UNIVERSIDAD MIGUEL HERNÁNDEZ DE ELCHE

ESCUELA POLITÉCNICA SUPERIOR DE ELCHE

INGENIERÍA DE TELECOMUNICACIONES

"COMPARACIÓN DE LOS MÉTODOS PCA Y WHITENING PARA EL RECONOCIMIENTO DE CARAS ANTE DIFERENTES PREPROCESADOS"

PROYECTO FIN DE CARRERA

MAYO 2017

AUTOR: EDGAR ROCHERA ORENGA

DIRECTOR: CÉSAR FERNÁNDEZ PERIS

VISTO BUENO Y CALIFICACIÓN DEL PROYECTO

Título proyecto:

Proyectante: _____

Director/es: _____

_____

VºBº director/es del proyecto: Fdo.: Fdo.:

Lugar y fecha: _____

CALIFICACIÓN NUMÉRICA ☐ MATRÍCULA DE HONOR ☐

Méritos justificativos en el caso de conceder Matrícula de Honor:

_____

_____

_____

_____

Conforme presidente: Fdo.:	Conforme secretario: Fdo.:	Conforme vocal: Fdo.:

Lugar y fecha: _____

ÍNDICE GENERAL

1. Objetivo del proyecto 1

2. Reconocimiento de caras 3

2.1. Detección vs reconocimiento 5

2.2. Técnicas alternativas de reconocimiento biométrico vs reconocimiento facial 8

3. Fundamentos teóricos de los métodos PCA y Whitening 11

3.1. Pca por maximización de la varianza 12

3.2. Pca por minimizacion del error cuadrático medio 21

3.3. Consideraciones sobre PCA 23

3.4. Blanqueo de datos, Whitening 24

4. Bases de datos utilizadas 27

4.1. Base de datos AR 27

4.2. Base de datos AT&T 30

4.3. Base de datos LFW 32

4.4. Base de datos FERET 35

4.5. Base de datos UMH 37

5. Preprocesados 39

5.1. Ecualización del histograma 39

5.2. Máscara elipsoidal 43

5.3. Preprocesado Tan-Triggs 44

5.4. Viola-Jones 48

6. Metodología 53

7. Experimentos 65

7.1. Experimento 1 65

7.1.1. Base de datos AR 65

7.1.2. Base de datos ORL 70

7.1.3. Base de datos UMH 75

7.1.4. Base de datos LFW 80

7.1.5. Base de datos FERET 85

7.1.6. Resumen experimento 1 90

7.2. Experimento 2 92

7.2.1. Base de datos AR 92

7.2.2. Base de datos ORL 95

7.2.3. Base de datos UMH 98

7.2.4. Base de datos LFW 101

7.2.5. Base de datos FERET 104

7.2.6. Resumen experimento 2 107

7.3. Experimento 3 109

7.3.1. Preprocesado 000 109

7.3.2. Preprocesado 001 111

7.3.3. Preprocesado 010 113

7.3.4. Preprocesado 011 115

7.3.5. Preprocesado 020 117

7.3.6. Preprocesado 021 119

7.3.7. Preprocesado 100 121

7.3.8. Preprocesado 101 123

7.3.9. Preprocesado 110 125

7.3.10. Preprocesado 111 127

7.3.11. Preprocesado 120 129

7.3.12. Preprocesado 121 131

7.3.13. Resumen experimento 3 133

7.4. Experimento 4 135

7.4.1. Ratio de reconocimiento por base de datos y preprocesado 135

7.4.2. Comparativa de preprocesados por base de datos 137

7.4.3. Mejor preprocesado por base de datos 141

8. Conclusiones 143

1. OBJETIVO DEL PROYECTO

El estudio que se desarrolla a continuación tiene como objetivo comparar dos técnicas utilizadas en el reconocimiento facial, PCA y Whitening. Estas técnicas serán aplicadas sobre unas bases de datos, posteriormente descritas, para estudiar cuál de las dos alcanza unas mejores tasas de reconocimiento. Sobre las bases de datos se aplicarán distintos preprocesados, siendo los mismos para ambas técnicas, con la finalidad de comprobar el comportamiento de las distintas técnicas aplicadas ante distintos preprocesados. Estos últimos también forman parte del presente estudio, donde se estudiará cuál o cuáles son los mejores a aplicar dependiendo de las técnicas utilizadas e independientemente de estas.

2. RECONOCIMIENTO DE CARAS

El primer sistema semi automático de reconocimiento de caras conocido fue creado por Woody Bledsow, Helen Chan Wolf y Charles Bisson durante los años 1964 y 1965. Este sistema requería de la interactuación de un administrador para localizar características como pueden ser los ojos, oídos, nariz y la boca, para después calcular las distancias y ratios a un punto común, para más tarde compararlos con datos de referencia.

En el 1971 los investigadores, A. J. Goldstein, L. D. Harmon y A. B. Les, utilizaron marcadores subjetivos como el color del cabello, grosor de labios, etc, para automatizar el proceso. El gran inconveniente de estas técnicas fue que las mediciones y las ubicaciones eran calculadas manualmente.

En 1988, L. Sirovich y M. Kirby aplicaron el concepto de análisis de componentes principales (PCA) para el problema del reconocimiento facial, demostrando que se necesitaban menos de 100 valores para codificar con precisión una imagen de una cara debidamente alineada y normalizada. Es considerado un hito en el campo del reconocimiento computacional.

Durante 1991, Turk y Pentland descubrieron que se podía utilizar el error residual para la detección de rostros mientras estaban utilizando las técnicas de eigenfaces. Gracias a ese hallazgo se permitió la puesta en marcha de los primeros sistemas en tiempo real automatizados. Pese a los problemas causados por la dependencia del entorno en la detección de rostros, llamó la atención de los desarrolladores.

Hoy en día la tecnología de reconocimiento facial es utilizada para combatir la falsificación de pasaportes, identificación de niños desaparecidos, minimizar el fraude en las prestaciones de identidad como en cajeros automáticos, accesos a estancias VIP, empresas…

Actualmente también tenemos otro tipo de usos que pueden poner en peligro la privacidad de los usuarios, ya que están empezando a proliferar aplicaciones como las que nos ofrece NameTag o LAMBDA LABS para GOOGLE GLASS donde con una fotografía del rostro es capaz de compararla con las imágenes en redes sociales como FACEBOOK, TWITTER, LINKED IN o incluso datos en bases de datos públicos y

registros civiles. Pero también es capaz de cotejarlas contra una base de datos de más de 450.000 acosadores sexuales. También existe dentro de su página web un apartado para aquellos que quieran preservar su privacidad quedando fuera de sus bases de datos.

El departamento de investigación de inteligencia artificial de Facebook, investigadores: Yaniv Taigman Ming, Yang Marc'Aurelio, Ranzato, junto con la universidad de Tel Aviv, investigador: Lior Wolf, han publicado el artículo: "*DeepFace: Closing the Gap to Human-Level Performance in Face Verification"*, donde presentan el sistema DeepFace. Este nuevo sistema de reconocimiento ha sido entrenado bajo una base de datos de 4 millones de imágenes sacadas de FACEBOOK. Dicho software utiliza técnicas de modelado 3D y de redes neuronales artificiales. Indican que, a diferencia de los actuales sistemas de reconocimiento facial, donde se utilizaron decenas de miles de descriptores de imágenes manuales, ellos directamente lo aplican a los valores RGB de los pixeles, produciendo un descriptor muy compacto. El sistema DeepFace trabaja bajo entornos no controlados, sin importar la iluminación, expresión facial, ángulo…

Bajo la base de datos LFW (Labeled Faces in the Wild) han obtenido una precisión de un 97.35%, reduciendo el error actual del reconocimiento facial en más de un 27%, acercándose estrechamente al reconocimiento humano.

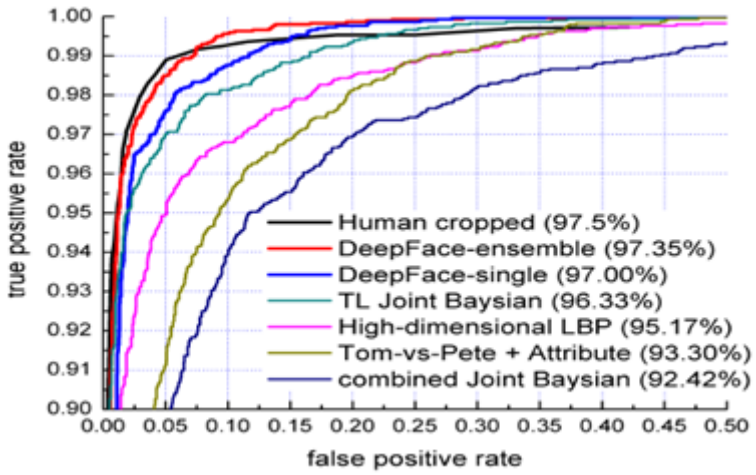


FIGURA 2.1

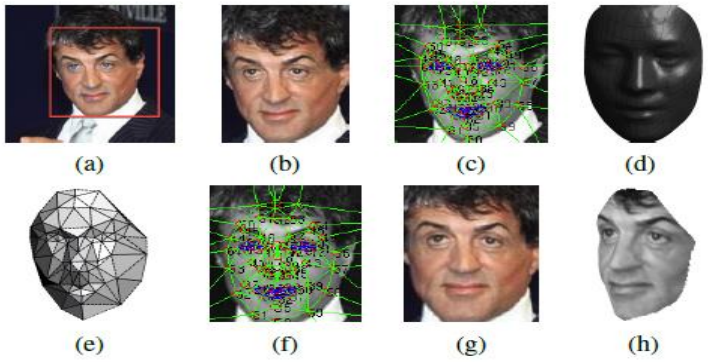


Figure 1. **Alignment pipeline.** (a) The detected face, with 6 initial fiducial points. (b) The induced 2D-aligned crop. (c) 67 fiducial points on the 2D-aligned crop with their corresponding Delaunay triangulation, we added triangles on the contour to avoid discontinuities. (d) The reference 3D shape transformed to the 2D-aligned crop image-plane. (e) Triangle visibility w.r.t. to the fitted 3D-2D camera; darker triangles are less visible. (f) The 67 fiducial points induced by the 3D model that are used to direct the piece-wise affine warpping. (g) The final frontalized crop. (h) A new view generated by the 3D model (not used in this paper).

FIGURA 2.2

2.1. DETECCIÓN VS RECONOCIMIENTO

En muchas ocasiones cuando se habla del término reconocimiento, este es confundido por el término llamado detección.

La detección en una imagen se puede definir como: dada una imagen, discernir si hay algún objeto en ella, en caso de haber alguno, indicar donde está ubicada dicho objeto y el tamaño del mismo. En el caso que se va a estudiar los objetos serán los rostros.

En la actualidad existen dispositivos electrónicos como terminales móviles, cámaras fotográficas, televisores… dotados de una cámara y software que son capaces de realizar esta detección de rostros bien en tiempo real, video o dada una imagen. El sistema de detección previamente debe ser entrenado para aprender el tipo de objeto a detectar. Se entrena con una base de datos donde hay presentes rostros junto con otra base de datos en la que hay otro tipo de objetos diferentes. La finalidad es que el sistema sea capaz de aprender a distinguir qué es y qué no es un rostro.

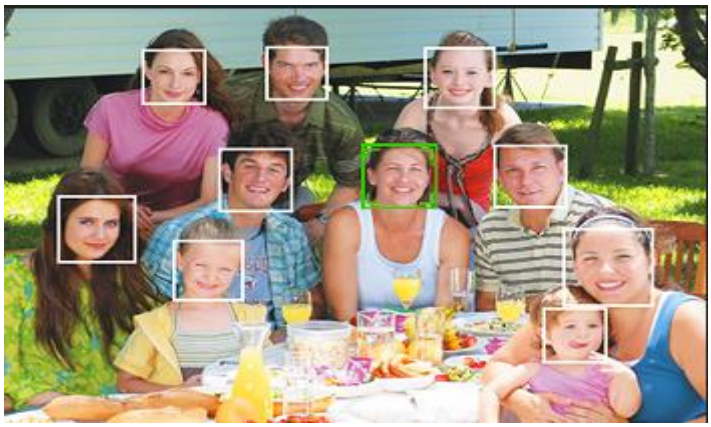


FIGURA 2.3: Detección de rostros

El reconocimiento se da cuando un sistema es capaz, dada una o varias imágenes de un rostro, de identificar a qué rostro dentro de una base de datos almacenada pertenece el rostro de la imagen facilitada. Dentro del proceso del reconocimiento la detección de rostros es una fase vital, ya que si no se detecta ningún rostro no se puede llevar a cabo las siguientes fases.

El reconocimiento de caras es un proceso que puede ser diferenciado en 4 fases clave:

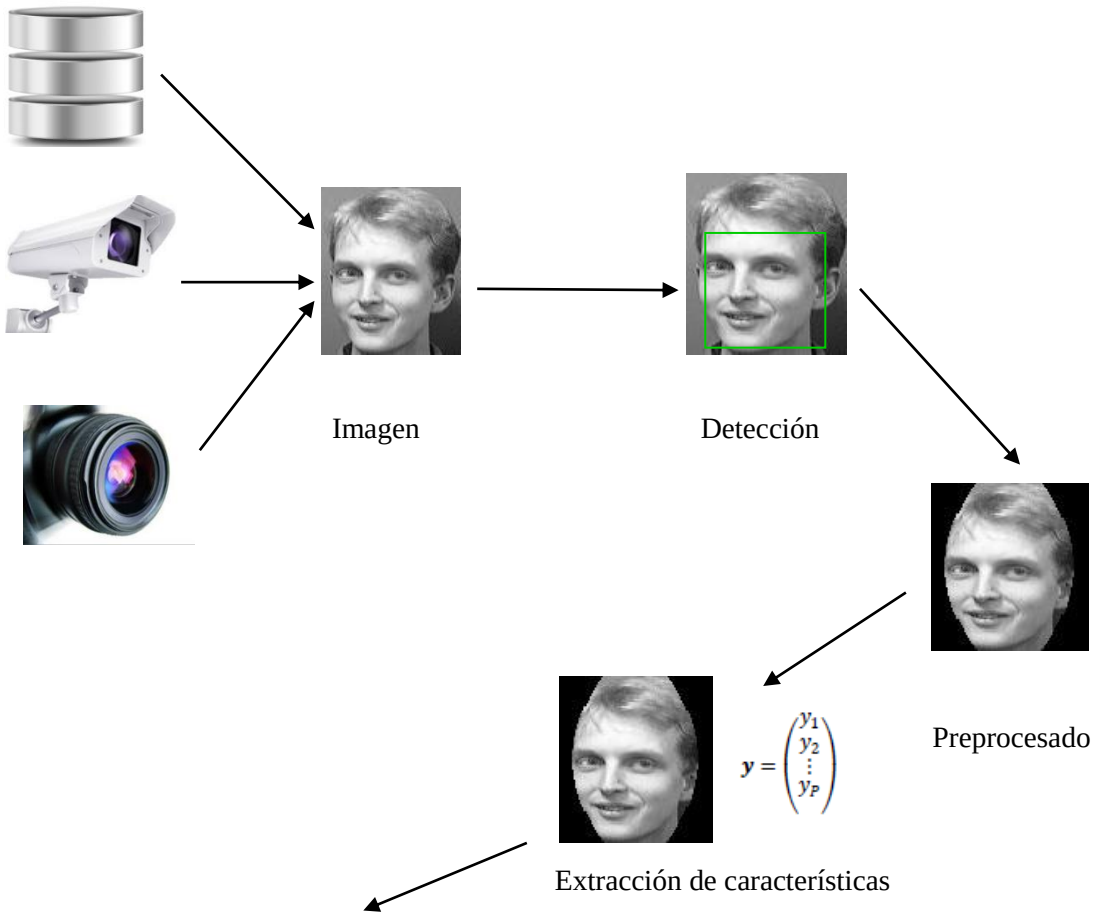
1. DETECCIÓN: En esta fase, como se ha definido anteriormente, es la primera del proceso de reconocimiento de caras ya que en ella se debe detectar un rostro dado una imagen en tiempo real o no.

