



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

TESIS

**DISEÑO DE UN SISTEMA DE RECONOCIMIENTO
FACIAL USANDO MODELOS OCULTOS DE MARKOV**

**QUE PARA OBTENER EL TÍTULO DE
INGENIERO EN COMPUTACIÓN**

PRESENTA:

MIQUEAS ESLI ALDAMA SÁNCHEZ

**DIRECTOR DE TESIS
M. I. JAIME ALFONSO REYES CORTÉS**

CIUDAD UNIVERSITARIA

Septiembre, 2013





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Agradecimientos

Agradezco a:

- Mi familia, por su apoyo en todo sentido y, además, por sus consejos a lo largo de todo este tiempo.
- Mi tutor de tesis, por su tiempo, sus sugerencias y su compromiso con este trabajo.
- Los sinodales asignados para la revisión de este trabajo, por sus observaciones y su tiempo.
- Los profesores de la Facultad de Ingeniería de la Universidad Nacional Autónoma de México, por sus enseñanzas, sus consejos y su disposición.
- La Universidad Nacional Autónoma de México, por su acogimiento durante el curso de la carrera.

Contenido

Agradecimientos	3
Contenido.....	5
Capítulo 1. Introducción	7
1.1 Objetivos	7
1.2 Alcances.....	8
1.3 Organización	8
Capítulo 2. Marco teórico de los HMMs.....	11
2.1. Conceptos generales	11
2.1.1. <i>Modelo de Markov</i>	11
2.1.2. <i>HMM</i>	12
2.1.3. <i>Elementos de un HMM</i>	13
2.2. Tipos de HMM.....	13
2.2.1. <i>Modelo Ergódico</i>	13
2.2.2. <i>Modelo Bakis</i>	14
2.3. Problemas principales en el uso de HMMs	15
2.3.1. <i>Evaluación</i>	15
2.3.2. <i>Decodificación</i>	17
2.3.3 <i>Aprendizaje</i>	17
2.4. Consideraciones adicionales en la implementación de HMMs	20
2.4.1 <i>Escalamiento</i>	20
2.4.2 <i>Uso de múltiples secuencias de observación</i>	20
Capítulo 3. Visión general del Reconocimiento de Rostros.....	23
3.1. Descripción general del Reconocimiento de Rostros	23
3.2. Etapas de procesamiento de un SRR	24
3.2.1. <i>Detección del rostro</i>	24
3.2.2. <i>Normalización</i>	24
3.2.3. <i>Extracción de características</i>	25
3.2.4. <i>Reconocimiento</i>	25
3.3. Descripción general de algunas técnicas empleadas en Reconocimiento de Rostros	
25	
3.3.1. <i>PCA y Eigenfaces</i>	25
3.3.2. <i>LDA</i>	26
3.3.3 <i>Modelos 3D</i>	27
3.3.4. <i>HMMs y Wavelets en Reconocimiento de Rostros</i>	27
3.4. Principales dificultades en el Reconocimiento de Rostros	28
3.4.1 <i>Cambios de iluminación</i>	28
3.4.2 <i>Cambios en la posición del rostro</i>	29
3.4.3 <i>Expresiones faciales</i>	29
3.4.4 <i>Edad</i>	29
3.4.5 <i>Accesorios y rasgos faciales</i>	30
3.4.6 <i>Cooperación del usuario</i>	30
3.5. Aplicaciones	30
3.5.1 <i>Seguridad</i>	30
3.5.2 <i>Control de acceso</i>	31
3.5.3 <i>Asistencia en departamentos de policía</i>	31
3.5.4 <i>Interacción humano-computadora</i>	31

Capítulo 4. Descripción del SRRP	33
4.1 Etapa de Entrenamiento	33
4.1.1 Definición de parámetros	34
4.1.2 Extracción de características	34
4.1.3 Obtención de secuencias de observación de entrenamiento	36
4.1.4 Obtención de HMMs entrenados	37
4.2 Etapa de Reconocimiento	38
4.2.1 Extracción de Características de la Imagen de Prueba	38
4.2.2 Obtención de la Secuencia de Observación de Prueba	38
4.2.3 Comparación de HMMs.....	39
4.3 Descripción de la interfaz gráfica del SRRP	39
Capítulo 5. Pruebas y Resultados.....	45
5.1 Resultados de las pruebas de reconocimiento hechas con el SRRP por tipo de subimagen.....	45
5.1.1 Resultados de las pruebas de reconocimiento para el tipo de subimagen 1 ...	45
5.1.2 Resultados de las pruebas de reconocimiento para el tipo de subimagen 2 ...	49
5.1.3 Resultados de las pruebas de reconocimiento para el tipo de subimagen 3 ...	52
5.1.4 Resultados de las pruebas de reconocimiento para el tipo de subimagen 4 ...	55
5.1.5 Resultados de las pruebas de reconocimiento para el tipo de subimagen 5 ...	59
5.1.6 Resultados de las pruebas de reconocimiento para el tipo de subimagen 6 ...	62
5.1.7 Resultados de las pruebas de reconocimiento para el tipo de subimagen 7 ...	65
5.2 Análisis general de las pruebas de reconocimiento hechas con el SRRP	69
5.3 Análisis de los porcentajes de reconocimiento global más altos.....	69
Capítulo 6. Conclusiones, mejoras y posibles aplicaciones	77
6.1 Conclusiones.....	77
6.2 Mejoras	78
6.3 Posibles aplicaciones.....	79
Referencias	81
Apéndice A. Descripción de la BDR ORL.....	85
Apéndice B. Cuantización Vectorial	87
B.1 Conceptos generales	87
B.2 Algoritmo de agrupamiento k-medias	88

Capítulo 1. Introducción

Los seres humanos logran identificarse entre sí de muchas formas, por ejemplo, a través del habla o de características corporales. En particular, la forma más común en la que podemos reconocer a personas que nos son familiares es a través de su rostro. Dicha habilidad de reconocimiento ha motivado a investigadores, no sólo de Ciencias de la Computación, sino de disciplinas como Neurociencia y Psicología [14] a enfocarse en un campo de estudio conocido como Reconocimiento de Rostros, el cual ha adquirido gran auge e importancia en las últimas décadas.

Con los avances tecnológicos en el área de la Computación, la interacción entre el humano y un sistema automático se ha hecho cada vez más viable y común, dando lugar a las que se denominan “interacciones inteligentes” [9]. Actualmente, la colaboración científica en el Reconocimiento de Rostros ha llevado a desarrollar métodos y herramientas de procesamiento y análisis en esta dirección, permitiendo la creación de Sistemas de Reconocimiento de Rostros (SRRs). El interés por estos sistemas se ha centrado en aquellas aplicaciones orientadas a la seguridad, la vigilancia y al control de acceso a lugares específicos. En muchas ocasiones su utilización puede ser de gran ayuda, por ejemplo en agencias de seguridad, para búsqueda e identificación de personas, etc.

Muchos SRRs se valen de bases de datos especiales, las cuales contienen imágenes de rostros de las personas que se pretenden reconocer de forma automática. En general, una base de datos de este estilo, que en adelante denominaremos como Base de Datos de Rostros (BDR), almacena varias imágenes del rostro de cada persona a reconocer, con el fin de dar mayor robustez al SSR al cual pertenece. Sin embargo, al trabajar con imágenes, uno de los problemas más frecuentes es el de la gran cantidad de datos que se deben manejar, por lo que se requiere de técnicas de reducción de datos y de extracción óptima de información, como la Cuantización Vectorial. Otra de las técnicas matemáticas que se ha utilizado para resolver este tipo de problemas es la Transformada Wavelet, la cual es relativamente reciente. Su versatilidad ha hecho posible su uso en muchas aplicaciones como procesamiento de señales, análisis de imágenes, sistemas de comunicación, entre otras [36], por lo cual se ha elegido estudiar algunos de sus conceptos, en cuestiones de implementación, para el presente trabajo. En relación a este último punto, en particular, se ha decidido analizar la aplicación de Filtros de Descomposición.

Al tratar con un problema como el del Reconocimiento de Rostros, se debe tomar en cuenta también la presencia de muchos factores que afectan la información que se pueda obtener para la construcción de un SRR. Lo anterior produce una gran variabilidad en los datos de análisis, es decir, en las imágenes de una BDR, por lo cual se necesita de un enfoque que, al menos, logre identificar qué datos son los más probables en cuanto a la representación de un determinado sujeto. Con respecto a este último planteamiento, en este trabajo se consideran los modelos estadísticos conocidos como Modelos Ocultos de Markov (HMMs), los cuales se han empleado en otros campos, como en reconocimiento de voz [25]. En el contexto del Reconocimiento de Rostros, los HMMs han sido probados desde las últimas dos décadas, aproximadamente, y han dado excelentes resultados, como en [4], [17], [18].

El diseño del SRR que aquí se propone, en adelante: Sistema de Reconocimiento de Rostros Propuesto (SRRP), se considera como una herramienta que está orientada a la identificación de rostros de sujetos, cuyas imágenes provienen de una BDR.

1.1 Objetivos

A partir de lo mencionado en los párrafos anteriores, se pueden definir a continuación los objetivos de este trabajo:

En primer lugar, se busca diseñar un sistema que sea capaz de efectuar el reconocimiento de rostros en imágenes provenientes de una BDR, con la cual se tendrá un mayor control en el momento de evaluar su funcionamiento y se podrán obtener estadísticas relevantes sobre su nivel de eficiencia. Esto último depende del buen desempeño de los algoritmos desarrollados para el SRRP.

En segundo lugar, se pretende desarrollar una metodología para el diseño de un SRR con las características antes mencionadas, con la cual se pueda contribuir al diseño y la implementación de algoritmos que son aplicables no sólo a SRRs, sino a sistemas que tratan de resolver otros problemas de reconocimiento de patrones, en general.

Finalmente, un objetivo más es el de permitir la extensión de este trabajo para el diseño de un SRR que se pueda usar en vídeo, a través de proponer teóricamente una forma de usar los HMMs para analizar el conjunto de imágenes que se pueden extraer de archivos de vídeo.

1.2 Alcances

El SRRP desarrollado en este trabajo está pensado principalmente para Reconocimiento y no para Detección de rostros, debido a que se cuenta con una BDR, cuyas imágenes comparten características comunes como formato, tamaño, forma de almacenamiento, entre otras. En particular, para la BDR Olivetti Research Laboratory (ORL) que se maneja en este trabajo, el Apéndice A describe con más detalle dichas características. Por tal motivo, no se contemplan algoritmos para Detección de rostros en imágenes, al menos para esta primera aproximación de un SRR.

Aunque el SRRP se limita a trabajar sólo con imágenes, en este caso, de la BDR mencionada en el párrafo anterior, se reconocen y plantean las extensiones que pueden realizarse a las etapas que componen dicho sistema para permitir trabajar con secuencias de imágenes provenientes de archivos de vídeo.

1.3 Organización

El presente trabajo está dividido en 6 capítulos y una sección que conjunta las referencias utilizadas, además de 2 apéndices en los que se encuentra información complementaria. A continuación se mencionan los temas que se abordan en cada uno de los capítulos y apéndices que conforman este trabajo.

La teoría básica sobre HMMs se presenta en el Capítulo 2, incluyendo algunos conceptos generales, los tipos de HMMs que se utilizan con mayor frecuencia, y los algoritmos necesarios para su implementación. También se abordan algunas consideraciones que deben hacerse con respecto a este último punto.

En el Capítulo 3 se hace una introducción al campo del Reconocimiento de Rostros, abarcando aspectos como las etapas generales de procesamiento de un SRR, técnicas relevantes usadas en SRRs y técnicas que hacen uso de HMMs en este sentido. También se tratan los principales problemas que surgen en esta área, así como sus aplicaciones más comunes.

La descripción del SRRP se concentra en el Capítulo 4, en el cual se explican las etapas que constituyen dicho sistema. La parte final del capítulo contiene una breve documentación de la interfaz gráfica desarrollada para el SRRP.

Las pruebas de reconocimiento hechas con el SRRP, junto con sus respectivos resultados, se incluyen en el Capítulo 5, las cuales se agrupan de acuerdo a los parámetros que se usaron para entrenar al sistema. Se analizan, en términos generales, los resultados de estas pruebas, discutiendo el impacto de los parámetros mencionados anteriormente sobre el desempeño del sistema.

El Capítulo 6 resume las conclusiones a las que se llegó con el desarrollo de este trabajo. En esta parte se indican las características más importantes del SRRP, en cuanto a su desempeño y sus posibles aplicaciones. Se plantean las mejoras que se pueden hacer a este sistema, además de las alternativas futuras sobre su extensión y enfoque.

El Apéndice A explica las características principales de la BDR a la cual se recurre en este trabajo. Por último, en el Apéndice B se abordan los conceptos generales de la Cuantización Vectorial ocupados en el SRRP.

Capítulo 2. Marco teórico de los HMMs

2.1. Conceptos generales

Para definir qué es un Modelo Oculto de Markov (HMM), primero se definirá lo que es un Modelo de Markov y los elementos que lo constituyen. Después de esto se podrá dedicar el resto del capítulo a analizar las características más relevantes de los HMMs, así como los algoritmos que se requieren para su implementación.

2.1.1. Modelo de Markov

Un Modelo de Markov es un proceso estocástico [25] en el cual, la variable aleatoria está definida sobre un conjunto conocido como *espacio de estados* [34]. Denotando dicho conjunto como S y la variable aleatoria como q , el valor de esta última representa uno de los posibles estados de S . En tiempo discreto, se suele aplicar el término *Cadena de Markov* a este proceso que, además, cumple una propiedad conocida como *Propiedad de Markov* [33]. Para un proceso de orden 1, ésta se define de la siguiente manera:

Propiedad de Markov. Dado el espacio de estados S y la secuencia de estados $\{q_1 = S_i, q_2 = S_j, \dots, q_t = S_k\}$, la probabilidad de que el estado presente (instante de tiempo $t+1$) sea el estado m depende sólo de dicho estado y del anterior. Esto se expresa como:

$$P(q_{t+1} = S_m \mid q_1 = S_i, q_2 = S_j, \dots, q_t = S_k) = P(q_{t+1} = S_m \mid q_t = S_k). \quad (2.1)$$

Los Modelos de Markov son usados para representar un sistema, del que se pueden observar ciertos eventos, a través del tiempo. A cada uno de estos eventos se les asocia un estado y se asume que el sistema modelado estará en uno de dichos estados en un instante de tiempo determinado. Para cambiar de un estado a otro, se definen probabilidades conocidas como *probabilidades de transición*. Teniendo presente lo anterior, una Cadena de Markov de orden 1 esta definida formalmente por los siguientes elementos:

S : Espacio de estados, con N elementos:

$$S = \{S_1, \dots, S_N\}. \quad (2.2)$$

$\boldsymbol{\pi}$: Probabilidades de estado inicial. $\boldsymbol{\pi} = [\pi_1, \dots, \pi_N]^T$;

$$\pi_i = P(q_1 = S_i). \quad (2.3)$$

Además, se satisface la restricción:

$$\sum_{i=1}^N \pi_i = 1; \quad (2.4)$$

\mathbf{A} : Matriz de transición. $\mathbf{A} = [a_{i,j}]$; $i, j = 1, \dots, N$;

$$a_{i,j} = P(q_t = S_j \mid q_{t-1} = S_i). \quad (2.5)$$

Los elementos de \mathbf{A} son las probabilidades de transición antes mencionadas, las cuales satisfacen la ecuación (2.6):

$$\sum_{j=1}^N a_{i,j} = 1; i = 1, \dots, N \quad (2.6)$$

A continuación se muestra un ejemplo sobre el uso de Cadenas de Markov para esclarecer los conceptos anteriores.

Suponiendo que se tiene un sistema, el cual puede exhibir 2 distintos comportamientos: funcionamiento normal o falla, se define una Cadena de Markov μ de orden 1, que modela su observación. Los elementos de dicha cadena son:

$S = \{S_1, S_2\}$, en el que cada estado representa uno de estos comportamientos, es decir: S_1 : Normal, S_2 : Falla.

$$\boldsymbol{\pi} = \begin{bmatrix} 1.0 \\ 0.0 \end{bmatrix}; \quad \mathbf{A} = \begin{bmatrix} 0.9 & 0.1 \\ 0.0 & 1.0 \end{bmatrix}.$$

Gráficamente, μ se puede representar como se muestra en la Figura 2.1.

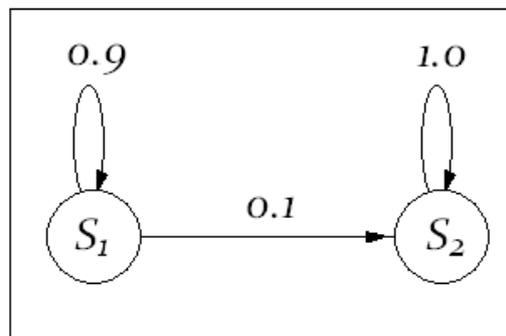


Figura 2.1. Ejemplo de Cadena de Markov con 2 estados.

Si se deseara realizar 4 pruebas al sistema en diferentes instantes de tiempo, ¿Cuál sería la probabilidad de obtener la secuencia de estados $O = \{S_1, S_1, S_1, S_2\}$? Para responder esta pregunta se debe calcular el valor $P(O | \mu)$. Utilizando los parámetros \mathbf{A} y $\boldsymbol{\pi}$, se obtiene:

$$\begin{aligned} P(O | \mu) &= P(S_1, S_1, S_1, S_2 | \mu) \\ &= P(S_1) \cdot P(S_1 | S_1) \cdot P(S_1 | S_1) \cdot P(S_2 | S_1) \\ &= \pi_1 \cdot a_{1,1} \cdot a_{1,1} \cdot a_{1,2} = 0.081 \end{aligned}$$

2.1.2. HMM

En el modelo anterior, se supone que cada estado es observable en cada instante de tiempo. Ahora se introduce una variante en la que los estados que definen al modelo están ocultos, por lo que el proceso estocástico representado también está oculto y sólo puede observarse asignando probabilidades a las respuestas que se obtienen del sistema modelado, a las cuales se referirá como *probabilidades de emisión*, para cada estado. He aquí el por qué se le da ese nombre a este modelo.

En el caso en el que las respuestas del sistema modelado se consideran valores discretos, se hace referencia a éstos como *símbolos de observación* y, por ello, se les asignan distribuciones

de probabilidad discretas a cada estado, dando lugar a un HMM discreto. Para este trabajo se consideran este tipo de HMM. Ahora, no se trabaja con secuencias de estados, sino con *secuencias de símbolos de observación* o, simplemente, *secuencias de observación*.

2.1.3. Elementos de un HMM

Un HMM discreto está definido por los siguientes elementos:

S : Espacio de estados, con N elementos.

π : Probabilidades de estado inicial.

A : Matriz de transición.

V : Conjunto de símbolos de observación, con M elementos.

$$V = \{v_i \mid v_i \in \{1, 2, \dots, M\}; i = 1, \dots, M\}. \quad (2.7)$$

B : Matriz de emisión. $B = [b_{i,k}]; i = 1, \dots, N; k = 1, \dots, M;$

$$b_{i,k} = P(v_k \text{ en } t \mid q_t = S_i). \quad (2.8)$$

Los primeros tres elementos se definen como en la sección 2.1.1. Como se aprecia, los últimos dos elementos hacen la diferencia entre un HMM y un Modelo de Markov. Los elementos de la matriz B son las probabilidades de emisión que se mencionaron anteriormente para cada estado, mientras que las respuestas del sistema modelado se definen en V . Además, para B se tiene la siguiente restricción:

$$\sum_{k=1}^M b_{i,k} = 1; i = 1, \dots, N \quad (2.9)$$

Una secuencia de observación, normalmente se denotan como:

$$O = \{o_t \mid o_t \in V; t = 1, \dots, T\}, \quad (2.10)$$

donde o_t es el símbolo de observación correspondiente al instante de tiempo t , y T es la longitud de dicha secuencia.

2.2. Tipos de HMM

Básicamente, los HMMs usados con mayor frecuencia en la práctica son los llamados *ergódicos* y los modelos *Bakis*. A continuación se describen las características de estos dos modelos.

2.2.1. Modelo Ergódico

Este modelo tiene la particularidad de permitir transiciones de un estado a cualquier otro, definido en el espacio de estados, sin restricciones. Esto evita que existan probabilidades de transición nulas y, por tanto, cada elemento de A será diferente de cero. Por ejemplo, si se definiera un HMM ergódico de 3 estados, su topología sería similar a la que se muestra en la Figura 2.2, en donde se aprecian todas las transiciones posibles.

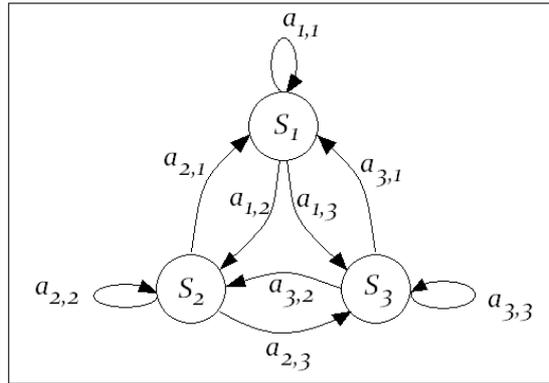


Figura 2.2. Ejemplo de un HMM ergódico de 3 estados.

2.2.2. Modelo Bakis

Para este modelo las probabilidades de transición tienen la siguiente propiedad:

$$a_{i,j} = 0, \text{ si } j < i; i, j = 1, \dots, N. \quad (2.11)$$

Lo anterior implica que no se permiten transiciones a estados con índices menores al del estado actual. En cuanto a las probabilidades de estado inicial, $\pi_1 = 1$, debido a que se establece que la secuencia de estados debe iniciar con el estado 1 y terminar en el estado N , causando que $\pi_i = 0$ para $i = 2, \dots, N$. Una restricción que se aplica a este modelo es el número de saltos permitidos para las transiciones de cada estado. Si este número se denota como Δ , entonces esta restricción se expresa formalmente como:

$$a_{i,j} = 0, \text{ si } j < i + \Delta. \quad (2.12)$$

Además, el valor máximo que Δ puede tomar es $N - 1$. A manera de ejemplo, si se define un HMM tipo Bakis con 4 estados y $\Delta = 2$, \mathbf{A} tendrá la forma:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 0 \\ 0 & a_{2,2} & a_{2,3} & a_{2,4} \\ 0 & 0 & a_{3,3} & a_{3,4} \\ 0 & 0 & 0 & a_{4,4} \end{bmatrix}.$$

La topología de este HMM de ejemplo se puede ver en la Figura 2.3. Además, para este tipo de modelos se observa que $a_{N,N} = 1$, para cumplir con la restricción dada en la ecuación (2.6). En el presente trabajo se hace uso de HMMs con topología Bakis.

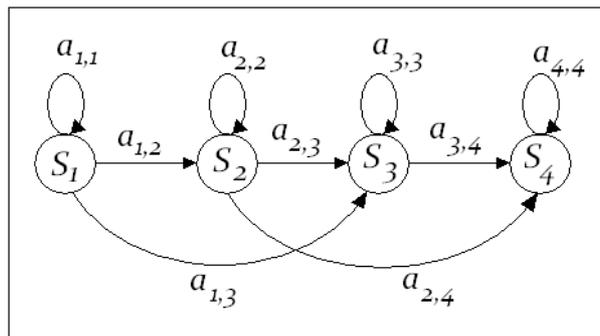


Figura 2.3. Ejemplo de HMM tipo Bakis, con $N = 4$ y $\Delta = 2$.

2.3. Problemas principales en el uso de HMMs

Se distinguen tres problemas a resolver para poder implementar los HMMs: *Evaluación*, *Decodificación* y *Aprendizaje*. A continuación se describe en qué consiste cada uno, así como los algoritmos que se necesitan para resolverlos. El nombre que se le da al segundo problema aparece como en [8] y, en adelante, se considera que un HMM se puede definir de la siguiente forma abreviada [25]:

$$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}). \quad (2.13)$$

2.3.1. Evaluación

Este problema consiste en evaluar la probabilidad de una secuencia de observación O , dado un HMM λ . Por lo tanto, se requiere calcular el valor de $P(O|\lambda)$ y, para lograrlo, existen dos algoritmos que son capaces de determinarlo: el *Algoritmo de Avance (Forward)* y el *Algoritmo de Retroceso (Backward)*.

Algoritmo de Avance. Para una secuencia de observación O y un HMM λ con N estados, este algoritmo consiste en el cálculo de un conjunto α de variables, denominadas *variables de avance*, el cual se definen como:

$$\alpha = \{\alpha_i(t); i = 1, \dots, N; t = 1, \dots, T\}, \quad (2.14)$$

donde

$$\alpha_i(t) = P(o_1, \dots, o_t, q_t = S_i | \lambda). \quad (2.15)$$

Es decir, $\alpha_i(t)$ representa la probabilidad de que se haya observado la secuencia o_1, \dots, o_t , y el estado actual sea S_i hasta el instante t , dado el modelo λ . En el Algoritmo 2.1 se presenta la obtención de estas variables.

Algoritmo de Retroceso. Para una secuencia de observación O y un HMM λ con N estados, este algoritmo calcula un conjunto β de variables, denominadas *variables de retroceso*, el cual se definen como:

$$\beta = \{\beta_i(t); i = 1, \dots, N; t = 1, \dots, T\}, \quad (2.16)$$

donde

$$\beta_i(t) = P(o_{t+1}, \dots, o_T | q_t = S_i, \lambda). \quad (2.17)$$

En este caso, $\beta_i(t)$ representa la probabilidad de que se observe la secuencia o_{t+1}, \dots, o_T , dados el modelo λ y el estado actual S_i , a partir del instante t . El Algoritmo 2.2 detalla la generación de estas variables.

Algoritmo 2.1. Algoritmo de Avance
<p>$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$: HMM a analizar.</p> <p>N: Número de estados del HMM.</p> <p>S: Espacio de estados. $S = \{S_1, \dots, S_N\}$.</p> <p>M: Número de símbolos de observación.</p> <p>V: Conjunto de símbolos de observación definidos. $V = \{v_i \mid v_i \in \{1, 2, \dots, M\}; i = 1, \dots, M\}$.</p> <p>$O$: Secuencia de observación. $O = \{o_t \mid o_t \in V; t = 1, \dots, T\}$</p> <p>$\alpha$: Conjunto de variables de avance. $\alpha = \{\alpha_i(t); i = 1, \dots, N; t = 1, \dots, T\}$</p>
<p>para $i = 1, \dots, N$ hacer</p> $\alpha_i(1) = \pi_i b_{i,o_1}$ <p>fin para</p> <p>para $t = 1, \dots, T$ hacer</p> <p>para $j = 1, \dots, N$ hacer</p> $\alpha_j(t+1) = \left(\sum_{i=1}^N \alpha_i(t) a_{i,j} \right) b_{j,o_{t+1}}$ <p>fin para</p> <p>fin para</p> $P(O \mid \lambda) = \sum_{i=1}^N \alpha_i(T)$ <p>devolver $\alpha, P(O \mid \lambda)$</p>

Algoritmo 2.2. Algoritmo de Retroceso
<p>$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$: HMM a analizar.</p> <p>N: Número de estados del HMM.</p> <p>S: Espacio de estados. $S = \{S_1, \dots, S_N\}$.</p> <p>M: Número de símbolos de observación.</p> <p>V: Conjunto de símbolos de observación definidos. $V = \{v_i \mid v_i \in \{1, 2, \dots, M\}; i = 1, \dots, M\}$.</p> <p>$O$: Secuencia de observación. $O = \{o_t \mid o_t \in V; t = 1, \dots, T\}$</p> <p>$\beta$: Conjunto de variables de retroceso. $\beta = \{\beta_i(t); i = 1, \dots, N; t = 1, \dots, T\}$</p>
<p>para $i = 1, \dots, N$ hacer</p> $\beta_i(T) = 1$ <p>fin para</p> <p>para $t = T-1, \dots, 1$ hacer</p> <p>para $i = 1, \dots, N$ hacer</p> $\beta_i(t) = \sum_{j=1}^N a_{i,j} b_{j,o_{t+1}} \beta_j(t+1)$ <p>fin para</p> <p>fin para</p> $P(O \mid \lambda) = \sum_{i=1}^N \pi_i b_{i,o_1} \beta_i(1)$ <p>devolver $\beta, P(O \mid \lambda)$</p>

La Figura 2.4 ilustra la obtención de las variables de avance y retroceso. Aunque cada uno de los dos algoritmos anteriores permite calcular el valor de $P(O \mid \lambda)$, la implementación de ambos es necesaria para resolver el problema de Aprendizaje, que se verá en la sección 2.3.3.

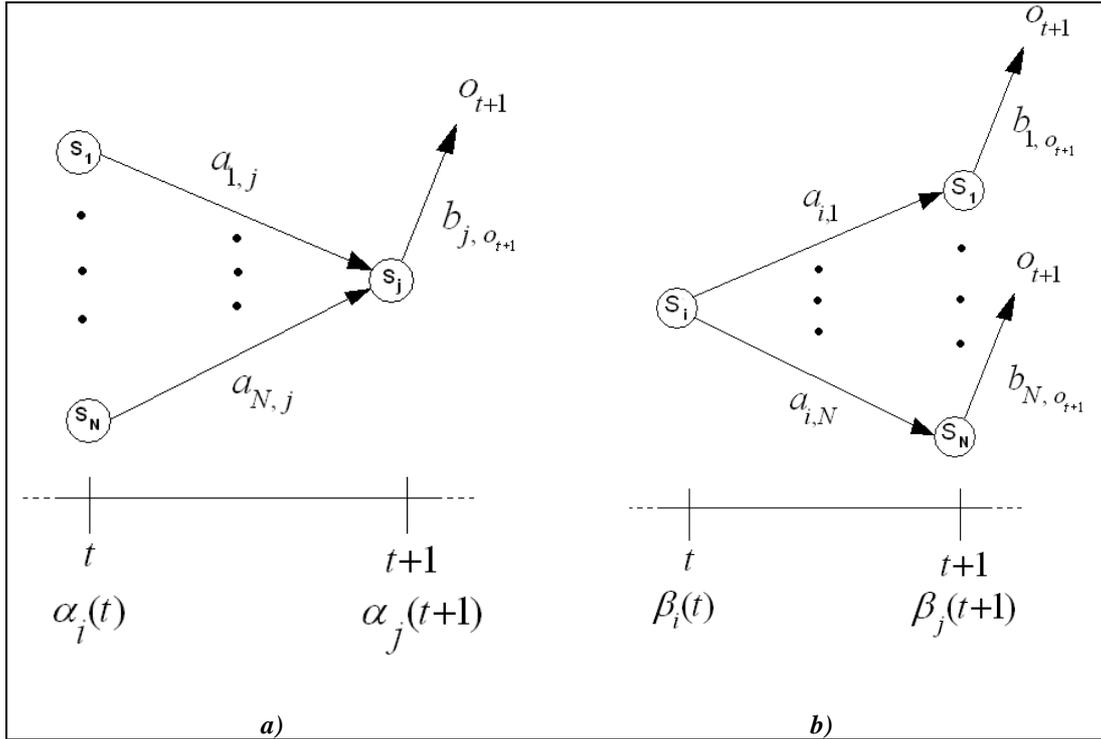


Figura 2.4. a) Obtención de las variables de avance $\alpha_j(t+1)$; b) Obtención de las variables de retroceso $\beta_i(t)$.

2.3.2. Decodificación

Dada una secuencia de observación O y un HMM λ con N estados, el problema de Decodificación tiene que ver con determinar la secuencia de estados que mejor describa a O . En otras palabras, se requiere obtener la secuencia más probable de estados para O dado λ y, para encontrarla, se utiliza el Algoritmo Viterbi. Véase Algoritmo 2.3.

