, sin embargo, el incremento de la resolución espacial reduce la resolución en frecuencia.

En el enfoque de representación discreta o estructural de la forma, el contorno se divide en segmentos llamados primitivas. La diferencia entre los métodos estructurales es el criterio de selección y organización de las primitivas. Los métodos de descomposición de contornos están basados en aproximación poligonal, descomposición de curvatura y ajuste de curvas [Pavlidis, 1982]. El resultado se codifica en una cadena de la forma general  $S = s_1, s_1, \dots, s_n$  en donde  $s_1$  puede ser un elemento de una cadena de código, el lado de un polígono, un arco cuadrático, una *spline*, etc. Al mismo tiempo  $s_1$  puede contener un número de atributos tales como longitud, curvatura media, curvatura máxima, orientación, etc. Esta cadena se puede usar directamente como descripción o como datos de entrada para un análisis sintáctico en un nivel superior.

Uno de los métodos más utilizados es la representación *chain code*, en donde una curva arbitraria se representa por una secuencia de pequeños vectores de longitud unidad y con un número limitado de posible direcciones [Freeman, 1961]. La representación *code chain* tiene dimensiones altas y es sensible al ruido. Comúnmente, se utiliza como entrada a sistemas que realizan análisis en niveles más altos.

En el método de descomposición de polígonos, el contorno es dividido en líneas de segmentos por aproximación polígonos, fijando los vértices de los polígonos como primitivas [Groskey y Mehrotra, 1990], [Groskey, 1992]. Las características de cada primitiva se expresa como una cadena de cuatro elementos que consisten en el ángulo interno, distancia hacia el próximo vértice y las coordenadas  $x$  e  $y$ . Como es obvio, estas características no son invariantes a la traslación, escalado y rotación.

La descomposición suave de curvas fue propuesta en [Barretti et al., 2000] como una extensión de [Groskey y Mehrotra, 1990], en donde se utilizan, para obtener las primitivas (también llamados *tokens*), los puntos de cruce por cero de la curva producto de un contorno suavizado por un filtro Gaussiano. La característica de cada token es la máxima curvatura y su orientación. Dado que la característica incluye orientación de la curva, no es invariante a la rotación. El método de escalado espacial utiliza un esquema de coincidencia modelo a modelo. En este enfoque las primitivas de la forma, se

obtienen primero de una técnica de sintonización del suavizado de la curvatura [Dudek y Tsotsos, 1997]. Un descriptor está compuesto por la longitud del segmento, la posición y el valor de sintonización de la curvatura. A continuación, se crea una cadena de descriptores de segmentos que describe la forma. Debido a que se incluye la longitud de los segmentos estos descriptores no son invariantes al escalado.

El análisis sintáctico representa la forma por un conjunto definido de primitivas llamado codebook, en donde cada primitiva es llamada *codeword* [Fu, 1974]. El método general, se basa en formular la representación como una cadena gramática. Cada primitiva se interpreta como un alfabeto de alguna gramática, la cual es el conjunto de reglas de sintaxis que gobiernan la generación de oraciones formadas por símbolos del alfabeto. Sin embargo no es un método práctico para aplicaciones generales, debido al hecho que no es posible inferir un patrón gramatical que que pueda generar solo patrones válidos. Además, este método requiere el conocimiento a priori de las bases de datos, para definir los *codewords* o el alfabeto.

### **1.2.2 Uso de los Modelos Ocultos de Markov**

El primer trabajo de la última década, en donde se introduce el concepto de kernel de los Modelos Ocultos de Markov (HMM) es el presentado en [Wang et al., 2002]. Los autores proponen el mapeo entre dos secuencias dinámicas usando un kernel HMM (KHMM). En este caso el KHMM es un HMM para el cual las funciones basadas en el kernel son usadas para modelar la densidad de las observaciones de los estados de los conjuntos de entrada y salida. Específicamente, la densidad de las observaciones de los estados es estimada empleando un esquema de probabilidad ponderada de muestreo.

Por otra parte, Cambell desarrolla un entorno que permite combinar SVMs con HMM [Campbell, 2003], en donde utiliza el algoritmo de Viterbi para generar un conjunto de sub-secuencias del vector de características. Cada sub-secuencia es puntuada usando SVMs con un kernel *Generalized Linear Discriminant Sequence* (GLDS).

Basados en esta idea Cambell, un año más tarde, se presenta un sistema para el reconocimiento de escritura manuscrita Thai *online* [Sanguansat et al., 2004], donde se combinan HMMs y SVMs mediante kernels de espacio de puntuación. Los kernel de espacio de puntuación generalizan los kernel de Fisher y están basados en modelos

generativos subyacentes tales como *Gaussian mixture models* (GMMs) los cuáles son la distribución de salida de cada estado en los HMMs.

Un aspecto que debe tomarse en cuenta cuando se utilizan HMMs, es el coste computacional. Así en [Andelic, 2006], se propone un nuevo algoritmo de entrenamiento de orden recursivo para kernels basados en discriminantes para hacerlo computacionalmente más eficiente. En este caso, se integró el método en un sistema de reconocimiento de voz basado en HMM mediante la traducción de las salidas del clasificador basado en kernel, en probabilidades de clase condicional.

Más tarde, Jin y Mokhtarian proponen un sistema de reconocimiento de objetos en movimientos basado en la dinámica de sus formas [Jin y Mokhtarian, 2006]. En esta propuesta, la deformación espacio-temporal de las formas es capturada por HMMs. Dado que en el marco tradicional de los HMMs los estados ocultos están típicamente emparejados con los datos de entrenamiento, lo que acarrea muchos problemas en el proceso de aprendizaje. Se introduce un enfoque de HMM no paramétrico, el cual utiliza salidas continuas del HMM con estados arbitrarios (que no están emparejadas con los datos de entrenamiento) para aprender la dinámica de las formas directamente de una gran cantidad de datos de entrenamiento. Para alcanzar este objetivo, se aplicó un algoritmo de estimación de densidad con un kernel no paramétrico para el aprendizaje de la densidad de distribución de las observaciones con el fin de compensar la incertidumbre introducida por los estados ocultos arbitrarios.

En 2008, Aran y Akarun proponen una estrategia de clasificación multi-clase basada en el kernel de Fisher [Aran y Akarun, 2008]. El kernel de Fisher combina la potencia de los clasificadores discriminativos y generativos mediante la asignación de secuencias de longitud variable a un nuevo espacio fijo de características. La asignación se basa en un modelo generativo único y un clasificador que es intrínsecamente binario. Dado que el problema de análisis trataba más de dos clases, los autores implementaron una clasificación multi-clase en lugar de una clasificación binaria, en cada espacio de puntuación de Fisher y finalmente, combinaron las decisiones de los clasificadores multi-clase.

Recientemente Hamdi et al. Construyen el kernel de una SVM utilizando HMMs [Hamdi et al., 2010]. Este enfoque se basa en modelos HMM y tiene dos pasos principales, en primer lugar se ajusta un HMM para cada una de las secuencias

individuales  $N$ . Para cada modelo ajustado, se evalúa la verosimilitud logarítmica de cada secuencia. Esto dará una matriz  $N \times N$  de verosimilitud logarítmica, que se adapta para servir como kernel del clasificador SVM. En el segundo paso, se entrena un clasificador SVM para el aprendizaje de la frontera de decisión, entre las muestras positivas y las muestras negativas.

A modo de resumen de esta revisión del estado del arte, se observa que los HMMs son utilizados como sistemas de clasificación y cuando son utilizados como herramientas de transformación, se utilizan normalmente las funciones de verosimilitud del clasificador, o bien, kernels Gaussianos o kernels basados en discriminante.

Los trabajos más representativos, en donde se ha utilizado el kernel de Fisher como transformador de los datos de entrada para combinarlos con SVMs, son los presentados por [Haussler et al., 2000], [Jaakkola et al., 1999], [Jaakkola et al., 2000]. En estos sistemas, se parte de un HMM entrenado con las muestras positivas para modelar una clase dada. Se utiliza este HMM para mapear cada nueva secuencia que se desea clasificar a un vector de longitud fija, definido por los autores como *las puntuaciones de Fisher*. El cálculo del kernel se realizó en base a la distancia Euclidiana entre el vector de puntuaciones obtenido y los vectores de puntuaciones de las muestras positivas y negativas. Dicho kernel se modeló con una respuesta Gaussiana, siendo la función discriminantes la misma usada para el SVM. Los autores definen a este resultado como el método SVM-Fisher.

Esta descripción del estado de arte deja muchas puertas abiertas a la investigación, y en ese sentido, en esta tesis se va a proponer una idea innovadora que trate de identificar y que sea útil sobre diversas aplicaciones.

### **1.3 La Tesis: su objetivo**

Los sistemas de reconocimiento de formas han sido profundamente estudiados durante los últimos años, siendo referentes para el establecimiento de estándares de video, aplicaciones industriales y reconocimiento de patrones en general sobre diversas aplicaciones. Muchos trabajos se han desarrollado en esta línea, y en esta tesis doctoral se diseña, analiza, desarrolla, aplica y se chequea un nuevo método para la detección de formas. El presente documento tratará de demostrar la siguiente hipótesis:

*“Es posible diseñar un sistema de identificación de formas a partir de diversos datos de su contorno mediante una transformación de los modelos ocultos de Markov como elemento parametrizador, con independencia de la variabilidad del contorno.”*

Por tanto, el objetivo de esta tesis es demostrar y aplicar la anterior hipótesis en aplicaciones reales. Además, en esta tesis, se ha tratado de desarrollar la demostración matemática que justifica el uso del HMM como un nuevo elemento parametrizador, así como una metodología de uso para que se pueda adaptar a cualquier aplicación.

El reto de generar una nueva teoría, no solo se basa en su la propia generación de nuevo conocimiento, sino en su aplicación de forma correcta, es por ello que se decidió aplicarlo a diversas aplicaciones sobre diversas bases de datos, las cuales tuviesen diferentes características en cuanto a su naturaleza y a su forma o contorno final. De esta manera, se podría modelar y dar a conocer el comportamiento de esta nueva parametrización frente a esas aplicaciones, para que se pudiera entender y demostrar su generalización a cualquier otra aplicación no estudiada en este documento.

### **1.4 Metodología**

La metodología general que se siguió fue la siguiente. En primer lugar, se realizó un preprocesado de las imágenes para acondicionarlas, generalizado para todas las bases de datos, pero con pequeñas particularizaciones en función de la idiosincrasia de cada una de ellas, para posteriormente obtener su contorno. A continuación, se llevo a cabo la parametrización del contorno basada en información angular para lograr invarianza en el tamaño, en la rotación y en la traslación.

Para analizar el poder de discriminación de la información del contorno, se aplicaron dos sistemas de clasificación diferentes. Estos sistemas fueron particularizados para un esquema de identificación, bajo una clasificación supervisada y usando la metodología de validación cruzada *Hold-out*. En primero de estos dos sistemas de clasificación fueron HMMs (sistema clásico), en donde se varió el número de parámetros de entrada y el número de estados. En segundo sistema que se implementó, fue la propuesta de la tesis, es decir, se llevó a cabo la transformación de los parámetros de contorno mediante el kernel desarrollado y el resultado de esta transformación, se utilizó como datos de entrada en los SVMs, donde se comparó los resultados obtenidos de implementar un kernel lineal y kernel RBF, propios del SVM. En el segundo enfoque, se variaron los parámetros de entrada, el número de estados, los puntos de control para la definición del contorno, el porcentaje de muestras utilizadas para entrenar los sistemas y los parámetros del kernel RBF. Cada una de estas variables será expuesta en detalle en el capítulo 3.

Para alcanzar con éxito el objetivo de esta tesis, se desarrolló una experimentación para tratar de generalizar la teoría propuesta, se planteó una validación *Hold-out* que permitiera validar la teoría propuesta y abordarla desde distintos grados de complejidad en función del tipo de identificación a abordar. Así, se estudiaron problemas de complejidad simple, media y alta.

Para desarrollar esta tesis, se validó la metodología con una base de datos sintética, que sirvió como campo de pruebas de la hipótesis propuesta. Posteriormente, se realizaron distintos estudios sobre 8 aplicaciones reales de distinta naturaleza y distinta complejidad en su forma, tres simples, dos moderadas y tres de alta complejidad; hasta un total de 12 bases de datos.

## **1.5 Contribuciones y Resultados**

La Tesis presenta diversas contribuciones y resultados científicos, los cuales avalan la consecución del objetivo marcado para esta tesis. Fundamentalmente, se puede afirmar que se ha desarrollado una nueva teoría de identificación de formas, la cual ha sido chequeada frente a nueve diferentes bases de datos, que presentaban tres grados de variabilidad en su forma, simple, moderada y compleja. Véase a continuación, las contribuciones generadas y como se han elaborado.



Se creó una base de datos sintética, en donde se demostró la viabilidad del sistema propuesto y además se presentaron los primeros resultados sobre la robustez del sistema frente al ruido en los contornos [Travieso et al., 2007].

En los trabajos realizados con bases de datos con variabilidad simple [Travieso et al., 2007], [Travieso et al., 2011], [Briceño et al., 2002], [Briceño et al., 2007], se alcanzó un alto porcentaje de acierto alcanzando incluso el 100%, cuando se encontraron los parámetros óptimos [Briceño et al., 2002], [Briceño et al., 2007]. Se confirma la robustez del sistema frente a la presencia de un alto grado de ruido en base de datos reales [Briceño et al., 2007].

En las bases de datos con contorno de variabilidad moderada [Briceño et al., 2010], [Briceño et al., 2011] se alcanzó, con los parámetros óptimos, porcentajes mayores al 99% de acierto utilizando en algunos casos, solo 2 y 1 muestra de entrenamiento respectivamente. Lo que proporciona un sistema eficiente desde del punto de vista funcional y computacional. Se demuestra además que el sistema resulta aún más eficiente cuando se redimensionan las imágenes a una escala de reducción optima, aumentando el porcentaje de éxito y bajando los tiempo de cómputo [Briceño et al., 2010].

Se generaliza el sistema para el problema de verificación en manos, el cual resulta ser muy prometedor dados los excelentes resultados obtenidos, alcanzando hasta un 100% de éxito. Este estudio se realiza utilizando bases de datos y públicas propias (generadas por el grupo de investigación donde se desarrolló la tesis) y otras bases de datos públicas, llevando a cabo experimentos incluso en el rango del infrarrojo cercano del espectro electromagnético, en donde se validan los resultados obtenidos en el rango del visible, el que normalmente se había usado.

Finalmente, se confirma la eficacia y eficiencia del sistema cuando se tratan bases de datos con contorno de variabilidad compleja tales como de palabras, de firmas [Travieso et al., 2003] y de signos dactilológicos [Briceño et al., 2009]. La tasa de acierto alcanzada fue mayor que 98% y 99%.

El sistema propuesto mediante la transformación de los datos utilizando el kernel desarrollado en esta tesis, es eficaz para el reconocimiento de formas utilizando solo la información de contorno con variabilidad simple, moderada y compleja. Se demuestra la robustez frente al ruido y la eficiencia en las operaciones y resultados según las

condiciones de contorno. Así mismo, se presenta un sistema extensible para tratar problemas de verificación.

## **1.6 Estructura de la Memoria**

Esta memoria está compuesta por un total de 6 capítulos, así como las referencias y un listado de acrónimos. Los capítulos están ordenados en esta memoria de la siguiente forma:

En el capítulo 1 se lleva a cabo la introducción en donde se repasa el estado del arte y se describe la motivación para llevar a cabo este trabajo. Se definen la hipótesis de la tesis y los objetivos que se persiguen, y se describe la metodología y las contribuciones y los resultados alcanzados.

El capítulo 2 desarrolla y demuestra matemáticamente la hipótesis planteada. Se describe en detalle la parametrización implementada a lo largo de todos los estudios realizados y finalmente, se justifica y describe el sistema de clasificación utilizado.

En el Capítulo 3 se presenta el esquema básico de los sistemas de reconocimiento de patrones implementado en los trabajos de esta tesis doctoral. Se describen en detalle cada uno de los módulos que lo componen y las bases de datos utilizadas en cada uno de los estudios. Se define la metodología utilizada en las etapas de experimentación.

El capítulo 4 presenta en detalle todos los resultados de los experimentos realizados en los distintos estudios.

En el capítulo 5 se analizan los resultados obtenidos en el capítulo anterior. Este análisis se lleva a cabo tomando en cuenta el tipo de estudio y el tipo de base de datos para extraer las conclusiones particulares que aporta cada uno de los estudios tomando en cuenta sus singularidades.

Finalmente, en el capítulo 6 se presentan las conclusiones y líneas futuras derivadas de esta tesis doctoral.

## CAPÍTULO 2

# Demostración de la hipótesis

---

En el presente capítulo, se demuestra matemáticamente, se desarrolla y se muestra como aplicar la hipótesis de la tesis doctoral, y finalmente, se establece un nuevo método que utiliza el contorno, como fuente de información para la clasificación y reconocimiento de formas simples<sup>1</sup>. En el enfoque abordado, el contorno de la forma se caracteriza por puntos de control que a su vez se expanden o transforman y se proyectan sobre espacios hiperdimensionales, mediante la aplicación de Modelos Discretos Ocultos de Markov (DHMM) [Rabiner ,1989]. Esta transformada, se consigue mediante el uso del “kernel” desarrollado en esta tesis. Finalmente se justifica y se usan Máquinas de Vectores de Soporte (SVM) [Vapnik, 1995] como el sistema clasificador.

---

<sup>1</sup> Entendiéndose como aquellas no compuestas por varias formas independientes.

## 2.1 Planteamiento de la hipótesis.

A partir de la información dada por puntos de control, que definen el contorno de una forma cualesquiera, se plantea su procesado usando la interpretación del *kernel* de Fisher sobre un Modelo Discreto Oculto de Markov; pasando a un espacio hiperdimensional y produciendo entonces, una herramienta eficiente para la identificación automática de formas.

## 2.2 Parametrización: Codificación de la forma

Con el fin de obtener los puntos de control, atributos de caracterización del contorno, y transformarlos a un espacio hiperdimensional, se implementa una parametrización basada en su codificación. En primer lugar, las imágenes de los objetos en estudio son binarizadas, aplicando el método de Otsu, mientras que el contorno se obtiene utilizando la detección del borde [Otsu, 1979]. Al final se obtiene un perímetro cerrado y esqueletizado, de un píxel de ancho. La caracterización del borde definida por las posiciones de los píxeles en el perímetro  $(\lambda, \mu)$ , se obtienen utilizando un proceso de sombreado (forma blanca sobre fondo negro), filtrando los puntos aislados y finalmente estableciendo una localización automática de los puntos del perímetro, con la codificación  $\lambda = \text{columna}$  y  $\mu = \text{fila}$ , por un proceso de seguimiento de puntos continuos.

El siguiente procedimiento podría iniciar en cualquier punto del borde aunque, por aspectos de normalización en una descripción ordenada, empieza en el punto más alto de la imagen. A continuación recorre los píxeles del borde en el sentido de las agujas del reloj, retrocede cuando se encuentra en bucles sin salida, obviando éstos como desechables, finaliza cuando todos los píxeles del perímetro han sido considerados. En esta fase, el resultado final es una descripción cartesiana de la localización de los puntos del perímetro  $\{(\lambda_i, \mu_i) \mid i = 1, \dots, n\}$ , representando el contorno cerrado de la forma del objeto en estudio con un píxel de ancho. La idea general es la de considerar los píxeles del perímetro  $(\lambda, \mu)$  como puntos de una relación de una dimensión, es decir el grafo de puntos  $(\lambda, F(\lambda))$  de una función de descripción  $F$ , 1-D, representando el contorno de una imagen en el plano 2-D.

Se hace la consideración arbitraria para las coordenadas  $\mu$  como  $\mu = F(\lambda)$ , por la forma en que el procedimiento de estructuración simplifica la selección de puntos:

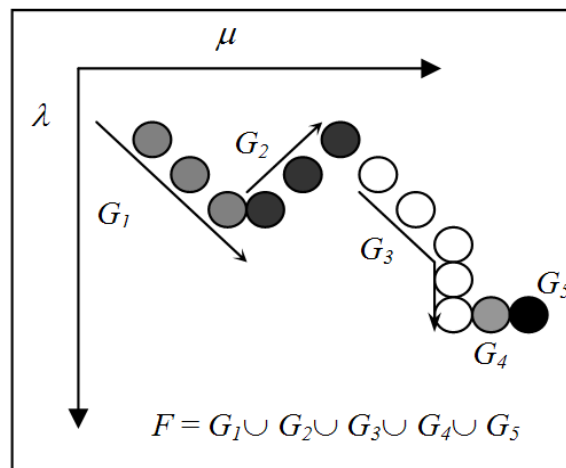
necesita del máximo tamaño sobre la ordenada  $\lambda$  por lo que, para efectos prácticos, se considera según sea el caso, la imagen rotada en  $90^\circ$ .

La reducción de atributos de caracterización, reduciendo el número de puntos de control, la normalización del tamaño y el control crítico en la selección de los puntos del perímetro, se obtienen fácilmente mediante la implementación del procedimiento de estructuración. La estructura resultante, que aglutina puntos con comportamientos similares, ayuda a mejorar la interpretación de los estados del HMM y facilita de esta forma su ajuste.

Este procedimiento está basado en la idea que, un trazo de píxeles en una imagen en blanco y negro, puede describirse como un grafo  $G_f$  de una aplicación  $f$  o trayectoria en una dimensión, si se conserva la definición de secuencias correctas; es decir conservando el orden establecido o de comportamiento monótono en la ordenada  $\lambda$ . Es decir  $G_f = \{(\lambda_i, \mu_i) \mid \mu_i = f(\lambda_i), i = 1, \dots, n\}$ , es un grafo que describe la relación, si los puntos de ordenadas  $\lambda_i$  del trazo  $f$ , deben estrictamente y exclusivamente ascender o descender; por lo tanto, los  $\lambda_i$ , deben ser de la forma:  $\lambda_i < \lambda_{i+1}$  o  $\lambda_{i+1} < \lambda_i$  para  $i = 1, \dots, n-1$ . Considerando el perímetro completo, se define su relación de descripción  $F$  por partes: como la definición parcial de aplicaciones de trayectoria 1-D (con grafos  $G_j$ ) conservando el comportamiento monótono. Esto implica que  $G$  es la relación general de un contorno tal que,  $G = \bigcup_{j \in J} G_j$  donde cada  $G_j$  es un conjunto relacional de puntos y cada uno de ellos es obviamente, por su calidad monótona, una parte relacional de información de contorno. Como resultado, las aplicaciones de restricciones de trayectoria (o parte codificada del borde)  $f_j = F_{\{\lambda_\alpha \mid \alpha \in J_j\}}$ , es decir: para puntos  $x_\alpha$  del contorno, de subíndice  $\alpha$ , pertenecientes a un conjunto ordenado  $J_j$  dependiente del trozo de contorno considerado, son de tal forma que el punto inmediatamente siguiente al último de  $G_j$ , es el primero de  $G_{j+1}$ . En consecuencia los grafos  $G_j$  son correctas trayectorias  $f_j$  de aplicaciones de descripción parcial del contorno total  $G$ .

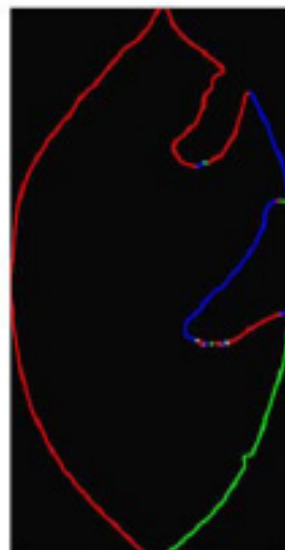
La construcción de los conjuntos  $G_j$  es una operación muy sencilla:

- A partir de un primer punto se incluye el siguiente de  $F$
- En el instante que este punto no conserva un comportamiento monótono se empieza con un nuevo  $G_{j+1}$ .
- El proceso se detiene cuando todos los puntos de  $F$  están asignados.



**Figura 2.1:** Ejemplo de un fragmento de un contorno de imagen descompuesto en  $G_j$  grafos.

Con el fin de simplificar y evitar que  $G_j$  se reduzca a un píxel, como es el caso de  $G_4$  y  $G_5$  en la figura. 2.1, se conserva solo el primer punto  $\lambda$  constante, de la serie de ordenadas. Se observa que la estructuración de  $G$ , por las descripciones parciales de grafos  $G_j$ , toma en cuenta los cambios bruscos de dirección en la descripción del perímetro y que pueden por lo tanto ser interpretados como estados de una cadena Markoviana [Ross, 2007].



**Figura 2.2:** Ejemplo de estructuración por grafos  $G_i$  en una hoja fuertemente depredada. La presentación de los grafos se hace siguiendo la dirección contraria de las manecillas del reloj, y en orden de secuencia de colores: rojo, verde, azul.

Con el fin de definir  $p \geq n$  codificando los puntos de control  $(\lambda_i, \mu_i)_{i=1, \dots, p}$ , después de construir los  $G_j$  se seleccionan todos los primeros  $n$  puntos de cada  $G_j, j= 1, \dots, n$  y luego se completa la descripción de los puntos del perímetro mediante  $k= n-p$  puntos, seleccionando una distribución uniforme para cada  $G_j$  y proporcional a su tamaño.

Por otra parte, con el propósito de crear una codificación independiente de la referencia de origen, rotación, traslación y escala, se lleva a cabo una transformación angular de la posición de los puntos de los contornos codificados de la forma descrita anteriormente. Para un contorno codificado dado, con  $p$  puntos posicionales de control  $G = \{X_i = (\lambda_i, \mu_i) | i=1, \dots, p\}$  siendo el punto  $C_0$  el centroide o baricentro y el  $X_0$  punto de inicio del borde; se definen  $\theta_i$  y  $\rho_i$  los ángulos referidos por  $C_0$  y  $X_i$  como:  $\theta_i = \text{angle}(C_0, X_i, X_{i+1})$  y  $\rho_i = \text{angle}(X_i, C_0, X_{i+1})$ . Se observa que tal codificación interpreta dependencias punto a punto de secuencias de cadenas Markovianas, ya que para posicionar el punto actual, necesito de la descripción angular del punto anterior.

A continuación la secuencia de puntos posicionales  $(\lambda_i, \mu_i) i=1, \dots, p$ , son transformados en secuencias de puntos de representación angular de origen independiente  $(\theta_i, \rho_i)$  donde  $i = 1, \dots, p-1$ . Nótese que, dado el centroide  $C_0$  y la elección del punto de inicio en la figura 2.3 toma en cuenta la escala y la orientación del objeto, así como las propiedades geométricas de similitudes triangulares, hacen de tal secuencia de codificación del objeto en estudio, independientes de la localidad, orientación y del tamaño de dicho objeto.

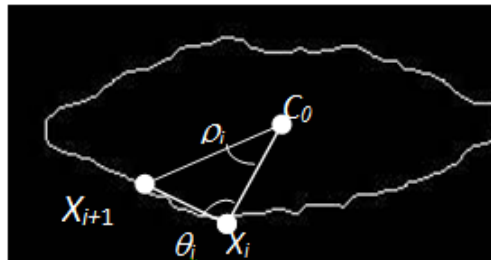


Figura 2.3: Ejemplo general de una codificación angular.