2. PREPROCESADOS: Fase en la cual se aplican varias técnicas sobre las imágenes, bien sea, para mejorar el procesado de los datos en fases posteriores o para intentar aumentar el éxito de un reconocimiento correcto, mejorando la imagen del rostro detectado en la fase de detección, devolviendo la ubicación y el tamaño de este. Para ello se emplean técnicas como: normalizado, suavizado, escala de grises, histograma, etc…

3. EXTRACCION DE CARACTERÍSTICAS: Fase donde se aplican técnicas como ICA, PCA… para obtener información de los rostros.

4. RECONOCIMIENTO: Fase en la cual se compara el vector facilitado tras el módulo de extracción de características con una base de datos previamente creada. El reconocimiento será positivo cuando, comparando con la base de datos, el porcentaje de semejanza supere cierto valor que previamente se ha configurado, siendo negativo cuando no supere dicho porcentaje.

En algunas ocasiones el sistema nos reconoce a la persona X como Y, devolviéndolo como un acierto, aunque realmente es un error. Este concepto se conoce como falso positivo.



Compara BBDD de características de otros sujetos para darnos el sujeto con el porcentaje más elevado

FIGURA 2.4: Diagrama de las fases que puede constar el proceso de reconocimiento facial desde la captación del rostro hasta el reconocimiento del mismo.

2.2.TÉCNICAS ALTERNATIVAS DE RECONOCIMIENTO BIOMÉTRICO VS RECONOCIMIENTO FACIAL

Desde hace miles de años el ser humano viene utilizando métodos de identificación biométrica, como puede ser la longitud de los brazos utilizada en la Antigua Egipto 2600AC, o una huella dactilar impresa en arcilla como firma de transacciones en los mercados de Babilonia 500AC.

Hoy en día existen diversos métodos de reconocimiento biométricos que son utilizados de forma individual o como complemento unos de otros para cubrir las deficiencias que puedan tener debido a su naturaleza.

Las técnicas de biometría existentes más utilizadas pueden ser: huella dactilar, reconocimiento facial, análisis del iris y/o retina, reconocimiento de voz, firma o escritura manuscrita, termograma de venas de rostros y manos, ADN... Dentro de los beneficios y contras de cada uno de ellos, se puede mencionar como ejemplo, el reconocimiento por análisis de iris, caso en el que no existen dos iguales, incluso en los gemelos idénticos son distintos. La textura del iris se forma en el desarrollo del feto y se estabiliza dentro de los dos primeros años de vida, con lo que una de las mayores ventajas es que no es necesario actualizar la información biométrica de un sujeto ya que no variará durante el tiempo. La desventaja notoria de esta tecnología es el alto coste de la misma, aunque con el paso de los años ha ido bajando. El mayor problema de esta técnica es el carácter intrusivo para el sujeto, el cual no tiene una buena aceptación generalmente. Comparando dicho reconocimiento con la del reconocimiento facial se puede destacar que una de las grandes ventajas de este último puede ser el grado de aceptación ya que no requiere de una intervención del sujeto, así como la posibilidad de identificar a múltiples sujetos al mismo tiempo. El coste de la tecnología utilizada en ambas técnicas, también es favorable a la técnica de reconocimiento facial debido al bajo coste de este. Como desventaja, se puede mencionar la necesidad de actualizar la información de la biometría facial debido al cambio del rostro durante el paso de los años o bien por los causados por operaciones de cirugía plástica. Otra de las desventajas puede ser la naturaleza de la actual tecnología utilizada, debido a la alta sensibilidad de las condiciones de iluminación, rotación, inclinaciones de los rostros, etc.

A continuación se adjunta una tabla donde se resumen, dentro de las técnicas de reconocimiento biométrico más utilizadas, la fiabilidad de las mismas, el grado de aceptación…

Técnica	Fiabilidad	Facilidad de uso	Prevención de ataques	Aceptación	Estabilidad
Ojo (Iris)	Muy alta	Media	Muy alta	Media	Alta
Ojo (Retina)	Muy Alta	Baja	Muy Alta	Baja	Alta
Huellas dactilares	Muy Alta	Alta	Alta	Alta	Alta
Vascular dedo	Muy Alta	Muy Alta	Muy Alta	Alta	Alta
Vascular mano	Muy Alta	Muy Alta	Muy Alta	Alta	Alta
Geometría de la mano	Alta	Alta	Alta	Alta	Media
Escritura y firma	Media	Alta	Media	Muy Alta	Baja
Voz	Alta	Alta	Media	Alta	Media
Cara 2D	Media	Alta	Media	Muy alta	Media
Cara 3D	Alta	Alta	Alta	Muy alta	Alta

FIGURA 2.5: RECONOCIMIENTO DE CARAS VS OTROS RECONOCIMIENTOS BIOMETRICOS

Existen distintos y nuevos métodos de reconocimiento biométrico, como puede ser el ritmo cardíaco, el olor corporal, el análisis de la córnea, etc. Estos se encuentran en una continua fase de desarrollo y verán la luz en cuestión de años junto con la aparición de nuevas tecnologías para llevarlos a cabo. También existen diversos métodos de reconocimiento biométrico en fase de desarrollo, como pueden ser el pulso de la sangre, las arrugas del dedo, la transmisión de sonido de los huesos, etc.

3. FUNDAMENTOS TEÓRICOS DE LOS MÉTODOS PCA Y WHITENING

La técnica de componentes principales es debida a Hotelling (1933), aunque sus orígenes se encuentran en los ajustes ortogonales por mínimos cuadrados introducidos por Karl Pearson (1901).

El Análisis de Componente Principales (PCA) es una técnica clásica muy utilizada para el análisis estadístico de datos, la extracción de características y la compresión de datos.

PCA es una transformación lineal desde un espacio origen de dimensión muy elevada a un espacio de dimensiones más reducidas. Se produce una rotación del sistema de referencia del espacio original, de manera que la diversidad y la información de los datos originales se pueden describir utilizando sólo un subconjunto de vectores.

Para estudiar las relaciones que se presentan entre p variables correlacionadas (que miden información común, como en el caso que se pretende estudiar sobre imágenes de rostros) se puede transformar el conjunto original de variables en otro conjunto de nuevas variables incorreladas entre sí (que no tenga repetición o redundancia en la información) llamado conjunto de componentes principales.

Las nuevas variables son combinaciones lineales de las anteriores y se van construyendo según el orden de importancia en cuanto a la variabilidad total que recogen de la muestra.

De modo ideal, se buscan $n < p$ variables que sean combinaciones lineales de las p originales y que estén incorreladas, recogiendo la mayor parte de la información o variabilidad de los datos.

Si los datos están muy correlados en el espacio original, PCA produce buenos resultados ya que una alta correlación de los vectores de entrada significa que existe

información redundante, como suele pasar en las imágenes de rostros. PCA realiza una disminución de la redundancia, incorrelando los vectores de entrada como se ha mencionado anteriormente. Por lo tanto, los vectores de entrada de dimensión elevada y correlados pueden ser eficientemente representados en un espacio de dimensiones inferior donde los vectores no están correlados, de modo que llegamos a realizar una compresión de los datos.

3.1. PCA POR MAXIMIZACIÓN DE LA VARIANZA

Supongamos que se dispone de N vectores x_j de dimensión M alineados en una matriz de datos $X \in R^{M \times N}$. Sea **u** una dirección de R^M. La proyección del vector x_j sobre el vector **u** puede calcularse de la siguiente forma:

$$p_j = < x_j, u > = u^T x_j = \sum_{i=1}^{M} u_i x_{ij}$$

El objetivo es encontrar la dirección *u* que maximiza la varianza de las proyecciones de todos los vectores de entrada $x_j, j = 1, \ldots . . N$, por lo tanto la función a maximizar es:

$$J^{PCA}(u) = \sigma^2(p_j) = E\{p_j^2\}$$

Si la media de las proyecciones se calcula según el estimador siguiente:

$$\bar{p} = \frac{1}{N} \sum_{j=1}^{N} p_j = \frac{1}{N} \sum_{j=1}^{N} \sum_{i=1}^{M} u_i x_{ij} = \sum_{i=1}^{M} u_i \mu_i$$

la varianza a maximizar es igual a:

$$J^{PCA}(u) = \sigma^2(p_j) = E\{p_j^2\} =$$

$$= \frac{1}{N}\sum_{j=1}^{N}(p_j - \bar{p})^2 = \frac{1}{N}\sum_{j=1}^{N}\left(\sum_{i=1}^{M} u_i x_{ij} - \sum_{i=1}^{M} u_i \mu_i\right)^2 =$$

$$= \frac{1}{N}\sum_{j=1}^{N}\left(\sum_{i=1}^{M} u_i \hat{x}_{ij}\right)^2 = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{M}\sum_{l=1}^{M} u_i \hat{x}_{ij} u_l \hat{x}_{lj} =$$

$$= \sum_{i=1}^{M}\sum_{l=1}^{M} u_i u_l \frac{1}{N} < \hat{x}_{i'}, \hat{x}_{l'} \geq \sum_{i=1}^{M}\sum_{l=1}^{M} u_i c_{il} u_l = u^T C_x u \qquad (3.1)$$

En la ecuación anterior, μ_i es la media de la fila *i* de la matriz de datos X y $\hat{x}_{ij}$ es el valor de x_{ij} sin el valor de la media μ_i. Así el vector μ está formado por las medias de todas las filas:

$$\mu = [\mu_1, \ldots., \mu_M]^T = \frac{1}{N}\sum_{j=1}^{N} x_j$$

por lo tanto:

$$\hat{X} = X - \mu 1_{1xN}$$

y C_x es la matriz de covarianzas de X:

$$C_x = \frac{1}{N}\hat{X}\hat{X}^T$$

Como ya se ha dicho, el objetivo es maximizar la función 3.1 y además bajo la restricción que $\|u\| = 1$. Puesto que la varianza depende tanto de la norma como de la orientación del vector *u***,** y aumenta sin límites si la norma también aumenta, cabe imponer la restricción de que la norma *u* sea constante y en la práctica, se elige igual a la unidad. Si se utiliza la técnica de los multiplicadores de Lagrange, la función a maximizar es la siguiente:

$$F(u, \lambda) = u^T C_x u - \lambda(u^T u - 1) = \sum_{i=j}^{M}\sum_{j=1}^{M} u_i c_{ij} u_j - \lambda\left(\sum_{i=1}^{M} u_i^2 - 1\right) \qquad (3.2)$$

Una posible solución a este problema de maximización puede ser la siguiente:

$$\frac{\partial F}{\partial u_l} = \sum_{j=1}^{M} c_{lj} u_j + \sum_{i=1}^{M} u_i c_{il} - \lambda 2 u_l = 0; \qquad l = 1 \dots M$$

$$\sum_{i=1}^{M} c_{li} u_i = \lambda u_l; \qquad l = 1 \dots M$$

$$C_x u = \lambda u$$

Por lo tanto, para encontrar los valores de *u* y *λ* que maximizan 3.2 se deben calcular los vectores propios y los valores propios de la matriz de covarianzas de los datos C_x. El valor propio mayor es igual a la máxima varianza, mientras que su correspondiente vector propio determina la dirección de la máxima varianza en los datos.

Para diagonalizar la matriz de covarianzas de los datos C_x se realiza la descomposición en valores singulares (SVD) de la matriz C_x :

$$C_x = U \mathbf{\Lambda} U^{\mathbf{T}} \qquad (3.3)$$

en la expresión anterior la matriz ortonormal $U = [u_1, \dots, u_R] \in \Re^{MxR}$ contiene en sus columnas a los vectores propios $u_1, \dots, u_R$ y la matriz diagonal $\mathbf{\Lambda} \in \Re^{RxR}$ contiene en la diagonal a los valores propios $\lambda_1, \dots, \lambda_R$, siendo R el rango de la matriz C_x . A efectos prácticos, los valores propios y sus correspondientes vectores propios se ordenan de mayor a menor según el valor de los valores propios, es decir, $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_R$. Con este orden, la mayor variabilidad en los datos está reflejada en los primeros vectores propios.
Por lo tanto, los vectores propios se suelen también denominar vectores principales (o también ejes principales o direcciones principales).

Esta forma de calcular los vectores principales es la más clara y la más utilizada. Sin embargo, si el tamaño del vector de datos, M, es muy grande, que es el caso de las imágenes en visión por computador, la matriz de covarianzas $C_x \in \Re^{RxR}$ es demasiado grande y su diagonalización no es factible computacionalmente. Pero existe una solución: si el número de vectores de datos es menor que el tamaño de estos vectores (N < M), puede realizarse el análisis PCA de la matriz de datos original utilizando el

método propuesto por Murakami y Kumar y conocido en la literatura por truco de Turk y Pentland y que consiste en un simple truco algebraico. En lugar de utilizar la matriz de covarianzas de los datos $C_x \in \Re^{MxM}$ (producto externo de la matriz de datos) se utiliza la matriz de covarianzas $C'_x \in \Re^{NxN}$ (producto interno de la matriz de datos):

$$C'_x = \frac{1}{N}\hat{X}^T\hat{X}$$

Los vectores propios y valores propios de la matriz de covarianzas C_x pueden calcularse a partir de los vectores propios u'_i y valores propios λ'_i de C'_x de la siguiente manera:

$$\lambda_i = \lambda'_i \qquad i = 1 \dots N$$

$$u_i = \frac{\hat{X}u'_i}{\sqrt{N\lambda_i}} \qquad i = 1 \dots N$$

La matriz C'_x es mucho menor que la original C_x cuando $N \ll M$. Por lo tanto, la descomposición en vectores y valores propios de la matriz original $C_x \in \Re^{MxM}$ se ha reducido a la descomposición más factible de $C'_x \in \Re^{NxN}$.