2.3.3 Aprendizaje

Este problema consiste en ajustar los parámetros de un HMM λ , con N estados, para que describan mejor una secuencia de observación O y, por tanto, se pueda maximizar el valor de $P(O|\lambda)$. En este contexto, O se puede denominar *secuencia de observación de entrenamiento*. El ajuste mencionado se logra a través de un proceso iterativo llamado *Algoritmo Baum-Welch*, el cual requiere de la aplicación de los algoritmos vistos para la solución del problema de Evaluación, en la sección 2.3.1. En la Figura 2.5 se pueden ver los elementos que intervienen en este algoritmo.

Primero se deben inicializar los parámetros de λ , ya sea con base en un análisis previo de los datos a modelar o, en su defecto, de forma aleatoria, respetando las restricciones (2.4), (2.6) y (2.9). Si se trabaja con un modelo Bakis entonces, a cada paso del algoritmo Baum-Welch, se aplican las fórmulas de reestimación (2.18) a (2.20) a los elementos de π , \mathbf{A} y \mathbf{B} , respectivamente.

$$\bar{\pi}_i = \begin{cases} 1 & \text{si } i = 1 \\ 0 & \text{si } i \neq 1 \end{cases} \quad (2.18)$$

$$\bar{a}_{i,j} = \frac{\sum_{t=1}^{T-1} \alpha_i(t) a_{i,j} b_{j,o_{t+1}} \beta_j(t+1)}{\sum_{t=1}^{T-1} \alpha_i(t) \beta_i(t)} \quad (2.19)$$

$$\bar{b}_{j,m} = \frac{\sum_{t=1}^T \alpha_j(t) \beta_j(t)}{\sum_{t=1}^T \alpha_j(t) \beta_j(t)} \quad (2.20)$$

El empleo de estas fórmulas genera, al finalizar el algoritmo, un HMM entrenado $\bar{\lambda} = (\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\boldsymbol{\pi}})$. La estructura general de este proceso se muestra en el Algoritmo 2.4, en donde la condición de paro está dada, por una parte, por el decremento en el valor de $P(O | \bar{\lambda})$ y, por otra parte, por el alcance de un número máximo de iteraciones permitido. En [31] se proporciona un ejemplo de aplicación de este algoritmo, y sus correspondientes resultados después de 100 iteraciones.

Algoritmo 2.3. Algoritmo Viterbi [25]
$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$: HMM a analizar. N : Número de estados del HMM. S : Espacio de estados. $S = \{S_1, \dots, S_N\}$. M : Número de símbolos de observación. V : Conjunto de símbolos de observación definidos. $V = \{v_i \mid v_i \in \{1, 2, \dots, M\}; i = 1, \dots, M\}$. O : Secuencia de observación. $O = \{o_t \mid o_t \in V; t = 1, \dots, T\}$ Q : Secuencia más probable de estados. $Q = \{q_t \mid q_t \in S; t = 1, \dots, T\}$ $\delta = \{\delta_i(t); i = 1, \dots, N; t = 1, \dots, T\}$ $\psi = \{\psi_i(t); i = 1, \dots, N; t = 1, \dots, T\}$
para $i = 1, \dots, N$ hacer $\delta_i(1) = \pi_i b_{i,o_1}$ $\psi_i(1) = 0$ fin para para $t = 2, \dots, T$ hacer para $j = 1, \dots, N$ hacer $\delta_j(t) = \left[\max_{i=1}^N (\delta_i(t-1) a_{i,j}) \right] b_{j,o_t}$ $\psi_j(t) = \arg \max_{i=1}^N (\delta_i(t-1) a_{i,j})$ fin para fin para $P = \max_{i=1}^n (\delta_i(T)); q_T = \arg \max_{i=1}^n (\delta_i(T))$ para $t = T-1, \dots, 1$ hacer $q_t = \psi_{q_{t+1}}(t+1)$ fin para devolver Q

Algoritmo 2.4. Algoritmo Baum-Welch
$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$: HMM a analizar. $\bar{\lambda} = (\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\boldsymbol{\pi}})$: λ entrenado. N : Número de estados del HMM. S : Espacio de estados. $S = \{S_1, \dots, S_N\}$. M : Número de símbolos de observación. V : Conjunto de símbolos de observación definidos. $V = \{v_i \mid v_i \in \{1, 2, \dots, M\}; i = 1, \dots, M\}$. O : Secuencia de observación de entrenamiento. $O = \{o_t \mid o_t \in V; t = 1, \dots, T\}$. <i>itersmax</i> : Número máximo de iteraciones.
Inicializar λ $i = 0$ hacer $i = i + 1$ Calcular $P(O \mid \lambda)$ Aplicar fórmulas de reestimación a $\lambda \rightarrow \bar{\lambda}$ Calcular $P(O \mid \bar{\lambda})$ $\lambda \leftarrow \bar{\lambda}$ mientras (($P(O \mid \bar{\lambda}) > P(O \mid \lambda)$) y ($i < \textit{itersmax}$)) devolver $\bar{\lambda}$

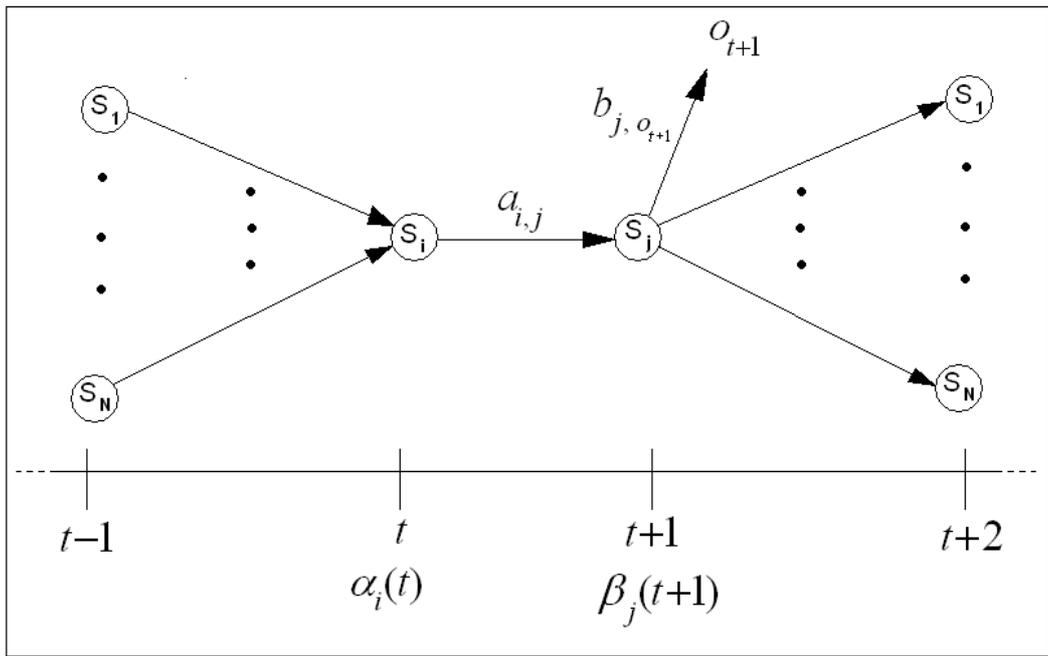


Figura 2.5. Elementos involucrados en el Algoritmo Baum-Welch.

Los algoritmos vistos hasta ahora manejan la probabilidad $P(O \mid \lambda)$ sin ninguna consideración especial. Sin embargo, en la práctica se suele usar el valor logarítmico de esta probabilidad. Existen también otros detalles importantes en cuanto a implementación de HMMs, los cuales se tratan en la sección 2.4.

2.4 Consideraciones adicionales en la implementación de HMMs

En las secciones 2.4.1 y 2.4.2 se presentan dos puntos que se deben tomar en cuenta en la implementación de HMMs. En cada una de estas secciones se muestran las respectivas modificaciones que se hacen a las fórmulas de reestimación vistas en la sección 2.3.3.

2.4.1 Escalamiento

En el cálculo de las variables de avance se detecta que su valor comienza a acercarse exponencialmente a cero conforme t se hace más grande, lo que llega a provocar que se exceda la precisión que una computadora puede proporcionar [25]. Por tal motivo, es necesario incluir factores de escala para las variables de avance y también para las de retroceso. Para $t=1, \dots, T-1$, primero se calculan los términos $\alpha_i(t)$ y posteriormente se obtienen sus valores escalados $\hat{\alpha}_i(t)$, de acuerdo a la ecuación (2.21):

$$\hat{\alpha}_i(t) = c_t \alpha_i(t) \quad (2.21)$$

donde c_t es el factor de escala definido como:

$$c_t = \frac{1}{\sum_{i=1}^N \alpha_i(t)} \quad (2.22)$$

Para el siguiente valor de t , los términos $\alpha_i(t)$ y el factor c_t ahora estarán en función de los valores $\hat{\alpha}_i(t-1)$. Los mismos factores de escala se aplican a las variables de retroceso. Así, para cada $\beta_j(t)$, su correspondiente valor escalado tendrá la forma $\hat{\beta}_j(t) = c_t \beta_j(t)$. El uso de estos factores modifica las fórmulas de reestimación de la sección 2.3.3. Para los elementos de $\boldsymbol{\pi}$ se sigue utilizando la ecuación (2.18) mientras que, para los elementos de \mathbf{A} y \mathbf{B} , se recurre a las ecuaciones (2.23) y (2.24), respectivamente.

$$\bar{a}_{i,j} = \frac{\sum_{t=1}^{T-1} \hat{\alpha}_i(t) a_{i,j} b_{j,o_{t+1}} \hat{\beta}_j(t+1)}{\sum_{t=1}^{T-1} \frac{\hat{\alpha}_i(t) \hat{\beta}_j(t)}{c_t}} \quad (2.23)$$

$$\bar{b}_{j,m} = \frac{\sum_{t=1}^T \frac{\hat{\alpha}_j(t) \hat{\beta}_j(t)}{c_t}}{\sum_{t=1}^T \frac{\hat{\alpha}_j(t) \hat{\beta}_j(t)}{c_t}} \quad (2.24)$$

2.4.2 Uso de múltiples secuencias de observación

Una de las restricciones que imponen los modelos tipo Bakis es que se necesita más de una secuencia de observación para entrenarlos correctamente. Lo anterior implica que ahora se

maneje un conjunto de K secuencias de observación de entrenamiento, el cual se expresa como:

$$O = \{O^k; k = 1, \dots, K\} \quad (2.25)$$

donde $O^k = \{o_t^k \mid o_t^k \in V; t = 1, \dots, T_k\}$. Al igual que en la sección 2.4.1, las fórmulas de reestimación de la sección 2.3.3 sufren modificaciones. Para los elementos de $\boldsymbol{\pi}$ se continúa utilizando la ecuación (2.18) pero, para los elementos de \mathbf{A} y \mathbf{B} , se usan las ecuaciones (2.26) y (2.27), respectivamente.

$$\bar{a}_{i,j} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \hat{\alpha}_i^k(t) a_{i,j} b_{j,o_{t+1}^k} \hat{\beta}_j^k(t+1)}{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \frac{\hat{\alpha}_i^k(t) \hat{\beta}_i^k(t)}{c_t^k}} \quad (2.26)$$

$$\bar{b}_{j,m} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k} \frac{\hat{\alpha}_j^k(t) \hat{\beta}_j^k(t)}{c_t^k}}{\sum_{k=1}^K \sum_{t=1}^{T_k} \frac{\hat{\alpha}_j^k(t) \hat{\beta}_j^k(t)}{c_t^k}} \quad (2.27)$$

Estas ecuaciones contemplan el escalamiento visto en la sección 2.4.1. Los términos c_t^k forman el conjunto de factores de escala correspondientes a la k -ésima secuencia de observación de entrenamiento para $t = 1, \dots, T_k$. Se puede encontrar una explicación más detallada de estas ecuaciones en [25], [29] y [31].

Capítulo 3. Visión general del Reconocimiento de Rostros

3.1. Descripción general del Reconocimiento de Rostros

Una de las principales formas en las que los seres humanos pueden reconocer a su semejantes es a través de sus rostros. Esta capacidad se ha intentado replicar en sistemas de cómputo desde los años 60s, aproximadamente [2], y ése es, precisamente, el interés del Reconocimiento de Rostros, el cual puede entenderse como el campo de estudio del conjunto de técnicas dedicadas a desarrollar Sistemas de Reconocimiento de Rostros (SRRs), y su objetivo es el reconocimiento automático de rostros a través del diseño de una metodología determinada. La evolución de estos sistemas, se ha caracterizado por un funcionamiento cada vez más autónomo y preciso. De acuerdo con [19], el primer SRR automático fue desarrollado por Takeo Kanade, en 1973.

En relación con años más cercanos, se ha notado que las investigaciones en este campo han tenido un gran auge en las últimas dos décadas. En la Figura 3.1 se puede apreciar claramente este hecho. Dicha gráfica compara el número de publicaciones que se han hecho para dos temas relacionados estrechamente: el análisis de imágenes de rostros (*face image analysis*) y el reconocimiento de rostros (*face recognition*).

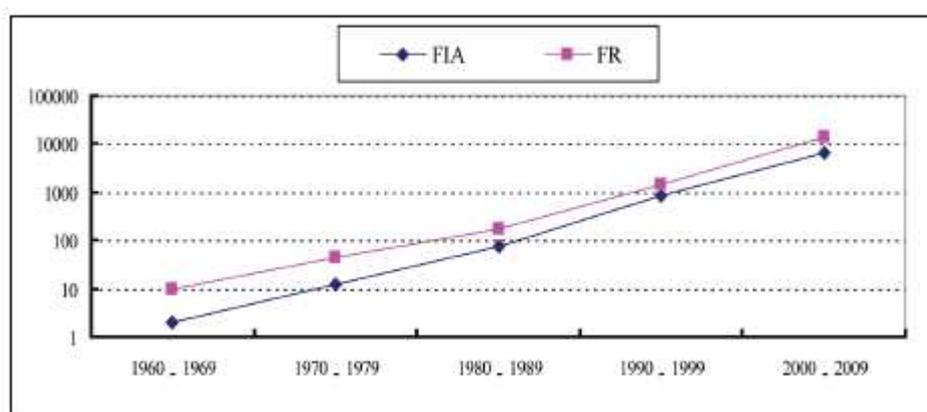


Figura 3.1. Número de publicaciones con “face image anlysis (FIA)” y “face recognition (FR)” [41].

El análisis de imágenes de rostros se refiere al procedimiento, a través del cual es posible obtener información que, entre otros aspectos, puede servir para clasificación de expresiones faciales, determinación de género, estimación de edad y reconocimiento de rostros, siendo este último el tema de investigación que, como se mencionaba anteriormente, ha adquirido más popularidad recientemente.

Un SRR puede trabajar básicamente de dos modos:

- Verificación (o autenticación) de rostros. Hace referencia a la comparación uno a uno de una imagen de un rostro de prueba con la imagen de un rostro buscado.
- Identificación (o reconocimiento) de rostros. Es la comparación de una imagen de un rostro de prueba con un conjunto de imágenes de rostros que se encuentran en una base de datos, con base en los que se debe encontrar la identidad del rostro de prueba.

En el caso del SRRP, se tiene como fin la identificación de rostros. El funcionamiento de dicho sistema se describe con más detalle en el Capítulo 4.

3.2. Etapas de procesamiento de un SRR

En la Figura 3.2 se presenta un diagrama sobre el flujo del procesamiento que, en general, se sigue en un SRR. En este diagrama, que está basado en el que aparece en el Capítulo 1 de [19], se distinguen 4 etapas principales de procesamiento, que se describirán en las secciones 3.2.1 a 3.2.4, siguiendo el mismo orden.

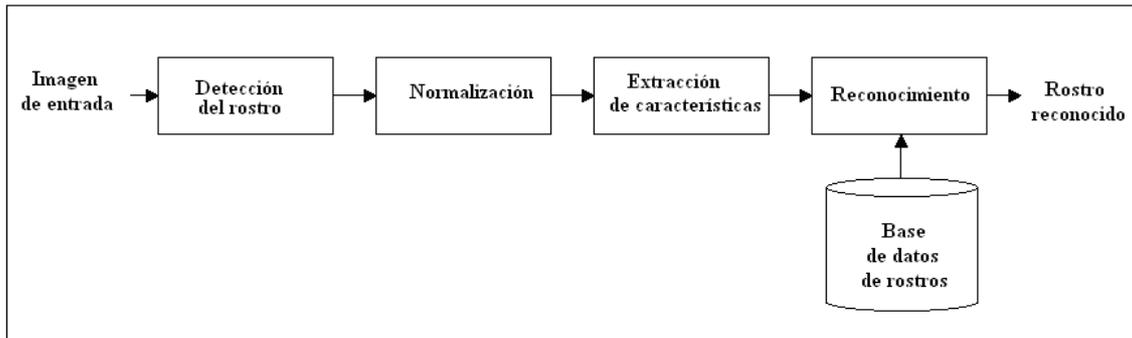


Figura 3.2. Flujo de procesamiento de un SRR.

3.2.1. Detección del rostro

Consiste en la localización de la región ocupada por el rostro en la imagen de entrada (ver Figura 3.3 a)). Con frecuencia se recurre al uso de técnicas de segmentación para aislar el rostro del resto de la imagen, con lo cual se puede determinar la posición y la extensión del área de interés. En esta etapa también se puede incluir un proceso de localización de ciertos rasgos faciales, como ojos, nariz, orejas, boca, entre otros, para un análisis más detallado (ver Figura 3.3 b)).

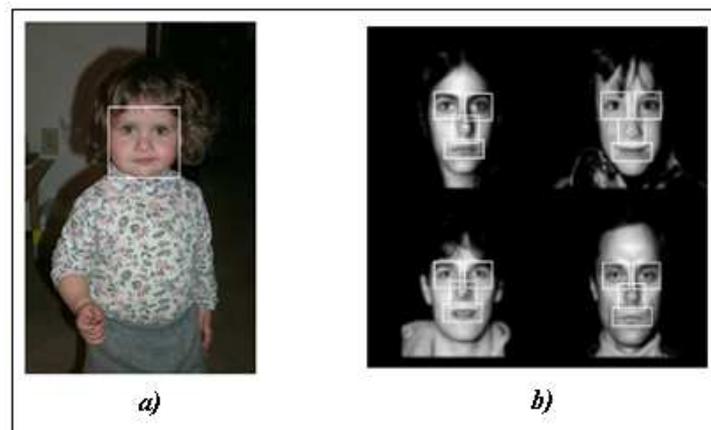


Figura 3.3. a) Ejemplo de detección de un rostro en una imagen [19]. b) Ejemplo de detección de rasgos faciales, para este caso: ojos, nariz y boca [35].

3.2.2. Normalización

Esta etapa tiene por objetivo obtener una imagen estandarizada tanto en escala como en propiedades que presentan variaciones importantes, como iluminación, posición del rostro y color. Esta última propiedad ha sido manejada muchas veces con una escala de grises. Cuando se cuenta con imágenes en las que el rostro presenta variaciones en su posición, se requiere aplicar un proceso de alineación, para obtener imágenes con posiciones de rasgos faciales lo más uniformes posible.

3.2.3. Extracción de características

En esta etapa se extrae de la imagen normalizada la información que se considera lo suficientemente relevante como para permitir distinguir rostros de diferentes personas. Entre las características que comúnmente se toman en cuenta están los colores, las texturas y las características geométricas, tales como distancias, ángulos y curvaturas de rasgos faciales.

Por lo general, el resultado de esta etapa arroja una cantidad considerable de datos que, en el caso de representarlos por medio de un vector conocido como *vector de características*, se traduce en una dimensionalidad alta de dicho vector. Lo anterior hace necesaria la aplicación de técnicas de reducción de la dimensionalidad, con las cuales se pueden evitar algunos problemas de cómputo y aumentar la precisión obtenida en la etapa de reconocimiento.

El desempeño que los SRRs exhiben, depende en gran medida de la información que resulta de esta etapa. Al respecto, uno de los problemas más importantes con los que aún se enfrenta este campo de estudio es encontrar formas adecuadas de medir la similitud entre características faciales, para realizar la comparación de éstas en la siguiente etapa de procesamiento.

3.2.4. Reconocimiento

En esta etapa se contempla la comparación de las características extraídas de la imagen de entrada con aquellas de las imágenes contenidas en la base de datos de rostros, a las cuales se les aplica el mismo proceso definido para la etapa de extracción de características.

Retomando los posibles modos en los que un SRR puede trabajar, planteados en la sección 3.1, el resultado de esta etapa puede tener las siguientes formas.

- Para Verificación, sólo se dará una respuesta positiva o negativa, dependiendo de si la imagen del rostro de entrada corresponde o no con la imagen del rostro del sujeto buscado.
- Para Identificación, se pueden considerar dos situaciones: a) si el sistema no considera un umbral de certeza, se obtendrá la identidad del sujeto cuyo rostro fue introducido al sistema o, en su defecto, la identidad del rostro más parecido; b) si se considera un determinado umbral de certeza, el sistema obtendrá la identidad del sujeto cuyo rostro fue introducido, si es que dicha identificación sobrepasa el umbral. De lo contrario el sistema responde que la identidad del rostro de la imagen de entrada es desconocida.

3.3. Descripción general de algunas técnicas empleadas en Reconocimiento de Rostros

3.3.1. PCA y Eigenfaces

El Análisis de Componentes Principales (PCA) es una técnica que surgió del trabajo de Sirivich y Kirby en 1987 [19], empleada en muchos problemas de Clasificación de Patrones, la cual permite reducir la dimensión del espacio en el que están definidos los datos a analizar. Los componentes principales de dichos datos se calculan a partir de los eigenvalores de su matriz de covarianza [26]. Es particularmente útil cuando dicha dimensión es muy alta, como ocurre en el procesamiento de imágenes. Con respecto al Reconocimiento de Rostros, el uso de PCA consiste en realizar una proyección de una imagen en un subespacio de menor dimensión, al que se le denomina *subespacio de rostros* [39], [19].

La aplicación de PCA al Reconocimiento de Rostros más relevante se conoce como Eigenfaces y fue propuesta por Turk y Pentland [37] en 1991. Para realizar cálculos con una imagen, los elementos de ésta se almacenan en un vector columna, concatenando los renglones de la matriz que la representa [26], [2]. Esta técnica establece que, a partir de un

conjunto X de imágenes de entrenamiento, se puede reconstruir una versión comprimida y de una imagen x de un rostro, por medio de su correspondiente imagen media \mathbf{m} y un pequeño conjunto de imágenes escaladas \mathbf{u}_i denominadas eigenfaces, donde $i = 1, \dots, M$. Esto se puede expresar por medio de la ecuación 3.1:

$$\mathbf{y} = \mathbf{m} + \sum_{i=0}^{M-1} a_i \mathbf{u}_i \quad (3.1)$$

Para una imagen x , los coeficientes a_i de la ecuación (3.1) se obtienen por medio de la ecuación (3.2):

$$a_i = (\mathbf{x} - \mathbf{m}) \cdot \mathbf{u}_i \quad (3.2)$$

Por otro lado, los términos \mathbf{u}_i son los eigenvectores de la matriz de covarianza del conjunto X , que resultan de la descomposición de eigenvalores de dicha matriz. Una explicación más detallada de la técnica Eigenfaces se puede encontrar en [35]. En la Figura 3.4 se muestra un ejemplo de la aplicación de esta técnica.

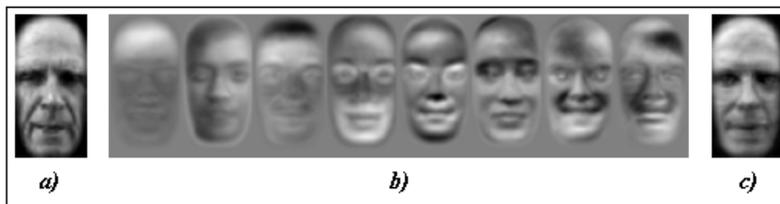


Figura 3.4. Ejemplo de aplicación de PCA al Reconocimiento de Rostros [35]. a) Imagen de entrada; b) primeras 8 eigenfaces; c) imagen de entrada reconstruida y comprimida.

3.3.2. LDA

Al igual que PCA, el Análisis Discriminante Lineal (LDA) es una técnica para reducción de la dimensionalidad y se emplea en problemas de clasificación cuyo objetivo es encontrar una dirección \mathbf{u}^* tal que la separación entre las clases sea la más alta. Esto se logra maximizando una relación llamada Criterio de Fisher, la cual está en función de una matriz de dispersión inter-clase S_B y una matriz de dispersión intra-clase S_W . La ecuación (3.3) expresa lo anterior [6], [35]:

$$\mathbf{u}^* = \arg \max_{\mathbf{u}} \frac{\mathbf{u}^T S_B \mathbf{u}}{\mathbf{u}^T S_W \mathbf{u}} \quad (3.3)$$

Así, se pretende maximizar la varianza inter-clase y minimizar la varianza intra-clase lo cual, en Reconocimiento de Rostros, permite clasificar imágenes de rostros de distintos sujetos [2]. Pero, cuando se trabaja con una dimensionalidad muy alta, LDA presenta inconvenientes [39]. La Figura 3.5 muestra un ejemplo de clasificación de imágenes de rostros de distintos sujetos usando LDA.



Figura 3.5. Clasificación de imágenes de rostros de 6 sujetos, aplicando LDA (tomada de [2]).

3.3.3 Modelos 3D

Como una alternativa para atacar algunas de las dificultades que se presentan en el Reconocimiento de Rostros, se han introducido modelos en 3D, ya que estos aportan información completa sobre la geometría del rostro y, en algunos casos, del cráneo humano. La representación más sencilla de un rostro en 3D se logra a través de una malla poligonal, como la de la Figura 3.6 a) [41]. Sin embargo, existen otros enfoques que se valen de tecnologías especiales de escaneo con las cuales se logran resultados más robustos, como en la Figura 3.6 b). Una revisión más detallada de éstas y otras técnicas referentes a este tema se puede encontrar en [41] y [19].

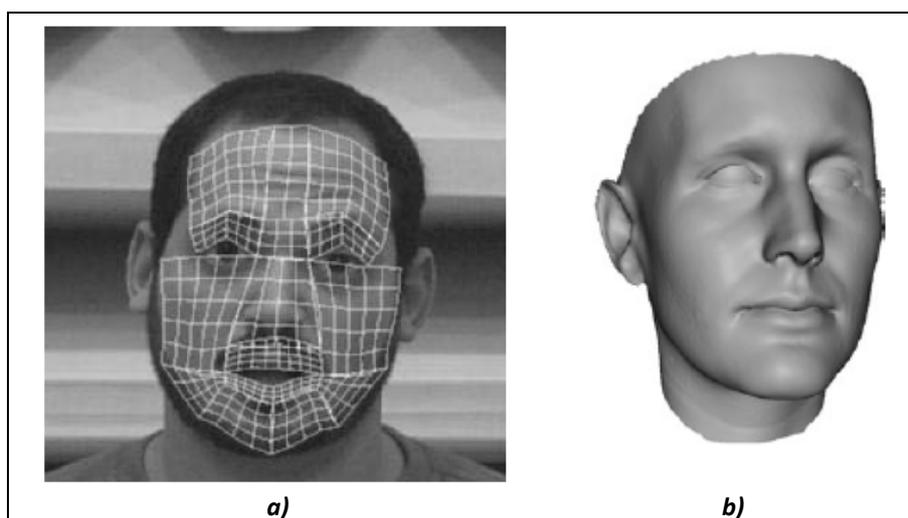


Figura 3.6. Ejemplos de modelos 3D utilizados en Reconocimiento de Rostros: a) malla para representación de un rostro en 3D (tomada de [27]); b) escaneo de un rostro en 3D (tomada de [19]).

3.3.4. HMMs y Wavelets en Reconocimiento de Rostros

Las técnicas para Reconocimiento de Rostros que resultan de mayor interés para el presente trabajo tienen que ver con las que se valen de los HMMs y de la Transformada Wavelet. Recientemente se han propuesto diversos enfoques que usan estos dos conceptos, por separado o en combinación. En [4], por ejemplo, se presenta un método en el que se utiliza dicha transformada para extraer la información de la imagen de un rostro, usando un esquema de muestreo que consiste en un barrido de izquierda a derecha y de arriba hacia abajo. Este mismo esquema aparece en [17], [22] y [26], con variaciones en relación a la codificación de los datos. El uso de la Transformada Wavelet, también conduce a algoritmos más eficientes que permiten una extracción de vectores de características más óptima, como en [11].

En cuanto a los HMMs, aproximadamente desde hace dos décadas se han creado métodos dirigidos al Reconocimiento de Rostros que han mostrado excelentes resultados. Por ejemplo, se ha logrado un reconocimiento del 100 % en la BDR ORL [3], usando HMMs continuos en [4] y discretos en [18]. En ambos trabajos, se han usado filtros de descomposición de tipo Haar que, a pesar de su sencillez, logra ser muy útil. Otro aspecto de los HMMs que se ha estudiado es su topología, encontrando diversas propuestas como la de [18], en la que se menciona el uso de modelos ergódicos de 5 estados. Se han empleado modelos tipo Bakis junto con barridos verticales de las imágenes para muestrearlas [26] (ver Figura 3.7 a)) y también se han ideado otros, llamados *Pseudo 2D HMMs* (P2D-HMMs) con esquemas de muestreo como el comentado en el párrafo anterior. En la Figura 3.7 b) se puede ver un ejemplo de este tipo de modelos. En [38] se analizan P2D-HMMs por medio de una comparación de su implementación continua y discreta, mostrando que incluso esta última proporciona buenos resultados.

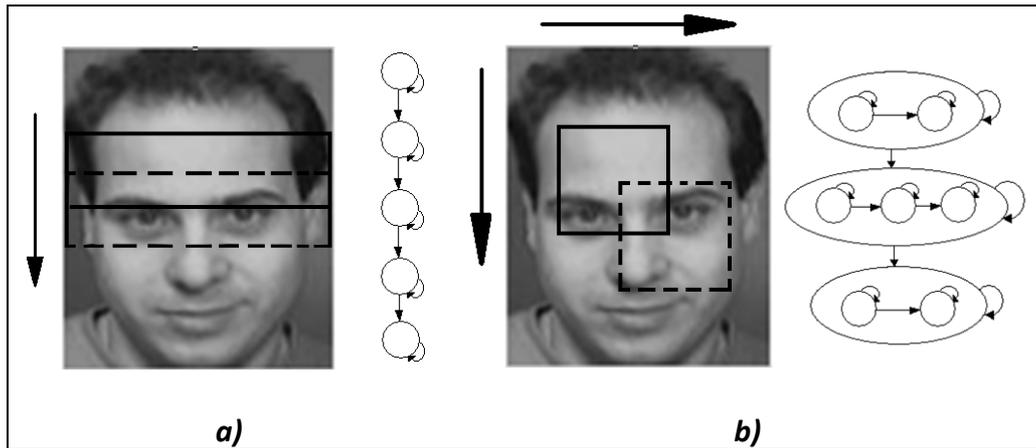


Figura 3.7. Ejemplos de HMMs y esquemas de muestreo usados para Reconocimiento de Rostros: a) HMM Bakis; b) P2D- HMM.

3.4. Principales dificultades en el Reconocimiento de Rostros

A continuación se describen las principales dificultades que, hasta ahora, se consideran como las que más influyen en el funcionamiento de un SRR.

3.4.1 Cambios de iluminación

Para la mayoría de las aplicaciones de Reconocimiento de Rostros, las condiciones en las que hay variaciones de iluminación hacen más difícil lograr un buen reconocimiento, ya que este factor altera considerablemente la etapa de extracción de características. Una solución que se ha propuesto es el uso de una tecnología de Reconocimiento de Rostros que se conoce como NIR (de *Near Infrared*), la cual consiste en utilizar cámaras infrarrojas para controlar la dirección de la iluminación y su intensidad.