### 2.3 Desarrollo del núcleo: Espacios Hiperdimensionales.

Los métodos probabilísticos como el HMM, tienen la ventaja de que permiten trabajar con secuencias de tamaño variable y con fragmentos desconocidos de estas secuencias. Por lo tanto, el HMM permite aprovechar la caracterización de los contornos, ya que éste puede almacenarse en forma vectorial. Los modelos Discretos Ocultos de Markov vienen definidos en [Rabiner ,1989] y [Rabiner y Juang, 1993].

La característica fundamental de un proceso de Markov es que la descripción probabilística del estado actual en el tiempo  $t$  únicamente depende del estado actual y del estado anterior. En los procesos estocásticos de cadenas Markovianas [Rabiner y Juang, 2003], los estados están determinísticamente asignados a los valores de secuencias observables. Esto produce dificultades a la hora de evaluar adecuadamente largas secuencias de estados frente a secuencias cortas, de manera de no sesgar la medida probabilística de la presencia de unos estados frente a la presencia de otros. De esta forma, un HMM tiene dos procesos estocásticos asociados, agregando un proceso estocástico subyacente que no es observable u oculto. Este proceso oculto únicamente puede ser visto a través de otro conjunto de procesos estocásticos que producirían la secuencia de observaciones. El otro es observable y está asociado con la probabilidad de obtener cada uno de los posibles valores de la salida.

#### 2.3.1 Modelos de Markov

Se considera un sistema como el de la figura 2.4, formado por un conjunto de estados  $S_1, S_2, \dots, S_N$ , con  $N=5$  (para simplificar). Este sistema sufrirá regularmente cambios de estados, de acuerdo a un conjunto de probabilidades asociado al estado. Los instantes en los que el sistema experimenta cambios, se les denomina como  $t=1,2\dots$  y se denota el estado actual en el tiempo  $t$  como  $q_t$ .

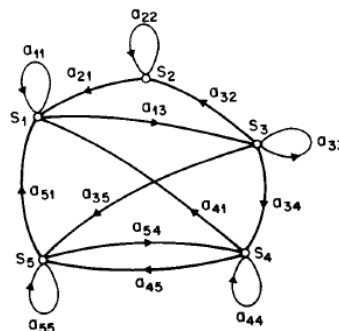


Figura 2.4: Modelo de estados de Markov.



La característica fundamental de un proceso de Markov es que la descripción probabilística del estado actual en el tiempo  $t$  únicamente depende del estado actual y del estado anterior. Por lo tanto las probabilidades de transición entre estados se podrán definir como:

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], \quad 1 \leq i, \quad j \leq N \quad (2.1)$$

Donde los coeficientes  $a_{ij}$  presentan las siguientes propiedades:

$$a \geq 0 \quad (2.2)$$

$$\sum_{j=1}^N a_{ij} = 1 \quad (2.3)$$

### 2.3.2 Extensión a los modelos ocultos de Markov

Se extiende el concepto de Markov para incluir aquellos casos donde la observación es una función probabilística del estado. Por lo tanto, el modelo resultante es un doble proceso estocástico, con un proceso estocástico subyacente que no es observable (oculto). Este proceso oculto únicamente puede ser visto a través de otro conjunto de procesos estocásticos que producen la secuencia de observaciones. De esta forma se puede modelar un experimento con estas características mediante infinitos modelos ocultos de Markov, cada vez con más grados de libertad.

### 2.3.3 Elementos de un HMM

Un modelo HMM se caracteriza por lo siguientes elementos:

- $N$ , es el número de estados del modelo. A pesar de que los estados son ocultos, para muchas aplicaciones prácticas hay algún significado físico que permite distinguirlos. Se denota el conjunto de estados como  $S = \{S_1, S_2, \dots, S_N\}$  y el estado del instante  $t$  será  $q_t$ . Para el ejemplo de las monedas, los estados son las monedas que se lanzan.

- $M$ , es el número de símbolos distintos por estado, es decir, el tamaño del alfabeto discreto. Los símbolos corresponden con la salida física del sistema. Se denota el conjunto de símbolos como:  $V=\{V_1, V_2, \dots, V_M\}$ . Para el ejemplo de las monedas, eran únicamente caras y cruces.
- La distribución de probabilidades de transición de estados  $A=\{a_{ij}\}$  será de la siguiente forma:

$$a_{ij} = P[q_{t+1} = S_j | q_t = S_i], \quad 1 \leq i, j \leq N \quad (2.4)$$

En el caso particular de que cualquier estado pueda alcanzar el resto de estados, se tendrá que  $a_{ij} > 0$  en todos los casos.

- La distribución de la probabilidad de símbolo en el estado  $j$ ,  $B=\{b_j(k)\}$ , tiene el siguiente aspecto:

$$b_{ij} = P[v_k \text{ at } t | q_t = S_j], \quad \begin{array}{l} 1 \leq j \leq N \\ 1 \leq k \leq M \end{array} \quad (2.5)$$

- La distribución inicial de estado será  $\pi=\{\pi_i\}$ , con:

$$\pi_j = P[q_1 = S_j], \quad 1 \leq j \leq N \quad (2.6)$$

Por lo tanto, dados valores adecuados de  $N$ ,  $M$ ,  $A$ ,  $B$  i  $\pi$ , el HMM puede proporcionar una secuencia de observación  $O=O_1, O_2, \dots, O_T$ , donde cada observación  $O_t$  es uno de los símbolos  $V_t$  y  $T$  es el número de observaciones en la secuencia.

Para realizar una especificación completa de un modelo HMM, se requiere lo siguiente:

- Especificación de los parámetros de modelo  $N$  y  $M$ .
- Especificación de los símbolos.
- Especificación de las tres probabilidades  $A$ ,  $B$  y  $\pi$ . Para denotar la correcta especificación de los parámetros de un HMM se utilizará  $\lambda(A, B, \pi)$ .

### 2.3.4 Los tres problemas básicos del HMM

Los modelos ocultos de Markov plantean tres problemas fundamentales que deben ser resueltos, para que el sistema sea aplicable a casos reales. A continuación se enumeran y se describe de forma breve su resolución.

#### Problema 1: problema de reconocimiento

Consiste en calcular de la secuencia de observaciones  $O=O_1, O_2, \dots, O_T$ , dado un modelo  $\lambda$ , es decir,  $P(O|\lambda)$ . Esta probabilidad se puede calcular como:

$$\begin{aligned}
 P(O|\lambda) &= \sum_{\text{all } Q} P(O|Q, \lambda) P(Q|\lambda) \\
 &= \sum_{q_1, q_2, \dots, q_T} \pi_{q_1} b_{q_1}(O_1) a_{q_1, q_2} b_{q_2}(O_2) \\
 &\quad \dots a_{q_{T-1}, q_T} b_{q_T}(O_T)
 \end{aligned} \tag{2.7}$$

Para realizar la operación mostrada en la ecuación. 2.7, el número total de operaciones es de orden exponencial, por esta razón, se utiliza el procedimiento de *adelante-atrás*. En primer lugar, en el procedimiento hacia delante, se define la variable hacia delante  $\alpha_t(i)$ :

$$\alpha_t(i) = P(O_1 O_2 \dots O_t, q_t = S_i | \lambda) \tag{2.8}$$

De forma análoga, se define la variable hacia atrás  $\beta_t(i)$ :

$$\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T, q_t = S_i | \lambda) \tag{2.9}$$

En este caso, lo que se define es la probabilidad conjunta de que la secuencia parcial de observaciones, desde el instante  $t$  al final  $T$ , y el estado  $S_i$  en el instante  $t$  sean producidos por el modelo  $\lambda$ . Finalmente, hay que decir que la diferencia entre el procedimiento hacia delante y el procedimiento hacia atrás radica en el paso de la inducción [Rabiner y Juang, 1993]

Problema 2: asignación de estados

Se aplica el criterio de elegir aquellos estados que son individualmente más probables. Para aplicar este criterio, se define la siguiente variable:

$$\gamma_t(i) = P(q_t = S_i | O, \lambda) \quad (2.10)$$

Es decir, la probabilidad de estar en un estado concreto dados una secuencia de observación y un modelo concretos. Esta variable también puede ser expresada en términos de las variables hacia delante y hacia atrás, definidas anteriormente, de la siguiente forma:

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O|\lambda)} = \frac{\alpha_t(i)\beta_t(i)}{\sum_{i=1}^N \alpha_t(i)\beta_t(i)} \quad (2.11)$$

Utilizando este parámetro definido, se puede decidir que el estado más probable será aquel que cumpla la relación que se muestra a continuación:

$$q_t = \underset{1 \leq i \leq N}{\operatorname{arg\,max}} [\gamma_t(i)], \quad 1 \leq t \leq T \quad (2.12)$$

De esta forma, se maximiza el número esperado de estados correctos. Finalmente, se buscará la secuencia única y mejor, o lo que es lo mismo, se maximizará  $P(Q|O, \lambda)$ , que es equivalente a maximizar  $P(Q, O|\lambda)$  utilizando el algoritmo de Viterbi [Rabiner, 1989].

Problema 3: optimización de los parámetros del modelo

El tercer problema es el de encontrar un método para ajustar los parámetros  $\lambda(A, B, \pi)$  de un modelo, para maximizar las probabilidades de que un modelo haya generado una determinada observación. Para esto se utilizan técnicas iterativas, en este caso se ha utilizado un método basado en el trabajo de Baum [Rabiner, 1989] y que se describe a continuación.

En primer lugar, se define el parámetro  $\xi_t(i, j)$  como la probabilidad de estar en el estado  $S_i$  en el instante  $t$  y en el estado  $S_j$  en el instante  $t+1$ , dados un modelo y un secuencia de observaciones concretos.

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O, \lambda) \quad (2.13)$$

Mediante las variables hacia delante y hacia detrás, definidas anteriormente, se puede definir esta nueva variable de la siguiente forma:

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(i)}{P(O | \lambda)} = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(i)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(i)} \quad (2.14)$$

Se puede observar que se ha realizado una normalización respecto a la definición inicial para obtener la probabilidad deseada. Anteriormente, se ha definido el parámetro  $\gamma_t(i)$  como la probabilidad de estar en el estado  $S_i$  en el instante  $t$ , por lo tanto, se puede relacionar este parámetro con  $\xi_t(i, j)$  de la siguiente forma:

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j) \quad (2.15)$$

Se define entonces para la reestimación de los parámetros  $\lambda$  de un modelo HMM:

$$\begin{aligned} \bar{\pi}_i &= \text{frecuencia esperada de visitas (número de visitas) } S_i \text{ en el instante } (t = 1) \\ &= \gamma_1(i) \end{aligned} \quad (2.16)$$

$$\begin{aligned} \bar{a}_{ij} &= \frac{\text{número esperado de transiciones del estado } S_i \text{ a } S_j}{\text{número esperado de transiciones desde el estado } S_i} \\ &= \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(j)} \end{aligned} \quad (2.17)$$

$$\begin{aligned} \bar{b}_j(k) &= \frac{\text{número esperado de veces en el estado } j \text{ observando el símbolo } v_k}{\text{número esperado de veces en el estado } j} \\ &= \frac{\sum_{t=1}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)} \end{aligned} \quad (2.18)$$

Para el cálculo de los definitivos parámetros  $\lambda$ , lo que se hará, será aplicar las fórmulas propuestas de forma iterativa, reestimando el valor de los parámetros del modelo, hasta que estos parámetros converjan a unos determinados valores.

Entre los HMM existentes, se eligió el tipo de HMM llamado *left-right* o modelo Bakis [Rabiner, 1989] porque es particularmente apropiado para la evaluación de secuencias de contornos de formas (véase la figura 2.5).



Figura 2.5: Estados del modelo *left-right*.

El modelo es llamado *left-right* fundamentalmente porque la secuencia de estados asociada con el modelo tiene la propiedad que: a medida que la variable  $t$  se incrementa, el índice del estado aumenta (o se mantiene en el mismo); es decir, los estados avanzan de izquierda a derecha. Claramente el tipo HMM *left-right* tiene la propiedad deseada de modelar fácilmente señales cuyas propiedades cambian a lo largo del tiempo de forma progresiva. La propiedad fundamental de todos los HMMs *left-right* es que los coeficientes de transición de estados tienen la propiedad.

$$a_{ij} = 0, \quad j < i \quad (2.19)$$

Es decir, no se permiten transiciones a estados cuyos índices son menores que el estado actual. Por lo tanto, las probabilidades del estado inicial tiene la propiedad:

$$\pi_i = \begin{cases} 0, & i=2, \dots, N \\ 1, & i=1 \end{cases} \quad (2.20)$$

dado que la secuencia de estados debe empezar por el estado 1 y terminar en el estado  $N$ . Normalmente cuando se usan modelos *left-right*, se fijan restricciones adicionales a los coeficientes de transición de estados para asegurar que no ocurran cambios, producto de saltos excesivos entre estados.

### 2.3.6 Adaptación del evaluador de Fisher

Definido los modelos ocultos de Markov, puesto que por éstos se produce la evaluación probabilística de secuencias de contorno y que a su vez, estos modelos están caracterizados por sus estados, se describe a continuación el procedimiento para calcular el evaluador de Fisher desarrollado, para transformarlos utilizando DHMM [Rabiner, 1989], [Rabiner y Juang, 1993].

Sólo se consideran los gradientes sobre las probabilidades de emisión del DHMM, puesto que se considera la ecuación:  $U_x = \nabla \log P(X|\lambda)$  donde se ha usado un modelo Bakis, debido al carácter secuencial del contorno. Para ello se van a dar los siguientes pasos:

Paso 1. Se define la probabilidad de emisión de un residuo  $x$  (de la cadena de observación  $X = (x_1, \dots, x_n)$ ) del alfabeto, mientras se está en el estado  $s \in \{s_1, \dots, s_n\}$  por  $P(x|s, \theta) = \theta_{x|s}$ .

a. Nótese que:

$$\theta_{x|s} = P(x|s, \lambda) = b_s(x) \quad (2.21)$$

b. Con la propiedad:

$$\forall s \quad \sum_x \theta_{x|s} = \sum_x P(x|s, \lambda) = \sum_x b_s(x) = 1 \quad (2.22)$$

por la implementación del DHMM

i. Donde por supuesto se define:  $b_s = P(x = v_k | s) 1 \leq k \leq M$  donde los  $v_k$  son las etiquetas de cuantificación del DHHH

Paso 2. Se define la probabilidad de transición del estado  $s$  al estado  $s'$  por:

$$P(s'|s, \tau) = \tau_{s'|s}.$$

Nótese que:

$$\tau_{s'|s} = P(s'|s) = a_{s's}. \quad (2.23)$$

Para simplificar, se asume **un único estado inicial**  $s_0$  es decir:  $\pi_{s_0} = 1$ . Nótese que esto, para efectos computacionales, no es una pérdida de generalidad ya que en términos de la derivación los  $\pi_{s_i}$  son constantes pues la variable considerada es con respecto a  $\partial X$ , para secuencias de contorno  $X$ , donde la  $x$  está caracterizada por las  $\theta_{x|s} = b_s(x)$  (es decir, sólo depende de los descriptores  $x$ ). A esta dependencia se le va a denominar Propiedad 1.

Paso 3. El  $\lambda$  DHMM definido de esta forma, asigna una probabilidad para toda secuencia  $X = (x_1, x_2, \dots, x_T)$  dada:

$$\begin{aligned} P(X|\theta, \tau) &= \\ &= \sum_{s_1, \dots, s_m} \prod_i \theta_{x_i|s_i} \tau_{s_i|s_{i-1}}. \end{aligned} \tag{2.24}$$

Paso 4. Y, según las ecuaciones 2.10 y 2.12;

$$\begin{aligned} P(X|\theta, \tau) &= P(X|\lambda) = \sum_{s_1, \dots, s_n} \prod_i b_{s_i}(x_i) a_{s_i, s_{i-1}} \\ &= \sum_{s_1, \dots, s_n} b_{s_1}(x_1) a_{s_1 s_2} b_{s_2}(x_2) \dots a_{s_{n-1} s_n} b_{s_n}(x_n). \end{aligned} \tag{2.25}$$

donde la suma se efectúa sobre todas las posibles secuencias de estados.

Nuestro interés es el de calcular las derivadas de  $\log P(X|\theta, \tau) = \log P(X|\lambda)$ , con respecto a las probabilidades de emisión  $\theta_{x|s} = b_s(x)$  y tal como se señala en la Propiedad 1. El resultado de la derivación serán las componentes del vector evaluador  $U_X$ .

Por la ecuación 2.11, los vectores  $\theta_{x|s}$  están ligados conjuntamente, puesto que su suma es 1 para cualquier estado fijo  $s$ . Para poder implementar operaciones de derivación independientes, se debe implementar una descripción independiente.

Para lograrlo, se debe escribir los términos  $\theta_{x|s}$  en términos de un conjunto de parámetros independientes.



Paso 5. Sean entonces:  $\theta_{x|s} = \frac{\theta_{x,s}}{\sum_{x'} \theta_{x',s}}$ , de forma que los valores  $\theta_{x,s}$  se establecen

tales que;

$$\left( \sum_{x'} \theta_{x',s} = 1 \right) \Rightarrow \theta_{x,s} = \theta_{x|s}. \quad (2.26)$$

Se recuerda que los  $\theta_{x|s} = P(x|s, \lambda) = b_s(x)$  las probabilidades de emitir el símbolo  $x$  estando en el estado  $s$ . Nótese que la propiedad de dependencia sobre los  $\theta_{x|s}$  ya no aplica, sino que depende de la elección de los  $\theta_{x,s}$ .

Suponiendo todos los estados considerados como posibles, es decir que  $P(s) \neq 0$ , nótese que,  $P(x|s, \lambda) = \frac{P(x \wedge s|\lambda)}{P(s)}$  se puede simplificar por la notación:  $P(x \wedge s|\lambda) = P(x, s|\lambda)$  y entonces:  $P(x, s|\lambda) = P(s)P(x|s, \lambda)$ , que en términos de nuestra notación, se traduce por:

$$\theta_{x,s} = P(s)\theta_{x|s} \quad (2.27)$$

Se considera:  $\frac{\theta_{x,s}}{\sum_{x'} \theta_{x',s}}$  para algunos símbolos  $x'$  y todo símbolo  $x$ .

Es decir, por (2.15) y (2.16):

$$\frac{\theta_{x,s}}{\sum_{x'} \theta_{x',s}} = \frac{P(s)\theta_{x|s}}{\sum_{x'} \theta_{x',s}} \quad (2.28)$$

Se define la Propiedad 2:

Se considera todos los símbolos posibles  $x'$  que se puedan observar en el estado  $s$ , y se define para todo  $s$ .

Es decir:  $P(s) = \sum_{x'} P(x', s|\lambda) = \sum_{x'} \theta_{x',s}$  que se realiza para algún DHMM y entonces:

$$\frac{P(s)\theta_{x|s}}{\sum_{x'}\theta_{x',s}} = \theta_{x|s} \quad (2.29)$$

Para los  $x'$  que cumplen la Propiedad 2 por (2.18) se tiene que:

$$\theta_{x,s} = \theta_{x|s} \quad \forall x, \forall s. \quad (2.30)$$

Y, como se enunció anteriormente, obtener (2.18) en  $\theta_{x|s} = \frac{\theta_{x,s}}{\sum_{x'}\theta_{x',s}}$  es equivalente

a suponer que  $\sum_{x'}\theta_{x',s} = 1$ .

Entonces los parámetros  $\theta_{x|s}$  así definidos son *independientes*.

Paso 6. Consecuentemente, derivando por partes:

$$\begin{aligned} \frac{\partial}{\partial \theta_{\tilde{x},\tilde{s}}} \log P(X|\theta, \tau) &= \frac{1}{P(X|\theta, \tau)} \sum_{s_1, \dots, s_n} \frac{\partial}{\partial \theta_{\tilde{x},\tilde{s}}} \prod_i \frac{\theta_{x_i, s_i}}{\sum_{x'} \theta_{x', s_i}} \tau_{s_i | s_{i-1}} \\ &= \frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} \log P(X|\lambda) = \frac{1}{P(X|\lambda)} \sum_{s_1, \dots, s_n} \frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} \prod_i \frac{b_{s_i}(\tilde{x}_i)}{\sum_{x'} b_{s_i}(x')} a_{s_i s_{i-1}} \end{aligned} \quad (2.31)$$

Paso 7. Se considera los términos de la sumatoria en (2.20):

$$\begin{aligned} \frac{\partial}{\partial \theta_{\tilde{x},\tilde{s}}} \prod_i \frac{\theta_{x_i, s_i}}{\sum_{x'} \theta_{x', s_i}} \tau_{s_i | s_{i-1}} &= \sum_k \left[ \frac{\partial}{\partial \theta_{\tilde{x},\tilde{s}}} \left( \frac{\theta_{x_k, s_k}}{\sum_{x'} \theta_{x', s_k}} \right) \right] \tau_{s_k | s_{k-1}} \prod_{i \neq k} \frac{\theta_{x_i, s_i}}{\sum_{x'} \theta_{x', s_i}} \tau_{s_i | s_{i-1}} \\ \frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} \prod_i \frac{b_{s_i}(x_i)}{\sum_{x'} b_{s_i}(x')} a_{s_i s_{i-1}} &= \sum_k \left[ \frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} \left( \frac{b_{s_k}(x_k)}{\sum_{x'} b_{s_i}(x')} \right) \right] a_{s_k s_{k-1}} \prod_{i \neq k} \frac{b_{s_i}(x_i)}{\sum_{x'} b_{s_i}(x')} \end{aligned} \quad (2.32)$$

Paso 8. Aplicando la regla de la cadena en 2.21;

$$\left( \text{si: } f(x) = \frac{u(x)}{v(x)} \text{ entonces: } f'(x) = \frac{u'(x)}{v(x)} - u(x) \frac{v'(x)}{v^2(x)} \right)$$

$$\frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} \left( \frac{b_{s_k}(x_k)}{\sum_{x'} b_{s_i}(x')} \right) = \frac{\frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} b_{s_k}(x_k)}{\sum_{x'} b_{s_i}(x')} - b_{s_k}(x_k) \frac{\sum_{x'} \frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} b_{s_i}(x')}{\left( \sum_{x'} b_{s_i}(x') \right)^2} \quad (2.33)$$

Paso 9. Puesto que las  $b_{s_i}(x_i)$  son discretas, sus derivadas son deltas de Dirac que pueden separarse en  $\tilde{x}$  y  $\tilde{s}$ , es decir usando (2.22), se reescribe como:

$$\frac{\frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} b_{s_k}(x_k)}{\sum_{x'} b_{s_i}(x')} - b_{s_k}(x_k) \frac{\sum_{x'} \frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} b_{s_i}(x')}{\left( \sum_{x'} b_{s_i}(x') \right)^2} = \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}}{\sum_{x'} b_{s_i}(x')} - b_{s_k}(x_k) \frac{\sum_{x'} \delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}}{\left( \sum_{x'} b_{s_i}(x') \right)^2} \quad (2.34)$$

Paso 10. Y, puesto que por  $\sum_{x'} b_{s_i}(x') = 1$ , y partiendo de (2.20), se obtiene,

$$\sum_k \left[ \delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} - b_{s_k}(x_k) \delta_{s_k, \tilde{s}} \right] a_{s_k s_{k-1}} \prod_{i \neq k} b_{s_i}(x_i) a_{s_i s_{i-1}}$$

$$= \sum_k \left[ \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}}{b_{s_k}(x_k)} - (x_k) \delta_{s_k, \tilde{s}} \right] \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}} \quad (2.35)$$

Aplicando (2.15) sobre (2.24), se obtiene;

$$\sum_k \left[ \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}}{b_{\tilde{s}}(\tilde{x})} - (x_k) \delta_{s_k, \tilde{s}} \right] \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}} \quad (2.36)$$

Paso 11. Retomando (2.25) sobre (2.20), se logra;

$$\frac{\partial}{\partial b_{\tilde{s}}(\tilde{x})} \log P(X|\lambda) = \sum_k \frac{1}{P(X|\lambda)} \sum_{s_1, \dots, s_n} \left[ \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}}{b_{\tilde{s}}(\tilde{x})} - \delta_{s_k, \tilde{s}} \right] \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}}$$
(2.37)

$$= \frac{1}{b_{\tilde{s}}(\tilde{x})} \sum_k \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}}}{P(X|\lambda)} - \sum_k \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}}}{P(X|\lambda)}$$

Definiendo la esperanza condicional de  $D$ , dado que  $Y = y$ , se define como [Ross, 2007];

$$E[D|Y = y] = \sum_d dP\{D = d|Y = y\}$$
(2.38)

Donde  $D$  es  $\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}$ ,  $Y$  es  $X$ ,  $d$  es  $k$ .

Se observa que en (2.27), se tiene la definición de esperanza condicional que aplicada sobre (2.26), se obtiene;

$$\frac{1}{b_{\tilde{s}}(\tilde{x})} \sum_k E[\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} | X, \lambda] - \sum_k E[\delta_{s_k, \tilde{s}} | X, \lambda]$$
(2.39)

Desarrollando la ecuación anterior, se tiene;

$$\frac{1}{b_{\tilde{s}}(\tilde{x})} \sum_k \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}}}{P(X|\lambda)} - \sum_k \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \prod_i b_{s_i}(x_i) a_{s_i s_{i-1}}}{P(X|\lambda)}$$
(2.40)

Se recuerda que en (2.40) las  $s_k, \tilde{s}$  denotan estados y las  $x_k, \tilde{x}$  denotan valores en la cadena de observación; las  $\delta_{s_k, \tilde{s}}, \delta_{x_k, \tilde{x}}$  son el resultado de la evaluación de observación conjunta de la transición de secuencia de estados  $s_k, \tilde{s}$  y de valores  $x_k, \tilde{x}$ .

Si se define la función  $g$  como esta evaluación, para una variable aleatoria  $X$  en los puntos de transición  $(s_k, \tilde{s})$  y  $(x_k, \tilde{x})$ . Por la propiedad del estadista inconsciente [Ross, 2007], se tendría que:

$$E[g(X)] = \sum_{\{(s_k, \tilde{s}), (x_k, \tilde{x})\}} g\{(s_k, \tilde{s}), (x_k, \tilde{x})\} p_X\{(s_k, \tilde{s}), (x_k, \tilde{x})\} \quad (2.41)$$

Donde  $p_X$  indica la función de probabilidad discreta de  $X$ .