PROPIEDADES DE PCA

La matriz ortonormal U que contiene los vectores principales puede utilizarse como una transformación lineal para proyectar los datos originales de un espacio de entrada de muy alta dimensión en otro espacio de más baja dimensión (subespacio) y viceversa. Las columnas de U son la base de vectores del nuevo subespacio de menor dimensión expresadas en coordenadas multidimensionales del espacio original. De esta manera, un vector de datos puede ser proyectado en el subespacio principal utilizando la matriz de transformación $U^T : \Re^M \rightarrow \Re^N$.

$$p = U^T \hat{x}$$

Los coeficientes p_j se calculan como las proyecciones de imagen de entrada sobre el vector principal:

$$p_j =< \hat{x}_j, u_j >= u_j^T \hat{x}_j = \sum_{i=1}^{M} u_{ij}\, \hat{x}_i; \qquad j = 1 \dots N$$

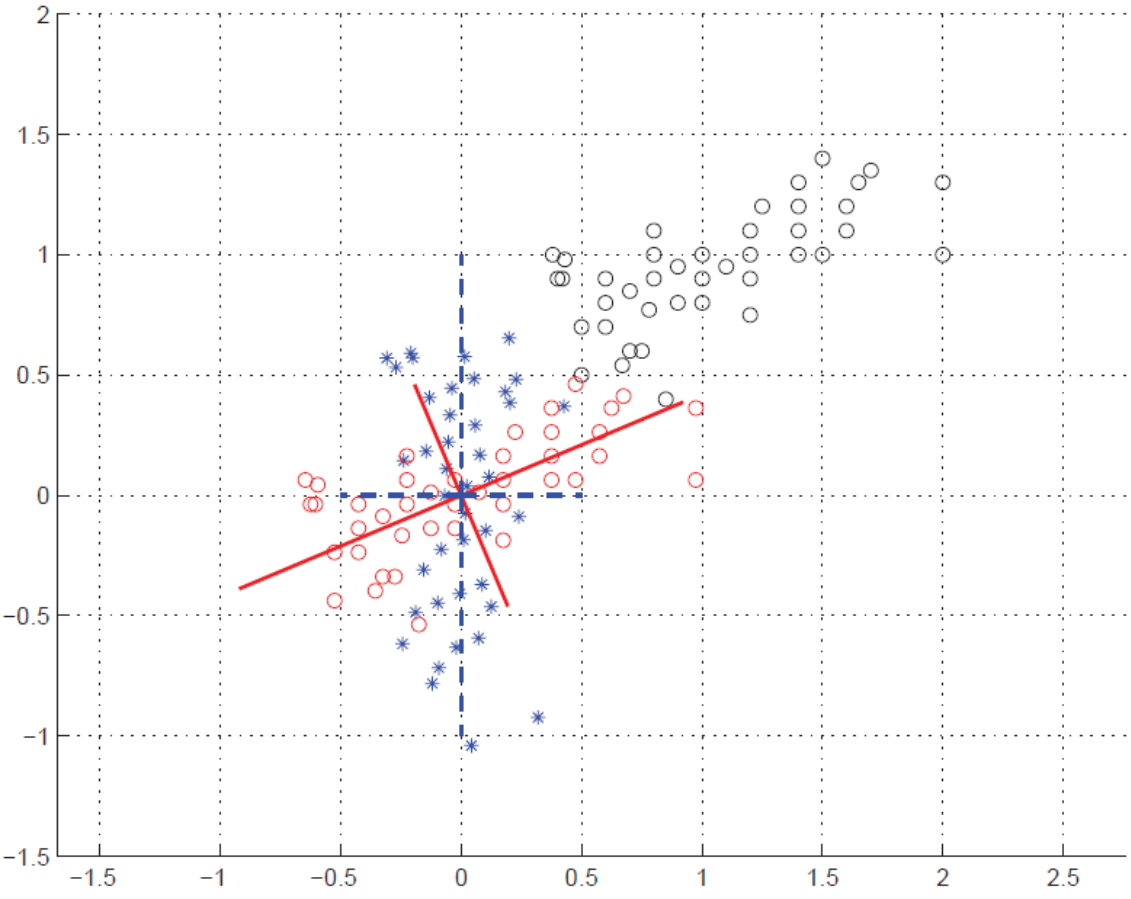


FIGURA 3.1: Ejemplo ilustrativo de la transformación PCA: datos originales (puntos negros), datos centrados (puntos rojos), datos proyectados en las direcciones de los vectores propios (puntos azules), los ejes principales están en azul.

Todos los vectores de entrada contenidos en la matriz de datos centrada, $\hat{X}$, pueden ser proyectados,

$$P = U^T \hat{X}$$

y la matriz que contiene las proyecciones P es una transformación ortonormal de los datos centrados $\widehat{X}$, así que la media de las proyecciones también está centrada:

$$\mu_P = \frac{1}{N}\sum_{j=1}^{N} p_j = \frac{1}{N}\sum_{j=1}^{N} U^{\mathbf{T}}\,\hat{x} = U^{\mathbf{T}}\frac{1}{N}\sum_{j=1}^{N}\hat{x} = 0$$

La figura 3.1 ilustra el proceso de la transformación PCA de unos datos artificiales: la información original son los puntos negros, los datos centrados se representan en rojo, los ejes principales (los vectores propios que definen la transformación) se muestran en azul y los datos proyectados en el nuevo espacio son los puntos azules.

Ahora, calculemos la matriz de correlación de las proyecciones:

$$C_P = \frac{1}{N}PP^T = \frac{1}{N}U^{\mathbf{T}}\widehat{X}\left(U^{\mathbf{T}}\widehat{X}\right)^T = U^{\mathbf{T}}\frac{1}{N}\widehat{X}\widehat{X}^T U =$$
$$= U^{\mathbf{T}}C_x U = U^{\mathbf{T}}U\boldsymbol{\Lambda}U^{\mathbf{T}}U = \boldsymbol{\Lambda}$$

En el desarrollo anterior se ha sustituido la matriz C_x por su forma diagonal (3.3) y se ha considerado la ortonormalidad de la matriz U (es decir que, $U^{\mathbf{T}}U = I$). Por lo tanto, la matriz de covarianzas de los datos transformados (o proyecciones) es igual a la matriz diagonal **Λ,** que contiene a los valores propios en su diagonal. Este hecho tiene dos consecuencias importantes, la primera de ellas, que demuestra que los vectores transformados están no correlados. Luego la redundancia debida a la correlación entre los vectores de entrada se ha eliminado. Y la segunda consecuencia es que se demuestra que la varianza en la dirección del eje principal i (la varianza de cada una de las componentes principales) es igual al valor del valor propio i, es decir, $\frac{1}{N}\sum_{j=1}^{N} p_{ij}^2 = \lambda_i$.

Una propiedad importante de la diagonalización (3.3) es que se preserva la traza de la matriz que ha sido diagonalizada. Dado que la suma de los elementos de la diagonal de la matriz de covarianzas es igual a la suma de las varianzas de los vectores de entrada, esto implica que la varianza total de los datos de entrada se ha conservado en la transformación y es igual a la suma de todos los valores propios:

$$\mathrm{VAR(X)} = \sum_{\mathrm{i=1}}^{\mathrm{M}} \frac{1}{\mathrm{N}} \hat{\mathrm{x}}\hat{\mathrm{x}}_{\mathrm{i}}^{\mathrm{T}} = \sum_{\mathrm{i=1}}^{\mathrm{M}} \mathrm{c_{ii}} =$$

$$\sum_{\mathrm{i=1}}^{\mathrm{N}} \lambda_{\mathrm{i}} = \sum_{\mathrm{i=1}}^{\mathrm{N}} \frac{1}{\mathrm{N}} \mathrm{p_{i'} p_{i'}^{T}} = \mathrm{VAR(P)}$$

RECONSTRUCCIÓN

A continuación se explica como utilizar la matriz de vectores propios U para proyectar un conjunto de coeficientes del subespacio en el espacio original; a esta operación se la denomina reconstrucción. El vector de coeficientes o proyecciones p puede ser reconstruido utilizando la matriz de transformación $\mathbf{U} : \Re^{\mathrm{N}} \rightarrow \Re^{\mathrm{M}}$:

$$\hat{y} = U_p = \sum_{j=1}^{N} p_j u_j$$

Si se utilizan los N vectores propios de $\mathbf{U} \in \Re^{\mathrm{M\ x\ N}}$para la reconstrucción, cada imagen de X puede ser perfectamente reconstruida sin ningún tipo de error en la reconstrucción. Pero lo que es más interesante es el hecho que una imagen pueda ser también reconstruida utilizando sólo un subconjunto de las componentes principales. Esto es posible debido a la distribución de la varianza en los ejes principales: la mayor parte de la varianza está contenida en los primeros vectores propios. Este hecho puede representarse utilizando una fracción de la varianza total del conjunto de datos que se suele denominar como la *energía* contenida en los primeros *k* vectores propios:

$$energía_k = \frac{\sum_{i=1}^{k} \lambda_{\mathrm{i}}}{\sum_{i=1}^{N} \lambda_{\mathrm{i}}}$$

Por lo tanto, es posible obtener una buena aproximación de las imágenes de entrada considerando solo un subconjunto de vectores propios (los asociados a los eigevalores de mayor magnitud). Luego, si se consideran sólo *k* vectores propios o componentes

principales, $k < N$, entonces el subespacio a utilizar se reduce a $U = [u_1 \dots u_k] \in \Re^{\mathrm{M\,x\,k}}$. Un vector de entrada puede ser proyectado en este subespacio de componentes principales de dimensión k utilizando la matriz de transformación $\mathbf{U^T} : \Re^{\mathrm{M}} \rightarrow \Re^{\mathrm{k}}$:

$$p = U^T \hat{x} = U^T (x - \mu)$$

$$p_j = < \hat{x}_j, u_j > = u_j^T \hat{x}_j = \sum_{i=1}^{M} u_{ij} \hat{x}_j = \sum_{i=1}^{M} u_{ij} (x_i - \mu_i); \qquad j = 1 \dots . k$$

y reconstruirse utilizando la transformación $\mathbf{U} : \Re^{\mathrm{k}} \rightarrow \Re^{\mathrm{M}}$:

$$\hat{y} = Up = \sum_{j=1}^{Mk} p_j u_j$$

$$y = \hat{y} + \mu$$

De esta manera una imagen puede aproximarse por una combinación lineal de los primeros k vectores propios.

La figura 3.1 ilustra el proceso completo de PCA descrito anteriormente: generación del subespacio y reconstrucción. En la parte superior, la primera fila recoge las imágenes en color de nueve conocidos pilotos de motociclismo (Valentino Rossi, Sete Gibernau, Dani Pedrosa, Jorge Lorenzo, Loris Capirossi, Max Biaggi, Toni Elías, Roberto Rolfo y Manuel Poggiali), son las imágenes de entrada utilizadas para generar el subespacio PCA, es decir forman la matriz de datos $\mathbf{X} : \Re^{\mathrm{M\,x\,N}}$, donde N=9 es el número de pilotos y $M = 273\ x\ 254\ x\ 3 = 208026$, es la dimensión de cada una de ellas. En la segunda fila, se muestran los ocho vectores propios obtenidos a partir del conjunto de imágenes anterior, **U**, y la media (última imagen de la derecha en la segunda fila) de éste, μ. En el resto de filas se muestran las reconstrucciones, y, de cada una de las imágenes originales a partir de subconjuntos del eigenespacio, es decir, se muestra la reconstrucción de cada imagen utilizando subconjuntos de vectores propios compuestos por 1,2,…8 vectores, la imagen situada a la izquierda es la reconstruida a partir de un sólo coeficiente y un sólo vector propio, y la situada más a la derecha es la reconstruida

a partir de todo el subespacio y por tanto es igual a la original. Al final de cada conjunto de imágenes reconstruidas se muestra el error de reconstrucción de cada una de ellas (se define en 3.4).

El vector de error de reconstrucción es igual a la diferencia entre el vector de entrada original y el reconstruido:

$$e = \hat{x} - \hat{y} = \sum_{j=1}^{N} p_j u_j - \sum_{j=1}^{k} p_j u_j = \sum_{j=k+1}^{N} p_j u_j$$

La medida de error más común es el cuadrado del vector del error de reconstrucción, el cual es igual a su norma L2 al cuadrado; si consideramos la ortonormalidad de los vectores propios, u_j, se obtiene:

$$e = \|e\|^2 = \left\| \sum_{j=k+1}^{N} p_j u_j \right\|^2 = \sum_{j=k+1}^{N} p_j^2 \qquad (3.4)$$

Por lo tanto, el error de reconstrucción es igual a la suma de los cuadrados de los coeficientes rechazados. En el ejemplo anterior, (figura 3.2), se muestra el error de cada una de las reconstrucciones de cada uno de los pilotos, y se aprecia como el error es diferente para cada uno de ellos, así para el primer piloto (Rossi), el error de reconstrucción es muy elevado en todas las reconstrucciones y sólo se elimina al utilizar el espacio de vectores propios completo, mientras que para otros pilotos, como el quinto (Capirossi), se obtiene una buena reconstrucción con solo 4 vectores propios.

Normalmente, los coeficientes rechazados no son conocidos, puesto que se desprecian sus vectores propios asociados y los coeficientes no se calculan, sin embargo, es posible estimar el valor del error de reconstrucción con los valores de las varianzas a lo largo de los vectores propios rechazados, que son iguales a sus valores propios correspondientes:

$$E\{e\} = \sum_{j=k+1}^{N} \lambda_j$$

De esta forma, la esperanza del error es igual a la suma de los valores propios despreciados en la reconstrucción. Esta última afirmación confirma el hecho que maximizando la varianza de los primeros vectores propios (los que se utilizan para la reconstrucción), el error de reconstrucción está siendo minimizado. Esta consideración da lugar a la posibilidad de plantear el problema de PCA desde el punto de vista de la minimización del error cuadrático medio, que se discutirá en el siguiente apartado.

3.2. PCA POR MINIMIZACION DEL ERROR CUADRÁTICO MEDIO

Como alternativa a la maximización de la varianza, los vectores principales pueden estimarse a partir de la minimización del error cuadrático medio:

$$J^{PCA}(u) = E\{\|e\|^2\} = E\{\|\hat{x} - \hat{y}\|^2\} = E\left\{\left\|\hat{x} - \sum_{j=1}^{N}(u_j^T \hat{x})u_j\right\|^2\right\}$$

debido a la ortonormalidad de los vectores u_j, la función anterior puede escribirse como:

$$J^{PCA}(u) = E\{\|\hat{x}\|^2\} - E\left\{\sum_{j=1}^{N}\left(u_j^T \hat{x}\right)^2\right\} =$$

$$= trace(C_x) - \sum_{j=1}^{N} u_j^T \, C_x u_j \quad (3.5)$$

En [43] se demuestra que un mínimo de (3.5), bajo la condición de ortonomalidad de los vectores u_j, viene dado la base ortonormal compuesta por los N primeros vectores propios de (3.3). Sin embargo, el criterio de minimización de (3.5) no especifica que la anterior solución sea la única y de hecho cualquier base ortonormal del subespacio de los vectores propios de (3.3) proporciona la misma propiedad de máxima compresión en relación al error cuadrático medio.