Figura 3.8. Imágenes capturadas (arriba) por una cámara a color y (abajo) por un sistema de procesamiento de imágenes NIR (tomada de [19]).

En la Figura 3.8 se puede ver como el procesamiento de imágenes NIR puede lidiar de una mejor manera con los cambios de iluminación ambiental. Esta tecnología ha sido usada en algunos sistemas en China desde el año 2005 [19].

3.4.2 Cambios en la posición del rostro

Mientras las variaciones en la posición del rostro son más marcadas, es más difícil llevar a cabo un correcto reconocimiento de éste, por lo que el buen funcionamiento de algunos SRRs se debe a que las imágenes de la BDR correspondiente son, en su mayoría, imágenes de rostros en posición frontal y sólo se contemplan ligeros cambios de dicha posición (ver Figura 3.9). Por otro lado, para SRRs más robustos, se procura tener una BDR lo suficientemente amplia para enfrentar esta limitante.



Figura 3.9. Ejemplos de imágenes del rostro de un sujeto de la BDR ORL con cambios de posición [3].

3.4.3 Expresiones faciales

Considerando que los rostros humanos pueden exhibir diferentes emociones como enojo, disgusto, miedo, felicidad, tristeza y sorpresa, éstas a su vez se traducen en expresiones faciales que alteran los rasgos que caracterizan a un rostro. Esto compromete en mayor o menor medida la capacidad del SRR, dependiendo del tipo de algoritmos que utilice. Cuando se limita la apariencia del rostro a una expresión neutral, generalmente se obtienen buenos niveles de reconocimiento. Cabe mencionar que, además del Reconocimiento de Rostros, existe otro campo de estudio conocido precisamente como Reconocimiento de Expresiones Faciales, cuyo objetivo no es reconocer rostros sino determinadas emociones, como las mencionadas al inicio de este párrafo (ver Figura 3.9).

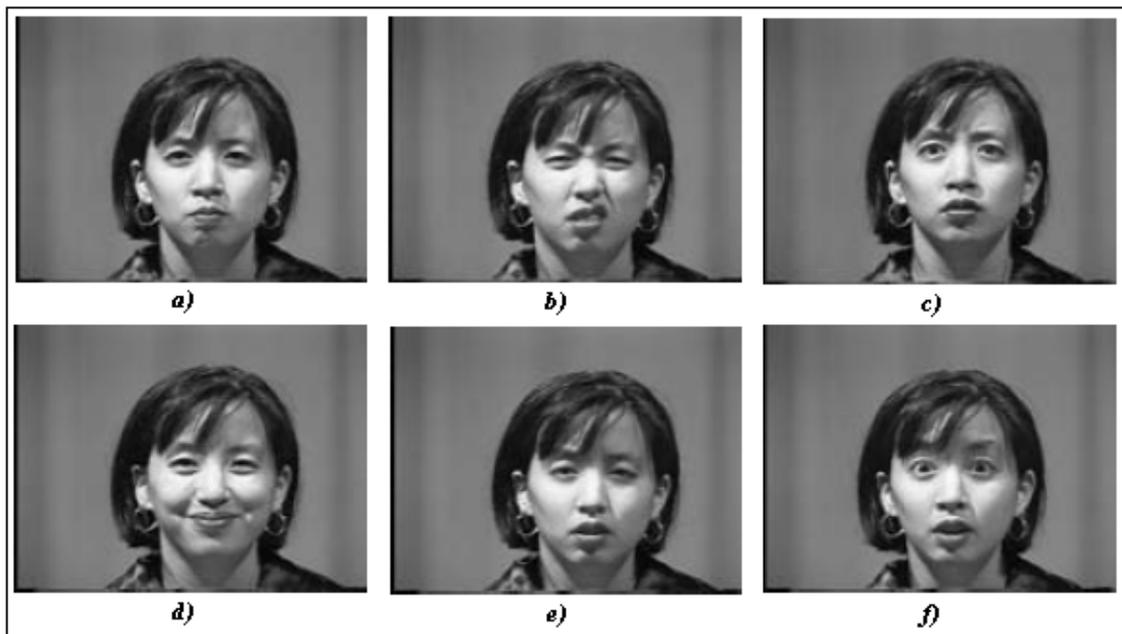


Figura 3.10. Imágenes de rostros de un mismo sujeto con diferentes expresiones faciales correspondientes a 6 emociones [27]: a) enojo, b) disgusto, c) miedo, d) felicidad, e) tristeza y f) sorpresa.

3.4.4 Edad

Si el tiempo transcurrido entre las pruebas que se realizan con un SRR es largo o si simplemente se le presenta a dicho sistema una imagen del rostro del sujeto a reconocer con una edad diferente a la actual, pueden presentarse variaciones importantes en el aspecto de

dicho rostro, las cuales no estarían contempladas por el sistema en cuestión. Dichas variaciones pueden reducir el desempeño de reconocimiento del sistema. Para solucionar este problema se requiere de actualizaciones constantes de las BDRs y, en algunos casos, del propio sistema, si se contemplan etapas de entrenamiento, por lo que esta alternativa resulta impráctica.

3.4.5 Accesorios y rasgos faciales

El uso de accesorios como lentes, cubre bocas, gorras, etc., dificulta el reconocimiento de rostros debido a que se introduce información no determinante sobre las características faciales de los sujetos a reconocer. En cuanto a los rasgos faciales, como la barba en hombres, si bien pueden jugar un papel importante al reconocer a una persona, también pueden provocar una extracción de características poco útil mientras más irregulares sean. Dicha extracción puede dificultarse aún más si se consideran tanto accesorios como rasgos faciales notorios. La Figura 3.11 muestra ejemplos de imágenes de rostros para estos casos.

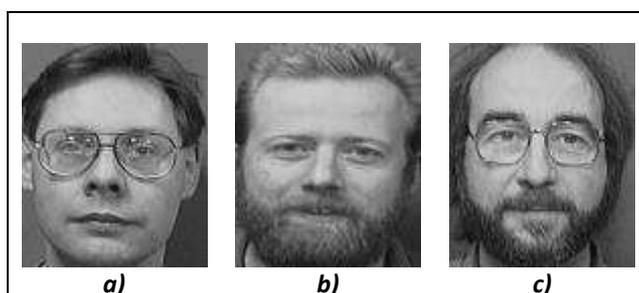


Figura 3.11. Ejemplos de: a) uso de accesorios (lentes); b) rasgos faciales notorios (barba); c) uso de accesorios (lentes) y presencia de rasgos faciales notorios (barba).

3.4.6 Cooperación del usuario

En las aplicaciones de Reconocimiento de Rostros se pueden presentar dos escenarios, dependiendo del tipo de interacción que exista entre el SRR y el usuario [19], [9]:

Escenarios con cooperación del usuario. Los SRR que operan en estos escenarios requieren que el usuario presente su rostro de una manera específica. Por lo general, se requiere de una posición frontal del rostro, así como una expresión neutral y una distancia razonablemente cercana a la cámara que capturará la imagen del rostro de la persona a ser reconocida. Lo anterior implica que las variaciones en la posición del rostro estén limitadas y que se dificulte el reconocimiento ante cambios en las condiciones ambientales.

Escenarios sin cooperación del usuario. En estos escenarios, el usuario no interactúa directamente con el SRR, pues éste debe trabajar bajo condiciones no controladas, es decir, sin restricciones en las posiciones del rostro, cambios de iluminación, ni uso de accesorios como lentes. El diseño de SRRs pensados para trabajar en escenarios sin cooperación del usuario resulta aún más complejo que el de los sistemas que requieren de la ayuda de éste, ya que, como se mencionó, requieren tomar en cuenta más factores.

3.5. Aplicaciones

Existe una gran variedad de problemas para los cuales los SRRs se pueden aplicar. En esta sección se explican brevemente las aplicaciones más comunes de dichos sistemas.

3.5.1 Seguridad

Las tecnologías de Reconocimiento de Rostros se han pensado usar en aeropuertos, estadios y otros lugares públicos, con el fin de identificar personas sospechosas o, simplemente, para

reforzar la seguridad y la vigilancia de los empleados y usuarios de dichos lugares. De hecho, algunas de ellas ya se usan actualmente (ver Figura 3.10). Sin embargo, debido a la enorme cantidad de personas a analizar y a las condiciones ambientales, los SRRs orientados a esta clase de aplicaciones no alcanzan aún un funcionamiento óptimo.



Figura 3.12. Uso de SRRs para verificación, en Beijing, China [19].

3.5.2 Control de acceso

Las aplicaciones de este tipo pueden servir ya sea para acceder a un lugar específico, por ejemplo una oficina o, simplemente, para identificación de usuarios de equipos de cómputo. Las particularidades que se pueden señalar en la mayoría de estas aplicaciones son básicamente el número relativamente pequeño de personas a ser reconocidas y las restricciones de las imágenes de éstas, tales como iluminación homogénea y rostro en posición frontal. No obstante, una ventaja de los SRRs para control de acceso es que necesitan poca cooperación del usuario [19].

3.5.3 Asistencia en departamentos de policía

Los SRRs están siendo utilizados cada vez más en departamentos de policía, debido a que con ellos es posible buscar e identificar rápidamente personas sospechosas, incluso con información incompleta de su identidad [19]. Las dificultades que estos sistemas pueden encontrar tienen que ver con la disponibilidad de BDRs de criminales o sospechosos y con la calidad de dichas imágenes, ya que muchas de ellas pueden haberse adquirido bajo condiciones no controladas de iluminación y pose.

3.5.4 Interacción humano-computadora

Actualmente, el desarrollo tecnológico de los equipos de cómputo ha aumentado considerablemente, por lo que el interés por interactuar de una forma más cercana con dichos equipos ha crecido también. Para lograr esta interacción, se requieren métodos de identificación de usuarios, de los cuales los más naturales son el Reconocimiento de voz y el Reconocimiento de Rostros [9]. Este último se caracteriza porque exige poca o, en algunos casos, nula cooperación del usuario aunque, como se comentó en la sección 3.4., mientras menos cooperación se contemple por parte del usuario, el diseño del sistema será más complejo.

Capítulo 4. Descripción del SRRP

En este capítulo se describe detalladamente el funcionamiento del sistema de reconocimiento de rostros propuesto el cual, como se mencionó en la sección 1.2, trabaja con la BDR ORL. Dicho sistema se compone de dos etapas principales: Entrenamiento y Reconocimiento. En la primera etapa, tienen lugar los procesos de extracción de características, la obtención de las secuencias de observación de entrenamiento y la obtención de HMMs entrenados. Estos procesos se aplican sobre un conjunto de imágenes de entrenamiento de la BDR. En la segunda etapa, se consideran los procesos de extracción de características para la imagen de prueba, obtención de la secuencia de observación de prueba y comparación de los HMMs, a partir de la cual se obtendrá el sujeto reconocido por el sistema para una imagen de prueba dada.

Cabe destacar que, al tener una etapa de entrenamiento y otra de reconocimiento o prueba, consecuentemente se estarán utilizando dos conjuntos diferentes de imágenes: un conjunto de imágenes de entrenamiento y otro de imágenes de prueba. La estructura de la BDR ORL permite esto, debido a que cada uno de sus sujetos cuenta con diferentes imágenes de su rostro. Por otro lado, las pruebas de reconocimiento pueden efectuarse sólo cuando el sistema haya sido entrenado, es decir, cuando se cuenta con los HMMs cuyos parámetros hayan sido ajustados. Las siguientes secciones se dedican a explicar en qué consisten cada uno de los procesos mencionados anteriormente, de acuerdo a la etapa a la que correspondan.

4.1 Etapa de Entrenamiento

En la Figura 4.1 se muestra un esquema de los procesos involucrados en la etapa de Entrenamiento del SRRP, así como el orden en que se llevan a cabo. Se hace énfasis en la parte de la BDR utilizada que, en este caso, se compone de las imágenes destinadas al entrenamiento del sistema.

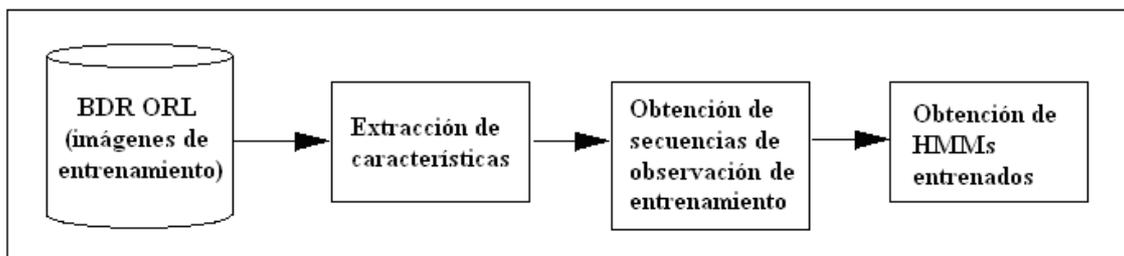


Figura 4.1. Procesos correspondientes a la etapa de Entrenamiento del SRRP.

El funcionamiento del SRRP requiere de un conjunto de parámetros que guiarán la etapa de entrenamiento y, a su vez, repercutirán en los resultados que se obtengan de la etapa de reconocimiento. Por ello, en primer lugar se aborda la definición de dichos parámetros en la sección 4.1.1 y, posteriormente, las secciones 4.1.2 a 4.1.4 se dedican a los procesos propios del entrenamiento del sistema.

4.1.1 Definición de parámetros

Los parámetros que se deben definir para que el SRRP realice los procesos de Entrenamiento y Reconocimiento se listan a continuación:

- Tipo de filtro paso-bajas a utilizar. Se consideran los siguientes tipos:
 - Haar
 - Biortogonal (B9/7)
 - Daubechies 4
- w_s y h_s : Ancho y alto de subimagen a utilizar, respectivamente. Estos parámetros se explican en la sección 4.1.2.
- N_s : Número de sujetos de la BDR a reconocer, igual al número de HMMs que se tendrán que considerar, ya que se busca representar a cada sujeto con un HMM. Al utilizar la BDR ORL completa, $N_s = 40$.
- N : Número de estados.
- Δ : Número de saltos permitidos entre estados.
- K : Número de imágenes de entrenamiento por sujeto. K debe ser menor al número de imágenes disponibles por sujeto (N_{ps}), para poder realizar pruebas de reconocimiento. En el caso de la BDR ORL, $N_{ps} = 10$, por lo que los valores mínimo y máximo de K serán 1 y 9, respectivamente.
- M : Número de símbolos de observación. El valor máximo de este parámetro se tomará igual a $N_s \cdot K$, lo cual se explica en la sección 4.1.3.
- Número máximo de iteraciones para el Algoritmo Baum-Welch.

Es importante recalcar que todos estos parámetros se aplicarán a todas las imágenes de entrenamiento de la BDR por igual. Por tal razón, todos los HMMs tendrán en común el número de estados, los saltos permitidos entre estados, los símbolos de observación a utilizar, así como el tipo de HMM que, para este trabajo, se toman como de tipo Bakis. El tipo de filtro elegido servirá para el proceso de Extracción de Características, el cual se trata en la sección 4.1.2.

4.1.2 Extracción de características

De cada imagen de entrenamiento se debe extraer la información que permita caracterizarla y, por tanto, reconocerla de entre las demás. Para ello se recurre al uso de subimágenes, cuya obtención es similar a la que se presenta en trabajos como [4], [17], [22] y [26]. La diferencia entre los enfoques anteriores y el del SRRP es que, para este último, primero se realiza una transformación de la imagen a analizar, con el fin de reducir la cantidad de datos que la representan.

En este proceso, se recurre a la aplicación de los Filtros de Descomposición. Concretamente, se obtiene el segundo nivel de descomposición de la sub-banda Paso-bajas/Paso-bajas de cada imagen de entrenamiento, es decir, se utilizan filtros paso-bajas junto con un proceso de submuestreo. En [18] se presenta un procedimiento similar, utilizando únicamente filtros de descomposición de tipo Haar. Sin embargo, para el SRRP se pretende utilizar los tres tipos de filtros paso-bajas indicados en la sección 4.1.1, cada uno de los cuales conducirá a una imagen de entrenamiento que se tomará como una versión reducida de la original. A partir de esta última, se procede a extraer las subimágenes que la componen, en forma de vectores columna. En la Tabla 4.1 se muestran los valores de los filtros mencionados.

En la Figura 4.2 se observa la forma en que se obtienen subimágenes de tamaño $w_s \times h_s$, a partir de un barrido de arriba hacia abajo y de izquierda a derecha de la versión reducida de una imagen de entrenamiento. Con la definición del tamaño de las subimágenes, se puede conocer el número de subimágenes que se generarán por cada imagen de entrenamiento que, para este trabajo, se considera igual a la longitud T de las secuencias de observación tanto de entrenamiento como de prueba.

Haar	Daubechies 4	Biortogonal 9/7
$1/\sqrt{2}$	$(1-\sqrt{3})/4\sqrt{2}$	0.02675
$1/\sqrt{2}$	$(3-\sqrt{3})/4\sqrt{2}$	-0.016864
	$(3+\sqrt{3})/4\sqrt{2}$	-0.078223
	$(1+\sqrt{3})/4\sqrt{2}$	0.266864
		0.602949
		0.266864
		-0.078223
		-0.016864
		0.02675

Tabla 4.1. Filtros paso-bajas de Descomposición: Haar [24], Daubechies 4 [24] y Biortogonal 9/7 [1].

Teniendo presentes los parámetros que aparecen en la sección 4.1.1, el Algoritmo 4.1 presenta de forma más clara el proceso de extracción de características. Para el SRRP, se suponen 7 tipos de subimagen a utilizar para este proceso, los cuales se muestran en la Tabla 4.2, junto con sus correspondientes valores de T. Cabe señalar que, para los últimos 3 tipos de subimagen, el barrido de la Figura 4.2 se limitará a un barrido vertical, debido al tamaño de dichas subimágenes.

Algoritmo 4.1. Extracción de características
$I_{hl \times wl}^{i,k}$: Imagen de entrenamiento k del sujeto i . $Y_{hX \times wX}^{i,k}$: Segundo nivel de descomposición de la sub-banda Paso-bajas/Paso-bajas de $I^{i,k}$. $hX = \frac{hI}{4}$; $wX = \frac{wI}{4}$. $J_{hj \times wJ}^{i,k}$: Matriz auxiliar. $hJ = \frac{hI}{2}$; $wJ = \frac{wI}{2}$. $\mathbf{y}_{hs \times ws}^{i,k,t}$: Subimagen t de $Y^{i,k}$. $\mathbf{x}^{i,k,t}$: Vector columna correspondiente a $\mathbf{y}^{i,k,t}$. X : Conjunto de vectores $\mathbf{x}^{i,k,t}$. $X = \{ \mathbf{x}^{i,k,t} \mid \mathbf{x}^{i,k,t} \in \mathbb{R}^{hs \cdot ws}; i = 1, \dots, Ns; k = 1, \dots, K; t = 1, \dots, T \}$.
para $i = 1, \dots, Ns$ hacer para $k = 1, \dots, K$ hacer Extraer sub-banda Paso-bajas/Paso-bajas de $I^{i,k} \rightarrow J^{i,k}$ Extraer sub-banda Paso-bajas/Paso-bajas de $J^{i,k} \rightarrow Y^{i,k}$ para $t = 1, \dots, T$ hacer Extraer subimagen t de $Y^{i,k} \rightarrow \mathbf{y}^{i,k,t}$ Ordenar los coeficientes de $\mathbf{y}^{i,k,t}$ en forma de vector columna $\rightarrow \mathbf{x}^{i,k,t}$ fin para fin para devolver X

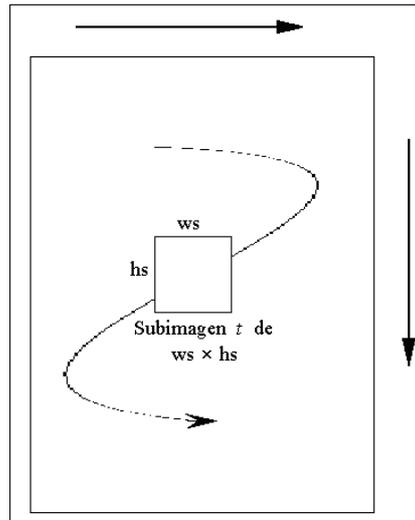


Figura 4.2. Barrido de la versión reducida de una imagen de entrenamiento para obtención de subimágenes.

Tipo de subimagen	Descripción	T	Ejemplo
1	4x4 sin traslape	35	
2	4x4 con 50 % de traslape	130	
3	8x8 sin traslape	6	
4	8x8 con 50 % de traslape	24	
5	23x1	28	
6	23x2 sin traslape	14	
7	23x2 con 50 % de traslape	27	

Tabla 4.2. Tipos de subimagen considerados para el SRRP. Para los ejemplos de la cuarta columna, el área sombreada representa el área de traslape.

4.1.3 Obtención de secuencias de observación de entrenamiento

Como se mencionó en la sección 2.1.2 se emplearán HMMs discretos, cuyo entrenamiento requiere de secuencias de observación de valores discretos. Así, el objetivo de este proceso es obtener, para cada imagen de entrenamiento, una secuencia de observación que la represente. En esta parte se hace uso de algunos conceptos de Cuantificación Vectorial, debido a que con este método es posible codificar los vectores reales resultantes de la extracción de características en un conjunto finito de valores enteros, por medio de un codebook previamente definido. Para más detalles consultar el Apéndice B.

Para el SRRP, la codificación mencionada en el párrafo anterior se plantea de la siguiente manera: a los vectores correspondientes a las subimágenes t de todas las imágenes de entrenamiento se les aplica el algoritmo de agrupamiento k-medias para obtener un conjunto de M vectores de código, el cual será el codebook correspondiente a dichas subimágenes y permitirá obtener el valor t de las secuencias de observación de cada imagen de entrenamiento. Considerando múltiples secuencias de observación para el sujeto i de la BDR, su conjunto de secuencias de observación de entrenamiento tiene la siguiente forma:

$$O^i = \{O^{i,k}; k = 1, \dots, K\}; i = 1, \dots, Ns \quad (4.1)$$

donde

$$O^{i,k} = \{o_t^{i,k} | o_t^{i,k} \in \{0, \dots, M-1\}; t = 1, \dots, T\} \quad (4.2)$$

lo que implica generar un total de T codebooks. En el caso en el que $M = Ns \cdot K$, no se aplicará el algoritmo k-medias, ya que dicho valor es equivalente al total de subimágenes t que se pueden extraer. Por tanto, este valor se toma como el máximo para M. El Algoritmo 4.2 muestra este proceso.

Algoritmo 4.2. Obtención de secuencias de observación de entrenamiento

X : Conjunto de vectores $\mathbf{x}^{i,k,t}$, obtenidos en el proceso de extracción de características.

$$X = \{\mathbf{x}^{i,k,t} | \mathbf{x}^{i,k,t} \in \mathbb{R}^{hs \cdot ws}; i = 1, \dots, Ns; k = 1, \dots, K; t = 1, \dots, T\}.$$

C : Conjunto de codebooks. $C = \{C^t; t = 1, \dots, T\}; C^t = \{\mathbf{c}^{t,i} | \mathbf{c}^{t,i} \in \mathbb{R}^{hs \cdot ws}; i = 0, \dots, M-1\}$.

D^t : Conjunto de vectores $\mathbf{x}^{i,k,t}$ correspondientes a todas las subimágenes t .

$$D^t = \{\mathbf{x}^{i,k,t}; i = 1, \dots, Ns; k = 1, \dots, K\}.$$

\mathcal{E}_t : Codificador vectorial asociado al codebook C^t .

O : Conjunto de secuencias de observación de entrenamiento. $O = \{O^i; i = 1, \dots, Ns\}$.

```

para  $t = 1, \dots, T$  hacer
  para  $i = 1, \dots, Ns$  hacer
    para  $k = 1, \dots, K$  hacer
       $D^t = D^t \cup \{\mathbf{x}^{i,k,t}\}$ 
    fin para
  fin para
  Aplicar algoritmo k-medias a  $D^t$  para M clases  $\rightarrow C^t$ 
fin para
para  $i = 1, \dots, Ns$  hacer
  para  $k = 1, \dots, K$  hacer
  para  $t = 1, \dots, T$  hacer
     $o_t^{i,k} = \mathcal{E}_t(\mathbf{x}^{i,k,t})$ 
     $O^{i,k} = O^{i,k} \cup \{o_t^{i,k}\}$ 
  fin para
   $O^i = O^i \cup \{O^{i,k}\}$ 
fin para
 $O = O \cup \{O^i\}$ 
fin para
devolver  $C, O$ 

```

4.1.4 Obtención de HMMs entrenados

Una vez que se cuenta con las secuencias de observación de entrenamiento para cada sujeto de la BDR, se procede a entrenar su HMM correspondiente, a través del Algoritmo Baum-Welch, descrito en la sección 2.3.3. Los HMMs resultantes de este proceso serán los que mejor representen a cada sujeto, es decir, los que obtengan la mayor probabilidad de emitir las secuencias de observación de entrenamiento dadas. Dichos modelos son los que se ocuparán en las pruebas de reconocimiento, que se describen en la sección 4.2.

Retomando la notación ocupada en el Capítulo 2 para representar un HMM, en la Figura 4.3 se ilustra el proceso de obtención de HMMs entrenados para los N_s sujetos de la BDR a reconocer. Para el sujeto i , λ_i y $\bar{\lambda}_i$ denotan el HMM inicial y el entrenado, respectivamente.

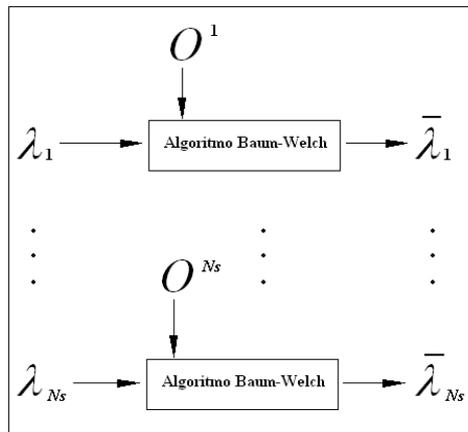


Figura 4.3. Obtención de HMMs entrenados para los N_s sujetos a reconocer.

4.2 Etapa de Reconocimiento

En la Figura 4.4 se muestran los procesos pertenecientes a la etapa de Reconocimiento del SRRP. Como se puede ver, existe una gran similitud con el esquema de la Figura 4.1, excepto que ahora se considera la parte de la BDR reservada para pruebas de reconocimiento. Al igual que en [4], [18] y [26], las $N_p - K$ imágenes restantes de cada sujeto serán las que se destinen para esta etapa. En las siguientes secciones se tratan cada uno de sus procesos.

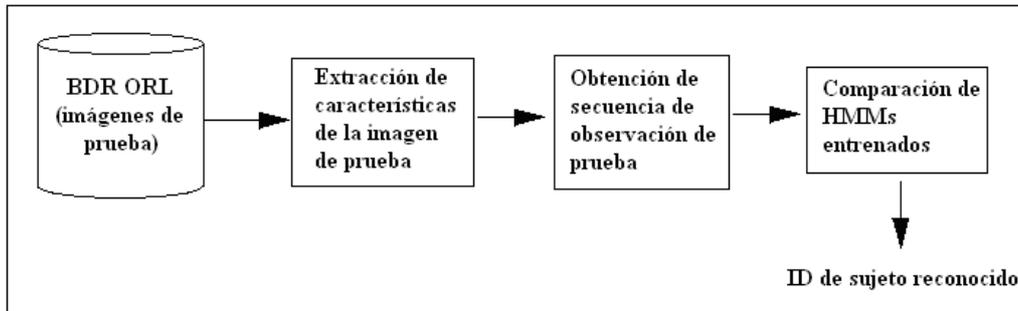


Figura 4.4. Procesos correspondientes a la etapa de Reconocimiento del SRRP.

4.2.1 Extracción de Características de la Imagen de Prueba

Dada una imagen de prueba, es decir, una imagen a reconocer, este proceso consiste en aplicar los filtros paso-bajas de descomposición, para obtener una versión reducida de dicha imagen y extraer sus subimágenes, de la misma forma que se describe en la sección 4.1.2. El resultado serán un conjunto de T vectores de longitud $h_s \cdot w_s$.

4.2.2 Obtención de la Secuencia de Observación de Prueba

De manera similar a como se generaron las secuencias de observación de entrenamiento, en este proceso, se realiza una codificación de los vectores que se obtuvieron en el proceso anterior, utilizando los mismos codebooks mencionados en la sección 4.1.3, con el fin de formar la secuencia O^p a la cual se le llamará secuencia de observación de prueba, definida como:

$$O^p = \{o_t^p | o_t^p \in \{0, \dots, M-1\}; t=1, \dots, T\} \quad (4.3)$$

4.2.3 Comparación de HMMs

Con una secuencia de observación de prueba dada y el conjunto de HMMs entrenados, se calcula para el HMM i el valor de $\ln[P(O^p | \bar{\lambda}_i)]$, con $i = 1, \dots, N_s$. Al final se puede determinar a qué sujeto de la BDR corresponde la imagen de prueba en cuestión, obteniendo el índice sr tal que

$$sr = \arg \max_{i=1}^{N_s} \{\ln[P(O^p | \bar{\lambda}_i)]\} \quad (4.4)$$

como se ejemplifica en la Figura 4.5.

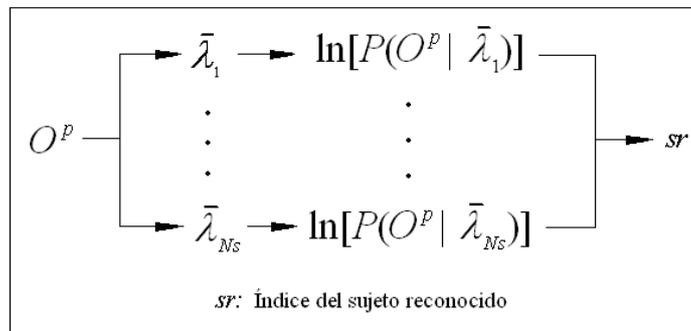


Figura 4.5. Obtención del sujeto reconocido por el SRRP.

4.3 Descripción de la interfaz gráfica del SRRP

En esta sección se describe brevemente la interfaz gráfica desarrollada para el SRRP. Dicha interfaz, junto con todos los algoritmos usados en este sistema, se implementaron con el lenguaje de programación C++. La ventana principal consta de una barra de menús, cuyas características se muestran a continuación, además de la información que proporcionan.

Menú Procesos. Este menú contiene 3 opciones: la primera, **Entrenamiento** (ver Figura 4.6), lleva a cabo la etapa de entrenamiento del sistema, explicada en la sección 4.1.

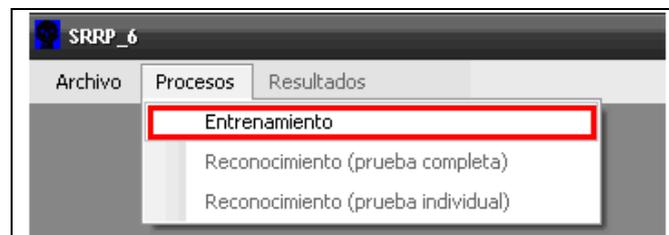


Figura 4.6. Menú Procesos del SRRP, opción: Entrenamiento.

En la Figura 4.7 a) se observa la ventana correspondiente a esta opción, en la que se introducen los valores a utilizar para los parámetros que se definieron en la sección 4.1.1. El botón **Preparar** se encarga de organizar las imágenes de entrenamiento y de prueba de la BDR de acuerdo al valor de K y, cuando esto se ha hecho, se despliega un mensaje como el de la Figura 4.7 b) informando que el SRRP está listo para el entrenamiento.