Luego  $b_{s_i}(x_i)a_{s_i, s_{i-1}}$  indican la evaluación de transición conjunta de  $s_i$  observando  $x_i$  en  $s_i$  según el modelo HMM  $\lambda$ , es decir se puede presentar la ecuación de densidad condicional de la variable aleatoria  $X$  de evaluación en HMM, según esa elección de variables como:

$$\frac{\prod_i b_{s_i}(x_i)a_{s_i, s_{i-1}}}{P(X|\lambda)} \quad (2.42)$$

Por lo tanto, usando la definición de esperanza condicional, es decir aplicando la esperanza matemática a la función condicional sobre la variable aleatoria, y usando la propiedad del estadista inconsciente.

$$E[g(X)|X, \lambda] = \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} \prod_i b_{s_i}(x_i)a_{s_i, s_{i-1}}}{P(X|\lambda)} \quad (2.43)$$

El segundo término de la ecuación tienen un desarrollo similar y entonces:

$$\begin{aligned} & \frac{1}{b_{\tilde{s}}(\tilde{x})} \sum_k \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} \prod_i b_{s_i}(x_i)a_{s_i, s_{i-1}}}{P(X|\lambda)} - \sum_k \sum_{s_1, \dots, s_n} \frac{\delta_{s_k, \tilde{s}} \prod_i b_{s_i}(x_i)a_{s_i, s_{i-1}}}{P(X|\lambda)} \\ &= \frac{1}{b_{\tilde{s}}(\tilde{x})} \sum_k E[\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}} | X, \lambda] - \sum_k E[\delta_{s_k, \tilde{s}} | X, \lambda] \end{aligned} \quad (2.44)$$

Interpretando las deltas como pulsos, resultando entonces como un muestreo, el producto  $\delta_{s_k, \tilde{s}} \delta_{x_k, \tilde{x}}$  se puede ver como las visitas al estado  $\tilde{s}$ , es decir, el producto de la observación  $\tilde{x}$  a partir del estado  $s_k$ . La esperanza evaluaría entonces, las frecuencias condicionadas esperadas de esas visitas que corresponde a la definición del  $\xi(\tilde{x}, \tilde{s})$  según la nomenclatura del HMM. Finalmente, bajo esa perspectiva, se puede interpretar;

$$U_X = \nabla \log P(X|\lambda) = \frac{\xi(\tilde{x}, \tilde{s})}{b_{\tilde{s}}(\tilde{x})} - \xi(\tilde{s})$$

donde:  $\xi(\tilde{s}) = \sum_{x'} \xi(x', \tilde{s})$  (2.45)

### 2.3.7 Sistemas de clasificación por hiperdimensionalidad: uso del SVM

En la literatura se pueden encontrar múltiples referencias donde se utiliza SVM [Cherkassky y Mulier, 2007], como sistema de clasificación para grandes volúmenes de datos [Xiaoou et al., 2010], [Cervantes et al., 2007a], [Thanh-Nghi y Fekete, 2007], [Cervantes et al., 2007b]. Basado en el estado del arte, se decidió que el sistema de clasificación, para la clasificación del espacio hiperdimensional dado por la aplicación del kernel de Fisher al DHMM.

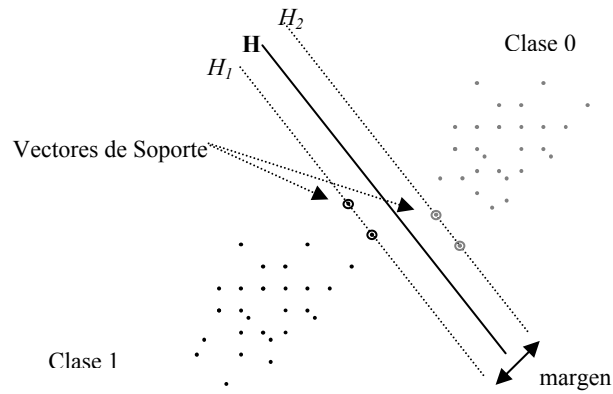
Por tanto, para llevar a cabo la clasificación de los datos transformados, se hace uso de las Máquinas de Vectores Soporte (SVM) desarrollado y generalizado por [Vapnik, 1995] basada en la minimización del riesgo estructural, del inglés *Structural Risk Minimization* (SRM).

La idea básica consiste en entrenar el sistema, para obtener dos conjuntos de vectores (en dos dimensiones correspondientes con los puntos de control) que representen la clase a identificar. Como consecuencia, se calcula el hiperplano de separación  $H$  (en dos dimensiones es un clasificador lineal) entre estos dos conjuntos. Los puntos correspondientes dentro del hiperplano deben satisfacer la siguiente ecuación [Burges, 1998]:

$$\omega \cdot x_i + \Psi = 0 \quad (2.46)$$

Donde  $\omega$  es normal al hiperplano,  $\Psi / \|\omega\|$  es la distancia perpendicular del plano al origen,  $\|\omega\|$  es la norma euclídeana de  $\omega$ ,  $\omega$  es el termino independiente y  $x_i$  es un punto en el hiperplano. La ecuación 2.29 es igual a 0 indicando el borde de decisión.

Por lo tanto se definen otros dos hiperplanos:  $H_1: x_i \cdot \omega + \psi = 1$  and  $H_2: x_i \cdot \omega + \psi = -1$ , los cuales contienen los vectores de soporte. La distancia entre  $H_1$  y  $H_2$  se denomina “margen”.



**Figura 2.6: Hiperplano de separación lineal para el caso de separación lineal.**

Una vez que el sistema ha sido entrenado, y por lo tanto, se ha obtenido la separación del plano, se debe determinar cuál es el límite de decisión (hiperplano localizado de forma equidistante entre  $H_1$  y  $H_2$ ). De acuerdo con esta decisión, se asigna la correspondiente etiqueta de clase, es decir, la clase de  $x$  será definida como  $\text{sgn}(\omega \cdot x_i + \Psi)$ . Esto implica que las muestras de test son etiquetadas con “+1” y las restantes con la etiqueta “-1”.

$$D^2(X, Y) = \frac{1}{2}(U_x - U_y)^T F^{-1}(U_x - U_y) \quad (2.47)$$

Donde  $F$  es la matriz de información y es equivalente a la matriz de covarianza de los vectores  $U_x$  y  $U_y$ .

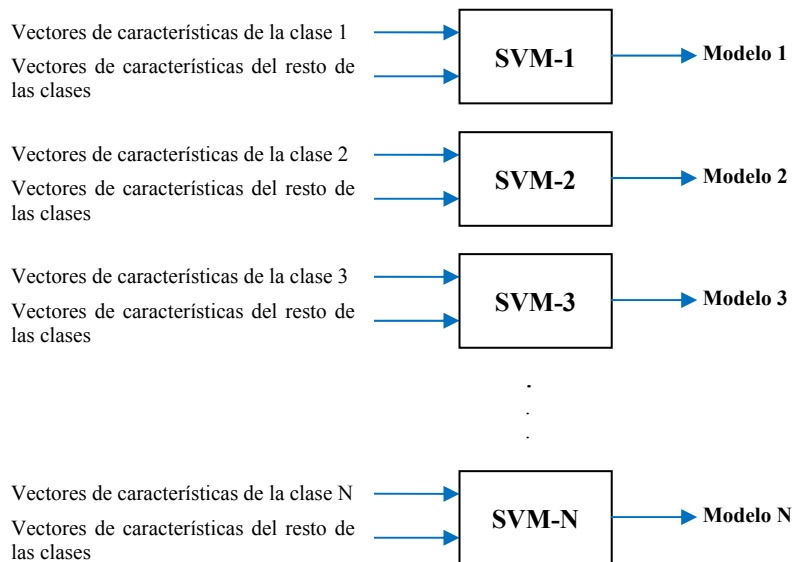
La gran ventaja que se consigue al utilizar las SVMs como clasificadores, es su criterio de elección del plano de separación, ya que calcula aquel plano que tiene el mayor margen posible con el conjunto de entrenamiento. Según la teoría de aprendizaje estadístico, la elección de un margen máximo de separación dará al sistema de clasificación una gran capacidad de generalización, como se ha demostrado en el estado del arte durante la última década.

Por definición, las SVM resuelven problemas biclase, sin embargo, existen un gran número de problemas que presentan más de dos clases. Es por ello, que se debe extender el sistema biclase de las SVM a un sistema multiclase.

Existen varias técnicas para realizar la clasificación multiclase haciendo uso de las SVM. Dichas técnicas son:

- Uno contra todos.
- Método de clasificación por parejas uno contra uno.
- Clasificación basada en un Gráfico Acíclico Dirigido.

Esta tesis se ha optado por el desarrollo de la técnica de uno contra todos [Shawe-Taylor y Cristianini, 2000] debido a los buenos resultados conseguidos en otros trabajos [Galar et al., 2011]. La técnica uno contra todos, se basa en la creación de  $M$  clasificadores binarios. Esto significa que cada clasificador se entrena para discriminar una clase de las  $M-1$  clases restantes. A una clase se le denomina la clase positiva y el resto serán clases negativas. En el caso en que se dé la clasificación de clase positiva en más de un clasificador, la clase asignada será aquella con la que se consiguió mayor margen. En la figura 2.7, se presenta el diagrama de bloques empleado en el modo de entrenamiento para el caso de un sistema SVMs multiclase, de  $N$  clases.



**Figura 2.7: Diagrama de bloques del clasificador SVM multiclase en modo entrenamiento para un sistema de  $N$  clases.**



## **CAPÍTULO 3**

# **Metodología Experimental**

---

En este capítulo se describe la metodología de experimentación utilizada para demostrar la hipótesis planteada en el anterior capítulo. En particular, se van a presentar las bases de datos y las variables del sistema propuesto para maximizar la eficiencia del sistema, referidas al preprocesado aplicado en general, la extracción de características, y su clasificación. Como base de datos, se tienen las colecciones de imágenes etiquetadas por el experto correspondiente, para su aplicación en una clasificación supervisada.

El esquema utilizado en este trabajo es el que se muestra en la figura 3.1. Está compuesto por 6 bloques:

- **Base de datos:** el primer bloque comprende un conjunto de bases de datos de distinta naturaleza que se utilizarán para comprobar la robustez de la transformación a espacios hiperdimensionales propuesto en esta tesis, para la correcta clasificación de formas.
- **Pre-procesado:** en el segundo bloque se realizan todas las operaciones necesarias sobre las imágenes de las bases de datos con el fin de eliminar el ruido y acondicionar las imágenes para los bloques posteriores.
- **Procesado:** en el tercer bloque se procesan las imágenes para localizar la forma que se va a analizar, y a continuación se extrae el contorno.
- **Parametrización:** en este bloque se extraen los parámetros característicos de los contornos obtenidos en el bloque anterior. En esta tesis, estos parámetros vienen definidos por su caracterización angular relativa, y distribuida dependiendo del tamaño del trazo del propio contorno.
- **Transformación de los datos:** en el siguiente bloque, se realiza la transformación de los parámetros extraídos del contorno a un espacio hiperdimensional. Se implementa y ejecuta el kernel desarrollado en esta tesis doctoral.
- **Clasificación:** Finalmente, se pasan los datos transformados al sistema de clasificación. Como se vio en el capítulo anterior, como sistemas de clasificación se utilizaron Máquinas de Soporte Vectorial (SVM), a las cuales se les proporcionó como datos de entrada, el resultado de la transformación basada en el kernel de Fisher. Además, la estrategia de modelado del sistema fue de aprendizaje supervisado. Este tipo de aprendizaje consiste en modelar una función, a partir de ejemplos etiquetados anteriormente, de manera que se establezca una correspondencia entre las entradas y las salidas deseadas del sistema. El sistema de aprendizaje trata de etiquetar automáticamente (clasificar) una serie de vectores, utilizando la correspondientes categorías (clases).

En la siguiente figura, se detallan cada uno de los bloques mencionados;

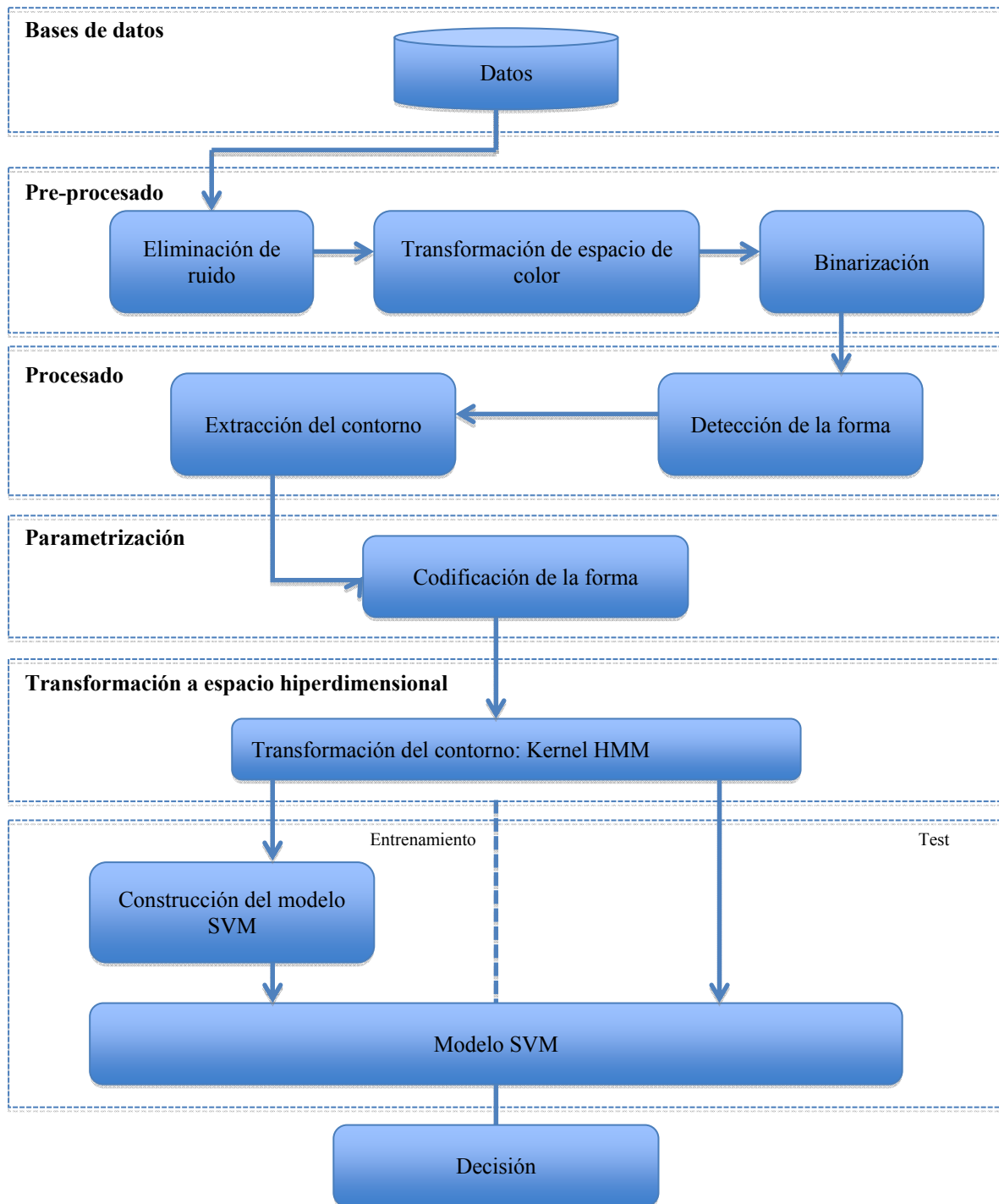


Figura 3.1: Esquema general de los bloques utilizados en este trabajo.

### **3.1 Bases de datos**

Las bases de datos representan un elemento crítico para el desarrollo y verificación de modelos y sistemas basados en técnicas biométricas. Es por esto, que las bases de datos utilizadas en este trabajo se dividen en dos categorías. La primera comprende un conjunto de imágenes generadas de forma sintéticas y que sirvió como punto de partida para realizar el modelado y comprobación del sistema propuesto. A esta base de datos se le denominó Base de datos Sintéticas.

La segunda categoría es un conjunto de bases de datos reales de distinta naturaleza, las cuales, según el grado de variabilidad de las formas de las imágenes y objetos, se categorizaron como: bases de datos con formas de variabilidad simple, bases de datos con formas de variabilidad moderada y bases de datos con formas de variabilidad complejas. La consideración de las bases de datos a cada grado de variabilidad, se analizó por las variaciones del contorno y simetrías. Las variabilidades simples no tienen cambios bruscos en la continuidad del contorno y se pueden encontrar múltiples simetrías. El grupo de bases de datos de variabilidad moderada, éstas presentan ciertos cambios bruscos y cierto grado de simetrías, sin ser totalmente simétricas. Finalmente, los contornos de variabilidad complejas, son asimétricas y presentan muchos cambios bruscos. Las bases de datos utilizadas según esta clasificación de complejidad de su contorno son:

- Base de datos sintética
- Variabilidad simple
  - Base de datos de Polen
  - Base de datos de Laurisilva de Canarias
  - Base de datos de Laurisilva de Costa Rica
- Variabilidad moderada
  - Base de datos de labios
  - Bases de datos de manos
- Variabilidad compleja
  - Base de datos de palabras manuscritas
  - Base de datos de firmas manuscritas
  - Base de datos de lenguaje dactilológico

### 3.1.1 Base de datos sintética

Se diseñó una base de datos sintéticas de 21 clases [Travieso et al., 2007], a partir de figura geométricas, tales como polígonos de varias caras, estrellas de varias puntas, etc. En particular, se han generado 100 figuras de cada clase, teniendo una base de datos global de 2100 imágenes. A estas figuras se les ha introducido un ruido aleatorio con una función de densidad uniforme, en la definición de su contorno, las transformaciones homomórficas, la variación de escala en sus dimensiones y además variaciones rotacionales. Finalmente, se han generado 100 figuras diferentes para cada una de las clases. La figura 3.2 muestra un ejemplo.

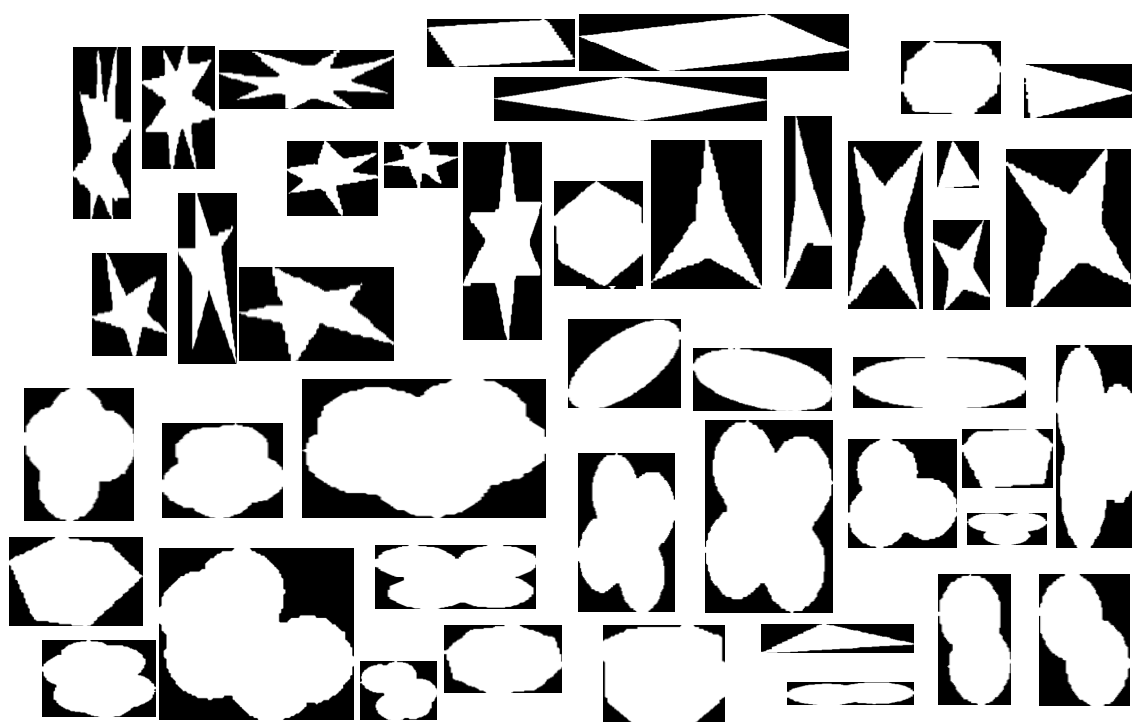


Figura 3.2: Algunas muestras de la base de datos sintética.

### 3.1.2 Granos de Polen: Variabilidad simple

La base de datos está compuesta por 564 imágenes de granos de polen cedidas por el Centro de Investigaciones Apícolas Tropicales (CINAT) [CINAT, 2012] de la Universidad Nacional de Costa Rica. Estas imágenes corresponden a 47 géneros y especies de 22 familias de plantas melíferas de Costa Rica. Para este trabajo se realizó una clasificación atendiendo al número de especies por lo cual el problema que se estudió cuenta con 47 clases. La tabla 3.1 y la tabla 3.2 enumeran la familia, género y

especie de los granos de polen estudiados. La figura 3.3 presenta algunos ejemplos de estas especies.

TABLA 3.1: FAMILIA, GÉNERO Y ESPECIE DE LA BASE DE DATOS DE GRANOS DE POLEN

Identificador	Familia	Género	Especie
1	<i>Acanthaceae</i>	<i>Bravaisia</i>	<i>intigerrima</i>
2	<i>Acanthaceae</i>	<i>Blechum</i>	<i>brownei</i>
3	<i>Anacardiaceae</i>	<i>Anacardium</i>	<i>excelsum</i>
4	<i>Anacardiaceae</i>	<i>Anacardium</i>	<i>occidentalis</i>
5	<i>Anacardiaceae</i>	<i>Spondias</i>	<i>purpurea</i>
6	<i>Asteraceae</i>	<i>Baltimora</i>	<i>recta</i>
7	<i>Asteraceae</i>	<i>Calea</i>	<i>urticifolia</i>
8	<i>Asteraceae</i>	<i>Critonia</i>	<i>morifolia</i>
9	<i>Asteraceae</i>	<i>Elephantopus</i>	<i>mollis</i>
10	<i>Asteraceae</i>	<i>Ageratum</i>	<i>purpureum</i>
11	<i>Asteraceae</i>	<i>Tridax</i>	<i>procumbens</i>
12	<i>Bignoneaceae</i>	<i>Tabebuia</i>	<i>rosea</i>
13	<i>Bignoneaceae</i>	<i>Tabebuia</i>	<i>impetiginosa</i>
14	<i>Bombacaceae</i>	<i>Bombacopsis</i>	<i>quinata</i>
15	<i>Boraginaceae</i>	<i>Cordia</i>	<i>alliodora</i>
16	<i>Boraginaceae</i>	<i>Heliotropium</i>	<i>indicum</i>
17	<i>Cochlospermaceae</i>	<i>Cochlospermum</i>	<i>vitifolium</i>
18	<i>Combretaceae</i>	<i>Combretum</i>	<i>fruticosum</i>
19	<i>Convolvulaceae</i>	<i>Ipomea</i>	<i>batatas</i>
20	<i>Euphorbiaceae</i>	<i>Croton</i>	<i>hirtus</i>
21	<i>Euphorbiaceae</i>	<i>Ricinus</i>	<i>comunis</i>
22	<i>Fabaceae</i>	<i>Aeschynomene</i>	<i>sensitiva</i>
23	<i>Fabaceae</i>	<i>Caesalpinia</i>	<i>pulcherrima</i>
24	<i>Fabaceae</i>	<i>Cassia</i>	<i>grandis</i>
25	<i>Fabaceae</i>	<i>Cassia</i>	<i>fistula</i>
26	<i>Fabaceae</i>	<i>Enterolobium</i>	<i>cyclocarpum</i>
27	<i>Fabaceae</i>	<i>Gliricidia</i>	<i>sepium</i>
28	<i>Fabaceae</i>	<i>Myrospermum</i>	<i>frutescens</i>
29	<i>Flacourtiaceae</i>	<i>Casearia</i>	<i>arguta</i>
30	<i>Lamiaceae</i>	<i>Hyptis</i>	<i>suavelons</i>
31	<i>Malpighiaceae</i>	<i>Banisteriopsis</i>	<i>muricata</i>
32	<i>Malpighiaceae</i>	<i>Bunchosia</i>	<i>cornifolia</i>
33	<i>Malpighiaceae</i>	<i>Byrsonima</i>	<i>crassifolia</i>
34	<i>Malvaceae</i>	<i>Herissanthia</i>	<i>crispa</i>
35	<i>Malvaceae</i>	<i>Malacra</i>	<i>Fasciata</i>
36	<i>Malvaceae</i>	<i>Sida</i>	<i>rhombifolia</i>
37	<i>Melastomataceae</i>	<i>Miconia</i>	<i>argentea</i>
38	<i>Myrsinaceae</i>	<i>Ardisia</i>	<i>revoluta</i>
39	<i>Polygonaceae</i>	<i>Antigonon</i>	<i>leptopus</i>
40	<i>Rubiaceae</i>	<i>Richardia</i>	<i>scabra</i>
41	<i>Sapindaceae</i>	<i>Cardiospermum</i>	<i>grandiflorum</i>
42	<i>Sapindaceae</i>	<i>Melicocca</i>	<i>bijuga</i>
43	<i>Sapotaceae</i>	<i>Sideroxylum</i>	<i>capiri</i>
44	<i>Sterculiaceae</i>	<i>Melochia</i>	<i>nodiflora</i>
45	<i>Sterculiaceae</i>	<i>Guazuma</i>	<i>umlifolia</i>
46	<i>Verbenaceae</i>	<i>Lantana</i>	<i>camara</i>
47	<i>Verbenaceae</i>	<i>Petrea</i>	<i>volubilis</i>

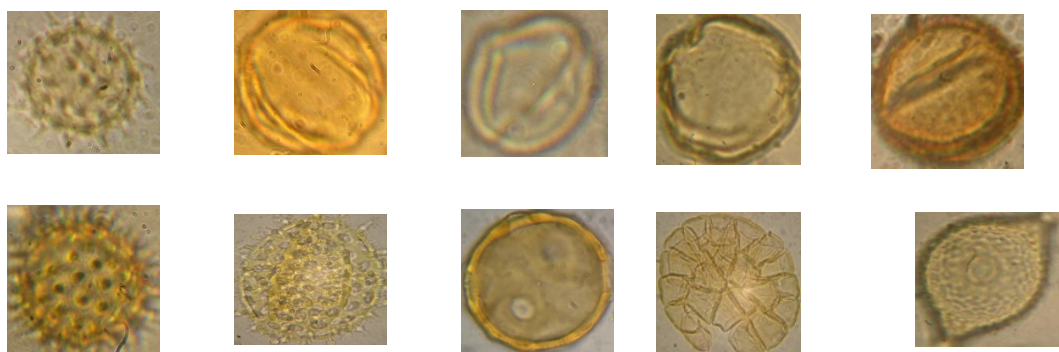


Figura 3.3: Ejemplos de la base de datos de polen.

### 3.1.3 Hoja de Árboles: Variabilidad simple

Se crearon dos bases de datos de hojas de árboles. La primera corresponde con La familia de Laurisilva de las islas Canarias y la segunda con la familia de Laurisilva Costaricense.