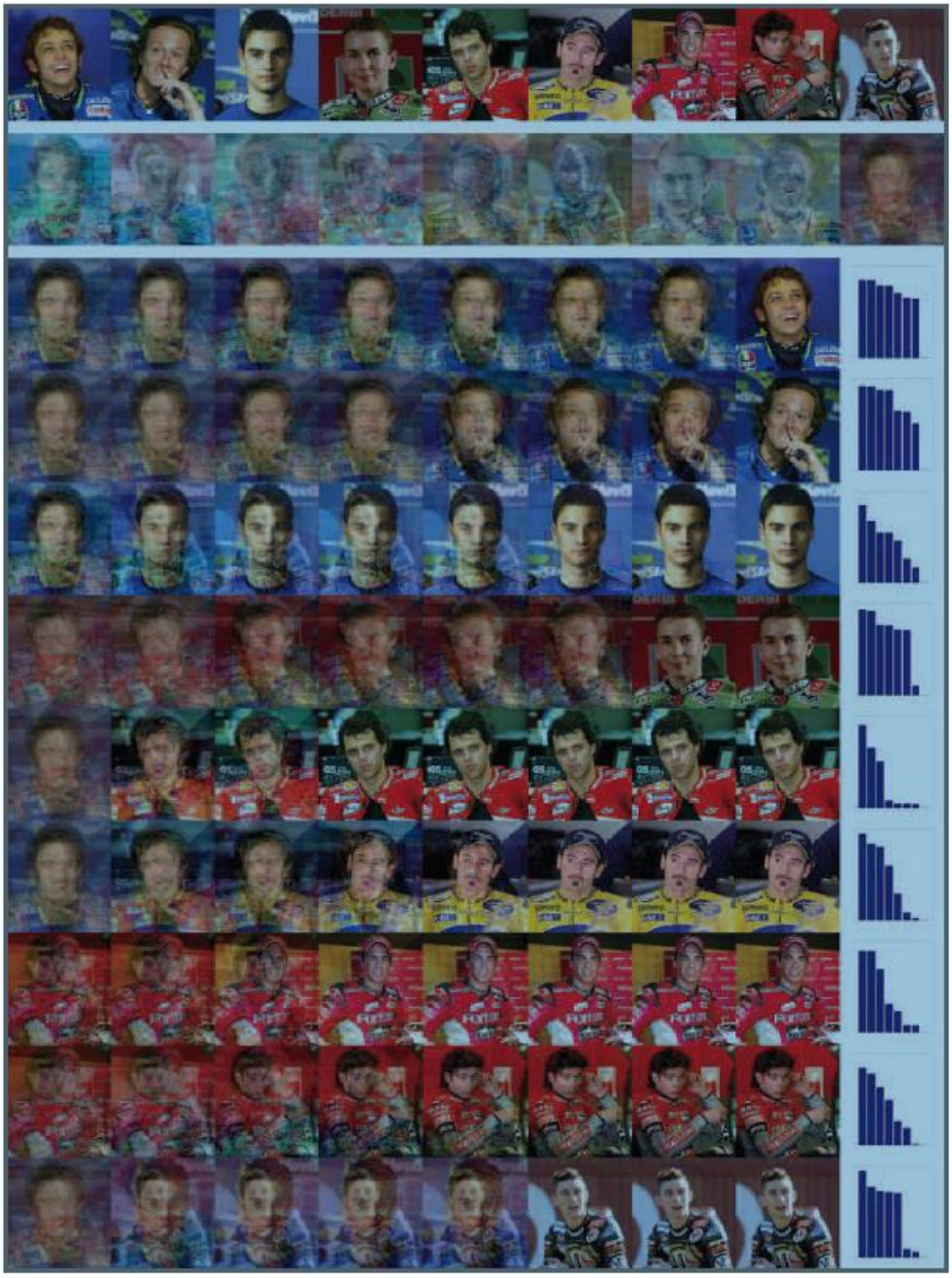


FIGURA 3.2: Esta figura ilustra el proceso de PCA. Arriba, la primera fila recoge las imágenes en color de 9 conocidos pilotos de motociclismo, y son las imágenes de entrada utilizadas para generar el eigenspacio. En la segunda fila, se muestran los ocho vectores propios obtenidos y la imagen media. En el resto de filas se muestran las reconstrucciones junto con su error de reconstrucción.

3.3. CONSIDERACIONES SOBRE PCA

PCA es una técnica de extracción de características poderosa y sencilla, pero con una serie de limitaciones debidas a su propia naturaleza. La transformación PCA no puede distinguir conjuntos de puntos cuyas estadísticas de segundo orden sean iguales. Esta limitación se ilustra en la figura 3.3(a), donde se muestran cuatro conjuntos de datos diferentes cuyas componentes principales son iguales.

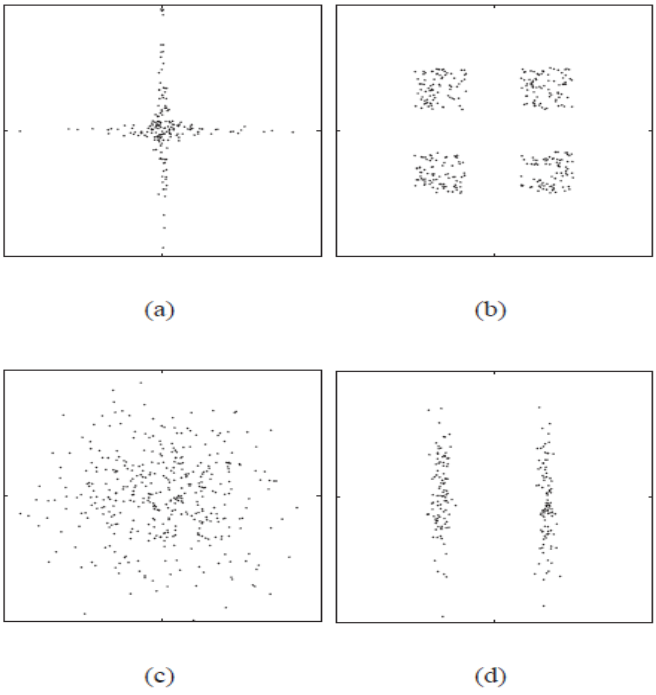


FIGURA 3.3: Ejemplos de conjuntos de datos diferentes cuyas componentes principales son iguales.

Además, puesto que se trata de una técnica no supervisada, esto es, la transformación es independiente de las clases existentes en los datos originales, los ejes principales no indican a priori las direcciones que mejor separan las clases. Por ejemplo, si consideramos el conjunto de datos de la figura 3.3(d) suponiendo que cada nube de puntos es una clase diferente, y pretendemos clasificar los puntos reduciendo previamente los datos a una sola dimensión, la proyección de los datos sobre la primera componente principal (el eje vertical) produce la confusión de las dos clases, mientras

que la proyección sobre la segunda componente principal (el eje horizontal) permite la separación completa entre las dos clases.

3.4. BLANQUEO DE DATOS, WHITENING

Siendo x un vector de datos con media cero. Siendo la matriz de covarianza,

$$\sum = E(xx^T)$$

Si los puntos de x están correlados, su covarianza, Σ, no será una matriz diagonal.

Con el fin de decorrelar los datos, se necesita realizar una transformación para que los datos transformados tengan una matriz de covarianza diagonal. Esta transformación se encuentra mediante la resolución del problema de los valores propios. Resolviendo la siguiente ecuación, se obtienen, los eigenvectores y sus eigenvalores asociados:

$$\sum \Phi = \Phi\Lambda$$

Λ es la matriz diagonal donde los elementos diagonales son los eigenvalores.
La matriz Φ diagonaliza la matriz de covarianza de x. Las columnas de Φ son los eigenvectores de la matriz de covarianzas.
También se puede expresar la covarianza diagonalizada como:

$$\Phi^T \sum \Phi = \Lambda$$

Si se quiere aplicar dicha transformada para la diagonalización a un solo vector de datos se puede hacer mediante:

$$y = \Phi^T x$$

Así, los datos de y son decorrelados, su covarianza, $E[yy^T]$ ahora es una matriz diagonal, Λ.

Los eigenvalores en Λ pueden ser iguales o distintos. Si se toman todos como iguales, esto se llama datos blanqueados, *whitening the data*. Debido que cada eigenvalor determina la longitud de los eigenvalores asociados, la covarianza corresponderá a una elipse donde los datos no estarán blanqueados, y a una esfera (tomando todas las dimensiones con la misma longitud, o uniforme) cuando los datos estarán blanqueados. Blanquear:

$$\Lambda^{-1/2}\Lambda\Lambda^{-1/2} = I$$

Igualmente, substituyendo en (1):

$$\Lambda^{-1/2}\Phi^T \sum \Phi\, \Lambda^{-1/2} = I$$

Entonces, para aplicar la transformada de blanqueado a y solo es necesario multiplicarlo por este factor de escalado, obteniendo los datos blanqueados w:

$$w = \Lambda^{-\frac{1}{2}} y = \Lambda^{-\frac{1}{2}}\Phi^T x$$

Ahora la covarianza de w no solamente es diagonal, sino también uniforme, puesto que la covarianza de w, $E[ww^T] = I$:

$$\mathrm{E}(\Lambda^{-1/2}\Phi^T x x^T \Phi\Lambda^{-1/2}) = I$$

4. BASES DE DATOS UTILIZADAS

4.1. BASE DE DATOS AR

AR face database fue creada por Aleix Martínez y Robert Benavente en el centro de visión por computador (CVC) de la Universidad Autónoma de Barcelona en junio de 1998. Contiene más de 4000 imágenes que corresponden a 126 sujetos, de los cuales 70 son hombres y 56 mujeres.

Se tomaron en el CVC bajo un entorno controlado. Las imágenes se tomaron frontalmente a los sujetos, con diferentes expresiones faciales y tres condiciones de iluminación distintas, así como varios complementos: bufandas, gafas o gafas de sol. Se realizaron dos sesiones de obtención de imágenes con los mismos sujetos, separados 14 días la una de la otra.

Se trata de imágenes con una resolución de 576 x 768 píxeles y profundidad de 24 bits, bajo el formato RGB RAW. Son almacenadas como:

Sujetos masculinos: M-xx-yy.raw
Sujetos femeninos: F-xx-yy.raw

'xx' identifica el número del sujeto, comprende desde el "00" a "70" en sujetos masculinos y de "00" a "56" en los femeninos
'yy' hace referencia a las características de la imagen, pudiendo ser:

01: Expresión Neutra
02: Sonrisa
03: Ira
04: Grito
05: Iluminación izquierda
06: Iluminación derecha
07: Todos los lados iluminados
08: Con gafas de sol

09: Con gafas de sol e iluminación izquierda

10: Con gafas de sol e iluminación derecha

11: Con bufanda

12: Con Bufanda e iluminación izquierda

13: Con Bufanda e iluminación derecha

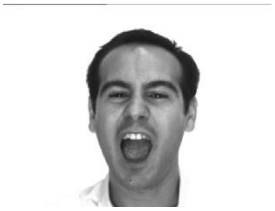
14-26: Mismas condiciones que de la 01 a 13 pero en la segunda sesión, 14 días más tarde.

En el estudio se utilizarán un total de 403 imágenes de esta AR face Database, 13 por sujeto (no se utilizan las imágenes tomadas en la segunda sesión, son siempre de la primera) de un total de 31, 19 hombres y 12 mujeres.

(01)

(02)

(03)



(04)

(05)

(06)

(07)

(08)

(09)

(10)

(11)

(12)

(13)

FIGURA 4.1: Imágenes de la AR Face Database.

4.2. BASE DE DATOS AT&T

Conocida formalmente como “The ORL Database of Faces”, se realizó entre abril de 1992 y abril de 1994 en el laboratorio de AT&T en Cambridge.

Se compone de un total de 10 imágenes por sujeto, de un total de 40. Para algunos sujetos las imágenes se tomaron en tiempos diferentes variando la iluminación, las expresiones faciales: ojos abiertos/cerrados, sonriendo/ sin sonrisa. Así como también con o sin gafas.

Las imágenes fueron tomadas bajo un fondo homogéneo negro, en posición vertical y de forma frontal al sujeto, con alguna pequeña rotación.

Se trata de imágenes con una resolución de 92 x 112 pixels en escala de grises, 255 niveles de gris por pixel.

Las imágenes se presentan en formato PGM y están organizadas en 40 directorios, 1 por sujeto, con el formato sX (X es el número del sujeto, que va del 1 al 40). Dentro de cada directorio el formato de las imágenes es Y.pgm, donde “Y” se corresponde al número de imagen del sujeto, que comprende desde la 1 a la 10.