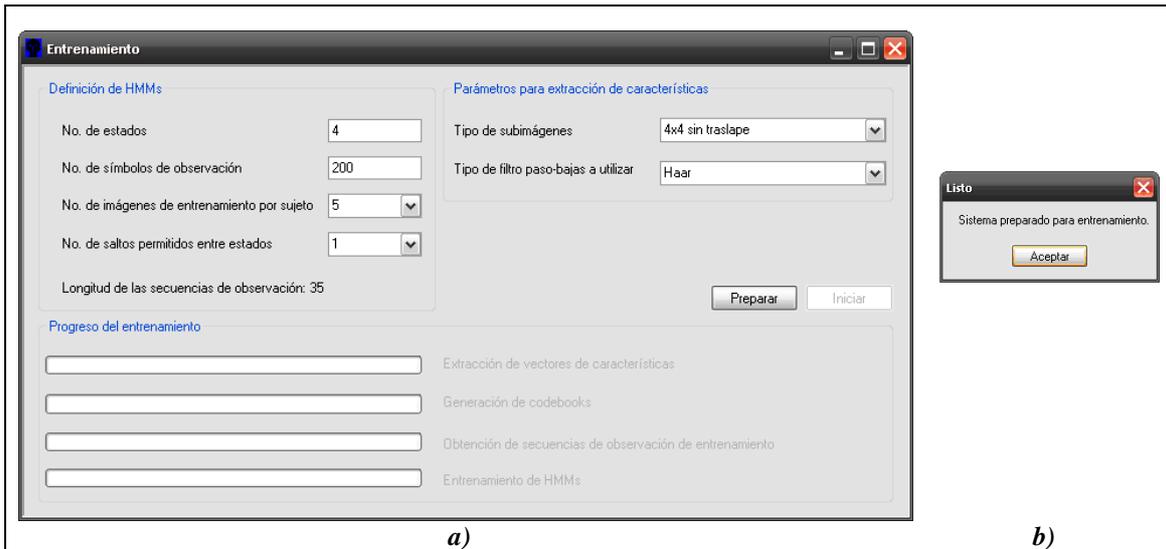


Figura 4.7. a) Ventana de Entrenamiento. b) Mensaje de preparación para el entrenamiento.

El botón **Iniciar** da comienzo a los procesos propios de esta etapa, al término de la cual la ventana de la Figura 4.7 a) se ve como en la Figura 4.8 a), acompañada de un mensaje de término del entrenamiento similar al de la Figura 4.8 b), en el que se indica el tiempo total en segundos requerido para éste.

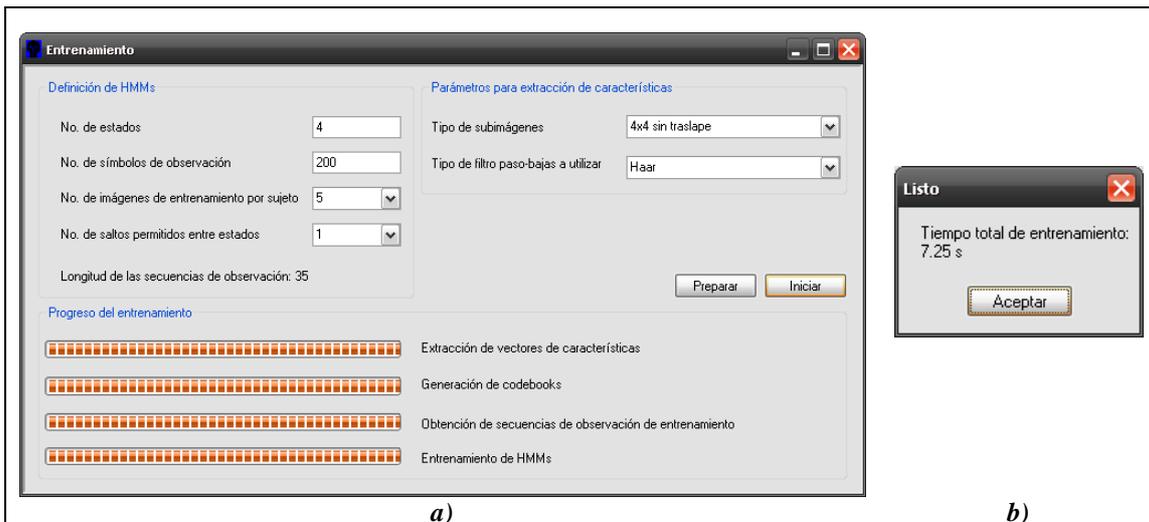


Figura 4.8. a) Ventana de Entrenamiento, al finalizar los procesos correspondientes. b) Mensaje de término del entrenamiento.

Es importante notar que, si no se tienen datos de un entrenamiento previo, la segunda y tercera opción del menú Procesos aparecerán deshabilitadas. La segunda opción, **Reconocimiento (prueba completa)** (ver Figura 4.9), realiza los procesos de reconocimiento, considerando todas las imágenes de prueba disponibles de la BDR ORL.

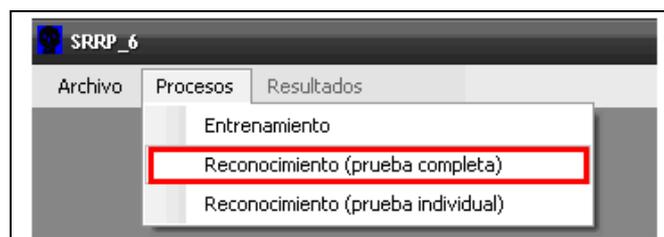


Figura 4.9. Menú Procesos del SRRP, opción: Reconocimiento (prueba completa).

Al elegirse esta opción aparece una ventana como la de la Figura 4.10 a), en la que el botón **Iniciar** hace que se empiecen a analizar todas las imágenes de prueba. La Figura 4.10 b) da un ejemplo del mensaje que indica el fin de dicho análisis. Este mensaje contiene el número de sujetos reconocidos correctamente, el número de imágenes de prueba, el tiempo promedio de reconocimiento en segundos y el porcentaje de reconocimiento global para la BDR, calculado con la ecuación (4.5).

$$\text{porcentaje de reconocimiento global} = \frac{nsr}{Ns \cdot (Nps - K)} \times 100 \quad (4.5)$$

donde *nsr* hace referencia al número de sujetos reconocidos correctamente.



Figura 4.10. a) Ventana correspondiente a la opción Reconocimiento (prueba completa). b) Mensaje de término de Reconocimiento (prueba completa).

La tercera opción del menú Procesos, **Reconocimiento (prueba individual)** (ver Figura 4.11), realiza los procesos de reconocimiento pero sólo para una imagen de prueba dada.

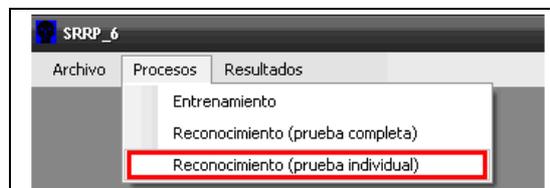


Figura 4.11. Menú Procesos del SRRP, opción Reconocimiento (prueba individual).

Esta opción despliega un cuadro de diálogo en el que se puede buscar la imagen a reconocer (ver Figura 4.12), con la condición de que se encuentre dentro del directorio llamado **Imágenes de Prueba**, el cual se crea como parte de la preparación del sistema para el entrenamiento.

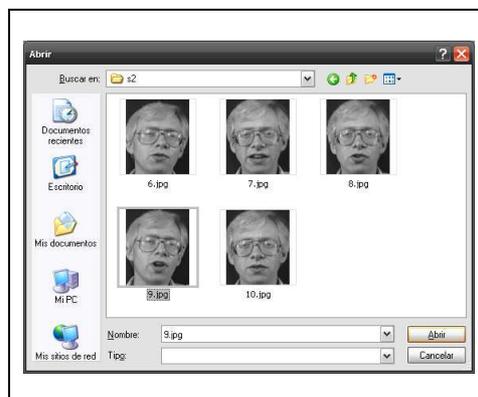


Figura 4.12. Cuadro de diálogo para elección de la imagen a reconocer.

Una vez seleccionada la imagen de prueba, se presenta una ventana con dicha imagen. Un ejemplo de la apariencia de esta ventana se aprecia en la Figura 4.13 a), en la que el botón Reconocer arroja la imagen reconocida por el SRRP (Figura 4.13 b)), junto con un mensaje de término que indica el tiempo de reconocimiento en segundos (Figura 4.13 c)).

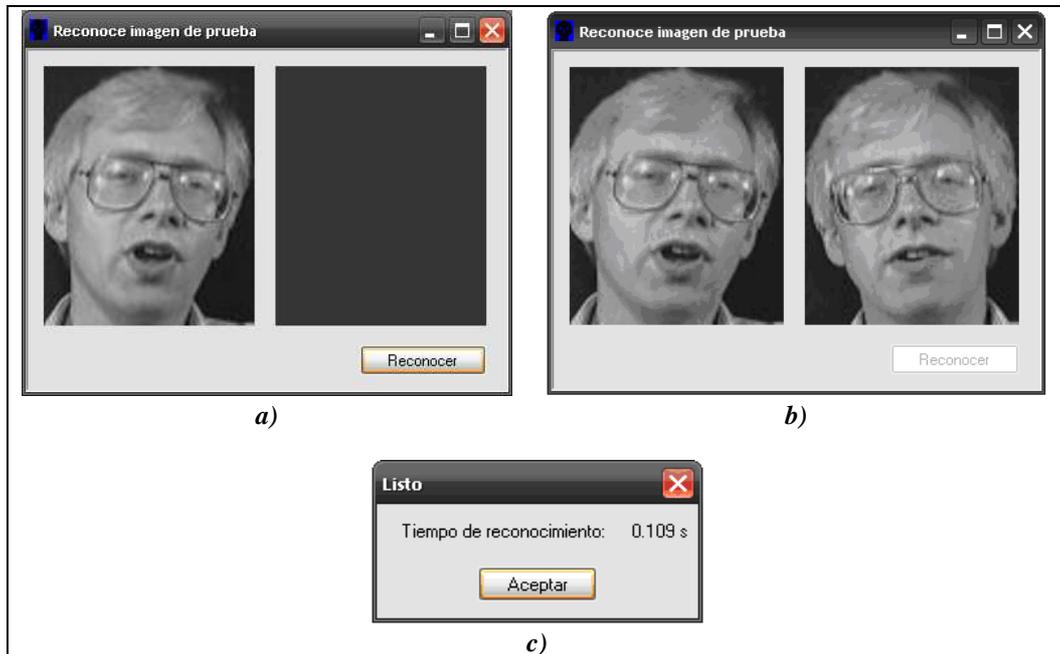


Figura 4.13 a) Ventana de Reconocimiento (prueba individual). b) Ventana de Reconocimiento (prueba individual) al final del reconocimiento. c) Mensaje de término de Reconocimiento (prueba individual).

Menú Resultados. Este menú tiene 2 opciones: la primera opción, **Resultados de Reconocimiento (prueba completa)** (Figura 4.14 a)), proporciona la ventana de la Figura 4.14 b), en la que es posible comparar las imágenes de prueba usada en la opción Reconocimiento (prueba completa) con las imágenes reconocida por el sistema.

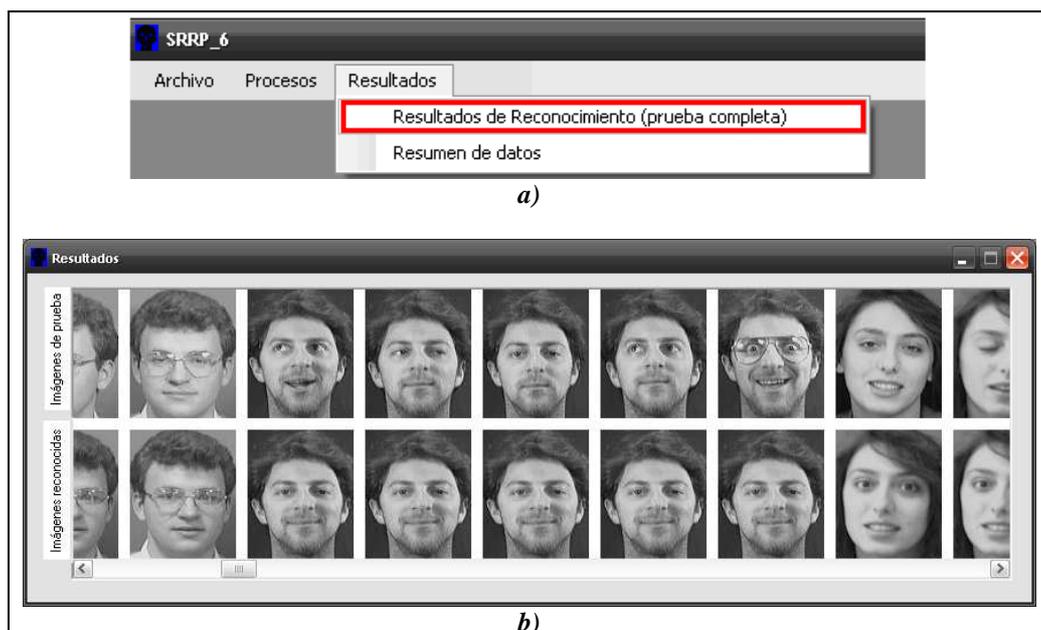


Figura 4.14. a) Menú Resultados, opción Resultados de Reconocimiento (prueba completa). b) Ventana de Resultados (prueba completa).

La opción **Resumen de datos** de este mismo menú (Figura 4.15 a)), permite ver los resultados generales que el SRRP obtuvo después de ejecutar la opción de Reconocimiento (prueba completa), los cuales se listan en una Ventana como la de la Figura 4.15 b).

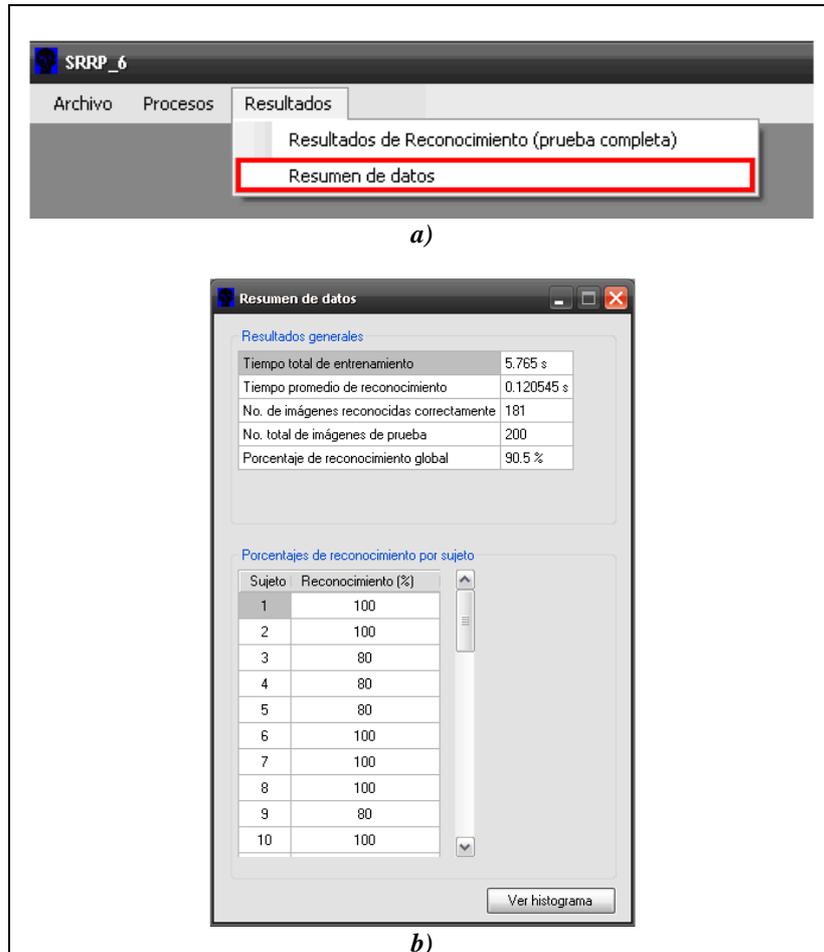


Figura 4.15 a) Menú Resultados, opción Resumen de datos. b) Ventana de Resumen de datos.

En la ventana anterior, los resultados que se pueden consultar son el tiempo total de entrenamiento, el tiempo promedio de reconocimiento, el número tanto de imágenes de prueba como de imágenes reconocidas correctamente por el sistema y el porcentaje de reconocimiento global. Se incluye también una tabla en la que es posible ver el porcentaje de reconocimiento de cada sujeto analizado de la BDR. Para el sujeto i , este porcentaje dependerá sólo del número de imágenes de prueba por sujeto y del número de éstas reconocidas correctamente. Si a este último número se le denota como nsr_i , entonces, el porcentaje de reconocimiento del sujeto i se define como se indica en la ecuación (4.6).

$$\text{porcentaje de reconocimiento del sujeto } i = \frac{nsr_i}{Nps - K} \times 100 \quad (4.6)$$

Además, en esta misma ventana se encuentra un botón con la leyenda **Ver histograma** que, al pulsarse, ofrece un histograma de los porcentajes de reconocimiento de cada sujeto (Figura 4.16).

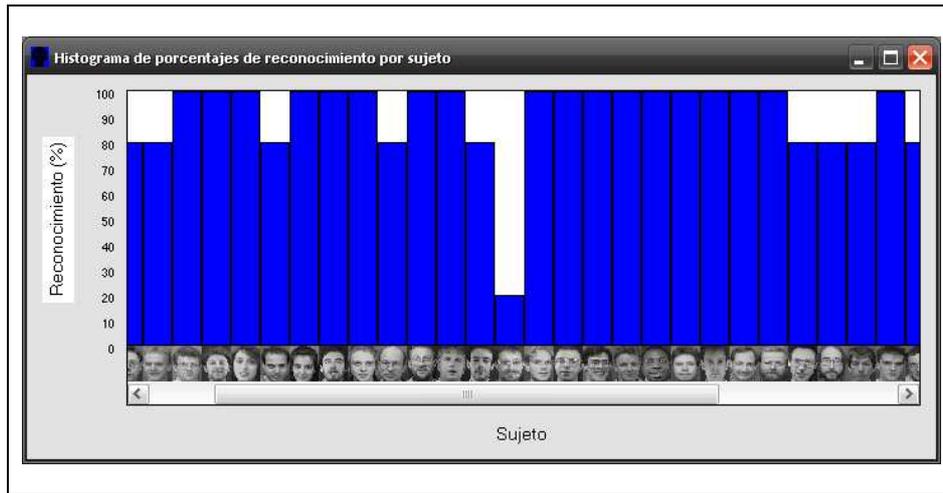


Figura 4.16. Ventana del histograma de porcentajes de reconocimiento por sujeto.

Menú Archivo. La opción **Información Adicional** de este menú (Figura 4.17 a) despliega el mensaje que aparece en la Figura 4.17 b), dentro del cual se incluye un enlace a la página Web mostrada en la Figura 4.17 c). En dicha página se encuentra toda la información relacionada con la BDR ORL [3]. Finalmente, la opción **Salir** del menú Archivo termina la ejecución del SRRP.

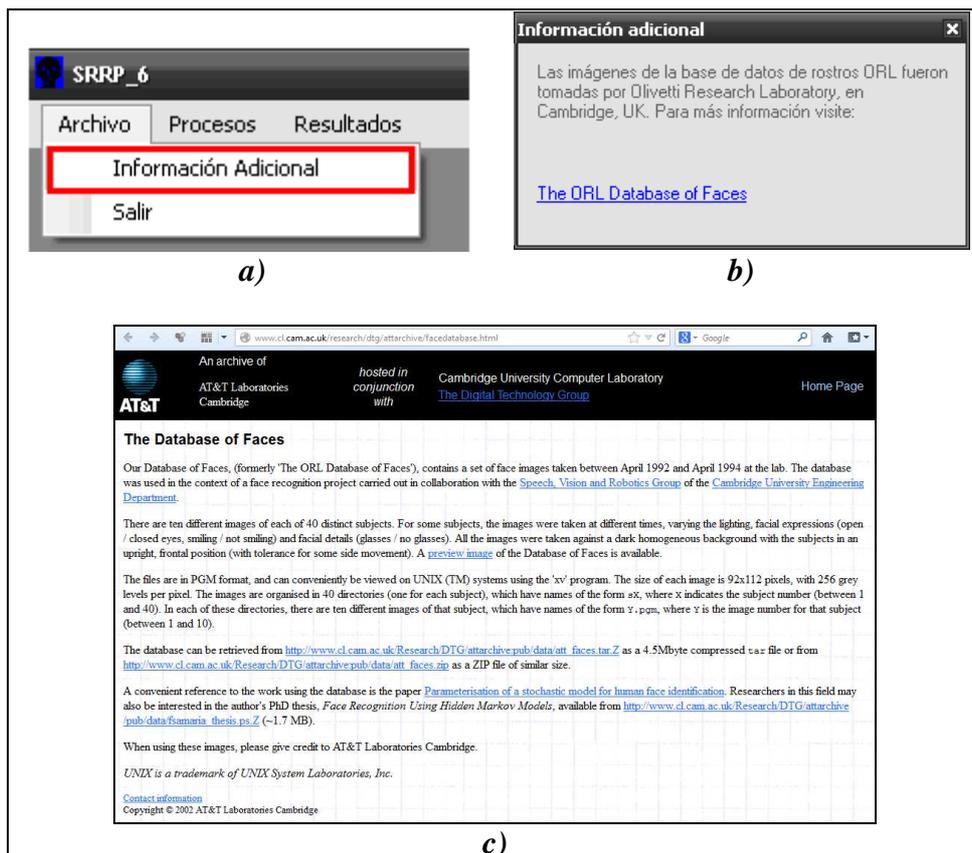


Figura 4.17. a) Menú Archivo, opción Información Adicional. b) Mensaje de Información Adicional. c) Página Web de la BDR ORL [3].

Capítulo 5. Pruebas y Resultados

En este capítulo se muestran los resultados obtenidos en relación al desempeño del SRRP, además de describir los aspectos más relevantes de las pruebas que condujeron a dichos resultados. Con el fin de acotar el número de parámetros que se consideran variables para las pruebas de reconocimiento, se hicieron las siguientes consideraciones: el valor de K se fijó en 5, por lo que las 5 imágenes restantes de cada sujeto fueron utilizadas para pruebas de reconocimiento [26], [4]; se trabaja con la BDR ORL completa, es decir, se contemplan 40 sujetos a reconocer; para todas las pruebas, $\Delta=1$ y, en cuanto a N y M, se suponen los siguientes valores: N = 2, 4, 8, 16, 32 y 64; M = 40, 80, 120, 160 y 200; el número máximo de iteraciones del algoritmo Baum-Welch se fijó en 50.

Los resultados que se pretenden analizar son principalmente los porcentajes de reconocimiento global para la BDR ORL, el tiempo total de entrenamiento y el tiempo promedio de reconocimiento de cada imagen de prueba. Estos resultados se dividen en 7 grupos de pruebas que corresponden a los 7 tipos de subimagen descritos en el Capítulo 4. La información anterior también se acompaña de los máximos porcentajes de reconocimiento global por subimagen, alcanzados con los tipos de filtros mencionados también en el Capítulo 4. Se tratarán con más detenimiento las pruebas que hayan conducido a los porcentajes de reconocimiento global más altos de todos los 7 grupos, discutiendo aspectos como las características de sus respectivos HMMs y el porcentaje de reconocimiento de cada sujeto de la BDR en cuestión. Las pruebas se realizaron en un equipo con sistema operativo Microsoft Windows XP, con un procesador Pentium 4 de 2 GHz.

5.1 Resultados de las pruebas de reconocimiento hechas con el SRRP por tipo de subimagen

Las secciones 5.1.1 a 5.1.7 contienen los resultados de las pruebas de reconocimiento con el SRRP de acuerdo al tipo de subimagen seleccionado. Cada sección incluye también un análisis de los porcentajes de reconocimiento global máximos, para los cuales se prefiere estudiar aquellas pruebas con un valor de N menor a otros valores que también hayan llevado a porcentajes iguales. En las Tablas 5.1 a 5.5, 5.7 a 5.11, 5.13 a 5.17, 5.19 a 5.23, 5.25 a 5.29, 5.31 a 5.35 y 5.37 a 5.41, las abreviaturas usadas son: Rec. (porcentaje de reconocimiento global), tentr (tiempo total de entrenamiento) y trec (tiempo promedio de reconocimiento). Para las Tablas 5.1 a 5.41, Filtro PB denota el tipo de filtro paso-bajas utilizado.

5.1.1 Resultados de las pruebas de reconocimiento para el tipo de subimagen 1

Las Tablas 5.1 a 5.5 listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 1 y las Figuras 5.1 a 5.5 muestran sus respectivas gráficas.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	63	3.609	0.042	55.5	4.797	0.049	62.5	2.578	0.031
	4	70.5	3.922	0.049	67.5	5.125	0.055	69.5	2.922	0.039
	8	73.5	5.594	0.064	74.5	6.484	0.072	76.5	4.766	0.059
	16	78.5	12.782	0.109	77.5	14.578	0.117	79.5	12.172	0.102
	32	81	30.906	0.227	78.5	31.797	0.237	82	30.172	0.216
	64	65.5	29.266	0.635	62	30.297	0.578	67	28.281	0.589

Tabla 5.1. Resultados de las pruebas de reconocimiento con M = 40 y tipo de subimagen 1.

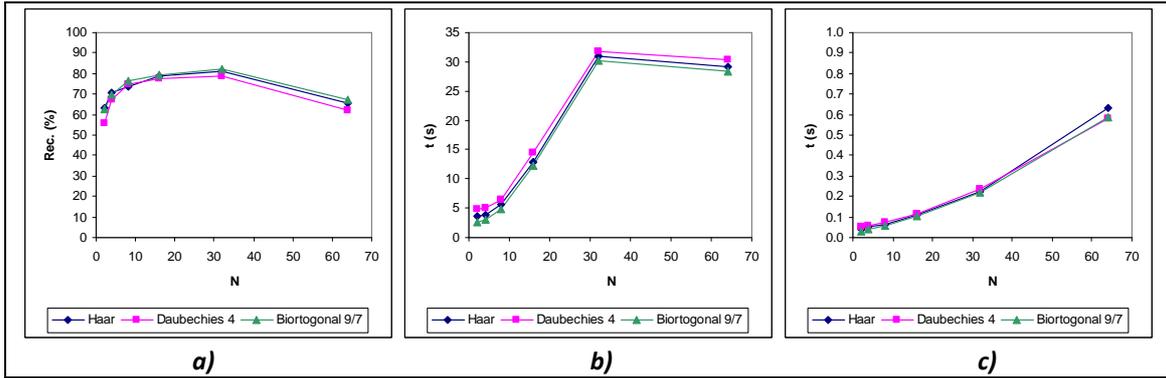


Figura 5.1. Gráficas de los resultados de la Tabla 5.1: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	63.5	4.125	0.047	68	5.36	0.052	65.5	3.079	0.039
	4	69	4.735	0.063	75.5	6.016	0.065	73	3.656	0.058
	8	76.5	7.297	0.094	78	8.515	0.1	77	6.578	0.086
	16	80	16.687	0.17	78.5	18.172	0.172	81.5	15.141	0.158
	32	81.5	38.375	0.371	81	39.937	0.341	82.5	37.328	0.378
	64	69	31.468	0.826	72.5	32.688	0.927	67.5	30.406	0.85

Tabla 5.2. Resultados de las pruebas de reconocimiento con M = 80 y tipo de subimagen 1.

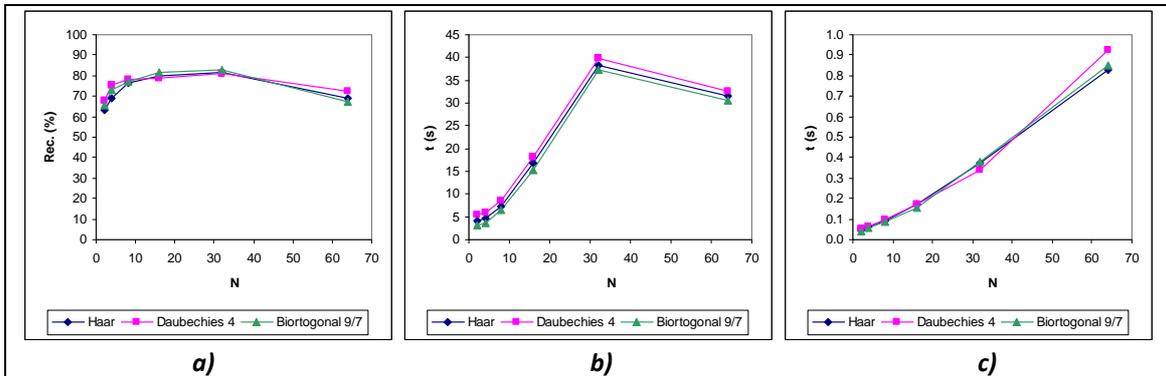


Figura 5.2. Gráficas de los resultados de la Tabla 5.2: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	69.5	4.578	0.057	70	5.672	0.061	65.5	3.437	0.049
	4	73.5	5.578	0.078	78	6.766	0.083	75.5	4.625	0.068
	8	75.5	8.828	0.123	77.5	9.859	0.126	77	7.953	0.113
	16	80	19.547	0.219	82	20.515	0.225	82	17.625	0.21
	32	82	42.063	0.462	85.5	41.578	0.503	85.5	41.922	0.493
	64	70	33.125	1.069	70	33.687	1.068	70.5	33.218	1.029

Tabla 5.3. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 1.

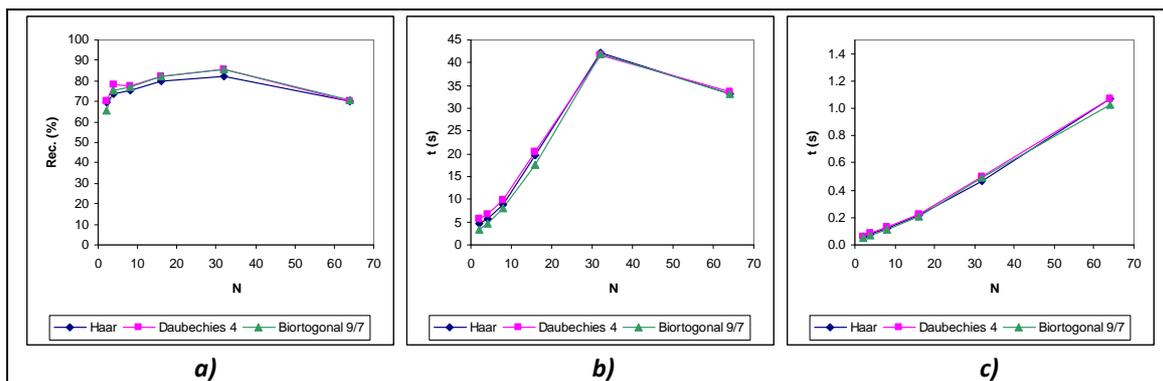


Figura 5.3. Gráficas de los resultados de la Tabla 5.3: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	77	5.109	0.065	76	6.063	0.067	77.5	3.735	0.056
	4	79	6.157	0.091	81.5	7.422	0.098	81	4.937	0.089
	8	81.5	10.234	0.158	82	11.062	0.169	81	8.516	0.157
	16	85	21.921	0.299	84	21.907	0.296	81.5	20.531	0.265
	32	86.5	45.828	0.554	84	40.688	0.591	83	45.296	0.6
	64	77.5	34.297	1.315	78.5	35.266	1.372	78	33.266	1.246

Tabla 5.4. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 1.