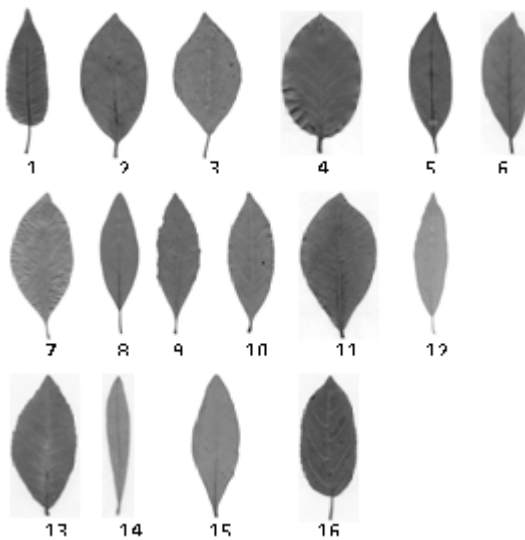
#### 3.1.3.1 Laurisilva Canaria

Para la construcción de la base de datos de Laurisilva Canaria se tomó en cuenta 16 de 20 variedades de hojas simples presentes en el archipiélago canario, con un total de 75 muestras por clase [Briceño et al., 2002]. Estas muestras han sido tomadas en las distintas islas, y se han escaneado en escalas de grises con una resolución de 300 dpi. Las variedades se enumeran en la tabla 3.2.

TABLA 3.2: VARIEDADES DE LA BASE DE DATOS DE HOJAS DE ÁRBOLES DE LAURISILVA CANARIA

Nº	Nombre de variedad	Nº	Nombre de variedad
1	Prunus lusitanica	9	Myrica fay
2	Ilex canariensis	10	Apollonias barbujana
3	Viburnum rigidum	11	Heberdenia excelsa
4	Ilex perado	12	Salis canariensis
5	Ocotea foetens	13	Visnea mocanera
6	Laurus azorica	14	Phillyrea angustifolia
7	Picconia excelsa	15	Mursyne canariensis
8	Persea indica	16	Rhamnus glandulosa

En la figura 3.4 se presentan un ejemplo de cada variedad.



**Figura 3.4:** Imágenes de las 16 variedades de *Laurisilva canariensis*.

### 3.1.3.2 Laurisilva de Costa Rica

La base de datos está compuesta por 7 géneros con diferentes especies (de uno a cinco especies), en total fueron 14 especies, y por cada especie se tomaron alrededor de 60 muestras [Briceño et al., 2007]. La base de datos de imágenes ha sido guardada en color y formato “jpeg” (*Joint Photographic Experts Group*) con compresión Huffman. Las imágenes fueron digitalizadas a 300 dpi. Los géneros considerados son: *Povedadaphne*, *Ocotea*, *Nectandra*, *Beilschmedia*, *Percea*, *Licaria* and *Cinnamomum*.



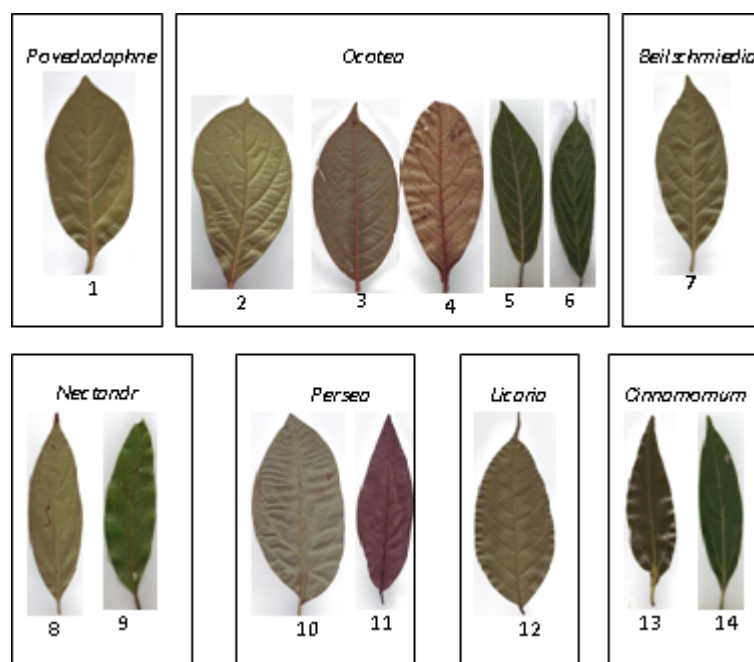


Figura 3.5: Géneros de Laurisilva de Costa Rica.

### 3.1.4 Labios: Variabilidad moderada

Esta base de datos está constituida por 50 usuarios y 10 muestras por usuario [Briceño et al., 2010]. Se utilizó una cámara compacta como medio de adquisición. Cada una de las imágenes se obtuvieron en tres sesiones diferentes para cada usuario. La tabla 3.4 muestra las características más importantes de esta base de datos.

TABLA 3.3: CARACTERÍSTICAS MÁS IMPORTANTES DE LA BASE DE DATOS CARAS PARA EL RECONOCIMIENTO DE LABIOS

Parámetros	Datos
Número de clases	50
Número de muestras por clase	10
Adquisición y Cuantificación	RGB (24 bits, 256 levels)
Tamaño	768×1024 píxeles


### 3.1.5 Manos: Variabilidad moderada

Se utilizaron tres bases de datos de manos de distintos usuarios para llevar a cabo esta experimentación: base de datos GPDS144 [Ferrer et. al, 2007], base de datos GPDS100hand3Band [Ferrer y Morales, 2011] y base de datos UST [Kumar et. al, 2003].

### 3.1.5.1 Base de datos GPDS144

Esta base de datos se ha formado con 144 usuarios, adquiriendo 10 muestras por clase en dos sesiones, haciendo un total de 1440 muestras. Se ha utilizado un escáner HP para adquirir cada muestra de la clase, lo cual minimiza los efectos sobre el entorno [Ferrer et. al, 2007]. La tabla siguiente muestra las características más importantes:

TABLA 3.4: CARACTERÍSTICAS MÁS IMPORTANTES DE LA BASE DE DATOS DE MANOS GPDS144


Parámetros	Datos
Número de clases	144
Número de muestras por clase	10
Adquisición y Cuantificación	Escala de grises (8 bits, 256 niveles)
Resolución	150 dpi
Tamaño	1403×1021 píxeles
Ejemplo	

### 3.1.5.2 Base de datos GPDS100hand3Band

Se encuentra formada por imágenes de tres bandas del espectro electromagnético: banda del rango visible y dos de las bandas del infrarrojo cercano, 850nm y 1470nm. Para este trabajo se utilizaron las imágenes del infrarrojo cercano adquiridas en la banda de 1470 nm.

La base de datos usada está compuesta por un total de 1000 imágenes de la mano derecha de 100 personas, adquiridas en 10 sesiones distintas. La mayoría de los usuarios tenían una edad entre los 23 y 40 años. Aproximadamente el 50 por ciento de los voluntarios eran hombres. Las imágenes en la banda de 1470 nm fueron adquiridas por una cámara XENICS modelo XEVA 1.7-320 con un sensor InGaAs, sensible en el rango 900- 1700 nm, con un lente filtro paso banda centrado en 1470 nm y un ancho de banda de 250 nm. Se utilizó una bombilla incandescente como medio de iluminación, teniendo ésta, un patrón de radiación de 400 nm a 2500 nm [Ferrer y Morales, 2011]. La tabla 3.6 muestra los detalles de esta base de datos.


TABLA 3.5: CARACTERÍSTICAS MÁS IMPORTANTES DE LA BASE DE DATOS DE MANOS GPDS100HAND3BAND

Parámetros	Datos
Número de clases	100
Número de muestras por clase	10 muestras de la mano derecha
Adquisición y Cuantificación	Escala de grises (8 bits, 256 niveles)
Resolución	150 dpi
Tamaño	320×256 píxeles
Ejemplo	

### 3.1.5.3 Base de datos UST

La última base de datos utilizada corresponde con una base de datos pública creada por la Universidad de Ciencia y Tecnología de Hong Kong. Está formada por 10 imágenes de la mano derecha y 10 imágenes de la mano izquierda de 287 personas. Las manos fueron adquiridas en un escenario sin contacto, utilizando una cámara digital Olympus C-3020 con una resolución de  $1280 \times 960$  píxeles. En esta base de datos no se utilizaron topos ni iluminación especial. Las características más importantes se muestran en la tabla 3.6:

TABLA 3.6: CARACTERÍSTICAS MÁS IMPORTANTES DE LA BASE DE DATOS DE MANOS UST

Parámetros	Datos
Número de clases	287
Número de muestras por clase	10 muestras de la manos izquierda y 10 muestras de la mano derecha
Adquisición y Cuantificación	Escala de grises (8 bits, 256 niveles)
Resolución	500 dpi
Tamaño	$1280 \times 960$ píxeles
Ejemplo	

### 3.1.6 Palabras: Variabilidad compleja

La base de datos de palabras está formada por un total de 7680 imágenes de palabras manuscritas. Este conjunto de imágenes presenta la siguiente estructura jerarquizada:

- 64 clases, cuyo contenido se muestra en la tabla 3.7.

TABLA 3.7: DESCRIPCIÓN DEL CONTENIDO DE LA BASE DE DATOS DE PALABRAS

CLASES (64)	DESCRIPCIÓN
1-30	Cifras manuscritas del 1 al 30
30-36	Cifras manuscritas correspondientes a las decenas de la “cuarenta” a la “noventa”
37-47	Cifras de las centenas, en femenino, de la “cien” a la “novecientas”
48	Cifra “mil”
49	Palabra manuscrita “millón”
50	Palabra manuscrita “un”
51-59	Cifras de las centenas, en masculino, desde “ciento” a “novecientos”
60	Palabra manuscrita “millones”
61	Palabra manuscrita “una”
62	Palabra manuscrita “y”
63	Palabra manuscrita “pesetas”
64	Símbolo del dólar

- Cada clase contiene 120 muestras.
- Estas 120 muestras se dividen en 40 grupos.
- En cada uno de estos 40 grupos se tienen imágenes de 3 escritores distintos.
- Todas las imágenes tienen una ocupación de memoria de 48KB.
- El tamaño de las imágenes en píxeles es de 109x438.
- Los píxeles presentan valores de 0 a 255 (blanco o negro). Por lo tanto, se utilizan 8 bits de codificación.

Finalmente, en cuanto a este apartado se refiere, se muestran en la figura 3.6 cuatro imágenes de muestra pertenecientes a diferentes clases de la base de datos.

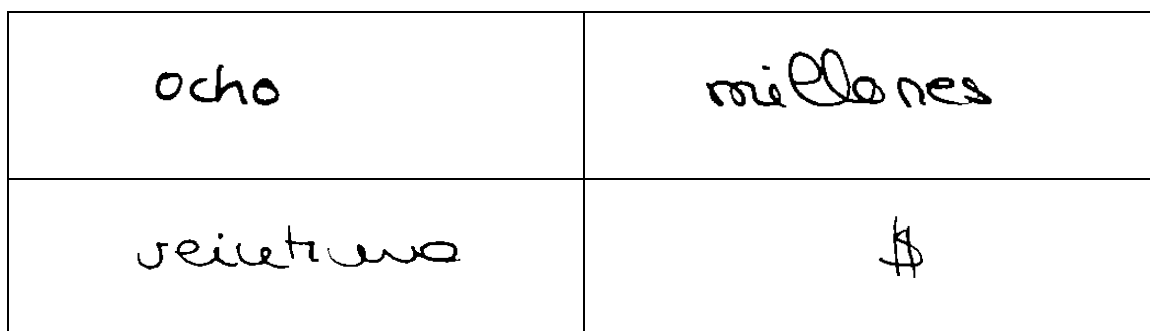


Figura 3.6 Imágenes de la base de datos de palabras.

### 3.1.7 Firmas: Variabilidad compleja

Para esta aplicación se usaron dos bases de datos públicas de firmas manuscritas *off-line*, la MYCT-75 [Fierrez-Aguilar et al., 2004] y la GPDS-800 [Briceño et al., 2009], usándose la versión de hasta 100 muestras en un primer experimento y el total de la base de datos para otro experimento.

La base de datos MYCT-75 es parte de una base de datos pública multimodal, creada por el Grupo de Reconocimiento Biométrico - ATVS de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid, y que se encuentra disponible en <http://atvs.ii.uam.es/databases.jsp>. El corpus MYCT-75 contiene 75 firmantes. Se disponen de 15 muestras genuinas y 15 falsificaciones para cada firmante. Las imágenes tienen una resolución de 600ppp. Es importante mencionar que todas las firmas con el mismo útil escritural y esta base de datos ha sido una base de datos de referencia de múltiples trabajos científicos como [Fierrez-Aguilar et al., 2004], o más recientemente [Vargas, 2010].

La base de datos pública GPDS-800 [Briceño et al., 2009] contiene 24 firmas auténticas de 800 individuos, por tanto, se tienen 18400 firmas auténticas. Para la adquisición de esta base de datos, se usó un escáner a una resolución de 300 dpi con una cuatificación de 8 bits en escala de grises. La construcción de esta base de datos se realizó en una sesión. Los usuarios firmaron una plantilla con 24 celdas de tamaños distintos. Todo el proceso se realizó bajo supervisión. La tabla 3.8 resume los valores característicos de esta base de datos.

TABLA 3.8: DISTRIBUCIÓN DEL CORPUS DE GPDS-800

	Genero (%)		Edad Media	Diestro	Izquierdo
	Masculino	Femenino			
Genuinas	50.25	49.75	24	91.5	8.5
Falsificaciones	54.63	45.37	25	98	2

Esta base de datos ha sido usada en múltiples trabajos, como se puede recoger en [Vargas, 2010].

### 3.1.8 Dactilológico: Variabilidad compleja

La base de datos dactilológica [Travieso et al., 2003] fue creada para este trabajo y está compuesta por 25 signos o patrones usando los dedos de la mano, recolectando 50 muestras de cada uno de los signos, todos de personas distintas. Estas imágenes fueron capturadas utilizando una cámara fotográfica digital comercial.

Se ha trabajado con imágenes estáticas, sin movimiento, dado que el alfabeto de signos tiene algunos signos que llevan a cabo movimientos tales como la CH, H, Ñ, V, X, Y y Z. Se eliminan entonces los movimientos en la combinación CH, y las letras H, X e Y. De la misma manera se eliminan algunos signos de manos, porque hay varios que son muy similares sin movimientos. En este caso se eliminan la Z, Ñ y la V.

Algunos datos estadísticos de esta base de datos, se enumeran a continuación;

- Un 74% de hombres y un 26% de mujeres.
- El 30% de las personas eran sordas.
- La distribución de las edades varía entre 20 a los 60 años.
- Las imágenes tienen un tamaño de 312×232 píxeles, con una resolución de 24 bits a color, en el modelo de color RGB, es decir, 8 bits por cada uno de los canales.

Un ejemplo de cada uno de los signos se presenta en la figura 3.7



**Figura 3.7: Signos que conforman la base de datos dactilológica.**

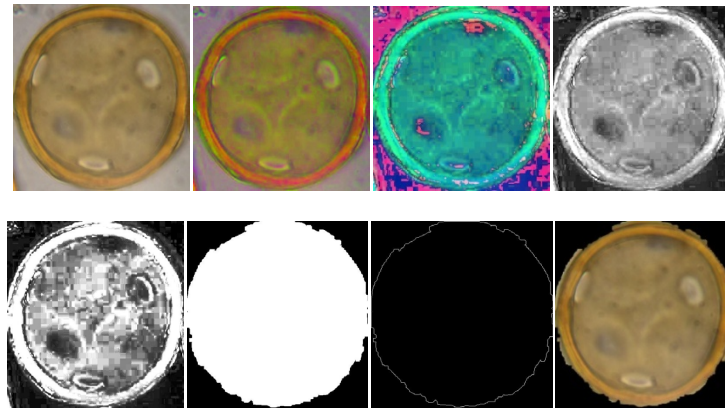
### **3.2 Pre-Procesado**

El acondicionamiento de las imágenes de las bases de datos representa un aspecto fundamental de cara a la extracción de la información de interés. Las operaciones de pre-procesado difieren según la base de datos, atendiendo a las particularidades de cada una. Como primer objetivo, se focalizó en minimizar el ruido provocado por las condiciones del entorno en el momento de la adquirir la imagen, introducido por el propio sistema de adquisición. Debido a que las bases de datos de granos de polen, de labios, de hojas de árboles de Laurisilva Costaricense y de los signos dactilológicos fueron guardadas en formato RGB, fue necesario realizar una transformación del modelo de color, a escala de grises en todas ellas. En el caso particular de la base de datos de los granos de polen, se aplicaron dos transformaciones específicas adicionales. La primera fue la decorrelación del color entre los canales RGB, y la segunda, una transformación al espacio de color HSV [Travieso et al., 2011]. Una vez conseguido el acondicionamiento máximo posible, las imágenes se binarizaron utilizando el método de umbralización de Otsu [Otsu, 1979].

### **3.3 Procesado y parametrización, transformación a espacios hiperdimensionales y clasificación**

Con las imágenes binarizadas, el bloque de procesado se encarga de eliminar cualquier resto de ruido mediante operaciones morfológicas matemáticas. En este punto, independientemente de la base de datos, la extracción del elemento de la imagen que se desea analizar se hace mediante una operación muy simple de aplicar en la mayoría de las bases de datos. En el caso particular de las bases de datos de palabras y firmas, se realiza un paso intermedio, que consiste en la unión de los caracteres que conforman una palabra o firma ya que el objeto de estudio es el contorno del conjunto.

Finalmente, el contorno se obtiene aplicando las operaciones morfológicas de dilatación y erosión y restando el resultado a la imagen original. La figura 3.8 muestra un ejemplo para el caso de la extracción de granos de polen.



**Figura 3.8:** Pre-procesado aplicado a la base de datos de granos de polen.

Las estrategias y métodos utilizados en los bloques de parametrización y clasificación son los explicados en el capítulo 2.

### 3.4 Metodología de experimentación

La metodología que se ha planteado, sigue una estructura clásica de sistemas de reconocimiento de patrones configurados en modo identificación con la finalidad de evitar multiplicar los experimentos con diversos modos de evaluación.

La tarea de todo sistema configurado en modo identificación, es la de clasificar correctamente una muestra dada la clase a la que pertenece. La metodología general que se aplicó a todas las bases de datos es la siguiente: primero, todas las imágenes fueron tratadas por el bloque de pre-procesado, en donde se aplicaron todas las operaciones necesarias para eliminar el ruido, ajustar el contraste, ajustar la luminosidad, equalizar la distribución de los píxeles de las imágenes según las características particulares, etc., aplicando algunas más funciones específicas, dependiendo de la base de datos. Una vez las imágenes fueron acondicionadas, se aplicaron distintas técnicas de procesamiento de imágenes para extraer el contorno, fundamentalmente fueron basadas en morfología matemática. Dado que la naturaleza de las bases de datos era variable, las técnicas implementadas también eran específicas según las particularidades de las bases de datos, sin embargo, el objetivo que se perseguía y que se alcanzó en todas las bases de datos, fue la obtención del contorno del objeto analizado, a partir de imágenes binarizadas. El siguiente paso implementado fue la parametrización de los contornos extraídos, mediante el método propuesto en la sección 2.2.



Llegados a este punto, las siguientes fases de la experimentación estuvieron basadas en la aplicación del modo identificación, basado en una clasificación supervisada. Se realizaron aplicando dos enfoques: primero se diseñaron sistemas en donde se utilizó un clasificador basado en HMM y utilizando el contorno como parámetros de entrada. Se implementó este enfoque, para tener una referencia común de comparación de los resultados obtenidos usando las distintas bases de datos y mostrar el grado discriminador del uso de los contornos de forma directa. Como se pudo comprobar, los resultados obtenidos en la implementación de este enfoque fueron bajos, debido a que solo la información de contorno es insuficiente para obtener una correcta clasificación. Se puede encontrar en el estado del arte trabajos en esta línea, donde solo la información de contorno frente a otras fuentes de información no es una buena opción [Wang and Liew, 2012].

Por esta razón, se implementó el segundo enfoque, siendo además la justificación práctica de esta tesis doctoral. Así, en el segundo enfoque se realizó la transformación de los parámetros de contorno usando el kernel desarrollado en esta tesis mediante HMMs (véase el apartado 2.3.6). El resultado dado por dicha transformación a espacios hiperdimensionales, se utilizó como datos de entrada a un sistema de clasificación basado en SVMs. Como se explicó en el apartado 2.3.7, el uso de SVMs queda avalado por los buenos resultados conseguidos en otras investigaciones [Cherkassky y Mulier, 2007], [Xiaoou et al., 2010], [Cervantes et al., 2007a], [Thanh-Nghi y Fekete, 2007], [Cervantes et al., 2007b]; en donde se trabajó con un número masivo de datos. En este caso, el número de datos de entrada a las SVMs se varió durante la experimentación, hasta obtener el número óptimo. Los parámetros del kernel y de regularización se han optimizado automáticamente por medio de iteraciones utilizando la metodología de validación *Hold-out* [Arlot, 2010], con diversos porcentajes, fundamentalmente valores inferiores al 50%.

La técnica de modelado de los sistemas comprende dos fases: entrenamiento y test. Por lo tanto, las bases de datos fueron divididas en dos grupos, los cuales fueron utilizados en cada una de las fases de modelado. Para obtener resultados fiables, en la división de los grupos se utilizó la técnica de validación cruzada *Hold-out* [Arlot, 2010], de forma que los sistemas fueron entrenados y testeados con muestras totalmente diferentes. Además, para estudiar la robustez de los sistemas, los porcentajes de entranamiento y test fueron variados del 50 por ciento hasta el 10 por ciento. Vale la

pena mencionar que los conjuntos de entrenamiento y de test de cada base de datos y de cada clase, fueron calculados de forma independiente, tomando en cuenta que cada base de datos y/o cada clase tienen un número distinto de muestras.

Finalmente, se realizaron experimentos singulares con el fin de probar la estabilidad, invariabilidad y robustez de los sistemas en escenarios más complejos. Estos experimentos comprenden el estudio de la robustez del sistema frente al ruido, comportamiento del sistema frente a la resolución de las imágenes, y la generalización del sistema para utilizarlo en modo verificación para algunas aplicaciones en concreto.

A continuación, se describen las fases de modelado de los sistemas así como la técnica de validación cruzada implementada.

### **3.4.1 Fase de entrenamiento**

En esta fase se determinan los valores de las variables de optimización necesarios para caracterizar cada una de las clases de forma que se desean clasificar. Estas variables de optimización son:

#### **Número de puntos de control:**

Con el fin de estandarizar el tamaño del vector de descripción se implementa una interpolación de los puntos de control del contorno. El criterio de selección de puntos se basa en conseguir un error de interpolación de la forma mínima y de inversión computacional aceptable.

#### **Número de estados del HMM**

El número de estado del HMM representa un parámetro fundamental a determinar durante la fase de modelado del sistema ya que define el grado de fiabilidad, robustez y estabilidad del sistema. Para determinar este número se aplicaron las técnicas explicadas en detalle en el capítulo dos.

#### **Varianza del kernel Gaussiano (gamma).**

Como se expuso en el capítulo 2, la versión utilizada en este trabajo es la SVM<sup>light</sup> con función kernel de base radial (RBF), definida por [Steinwart and Christmann, 2008]:

$$K(x, y) = e^{-\gamma\|x-y\|^2} = e^{-\frac{\|x-y\|^2}{2\sigma^2}} \quad (3.1)$$

Siendo  $\sigma^2$  la varianza del Kernel Gaussiano. Con esta configuración el parámetro clave a determinar es la varianza de la gaussiana (gamma ( $\gamma$ )), la cual es inversamente proporcional a la varianza del kernel Gaussiano y corresponde al ancho del kernel RBF.

### **Coste de entrenamiento (c).**

Esta constante permite regular la solución de compromiso entre complejidad y coste computacional. La determinación de este valor en una aplicación real, es en ocasiones más compleja, que la elección del kernel a emplear, ya que éste puede estar determinado por la naturaleza de los datos. Se utilizó un valor de 10 en todos los experimentos.

### **3.4.2 Fase de test**

Después de este proceso de entrenamiento, el clasificador irá tomando, elemento a elemento, los datos que se introducen, procedentes de la caracterización. Sólo se testearán aquellas muestras que no hayan sido utilizadas en el entrenamiento. Utilizando diferentes operaciones y criterios de decisión, se realizará una comparación con los rasgos distintivos de cada clase que se han establecido en el entrenamiento. Finalmente decidirá a que clase pertenece la muestra.

La fase de test es el proceso final que mide la capacidad del sistema de caracterización. Lo que se hace en esta fase es anotar, una a una, las predicciones realizadas por el clasificador, para cada una de las muestras evaluadas. Una vez anotados todos los resultados, se realiza un procesado que permite conocer los porcentajes de acierto y de error para cada clase.

### **3.4.3 Validación cruzada:**

Los resultados serán presentados con la media y desviación estándar de las tasas de acierto de los distintos números de experimentos realizados. Para que estos resultados sean estadísticamente independientes de las muestras de los conjuntos, se aplicó el

método de validación cruzada *Hold-out* [Arlot, 2010]. Con este método el conjunto de datos es dividido en dos subconjuntos. La ventaja de este método es que es muy rápido en el momento del cómputo, y por lo tanto, es viable la realización de múltiples experimentos. La desventaja que presenta, es que al realizar una sola iteración, la evaluación puede depender de las muestras elegidas en un determinado caso; y el conjunto de test, puede ser muy fáciles o muy difíciles de clasificación. Este problema se ha resuelto realizando un número significativo de iteraciones para que los resultados obtenidos sean estadísticamente independientes de las muestras de los subconjuntos [Nieves y Dominguez, 2010].

## **CAPÍTULO 4**

# **Experimentos y resultados**

---

En este capítulo se describen los experimentos realizados con las distintas bases de datos y se presentan los resultados obtenidos. El grado de complejidad de los problemas abordados siguió una secuencia progresiva. Como punto de partida para la implementación de la estrategia propuesta, se utilizó la base de datos sintética. A continuación, se utilizaron las bases de datos reales, empezando con el uso de las bases de datos de variabilidad simple, continuando con la bases de datos de variabilidad moderada y terminando con la utilización de las bases de datos de variabilidad compleja. El ajuste de los parámetros se hizo con un barrido muy amplio, pero se muestran los resultados del ajuste fino sobre los resultados más significativos, para el modelado que finalmente ofrecía mejor tasa de éxito.

#### 4.1 Utilización base de datos sintética.

Con el fin obtener una primera aproximación de la efectividad y robustez de la transformación de los datos para su posterior clasificación, en el primer estudio, se utilizó la base de datos sintética. Para tener un punto de referencia con el cual hacer una comparación de los resultados finales; como primer paso, se clasificaron los datos utilizando un clasificador HMM. Para este experimento, el porcentaje de muestras usadas para entrenar, se varió desde el 15% hasta el 30% de la base de datos creada y para cada experimento. Se escogieron aleatoriamente las muestras para entrenar con una distribución de densidad uniforme. El resto de las muestras, del 85% hasta el 70%, se utilizaron en la fase de test. La tabla 4.1 muestra los resultados de este experimento, donde se puede observar que en ningún caso, se superó el 30% de tasa de acierto.