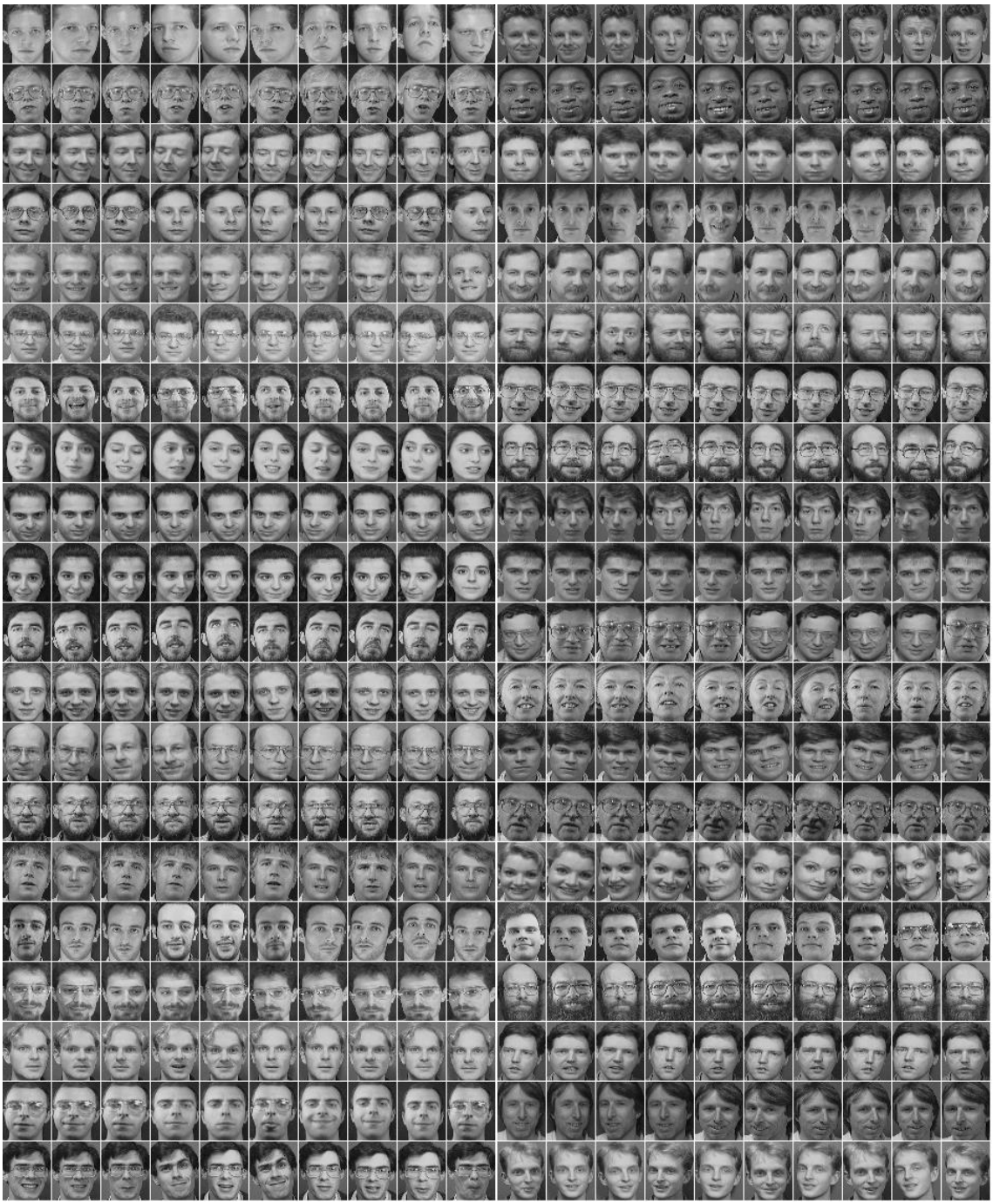


FIGURA 4.2: Imágenes de The ORL Database of Faces

4.3. BASE DE DATOS LFW

LFW database (Labeled Faces in the Wild) fue creada en octubre de 2007 en la Universidad de Massachussets. Sus autores son: Gary B. Huang, Manu Ramesh, Tamara Berg y Erik Learned-Miller.

La BD se compone de un total de 13233 imágenes de 5749 sujetos, de los cuales existen 1680 sujetos con dos o más imágenes y el resto de los 4069 solo hay una imagen.

Las imágenes han sido tomadas de Internet y preprocesadas y centradas por el detector Viola-Jones.

Se trata de imágenes con una resolución de 250x250 píxeles, la mayor parte son en color, aunque también existen algunas en escala de grises.

Las imágenes se presentan en formato JPG y están organizadas por carpetas, donde una carpeta corresponde a un sujeto bajo el formato X_Y_YN1_YN2…, donde X se corresponde con el nombre del sujeto, Y con el apellido e YN1, YN2 etc, con los segundos apellidos o preposiciones, por ejemplo:

Sergio_Vieira_De_Mello
X: Sergio
Y: Vieira
YN1: De
YN2: Mello

Dentro de cada carpeta de sujeto tenemos las imágenes del sujeto que se comprenden desde la mínima por sujeto, una, hasta la máxima, correspondiente al sujeto George_W_Bush con 530 imágenes.

El formato de las imágenes es el siguiente:

x_y_yn[0…N]_zzzz

x: Nombre

y: Primer apellido

yn: apellidos o preposiciones pudiendo tener más de un yn como en ejemplo anterior

zzzz: desde 0001 hasta un máximo de 0530, dependiendo del número de imágenes del sujeto.

Para el estudio se ha realizado una selección de los sujetos que tienen igual o más de 10 imágenes y se ha seleccionado solo las 10 primeras de estos. También se ha pasado a escala de grises (255 niveles de gris por pixel) y se ha recortado a 150x150 píxeles.

George_W_Bush_0001

George_W_Bush_0002

George_W_Bush_0003

George_W_Bush_0004

George_W_Bush_0005

George_W_Bush_0006

George_W_Bush_0007

George_W_Bush_0008

George_W_Bush_0009

George_W_Bush_010

FIGURA 4.3: Imágenes de LFW FACE DATABASE

4.4. BASE DE DATOS FERET

FERET (Face Recognition Technology) Database fue creada en 15 sesiones en un entorno semicontrolado entre agosto de 1993 y julio de 1996. El autor principal es el Dr. Harry Wechsler de la Universidad George Mason, en colaboración con el Dr. P. Jonathon Phillips.

Dicha base de datos contiene un total de 1564 conjuntos de imágenes, con un total de 14126 imágenes, de las cuales 1199 son individuales y 365 son duplicadas. Los conjuntos duplicados pertenecen a sujetos que ya están presentes en el conjunto de imágenes individuales y que fueron tomadas, generalmente, con algún día de diferencia.

Algunas de las imágenes tomadas de un mismo sujeto, difieren en el tiempo un par de años, con lo que se puede utilizar para tratar también los cambios en el rostro que aparecen con el paso del tiempo.

Las imágenes tienen una profundidad de 24bits, RGB, por lo que se tratan de imágenes en color, con una resolución de 512x768 píxeles. La extensión de las imágenes es *.ppm* (Portable Pixel Map)

La nomenclatura de las imágenes de la base de datos es nnnnn_yymmdd_xx_q.ext, donde:

nnnnn: Corresponde al indicador único, de cinco dígitos enteros, de un solo sujeto.
yymmdd: Fecha de en la cual se tomó la fotografía.
xx: indica el tipo de fotografía tomada:

fa: Fotografía frontal con expresión regular

fb: Fotografía frontal con expresión alternativa tomada segundos después de fa.

q: es un modificador que no está presente en todas las imágenes:

a: Gafas. Cabe mencionar que no todos los sujetos con gafas tienen este marcador.

Para el estudio, solo se han tenido presente en la modalidad de tipo de fotografía, marcador *xx*, las antes mencionadas, ya que la base de datos tiene un gran repertorio en cuanto a expresiones faciales, así como ángulos de rotación de la cabeza respecto a la cámara. El marcador, q, que no está presente en todas las imágenes, para el estudio solo se ha utilizado el descrito, ya que hay en la base de datos existen otros marcadores que hacen referencia a distintas longitudes de cabello, reducciones del tamaño de la imagen, histograma ajustado, reducciones de brillo etc…

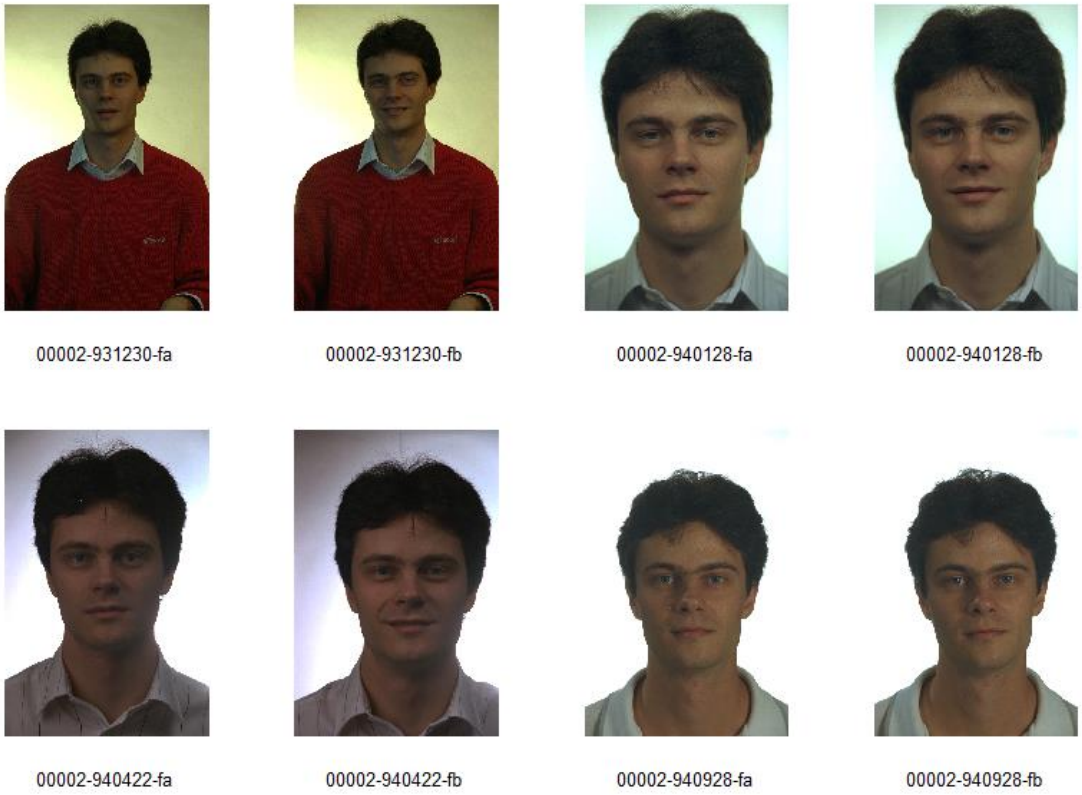


FIGURA 4.4: Ejemplo sujeto 2.

4.5. BASE DE DATOS UMH

Base de datos creada en la Universidad Miguel Hernández en el año 2006. Los autores son César Fernández Peris y María Asunción Vicente Ripoll.

Contiene un total de 510 imágenes, correspondientes a 17 sujetos. De cada sujeto existen 30 imágenes.

Las imágenes fueron tomadas bajo condiciones de luz no controladas. La distancia de un mismo sujeto a la cámara es diferente, así como también la iluminación, expresiones faciales, objetos en las mismas, etc…

Esta DB fue creada con distintas cámaras, por lo que las imágenes tomadas para un mismo sujeto no están tomadas siempre con la misma cámara fotográfica.

Se trata de imágenes en color con una resolución de 121x151 pixels, bajo el formato JPG. Son almacenadas y nombradas como:

- CaraX_Y.jpg
X = Numero de Sujeto desde 1 hasta 17
Y = Número de Imagen del Sujeto X. Desde la 1 a la 30 siguiendo el siguiente orden:

De la 1 a la 10 de un mismo sujeto fueron tomadas bajo un fondo blanco homogéneo
De la 11 a la 20 de un mismo sujeto fueron tomadas bajo un fondo complejo.
De la 21 a la 25 con oclusiones parciales bajo fondo blanco homogéneo
De la 26 a la 30 con oclusiones parciales bajo fondo complejo

FIGURA 4.5: Imágenes de UMH DATABASE

FIGURA 4.6: Imágenes de UMH DATABASE

5. PREPROCESADOS

5.1. ECUALIZACIÓN DEL HISTOGRAMA

El Histograma, de una imagen en escala de grises, es la representación gráfica de la frecuencia con las que aparecen los distintos niveles de gris en una imagen, es decir, la cantidad de pixeles de una imagen con un determinado nivel de gris. El eje de abscisas representa el nivel de gris entre 0 y 255, el de ordenadas la cantidad de pixels que hay en determinado nivel de gris. Ejemplo:

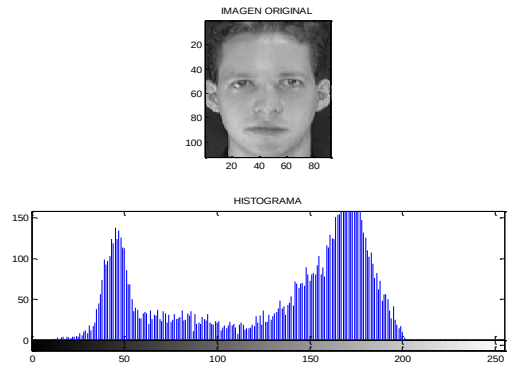


FIGURA 5.1: Cara de BD ORL sujeto 1 imagen 1.

El histograma, a su vez, nos informa del brillo y contraste de una imagen. Una imagen será oscura cuando tenga los componentes centrados en la parte baja del histograma, donde se encuentran los niveles de gris más oscuros. Cuando suceda lo contrario, cuando se observe en el histograma que hay más componentes centrados en la parte alta del histograma, donde se encuentran los grises más claros, se indica que la imagen es más brillante. Respecto al contraste, se indica que una imagen es de bajo contraste cuando en el histograma los componentes están agrupados y centrados en el centro de los niveles de gris. Finalmente, una imagen de alto contraste es aquella que cubre casi todos los niveles de grises y cuyo histograma se presenta casi homogéneo en cuanto a número de pixeles por los niveles de gris.

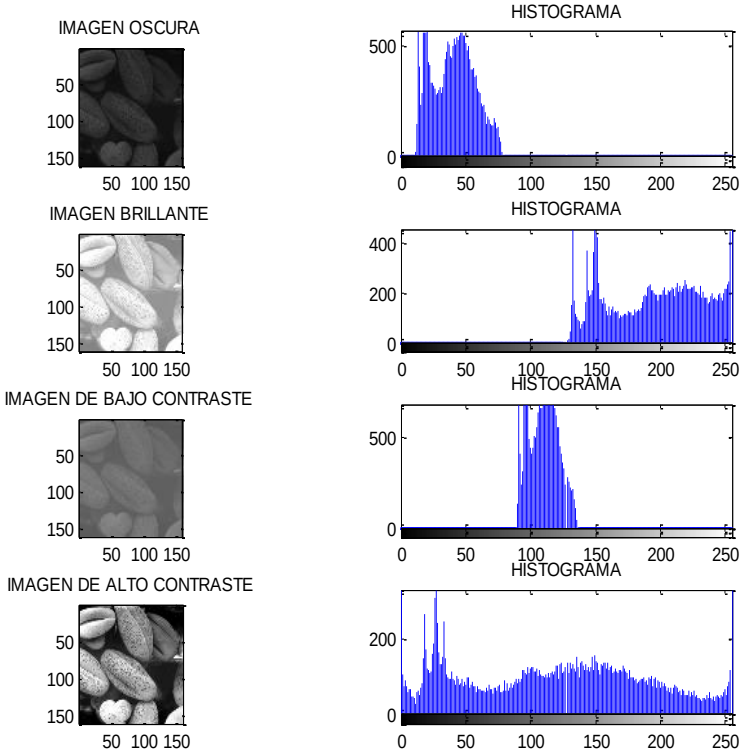


FIGURA 5.2: Ejemplos de la misma imagen con distintos tipos de histogramas

Se define el brillo de una imagen como la media de la imagen:

$$Brillo = \bar{g} = \frac{1}{M * N} \sum_{i=1}^{M} \sum_{j=1}^{N} f(i,j)$$

$$N = Numero\ de\ Filas \quad M = Numero\ de\ Columnas$$

$$f(i,j) = Valor\ de\ gris\ del\ pixel\ en\ las\ coordenadas(i,j)$$

Una imagen brillante tiene una media alta, al contrario, la imagen oscura tiene una media baja.