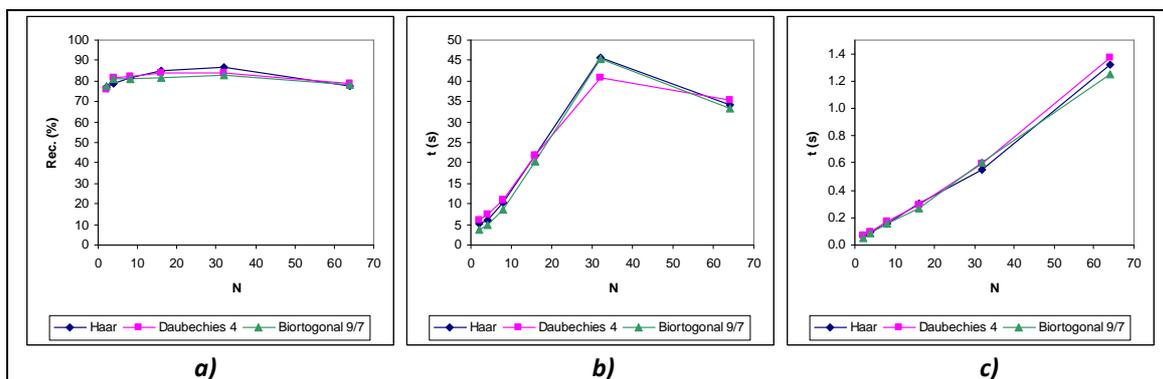


Figura 5.4. Gráficas de los resultados de la Tabla 5.4: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	90	4.047	0.067	90	5.515	0.075	88.5	3.187	0.063
	4	90	4.484	0.103	90	5.531	0.11	88.5	3.515	0.095
	8	90	4.922	0.185	90	6.031	0.181	88.5	4.032	0.166
	16	90.5	6.578	0.351	90	7.297	0.364	88.5	5.063	0.313
	32	90.5	8.891	0.656	90	9.547	0.661	88.5	7.328	0.645
	64	90.5	35.391	1.5	90	36.89	1.497	88.5	34.672	1.556

Tabla 5.5. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 1.

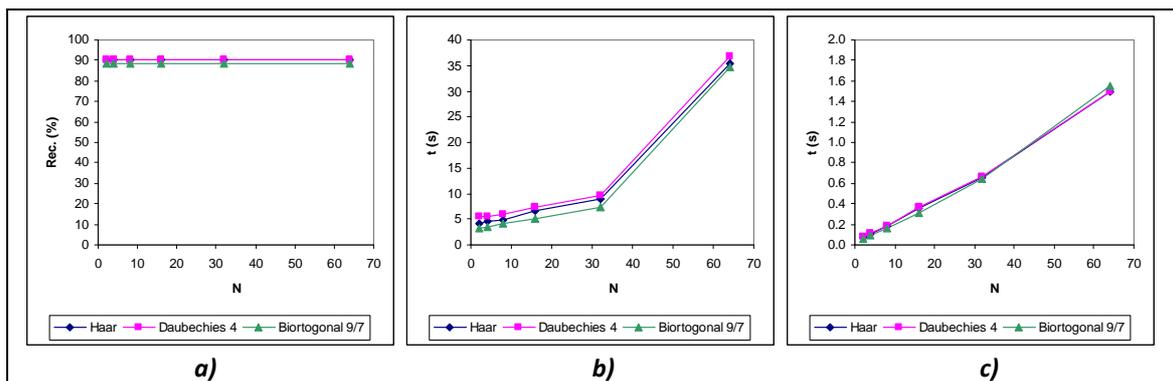


Figura 5.5. Gráficas de los resultados de la Tabla 5.5: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.6 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 1. Dichos porcentajes se presentan gráficamente en la Figura 5.6.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	81	78.5	82
80	81.5	81	82.5
120	82	85.5	85.5
160	86.5	84	83
200	90.5	90	88.5

Tabla 5.6. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 1.

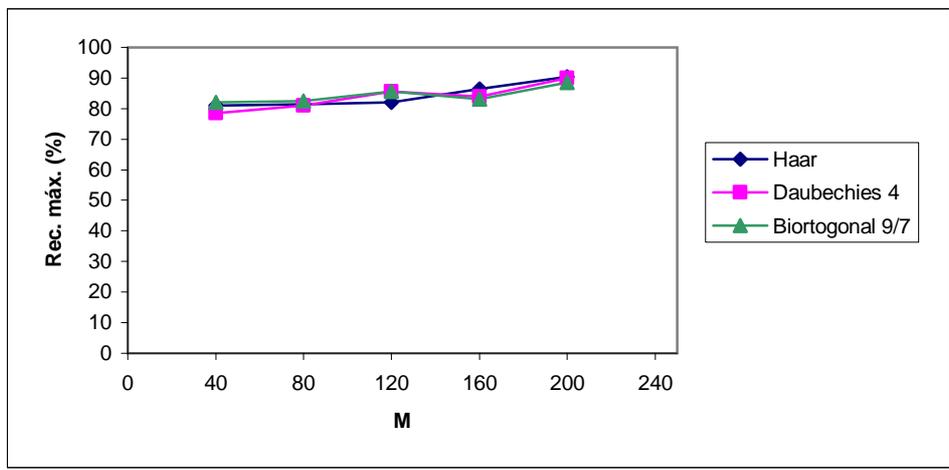


Figura 5.6. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 1 (Rec. máx.: reconocimiento global máximo).

Las Figuras 5.1 a 5.4 permiten ver que, en general, para $M \in \{40, 80, 120, 160\}$ hay un máximo en el porcentaje de reconocimiento global cuando $N = 32$ y, si $N = 64$, este porcentaje decrece. Los tiempos de entrenamiento no rebasan los 50 s. La Tabla 5.6 indica que el porcentaje de reconocimiento global más alto para este grupo de resultados fue de 90.5 %, para $M = 200$, $N = 16$ y un filtro paso-bajas de Haar. El tiempo total de entrenamiento para este porcentaje fue de 6.578 s, mientras que el tiempo promedio de reconocimiento fue de 0.351 s.

5.1.2 Resultados de las pruebas de reconocimiento para el tipo de subimagen 2

En las Tablas 5.7 a 5.11 se listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 2 y en las Figuras 5.7 a 5.11 se muestran las gráficas de estos resultados, respectivamente.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	58.5	6.5	0.041	61.5	7.859	0.047	62	5.484	0.036
4	64	8.343	0.052	67.5	10.031	0.058	65	7.859	0.044
8	72.5	16.688	0.078	72	19.344	0.082	72	18.219	0.073
16	79.5	52.234	0.141	78	57.078	0.157	77	59.031	0.139
32	81.5	226.766	0.316	78	203.984	0.329	78	200.094	0.319
64	85	794.656	0.94	82	909.75	1.016	81	847.688	1.144

Tabla 5.7. Resultados de las pruebas de reconocimiento con M = 40 y tipo de subimagen 2.

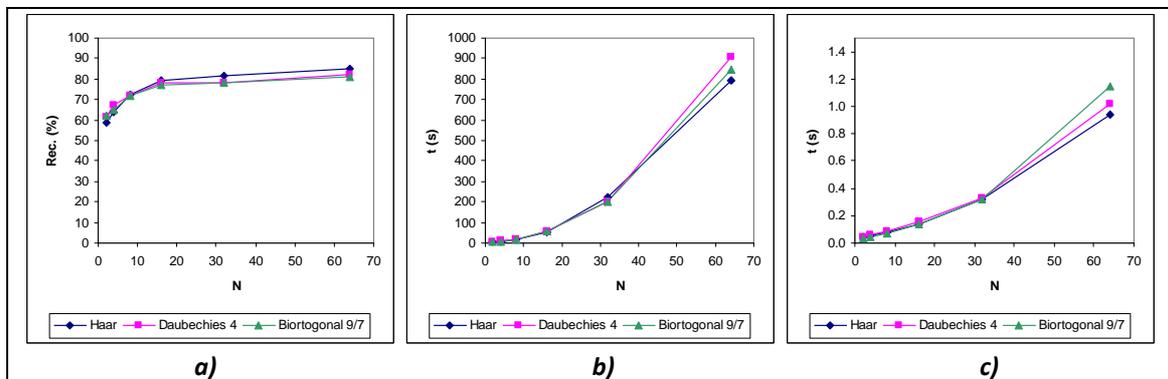


Figura 5.7. Gráficas de los resultados de la Tabla 5.7: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	65	9.296	0.049	64.5	9.297	0.056	61.5	6.922	0.042
4	71	12.844	0.065	72	15.875	0.077	71.5	10.828	0.06
8	72	27.656	0.104	75.5	30.062	0.11	76	23.109	0.097
16	78	73.094	0.197	78.5	72.828	0.208	76.5	71.781	0.205
32	80.5	308.14	0.428	81	251.235	0.428	80.5	262.907	0.415
64	83.5	860.172	1.325	83.5	852.406	1.243	84.5	913.25	1.286

Tabla 5.8. Resultados de las pruebas de reconocimiento con M = 80 y tipo de subimagen 2.

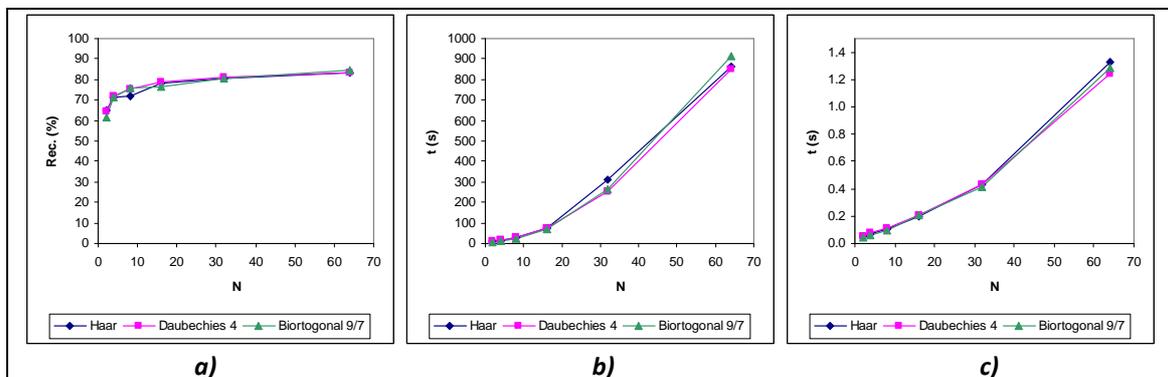


Figura 5.8. Gráficas de los resultados de la Tabla 5.8: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	64.5	11.312	0.059	68.5	10.703	0.063	67	8.797	0.051
4	71.5	17.578	0.082	75	17.485	0.089	69	14.703	0.074
8	78.5	34.594	0.134	77	35.609	0.141	76.5	34.344	0.14
16	79.5	100.282	0.264	81	100.891	0.257	79	117.25	0.246
32	82	343.562	0.522	82.5	346.938	0.543	79	365.656	0.576
64	85.5	802.578	1.53	86	840.125	1.692	85	923.578	1.7

Tabla 5.9. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 2.

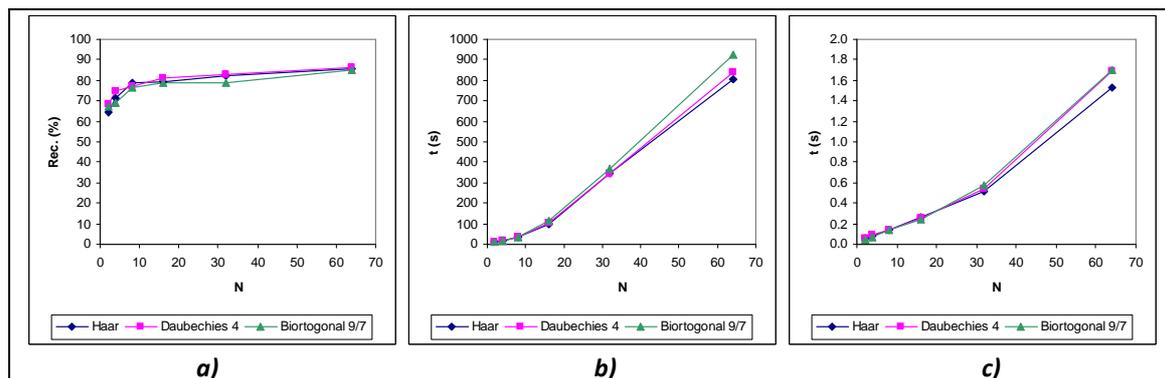


Figura 5.9. Gráficas de los resultados de la Tabla 5.9: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	79.5	10.75	0.066	79.5	13.063	0.077	79.5	9.078	0.06
4	80.5	17.375	0.096	79.5	16.891	0.1	79.5	14.672	0.09
8	82.5	35.672	0.162	82.5	36.422	0.17	82.5	40.687	0.155
16	84.5	104.86	0.301	84.5	116.703	0.312	83.5	88.766	0.296
32	87	342.078	0.688	85.5	347.609	0.705	85.5	312.906	0.746
64	87	537.406	1.992	87.5	485.578	1.908	87.5	664.953	1.706

Tabla 5.10. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 2.

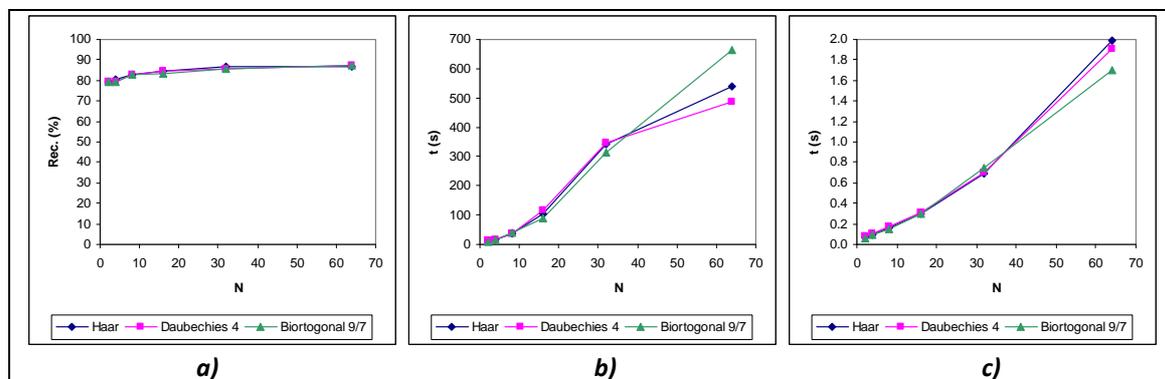


Figura 5.10. Gráficas de los resultados de la Tabla 5.10: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	91.5	7.641	0.077	90	8.938	0.083	90.5	7.187	0.073
4	91.5	8.312	0.111	90	10.156	0.118	90.5	7.453	0.103
8	91.5	9.5	0.187	90	10.797	0.192	90.5	8.718	0.18
16	91.5	12.312	0.382	90	13.375	0.348	90.5	10.953	0.392
32	91.5	18.25	0.858	90	19.875	0.729	90.5	17.735	0.808
64	91.5	45.89	1.938	90	41.203	1.854	90.5	40.563	1.88

Tabla 5.11. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 2.

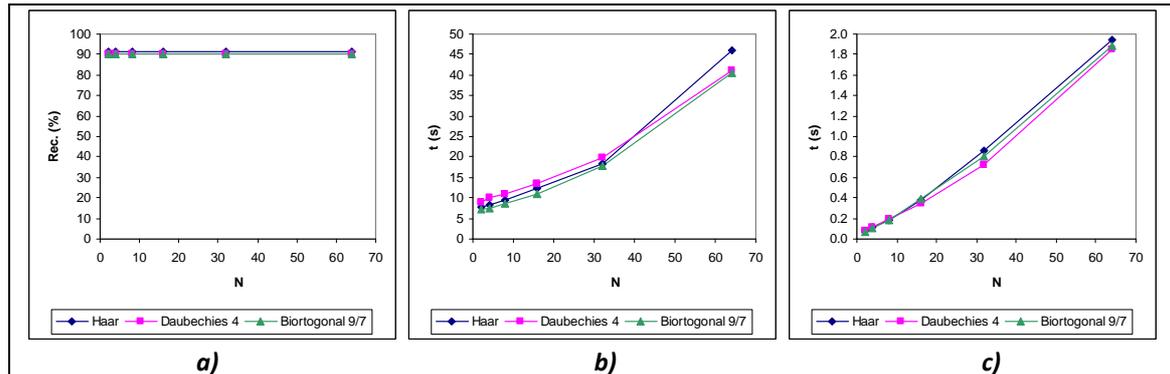


Figura 5.11. Gráficas de los resultados de la Tabla 5.11: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.12 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 2. Estos porcentajes se presentan gráficamente en la Figura 5.12.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	85	82	81
80	83.5	83.5	84.5
120	85.5	86	85
160	87	87.5	87.5
200	91.5	90	90.5

Tabla 5.12. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 2.

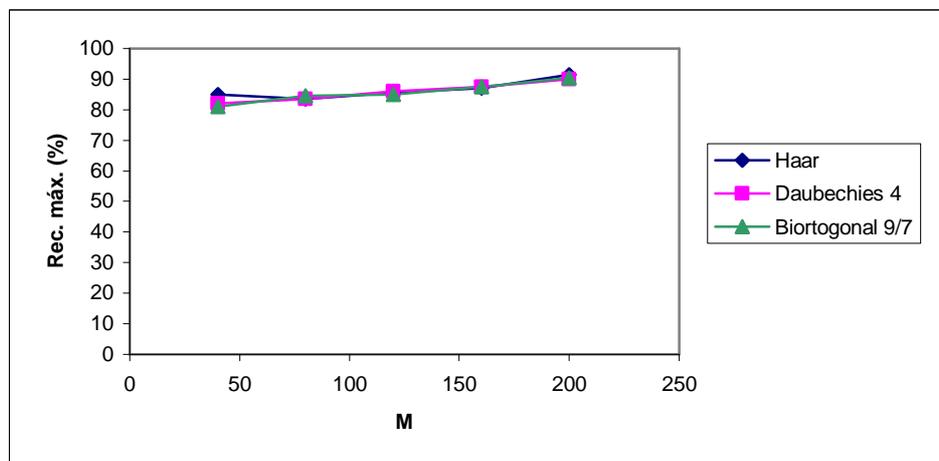


Figura 5.12. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 2 (Rec. máx.: reconocimiento global máximo).

En las Figuras 5.7 a 5.10 se observa que se llega máximo porcentaje de reconocimiento global cuando $N = 64$, si $M \in \{40, 80, 120, 160\}$. Los tiempos de entrenamiento más altos llegan a ser de alrededor de 15 minutos, debido a que el tipo de subimagen que se considera hace que el barrido de las imágenes de entrenamiento sea el más exhaustivo de todos los 7. Según la Tabla 5.12, el porcentaje de reconocimiento global más alto para este grupo de resultados fue de 91.5 %, para $M = 200$, $N = 2$ y un filtro paso-bajas de Haar. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento de este porcentaje fueron de 7.641 s y 0.077 s, respectivamente.

5.1.3 Resultados de las pruebas de reconocimiento para el tipo de subimagen 3

En las Tablas 5.13 a 5.17 se listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 3 y en las Figuras 5.13 a 5.17 se muestran las gráficas de estos resultados, respectivamente.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	56.5	3.234	0.039	57	4.406	0.047	65	2.266	0.031
	4	64	3.328	0.047	64	4.547	0.052	71.5	2.391	0.039
	8	55	3.265	0.063	58.5	4.453	0.073	61	2.359	0.056
	16	55	3.672	0.106	58.5	4.89	0.11	61	2.734	0.098
	32	55	4.656	0.212	58.5	5.969	0.241	61	3.796	0.205
	64	55	8.422	0.581	58.5	9.703	0.615	61	7.5	0.54

Tabla 5.13. Resultados de las pruebas de reconocimiento con $M = 40$ y tipo de subimagen 3.

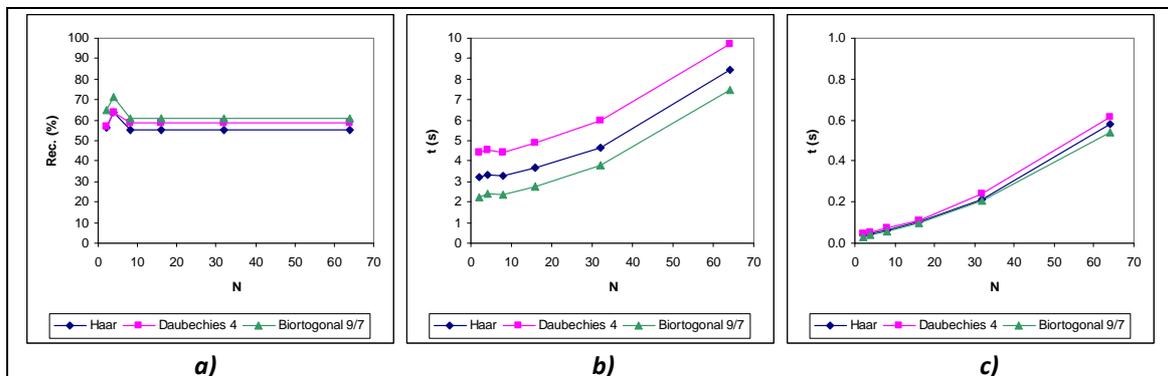


Figura 5.13. Gráficas de los resultados de la Tabla 5.13: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	69	3.547	0.047	69	4.671	0.051	66.5	2.672	0.038
	4	71	3.828	0.061	69.5	4.89	0.065	68	2.922	0.055
	8	67.5	3.641	0.098	63	4.766	0.096	65.5	2.766	0.087
	16	67.5	4.078	0.171	63	5.203	0.169	65.5	3.219	0.151
	32	67.5	5.438	0.362	63	6.719	0.336	65.5	4.578	0.346
	64	67.5	9.672	0.855	63	11.297	0.793	65.5	8.844	0.82

Tabla 5.14. Resultados de las pruebas de reconocimiento con $M = 80$ y tipo de subimagen 3.

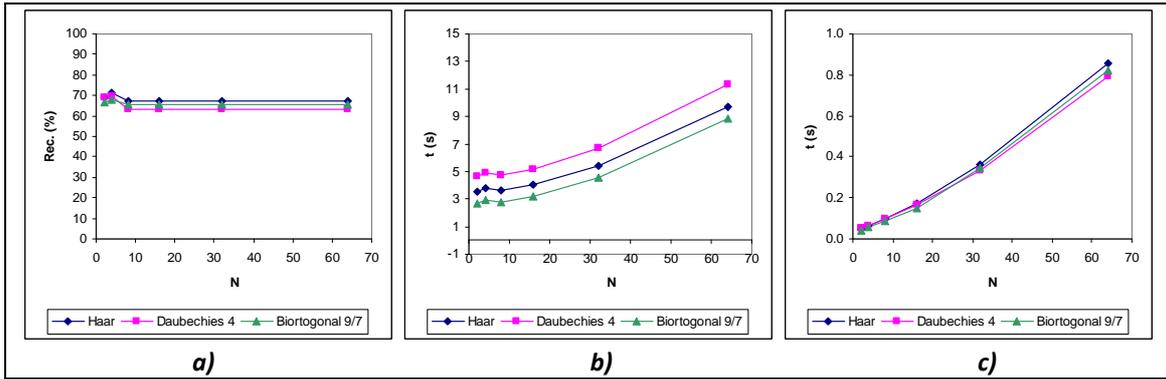


Figura 5.14. Gráficas de los resultados de la Tabla 5.14: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	72	3.938	0.053	67.5	5.109	0.061	72.5	2.969	0.046
	4	74	4.281	0.077	73.5	5.547	0.083	78	3.375	0.068
	8	66.5	4.063	0.12	65	5.985	0.165	69.5	3.141	0.121
	16	66.5	4.64	0.223	65	6.094	0.244	69.5	3.688	0.21
	32	66.5	6.656	0.466	65	7.75	0.458	69.5	5.25	0.421
	64	66.5	10.781	1.091	65	12.016	1.033	69.5	10.297	0.984

Tabla 5.15. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 3.

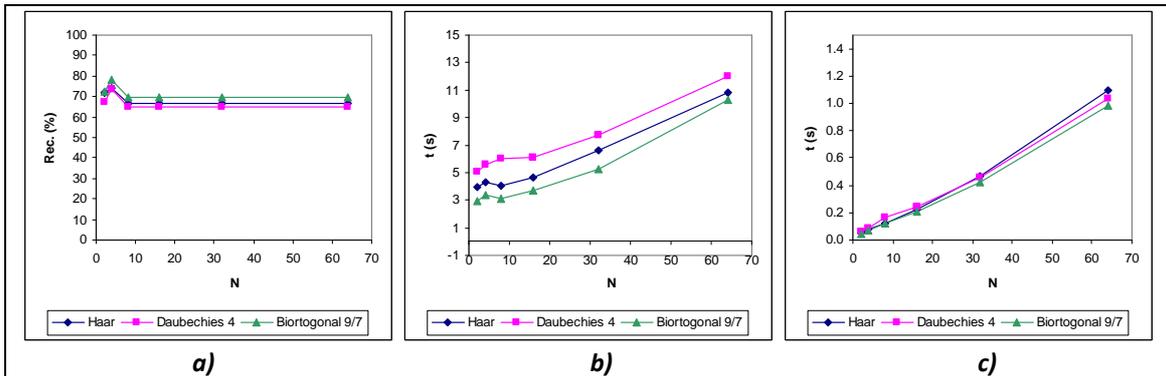


Figura 5.15. Gráficas de los resultados de la Tabla 5.15: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	75.5	4.094	0.062	74	5.375	0.065	78	3.14	0.053
	4	75	4.609	0.09	75	5.86	0.094	76.5	3.734	0.082
	8	75	4.359	0.148	72	6.359	0.154	76	3.406	0.139
	16	75	5.015	0.268	72	6.281	0.289	76	4.079	0.266
	32	75	6.828	0.594	72	8.078	0.545	76	5.828	0.54
	64	75	11.875	1.327	72	13.109	1.22	76	11.266	1.274

Tabla 5.16. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 3.

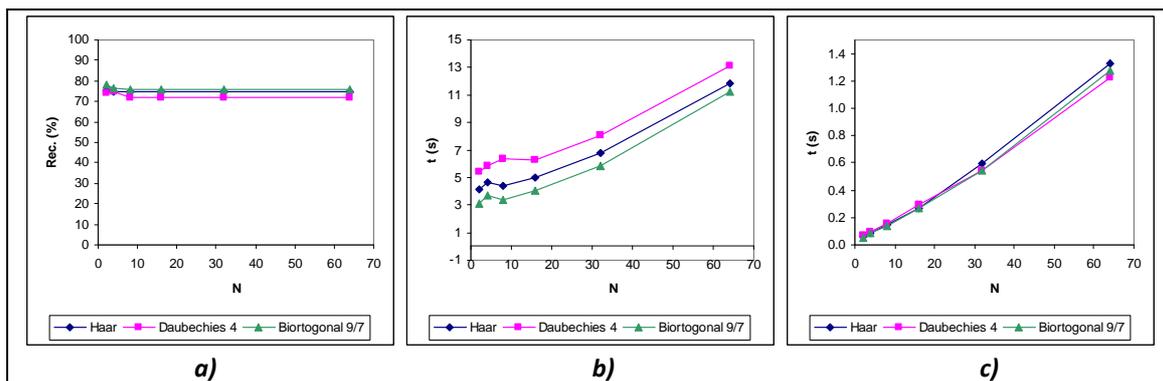


Figura 5.16. Gráficas de los resultados de la Tabla 5.16: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	76	3.5	0.067	76	4.703	0.077	75	2.61	0.062
	4	76	3.609	0.102	76	4.812	0.112	76	2.781	0.094
	8	76	3.953	0.172	76	5.125	0.178	76	3.109	0.165
	16	76	4.75	0.339	76	5.985	0.333	76	3.891	0.359
	32	76	6.75	0.722	76	8.766	0.773	76	5.953	0.737
	64	76	12.453	1.489	76	13.968	1.537	76	11.515	1.523

Tabla 5.17. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 3.

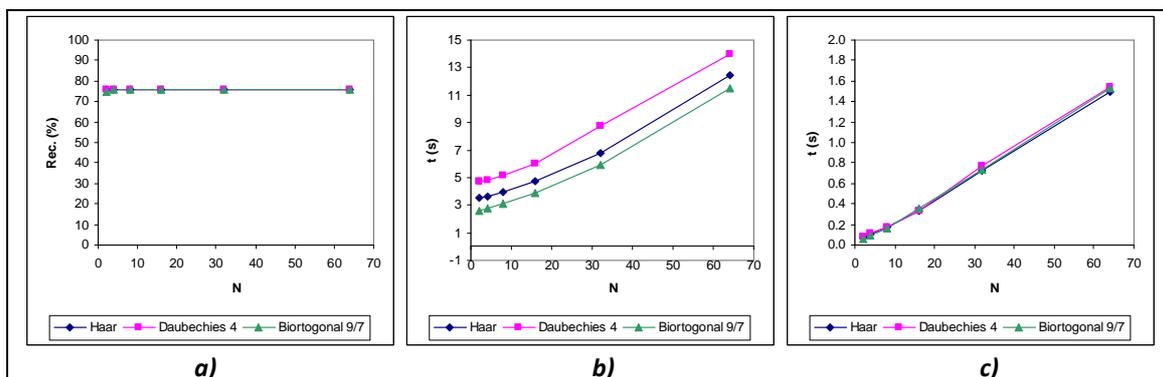


Figura 5.17. Gráficas de los resultados de la Tabla 5.17: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.18 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 3. Estos porcentajes se presentan gráficamente en la Figura 5.18.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	64	64	71.5
80	71	69.5	68
120	74	73.5	78
160	75.5	75	78
200	76	76	76

Tabla 5.18. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 3.

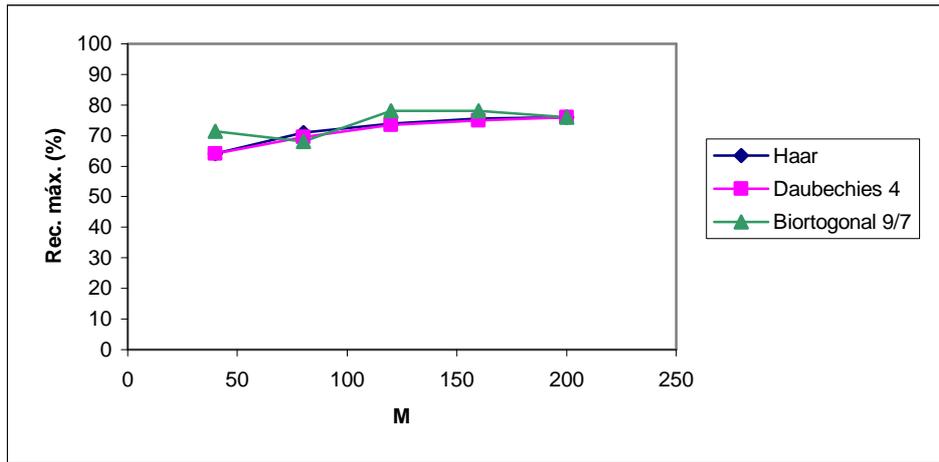


Figura 5.18. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 3 (Rec. máx.: reconocimiento global máximo).