TABLA 4.1: RESULTADOS CON HMM USANDO 40, 50 Y 60 ESTADOS

% de entrenamiento	HMM con 40 estados	HMM con 50 estados	HMM con 60 estados
15	24.30% ± 1.33	25.59% ± 2.29	27.06% ± 1.97
20	24.00% ± 2.14	28.37% ± 2.36	27.16% ± 1.99
25	24.48% ± 1.12	27.31% ± 1.66	27.30% ± 2.05
30	28.08% ± 2.42	26.90% ± 2.05	26.53% ± 1.96

Una vez obtenidos estos resultados, se procedió a realizar la transformación de los datos mediante el kernel propuesto a un espacio hiperdimensional. Como se describió en el capítulo 3, en esta fase se utilizaron SVMs como sistemas de clasificación. Para realizar una comparación a este nivel, se utilizó el kernel lineal y el kernel RBF.

En el segundo experimento, la división de las muestras partió del 5% de las muestras para entrenar y 95% para el test, hasta el 30% de las muestras para entrenar y el resto para el test. En la tabla 4.2, se muestran los resultados, donde se observa que se alcanzó una tasa de acierto del 99,71% de media.

TABLA 4.2: RESULTADOS DEL CLASIFICADOR SVM CON LOS DATOS TRANSFORMADOS CON EL KERNEL HMM CON 50 ESTADOS

% entrenamiento	Kernel Lineal	Kernel RBF	Valor de g (RBF)
3	50.08% ± 41.78	50.14% ± 40.81	$3 \times 10^{-7}$
4	89.95% ± 2.83	87.23% ± 4.09	$2 \times 10^{-7}$
7	94.77% ± 1.39	92.31% ± 0.68	$2 \times 10^{-7}$
11	96.77% ± 1.62	95.25% ± 1.62	$3 \times 10^{-7}$
14	98.23% ± 0.34	97.18% ± 0.29	$2 \times 10^{-7}$
18	98.89% ± 0.46	98.38% ± 0.78	$2 \times 10^{-7}$
21	99.48% ± 0.13	99.05% ± 0.28	$2 \times 10^{-7}$
25	99.58% ± 0.27	99.39% ± 0.39	$1 \times 10^{-7}$
30	99.71% ± 0.14	99.59% ± 0.17	$3 \times 10^{-7}$

## 4.2 Variación Simple

### Base de datos de granos de polen

Al igual que en el experimento anterior, se aplicaron dos enfoques utilizando el 50% de las muestras en la fase de entrenamiento. El número de estados del HMM en el primer enfoque se varió entre 30 y 70 estados. En el primer enfoque, la tasa de acierto alcanzada fue menor al 19%, para los distintos números de estados utilizados. En la tabla 4.3 se muestran los mejores resultados de este experimento.

En el segundo enfoque, se mejora los primeros resultando introduciendo la transformación mediante el kernel propuesto utilizando el HMM y SVMs. En la tabla 4.4, se muestran las mejores tasas de acierto conseguidas, siendo el número de estados óptimo en este caso de 50, donde gamma (g) es el parámetro a ajustar en el kernel RBF. Finalmente, la tabla 4.5 muestra las mejores tasas de acierto utilizando 300 puntos de control del contorno frente a diferentes números de estados, en particular, entre 30 y 70 estados.

TABLA 4.3: TASA DE ÉXITO USANDO EL CLASIFICADOR HMM

Puntos de contorno/Números de estados	30	50	70
150	8.87% ± 0.35	10.52% ± 1.14	7.92% ± 1.43
200	12.17% ± 1.14	9.10% ± 1.68	10.76% ± 1.43
250	13.71% ± 0.54	13.23% ± 0.89	10.87% ± 1.75
300	13.83% ± 1.84	14.30% ± 2.17	14.18% ± 1.77
350	14.78% ± 0.58	15.48% ± 1.34	14.30% ± 0.54
400	19.15% ± 0.71	17.97% ± 3.37	17.85% ± 0.54

TABLA 4.4: TASA DE ACIERTO DE LOS DATOS TRANSFORMADOS UTILIZANDO SVM COMO CLASIFICADOR

Número de estados-Puntos de control	Número de estados		
	Kernel Lineal	Kernel RBF	g
50-150	89.01% ± 2.81	88.30% ± 3.99	$1 \times 10^{-6}$
50-200	89.72% ± 0.95	89.72% ± 0.95	$1 \times 10^{-6}$
50-250	91.72% ± 1.95	90.66% ± 1.08	$4 \times 10^{-7}$
50-300	93.85% ± 1.43	93.38% ± 1.47	$4 \times 10^{-7}$
50-350	93.38% ± 2.14	93.02% ± 2.14	$4 \times 10^{-7}$
50-400	90.43% ± 3.15	89.48% ± 3.20	$6 \times 10^{-7}$

En la tabla 4.4 se muestran los resultados obtenidos variando el número de puntos de control con un número fijo de estados, en particular 50. El mejor resultado se consigue con 300 puntos, alcanzando un 93.85% de tasa de acierto.

TABLA 4.5: TASA DE ACIERTO UTILIZANDO EL KERNEL HMM Y SVM VARIANDO EL NÚMERO DE ESTADOS

Número de estados-Puntos de control	Número de estados		
	Kernel Lineal	Kernel RBF	g
30-300	91.13% ± 2.21	89.13% ± 3.55	$4 \times 10^{-6}$
40-300	93.03% ± 1.34	93.29% ± 0.84	$4 \times 10^{-6}$
50-300	93.85% ± 1.43	93.38% ± 1.47	$4 \times 10^{-7}$
60-300	92.79% ± 0.94	92.43% ± 1.95	$6 \times 10^{-7}$
70-300	90.43% ± 3.15	89.47% ± 3.20	$2 \times 10^{-7}$

En la tabla 4.5, se ha finalizado los experimentos variando el número de estados. El mejor resultado se consiguió con 50 estados y 300 puntos.

### Laurisilva Costa Rica.

Se utilizó un porcentaje de entrenamiento entre 5% y 30%. También se aplicó el kernel lineal y RBF en las SVMs. Cada experimento se repitió cinco veces, para conseguir independencia estadística de los resultados. En las siguientes tablas, se muestran los resultados obtenidos con 40, 50 y 60 estados en el HMM. Se observa que el mejor resultado, se alcanza utilizando 50 o 60 estados, y que se necesitan pocas muestras de entrenamiento. Entrenando solamente con 4 hojas (5% de la base de datos), se obtiene una tasa de acierto del 100%. Se observa, usando 50 estados en el HMM, es un buen valor para empezar las operaciones del sistema. También se desprende, que el resultado con el kernel RBF muestra una ligera mejora frente al kernel lineal. Además del RBF, el



valor óptimo de la variable “g” está localizada siempre alrededor  $5 \times 10^{-7}$ , lo cual muestra estabilidad frente a las demás variables de optimización.

TABLA 4.6: RESULTADOS USANDO 40 ESTADOS DEFINIDOS EN HMM

% de entrenamiento	Tasa de éxito (Lineal)	Tasa de éxito (RBF)	Valor de g (RBF)
1 ( 1 hoja)	99.17% ± 1.44	99.17% ± 1.44	$5 \times 10^{-7}$
2 ( 2 hojas)	99.67% ± 0.74	99.67% ± 0.74	$5 \times 10^{-7}$
4 ( 3 hojas)	99.33% ± 1.09	99.67% ± 0.74	$2 \times 10^{-8}$
5 ( 4 hojas)	99.83% ± 0.37	99.83% ± 0.37	$5 \times 10^{-7}$
10 (7 hojas)	99.83% ± 0.37	99.83% ± 0.37	$5 \times 10^{-7}$
15 ( 11 hojas)	99.83% ± 0.37	99.83% ± 0.37	$5 \times 10^{-7}$
20 ( 14 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
25 ( 18 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
30 ( 21 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$

TABLA 4.7: RESULTADOS USANDO 50 ESTADOS DEFINIDOS EN HMM

% de entrenamiento	Tasa de éxito (Lineal)	Tasa de éxito (RBF)	Valor de g (RBF)
1 ( 1 hoja)	99.83% ± 0.37	99.83% ± 0.37	$7 \times 10^{-7}$
2 ( 2 hojas)	99.00% ± 0.69	99.00% ± 0.69	$5 \times 10^{-7}$
4 ( 3 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
5 ( 4 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
10 (7 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
15 ( 11 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
20 ( 14 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
25 ( 18 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
30 ( 21 hojas)	99.83% ± 0.37	100% ± 0	$5 \times 10^{-6}$

TABLA 4.8: RESULTADOS USANDO 60 ESTADOS EN EL HMM

% de entrenamiento	Tasa de éxito (Lineal)	Tasa de éxito (RBF)	Valor de g (RBF)
1 ( 1 hoja)	99.83% ± 0.37	99.83% ± 0.37	$5 \times 10^{-7}$
2 ( 2 hojas)	99.83% ± 0.37	100% ± 0	$1 \times 10^{-6}$
4 ( 3 hojas)	98.17% ± 2.79	98.67% ± 2.17	$6 \times 10^{-6}$
5 ( 4 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
10 (7 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
15 ( 11 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
20 ( 14 hojas)	99.83% ± 0.37	100% ± 0	$5 \times 10^{-7}$
25 ( 18 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$
30 ( 21 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$

### Base de datos Laurisilva Canaria

El proceso seguido fue el mismo que para la base de datos de Laurisilva de Costa Rica. Por lo tanto, el número de estados se varió entre 40 y 60. Comparando los resultados anteriores con el 5% de las muestras de entrenamiento (4 hojas), se obtiene un 99.98% para el mejor de los casos con 60 estados, y un 99.90% con 50 estados. Estos valores

son similares o comparables al obtenido con la Laurisilva de Costa Rica, y con similar kernel (RBF) son similar valor o igual valor de g.

TABLA 4.9: RESULTADOS CON SVM USANDO UN HMM CON 40 ESTADOS

% de entrenamiento	Tasa de éxito (Lineal)	Tasa de éxito (RBF)	Valor de g (RBF)
5 (4 hojas)	99.81 % ± 0.20	99.86% ± 0.15	$8 \times 10^{-7}$
10 (8 hojas)	99.94% ± 0.08	99.96% ± 0.08	$7 \times 10^{-7}$
15 (12 hojas)	99.94% ± 0.09	99.92% ± 0.08	$7 \times 10^{-7}$
20 (15 hojas)	99.96% ± 0.09	99.96% ± 0.09	$7 \times 10^{-7}$
25 (19 hojas)	100% ± 0	100% ± 0	$8 \times 10^{-7}$
30 (23 hojas)	100% ± 0	100% ± 0	$7 \times 10^{-7}$

TABLA 4.10: RESULTADOS CON SVM USANDO UN HMM CON 50 ESTADOS

% de entrenamiento	Tasa de éxito (Lineal)	Tasa de éxito (RBF)	Valor de g (RBF)
5 (4 hojas)	99.93% ± 0.11	99.93% ± 0.11	$5 \times 10^{-7}$
10 (8 hojas)	99.90% ± 0.12	99.90% ± 0.12	$7 \times 10^{-7}$
15 (12 hojas)	100% ± 0	100% ± 0	$7 \times 10^{-7}$
20 (15 hojas)	99.96% ± 0.09	99.98% ± 0.05	$5 \times 10^{-7}$
25 (19 hojas)	100% ± 0	100% ± 0	$7 \times 10^{-7}$
30 (23 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$

TABLA 4.11: RESULTADOS CON SVM USANDO UN HMM CON 60 ESTADOS

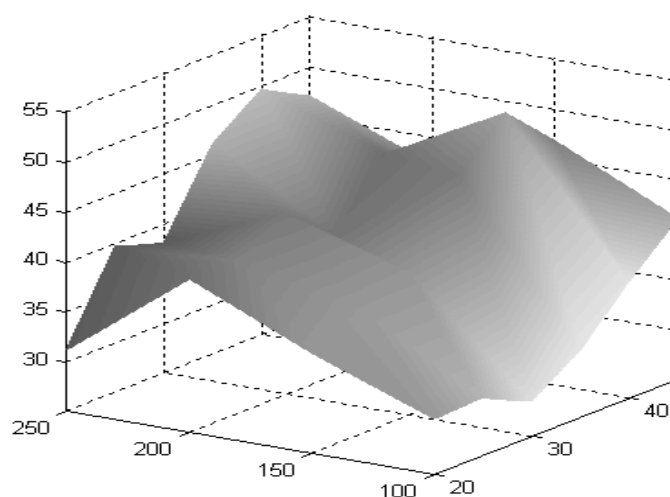
% de entrenamiento	Tasa de éxito (Lineal)	Tasa de éxito (RBF)	Valor de g (RBF)
5 (4 hojas)	99.98% ± 0.01	99.98% ± 0.01	$5 \times 10^{-7}$
10 (8 hojas)	99.96% ± 0.05	99.98% ± 0.04	$2 \times 10^{-7}$
15 (12 hojas)	99.98% ± 0.04	100% ± 0	$7 \times 10^{-7}$
20 (15 hojas)	100% ± 0	100% ± 0	$7 \times 10^{-7}$
25 (19 hojas)	99.98% ± 0.05	100% ± 0	$7 \times 10^{-7}$
30 (23 hojas)	100% ± 0	100% ± 0	$5 \times 10^{-7}$

### 4.3 Variaciones moderadas

#### 4.3.1 Base de datos de labios

Los experimentos se basaron en los parámetros calculados del contorno del labio. En este experimento, el número de estados del HMM se varió entre 20 y 45 estados, se usaron entre 100 y 200 puntos de control, como número de parámetros que definen la forma del labio. En el primer experimento, se utilizó el HMM como clasificador. En la figura 4.1, se observa este experimento usando 5 muestras de entrenamiento para cada caso, y 5 muestras independientes de test

para su evaluación. La tasa de acierto alcanzada fue de 51% para la clasificación con HMM usando entre 20 y 45 estados como puede verse en la tabla 4.12.



**Figura 4.1: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento.**

TABLA 4.12: TASA DE ACIERTO USANDO EL HMM COMO CLASIFICADOR CON 5 MUESTRAS DE ENTRENAMIENTO CON EL TAMAÑO ORIGINAL

Número de estados	Número de puntos de control			
	100	150	175	200
20	30.71% ± 4.36	35.22% ± 4.13	40.47% ± 2.21	31.26% ± 3.73
25	30.92% ± 3.78	37.58% ± 2.52	40.43% ± 1.94	39.72% ± 2.91
30	28.58% ± 3.95	39.42% ± 2.09	43.31% ± 1.86	38.23% ± 2.36
35	31.84% ± 3.36	47.33% ± 1.57	39.02% ± 2.04	46.29% ± 1.92
40	36.73% ± 4.05	51.64% ± 1.37	45.39% ± 1.65	49.80% ± 1.43
45	41.20% ± 3.26	46.83% ± 1.52	42.25% ± 1.83	47.16% ± 1.74

A continuación, se procedió a utilizar la transformación de los datos utilizando el kernel propuesto en el HMM y la posterior clasificación utilizando SVMs. En este experimento se decidió utilizar solamente el kernel Gaussiano, dado que en el primer enfoque arrojó mejores resultados que el kernel lineal.

Los valores de partida de las variables de optimización, fueron aquellos que mostraron los mejores resultados clasificando los datos utilizando el HMM. Así, utilizando las imágenes originales, el número de estados se fijó en 40 y los de

puntos de control, en 150. Los resultados obtenidos se pueden ver en la tabla 4.13.

TABLA 4.13: TASA DE ACIERTO USANDO SVM CON UN KERNEL RBF PARA DIFERENTES NÚMEROS DE MUESTRAS DE ENTRENAMIENTO CON IMÁGENES DE TAMAÑO ORIGINAL

Número de puntos de control	Número de estados	Número de muestras de entrenamiento	Tasa de acierto	g
100	40	5	100% ± 0	$1 \times 10^{-6}$
100	40	4	99.54% ± 0.08	$3 \times 10^{-6}$
100	40	3	99.29% ± 0.56	$3 \times 10^{-6}$
100	40	2	99.01% ± 0.90	$4 \times 10^{-6}$
150	40	5	100% ± 0	$1 \times 10^{-6}$
150	40	4	99.67% ± 0.15	$3 \times 10^{-6}$
150	40	3	99.31% ± 0.40	$3 \times 10^{-6}$
150	40	2	99.17% ± 0.68	$3 \times 10^{-6}$
175	40	5	100% ± 0	$1 \times 10^{-6}$
175	40	4	99.56% ± 0.20	$3 \times 10^{-6}$
175	40	3	99.43% ± 0.58	$3 \times 10^{-6}$
175	40	2	99.37% ± 0.78	$4 \times 10^{-6}$
200	40	5	100% ± 0	$1 \times 10^{-6}$
200	40	4	99.56% ± 0.20	$3 \times 10^{-6}$
200	40	3	99.36% ± 0.52	$3 \times 10^{-6}$
200	40	2	99.15% ± 0.71	$3 \times 10^{-6}$

#### 4.3.2 Base de datos manos

En este estudio [Briceño et al. 2011], se utilizó la base de datos GPDS144hand. Todos los experimentos han sido repetidos cinco veces. Se han usado de 4 a 1 muestras de manos para el entrenamiento, es decir, del 40 al 10 % de las muestras.

Se aplican de igual forma los enfoques implementados en los experimentos anteriores, de forma que se clasifican primero los datos utilizando el HMM de forma aislada y posteriormente, se transforman mediante el kernel propuesto y se clasifica con SVMs utilizando el kernel lineal y Gaussiano. Para el HMM, el número de estados se varió entre 20 y 140 obteniendo una tasa de acierto menor que 85%. En la tabla 4.14, se muestran el mejor resultado con los valores de HMM correspondiente a 60 estados.

TABLA 4.14: TASA DE ACIERTO PARA UN CLASIFICADOR HMM, USANDO INFORMACIÓN DEL CONTORNO.

Número de puntos de control	Número de estados HMM	Número de muestras de entrenamiento	Tasa de acierto
750	60	4	46.61% ± 5.94
		3	38.09% ± 4.73
		2	32.41% ± 4.67
		1	23.55% ± 3.73
300	60	4	72.50% ± 4.77
		3	71.90% ± 4.36
		2	61.04% ± 6.48
		1	54.37% ± 5.71
200	60	4	84.17% ± 6.40
		3	82.33% ± 4.33
		2	79.79% ± 4.03
		1	73.92% ± 5.64
100	60	4	73.12% ± 6.21
		3	71.00% ± 7.84
		2	70.33% ± 2.33
		1	65.63% ± 5.63

En la Tabla 4.15, se muestran las tasas de acierto de los resultados del SVM una vez transformados los datos con el kernel HMM propuesto. En este caso se utilizaron 200 puntos de control con 60 estados de HMM y se usaron kernels lineal y Gausiano en las SVMs. La tasa de acierto es la misma en ambos casos, por lo tanto, se trabajó con el kernel lineal porque es ligeramente más rápido. Finalmente, con una mano de entrenamiento, se ha alcanzado tasas de éxito mayores que 99.9%.

TABLA 4.15: TASA DE ACIERTO USANDO EL CLASIFICADOR SVM CON LOS DATOS TRANSFORMADOS POR EL KERNEL HMM

Número de puntos de control	Número de muestras de entrenamiento	Tasa de acierto usando SVM		
		Lineal kernel	RBF kernel	gamma
750	4	100% ± 0	100% ± 0	1×10 <sup>-6</sup>
	3	99.96% ± 0.08	99.96% ± 0.08	1×10 <sup>-6</sup>
	2	99.96% ± 0.09	99.96% ± 0.09	1×10 <sup>-5</sup>
	1	99.85% ± 0.16	99.85% ± 0.16	1×10 <sup>-6</sup>
300	4	100% ± 0	100% ± 0	1×10 <sup>-6</sup>
	3	99.93% ± 0.15	99.93% ± 0.15	1×10 <sup>-6</sup>
	2	99.96% ± 0.09	99.96% ± 0.09	1×10 <sup>-6</sup>
	1	99.85% ± 0.16	99.85% ± 0.16	1×10 <sup>-6</sup>
200	4	100% ± 0	100% ± 0	1×10 <sup>-6</sup>
	3	99.95% ± 0.11	99.95% ± 0.11	5×10 <sup>-6</sup>
	2	99.96% ± 0.09	99.96% ± 0.09	8×10 <sup>-8</sup>
	1	99.92% ± 0.10	99.92% ± 0.10	4×10 <sup>-8</sup>
100	4	99.96% ± 0.09	99.96% ± 0.09	1×10 <sup>-6</sup>
	3	99.95% ± 0.11	99.95% ± 0.11	1×10 <sup>-6</sup>
	2	99.96% ± 0.09	99.96% ± 0.09	1×10 <sup>-6</sup>
	1	99.87% ± 0.14	99.87% ± 0.14	1×10 <sup>-6</sup>

## 4.4 Variaciones Complejas

### 4.4.1 Base de datos de palabras

Finalmente se evaluó el sistema en bases de datos con variaciones complejas. De la misma forma se aplicaron los dos enfoques. En el primer caso se varió el número de estados entre 30 y 50 estados y el número de puntos de control entre 75 y 250, dejando fijo el porcentaje de entrenamiento al 20% de las muestras. En la tabla 4.16 se muestran los resultados de ambos enfoques.

TABLA 4.16: TASA DE ACIERTO USANDO CLASIFICADORES HMM Y SVM PARA DATOS SIN TRANSFORMAR Y TRANSFORMADOS, USANDO UN 20% DE LAS MUESTRAS PARA EL ENTRENAMIENTO

Puntos	Estados	HMM	Lineal	RBF	Gamma
75	30	10.71% ± 0.32	80.43% ± 18.28	94.48% ± 2.30	8 × 10 <sup>-6</sup>
75	40	10.15% ± 0.92	65.66% ± 2.26	94.78% ± 0.88	2 × 10 <sup>-5</sup>
75	50	10.29% ± 2.44	69.86% ± 8.56	94.59% ± 0.62	8 × 10 <sup>-5</sup>
100	30	10.51% ± 2.62	89.03% ± 0.32	92.70% ± 0.27	6 × 10 <sup>-5</sup>
100	40	10.79% ± 1.21	86.03% ± 4.96	94.95% ± 0.50	6 × 10 <sup>-5</sup>
100	50	10.03% ± 0.55	76.59% ± 0.87	94.27% ± 1.88	6 × 10 <sup>-5</sup>
150	30	13.09% ± 2.48	81.09% ± 0.60	94.86% ± 0.18	3 × 10 <sup>-6</sup>
150	40	11.75% ± 0.51	86.56% ± 3.08	95.12% ± 1.01	3 × 10 <sup>-6</sup>
150	50	10.69% ± 1.89	89.35% ± 2.81	94.56% ± 0.84	2 × 10 <sup>-6</sup>
200	30	12.40% ± 3.04	78.61% ± 6.12	93.93% ± 0.97	6 × 10 <sup>-7</sup>
200	40	11.10% ± 1.23	81.25% ± 9.65	95.25% ± 2.39	2 × 10 <sup>-6</sup>
200	50	15.44% ± 3.77	91.63% ± 2.09	94.21% ± 1.20	2 × 10 <sup>-6</sup>
250	30	13.35% ± 4.05	77.44% ± 21.68	94.43% ± 1.33	6 × 10 <sup>-7</sup>
250	40	15.30% ± 1.10	81.45% ± 14.18	94.66% ± 1.57	6 × 10 <sup>-7</sup>
250	50	12.50% ± 1.84	81.25% ± 17.67	94.56% ± 1.52	2 × 10 <sup>-6</sup>

En la tabla 4.17, se muestran los resultados de ambos enfoques variando el porcentaje de entrenamiento y el número de estados del HMM, buscando la mejor tasa de acierto, fijando el número de puntos de control a 200, como mejor resultado del experimento anterior (véase la Tabla 4.16).

TABLA 4.17: TASA DE ACIERTO USANDO CLASIFICADORES HMM Y SVM PARA DATOS SIN TRANSFORMAR Y TRANSFORMADOS, USANDO 200 PUNTOS DE CONTROL

% Training	Estados	HMM	Lineal	RBF	Gamma
20	30	12.40% ± 3.04	78.61% ± 6.12	93.93% ± 0.97	6×10 <sup>-7</sup>
20	40	11.10% ± 1.23	81.25% ± 9.65	95.25% ± 2.39	2×10 <sup>-6</sup>
20	50	15.44% ± 3.77	91.63% ± 2.09	94.21% ± 1.20	2×10 <sup>-6</sup>
30	30	18.30% ± 4.42	81.95% ± 5.52	95.05% ± 2.78	8×10 <sup>-6</sup>
30	40	18.71% ± 1.21	90.18% ± 0.10	95.13% ± 0.99	6×10 <sup>-6</sup>
30	50	15.62% ± 3.05	85.65% ± 4.18	95.68% ± 2.94	1×10 <sup>-6</sup>
40	30	19.57% ± 0.80	90.22% ± 1.69	94.70% ± 1.47	2×10 <sup>-6</sup>
40	40	20.35% ± 0.18	90.67% ± 1.04	95.79% ± 0.79	6×10 <sup>-6</sup>
40	50	21.92% ± 0.92	84.55% ± 2.09	97.74% ± 0.98	2×10 <sup>-6</sup>
40	60	21.09% ± 1.35	81.12% ± 6.08	97.92% ± 0.73	4×10 <sup>-7</sup>
40	70	20.74% ± 1.60	85.11% ± 7.18	96.53% ± 1.72	4×10 <sup>-7</sup>
50	30	21.51% ± 2.58	82.13% ± 8.17	97.03% ± 1.25	4×10 <sup>-6</sup>
50	40	23.02% ± 2.21	87.55% ± 1.54	96.51% ± 0.36	6×10 <sup>-6</sup>
50	50	23.49% ± 0.37	89.48% ± 3.09	96.41% ± 0.66	8×10 <sup>-7</sup>
50	60	24.06% ± 0.15	84.43% ± 0.96	96.93% ± 1.55	8×10 <sup>-6</sup>
60	30	22.72% ± 0.28	81.44% ± 3.04	96.48% ± 0.09	6×10 <sup>-6</sup>
60	40	26.95% ± 1.10	85.22% ± 0.46	97.46% ± 0.46	6×10 <sup>-6</sup>
60	50	28.32% ± 2.30	90.36% ± 0.92	98.50% ± 0.82	6×10 <sup>-6</sup>
60	60	26.63% ± 3.04	86.52% ± 2.67	97.72% ± 1.38	6×10 <sup>-7</sup>

Finalmente se observa, que con 50 estados se obtiene la mejor precisión, y cuando mayor es el porcentaje de muestras de entrenamiento, mayor será la bondad del sistema.

#### 4.4.2 Base de datos de firmas

En estas dos bases de datos usadas, cuando se utilizó los HMMs como sistema clasificador, se varió el número de puntos de control entre 100 y 250 y el número de estados, entre 40 y 80, como puede verse en la Tabla 4.18. Se utilizaron algo más de 20% de las muestras para el entrenamiento, en particular, se usan 6 muestras para entrenar el sistema. Se observa, que la mayor tasa de acierto se consigue con 250 puntos y 50 estados, logrando un éxito de 67.56% ± 1.69, para la base de datos MYCT-75. Mientras que para la base de datos GPDS100, se logró un 69.62% ± 1.52, con 150 puntos y 40 estados.