Queda definido el contraste como la varianza de una imagen:

$$Contraste = \sigma^2 = \frac{1}{M * N} \sum_{i=1}^{M} \sum_{j=1}^{N} (f(i,j) - \bar{g})^2$$

Un valor de varianza alto indica un contraste alto, y viceversa.

La ecualización del histograma, trata de transformar a un histograma con una distribución homogénea, convirtiendo todos los valores de gris, de 0 a 255 niveles de gris, con el mismo número de ocurrencias. Con esta transformación se consigue mejorar el contraste y se distinguen mejor los objetos de la misma pero también sufre el inconveniente de aumentar el ruido y la segmentación en la imagen tratada.

La ecualización del histograma está ligada a la función de distribución acumulada, FDA o CFD, debido a que queda uniforme tras la ecualización.

Partiendo de la función de distribución:

$$y' = H(x) = \sum_{i=0}^{x} p(i) \quad , donde:$$

$$p(i) = \frac{n_i}{\sum_{i=0}^{N-1} n_i} = \frac{n_i}{N} \quad y \quad \sum_{i=0}^{N-1} p\,(i) = 1$$

Quedando:

$$y' = \sum_{i=0}^{x} \frac{n_i}{N} \qquad x = 1,2,3 \ldots . N-1 \; (N \; Numero \; total \; de \; grises)$$

La función y' queda definida entre $0 < y' < 1$, debiendo ser transformada a los niveles de grises, $0 < y < N - 1$, mediante la transformada:

$$y = y'(N-1) = (N-1) \sum_{i=0}^{x} \frac{n_i}{N}$$

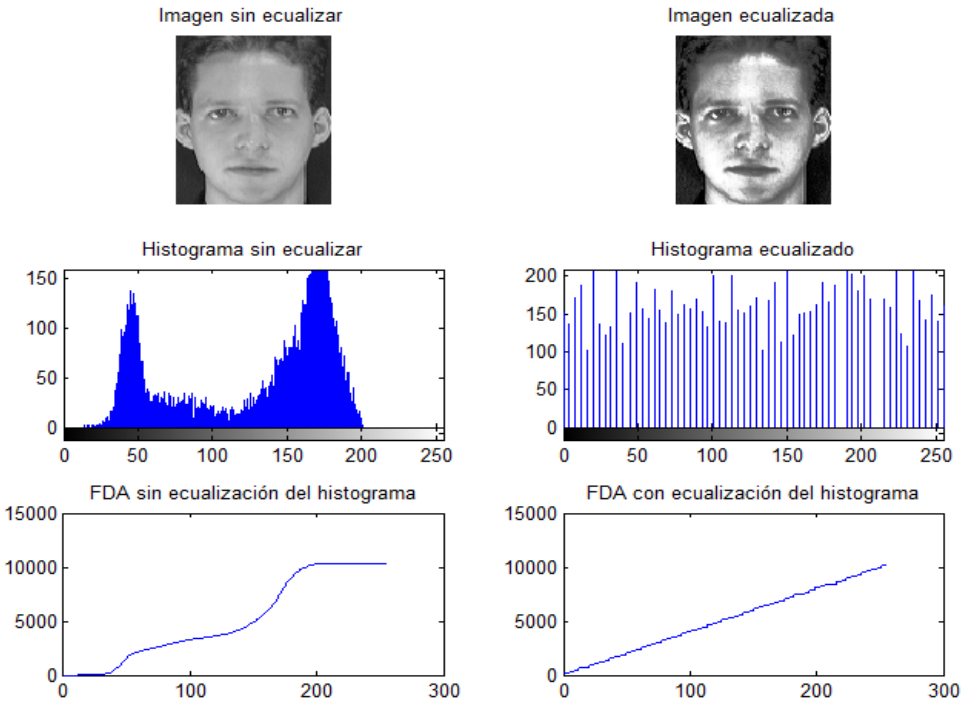


FIGURA 5.3: Ejemplo de ecualizado de histograma

5.2. MÁSCARA ELIPSOIDAL

La máscara Elipsoidal es un preprocesado cuya finalidad es, dada una imagen de tamaño *xy* donde existe algún rostro en ella, devuelve una imagen donde se recorta una elipse alrededor de esta, devolviendo los pixels internos a esta como un vector y eliminando el resto de información.

Este preprocesado es muy útil para tratar únicamente los objetos que interesan, las caras en el caso que se está estudiando, despreciando objetos que puedan provocar ruido en el estudio.

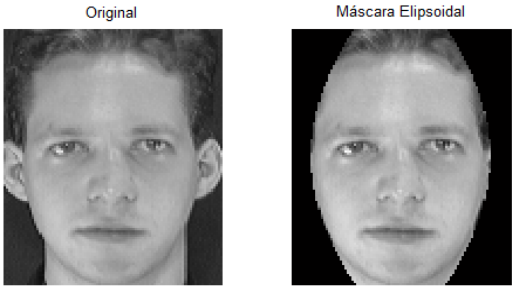


FIGURA 5.4: Ejemplo de imagen tras aplicar el preprocesado de la Máscara Elipsoidal.

5.3. PREPROCESADO TAN-TRIGGS

El preprocesado Tan-Triggs recibe el nombre de los autores, Xiaoyang Tan and Bill Triggs, que trata de estudiar los efectos de la variación de la iluminación, sombras locales, así como también del brillo. Todo ello manteniendo los elementos esenciales de la apariencia visual.

Este preprocesado se basa en la aplicación de tres técnicas: en la corrección gamma, un filtrado DoG (Filtro Gaussiano) y una ecualización del histograma:

1-Corrección Gamma: Se trata de una transformación no lineal de los niveles de gris que remplaza el nivel de gris I por I^{γ}, para γ>0, o log(I) si γ=0, donde γ ∈ [0,1]:

$$I_{salida} = A * I_{entrada}^{\gamma} \quad Ec.x$$

$$A = Constante, tipicament\ de\ valor\ a\ la\ unidad$$

El nivel de gris I debe estar normalizado entre [0,1]

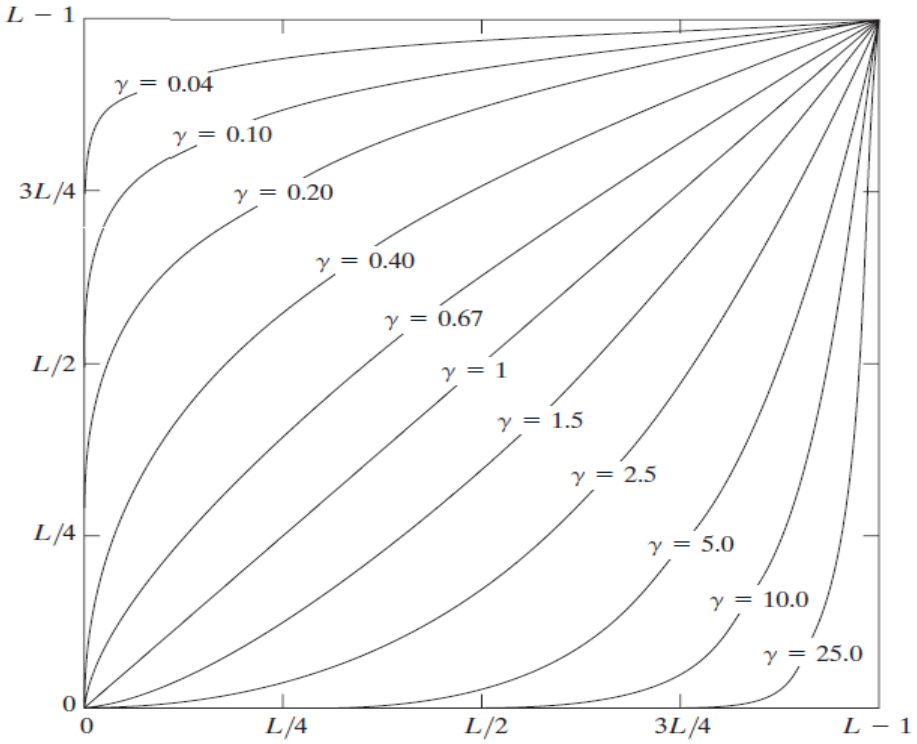


Figura 5.5: Representación gráfica de la ecuación $I_s = A * I_e^{\gamma}$ (A= 1 para todos los casos) para distintos niveles de gris L.

Con esta transformación se mejora el rango dinámico local de la imagen en zonas oscuras o con sombras, mientras que lo comprime en zonas brillantes y destacadas. Se rige por el principio de que la intensidad de la luz reflejada por el objeto es el producto de la entrada de la iluminación L y de la reflectancia de la superficie R.

Con esta transformada se pretende recuperar la información a nivel de objeto independientemente de la iluminación. Esta transformación de los niveles de gris solo modifica el rango dinámico de los mismos, no incluye o quita información de la imagen.

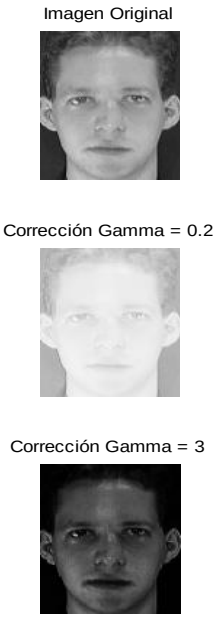


FIGURA 5.6: Ejemplo con distintos valores de Gamma.

2-Filtrado Gaussiano

El filtrado Gaussiano recibe este nombre debido a que los pesos del filtro son calculados siguiendo a una distribución Gaussiana:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad Ec.x$$

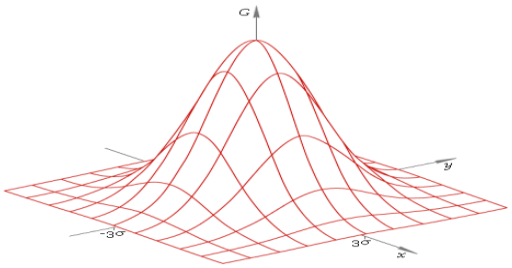


FIGURA 5.7

Este tipo de filtrado se comporta como un filtro paso bajo, atenuando las componentes de alta frecuencia. Tiene simetría rotacional. No tiene lóbulos laterales, ya que como se puede ver en la figura 5.7, tiene un único lóbulo, donde el peso de los pixels decrece respecto de la distancia al centro, así que cuanto más alejado esté un pixel, menos significativo es. Tras aplicar el filtrado se puede observar como la imagen queda suavizada a medida que se aumenta la varianza, σ^2, perdiendo el detalle de la misma pero eliminando el ruido de tipo Gaussiano en la imagen. Cuanto mayor es la varianza más se tienen en cuenta los puntos más alejados de la media, es lo mismo que decir que a medida que se aumenta la varianza se van teniendo en cuenta vecinos cada vez más alejados.

Este tipo de filtrado se utiliza para eliminar ruido de tipo Gaussiano, para el ruido de tipo impulsional, 'sal y pimienta', no es efectivo este tipo de filtrado. El ruido de tipo Gaussiano es aquel ruido que se produce, normalmente, por los componentes electrónicos que obtienen la imagen.

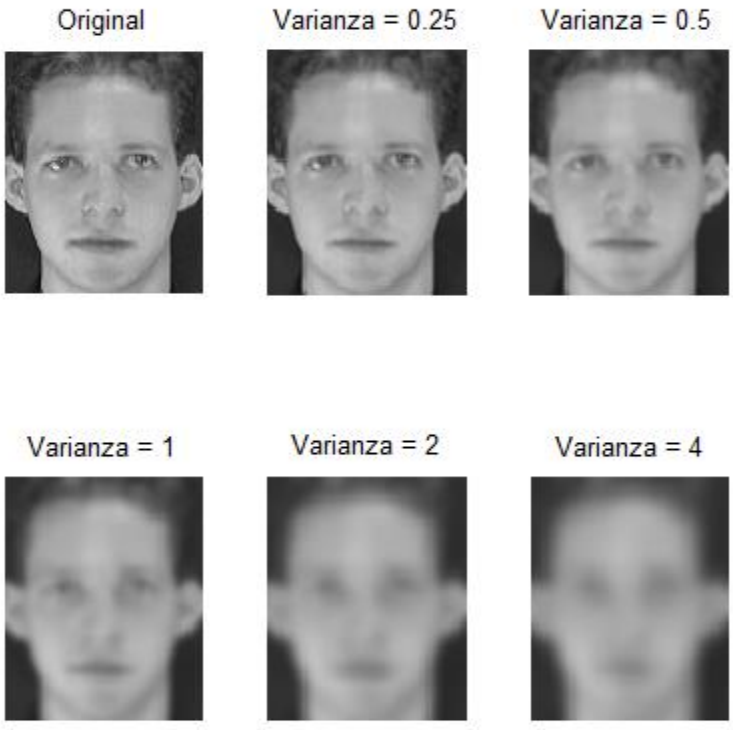


FIGURA 5.8: Sujeto de base de datos ORL con coeficientes de varianza distintos

3-Ecualizado del contraste: Se aplica para eliminar valores extremos como los reflejos, pequeñas regiones oscuras como las fosas nasales, los efectos de los bordes etc…

En vez de utilizar la media del valor absoluto de la imagen para normalizar el brillo, se utiliza una aproximación de esta basada en dos fases:

$$I(x,y) \leftarrow \frac{I(x,y)}{(media(|I(x',y')|^{\alpha}))^{1/\alpha}} \qquad Ec.\,5.3.1$$

$$I(x,y) \leftarrow \frac{I(x,y)}{(media(\min(\tau,|I(x',y')|)^{\alpha}))^{1/\alpha}} \qquad Ec.\,5.3.2$$

En las ecuaciones, α, es un exponente de compresión que reduce los valores extremos, y τ es un umbral utilizado para truncar los valores extremos tras la primera fase de la normalización. La media se realiza sobre la imagen sin enmascarar.