Las Figuras 5.13 a 5.16 indican que, en general, con valores de 2 y 4 para N, se llega al porcentaje de reconocimiento global máximo, si $M \in \{40, 80, 120, 160\}$. Los tiempos de entrenamiento son menores de 14 s, aproximadamente. La Tabla 5.18 indica que el porcentaje de reconocimiento global más alto para este grupo de resultados fue de 78 %, obtenido en dos pruebas: 1) $M = 120$ y $N = 4$; 2) $M = 160$ y $N = 2$, con un filtro paso-bajas del tipo Biortogonal 9/7. Para la primera de estas pruebas, el tiempo total de entrenamiento fue de 3.375s, mientras que el tiempo promedio de reconocimiento fue de 0.068 s. Para la segunda prueba, el tiempo total de entrenamiento fue de 3.14 s y el tiempo promedio de reconocimiento fue de 0.053 s.

5.1.4 Resultados de las pruebas de reconocimiento para el tipo de subimagen 4

En las Tablas 5.19 a 5.23 se listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 4 y en las Figuras 5.19 a 5.23 se muestran las gráficas de estos resultados, respectivamente.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	59.5	4.75	0.047	67	6.203	0.047	60	3.938	0.032
	4	66.5	5.219	0.049	68.5	6.688	0.054	68.5	4.219	0.04
	8	70.5	6.594	0.064	73	8.032	0.071	72	6.125	0.058
	16	74.5	11.282	0.109	77	12.672	0.113	75.5	9.719	0.1
	32	58.5	10.078	0.234	63	11.281	0.286	59	8.515	0.253
	64	58.5	22.719	0.61	63	24.625	0.651	59	22.219	0.653

Tabla 5.19. Resultados de las pruebas de reconocimiento con M = 40 y tipo de subimagen 4.

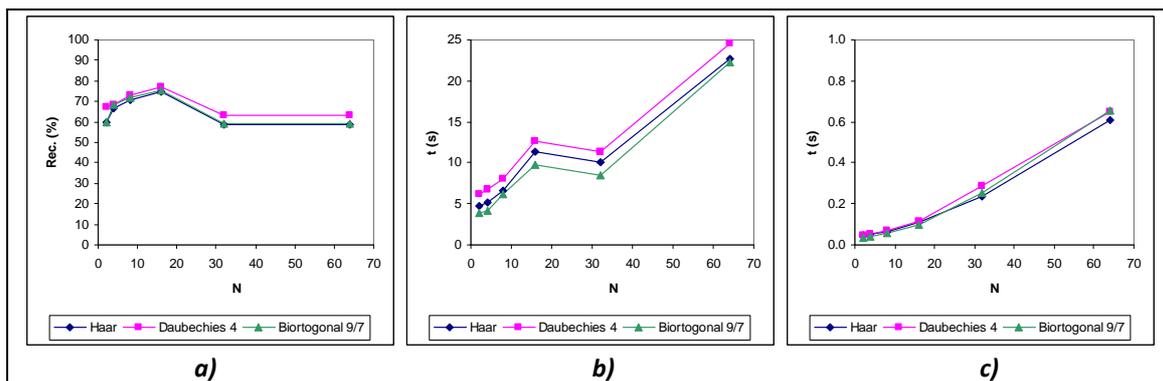


Figura 5.19. Gráficas de los resultados de la Tabla 5.19: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	73	6.391	0.047	70	7.766	0.053	69	5.141	0.04
	4	73	7.484	0.063	75.5	8.234	0.068	75	5.953	0.06
	8	79.5	9.422	0.094	75.5	10.297	0.1	78	7.969	0.087
	16	78.5	14.844	0.165	80.5	18.547	0.172	81.5	13.812	0.157
	32	73	12.031	0.349	74	13.266	0.371	69.5	10.688	0.363
	64	73	25.375	0.931	74	26.922	0.885	69.5	26.157	0.792

Tabla 5.20. Resultados de las pruebas de reconocimiento con M = 80 y tipo de subimagen 4.

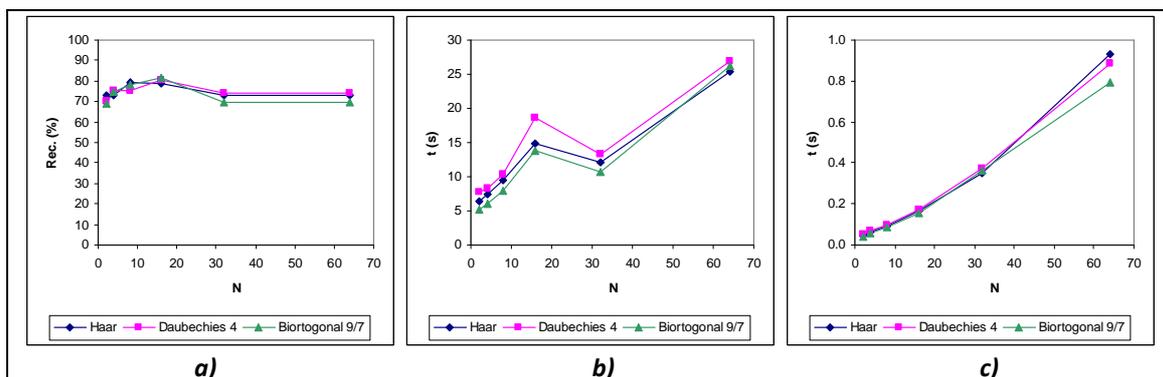


Figura 5.20. Gráficas de los resultados de la Tabla 5.20: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	68	7.203	0.056	67.5	8.656	0.063	67	6.188	0.048
	4	75	8.313	0.077	72	10.047	0.084	75	7.218	0.07
	8	78	10.969	0.125	79.5	13.219	0.129	79	10.11	0.116
	16	79.5	17.719	0.22	82	19.36	0.226	82	17.047	0.214
	32	72.5	13.36	0.507	72	14.25	0.504	71.5	12.328	0.45
	64	72.5	27.625	1.129	72	29.141	1.137	71.5	26.594	1.157

Tabla 5.21. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 4.

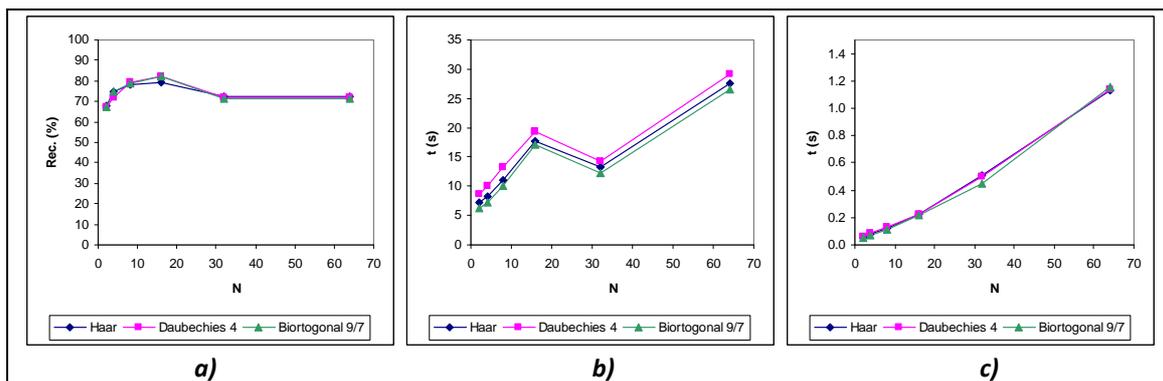


Figura 5.21. Gráficas de los resultados de la Tabla 5.21: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	78	7.859	0.063	78.5	9.625	0.07	79	6.922	0.056
	4	79	9.203	0.091	78.5	10.812	0.098	80	8.172	0.085
	8	82	12.781	0.151	80.5	14.391	0.156	81.5	11.891	0.144
	16	83	19.172	0.275	81	21.032	0.281	82	18.875	0.267
	32	79.5	14.609	0.681	78	15.718	0.631	80.5	14.766	0.623
	64	79.5	29.656	1.488	78	31.922	1.261	80.5	28.328	1.42

Tabla 5.22. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 4.

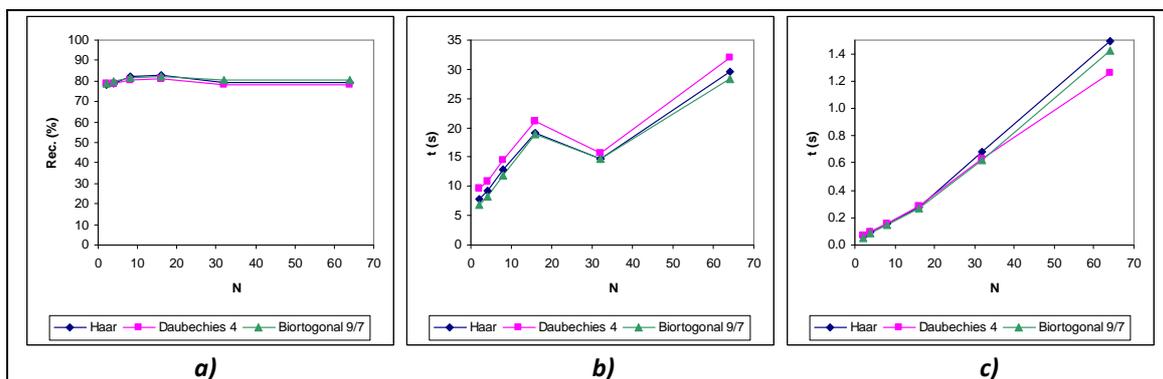


Figura 5.22. Gráficas de los resultados de la Tabla 5.22: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	87	6.125	0.072	86	7.281	0.078	84.5	5.265	0.064
	4	87	6.578	0.108	86	7.594	0.11	84	5.547	0.098
	8	87	7.438	0.177	86	8.312	0.184	84	6.422	0.169
	16	87	7.469	0.326	86	8.875	0.331	84	7.547	0.317
	32	87	13.422	0.75	86	14.453	0.75	84	12.516	0.666
	64	87	32.156	1.513	86	31.016	1.491	84	33.328	1.627

Tabla 5.23. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 4.

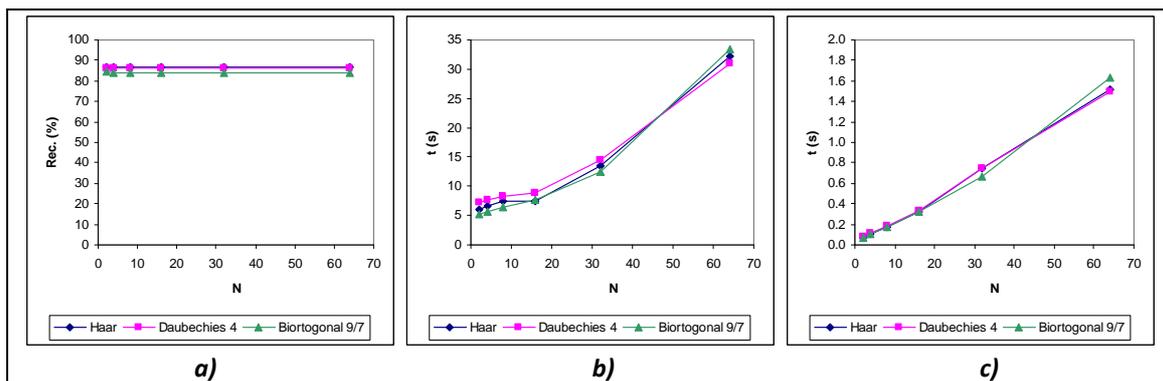


Figura 5.23. Gráficas de los resultados de la Tabla 5.23: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.24 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 4. Estos porcentajes se presentan gráficamente en la Figura 5.24.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	74.5	77	75.5
80	79.5	80.5	81.5
120	79.5	82	82
160	83	81	82
200	87	86	84.5

Tabla 5.24. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 4.

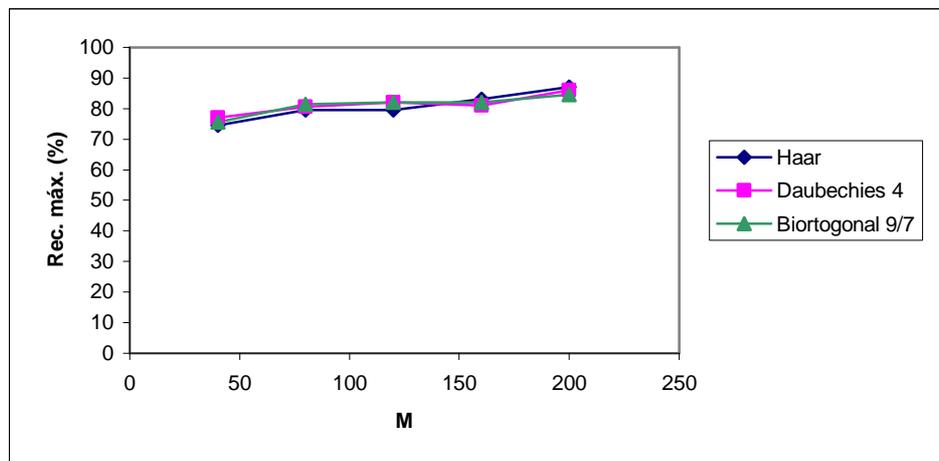


Figura 5.24. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 4 (Rec. máx.: reconocimiento global máximo).

Las Figuras 5.19 a 5.22 señalan un máximo en el porcentaje de reconocimiento global cuando $N = 16$ y $M \in \{40, 80, 120, 160\}$, para la mayoría de las pruebas de este grupo. Los tiempos de entrenamiento son menores a 35 s. El porcentaje de reconocimiento global más alto para este grupo de resultados fue de 87 %, con $M = 200$, $N = 2$ y un filtro paso-bajas de Haar, de acuerdo a la Tabla 5.24. El tiempo total de entrenamiento para este porcentaje fue de 6.125 s, y el tiempo promedio de reconocimiento fue de 0.072 s.

5.1.5 Resultados de las pruebas de reconocimiento para el tipo de subimagen 5

En las Tablas 5.25 a 5.29 se listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 5 y en las Figuras 5.25 a 5.29 se muestran las gráficas de estos resultados, respectivamente.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	70.5	3.671	0.046	65	4.922	0.046	65.5	2.672	0.033
4	76	4.047	0.047	70.5	5.312	0.055	70	3.094	0.042
8	80.5	5.579	0.063	77.5	7.187	0.073	76.5	4.359	0.057
16	86.5	11.266	0.109	85.5	12.187	0.112	81	9.828	0.101
32	67.5	9.453	0.232	66	10.625	0.231	62.5	8.407	0.221
64	67.5	24.609	0.66	66	26.046	0.644	62.5	26.907	0.674

Tabla 5.25. Resultados de las pruebas de reconocimiento con M = 40 y tipo de subimagen 5.

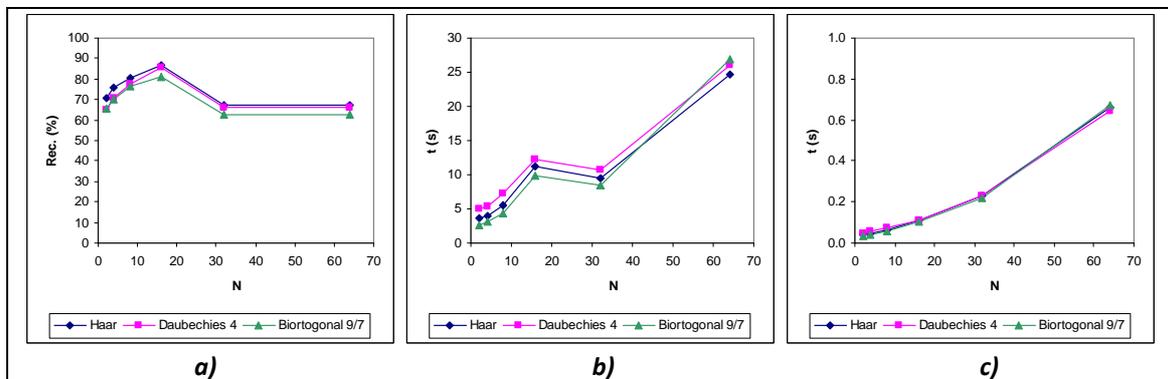


Figura 5.25. Gráficas de los resultados de la Tabla 5.25: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	71.5	4.359	0.047	76	6.031	0.051	76.5	3.266	0.039
4	79.5	4.984	0.063	81	6.125	0.066	82	4.109	0.056
8	82.5	7.078	0.093	81	8.125	0.099	84	6.547	0.086
16	85	13.968	0.168	84	14.937	0.171	84	13.422	0.156
32	75.5	10.656	0.336	77	11.703	0.345	78	9.344	0.326
64	75.5	26.453	0.795	77	27.594	0.93	78	25.406	0.887

Tabla 5.26. Resultados de las pruebas de reconocimiento con M = 80 y tipo de subimagen 5.

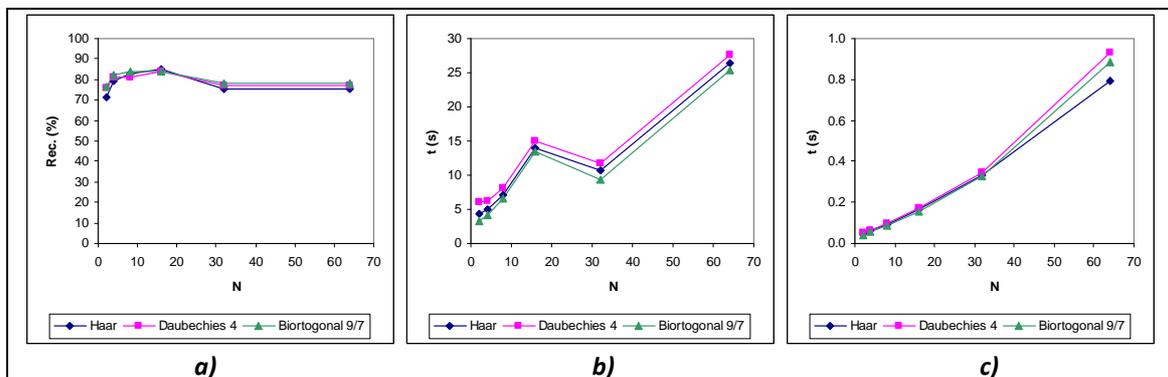


Figura 5.26. Gráficas de los resultados de la Tabla 5.26: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	78.5	4.812	0.054	78	5.969	0.061	76	3.906	0.046
4	84.5	5.657	0.08	82	7.672	0.08	82	4.844	0.068
8	86	8.782	0.122	82.5	10.578	0.125	84.5	7.625	0.114
16	86	16.984	0.219	83.5	17.625	0.222	86.5	16.312	0.212
32	76.5	11.313	0.446	77.5	12.5	0.496	78.5	10.359	0.503
64	76.5	28.016	1.194	77.5	32.422	1.05	78.5	29.594	1.179

Tabla 5.27. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 5.

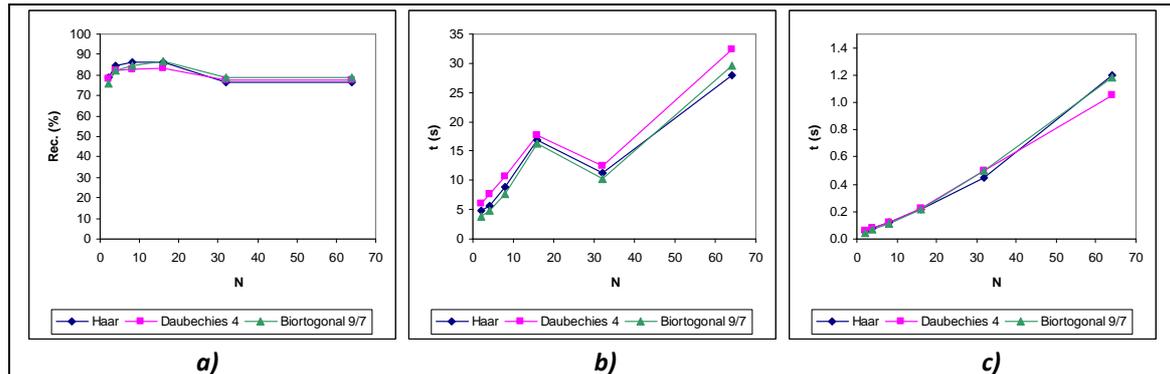


Figura 5.27. Gráficas de los resultados de la Tabla 5.27: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	82.5	5.266	0.072	82	6.437	0.066	84	4.328	0.053
4	88	6.046	0.091	84.5	7.516	0.095	86	5.453	0.082
8	86	9.547	0.152	86	11.141	0.156	86	8.922	0.142
16	86	18.859	0.274	87.5	19.593	0.329	88	16.234	0.268
32	84	12.328	0.59	84	13.453	0.583	85.5	12.532	0.55
64	84	33.5	1.426	84	31.375	1.33	85.5	28.89	1.418

Tabla 5.28. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 5.

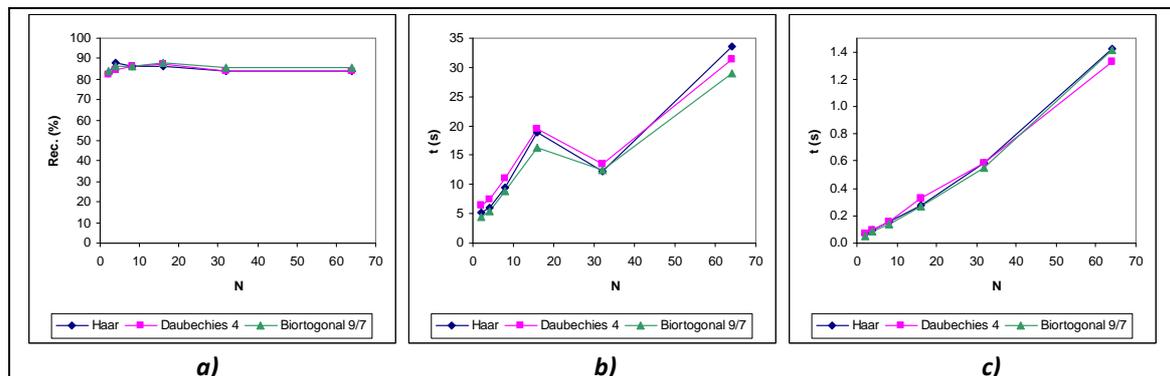


Figura 5.28. Gráficas de los resultados de la Tabla 5.28: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	90.5	4.219	0.068	87.5	5.86	0.076	89	3.313	0.062
4	91	4.578	0.103	87.5	6.047	0.109	89.5	3.609	0.096
8	91	4.953	0.172	87.5	6.062	0.211	89.5	4.078	0.167
16	91	5.672	0.323	87	6.922	0.378	89.5	4.922	0.317
32	91	12.047	0.697	87	13.047	0.66	89.5	11.032	0.667
64	91	29.672	1.443	87	31.078	1.451	89.5	28.766	1.537

Tabla 5.29. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 5.

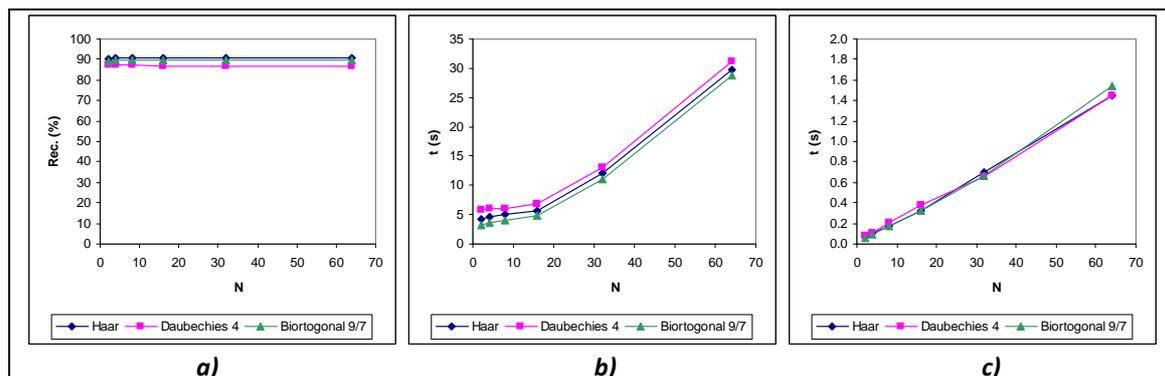


Figura 5.29. Gráficas de los resultados de la Tabla 5.29: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.30 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 5. Estos porcentajes se presentan gráficamente en la Figura 5.30.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	86.5	85.5	81
80	85	84	84
120	86	83.5	86.5
160	88	87.5	88
200	91	87.5	89.5

Tabla 5.30. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 5.

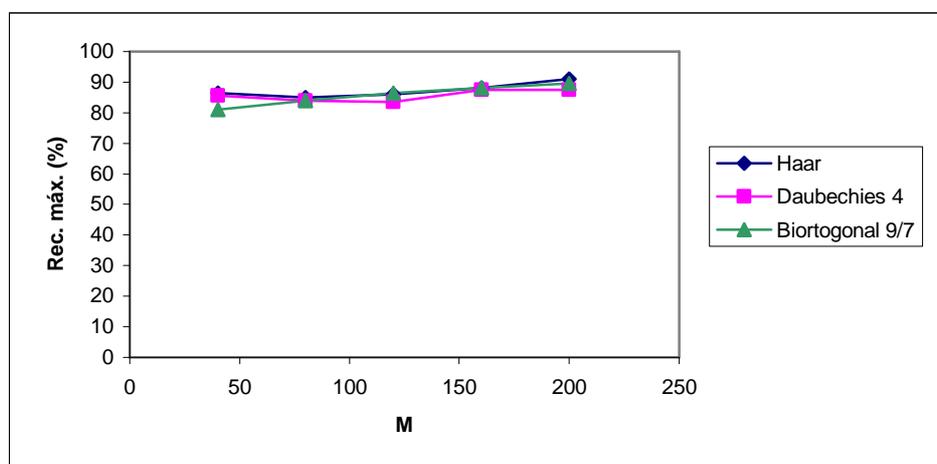


Figura 5.30. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 5 (Rec. máx.: reconocimiento global máximo).

En las Figuras 5.25 a 5.28 se puede apreciar que, en términos generales, se alcanza el máximo porcentaje de reconocimiento global cuando $N = 16$, si $M \in \{40, 80, 120, 160\}$. Al igual que en el grupo de resultados anteriores, los tiempos de entrenamiento no rebasan los 35 s. Como lo señala la Tabla 5.30, el porcentaje de reconocimiento global más alto para este grupo de resultados fue de 91 %, con $M = 200$, $N = 4$ y un filtro paso-bajas de Haar. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento correspondientes a este porcentaje fueron de 4.578 y 0.103 s, respectivamente.

5.1.6 Resultados de las pruebas de reconocimiento para el tipo de subimagen 6

En las Tablas 5.31 a 5.35 se listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 6 y en las Figuras 5.31 a 5.35 se muestran las gráficas de estos resultados, respectivamente.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	66.5	3.594	0.041	64	4.797	0.047	71	2.563	0.033
4	72.5	3.938	0.047	73.5	5.125	0.052	77.5	2.938	0.039
8	80	4.875	0.063	78	6.125	0.07	77.5	3.922	0.057
16	63.5	4.391	0.107	68.5	5.656	0.111	66	3.39	0.099
32	63.5	7.125	0.228	68.5	7.735	0.221	66	5.594	0.207
64	63.5	14.688	0.626	68.5	15.766	0.633	66	13.609	0.631

Tabla 5.31. Resultados de las pruebas de reconocimiento con $M = 40$ y tipo de subimagen 6.

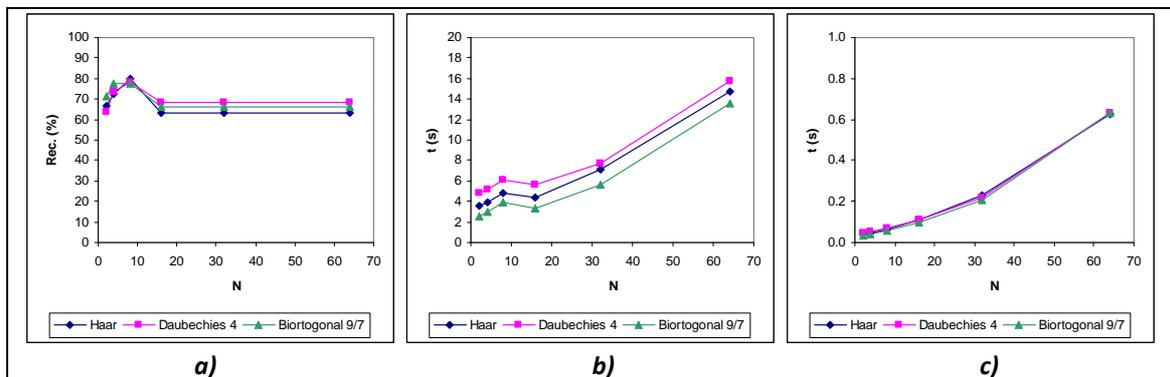


Figura 5.31. Gráficas de los resultados de la Tabla 5.31: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7		
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)
2	79.5	4.125	0.047	77.5	5.532	0.055	75.5	3.218	0.039
4	80	4.781	0.063	79.5	6.063	0.066	78.5	3.812	0.053
8	85	6.203	0.094	81.5	7.656	0.098	83	5.297	0.088
16	75	5.109	0.16	76.5	6.484	0.168	74.5	4.187	0.154
32	75	7.593	0.323	76.5	8.922	0.329	74.5	7.047	0.316
64	75	16.125	0.766	76.5	17.593	0.899	74.5	15.594	0.797

Tabla 5.32. Resultados de las pruebas de reconocimiento con $M = 80$ y tipo de subimagen 6.

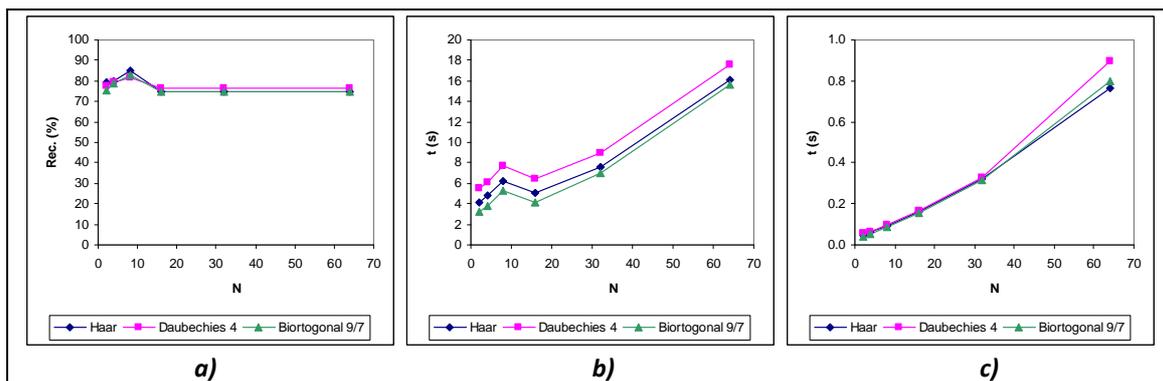


Figura 5.32. Gráficas de los resultados de la Tabla 5.32: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	79	4.609	0.054	82	5.969	0.061	78.5	3.718	0.046
	4	81	5.375	0.08	83	6.719	0.08	82.5	4.468	0.068
	8	84	7.297	0.122	85	9.031	0.125	85	6.672	0.113
	16	77	5.719	0.217	80.5	7.063	0.221	76.5	4.813	0.209
	32	77	8.453	0.432	80.5	9.75	0.447	76.5	7.828	0.425
	64	77	17.718	1.025	80.5	19.094	1.01	76.5	16.594	1.042

Tabla 5.33. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 6.