TABLA 4.18: TASA DE ACIERTO USANDO CLASIFICADORES HMM PARA FIRMAS MANUSCRITAS

Número de Puntos	Estados	Éxito para MYCT-75	Éxito para GPDS100
100	40	59.96% ± 5.81	59.65% ± 2.14
100	50	62.81% ± 3.45	60.01% ± 2.24
100	60	64.67% ± 2.15	62.61% ± 1.17
100	70	64.92% ± 2.39	60.58% ± 2.01
100	80	65.02% ± 2.48	63.14% ± 0.25
150	40	58.35% ± 0.67	69.62% ± 1.52
150	50	60.93% ± 1.42	59.68% ± 1.14
150	60	63.42% ± 2.85	59.73% ± 2.05
150	70	64.13% ± 2.37	56.86% ± 5.33
150	80	64.89% ± 1.07	55.74% ± 3.36
200	40	62.71% ± 3.05	59.32% ± 1.04
200	50	63.78% ± 0.50	60.03% ± 4.28
200	60	59.33% ± 3.75	60.35% ± 2.26
200	70	63.42% ± 2.85	59.95% ± 2.82
200	80	64.13% ± 2.37	56.01% ± 1.19
250	40	59.42% ± 5.66	63.23% ± 1.09
250	50	67.56% ± 1.69	62.03% ± 1.56
250	60	65.82% ± 5.12	62.72% ± 3.10
250	70	64.89% ± 1.07	60.26% ± 0.81
250	80	63.33% ± 2.72	59.68% ± 1.17

Se repitió el experimento pero aplicando la transformación y las SVMs, los valores de los puntos de control se variaron en el mismo rango como se puede ver en las Tablas 4.19 y 4.20. Se usaron tanto el kernel lineal como el kernel RBF. Para la base de datos MYCT-75, se alcanzó con 200 puntos de control y 80 estados, un total de  $98.80\% \pm 0.53$  usando el kernel lineal y  $98.58\% \pm 0.55$ , usando el kernel RBF.

TABLA 4.19: TASA DE ACIERTO USANDO CLASIFICADORES SVM PARA FIRMAS MANUSCRITAS DE LA BASE DE DATOS MYCT-75

Número de Puntos	Estados	Linear SVM	RBF SVM	G
100	40	97.77% ± 0.93	97.67% ± 0.94	$6 \times 10^{-7}$
	50	98.09% ± 0.79	98.02% ± 0.84	$6 \times 10^{-7}$
	60	98.21% ± 0.57	98.11% ± 0.62	$8 \times 10^{-7}$
	70	98.35% ± 0.48	98.27% ± 0.55	$6 \times 10^{-7}$
	80	98.41% ± 0.31	98.30% ± 0.42	$4 \times 10^{-7}$
150	40	97.89% ± 0.67	97.77% ± 0.81	$6 \times 10^{-6}$
	50	98.04% ± 0.61	98.01% ± 0.73	$6 \times 10^{-6}$
	60	98.15% ± 0.58	98.23% ± 0.62	$6 \times 10^{-6}$
	70	98.22% ± 0.41	98.31% ± 0.43	$6 \times 10^{-6}$
	80	98.34% ± 0.47	98.29% ± 0.61	$6 \times 10^{-6}$
200	40	97.96% ± 0.68	98.22% ± 0.43	$2 \times 10^{-7}$
	50	97.82% ± 0.54	98.40% ± 0.92	$2 \times 10^{-7}$
	60	98.71% ± 0.51	98.48% ± 0.68	$2 \times 10^{-7}$
	70	98.67% ± 0.48	98.58% ± 0.47	$2 \times 10^{-7}$
	80	98.80% ± 0.53	98.58% ± 0.55	$2 \times 10^{-7}$
250	40	98.18% ± 0.38	97.91% ± 0.51	$2 \times 10^{-7}$
	50	98.58% ± 0.07	98.53% ± 0.54	$6 \times 10^{-7}$
	60	97.87% ± 0.13	98.04% ± 0.07	$4 \times 10^{-7}$
	70	98.21% ± 0.39	98.19% ± 0.32	$6 \times 10^{-7}$
	80	98.49% ± 0.43	98.36% ± 0.34	$4 \times 10^{-7}$



Para la base de datos GPDS100, se alcanzó 99.16%  $\pm$ 0.09 usando el kernel lineal y 99.02%  $\pm$  0.20 cuando se implementó el kernel RBF, con 200 puntos de control y 80 estados, tal y como se puede observa en la Tabla 4.20.

TABLA 4.20: TASA DE ACIERTO USANDO CLASIFICADORES SVM PARA FIRMAS MANUSCRITAS DE LA BASE DE DATOS GPDS100

Número de Puntos	Estados	Lineal SVM	RBF SVM	G
100	40	98.81% $\pm$ 0.40	98.59% $\pm$ 0.50	$4 \times 10^{-7}$
	50	99.03% $\pm$ 0.24	98.98% $\pm$ 0.24	$8 \times 10^{-7}$
	60	98.77% $\pm$ 0.08	98.82% $\pm$ 0.22	$8 \times 10^{-7}$
	70	98.91% $\pm$ 0.25	98.74% $\pm$ 0.34	$8 \times 10^{-7}$
	80	98.89% $\pm$ 0.32	98.94% $\pm$ 0.24	$4 \times 10^{-7}$
150	40	98.09% $\pm$ 0.81	98.16% $\pm$ 0.77	$2 \times 10^{-6}$
	50	98.42% $\pm$ 0.51	98.39% $\pm$ 0.51	$2 \times 10^{-6}$
	60	98.37% $\pm$ 0.33	98.37% $\pm$ 0.33	$2 \times 10^{-6}$
	70	98.17% $\pm$ 0.29	98.08% $\pm$ 0.39	$2 \times 10^{-6}$
	80	98.56% $\pm$ 0.50	98.67% $\pm$ 0.55	$2 \times 10^{-6}$
200	40	97.88% $\pm$ 0.11	97.70% $\pm$ 0.11	$8 \times 10^{-7}$
	50	98.65% $\pm$ 0.26	98.56% $\pm$ 0.29	$6 \times 10^{-7}$
	60	98.77% $\pm$ 0.13	98.72% $\pm$ 0.11	$8 \times 10^{-7}$
	70	98.79% $\pm$ 0.32	98.80% $\pm$ 0.34	$8 \times 10^{-7}$
	80	99.16% $\pm$ 0.09	99.02% $\pm$ 0.20	$8 \times 10^{-7}$
250	40	98.93% $\pm$ 0.12	99.88% $\pm$ 0.29	$8 \times 10^{-7}$
	50	98.67% $\pm$ 0.12	98.61% $\pm$ 0.16	$2 \times 10^{-7}$
	60	98.92% $\pm$ 0.33	98.68% $\pm$ 0.29	$6 \times 10^{-7}$
	70	98.96% $\pm$ 0.30	98.93% $\pm$ 0.29	$4 \times 10^{-7}$
	80	98.77% $\pm$ 0.67	98.81% $\pm$ 0.74	$6 \times 10^{-7}$

#### 4.4.3 Base de datos de signos dactilológicos

En este experimento se varió el número de estados y el porcentaje de las muestras de entrenamiento para conseguir la máxima tasa de acierto. Los resultados son los siguientes:

TABLA 4.21: TASA DE ACIERTO USANDO CLASIFICADORES HMM PARA EL LENGUAJE DACTILOLÓGICO.

Muestras de entrenamiento	Número de estados				
	20	45	55	65	100
20%	71.64% $\pm$ 21.81	82.58% $\pm$ 2.32	83.72% $\pm$ 1.46	82.36% $\pm$ 6.65	60.2% $\pm$ 45.64
40%	74.1% $\pm$ 11.85	85.65% $\pm$ 2.64	87.43% $\pm$ 6.22	87.77% $\pm$ 0.82	68.1% $\pm$ 25.20
60%	75.76% $\pm$ 9.20	87.8% $\pm$ 1.95	88.32% $\pm$ 3.24	87.84% $\pm$ 3.76	64.08% $\pm$ 24.86
80%	75.32% $\pm$ 1.41	89.6% $\pm$ 13.17	90.18% $\pm$ 7.25	89.48% $\pm$ 13.14	70.64% $\pm$ 15.97

De la Tabla 4.21, se deduce que el mejor porcentaje de entrenamiento es para 80% y 55 estados, con una tasa de acierto de 90.18%, presentando una pequeña varianza. De este modelo se genera el kernel HMM, y se le aplicará SVM, tanto con un kernel de función lineal y función Gaussiana (RBF), Los resultados de estos test se presentan en la Tabla 4.22:

TABLA 4.22: TASA DE ACIERTO USANDO SVM CON KERNEL LINEAL Y RBF.

Kernel utilizado en las SVM	Parámetro	Media±Desviación Estándar
Lineal	-----	94.56% ± 11.65
RBF	$g=1.28 \times 10^{-14}$	92.08% ± 3.15
RBF	$g=6.28 \times 10^{-16}$	97.52% ± 1.87
RBF	$g=8.28 \times 10^{-16}$	98.34% ± 0.37
RBF	$g=1.28 \times 10^{-17}$	97.36% ± 2.13
RBF	$g=1.28 \times 10^{-18}$	93.44% ± 1.67

## **CAPÍTULO 5**

### **Análisis de resultados**

---

En este capítulo, se lleva a cabo el análisis de los resultados obtenidos en los experimentos descritos en el capítulo 4. Mediante este análisis, se caracterizará el sistema propuesto frente a los distintos grados de complejidad de los contornos estudiados.

#### **5.1 Análisis según BBDD**

Se va a estudiar el comportamiento del método propuesto, en función de cada base de datos.

### **5.1.1 Análisis de la base de datos Sintética.**

En esta base de datos, los números de estados utilizados fueron 40, 50 y 60, siendo el de 50 estados el que dio los mejores resultados. Se usaron 360 puntos de control y el tamaño de la base de datos es de 21 clases con 100 muestras por clase. Respecto al porcentaje de muestras de entrenamiento, se observa que es necesario al menos 21% de número muestras para alcanzar una tasa de acierto superior al 99%, de forma que se logró alcanzar un máximo de  $99.48\% \pm 0.13$  con kernel lineal y un  $99.05\% \pm 0.28$  con kernel RBF, este último con una gamma (g) de  $2 \times 10^{-7}$ .

A medida que se aumenta el porcentaje de entrenamiento, se va alcanzado un mejor resultado, siendo en el máximo para nuestros experimentos de un 30% de muestras de entrenamiento ( $99.71\% \pm 0.14$  con kernel lineal y un  $99.59\% \pm 0.17$  con kernel RBF y una gamma (g) de  $3 \times 10^{-7}$ ). El kernel lineal presenta mejor comportamiento frente al kernel RBF, si bien las diferencias no superan el medio punto cuando la tasa de acierto es superior al 99%. El valor de la g parece no afectar al sistema en el caso del kernel RBF. La desviación estándar es estable y siempre es menor que 0.4 cuando la tasa de acierto supera el 99%, tanto con kernel RBF como para el kernel lineal.

### **5.1.2 Análisis de la base de datos Polen**

La base de datos está compuesta por 47 clases con un total de 564 muestras. El intervalo de puntos utilizados para el estudio fue de 150 a 400 puntos con saltos de 50 puntos. El número de estados fijados fue de 30, 50 y 70, obteniéndose el mejor resultado con 30 estados en el primer enfoque, utilizando HMM como clasificador, en donde se obtuvo un máximo de  $19.15\% \pm 0.71$ , con una desviación estándar que osciló entre 0,35 y 2,17. En el segundo enfoque, realizando la transformación de los datos y usando SVMs como sistema de clasificación se realizaron dos experimentos, primero se varió el número de puntos entre 150 a 400 con saltos de 50 puntos y se fijó el número de estados a 50. Los mejores resultados se consiguieron en 300 y 350 puntos con un valor de gamma de  $4 \times 10^{-7}$ . Los máximos alcanzados  $93.85\% \pm 1.43$  usando kernel lineal y  $93.38\% \pm 1.47$  usando kernel RBF. Se observa que el kernel lineal presenta una ligera mejora, aunque ésta es menor a 0.5%. El valor de la gamma parece no ejercer ninguna influencia en los resultados, los cuales son dependientes de los números de puntos de control utilizados. La desviación estándar es relativamente grande variando entre 1.43 y 2.14 en sus máximos y no es estable en el resto de configuraciones.

En el segundo experimento del segundo enfoque se fijó el número de puntos a 300 puntos y se varió el número de estados entre 30 y 70 con saltos de 10 estados. El mejor porcentaje de acierto conseguido fue de  $93.85\% \pm 1.43$  utilizando kernel lineal y  $93.38\% \pm 1.47$  utilizando kernel RBF con 50 estados. De igual forma, el kernel lineal presenta una ligera mejora aunque también menor que 0.5% en sus máximos. El comportamiento de la desviación estándar es similar que en el primer experimento, sin embargo es menor que 0.05 en sus dos máximos. De igual forma la gamma no presenta ninguna influencia sobre los resultados y el comportamiento global del sistema.

### **5.1.3 *Análisis de la base de datos Laurisilva de Costa Rica***

En la base de datos de Laurilva de Costa Rica solamente se aplicó el segundo enfoque. Esta base de datos consta de 14 clases con 60 muestras por clase. Se fijó el número de puntos a 360 y se realizó el estudio fijando el número de estados a 40, 50 y 60 estados. En el caso de usar 40 estados fue necesario al menos un 20 % de muestras para la fase de entrenamiento para alcanzar  $100\% \pm 0$ , mientras que con 50 estados con fue suficiente entrenar el sistema con un 4% de las muestras, utilizando tanto el kernel lineal como el kernl RBF. Fijando a 60 estados, los resultados obtenidos fueron variables. Implementando el kernel RBF, se alcanzó el  $100\% \pm 0$  exceptuando cuando usan entre el 1% y 4% de las muestras de entrenamiento. Por otro lado, usando el kernel lineal, fue necesario utilizar al menos un 5% y cuando el porcentaje de entranamiento fue 20 % el porcentaje de éxito bajo ligeramente a  $99.83\% \pm 0.37$ . La gamma no ejerció ningún efecto y el valor óptimo de fue de  $5 \times 10^{-7}$ . Se concluye que el kernel RBF presenta mejor respuesta ya que con 50 y 60 estados necesita menor porcentaje de entrenamiento. La desviación estándar es pequeña y estable en los casos en los que el sistema no alcanza el 100% de acierto.

### **5.1.4 *Análisis de la base de datos Laurisilva de Canarias***

En este experimento también se utilizó el segundo enfoque. La base de datos de Laurisilva Canaria consta de 16 clases y 75 muestras por clase. El número de puntos de control fue fijado a 360 y el estudio se realizó con 40, 50 y 60 estados. Se alcanzó el 100% de éxito con 40 estados y al menos 25% de las muestras de entrenamiento. En el caso de utilizar 50 estados fue necesario al menos entre el 15% y el 30%, para coseguir

el mismo resultado. Cuando se fijó el sistema a 60 estados con kernel lineal, el 100% de éxito se consiguió utilizando 20% y 30% de muestras de entrenamiento. Los casos restantes arrojaron un resultado de 99,98%. En el mismo escenario, es decir, con 60 estados pero esta vez aplicando el kernel RBF, fue necesario al menos 15% para alcanzar el 100%. Puede concluirse que el kernel RBF presenta mejor comportamiento, en donde la gamma parece tener una influencia significativa sobre los resultados obtenidos. Cuando no se alcanzó el 100% de éxito, la desviación estándar fue muy pequeña y además muy estable.

### **5.1.5 *Análisis de la base de datos Labios***

La base de datos de labios está compuesta por 50 clases y 10 muestras por clase. En este estudio se implementaron los dos enfoques. El número de puntos utilizados fue de 100, 150, 175 y 200 y el número de estado se varió entre 20 y 45 estados. En el caso del primer enfoque, el mejor resultado correspondió a  $51.64\% \pm 1.37$  utilizando 40 estados y 150 puntos. En estos experimentos se entrenó el sistema con el 50% de las muestras. En este enfoque la desviación estándar fue grande y muy inestable, oscilando entre 1.43 y 4.05.

En el segundo enfoque se utilizó solamente el kernel RBF y se fijó el número de estados a 40, ya que en el primer enfoque el sistema presentó el mejor comportamiento. En este estudio se varió el porcentaje de entrenamiento del 20% al 50% de las muestras para entrenar. Independientemente del número de puntos utilizados, siempre se alcanzó el 100% de éxito en la clasificación cuando se entrenó el sistema con el 50% de las muestras y con una gamma igual a  $1 \times 10^{-6}$ . El valor de gamma parece ejercer influencia significativa en el resultado, si bien es cierto, también es dependiente del porcentaje de entrenamiento.

### **5.1.6 *Análisis de la base de datos Manos***

Esta base de datos está formada por 144 clases y 10 muestras por clase. Se implementaron los dos enfoques y se fijaron los puntos de control a 100, 200, 300 y 750. El porcentaje de muestras de entrenamiento se varió entre 10% y 40% con saltos de 10%. En este enfoque el mejor resultado se obtuvo con 200 puntos y 60 estados

consiguiendo hasta un  $84.17\% \pm 6.40$  de porcentaje de éxito. La desviación estándar en este enfoque es grande y muy inestable, oscilando entre 2.33 y 7.84.

En el segundo enfoque se utilizaron el kernel lineal y el kernel RBF, fijando el número de estados a 60. El resultado obtenido con ambos kernel fue de 100% siempre que se entrenaron los sistemas con el 40% de las muestras. La gamma no afecta el comportamiento de los sistemas en ninguno de los casos. La desviación estándar es muy pequeña y estable cuando no se consigue el 100% de éxito. Vale la pena mencionar que los sistemas fueron capaces de clasificar con más 99% de éxito con solo el 10% de las muestras, es decir, con una muestra.

### ***5.1.7 Análisis de la base de datos Palabras***

La base de datos de palabras está compuesta por 64 clases y 120 muestras por clase. Se implementaron los dos enfoques. En primer lugar se entrenaron los sistemas usando el 20 % de las muestras. Los números de puntos que se usaron, fueron 75, 100, 150, 200 y 250, y el número de estados se varió entre 30 y 50 con saltos de 10 estados. El mejor resultado se alcanzó con 200 puntos y 50 estados. En el primer enfoque se alcanzó un máximo de  $15.44\% \pm 3.77$  de porcentaje de éxito y en el segundo enfoque,  $91.63\% \pm 2.09$  usando el kernel lineal y  $94.21\% \pm 1.20$  usando el kernel RBF. En el primer enfoque la desviación estándar fue muy variable y en el segundo enfoque el comportamiento fue más estable cuando se utilizó el kernel RBF. De esta forma, en la segunda parte de este estudio se fijó el número de puntos de control a 200 y se varió el número de estados entre 30 a 70 estados y el porcentaje de entrenamiento entre 20% y 60% de las muestras. En este caso, en el primer enfoque se alcanzó un  $28.32\% \pm 2.30$  con 50 estados y 60% de entrenamiento. Con estos parámetros en el segundo enfoque se alcanzó un  $90.36\% \pm 0.92$  de porcentaje de éxito usando el kernel lineal y un  $98.50\% \pm 0.82$  de éxito usando el kernel RBF. La desviación estándar se presentó inestable, variando sus valores entre 0.18 y 4.42 en el primer enfoque y entre 0.09 y 9.65 en el segundo enfoque si bien fue menor que un punto en los valores máximos.

### **5.1.8 Análisis de la base de datos Firmas.**

Se encuentra formada por 800 clases y 24 muestras por clase, pero para este experimento se usaron 100 clases, llamada base de datos GPDS100, y muy usada de esta manera en otros trabajos científicos, como se recoge en [Vargas, 2010]. Se aplicaron los dos enfoques en los cuáles se varió el número de puntos de control entre 100 y 250 puntos y entre 40 y 80 estados. En todos los experimentos se aplicó la técnica de validación cruzada *Hold-Out* algo superior al 20% (6 muestras para el entrenamiento). El mayor porcentaje de éxito alcanzado en el primer enfoque fue de  $69.62\% \pm 1.52$  con 150 puntos y 40 estados mientras que en el segundo enfoque, el porcentaje de éxito correspondió a  $99.16\% \pm 0.09$  con kernel lineal y  $99.02\% \pm 0.20$  con kernel RBF para la base de datos GPDS100. La varianza en el primer enfoque es inestable variando entre 0.25 y 5.33. En el segundo enfoque fue inestable y siempre menor que un punto porcentual. La gamma no parece tener influencia en los resultados cuando se aplicó el kernel RBF.

### **5.1.9 Análisis de la base de datos Dactilológico.**

Esta base de datos está compuesta por 25 clases y 50 muestras por clase. En este estudio se varió el porcentaje de entrenamiento entre el 20% y el 80% de las muestras y entre 20 y 100 el número de estados. El número de puntos de control fue fijado a 360 puntos. En el primer enfoque se alcanzó un máximo de porcentaje de éxito de  $90.18\% \pm 7.25$  usando el 80% de las muestras para el entrenamiento y fijando a 55 el número de estados. Fijando estos valores para el segundo enfoque, se alcanza  $94.16\% \pm 11.64$  utilizando el kernel lineal y  $98.34\% \pm 0.37$  cuando se aplicó el kernel RBF con una gamma igual a  $8.28 \times 10^{-16}$ . En el primer enfoque la desviación estándar es grande e inestable, variando entre 0.82 y 45.64. En el segundo enfoque continúa siendo bastante inestable pero pequeña en el valor máximo siendo de 0.37.

Finalmente, se ha sido construida una tabla resumen con los valores más representativos de cada una de las bases de datos. Esta tabla nos servirá para analizar los resultados, en función de la variabilidad de la forma para cada base de datos (véase la Tabla 5.1).



TABLA 5.1: DESCRIPCIÓN DE LOS VALORES MÁS SIGNIFICATIVOS DE CADA BASE DE DATOS

BBDD	# Clases (Muestras)	Variabilidad	# puntos	# estados	Kernel	Training (%)	Resultado
Sintética	21 (100)	Sintética	360	50	Lineal	30	99.71% $\pm$ 0.14
Polen	100 (12)	Simples	300	50	Lineal	50	93.85% $\pm$ 1.43
Hojas	16 (75)		360	60	Lineal RBF	10	99.98% $\pm$ 0.04
Hojas CR	14 (60)		360	50	Lineal RBF	10	100%
Labios	50 (10)	Moderada	175	50	RBF $3 \times 10^{-6}$	30	99.43% $\pm$ 0.58
Manos	100 (10)		200	60	RBF $5 \times 10^{-6}$	30	99.95% $\pm$ 0.11
Palabras	64 (120)	Compleja	200	50	RBF $1 \times 10^{-6}$	30	95.68% $\pm$ 2.94
Firmas	son 800 (24)		200	80	Lineal	20	99.16% $\pm$ 0.09
Signos	25 (50)		360	55	RBF $1 \times 10^{-9}$	80	99.68% $\pm$ 0.11

## 5.2 Análisis según la variabilidad del contorno de las base de datos.

Se analizarán las diferentes bases de datos en función del tipo de base de base y en función del número de puntos, número de estados, tipo de kernel de la máquina de soporte vectorial, las muestras de entrenamiento y otros aspectos particulares de cada base de datos.

### 5.2.1 Análisis de la bases de datos de tipo simple.

Este grupo está conformado por las bases de datos de polen, de hojas de Costa Rica y de hojas de Canarias. Se puede observar de la tabla 5.1, que el comportamiento para los mejores resultados es bastante estable. El número de estados varía entre 50 y 60, el número de puntos de control es entre 300 y 360, el clasificador SVM es modelado con un kernel lineal, todo ello usando aproximadamente 6 o 7 muestras para entrenar el sistema. Esto equivale a tomar un punto de control por grado angular, y por estado trata de incluir 6 o 7 puntos de control de media. De esta forma queda bien caracterizado un contorno simple para esta propuesta. La estabilidad del sistema es apreciable en la desviación típica de la tasa éxito en toda las base de datos. De esta forma, éstos serían los parámetros de configuración del sistema para estos contornos simples.

### **5.2.2 *Análisis de la bases de datos de tipo moderada.***

Este grupo está compuesto por las bases de datos de labios y manos. La tasa de éxito se logra con pocas muestras de entrenamiento, con un 30% de ellas (tres muestras), manteniendo la estabilidad del sistema con una muy baja desviación estándar (véase Tabla 5.1). El número de estados está entre 40 y 60, usando 175 y 200 puntos, respectivamente. Los valores de éxito entorno a estos parámetros, son ligeramente inferiores, perdiendo solo unas décimas de porcentaje de éxito. En este caso, el SVM utiliza un kernel RBF con una gamma de  $3 \times 10^{-6}$  y  $5 \times 10^{-6}$ , siendo valores muy similares. El comportamiento es muy similar en ambas aplicaciones a pesar del que el números de clases es el doble uno del otro, de 50 a 100 clases.

Para contornos de complejidad moderada, hacen falta menos estados y menos puntos de control que para las formas simples, y al tener más cambios que las simples, solo con esos puntos de control se identifica mejor la forma.

### **5.2.3 *Análisis de la bases de datos de tipo compleja.***

Este grupo está formado por las bases de datos de palabras, firmas, y signos dactilológicos. En este caso, las similitudes son más amplias. Se puede ver que las palabras y los signos siguen una similar pauta, debido a las hampas y jambas de la escritura y a los dedos, respectivamente. De hecho los valores para este tipo de contorno son muy similares, con un número de estados entre 50 y 55 y con kernel RBF para el SVM. En cambio, las firmas, tienen muchísima más variabilidad, y no hay un patrón de similitud entre ellas. Esto queda reflejado en el número de estados, alcanzando hasta 80 estados para el mejor resultado, aunque con 50 estados, presenta un buen comportamiento, perdiendo solo unas décimas de porcentaje de éxito. Además el kernel del SVM, es lineal, diferente a las anteriores bases de datos.

También se observa, que el número de puntos de control es similar a los de variabilidad moderada, quizás ligeramente mayor. Y por tanto, se ve que la complejidad de la forma necesita similares números de puntos o algo superiores, pero menos que las formas sencillas; donde la gran similitud de formas obliga a usar más puntos para lograr identificarla de forma correcta.

#### **5.2.4 Análisis de todas las bases de datos.**

En general, las tasas de éxito presentan una buena respuesta y una gran estabilidad, mostrada por los bajos valores de la desviación estándar (véase la Tabla 5.1). Las muestras de entrenamiento son relativamente bajas, si bien es cierto, que al aumentar el número de muestras de entrenamiento, el éxito mejora.

Si se tratase de generalizar el método, se observa que para los mejores resultados, los valores de los parámetros están todos dentro de un rango. El número de puntos de control está entre 175 y 360. El número de estados está entre 40 y 80. Y en general, considerando el número de muestras de entrenamiento, se puede considerar bajo, aproximadamente entre 3 y 6 muestras para contornos de variabilidad simple y moderada; y entre 6 y 40 para variabilidades complejas del contorno. Si estos valores se tratasen ajustar a unos valores cerrados, por ejemplo, 250 puntos de control, para 50 estados, usando un kernel RBF, se obtendría un buen resultado, aunque no el mejor, según las tablas de los resultados de todas las bases de datos. Se podría considerar como un punto de partida, para la optimización del sistema (o ajuste fino) en función de la aplicación. Con esto, se podría reducir en gran medida, la experimentación a realizar para futuros usos de otras aplicaciones a formas de cualquier variabilidad.