La imagen tras la normalización puede continuar albergando valores extremos en la misma, para ello se realiza un mapeado no lineal para reducir los efectos de los valores extremos en las imágenes:

$$I(x,y) \leftarrow \tau * \tanh\left(\frac{I(x,y)}{\tau}\right), I \in (-\tau,\tau)$$

Los valores utilizados en el preprocesado Tan-Triggs son:

- $Corrección\ Gamma$: $\gamma = 2\ en\ Ec.x$
- Filtrado Gaussiano: $\sigma_0 = 1.0\ y\ \sigma_1 = 2.0\ en\ Ec.x$
- $Ecualización\ del\ Contraste$: $\alpha = 0.1\ y\ \sigma = 10\ en\ Ec.x$

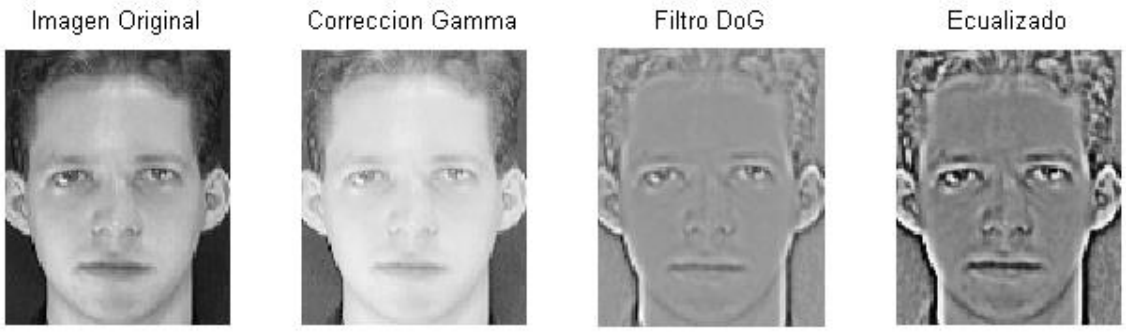


FIGURA 5.9.

5.4. VIOLA-JONES

El algoritmo de Viola-Jones busca detectar con fiabilidad cualquier tipo de objeto de un modo rápido y lo más fiable posible. Normalmente se aplica para detectar caras en orientación aproximadamente frontal a la cámara, pero puede ser entrenado para localizar caras en otra orientación o cualquier objeto del que se disponga de imágenes de entrenamiento. El algoritmo proporciona resultados fiables con independencia de la escala (la figura 5.10 muestra un ejemplo de funcionamiento del algoritmo donde hay grandes variaciones de tamaño entre unas caras y otras).

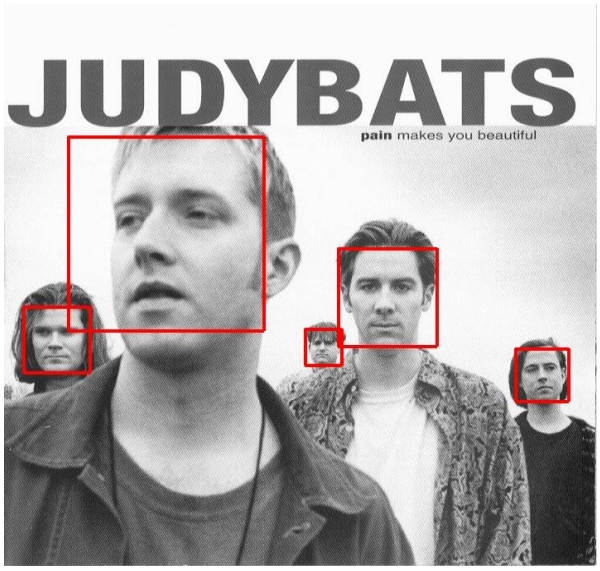


FIGURA 5.10: Ejemplo de funcionamiento del algoritmo de Viola-Jones.

El principio de funcionamiento se basa en la utilización de múltiples características o atributos (160.000 por cada subventana de la imagen a procesar) extremadamente simples, pertenecientes a 4 tipos básicos que se muestran en la figura 5.11, indicados como A, B, C y D. Estas características se repiten en múltiples posiciones y escalas, siempre sobre subventanas de 24x24 píxeles extraídas de las imágenes.

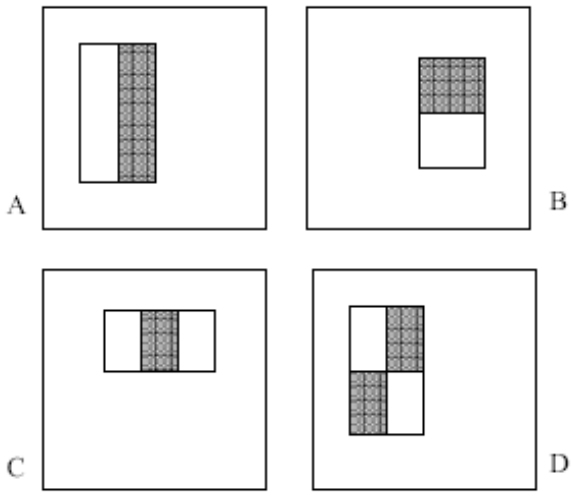


FIGURA 5.11: Características básicas utilizadas por el detector de Viola-Jones.

En todos los casos, el valor del atributo se obtiene como la suma del valor de los píxeles del área blanca menos la suma del valor de los píxeles del área negra. La imagen de la figura 5.12 muestra un ejemplo de zonas de la imagen que obtendrían un valor alto para dos de estos atributos.

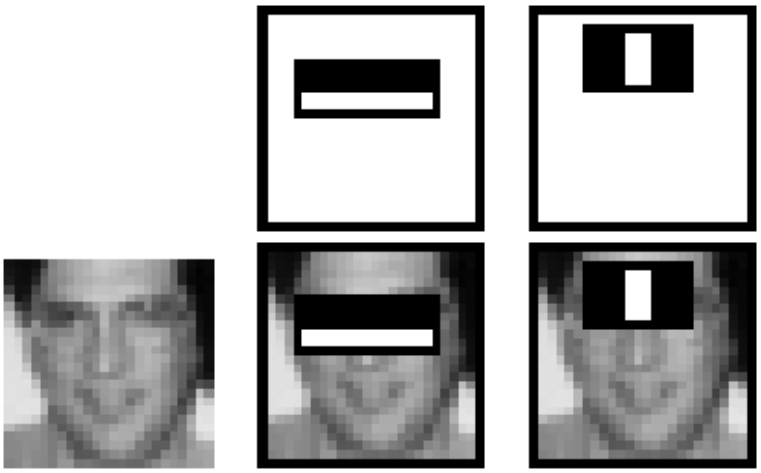


FIGURA 5.12: Dos de los atributos evaluados sobre una imagen.

Para que el cálculo de estos atributos sea rápido, se utiliza la llamada imagen integral, en la cual el valor de cada píxel se obtiene como la suma de los valores de los píxeles de

la imagen original contenidos en el rectángulo comprendido entre el origen de la imagen y el píxel a evaluar. La figura 5.13 muestra un ejemplo de cálculo de esta imagen integral a la izquierda; y un ejemplo de obtención de la suma de píxeles de un rectángulo cualquiera de la imagen a la derecha (la suma de los píxeles contenidos en el recuadro D sería igual a la siguiente operación sobre los valores de la imagen integral: 4+1-2-3.

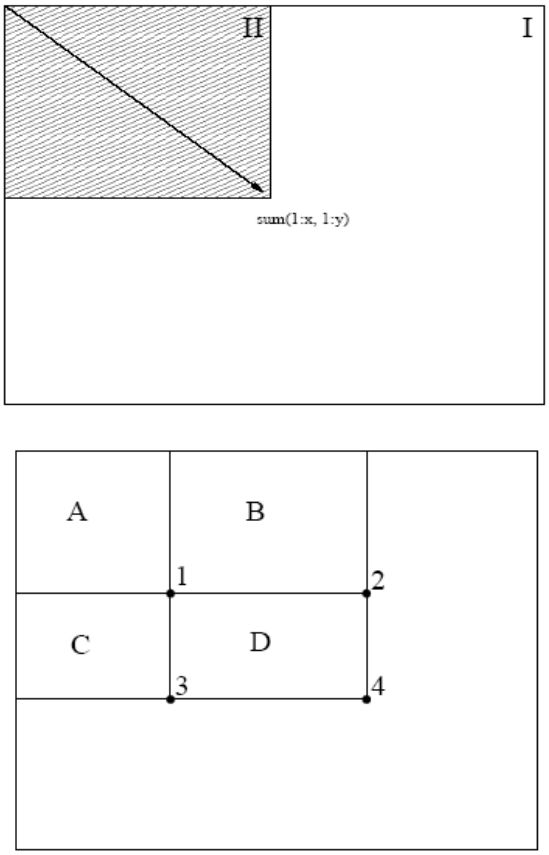


FIGURA 5.13: Cálculo y utilización de la imagen integral

Se utiliza Adaboost para poder trabajar con un número tan elevado de atributos sin incurrir en un sobreajuste del clasificador. Se ajusta un clasificador para cada característica, y los clasificadores utilizados también son extremadamente sencillos (perceptrón simple con una sola entrada).

El proceso de entrenamiento se basa en ajustar los clasificadores para un gran conjunto de imágenes de entrenamiento positivas (que contienen caras en la orientación buscada) y negativas (que no la contienen). En un procesador estándar, este proceso puede durar

varias semanas. Sin embargo, el resultado una vez que el sistema ha sido entrenado es extremadamente rápido: se puede analizar una imagen de 384x288 píxeles en 0.067 segundos y detectar todas las caras presentes en ella. Para conseguir esta alta velocidad se utiliza una cascada de clasificadores, en orden creciente de exigencia, de modo que cada uno de ellos descarta los falsos positivos (falsas detecciones de caras) propuestos por el anterior.

6. METODOLOGÍA

Para el estudio se ha utilizado el software MATLAB®, cuyo nombre viene de la abreviatura de MATrix LABoratory, laboratorio de matrices. Dicho software fue creado por Cleve Moler y Ramsés García Sánchez en 1984, partiendo con la premisa de emplear paquetes de subrutinas escritas con Fortran para el álgebra lineal y análisis numérico, sin tener la necesidad de escribir programas en lenguaje Fortran.

El origen de MATLAB surge debido a que durante muchos años el único lenguaje de alto nivel que existía para las aplicaciones científicas y de ingeniería era el lenguaje FORTRAN, el cual estaba especialmente diseñado para la computación numérica. Debido a la ineficiencia de repetir los mismos algoritmos una y otra vez el gobierno de los EEUU encargó a analistas numéricos formular algoritmos para las tareas comunes. De dicho encargo nacieron las librerías de las subrutinas, que se pueden encontrar para su uso gratuito en www.netlib.org, como son:

- BLAS: operaciones con vectores, como añadir vectores, hacer multiplicaciones, normalizar.
- LINPACK: Subrutinas de algebra lineal para operaciones con vectores y/o matrices, soluciones de sistemas lineales, inversiones de matrices, potencias de matrices. Esta fue más tarde remplazada por LAPACK.
- EISPACK: Calculo de eigenvalores y eigenvectores de matrices

Cleve Moler uno de los científicos que participaron en el desarrollo de estas librerías, pensó que el proceso todavía continuaba requiriendo de mucho trabajo. Se tenía que escribir FORTRAN, compilarlo, depurarlo, compilarlo… por lo que optó por escribir MATLAB para que sus estudiantes, debido a que era profesor en la universidad de Carolina del Sur, tuviesen un acceso fácil a LINPACK.

Durante una conferencia en la universidad de Santford, Jack Little, vio la ponencia del profesor y con su permiso, rescribió MATLAB EN C, añadió las M-Files (programas ya guardados), y otras características y librerías para más tarde fundar THE MATHWORKS y venderlo.

MATLAB es un sistema interactivo cuyos elementos básicos son matrices que no requieren de un dimensionamiento previo. Gracias a ellos puede resolver muchos

problemas de la técnica computacional, especialmente en aquellos con formulaciones de matrices y vectores, en cuestión de una fracción de tiempo, que requeriría de más tiempo si se tuviese que resolver en un lenguaje no interactivo como FORTRAN o C.

Una de las soluciones más atractivas de MATLAB son las denominadas TOOLBOXES, que son un conjunto de colecciones completas de las funciones de MATLAB(M-files) que aportan soluciones a ciertas áreas específicas, como puede ser para redes neuronales, procesamiento de señales, sistemas de control, identificación, etc.

Para el estudio se ha utilizado el toolbox, *Image Processing Toolbox*, que contiene funciones que han sido utilizadas como:

- *Imshow* Muestra una imagen
- *Imread* Realiza una lectura de un archivo de imagen
- *Imresize* Cambia el tamaño de una imagen
- *Rgb2gray* Convierte una imagen de color RGB o mapeado de color a escala de grises
- *Histeq* Mejora el contraste utilizando ecualización del histograma
- *Imfilter* Realizar un filtro de la media

El trabajo se ha dividido en distintos archivos, "*.m"*, para dividir el tipo de tarea que se ejecuta en cada uno de ellos:

save_"BBDD".m: La finalidad de este programa es salvar las imágenes en un vector que son almacenadas en un fichero *.mat* donde se han aplicado los distintos preprocesados

read_"BBDD".m: Programa llamado por el descrito anteriormente para leer las distintas bases de datos, las cuales no se han tratado de la misma forma debido a la naturaleza de las mismas, ya que no todas las imágenes que contienen estas BBDD tienen el mismo formato y tampoco el toolbox utilizado tiene soporte para alguna de ellas.

Prepro.m: Dicho programa es llamado por *read_"BBDD".m* donde se ejecutan los distintos preprocesados descritos anteriormente.