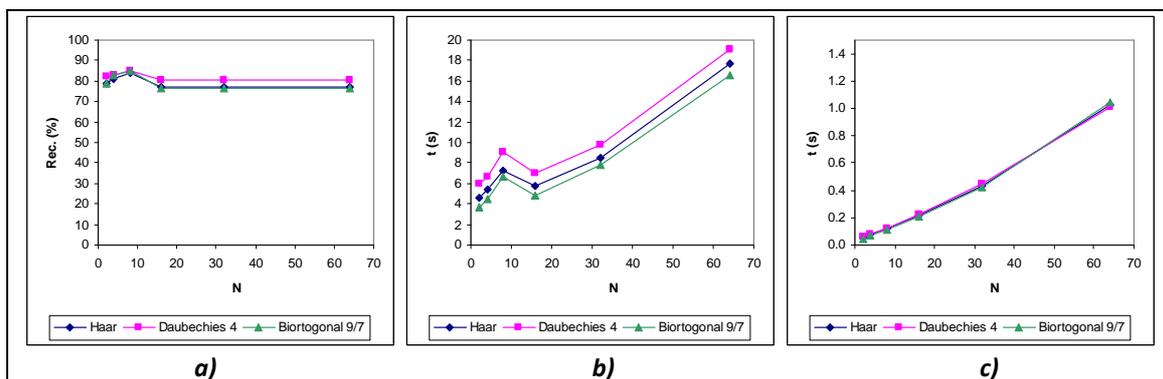


Figura 5.33. Gráficas de los resultados de la Tabla 5.33: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	79.5	4.922	0.063	84	6.422	0.066	85	4.046	0.054
	4	82	5.75	0.091	84.5	7.375	0.094	86	4.859	0.082
	8	85.5	7.391	0.149	85.5	9.5	0.156	86.5	6.875	0.141
	16	81	6.235	0.27	84.5	7.687	0.277	82.5	5.391	0.263
	32	81	9.188	0.594	84.5	10.64	0.558	82.5	8.781	0.569
	64	81	18.922	1.356	84.5	20.953	1.31	82.5	18.359	1.382

Tabla 5.34. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 6.

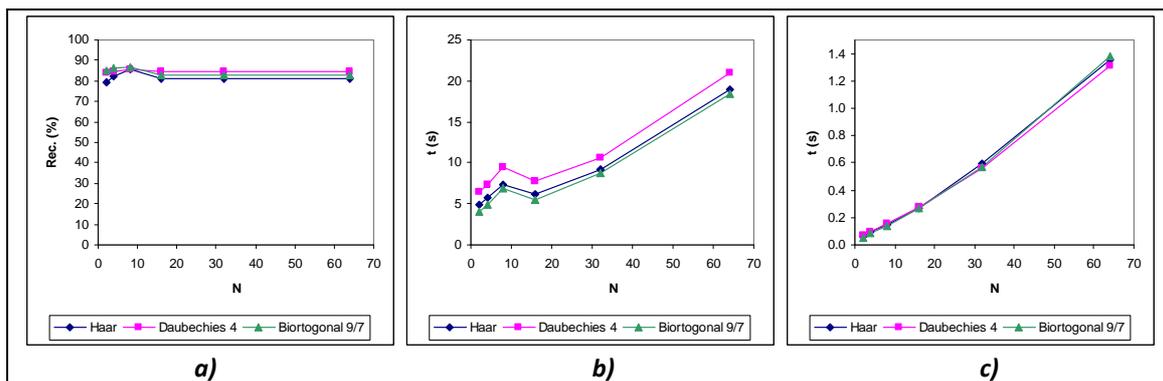


Figura 5.34. Gráficas de los resultados de la Tabla 5.34: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	87.5	4.093	0.068	86	5.344	0.076	89.5	3.25	0.061
	4	87.5	4.344	0.103	86	5.437	0.11	89.5	3.391	0.096
	8	87.5	4.641	0.173	86	5.828	0.18	89.5	3.719	0.166
	16	87.5	5.86	0.321	86	7.297	0.328	89.5	4.985	0.314
	32	87.5	9.031	0.648	86	10.656	0.651	89.5	8.593	0.707
	64	87.5	19.516	1.591	86	21.265	1.443	89.5	19.188	1.617

Tabla 5.35. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 6.

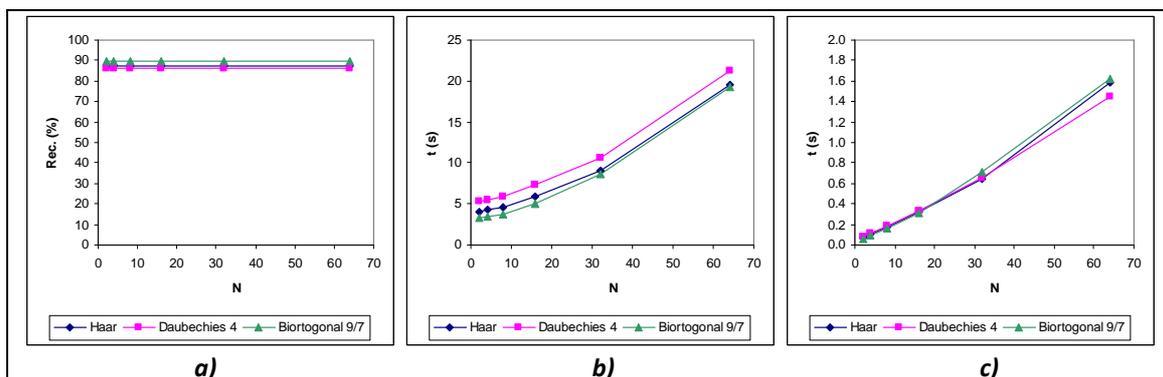


Figura 5.35. Gráficas de los resultados de la Tabla 5.35: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.36 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 6. Estos porcentajes se presentan gráficamente en la Figura 5.36.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	80	78	77.5
80	85	81.5	83
120	84	85	85
160	85.5	85.5	86.5
200	87.5	86	89.5

Tabla 5.36. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 6.

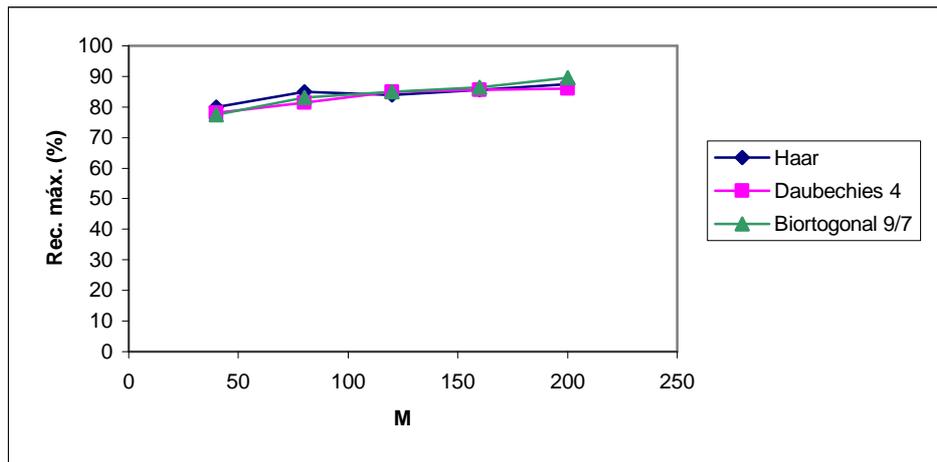


Figura 5.36. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 6 (Rec. máx.: reconocimiento global máximo).

Para las Figuras 5.31 a 5.34, se distingue que el máximo porcentaje de reconocimiento global se logra cuando $N = 8$, si $M \in \{40, 80, 120, 160\}$. Los tiempos de entrenamiento están por debajo de los 25 s. De la Tabla 5.36, se puede ver que el porcentaje de reconocimiento global más alto para este grupo de resultados fue de 89.5 %, con $M = 200$, $N = 2$ y un filtro paso-bajas de tipo Biortogonal 9/7. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento correspondientes a este porcentaje fueron de 3.25 y 0.061 s, respectivamente.

5.1.7 Resultados de las pruebas de reconocimiento para el tipo de subimagen 7

En las Tablas 5.37 a 5.42 se listan los resultados de las pruebas de reconocimiento con el SRRP para el tipo de subimagen 7 y en las Figuras 5.37 a 5.42 se muestran las gráficas de estos resultados, respectivamente.

M=40	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	68	4.547	0.043	67	5.813	0.047	68.5	3.438	0.032
	4	76	5.016	0.049	71	6.203	0.055	77	3.906	0.04
	8	79.5	6.828	0.068	76.5	7.563	0.072	75	5.469	0.058
	16	82	11.532	0.109	78	12.609	0.113	83.5	10.282	0.101
	32	69.5	10.219	0.239	68.5	11.094	0.228	68.5	11.812	0.229
	64	69.5	24.703	0.576	68.5	28.234	0.594	68.5	29.406	0.642

Tabla 5.37. Resultados de las pruebas de reconocimiento con $M = 40$ y tipo de subimagen 7.

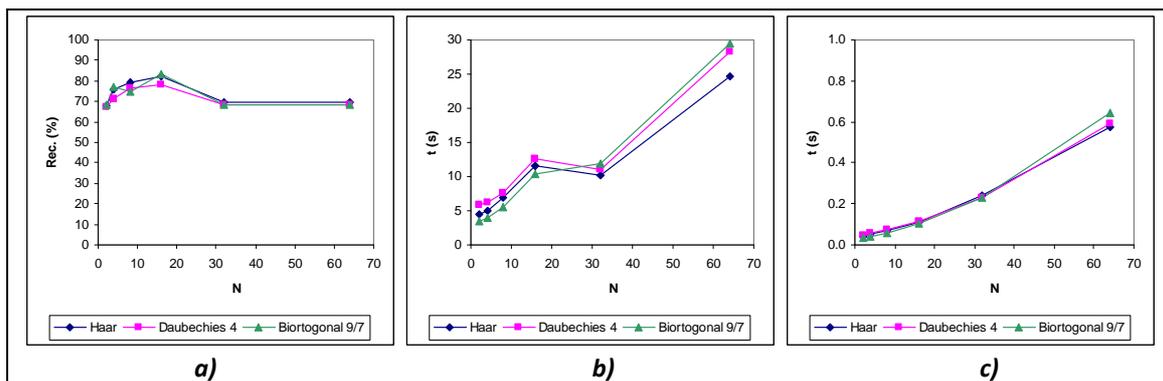


Figura 5.37. Gráficas de los resultados de la Tabla 5.37: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=80	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	76	5.984	0.052	75	7.032	0.053	76	4.562	0.042
	4	80.5	6.312	0.062	78.5	7.61	0.068	78.5	5.25	0.054
	8	83	8.344	0.094	80	9.594	0.099	82.5	7.625	0.086
	16	86.5	15.078	0.164	83	17.344	0.171	85.5	13.968	0.16
	32	74.5	11.359	0.37	75	12.829	0.359	77	10.344	0.363
	64	74.5	30.61	0.921	75	31.063	0.948	77	25.734	0.846

Tabla 5.38. Resultados de las pruebas de reconocimiento con M = 80 y tipo de subimagen 7.

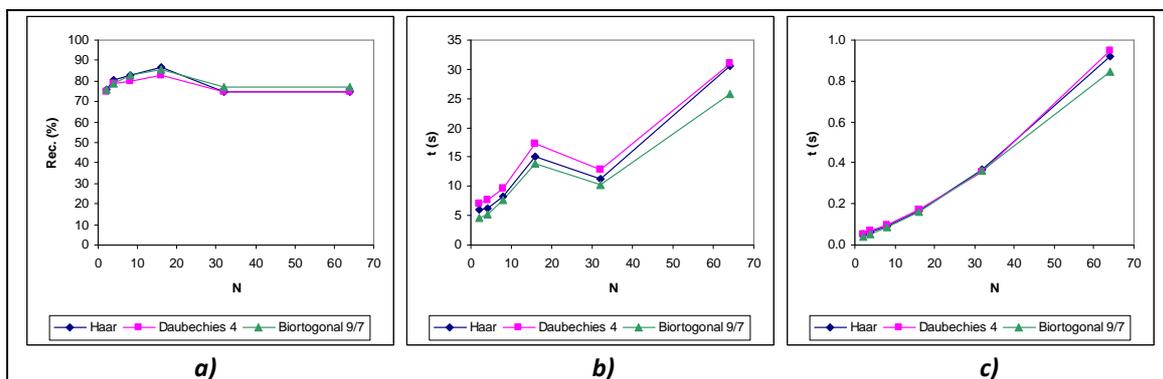


Figura 5.38. Gráficas de los resultados de la Tabla 5.38: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=120	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	79.5	6.891	0.055	78	7.719	0.078	80.5	5.625	0.048
	4	82	7.812	0.077	83	8.796	0.082	81.5	6.672	0.071
	8	83	9.937	0.125	83.5	13.813	0.127	82.5	9.953	0.115
	16	86	18.391	0.219	85	20.297	0.225	86	17.422	0.212
	32	77.5	12.625	0.443	79	13.937	0.447	77	11.765	0.456
	64	77.5	29.015	1.15	79	30.329	1.138	77	28.719	1.161

Tabla 5.39. Resultados de las pruebas de reconocimiento con M = 120 y tipo de subimagen 7.

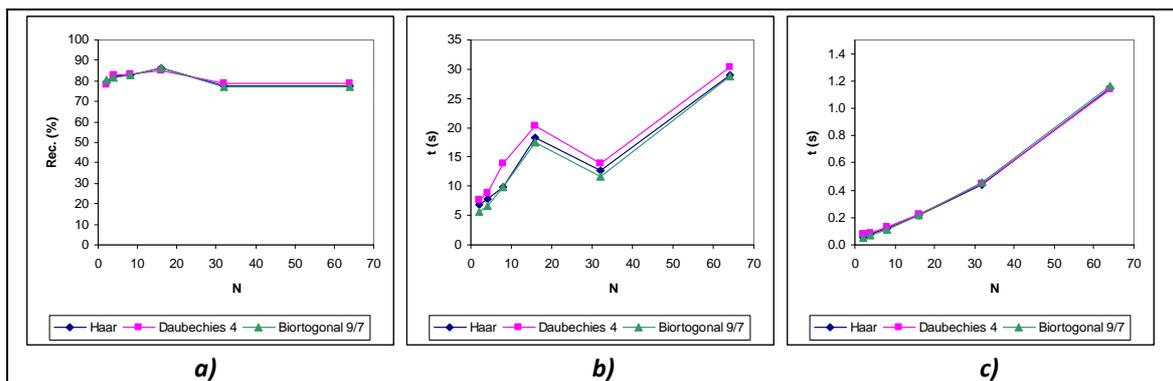


Figura 5.39. Gráficas de los resultados de la Tabla 5.39: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=160	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	83.5	7.593	0.063	85	8.954	0.073	83.5	6.437	0.059
	4	84.5	8.328	0.094	84.5	9.672	0.111	85.5	7.563	0.091
	8	85	11.672	0.169	86	13.375	0.16	87.5	10.813	0.149
	16	86	19.687	0.276	89	21.265	0.281	89	18.672	0.267
	32	83	13.875	0.558	86.5	15.11	0.622	84.5	13.203	0.543
	64	83	30.844	1.394	86.5	32.5	1.374	84.5	30.36	1.394

Tabla 5.40. Resultados de las pruebas de reconocimiento con M = 160 y tipo de subimagen 7.

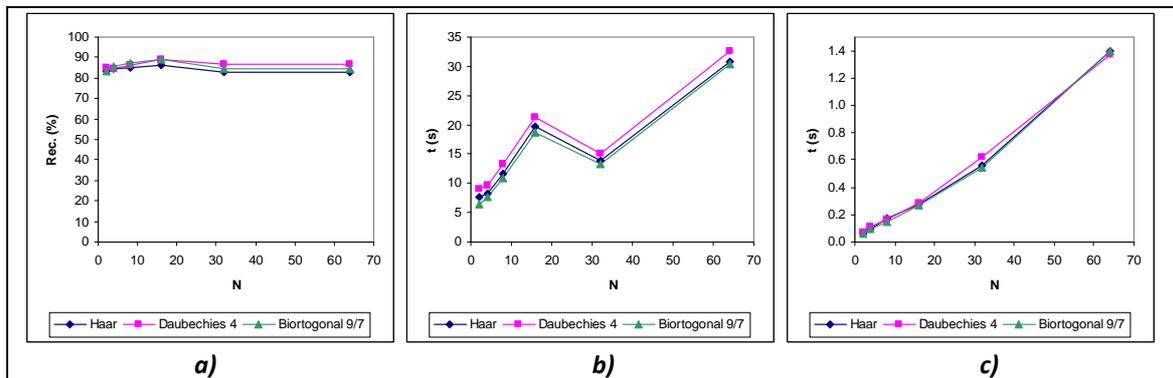


Figura 5.40. Gráficas de los resultados de la Tabla 5.40: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

M=200	Filtro PB: Haar			Filtro PB: Daubechies 4			Filtro PB: Biortogonal 9/7			
	N	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)	Rec. (%)	tentr (s)	trec (s)
	2	89.5	5.516	0.072	88.5	7.234	0.08	91	4.469	0.065
	4	89.5	6.297	0.113	88.5	7.86	0.112	91	4.797	0.1
	8	89.5	6.796	0.18	88.5	8.297	0.186	91	5.25	0.169
	16	89.5	7.703	0.326	89	8.797	0.329	90.5	6.547	0.367
	32	89.5	13.859	0.731	89	14.312	0.72	91	12.219	0.644
	64	89.5	37.547	1.459	89	35.094	1.573	91	30.734	1.638

Tabla 5.41. Resultados de las pruebas de reconocimiento con M = 200 y tipo de subimagen 7.

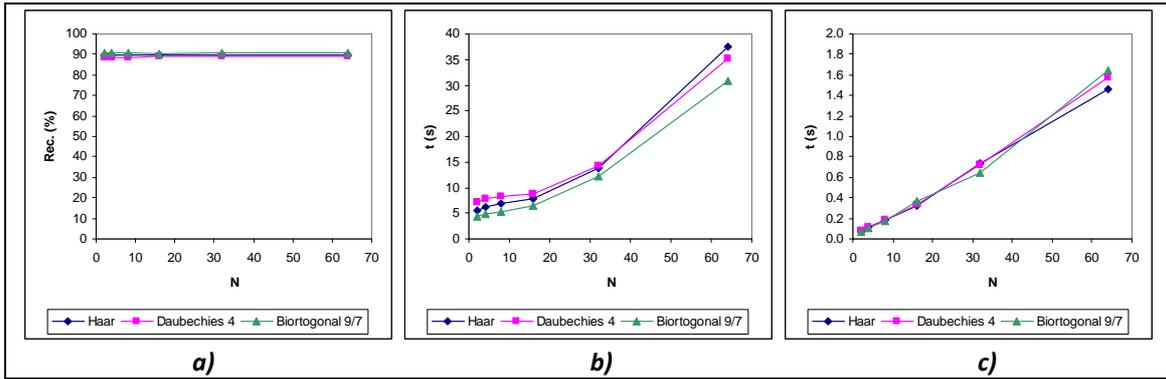


Figura 5.41. Gráficas de los resultados de la Tabla 5.41: a) Porcentaje de reconocimiento global, b) Tiempo total de entrenamiento, c) Tiempo promedio de reconocimiento.

En la Tabla 5.42 se muestran los porcentajes de reconocimiento global máximos para el tipo de subimagen 7. Estos porcentajes se presentan gráficamente en la Figura 5.42.

M	Reconocimiento global máximo (%)		
	Filtro PB: Haar	Filtro PB: Daubechies 4	Filtro PB: Biortogonal 9/7
40	82	78	83.5
80	86.5	83	85.5
120	86	85	86
160	86	89	89
200	89.5	89	91

Tabla 5.42. Porcentajes de reconocimiento global máximos utilizando el tipo de subimagen 7.

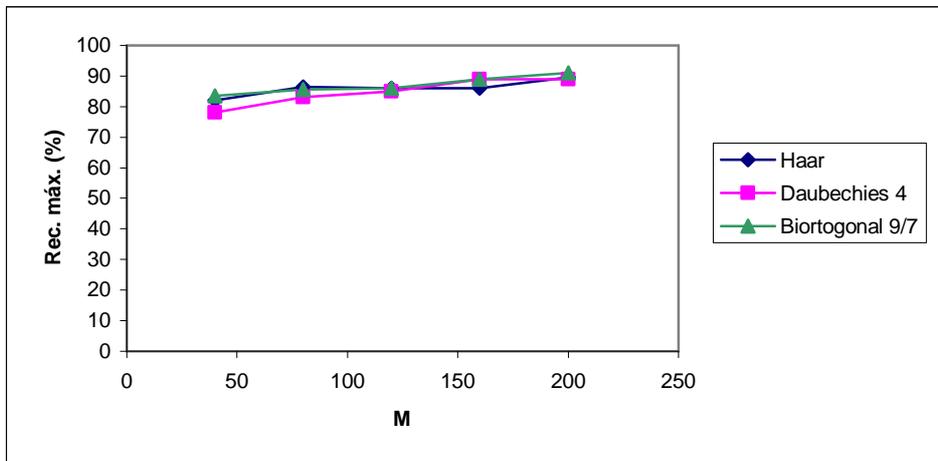


Figura 5.42. Comportamiento del porcentaje de reconocimiento global máximo en función de M para el tipo de subimagen 7 (Rec. máx.: reconocimiento global máximo).

Para las Figuras 5.37 a 5.40, el máximo porcentaje de reconocimiento global aparece cuando $N = 16$, si $M \in \{40, 80, 120, 160\}$. Los tiempos de entrenamiento son menores a los 40 s. La Tabla 5.42 indica que el porcentaje de reconocimiento global más alto para este grupo de resultados fue de 91 %, con $M = 200$, $N = 2$ y un filtro paso-bajas de tipo Biortogonal 9/7. Para esta prueba, el tiempo total de entrenamiento fue de 4.469 s y el tiempo promedio de reconocimiento fue de 0.065 s.

5.2 Análisis general de las pruebas de reconocimiento hechas con el SRRP

De acuerdo a los resultados de la sección 5.1, si se asume que los porcentajes de reconocimiento global más altos son aquellos superiores al 90 %, entonces, se puede apreciar que los tipos de subimagen que permitieron llegar a estos porcentajes fueron los tipos 1, 2, 5 y 7. Estas pruebas se analizarán con más detalle en la sección 5.3.

Las gráficas de las Figuras 5.1 a 5.42 dejan ver cómo es que los tres tipos de filtros paso-bajas tienen un comportamiento muy similar aunque, para las gráficas de las Figuras 5.6, 5.12, 5.18, 5.24, 5.30, 5.36 y 5.42, la tendencia que siguen presenta algunas variaciones. Se observa que con el máximo valor de M se obtiene el porcentaje de reconocimiento global más alto, excepto para el caso de la Figura 5.18, en el cual con el filtro paso-bajas de tipo Biorotogonal 9/7 se llega al máximo porcentaje de reconocimiento global (78 %) de la prueba correspondiente, con $M = 120$ y 160.

En cuanto a las gráficas de las Figuras 5.5, 5.11, 5.17, 5.23, 5.29, 5.35 y 5.41, el valor del mismo porcentaje se vuelve prácticamente constante. Esto último es debido a que, en la etapa de entrenamiento, no se encontraron HMMs que representaran mejor las secuencias de observación de entrenamiento, incluso si N aumentaba. Una causa de esto puede ser que el Algoritmo Baum-Welch haya encontrado un máximo local, evitando una nueva reestimación de parámetros, para todos los HMMs contemplados.

Los grupos de resultados 3 y 6 implican barridos menos exhaustivos de las imágenes de entrenamiento y de prueba, lo cual causa que los tiempos de entrenamiento sean más pequeños, en comparación con los otros grupos. No obstante, ninguno de estos tipos de subimagen permitió llegar a porcentajes de reconocimiento global superiores al 90 %. De hecho, al usar el tipo de subimagen 3 se obtienen los porcentajes de reconocimiento global más bajos, llegando a encontrar resultados de hasta 55 %.

Las variaciones en la tendencia del tiempo total de entrenamiento en función de N están directamente relacionadas con el Algoritmo Baum-Welch el cual, dependiendo de los parámetros iniciales de los HMMs y las secuencias de observación de entrenamiento, puede tener más o menos iteraciones, repercutiendo en un entrenamiento más tardado o más rápido.

5.3 Análisis de los porcentajes de reconocimiento global más altos

En esta sección se estudiarán con más detalle aquellas pruebas que produjeron los porcentajes de reconocimiento global más altos. Como se mencionó al final de la sección 5.2, los tipos de subimagen que llevaron a alcanzar estos porcentajes fueron los tipos 1, 2, 5 y 7, para cada uno de los cuales se eligió el porcentaje de reconocimiento global más alto de entre otros que también llegaron a rebasar el 90 %, correspondientes al mismo tipo de subimagen.

A continuación se analizan estas 4 pruebas de reconocimiento con mayor detenimiento, denotándolas como Prueba M_i , con $i = 1, 2, 3, 4$. Otras pruebas conducen también a los mismos porcentajes, sin embargo, se seleccionaron aquellas con los valores de N más pequeños.

Prueba M1. Para esta prueba, se usaron el tipo de subimagen 1, el filtro paso-bajas de Haar, un valor de $M = 200$ y $N = 16$. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento obtenidos fueron 6.578 s y 0.351 s, respectivamente. Se lograron reconocer correctamente 181 imágenes de prueba de las 200 disponibles, dando un porcentaje de reconocimiento global de 90.5 %. En la Tabla 5.43 aparecen los porcentajes de reconocimiento de cada sujeto de la BDR ORL y en la Figura 5.43 el histograma de estos porcentajes.

Sujeto	Reconocimiento (%)	Sujeto	Reconocimiento (%)
1	100	21	100
2	100	22	100
3	80	23	100
4	80	24	100
5	80	25	100
6	100	26	100
7	100	27	80
8	100	28	80
9	80	29	80
10	100	30	100
11	100	31	80
12	100	32	80
13	80	33	100
14	100	34	100
15	100	35	100
16	80	36	100
17	20	37	100
18	100	38	100
19	100	39	80
20	100	40	40

Tabla 5.43. Porcentajes de reconocimiento por sujeto para la Prueba M1.

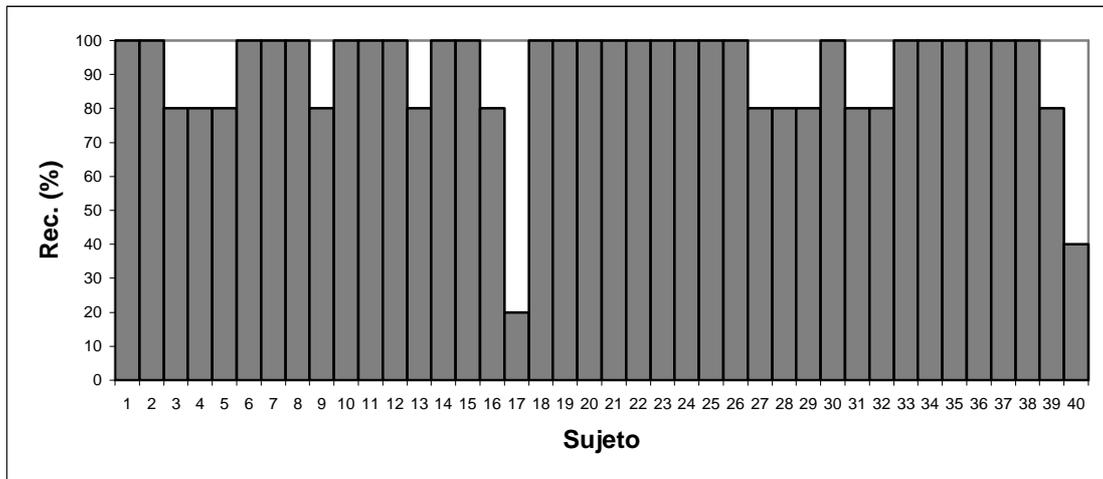


Figura 5.43. Histograma de porcentajes de reconocimiento por sujeto para la Prueba M1.

Prueba M2. Esta prueba considera el tipo de subimagen 2, el filtro paso-bajas de Haar, un valor de $M = 200$ y $N = 2$. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento obtenidos fueron 7.641 s y 0.077 s, respectivamente. Se reconocieron correctamente 183 imágenes de prueba de las 200 disponibles, dando un porcentaje de reconocimiento global de 91.5 %. En la Tabla 5.44 aparecen los porcentajes de reconocimiento de cada sujeto de la BDR ORL y en la Figura 5.44 el histograma de estos porcentajes.

Sujeto	Reconocimiento (%)	Sujeto	Reconocimiento (%)
1	100	21	100
2	100	22	100
3	80	23	100
4	100	24	100
5	80	25	100
6	100	26	100
7	100	27	100
8	100	28	80
9	80	29	100
10	100	30	100
11	100	31	80
12	100	32	100
13	100	33	100
14	60	34	100
15	100	35	80
16	60	36	100
17	0	37	100
18	100	38	100
19	100	39	100
20	100	40	60

Tabla 5.44. Porcentajes de reconocimiento por sujeto para la Prueba M2.

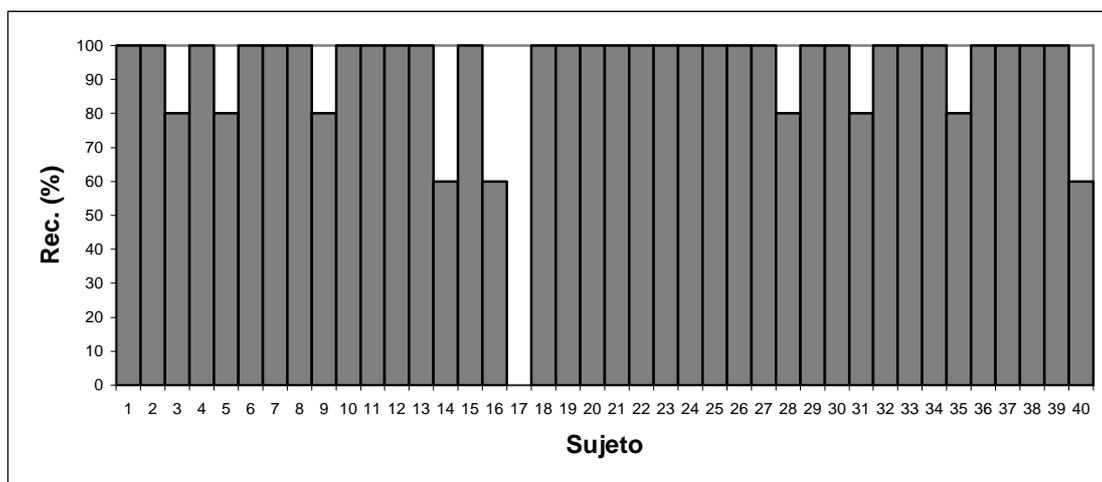


Figura 5.44. Histograma de porcentajes de reconocimiento por sujeto para la Prueba M2.

Prueba M3. Para esta prueba, se usaron el tipo de subimagen 5, el filtro paso-bajas de Haar, $M = 200$ y $N = 4$. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento fueron 4.578 s y 0.103 s, respectivamente. Se reconocieron correctamente 182 imágenes de prueba de las 200 disponibles, dando un porcentaje de reconocimiento global de 91 %. En la Tabla 5.45 se observan los porcentajes de reconocimiento de cada sujeto de la BDR ORL y en la Figura 5.45 el histograma de estos porcentajes.

Sujeto	Reconocimiento (%)	Sujeto	Reconocimiento (%)
1	60	21	100
2	100	22	100
3	100	23	100
4	100	24	100
5	100	25	100
6	100	26	80
7	100	27	80
8	100	28	80
9	60	29	100
10	100	30	80
11	100	31	80
12	100	32	60
13	100	33	100
14	100	34	100
15	100	35	100
16	80	36	80
17	40	37	100
18	100	38	100
19	80	39	100
20	100	40	80

Tabla 5.45. Porcentajes de reconocimiento por sujeto para la Prueba M3.