### **5.3 Extensión de la usabilidad del modelo y otras consideraciones.**

En esta sección se van a considerar diversos aspectos de las bases de datos, como el coste computacional, y particularidades de cada aplicación que se salen de un estudio generalizado.

#### **5.3.1 Coste computacional del sistema.**

Con respecto al coste computacional, como regla general, se observa que el entrenamiento del sistema es muy lento, pudiendo tardar varias horas para la creación del modelo, y tiempos relativamente rápidos para la evaluación de una muestra.

Durante estos años de tesis, se han usado diferentes tipos de ordenadores sobre diferentes sistemas operativos. Como ejemplo y referencia general al uso de este método, se va a mostrar la ejecución para la base de datos de manos, en la que utilizando un Procesador Intel® Core™ i5-460M (3M Cache, 2.53 GHz) con 4GB de memoria

RAM, en la que se necesitó 327 milisegundos para evaluar una muestra. Para generar el modelo, se empleó 20 minutos usando dos muestras para entrenar, y bajo un sistema operativo Windows 7, ejecutando el modelo sobre el lenguaje de programación MATLAB. Esto se puede generalizar para el resto de bases de aplicaciones usadas, a grandes rasgos.

Hay que considerar que el sistema propuesto ha sido chequeado en MATLAB, y no ha sido la finalidad de esta tesis, la reducción de coste computacional con técnicas de computación en paralelo o la utilización de GPUs sino la creación del sistema y mostrar su comportamiento frente a diferentes aplicaciones. En el caso de trabajar con estos sistemas, los tiempos se podrían ver drásticamente reducidos.

Al tiempo usado para la evaluación de la muestras en el sistema de clasificación, hay que sumarle el tiempo de preprocesado y parametrización para la obtención de puntos de control. Continuando con el ejemplo de la base de datos de manos, el preprocesado aplicado consta de la detección de la mano en la imagen, su binarización y detección de contorno, usando un tiempo de 274 milisegundos. Además este tiempo es incrementado por el cálculo de 200 puntos de control por la parametrización angular, empleando un tiempo de 942 milisegundos, para dicha muestra. Por tanto, el tiempo para evaluar una muestra es de 1543 milisegundos.

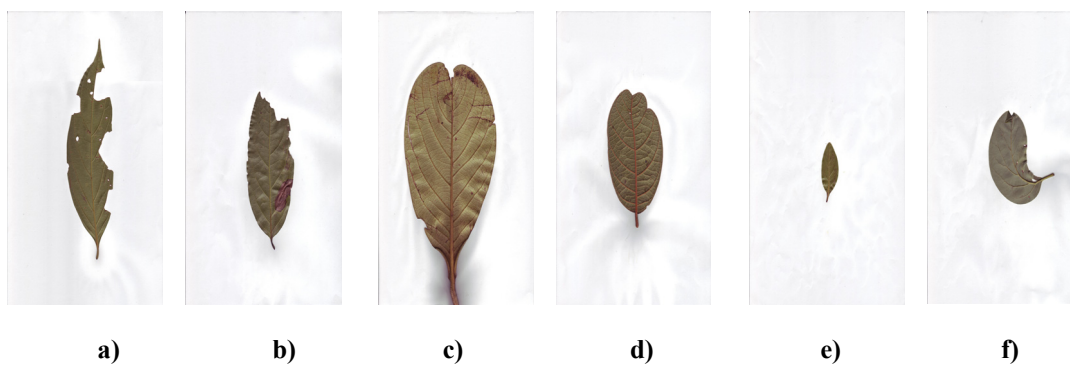
### **5.3.2 Robustez frente a ruido.**

Un aspecto muy importante de los sistemas de clasificación es la robustez que éstos tengan frente a la presencia de ruido en la fuente de información. Este fenómeno es inevitable ya viene impuesto, entre otros, por las condiciones y el entorno de adquisición, el tipo de dispositivos que lleva a cabo la captura de datos, el protocolo utilizado, la naturaleza misma de la fuente información, etc. Por este motivo dentro de los estudios realizados, se utilizaron dos bases de datos cuya influencia del ruido era mayor de forma significativa en comparación con el resto de la base de datos.

El primer estudio corresponde con la base de datos sintética, donde se observa que el sistema es muy robusto, logrando una tasa de éxito de  $89.95\% \pm 2.83$  con tan solo 4% de las muestras de entrenamiento y hasta  $99.71\% \pm 0.14$  cuando este porcentaje es del 30%. Dado que el tipo de imágenes son de variación simple, la dimensionalidad

conseguida fue tal, que los kernel usados en las SVMs presentaron respuestas similares, siendo el kernel lineal el que mostró una ligera mejora.

La segunda base de datos donde se estudió de forma particular este aspecto, corresponde con muestras de hojas de la base de datos de Costa Rica, donde se utilizaron muestras con el contorno afectado de forma notable por insectos predadores, enfermedades degenerativas o deformaciones circunstanciales. La figura 5.1 se observan varios ejemplos de este tipo de muestras.



**Figura 5.1:** (a, b) Ejemplos de muestras de hojas con ruido: muestras de comidas por insectos. (c, d) Mutación o deformaciones relacionadas con enfermedades. (e) Una muestra joven. (f) deformación circunstancial.

En este caso se consigue una tasa de acierto de más del 99% utilizando solamente una muestra de entrenamiento, y un 100% cuando este porcentaje aumentaba. El número de estados estudiados fueron 40, 50 y 60 y en todos ellos el éxito solamente dependió solamente del porcentaje de entranamiento ya que el número de puntos de control se fijó a 360 todos los casos. De esta forma se demuestra que el sistema es robusto frente a grandes influencias de ruido sobre la fuente de información.

### **5.3.3 Generalización del sistema para verificación en manos.**

Una vez comprobado que el sistema proporciona una buena respuesta en el problema de clasificación, se realiza un estudio configurando el sistema en modo verificación, es decir, de forma que el sistema verifique si una muestra dada, pertenece a la clase que dice pertenecer. En este estudio se utilizaron todas las clases de las bases de datos de manos GPDS144hand database [Ferrer et. al, 2007], las clases en el rango

infrarrojo de la GPDS100hand3Band database [Ferrer y Morales, 2011] y todas las clases de la UST database [Kumar et. al, 2003].

En la primera parte de este estudio se utilizaron las bases de datos en el rango visible y en la segunda parte se utilizó la base de datos en el rango del infrarrojo. La medida utilizada en los sistemas configurados en modo verificación es el ratio de igual error (EER o *Equal Error Rate*), la cual corresponde a la tasa en la que el sistema acepta y rechaza los errores es igual, es decir el ratio de falsa aceptación (FAR) y el de falso rechazo (FRR) se igualan. Cuanto más bajo es el EER se considera que el sistema es más exacto. Se define también la curva ROC (acrónimo de Receiver Operating Characteristic, o Característica Operativa del Receptor) como una representación gráfica de la sensibilidad frente a  $(1 - \text{especificidad})$  para un sistema clasificador binario según se varía el umbral de discriminación, siendo este el caso de los sistemas configurados en modo verificación. En la figura 5.2, se representan las curvas ROC utilizando las bases de datos GPDS144 y UST. Se tiene que entrenando el sistema con dos muestras se obtienen un EER= 0.017% cuando se uso la base de datos GPDS144 y un EER=0.31% cuando se uso la base de datos UST.

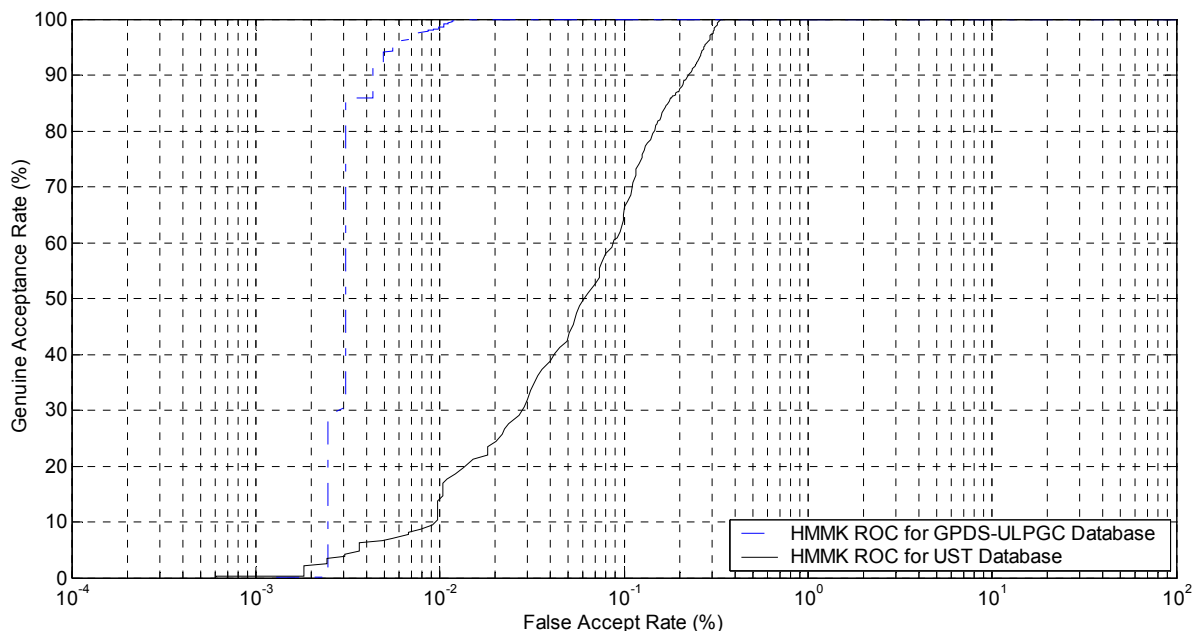


Figura 5.2: Curva ROC utilizando las bases de datos GPDS144 y UST usando dos muestras de entrenamiento.

En la segunda parte de este estudio se repitieron los experimentos utilizando las imágenes en el rango del infrarrojo (1470 nm) de la base de datos GPDS100hand3Band para validar el método propuesto. En este caso, cuando el sistema se entrenó con tres muestras, se obtuvo un EER=0%, confirmando dicha validación.

Con los resultados obtenidos para las tres bases de datos de manos, se prueba que el sistema puede funcionar bajo un sistema de verificación en manos, y vemos su independencia para diferentes tipos de rangos del espectro. Se generaliza así el sistema y se valida el modelo desarrollado.

#### 5.3.4 Reducción de la resolución en la base de datos de labios.

En este estudio se analizó el comportamiento del sistema frente a la reducción de resolución de las imágenes en modo clasificación. El objeto fue determinar si el sistema era capaz de obtener la información discriminante de esas imágenes independientemente de la resolución. Las imágenes originales se redimensionaron a una escala de reducción de 3:1, 4:1 y 5:1.

En este estudio se aplicó la misma metodología usada sobre las imágenes originales. Por lo tanto, primero se varió el número de puntos de control entre 100, 150, 175 y 200. El número de estados se varió entre 20 y 45 estados con saltos de 5 estados. En la figura 5.3 se aprecia la variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 3:1 del tamaño original. La tabla 5.2 resume los resultados obtenidos en donde el mejor resultado es de  $67.88\% \pm 3.29$  para una configuración de 150 puntos de control y 40 estados.

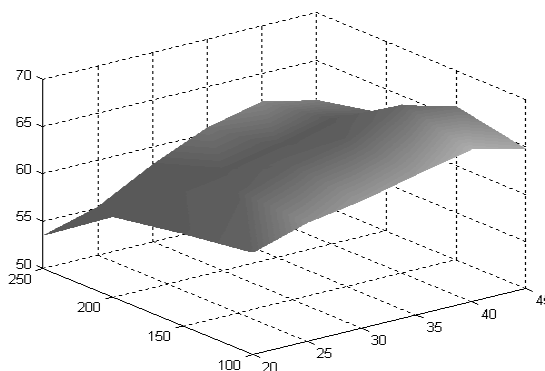


Figura 5.3: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 3:1 del tamaño original.

TABLA 5.2: TASA DE ACIERTO USANDO EL HMM COMO CLASIFICADOR CON 5 MUESTRAS DE ENTRENAMIENTO CON LAS IMÁGENES A UNA ESCALA 3:1

Número de estados	Número de puntos de control			
	100	150	175	200
20	60.83% ± 3.04	59.76% ± 9.32	58.51% ± 5.71	53.44% ± 2.77
25	62.65% ± 2.87	62.82% ± 3.67	59.76% ± 9.32	55.20% ± 3.30
30	63.52% ± 2.19	64.16% ± 3.13	60.73% ± 3.81	58.20% ± 4.28
35	64.88% ± 2.09	65.73% ± 3.81	61.88% ± 3.29	60.72% ± 3.51
40	66.16% ± 2.79	67.88% ± 3.29	63.28% ± 3.92	62.08% ± 3.76
45	64.80% ± 2.06	66.21% ± 3.29	62.45% ± 4.27	60.87% ± 3.89

Se repitieron los experimentos para una escala de reducción 4:1 de la imagen original, en este caso se consiguió un 73.60% ± 3.03 fijando a 175 puntos de control y con 40 estados. La figura 5.4 muestra el la variación del porcentaje de acierto que el sistema alcanzó con 5 muestras de entrenamiento, y la Tabla 5.3 muestra todos los resultados obtenidos con esta configuración.

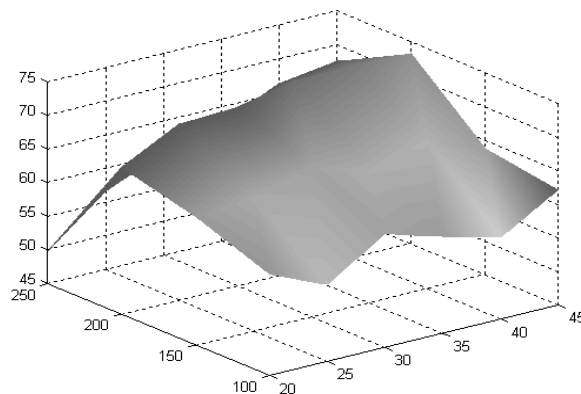


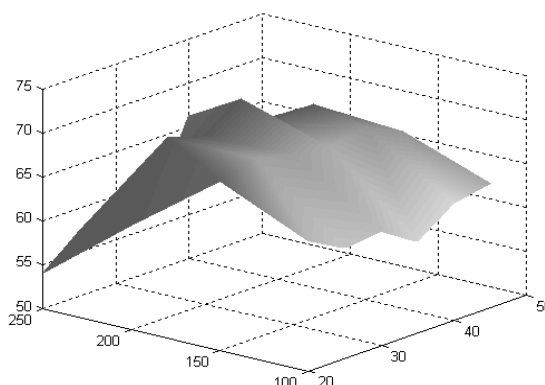
Figura 5.4: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 4:1 del tamaño original.

TABLA 5.3: TASA DE ACIERTO USANDO EL HMM COMO CLASIFICADOR CON 5 MUESTRAS DE ENTRENAMIENTO CON LAS IMÁGENES A UNA ESCALA 4:1

Número de estados	Número de puntos de control			
	100	150	175	200
20	60.08% ± 4.89	64.16% ± 1.46	66.88% ± 1.78	49.76% ± 5.74
25	56.40% ± 4.19	64.80% ± 2.42	71.20% ± 3.02	56.64% ± 3.52
30	61.92% ± 2.34	67.28% ± 3.01	71.68% ± 3.58	60.32% ± 1.48
35	59.36% ± 2.84	64.56% ± 2.88	73.52% ± 3.66	62.40% ± 1.90
40	57.12% ± 3.89	66.96% ± 1.51	73.60% ± 3.03	66.40% ± 4.03
45	62.16% ± 1.25	63.84% ± 2.18	73.38% ± 4.14	67.60% ± 2.53



Finalmente se llevó a cabo la experimentación con las imágenes reducidas a una escala 5:1 de la imagen original. En este caso se alcanzó un porcentaje de acierto de  $72.00\% \pm 3.76$  fijando el sistema a 175 puntos de control y con 35 estados. La figura 5.5 muestra el la variación del porcentaje de acierto del sistema con 5 muestras de entrenamiento. La tabla 5.4 muestra todos los resultados obtenidos en estos experimentos.



**Figura 5.5: Variación de la tasa de acierto usando el HMM como clasificador con 5 muestras de entrenamiento con una escala 5:1 del tamaño original.**

TABLA 5.4: TASA DE ACIERTO USANDO EL HMM COMO CLASIFICADOR CON 5 MUESTRAS DE ENTRENAMIENTO CON LAS IMÁGENES A UNA ESCALA 5:1

Números de estados	Números de puntos de control			
	100	150	175	200
20	64.80% ± 3.19	69.20% ± 1.27	61.92% ± 2.41	54.08% ± 4.35
25	62.56% ± 4.46	70.46% ± 1.42	70.48% ± 1.70	57.52% ± 5.58
30	63.04% ± 2.18	69.40% ± 1.59	69.20% ± 1.47	59.84 % ± 3.08
35	60.40% ± 2.32	68.80% ± 1.71	72.00% ± 3.76	57.28% ± 0.66
40	63.32% ± 3.04	69.68% ± 1.10	67.68% ± 1.96	66.32% ± 3.51
45	64.16% ± 2.71	67.76% ± 1.74	68.48% ± 2.67	63.20% ± 2.04

Como puede observarse, los mejores resultados se obtuvieron cuando las imágenes se redujeron a una escala de 4:1 de la imagen original, por lo tanto el segundo enfoque se aplicó con este tamaño de resolución y usando un kernel RBF en las SVMs. En este caso se fijó el número de estados a 40, y el número de puntos de control se varió entre 25, 50, 100 y 150. También se estudió el comportamiento del sistema variando el porcentaje de entrenamiento, concretamente entre el 20% y el 50% de las muestras, como puede verse en la tabla 5.5. Se observa que el porcentaje de acierto es muy alto, más del 99% en todos los casos. En relación al número de puntos de control y números

de muestras de entrenamiento, el mejor resultado se consigue con 100 puntos de control y 3 muestras de entrenamiento, lográndose en este caso un 100% de tasa de acierto.

TABLA 5.5: TASA DE ACIERTO USANDO SVM CON UN KERNEL RBF PARA DIFERENTES NÚMEROS DE MUESTRAS DE ENTRENAMIENTO CON UNA ESCALA 4:1 DEL TAMAÑO ORIGINAL

Número de puntos de control	Número de estados	Número de muestras de entrenamiento	Tasa de acierto	g
25	40	5	100% ± 0	$4 \times 10^{-6}$
25	40	4	100% ± 0	$4 \times 10^{-6}$
25	40	3	99.95% ± 0.12	$1 \times 10^{-6}$
25	40	2	99.80% ± 0.28	$1 \times 10^{-6}$
50	40	5	100% ± 0	$4 \times 10^{-6}$
50	40	4	100% ± 0	$4 \times 10^{-6}$
50	40	3	100% ± 0	$1 \times 10^{-6}$
50	40	2	99.83% ± 0.26	$2 \times 10^{-6}$
100	40	5	100% ± 0	$4 \times 10^{-6}$
100	40	4	100% ± 0	$4 \times 10^{-6}$
100	40	3	100% ± 0	$4 \times 10^{-6}$
100	40	2	99.92% ± 0.14	$4 \times 10^{-6}$
150	40	5	100% ± 0	$4 \times 10^{-6}$
150	40	4	100% ± 0	$4 \times 10^{-6}$
150	40	3	100% ± 0	$4 \times 10^{-6}$
150	40	2	99.87% ± 0.20	$3 \times 10^{-6}$

Se demuestra con este estudio que el sistema presenta un mejor rendimiento cuando se redimensionan las imágenes a un tamaño óptimo. Esto se debe a que en el proceso de redimensionado se elimina información redundante. Por el contrario al caso anterior, si se da una excesiva reducción de las imágenes, se elimina información discriminante necesaria para la correcta identificación. Un aspecto importante a destacar es que, una vez encontrado el tamaño óptimo de análisis, el sistema es más independiente del número de puntos de control, dependiendo solamente del número de estados y del porcentaje de acierto. Finalmente, el valor de la gamma en el kernel RBF es estable y no afecta el comportamiento del sistema.

### 5.3.5 Evaluación del sistema de parametrización de contornos con otros clasificadores aplicado a firmas manuscritas.

En este experimento las muestras se dividieron en 50% para el entrenamiento y el 50% para el proceso de test. Además, todos los experimentos se repitieron 5 veces.

Todos los experimentos se hicieron con firmas genuinas de nuestra base de datos. La detección de contorno ha sido implementada usando morfología matemática. Se ha rellenado toda la firma, y a continuación, la imagen resultante, ha sido sustraída con la imagen erosionada, obteniéndose así el contorno.

Se puede observar en la Tabla 5.6, el mejor clasificador para los mismos datos es HMM con la combinación de contornos. Por esta razón, los autores decidieron usar solamente el clasificador HMM, aunque los resultados con NN y HMM son muy similares. Por lo tanto, se han comparado estos resultados con nuestra propuesta (véase la Tabla 5.7). En conclusión, este método presenta una mejora en la tasa de acierto, usando solo la información de los ángulos, usando HMM como clasificador.

TABLA 5.6: TASAS DE ÉXITO EN LA BASE DE DATOS DE FIRMAS USANDO DIVERSOS CLASIFICADORES

Clasificador	Tipo de contorno	Media $\pm$ Desviación Estándar
KNN	Puntos del contorno más cercano	60.77% $\pm$ 11.78
	Puntos del contorno más lejano	60.90% $\pm$ 11.58
	Combinación	61.37% $\pm$ 12.04
NN	Puntos del contorno más cercano	76.77% $\pm$ 0.17
	Puntos del contorno más lejano	76.44% $\pm$ 0.17
	Combinación	78.24% $\pm$ 0.48
HMM	Puntos del contorno más cercano	75.23% $\pm$ 0.78
	Puntos del contorno más lejano	77.94% $\pm$ 0.61
	Combinación	79.18% $\pm$ 0.54

TABLA 5.7: TASA DE ÉXITO PARA EL MISMO CLASIFICADOR VARIANDO LA PARAMETRIZACIÓN DE LAS FIRMAS

Clasificador	Tipo de método	Media $\pm$ Desviación Estándar
HMM	Método clásico	79.18% $\pm$ 0.54
HMM	Método propuesto	84.64% $\pm$ 0.20

En trabajos futuros, se trabajará con firmas falsificadas y auténticas.

### 5.3.6 *Uso dos de bases de datos en la identificación de usuarios a partir de su firma manuscrita.*

Se han usado dos bases de datos públicas en los experimentos de identificación firmas manuscritas *off-line*, la base de datos MYCT-75 y la base de datos GPDS100. Los resultados se muestran en las Tablas 4.19 y 4.20, observándose que el modelado

para ambos casos coincide en números de puntos de control y números de estado del HMM. Los valores obtenidos son relativamente similares entre el 98% y el 99% de éxito. Esta ligera variación puede ser debida a la diferencia de número de usuarios y a la diferencia de la resolución de las imágenes. Aunque se observa que el sistema es estable a pesar de dichas diferencias. Esto muestra la estabilidad del sistema propuesto, para esta aplicación biométrica.

#### **5.4 Comparativa con el estado del arte.**

Resulta complicado realizar una comparativa de los sistemas propuestos con el estado del arte, debido a que la mayoría de las bases de datos son bases de datos propias. Por esa razón, se usaron algunas aplicaciones con bases de datos públicas, para poder realizar algunas comparaciones directas con otros métodos propuestos en el estado del arte.

A pesar de ello, las bases de datos propias han reflejado en las diversas publicaciones logradas, como los resultados alcanzados han sido competitivos con el estado del arte, igualando o mejorando el estado del arte, con base de datos similares, pero no iguales. En esta línea, cabe destacar los resultados alcanzados en las aplicaciones con granos de polen, biometría de los labios y en la identificación de especies de árboles a partir de sus hojas.

Centrados en las bases de datos, donde si se puede hacer una comparación directa, se va a analizar el caso para la identificación biométrica basada en el contorno de la mano. En este trabajo se han usado tres bases de datos públicas, y una comparación con el estado del arte se puede observar en la Tabla 5.8.

En la Tabla 5.8, se observa que la propuesta de este trabajo tiene la mejor respuesta frente al estado del arte para las bases de datos GPDS144 y GPDS100hand3Band. En Base de datos UST, el enfoque presenta la segunda mejor precisión. En este caso, el entrenamiento se realizó con 2 muestras por usuario, y en la referencia que tiene resultado ligeramente mejor [Zhenan et al., 2005], no indica las muestras de entrenamiento usadas. Por lo tanto y en líneas generales, este sistema propuesto es una buena y robusta opción para esta modalidad biométrica.

TABLA 5.8: COMPARATIVA DEL TRABAJO PROPUESTO PARA LA BIOMETRÍA DE LA MANO FRENTE AL ESTADO DEL ARTE

Base de datos	Referencia	Método	Éxito
UST	Este trabajo	DHMMK+SVM (2 samples for training)	EER=0.31%
	[Ramalho et al., 2012]	Fusión de 20 anchos de dedos (cuatro de cada dedo), cinco longitudes de los dedos, cinco dedos perímetros y cinco dentro de la palma de las distancias y las impresiones palmares con LDA	EER=0.43%
	[de-Santos-Sierra et al., 2011]	Ancho de dedos clasificados por la distancia euclídea	EER=1.40%
	[Zhenan et al., 2005]	Carácterísticas de códigos ordinales basada en distancia Gabor y Hamming	EER=0.22%
	[Rufeng et al., 2007]	Aprendizaje AdaBoost reducido mediante LDA	EER=0.35%
GPDS144	Este trabajo	DHMMK+SVM (2 muestras para entrenar)	EER=0.017%
	[Ramalho et al., 2012]	Fusión de 20 anchos de dedos (cuatro de cada dedo), cinco longitudes de los dedos, cinco dedos perímetros y cinco dentro de la palma de las distancias y las impresiones palmares con LDA	EER=0.047%
	[Mansoor et al., 2008]	Información de tectura para Palmprint usando la transformada Contourlet	EER=0.70%
	[Polat y Yildirim, 2008]	Geometría de las manos clasificada con red neural regresiva	93.33%
	[Aghili y Sadjedi, 2010]	Veinticuatro características de cuatro dedos usando distancia euclidiana	EER=0.17%
	[Fuertes et al., 2012]	80 anchos de 4 dedos cada dedo utilizando SVM-RBF	99.90%
GPDS100 hand3Band	Este trabajo	DHMMK+SVM (2 muestras para entrenar)	EER=0%
	[Morales et al., 2012]	100 anchos de cada dedo	EER=0.26%
	[Ferrer y Morales, 2011]	Anchos y longitudes de dedos y contorno de la mano usando LS-SVM	EER=0.13%

La otra aplicación donde se puede realizar una comparación directa, es para las bases de datos de firmas manuscritas *off-line*. En la siguiente tabla, se muestran para ambas bases de datos, algunas comparaciones;

TABLA 5.9: COMPARATIVA DEL TRABAJO PROPUESTO PARA LA BIOMETRÍA DE LA FIRMA FRENTE AL ESTADO DEL ARTE

Base de datos	Referencia	Éxito
MYCT-75	Este trabajo	Error=1.20%
	[Vargas, 2010]	EER=12.67%
	[Gilperez et al., 2008]	EER=10.18%
GPDS100	Este trabajo	Error=0.85%
	[Vargas, 2010]	EER=11.71%

Se observa en la tabla anterior, como los resultados alcanzados mejoran el estado del arte, aunque en este caso no se han mostrado un esquema de verificación, sino de identificación. La diferencia del error entre la propuesta y el estado del arte, presagia que este sistema será una muy buena opción para este tipo de biometría.



## CAPÍTULO 6

# Conclusiones y Líneas futuras

---

En este capítulo, se van a mostrar las conclusiones y líneas futuras de trabajo, generadas a partir del análisis y desarrollo de esta tesis doctoral.

### 6.1 Conclusiones

La presente tesis ha tratado de validar la siguiente hipótesis;

*“Es posible diseñar un sistema de identificación de formas a partir de diversos datos de su contorno mediante una transformación de los modelos ocultos de Markov como elemento parametrizador, con independencia de la variabilidad del contorno.”*

A raíz de los resultados obtenidos y analizado en el capítulo 5, se puede decir que se logrado satisfacer dicha hipótesis, así como los objetivos planteados para demostrar la hipótesis de esta tesis.

A continuación, se presentan una serie de conclusiones directamente relacionadas con cada una de las etapas seguidas en el desarrollo de la metodología presentada en este documento, después de haber sido usada para una base de datos sintética y en 8 aplicaciones reales usando 12 bases de datos diferentes:

1. Se ha comprobado que la metodología propuesta tiene un comportamiento estable para las diferentes formas bidimensionales con independencia de la complejidad de su forma.
2. El ajuste de los parámetros del sistema propuesto para llegar a un modelado con un ratio de identificación óptimo, presenta rangos de valores relativamente cercanos, con lo que su sintonización para encontrar el mejor porcentaje de éxito, se reduce a un ajuste fino de los parámetros, reduciendo el tiempo búsqueda de dicho modelado. A pesar de eso, la primera simulación dentro de los rangos establecidos dará buenos resultados de identificación, mejorándose con dicho ajuste fino. Además, los valores de éxito logrados entorno a los parámetros localizados para obtener la mejor tasa de éxito, son relativamente similares, mostrando una gran estabilidad entorno a ese valor óptimo de funcionamiento del modelo.
3. Si bien no es siempre posible plantear una comparativa real con sistema implementados por otros autores, depeden si la base de datos sea pública o no, en los casos que si se puede realizar una comparativa, los resultados han sido mejores o similares al estado del arte actual.
4. La combinación de los diversos métodos aplicados para realizar toda la metodología ha presentado un buen resultados, desde la elección de una caracterización angular de la forma y su parametrización mediante la transformación del modelo oculto de Markov, hasta su clasificación supervisada basada en las máquinas de soporte vectorial.
5. Se han probado otro esquema adicional a la metodología seguida bajo un esquema de identificación para el caso de la biometría del contorno de la mano, como ha sido la verificación, y los resultados encontrados superan y/o igual al estado del arte, validándose este tipo de uso con tres bases de datos públicas.



6. Se observa una robustez frente al ruido, como se ha podido observar en la base de datos de hojas y la sintética. También, esta robustez es validada en las posibles variabilidades que se puedan observar en las diversas muestras de una misma clase para el resto de las bases de datos, como es de forma más acentuada en las de firmas y en la dactilológica. Por tanto, la robustez frente al ruido y a variabilidades de la forma queda patente en la obtención de una buena tasa de éxito.
7. El uso de 8 aplicaciones reales y una sintética, en total 13 bases de datos, de las cuales algunas de ellas usan bases de datos públicas, muestran la validación del método para su uso generalizado en otro tipo de aplicaciones distintas a las estudiadas. Se considera que este número de bases de datos usadas, son adecuadas para aprobar la generalización de esta propuesta y su metodología.
8. Finalmente, se observa en la comparativa que este sistema, en la mayoría de las veces es más robusto que los planteamientos que se pueden encontrar en el estado del arte. Esto muestra el potencial de este método y su aplicabilidad en otros casos de formas.

## **6.2 Líneas futuras**

A continuación se plantean algunas líneas futuras que se desprenden del trabajo realizado en el desarrollo de la presente Tesis.

1. Estudiar el comportamiento del sistema frente a nuevos métodos de definición de contornos y observar su comportamiento y respuesta. Esto puede abrir la puerta a generar fusiones siempre que la definición de contornos estén incorreladas.
2. Utilización de otro tipo de clasificaciones que soporte un gran volumen de datos que pueda generar una función de verosimilitud incorrelada a la obtenida con las máquinas de soporte vectorial, para abrir la puerta a posibles fusiones de sistemas, y mejorar su eficiencia.
3. Integración del sistema en ambientes más complejos, como puede ser su uso en ambientes reales de trabajo y/o su uso en aplicaciones con vídeo, donde se deba de identificar de forma on-line, diversas formas bidimensionales.
4. Estudiar la resolución de las imágenes para ver su comportamiento y buscar su valor más adecuado para este sistema propuesto. En esta tesis se han usado varias resoluciones, desde 150ppp hasta 600ppp, pudiéndose comentar de forma muy

general, que la resolución es algo invariante al sistema, aunque esto se debe de estudiar en profundidad.

5. Utilización de programación paralela y/o uso de GPUs para tratar de acelerar la respuesta del sistema y poderlo generar como una aplicación abierta que se pueda ejecutar sobre cualquier plataforma.
6. Reducir la interacción ajuste del modelado y usuario, para tratar que el sistema sea lo más independiente a su manipulación. Se propone el uso de algoritmos genéticos o bien una búsqueda exhaustiva dependiente de la necesidad y urgencia en la que se quiera el modelo.
7. Estudiar y chequear sobre este modelo, el efecto de usar muestras falsificadas en aplicaciones biométricas, como puede ser en firmas manuscritas off-line y en moldes de manos que simulen el contorno.
8. Extender el uso del sistema propuesto y la experimentación a un esquema basado en la verificación, como se ha realizado un primer paso para las bases de datos de manos.

## **Bibliografía**

---

- [**Aghili y Sadjedi, 2010**] Aghili, B.; Sadjedi, H.; "Personal identification/verification by using four fingers", 2010 3rd International Congress on Image and Signal Processing, vol. 6, pp. 2619-2623, 2010
- [**Andelic, 2006**] Andelic, E.; Schaffoner, M.; Katz, M.; Kruger, S.E.; Wendemuth, A., "A Hybrid HMM-Based Speech Recognizer Using Kernel-Based Discriminants as Acoustic Models," 18th International Conference on Pattern Recognition (ICPR 2006), vol.2, pp.1158-1161, 2006
- [**Aran y Akarun, 2008**] Aran, O.; Akarun, L.; "Multi-class classification strategies for Fisher scores of gesture and sign sequences," 19th International Conference on Pattern Recognition (ICPR 2008), pp.1-4, 2008
- [**Arbter et al., 1990**] Arbter, K.; Snyder, W.E.; Burkhardt, H.; Hirzinger, G.; "Application of affine-invariant Fourier descriptors to recognition of 3-D objects", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.12, nº 7, pp. 640–647,1990

- [**Arbter, 1989**] Arbter, K.; “Affine-invariant Fourier descriptors”, in: J.C. Simon (Ed.), *From Pixels to Features*, Elsevier Science Publishers B.V. (North-Holland), Amsterdam, pp. 153–164, 1989
- [**Arlot, 2010**] Arlot, S.; “A survey of cross-validation procedures for model selection”, *Statistics Surveys*, vol. 4, pp. 40–79, 2010
- [**Asada y Brady, 1984**] Asada, H.; Brady, M.; “The curvature primal sketch”, MIT AI Memo 758, 1984.
- [**Asada y Brady, 1986**] Asada, H.; Brandy, M.; “The curvature primal sketch”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 8, n° 1, pp. 2–14, 1986
- [**Bai et al., 2010**] Xiang Bai; Xingwei Yang; Longin Jan Latecki; Wenyu Liu; Zhuowen Tu; “Learning Context Sensitive Shape Similarity by Graph Transduction”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, n° 5, pp. 861-874, 2010.
- [**Barretti et al., 2000**] Berretti, S., Bimbo, A.D., Pala, P., Retrieval by shape similarity with perceptual distance and effective indexing, *IEEE Trans. Multimedia* 2 (4), 225–239, 2000
- [**Briceño et al., 2002**] Briceño, J.C.; Travieso, C.M.; Ferrer, M.A.; “Automatic Recognition of Simple *Laurisilva canariensis* Leaves by Perimeter Characterization”, *IASTED International Conference on Signal Processing, Pattern Recognition, and Applications*, pp. 249–254, 2002
- [**Briceño et al., 2007**] Briceño, J.C.; Travieso, C.M.; Ferrer, M.A.; Alonso, J.B.; Briceño, R.D.; “A genus recognition system for the Costa Rica Lauraceae Family, using a Support Vector Machine”, *XXIII Conferencia Latinoamericana de Informática*, pp. 120, 2007
- [**Briceño et al., 2009**] Briceño, J.C.; Travieso, C.M.; Ferrer, M.A.; Alonso, J.B.; Vargas, F.; “Angular contour parameterization for signature identification”, *Computer Aided Systems Theory-EUROCAST 2009*, pp. 358-365, 2009
- [**Briceño et al., 2010**] Briceño, J.C.; Travieso, C.M.; Alonso, J.B.; Ferrer, M.A.; “Robust identification of persons by lips contour using shape transformation”, in *14th International Conference on Intelligent Engineering Systems (INES 2010)*, pp. 203-207, 2010
- [**Briceño et al., 2011**] Briceño, J.C.; Travieso, C.M.; Alonso, J.B.; C.M, Ferrer, M.A.; “Biometric Identification Based on Hand-Shape Features Using a HMM Kernel”, *2011 International Conference on Hand-Based Biometrics (ICHB 2011)*, pp. 1-6, 2011
- [**Brill, 1968**] Brill, E.L.; “Character recognition via Fourier descriptors”, *WESCON*, Session 25, *Qualitative Pattern Recognition Through Image Shaping*, Los Angeles, CA, pp. 1–10, (25/3), 1968

- [**Burges, 1998**] Burges, C.J.C.; “A Tutorial on Support Vector Machines for Pattern Recognition”, *Data Mining and Knowledge Discovery*, vol. 2, nº 2, pp. 121-167, 1998
- [**Campbell, 2003**] Campbell, W.M.; "A SVM/HMM system for speaker recognition," in *Proceedings of the 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '03)*, vol.2, pp. 209-212, 2003
- [**Cervantes et al., 2007a**] Cervantes, J.; Xiaou Li; Wen Yu; “Support Vector classification for large data sets by reducing training data with change of classes”, *IEEE International Conference on Systems, Man and Cybernetics*, pp. 2609 – 2614, 2008.
- [**Cervantes et al., 2007b**] Cervantes, J.; Xiaou Li; Wen Yu; Bejarano, J; “Multi-Class Support Vector Machines for Large Data Sets via Minimum Enclosing Ball Clustering”, *4th International Conference on Electrical and Electronics Engineering*, pp. 146 – 149, 2007
- [**Chellappa y Bagdazian**] Chellappa, R.; Bagdazian, R.; “Fourier coding of image boundaries”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 6, nº 1, pp. 102–105, 1984
- [**Cherkassky y Mulier, 2007**] Cherrassky, Vladimir and Filip Mulier, *Learning from Data: Concepts, Theory and Methods*, Ed. John Wiley & Sons, 2007
- [**Chetverikov y Khenokh, 1999**] Chetverikov, D.; Khenokh, Y.; “Matching for shape defect detection”, *Lecture Notes in computer Science*, vol. 1689, Springer, Berlin, pp. 367–374, 1999
- [**CINAT, 2012**] Centro de Investigaciones Apícolas Tropicales (CINAT9 de la Universidad Nacional de Costa Rica (UNA) <http://www.cinat.una.ac.cr/> (última visita: 9 de noviembre de 2012).
- [**Das y Paulink, 1990**] Das, M.; Paulik, M.J.; Loh, N.K.; “A bivariate autoregressive modeling technique for analysis and classification of planar shapes”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 12, nº 1, pp. 97–103, 1990
- [**Davies, 1997**] Davies, E.R.; “Machine Vision: Theory, Algorithms, Practicalities”, Academic Press, New York, pp. 171–191, 1997
- [**de-Santos-Sierra et al., 2011**] de-Santos-Sierra, A.; Sánchez-Ávila, C.; del Pozo, G.B.; “Unconstrained and Contactless Hand Geometry Biometrics”, *Sensors*, vol. 11, pp. 10143-10164, 2011
- [**Dubois y Glanz, 1986**] Dubois, S.R; Glanz, F.H.; “An autoregressive model approach to two-dimensional shape classification”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 627–637, 1986
- [**Dudek y Tsotsos, 1997**] Dudek, G.; Tsotsos, J.K.; “Shape representation and recognition from multiscale curvature”, *Computer Vision and Image Understanding*, vol. 68, nº 2, pp. 170–189, 1997

- [**Eom y Park, b 1990**] EOM, K.; Park, J.; “Recognition of shape by statistical modeling of centroidal profile”, Proceedings of the Tenth International Conference on Pattern Recognition, vol. 1, pp. 860–864, 1990
- [**Ferrer et. al, 2007**] Ferrer, M.A.; Morales, A.; Travieso, C.M.; Alonso, J.B.; “Low Cost Multimodal Biometric Identification System based on Hand Geometry, Palm and Finger Textures”, in 41st Annual IEEE International Carnahan Conference on Security Technology, pp. 52-58, 2007.
- [**Ferrer y Morales, 2011**] Ferrer, M.A.; Morales, A.; “Hand Shape Biometrics combining the visible and Short Wave In fraReds Bands”, in IEEE Transactions on Information Forensics & Security, vol. 6, n° 4, pp. 1305 – 1314, 2011.
- [**Freeman, 1961**] Freeman, H.; “On the encoding of arbitrary geometric configurations”, IRE Transaction on Electronic Computers vol. EC-10, pp. 260–268, 1961
- [**Fierrez-Aguilar et al., 2004**] Fierrez-Aguilar, J.; Alonso-Hermira, N.; Moreno-Marquez, G.; Ortega-Garcia J.; “An Off-line Signature Verification System. Based on Fusion of Local and Global Information”, D. Maltoni and A.K. Jain (Eds.): BioAW 2004, Springer-Verlag, LNCS, vol. 3087, pp. 295–306, 2004.
- [**Fu, 1974**] Fu, K.S.; “Syntactic Methods in Pattern Recognition”, Academic Press, New York, 1974.
- [**Fuertes et al., 2012**] Fuertes, J.J.; Travieso, C.M.; Naranjo, V.; “2-D Discrete Wavelet Transform for Hand Palm Texture Biometric Identification and Verification”, in book Wavelet Transforms and Their Recent Applications in Biology and Geoscience, (Ed. Dumitru Baleanu) pp. 103-118, 2012
- [**Galar et al., 2011**] Galar, M.; Fernández, V.A.; Barrenechea, E.; Bustince, H.; Herrera, F.; “An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes”, Pattern Recognition, vol. 44, n° 8, pp. 1761–1776, 2011.
- [**Gilperez et al., 2008**] Gilperez, A.; Alonso-Fernandez, F.; Pecharroman, S.; Fierrez, J.; Ortega-García, J.; “Off-line Signature Verification Using Contour Features”, in Proceedings of the International Conference on Frontiers in Handwriting Recognition, 2008
- [**Groskey y Mehrotra, 1990**] Groskey, W.I.; Mehrotra, R.; “Index-based object recognition in pictorial data management”, Computer Vision, Graphics, and Image Processing, vol. 52, n° 3, pp. 416–436, 1990
- [**Groskey, 1992**] Groskey, W.I.; Neo, P.; Mehrotra, R.; “A pictorial index mechanism for model-based matching”, Data & Knowledge Engineering, vol. 8, pp. 309–327, 1992
- [**Hamdi et al., 2010**] Hamdi, A.; Missaoui, O.; Frigui, H.; "AN SVM classifier with HMM-based kernel for landmine detection using ground penetrating radar," 2010 IEEE International on Geoscience and Remote Sensing Symposium (IGARSS), pp. 4196-4199, 2010

- [**Haussler et al., 2000**] Haussler, D.; Diekhans, M.; Jaakkola, T.; “A Discriminative Framework for Detecting Remote Protein Homologies”, *Journal of Computational Biology*, vol. 7, nº 1-2, pp. 95-114, 2000
- [**He y Kundu, 1991**] He, Y.; Kundu, A.; “2-D shape classification using hidden Markov model”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 13, nº 11, pp. 1172–1184, 1991
- [**Jaakkola et al., 1999**] Jaakkola, T.; Diekhans, M.; Haussler, D.; “A Discriminative Framework for Detecting Remote Protein Homologies”, in *Proceedings of the Seventh International Conference on Intelligent Systems for Molecular Biology*, pp. 149-158, 1999
- [**Jaakkola et al., 2000**] Jaakkola, T., Diekhans, M., Haussler, D., A Discriminative Framework for Detecting Remote Protein Homologies, *Journal of Computational Biology - JCB* , vol. 7, no. 1-2, pp. 95-114, 2000
- [**Jin y Mokhtarian, 2006**] Ning Jin; Mokhtarian, F.; "A Non-Parametric HMM Learning Method for Shape Dynamics with Application to Human Motion Recognition," 18th International Conference on Pattern Recognition (ICPR 2006), vol.2, pp.29-32, 2006
- [**Joachims, 2008**] SVM<sup>light</sup> Support Vector Machine. <http://svmlight.joachims.org/>. (Última visita: Noviembre de 2012)
- [**Kashyap y Chellappa**] Kashyap, K.L.; Chellappa, R.; “Stochastic models for closed boundary analysis: representation and reconstruction”, *IEEE Transaction on Information Theory*, vol. 27, pp. 627–637, 1981
- [**Kumar et. al, 2003**] Kumar, A.; Wong, D.C.M.; Shen, H.C.; Jain, A.K.; “Personal Verification Using Palmprint and Hand Geometry Biometrics”, in *Proceedings of the 4th International Conference on Audio and Video Based Biometric Person Authentication*, J. Kittler and M.S. Nixon (Eds.): AVBPA 2003, LNCS 2688, Springer-Verlag, pp. 668-678, 2003.
- [**Mansoor et al., 2008**] Mansoor, A.B.; Mumtaz, M.; Masood, H.; Butt, M.A.; Shoab, Khan, A.; “Personal Identification Using Palmprint and Contourlet Transform”, *Advances in Visual Computing, Lecture Notes in Computer Science*, vol. 5359, pp. 521-530, 2008
- [**Marine, 2000**] Marine, F.J.S.; “Automatic recognition of biological shapes with and without representation of shape”, *Artificial Intellegence in Medicine*, vol. 18, pp. 173–186, 2000
- [**Mokhtarian y Mackworth, 1986**] Mokhtarian, F.; Mackworth, A.; “Scale-based description and recognition of planar curves and two-dimensional shapes”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 8, nº 1, pp. 34–43, 1986
- [**Morales et al., 2012**] Morales, A.; González, E.; Ferrer, M.A.; “On the Feasibility of Interoperable Schemes in Hand Biometrics”, *Sensors*, vol. 12, nº 2, pp. 1352-1382, 2012

- [**MPEG, 2001**] MPEG, MPEG-7 Interoperability, Conformance Testing and profiling, version 2, ISO / IEC JTC1 / SC29 / WG11 / N4039, March 2001
- [**Nieves y Dominguez, 2010**] Nieves, A.; Dominguez, F.C.; “Probabilidad y estadística para ingeniería: un enfoque moderno”, Ed. McGrawHill, 2010
- [**Ohm et al., 2000**] Ohm, J.R.; Bunjamin, F.B. ; Liebsch, W.; Makai, B.; Muller, K.,; Somlic, A.; Zier, D.; “A set of visual feature descriptors and their combination in a low-level description scheme”, Signal Processing: Image Communication, vol. 16, pp. 157–179, 2000
- [**Otsu, 1979**] Otsu, N.; “A threshold selection method from gray level histograms”, IEEE Transaction on Systems Man & Cybernetics, vol. 9, pp. 62–66, 1979
- [**Pavlidis, 1982**] Pavlidis, T.; “Algorithms for Graphics and Image processing”, Computer Science Press, Rockville, MD, pp.143, 1982
- [**Peura y Livarinen, 1997**] Peura, M.; Iivarinen, J.; “Efficiency of simple shape descriptors”, Proceedings of the Third International Workshop on Visual Form, pp. 443–451, 1997
- [**Polat y Yildirim, 2008**] Polat, Ö.; Yildirim, T.; “Hand geometry identification without feature extraction by general regression neural network”, Expert Systems with Applications, vol. 34, n° 2, pp. 845-849, 2008
- [**Rabiner y Juang, 1993**] Rabiner, L.; Juang B.H.; “Fundamentals of Speech Recognition”. Ed. Prentice Hall, 1993.
- [**Rabiner, 1989**] Rabiner, L.; “A tutorial on Hidden Markov models and Selected Applications in Speech Recognition”. Proceedings of the IEEE, vol. 77, n° 2, pp. 257-286, 1989.
- [**Ramalho et al., 2012**] Ramalho, M.B.; Correia, P.L.; Soares, L.D.; “Hand-based multimodal identification system with secure biometric template storage”, IET Computer Vision, vol. 6, n° 3, pp. 165-173, 2012
- [**Ross, 2007**] Sheldon M. Ross; “Introduction to Probability Models”, Academic Press, 2007
- [**Rucklidge, 1997**] Rucklidge, W.J.; Efficient locating objects using Hausdorff distance, International Journal on Computer Vision, vol. 24, n° 3, pp. 251–270, 1997
- [**Rufeng et al., 2007**] Rufeng, C.; Zhen, L.; Yufei, H.; Ran, H.; Stan, Z.L.; “Learning Gabor Magnitude Features for Palmprint Recognition”, in Computer Vision – ACCV 2007, Lecture Notes in Computer Science, vol. 4844, pp. 22-31, 2007
- [**Sanguansat et al., 2004**] Sanguansat, P.; Asdornwised, W.; Jitapunkul, S.; "Online Thai handwritten character recognition using hidden Markov models and support vector machines," IEEE International Symposium on Communications and Information Technology (ISCIT 2004), vol.1, pp. 492- 497, 2004



- [**Sekita et al., 1992**] Sekita, I., Kurita, T., N. Otsu, Complex autoregressive model for shape recognition, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 14, pp. 489–496, 1992
- [**Shawe-Taylor y Cristianini, 2000**] John Shawe-Taylor; Nello Cristianini; “Support Vector Machines and other kernel-based learning methods”, Cambridge University Press, 2000.
- [**Sonka et al., 1993**] Sonka, M., Hlavac, V., Boyle, R., Image Processing, Analysis and Machine Vision, Chapman & Hall, London, UK, NJ, , pp. 193–242, 1993
- [**Steinwart y Christmann, 2008**] Steinwart, I.; Christmann, A.; “Support Vector Machines”. Ed. Springer, 2008.
- [**Thanh-Nghi y Fekete, 2007**] Thanh-Nghi Do; Fekete, J.D.; “Large Scale Classification with Support Vector Machine Algorithms”, Sixth International Conference on Machine Learning and Applications, pp. 7 – 12, 2007
- [**Tieng y Boles, 1997**] Tieng, Q.M., Boles, W.W., Recognition of 2D object contours using the wavelet transform zero-crossing representation, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 19, n° 8, pp. 910–916, 1997
- [**Travieso et al., 2003**] Travieso, C.M.; Alonso, J.B.; Ferrer, M.A.; , "Sign language to text by SVM", Proceedings of the Seventh International Symposium Signal Processing and Its Applications, vol.2, pp. 435- 438, 2003
- [**Travieso et al., 2007**] Carlos M. Travieso; Juan C. Briceño; Miguel A. Ferrer; Jesús B. Alonso; “Using Fisher Kernel on 2D-Shape Identification”, LNCS (Computer Aided Systems Theory – EUROCAST 2007) (Eds. R. Moreno-Diaz et al.), Volume 4739/2007, pp. 740-746, 2007
- [**Travieso et al., 2011**] Travieso, C.M.; Briceño, J.C.; Ticay-Rivas, J.R.; Alonso, J.B.; “Pollen Classification Based on Contour Features.” 15th IEEE International Conference on Intelligent Engineering Systems, pp.5-7, 2011
- [**Van Otterloo, 1991**] Van Otterloo, P.J.; “A Contour-Oriented Approach to Shape Analysis”, Prentice-Hall International (UK) Ltd, Englewood Cliffs, NJ, pp. 90–108, 1991
- [**Vapnik, 1995**] Vladimir Vapnik. “The Nature of Statistical Learning Theory”. Springer-Verlag, 1995.
- [**Vargas, 2010**] Vargas-Bonilla, F.; “Metodología para el uso de los niveles de Gris en verificación de firmas manuscritas”, Tesis Doctoral, Universidad de Las Palmas de Gran Canaria, Noviembre 2010.
- [**Wang et al., 2002**] Tianshu Wang; Nan-Ning Zheng; Yan Li; Ying-Qing Xu; Heung-Yeung Shum; "Learning Kernel-based HMMs for dynamic sequence synthesis," in Proceedings of the 10th Pacific Conference on Computer Graphics and Applications, pp. 87- 95, 2002

- [**Wang y Liew, 2002**] Wang, S.L.; Liew, A.W.C.; “Physiological and behavioral lip biometrics: A comprehensive study of their discriminative power”, *Pattern Recognition*, vol. 45, pp. 3328–3335, 2012
- [**Xiaoou et al., 2010**]: Xiaoou Li; Cervantes, J.; Wen Yu; “A Novel SVM Classification Method for Large Data Sets”, 2010 IEEE International Conference on Granular Computing, pp. 297 – 302, 2010.
- [**Yang et al., 1998**] Yang, H.S., Lee, S.U., Lee, K.M., Recognition of 2D object contours using starting-point-independent wavelet coefficient matching, *Journal on Visual Communication Image Represent*, vol. 9, nº 2, pp. 171–181, 1998
- [**Young et al., 1974**] Yong, I.; Walker, J.; Bowie, J.; “An analysis technique for biological shape”, *Comput. Graphics Image Processing*, vol. 25, pp. 357–370, 1974.
- [**Zhang y Lu, 2001**] Zhang, D.S.; Lu, G.; “A comparison of shape retrieval using Fourier descriptors and short-time Fourier descriptors”, in *Proceedings of the Second IEEE Pacific-Rim Conference on Multimedia (PCM01)*, pp. 855–860, 2001
- [**Zhang y Lu, 2002**] Zhang, D.S.; Lu, G.; “A comparative study of Fourier descriptors for shape representation and retrieval”, in *Proceedings of the Fifth Asian Conference on Computer Vision (ACCV02)*, pp. 646–651, 2002
- [**Zhenan et al., 2005**] Zhenan, S.; Tieniu, T.; Yunhong, W.; Li, S.Z.; “Ordinal palmprint representation for personal identification”, in: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 279-284, 2005