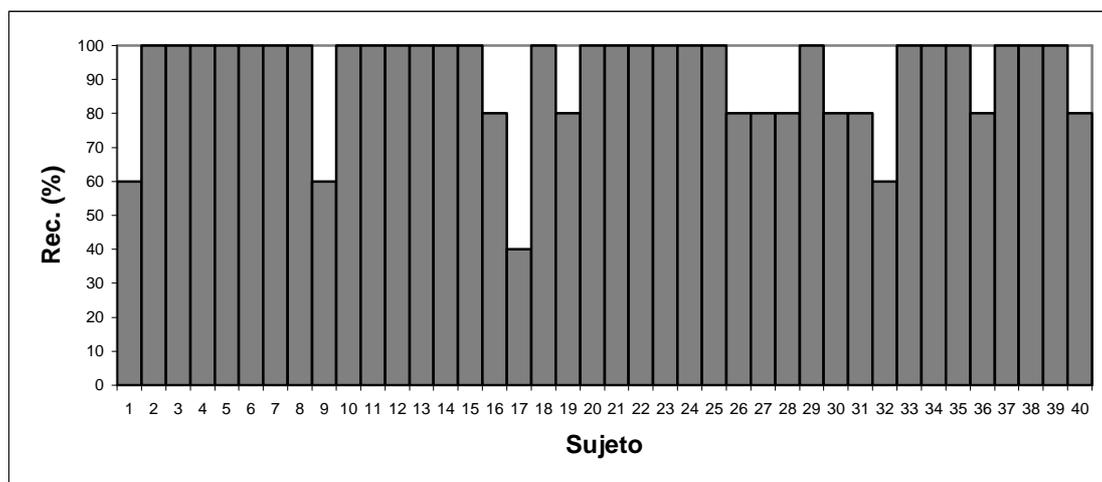


Figura 5.45. Histograma de porcentajes de reconocimiento por sujeto para la Prueba M3.

Prueba M4. En esta prueba se seleccionó el tipo de subimagen 7 y el filtro paso-bajas de tipo Biortogonal 9/7, y se manejaron los valores de $M = 200$ y $N = 2$. El tiempo total de entrenamiento y el tiempo promedio de reconocimiento fueron 4.469 s y 0.065 s, respectivamente. Se reconocieron correctamente 182 imágenes de prueba de las 200 disponibles, dando un porcentaje de reconocimiento global de 91 %. La Tabla 5.46 lista los porcentajes de reconocimiento de cada sujeto de la BDR ORL y en la Figura 5.46 se muestra el histograma de estos porcentajes.

Sujeto	Reconocimiento (%)	Sujeto	Reconocimiento (%)
1	80	21	100
2	100	22	100
3	80	23	100
4	100	24	100
5	100	25	100
6	100	26	80
7	100	27	80
8	100	28	80
9	80	29	100
10	100	30	80
11	100	31	80
12	100	32	60
13	100	33	100
14	80	34	100
15	100	35	80
16	80	36	100
17	40	37	100
18	100	38	100
19	80	39	100
20	100	40	80

Tabla 5.46. Porcentajes de reconocimiento por sujeto para la Prueba M4.

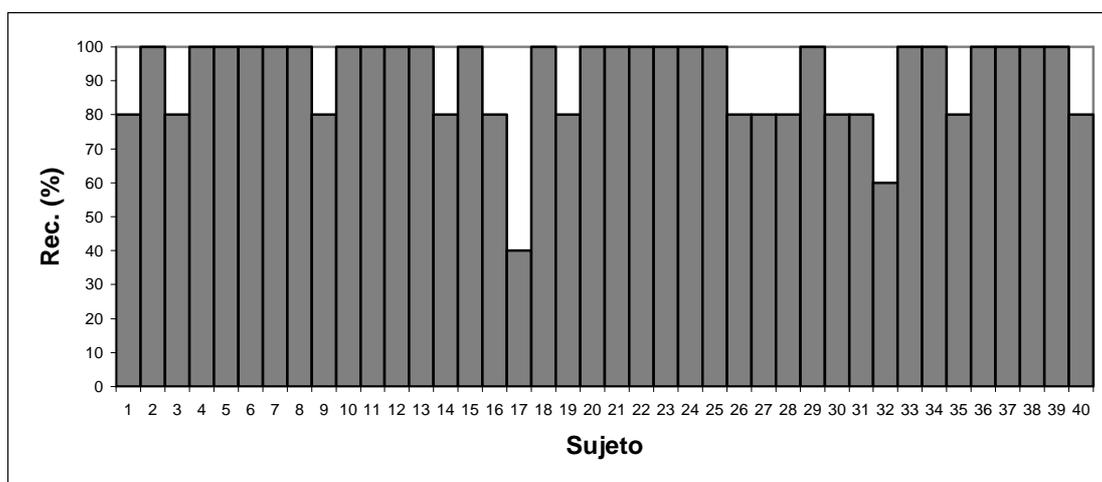


Figura 5.46. Histograma de porcentajes de reconocimiento por sujeto para la Prueba M4.

Como se puede ver en los histogramas de las Figuras 5.43 a 5.46, algunos sujetos de la BDR ORL obtuvieron el 100 % de reconocimiento en las 4 pruebas anteriores. Dichos sujetos se presentan en la Tabla 5.47. Se puede decir entonces que, para las pruebas descritas en este capítulo y con las consideraciones antes mencionadas, estos son los sujetos mejor reconocidos por el SRRP.

Por el contrario, cabe destacar que el sujeto con el porcentaje de reconocimiento más bajo en las mismas 4 pruebas fue el número 17, lo cual es más evidente en la Figura 5.44 correspondiente a la prueba M2, en la que obtuvo 0 % de reconocimiento. La Figura 5.47 contiene una imagen del rostro de este sujeto. Podría decirse que este fue el sujeto más difícil de reconocer por el SRRP, en las pruebas M1 a M4.

ID sujeto	Imagen	ID sujeto	Imagen	ID sujeto	Imagen	ID sujeto	Imagen
2		11		21		33	
6		12		22		34	
7		15		23		37	
8		18		24		38	
10		20		25			

Tabla 5.47. Sujetos mejor reconocidos por el SRRP en las pruebas M1 a M4.



Figura 5.47. Sujeto más difícil de reconocer por el SRRP en las pruebas M1 a M4.

En la Tabla 5.48 se muestran cuales fueron los sujetos que se reconocieron en lugar del sujeto 17, para las pruebas M1 a M4. De estos resultados, se aprecia que el sistema reconoció a los sujetos 1, 24 y 36 a partir de las imágenes de prueba del sujeto 17.

Imagen de prueba (sujeto 17)	ID sujeto reconocido (Prueba M1)	ID sujeto reconocido (Prueba M2)	ID sujeto reconocido (Prueba M3)	ID sujeto reconocido (Prueba M4)
1	1	1	24	36
2	1	1	24	24
3	17	1	17	17
4	1	36	17	17
5	1	1	24	24

Tabla 5.48. Resultados de reconocimiento de las pruebas M1 a M4 para el sujeto 17 de la BDR ORL.

Capítulo 6. Conclusiones, mejoras y posibles aplicaciones

Este capítulo está destinado a presentar las conclusiones a las que el presente trabajo condujo, además de explicar algunos aspectos que podrían mejorar el funcionamiento del SRRP. Se comentan también las posibles aplicaciones que dicho sistema puede tener y los factores que se deben tomar en cuenta en cada una.

6.1 Conclusiones

- 1) Se desarrolló un SRR aplicado a imágenes provenientes de la BDR ORL. Este sistema permite establecer los valores de los parámetros que se han de usar para su entrenamiento. Para la extracción de características, se proveen tres tipos de filtros paso-bajas de descomposición y 7 formas de analizar cada imagen. En el caso de la definición de los HMMs, se pueden elegir los valores de N , M y Δ , suponiendo un HMM discreto con topología Bakis.
- 2) Con respecto a los tipos de filtros utilizados, se muestra cómo el filtro de Haar es el que da los mejores resultados, seguido del Biortogonal 9/7. El filtro Daubechies 4 no figura en las pruebas que condujeron a los porcentajes de reconocimiento global más altos. Esto corrobora el hecho de que en otras aplicaciones en las que se han utilizado filtros de descomposición, se prefiera el de Haar, además de su sencillez.
- 3) De las pruebas de reconocimiento hechas con el SRRP y utilizando la BDR ORL, 4 de ellas condujeron a los porcentajes de reconocimiento global más altos: 90.5 %, 91 % y 91.5 %. A partir de estas pruebas, se identificaron 17 sujetos de la BDR que pueden ser identificados como los sujetos mejor reconocidos por el sistema, ya que el porcentaje de reconocimiento por sujeto en estos casos fue del 100 %.
- 4) Un punto importante que se considera en los SRRs es el tiempo de reconocimiento. Para el SRRP, de los porcentajes de reconocimiento global más altos, el mayor tiempo promedio de reconocimiento fue de 0.351 s, correspondiente a la Prueba M1 que se estudia en el Capítulo 5.
- 5) Se pudieron identificar correctamente sujetos cuyas imágenes de prueba exhibían cambios muy marcados en los rasgos faciales, como sonrisas, ojos cerrados y muecas de tristeza y de sorpresa. En algunos casos, la identificación fue satisfactoria incluso al considerar accesorios, como lentes.
- 6) Los resultados de la sección 5.3 permitieron distinguir que sólo el sujeto 17 presentó más complicaciones en su reconocimiento. Es importante recordar que las variables que repercuten más en el reconocimiento facial son los cambios de iluminación, los cambios en la posición del rostro y rasgos faciales no uniformes. Con respecto a este último punto, los rasgos que quizá dificultaron una buena identificación del sujeto antes mencionado fueron el vello facial y el uso de lentes, además de tomar en cuenta que las imágenes se analizan incluyendo el fondo, lo que dificulta más la extracción de características.
- 7) Las limitaciones del SRRP son, principalmente, que se requiere una BDR lo suficientemente amplia. Además, las imágenes de los sujetos de la BDR a utilizar deben presentar un rostro en una pose aproximadamente frontal, lo cual no siempre se logra en dichas imágenes.

6.2 Mejoras

- 1) En cuanto a los parámetros de los HMMs utilizados y con base en los resultados del Capítulo 5, se observa que no necesariamente mientras N es más grande, los HMMs representan mejor a las imágenes de los rostros de los sujetos de la BDR. Para el valor de M , esto no se cumple, debido a que un mayor número de símbolos de observación produce mejores resultados, en general. Sin embargo, al aumentar los valores de M , aumenta el número de elementos en los parámetros de los HMMs, principalmente en la matriz de emisión, causando que la etapa de entrenamiento sea más tardada. En el presente trabajo se logran buenos resultados de reconocimiento con un valor pequeño de N , sin embargo, también se podrían estudiar métodos para lograr una reducción en el valor de M .
- 2) Otro aspecto relevante de los HMMs es el tipo de topología usada. Como ya se mencionó, se implementaron HMMs Bakis discretos. Debido a la cantidad de pruebas que provocaría variar el valor de Δ , se fijó dicho valor en 1, con el fin de tener un primer panorama del comportamiento de estos modelos en el contexto del Reconocimiento de Rostros. Podrían realizarse más pruebas sobre esta misma topología y para distintos valores de Δ , pero convendría más considerar topologías distintas, como las de los modelos ergódicos y, también, no dejar de lado el uso de HMMs continuos. Las observaciones que se tendrían que hacer sobre estos últimos modelos serían el número de parámetros que se requerirían, así como las formulaciones a implementar.
- 3) Para la etapa de entrenamiento del SRRP, mientras más imágenes se tengan de un sujeto, su porcentaje de reconocimiento puede aumentar, siempre y cuando dichas imágenes describan de la mejor manera los posibles cambios en la posición de su rostro y en sus expresiones faciales lo cual, a su vez, implica contar con BDRs suficientemente extensas.
- 4) Con relación al punto anterior, las pruebas hechas con el SRRP consideraron sólo 5 imágenes de entrenamiento, lo cual de primera mano llevaría a pensar que se debe considerar un número mayor de dichas imágenes para alcanzar porcentajes de reconocimiento más altos. No obstante, se debe tener en cuenta la limitada cantidad de imágenes por sujeto que posee la BDR utilizada y, por tanto, la cantidad de imágenes de prueba disponibles para generar estadísticas útiles. Incluso, si se tuviera una BDR muy amplia, el desarrollo de un SRR eficiente y práctico debería concentrarse en la obtención de porcentajes de reconocimiento razonablemente altos con la menor cantidad de información posible.
- 5) Se podrían explorar otras alternativas para el proceso de extracción de características que, tal vez, no necesariamente tengan que valerse de analizar una imagen por bloques sino, únicamente, de la aplicación de técnicas como las de segmentación, para hacer este proceso más eficiente.
- 6) Sobre la inicialización de parámetros de los HMMs, en [31] se menciona que una inicialización aleatoria puede conducir a mejores resultados, al evitar caer en máximos locales. Sin embargo, en las pruebas del Capítulo 5 se consideraron valores equiprobables para las probabilidades de transición, de estado inicial y de emisión, con el fin de obtener porcentajes de reconocimiento consistentes al repetir cada prueba de reconocimiento y para tener una primera estimación del comportamiento de los HMMs.
- 7) El SRRP da como resultado el sujeto de la BDR ORL que se asocia con mayor probabilidad a una imagen de prueba dada. Es decir, si se diera al SRRP una imagen del rostro de un sujeto que no se encuentra en dicha BDR, el sistema trataría de reconocer al sujeto más probable de la BDR ORL. A este respecto, podría explorarse la posibilidad de que, a partir de ciertos umbrales, el sistema sea capaz de informar si ese sujeto pertenece o no a la BDR ocupada.
- 8) Por otro lado, un grupo de aplicaciones de particular interés tiene que ver con SRRs en vídeo. Se puede plantear una posible extensión del SRRP para funcionar con vídeos

de rostros humanos, los que a su vez se pueden tratar como secuencias de imágenes (*cuadros*) [21]. Para ello, resulta necesario recurrir a métodos de segmentación y, posteriormente, a métodos propios del procesamiento de vídeo que permitan localizar y aislar, a partir de dichas secuencias, las imágenes a utilizar para el reconocimiento de sujetos de una BDR. Uno de tales métodos puede ser la denominada Estimación de Movimiento, ya que haría posible predecir la ubicación del rostro y, consecuentemente, la extracción de la imagen con la región de interés para usar por el sistema. Al obtener estas imágenes, se podrían aplicar los procesos propios del SRRP, vistos en el Capítulo 4. En la Figura 6.1 se ilustran las ideas anteriores. También, se podría mejorar el desempeño de esta versión del SRRP al procurar detectar posiciones frontales del rostro de los sujetos.

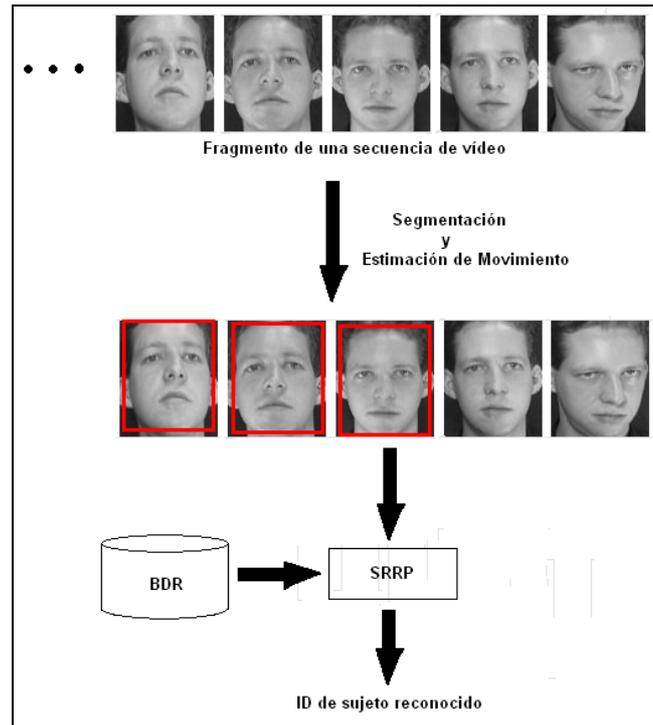


Figura 6.1. Posible extensión del SRRP al uso secuencias de vídeo.

6.3 Posibles aplicaciones

El SRRP puede ser aplicado de distintas formas, aunque se deben tomar en cuenta las condiciones y la información disponible con las que se estará trabajando. Las condiciones a las que se hace referencia tienen que ver con el tipo de cooperación que se espera por parte de los usuarios, así como el modo en el cual el sistema estaría funcionando: identificación o autenticación. La información disponible se refiere principalmente a las BDRs, las cuales tendrían que ser adaptadas a dicho sistema.

Un ejemplo de aplicación del SRRP puede ser el control de acceso a lugares con usuarios recurrentes, como instituciones académicas, oficinas y laboratorios de investigación. En un SRR con mayor cantidad de datos disponibles para entrenamiento, es decir, con BDRs más extensas, se podría pensar en aplicaciones a lugares más concurridos.

Otra aplicación es la autenticación de usuarios de equipos de cómputo. En este caso y en el control de acceso, en la etapa de reconocimiento del SRRP se pueden definir las condiciones que necesitan cumplir las imágenes de prueba. Esto está relacionado con la cooperación que el usuario debe mostrar en esta aplicación.

Existen sistemas cuya implementación está orientada al reconocimiento de algún rasgo humano en particular, como el iris o la huella digital. El SRRP podría usarse como complemento a dichos sistemas.

Finalmente, cabe señalar que los conceptos ocupados por el SRRP pueden utilizarse en otras aplicaciones que están fuera del campo del Reconocimiento de Rostros, como el reconocimiento de objetos en imágenes, en general, y el reconocimiento de voz.

Referencias

- [1] Grupo de Redes de Computadores. "Compresores Basados en la Transformada Wavelet". Universidad de Valencia. España.
<http://www.grc.upv.es/docencia/tdm/practicas/P4.pdf>
- [2] National Science and Technology Council. Subcommittee on Biometrics. "Face Recognition". 2006. <http://www.biometrics.gov/Documents/facerec.pdf>
- [3] Cambridge University Computer Laboratory. AT&T Laboratories. "The Database of Faces". Cambridge. U. K.
<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>
- [4] Bicego, M.; Castellani, U.; Murino, V. "Using Hidden Markov Models and Wavelets for face recognition". Proceedings of the 12th International Conference on Image Analysis and Processing, pp. 52 – 56, 2003.
- [5] Bovik, Al. "Handbook of Image and Video Processing". Academic Press. C. A., U. S. A., 2000.
- [6] Chen, C. H.; Wang, P. S. P. "Handbook of Pattern Recognition & Computer Vision". 3rd edition. World Scientific Publishing. Singapore, 2005.
- [7] Cova, Walter J. D.; Caballero, Rodolfo A. "Sobre Wavelets e Imágenes". Centro Universitario de Desarrollos en Automación y Robótica. Universidad Tecnológica Nacional. Argentina, 2006.
http://www.edutecne.utn.edu.ar/DOCUMENTOS/Sobre%20Wavelets%20e%20Imagenes_R1.pdf
- [8] Duda, R. O.; Hart, D. E.; Stork, D. G. "Pattern Classification". 2nd edition. John Wiley and Sons. New York, U. S. A., 2001.
- [9] Ekenel, H.K.; Stallkamp, J.; Gao, H.; Fischer, M.; Stiefelhagen, R. "Face Recognition for Smart Interactions". IEEE International Conference on Multimedia and Expo, pp. 1007 - 1010, 2007.
- [10] Gales, M.; Young, S. "The Application of Hidden Markov Models in Speech Recognition". Foundations and Trends® in Signal Processing, Vol. 1, No. 3, pp. 195–304, 2007.
- [11] Garcia C.; Zikos, G.; Tziritas, G. "A Wavelet-based Framework for Face Recognition". Int. Workshop on Advances in Facial Image Analysis and Recognition Technology, 5th European Conference on Computer Vision, pp. 84 - 92, 1998.
- [12] García M., Borja J.; Mancilla A., Eva; Montes F., Ruth. "Optimización de la Transformada Wavelet Discreta (DWT)". Universidad Complutense de Madrid. España, 2005. <http://eprints.ucm.es/9017/1/TC2005-39.pdf>
- [13] Gonzalez, Rafael C.; Woods, Richard E.; Eddins, Steven L. "Digital Image Processing Using MATLAB®". 2nd edition. Prentice Hall. New Jersey, U. S. A., 2004.
- [14] Grgic, Mislav; Delac, Kresimir. "Face Recognition Home Page". <http://www.face-rec.org>
- [15] Häggström, Olle. "Finite Markov Chains and Algorithmic Applications". Cambridge University Press. New York, U. S. A., 2002.

- [16] Jain, A. K. "Fundamentals of Digital Image Processing". Prentice Hall. New Jersey, U. S. A., 1989.
- [17] Kohir, Vinayadatt V.; Desai, U. B. "Face Recognition Using a DCT-HMM Approach". Proceedings of the 4th IEEE Workshop on Applications of Computer Vision, pp. 226-231, 1998.
- [18] Le, Hung-Son; Li, Haibo. "Simple 1D Discrete Hidden Markov Models for Face Recognition". Proceeding of Visual Content Processing and Representation, 8th International Workshop, pp. 41-49. Madrid, Spain, 2003.
- [19] Li, Stan Z.; Jain, Anil K. "Handbook of Face Recognition". 2nd edition. Springer. New York, U. S. A., 2011.
- [20] Liu, Xiaoming; Chen, Tsuhan. "Video-Based Face Recognition Using Adaptive Hidden Markov Models". Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 340-345, 2003.
- [21] Marques, Oge. "Practical Image and Video Processing Using MATLAB®". John Wiley & Sons, Inc. New Jersey, U. S. A., 2011.
- [22] Nefian, Ara V.; Hayes III, Monson H. "Face Recognition Using An Embedded HMM". Proceedings of the IEEE Conference on Audio and Video-based Biometric Person Authentication, pp. 19-24, 1999.
- [23] Ojeda, Ignacio; Gago, Jesús. "Métodos Matemáticos para Estadística". Universidad de Extremadura. Servicio de Publicaciones. España, 2008.
- [24] Pajares, G.; De la Cruz, Jesús M. "Visión por computador: imágenes digitales y aplicaciones". 2a edición. Alfaomega. México, 2008.
- [25] Rabiner, Lawrence R. "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition". Proceedings of IEEE, Vol. 77, No. 2, pp. 257-286, 1989.
- [26] Samaria, Ferdinando S. "Face Recognition Using Hidden Markov Models". PhD. Thesis. University of Cambridge. U. K., 1994.
- [27] Sebe, Nicu; Cohen, Ira; Garg, Ashutosh; Huang, Thomas S. "Machine Learning in Computer Vision". Springer. The Netherlands, 2005.
- [28] Sebe, Nicu; Lew, Michael S. "Robust Computer Vision. Theory and Applications". Kluwer Academic Publishers. The Netherlands, 2003.
- [29] Shen, Dawei. "Some Mathematics for HMM". Massachusetts Institute of Technology. 2008. <http://courses.media.mit.edu/2010fall/mas622j/ProblemSets/ps4/tutorial.pdf>
- [30] Shermina, J. "Illumination Invariant Face Recognition Using Discrete Cosine Transform and Principal Component Analysis". International Conference on Emerging Trends in Electrical and Computer Technology, pp. 826-830, 2011.
- [31] Stamp, Mark. "A Revealing Introduction to Hidden Markov Models". San Jose State University. Department of Computer Science. 2012. <http://www.cs.sjsu.edu/~stamp/RUA/HMM.pdf>
- [32] Steeb, Willi-Hans; Hardy, Yorick; Stoop, Ruedi. "The Nonlinear Workbook". 3rd edition. World Scientific Publishing. Singapore, 2005.
- [33] Strook, Daniel W. "An Introduction to Markov Processes". Springer. New York, U. S. A., 2005.

- [34] Sushov, Yuri; Kelbert, Mark. "Probability and Statistics by Example: II Markov Chains: a Primer in Random Processes and their Applications". Cambridge University Press. New York, U. S. A., 2008.
- [35] Szeliski, Richard. "Computer Vision. Algorithms and Applications". Springer. New York, U. S. A., 2011.
- [36] Tang, Yuan Y. "Wavelet Theory Approach to Pattern Recognition". 2nd edition. World Scientific Publishing. Singapore, 2009.
- [37] Turk, Matthew A.; Pentland, Alex P. "Face Recognition Using Eigenfaces". Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 586-591, 1991.
- [38] Walhoff, Frank; Eickeler, Stefan; Rigoll, Gerhard. "A Comparison of Discrete and Continuous Output Modeling Techniques for a Pseudo-2D Hidden Markov Model Face Recognition System". Proceedings of International Conference on Image Processing, pp. 685-688, 2001.
- [39] Yu, Hua; Yang, Jie. "A Direct LDA Algorithm for High-Dimensional Data - with Application to Face Recognition". Pattern Recognition, Vol. 34, No. 10, pp. 2067-2070, 2001.
- [40] Zaho, W.; Chellappa, R.; Phillips, P. J.; Rosenfeld, A. "Face Recognition: A Literature Survey". ACM Computing Surveys, Vol. 35, No. 4, pp. 399 - 458, 2003.
- [41] Zhang, Yu-Jin. "Advances in Face Image Analysis: Techniques and Technologies". Medical Information Science Reference. New York, U. S. A., 2011.
- [42] Zhao, Ming; Chua, Tat-Seng. "Markovian Mixture Face Recognition with Discriminative Face Alignment". 8th IEEE International Conference on Automatic Face & Gesture Recognition, pp. 1-6, 2008.
- [43] Zucchini, Walter; MacDonald, Iain L. "Hidden Markov Models for Time Series: An Introduction Using R". CRC Press. New York, U. S. A., 2009.

Apéndice A. Descripción de la BDR ORL

En este trabajo se utiliza la BDR Olivetti Research Laboratory (ORL) [3]. Las imágenes fueron tomadas por los *AT&T Laboratories Cambridge* y han sido usadas por el Departamento de Ingeniería de la Universidad de Cambridge, y en trabajos como [26], [4] y [17], entre otros.

Esta BDR está formada por 40 sujetos, cada uno con 10 imágenes de sus rostros en posición frontal con pequeños movimientos laterales. Las imágenes tienen un tamaño de 92x112 y tienen formato pgm (portable gray map) con almacenamiento binario y considerando 256 niveles de gris.

Entre otros aspectos de esta BDR, se puede mencionar que, para algunos sujetos, se introdujeron ligeras variaciones de iluminación, expresiones faciales y accesorios como el uso de lentes, para algunas de las diez posiciones. Todas las imágenes fueron tomadas con un fondo oscuro homogéneo. En la Tabla A.1 se muestran todos los sujetos de la BDR ORL.

ID	Imagen	ID	Imagen	ID	Imagen	ID	Imagen	ID	Imagen
1		9		17		25		33	
2		10		18		26		34	
3		11		19		27		35	
4		12		20		28		36	
5		13		21		29		37	
6		14		22		30		38	
7		15		23		31		39	
8		16		24		32		40	

Tabla A.1. Sujetos de la BDR ORL.

Apéndice B. Cuantización Vectorial

B.1 Conceptos generales

La Cuantización Vectorial es un método para reducción de datos, el cual consiste en representar un conjunto de n muestras de entrenamiento a través de otro conjunto finito de m muestras definidas en el mismo espacio que las muestras de entrenamiento, donde $m < n$. Para realizar una Cuantización Vectorial es necesario definir lo que se conoce como un *cuantizador vectorial*.

Cuantizador vectorial. Es una proyección del espacio \mathbb{R}^d en un subconjunto finito C del mismo espacio, definida como [24]:

$$Q: \mathbb{R}^d \rightarrow C \quad (\text{B.1})$$

donde $C = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbb{R}^d; i = 1, \dots, M\}$. Este conjunto se denomina libro de códigos (*codebook*), con dimensión d y tamaño M , es decir, está formado por M elementos llamados vectores de código (*codevectors*). Para un cuantizador vectorial se definen, a su vez, las siguientes dos funciones.

Codificador Vectorial. Función definida como:

$$\varepsilon: \mathbb{R}^d \rightarrow I \quad (\text{B.2})$$

donde $I = \{1, \dots, M\}$ es el conjunto de índices de los elementos de C . Este mapeo requiere de una medida que permita comparar los vectores de \mathbb{R}^d con los vectores de código. Para este trabajo se toma dicha medida como una función de distancia d .

Decodificador Vectorial. Función definida como:

$$\delta: I \rightarrow C \quad (\text{B.3})$$

donde I es el mismo conjunto utilizado en el Codificador vectorial y C es el codebook correspondiente. δ es un mapeo que asocia a cada índice $i \in I$ con el vector de código $\mathbf{c}_i \in C$.

Por lo general, se recurre a algoritmos de agrupamiento para la obtención de C , a partir de vectores de entrenamiento. Uno de los algoritmos más usados es el llamado *k-medias*, el cual se explica en la sección B.2.

B.2 Algoritmo de agrupamiento k-medias

A continuación, se describe el algoritmo de agrupamiento k-medias, el cual clasifica un conjunto de muestras en k clases [24], [8], de acuerdo con una función de distorsión determinada (d) y un conjunto de vectores promedio conocidos como centroides. Como resultado, además de dicha agrupación de muestras, se obtiene para cada grupo su centroide actualizado.

Algoritmo B.1. Algoritmo k-medias
<p>X: Conjunto de m muestras a agrupar. $X = \{\mathbf{x}_i \mid \mathbf{x}_i \in \mathbb{R}^n; i = 1, \dots, m\}$</p> <p>$k$: Número de clases.</p> <p>d: Función de distancia.</p> <p>C: Conjunto de centroides iniciales. $C = \{\mathbf{c}_i \mid \mathbf{c}_i \in \mathbb{R}^n; i = 1, \dots, k\}$</p> <p>$C^*$: Conjunto de centroides actualizados. $C^* = \{\mathbf{c}_i^* \mid \mathbf{c}_i^* \in \mathbb{R}^n; i = 1, \dots, k\}$</p> <p>$G_i$: Conjunto de muestras de la clase i; $i = 1, \dots, k$.</p> <p>N_i: Número de muestras asignadas al grupo G_i; $i = 1, \dots, k$.</p>
<p>$C^* = C$</p> <p>hacer</p> <p style="padding-left: 2em;">$G_i = \emptyset; i = 1, \dots, k$</p> <p style="padding-left: 2em;">$C = C^*$</p> <p style="padding-left: 2em;">para $i = 1, \dots, m$ hacer</p> <p style="padding-left: 4em;">$j = \arg \min_{l=1}^k (d(\mathbf{x}_i, \mathbf{c}_l))$</p> <p style="padding-left: 4em;">$G_j = G_j \cup \{\mathbf{x}_i\}$</p> <p style="padding-left: 2em;">fin para</p> <p style="padding-left: 2em;">para $i = 1, \dots, k$ hacer</p> <p style="padding-left: 4em;">$\mathbf{c}_i^* = \frac{1}{N_i} \sum_{\mathbf{x}_j \in G_i} \mathbf{x}_j$</p> <p style="padding-left: 2em;">fin para</p> <p>mientras ($C^* \neq C$)</p> <p>devolver $G_i; i = 1, \dots, k; C^*$</p>

Para este trabajo se utilizó como función de distorsión a la distancia Euclídea, definida de la siguiente forma:

Dados dos vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$,

$$d(\mathbf{x}, \mathbf{y}) = \left[\sum_{i=1}^n (x_i - y_i)^2 \right]^{\frac{1}{2}} \quad (\text{B.4})$$