Al finalizar el programa *save_"BBDD".m* se guardan los ficheros dependiendo de los preprocesados utilizados. Los cuales han sido guardados con el siguiente formato:

Nombre archivos	**Descripción Preprocesado**
BBDD_000.mat	Imágenes Originales
BBDD_001.mat	Máscara Elipsoidal
BBDD_010.mat	Histograma Ecualizado
BBDD_011.mat	Histograma Ecualizado + Máscara Elipsoidal
BBDD_020.mat	Ecualizado Tan-Triggs
BBDD_021.mat	Ecualizado Tan-Triggs + Máscara Elipsoidal
BBDD_100.mat	Detector Viola-Jones
BBDD_101.mat	Detector Viola-Jones + Máscara Elipsoidal
BBDD_110.mat	Detector Viola-Jones + Histograma Ecualizado
BBDD_111.mat	Detector Viola-Jones + Histograma Ecualizado + Máscara Elipsoidal
BBDD_120.mat	Detector Viola-Jones + Ecualizado Tan-Triggs
BBDD_121.mat	Detector Viola-Jones + Ecualizado Tan-Triggs + Máscara Elipsoidal

TABLA 6.1

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Lectura y preprocesado de las imágenes de la librería ORL
% Las imágenes se almacenan en forma de vector en un fichero *.MAT
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% ORL FACE DATABASE
% http://people.cs.uchicago.edu/~dinoj/vis/ORL.zip

nc=40; %número de clases (personas)
np=10; %número de imágenes por cada clase


tx_viola=80; %tamaño del rectángulo de Viola-Jones
ty_viola=80;

see_interno=0; % see_interno ---> ver o no los plots de las caras: (NO
==> 0;  YES ==> 1)
see_externo=1; % see_externo ---> ver o no los plots de las caras: (NO
==> 0;  YES ==> 1)
% see_externo muestra todas las caras de cada sujeto con
% todos los preprocesados

cont_figura=1;  %contador para visualizar las imágenes de cada
preprocesado en figuras diferentes


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%
%parámetros del preprocesado: viola ecu elip
%
% viola --> detector de Viola-Jones
%            0: no se utiliza
%            1: si se utiliza
% ecu ----> tipo de ecualizado:
%            0: no se hace ningún ecualizado
%            1: ecualización de histograma
%            2: ecualización de Tan/Triggs
%            3: ...
% elip ---> recorte elíptico: devuelve sólo los pixeles contenidos
% en una elipse alrededor de la cara. Normalmente sólo debe ser
utilizado
% si antes se ha utilizado Viola-Jones
%            0: no se hace el recorte
%            1: sí se hace el recorte
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%



%for viola =[0,1]
for viola =[1]
    %for ecu=[0, 1, 2]
    for ecu=[0]
        %for elip=[0, 1]
        for elip=[0]

            %lectura de las imágenes y preprocesado
            [x, class_x, cont_ok]=read_orl(nc, np, tx_viola, ty_viola,
```

```
viola, ecu, elip, see_interno, see_externo, cont_figura);

            %guardar en *.mat
            nombre_fichero=sprintf('orl_%d%d%d',viola,ecu,elip);
            save(nombre_fichero, 'x', 'viola' ,'ecu', 'elip', 'nc'
,'np')

            %%imprime los valores de los parámetros seleccionados
            fprintf('Numero de imágenes leídas de la librería
ORL=%d\n',size(x,2))
            if viola==1;
                fprintf('Numero de imágenes detectadas por Viola
Jones=%d\n',round(cont_ok))
                fprintf('Detector de Viola\n')
            end
            if ecu==1
                fprintf('Ecualización del histograma\n')
            end
            if ecu==2
                fprintf('Ecualización TAN-TRIGGS\n')
            end
            if elip==1
                fprintf('Elipse\n')
            end

            cont_figura=cont_figura+1;

            pause(2)
        end
    end
end
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
function [x, class_x, cont_ok]=read_orl(total_faces, obj_p_class, tx,
ty, viola, ecu, elip, see_interno, see_externo, cont_figura)
%%% ORL_FACES
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Esta función realiza la lectura y el preprocesado de las imágenes de
ORL
%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
%
% tx --> ancho del recuadro a devolver
% ty --> alto del recuadro a devolver
% viola --> detector de Viola-Jones
%             0: no se utiliza
%             1: si se utiliza
% ecu ----> tipo de ecualizado:
%             0: no se hace ningún ecualizado
%             1: ecualización de histograma
%             2: ecualización de Tan/Triggs
%             3: ...
% elip ---> recorte elíptico: devuelve sólo los pixeles contenidos
% en una elipse alrededor de la cara. Normalmente sólo debe ser
utilizado
% si antes se ha utilizado Viola-Jones
```

```
%             0: no se hace el recorte
%             1: sí se hace el recorte
% see --> igual a 0 si no se desea mostrar resultados
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%

disp('Leyendo las imágenes de la librería ORL')

cont=1;
cont_ok=0;  %% contador de ok del Viola

for i=1:total_faces

    for j=1:obj_p_class
        comando=sprintf('y=imread(''../ORL/s%d/%d.pgm'');',i,j);
        eval(comando);
        fprintf('Leyendo..../ORL/s%d/%d.pgm \n',i,j);

        %prepro devuelve la imagen de la cara en forma de vector
(im_vector)
        %y en forma de matriz (im_matriz)
        [im_vector, im_matriz, ok] = prepro(y, viola, ecu, elip, tx,
ty, see_interno);

        %matriz donde se guardan las imágenes (cada columna es una
cara)
        x(:,cont)=im_vector;
        %clase de cada vector columna (sujeto al que pertenece la
cara)
        class_x(cont)=i;

        if ok==1
            cont_ok=cont_ok+1;  %aciertos de Viola
        end
        cont=cont+1;

        %para visualizar todas las imágenes de una misma persona
        if see_externo
            %figure(cont_figura);
            figure(i);
            subplot(1,obj_p_class,j); %en una sola fila todas las
caras de una misma persona
            titulo = sprintf('%d, %d', i,j);
            muestra(im_matriz,titulo);
        end

    end

end
```

```
% realiza el preprocesado completo de una imagen
% ------------------------------------------------
% permite elegir las siguientes opciones:
% - Viola-Jones: puede lanzarse opcionalmente para recortar la parte
% de la imagen correspondiente a la cara (suponemos imágenes con
% una única cara). La imagen se recorta y se escala a un tamaño fijo
% - Preprocesado: puede realizarse un ecualizado de histograma, un
% ecualizado de Tan-Triggs o ningún ecualizado
% - Elipse: puede definirse una zona elíptica alrededor de la cara
% y devolver únicamente los píxeles contenidos en esa zona
%
% En todos los casos, la imagen se devuelve como un vector, para
% utilizar posteriormente PCA, LDA o ICA.

% probar como:
% imagen = rgb2gray(imread('imagen_1.jpg')); % (o imagen_2.jpg);
% [im_vector, ok] = prepro(imagen, 1, 2, 1, 100, 100, 1);

function [im_vector, im_matriz, ok] = prepro(imagen, viola, ecu, elip,
tx, ty, see)
% imagen --> imagen a preprocesar
% viola --> detector de Viola-Jones
%            0: no se utiliza
%            1: si se utiliza
% ecu ----> tipo de ecualizado:
%            0: no se hace ningún ecualizado
%            1: ecualización de histograma
%            2: ecualización de Tan/Triggs
%            3: ...
% elip ---> recorte elíptico: devuelve sólo los pixeles contenidos
% en una elipse alrededor de la cara. Normalmente sólo debe ser
utilizado
% si antes se ha utilizado Viola-Jones
%            0: no se hace el recorte
%            1: sí se hace el recorte
% tx -----> ancho del recuadro a devolver si se usa Viola-Jones
% ty -----> alto del recuadro a devolver si se usa Viola-Jones
% see ----> igual a 0 si no se desea mostrar resultados

% SE MUESTRAN LOS RESULTADOS SI SE PIDE
if see
    figure(100);
    subplot(1,4,1);
    muestra(imagen, 'original');
end

% RECORTE DE LA IMAGEN SEGUN VIOLA-JONES
%---------------------------------------
if viola

    % se llama al programa de detección de caras de Viola-Jones y
    % se devuelven las esquinas del rectángulo que contiene la imagen


    extra = 20; %tamaño del cuadrado extra para que el detector Viola-
Jones funcione mejor
    [x, y] = size(imagen);
    new_im = uint8(zeros(x+2*extra, y+2*extra));
    new_im(extra+1:extra+x,extra+1:extra+y) = imagen;
```

```
    [x1 y1 x2 y2 ok] = violajones(new_im);

    x1=x1+1; %%% correccion del 0
    y1=y1+1;
    x2=x2+1;
    y2=y2+1;

    % se recorta el recuadro y se escala a un tamaño fijo
    if ok==1
        imagen = new_im(y1:y2,x1:x2);
        imagen = imresize(imagen, [ty tx], 'bilinear');

    else %%si el falla el detector de Viola
        imagen = imagen;
        imagen = imresize(imagen, [ty tx], 'bilinear');

    end

else
    ok=1;
end

% se muestra el resultado
if see
    figure(100);
    subplot(1,4,2);
    titulo = sprintf('Viola-Jones: %d', viola);
    muestra(imagen, titulo);
end

% ECUALIZADO DE LA IMAGEN
% -----------------------
if ecu==1

    % ecualización de histograma
    imagen = histeq(imagen);

elseif ecu==2

    % ecualización de tan/Triggs
    imagen = tantriggs(double(imagen), 0);

end

% se muestra el resultado
if see
    figure(100);
    subplot(1,4,3);
    titulo = sprintf('Ecualizado: %d', ecu);
    muestra(imagen, titulo);
end

% RECORTE ELIPTICO ALREDEDOR DE LA CARA
if elip

    % se recorta la elipse y se devuelve como vector
    [imagen, imagen2D] = rec_elipse(double(imagen), 0);
```

```
else

    % se devuelve la imagen original como vector
    elipse = imagen;
    imagen2D=imagen;  %matriz imagen
    imagen = imagen(:);%vector imagen


end

if see
    figure(100);
    subplot(1,4,4);
    titulo = sprintf('Elipse: %d', elip);
    muestra(imagen2D, titulo);
end

% valor de retorno
im_vector = imagen;
im_matriz = imagen2D;
return
```

FIGURA 6.1: Ejemplo código empleado para el estudio

Una vez terminada la primera fase, donde se obtienen todas las BBDD con todos los preprocesados aplicados, se lanzan una segunda fase donde se realiza el proceso de PCA o blanqueado de datos, WHITENING, para posteriormente testear el reconocimiento facial con los distintos preprocesados

En la segunda fase tiene mucha importancia el clasificador utilizado. Se ha utilizado un clasificador basado en distancias, denominado, el vecino más cercano (k-NN, con k=1). Se ha utilizado para el cálculo de la distancia, la distancia Euclidea.

Distancia Euclidea: $= \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$

El clasificador K-NN, es un *clasificador* denominado a *posteriori o con aprendizaje*, debido a que este clasificador se construye siguiendo un procedimiento iterativo, donde el clasificador aprende de una forma progresiva los patrones de la muestra de aprendizaje. A diferencia de los clasificadores a priori, donde el clasificador se construye en un solo paso, ya que se utiliza la muestra de aprendizaje del cual se conoce su clase cierta para el cálculo de las funciones discriminantes.

Existen dos tipos de clasificadores si se tiene en cuenta la información que se proporciona para el proceso de construcción del clasificador, es decir, para el aprendizaje:

- Clasificador supervisado: Es aquel clasificador donde un agente externo, supervisor o maestro, es quien divide la muestra en las distintas clases ciertas conocidas en las que se quiere clasificar. Este conjunto donde se conoce la clase cierta, se denomina conjunto de entrenamiento. El agente debe saber en qué momento debe parar el proceso de aprendizaje, ya que un aprendizaje excesivo puede provocar sobreajustes.

- Clasificador no supervisado: Para este tipo de clasificador no se requiere de ningún agente externo, se realiza de manera automática, gracias a técnicas de agrupamientos, donde el sistema aprende y selecciona los patrones con características similares, con lo cual las clases son determinadas automáticamente.

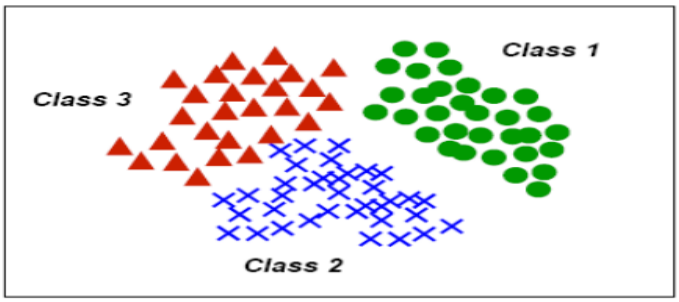


FIGURA 6.2: 3 tipos de clases: círculos, triángulos y cruces

Considerando un caso ideal donde cada agrupamiento representa a una clase y cada una de estas clases tienen asociado un agrupamiento donde se diferencia correctamente de las otras clases, uno de los problemas de la clasificación supervisada puede ser la búsqueda de las superficies que separan los distintos agrupamientos, más conocida como superficies de decisión. Esta superficie determina la región de decisión, por lo que cada clase tiene una región de decisión en el espacio X, donde para un nuevo patrón a clasificar se realizará en base a la región en la que se encuentre dentro del espacio X.

Dependiendo de como se realice la búsqueda de esta región y/o superficie de decisión, basándose en si se conoce o se supone un modelo estadístico determinado para la clase, los clasificadores pueden ser separados en dos modelos:

- Clasificadores paramétricos: A priori se tiene un conocimiento completo de la estructura estadística de la clase. El aprendizaje consiste en una estimación de los parámetros que determinan las funciones de densidad de probabilidad de la clase. Las fronteras de decisión quedan definidas por las distribuciones de probabilidad de las clases.
- Clasificadores no paramétricos: Las fronteras de decisión están definidas por los prototipos, debido a que no se supone un modelo estadístico, bien sea por la imposibilidad de asumir un modelo paramétrico adecuado o por desconocimiento del mismo.

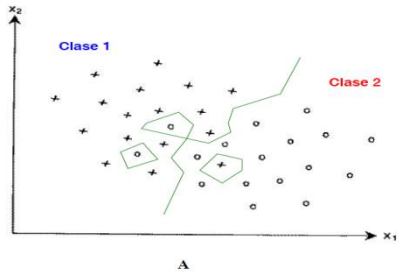


FIGURA 6.3. Ejemplo de sobreajustes de frontera

Quedando definidos los términos anteriores se puede decir que el clasificador K-NN, es del tipo supervisado no paramétrico.

En el estudio, el conjunto de imágenes de entrenamiento para el clasificador 1-NN, representa el 50% del total de cada una de las BBDD, siendo este 50%, el primer 50% del total de imágenes de un sujeto, por ejemplo, en un sujeto donde tiene un total 10 imágenes, las 5 primeras son tomadas como imágenes de entrenamiento y las otras 5 imágenes para realizar los test. El 50% restante, representan el conjunto de imágenes para realizar los Test de validación.

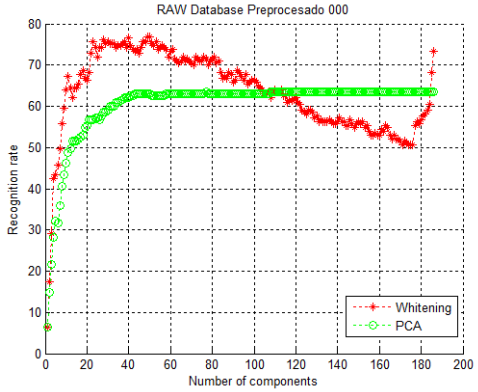
Una vez se tienen almacenadas el conjunto de imágenes de test, junto a la clase asociada, ante una nueva imagen a clasificar, se calcula la distancia euclidea con respecto al total del conjunto de imágenes de test, y se consideran los k vecinos más cercanos, clasificando esta nueva imagen según la clase de los ejemplos k más cercanos. Para el caso del estudio, al ser utilizado el vecino más cercano, se clasifica según a la clase (sujeto) del único ejemplo más cercano.

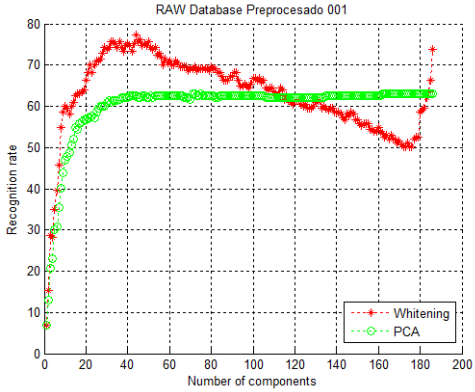
7. EXPERIMENTOS

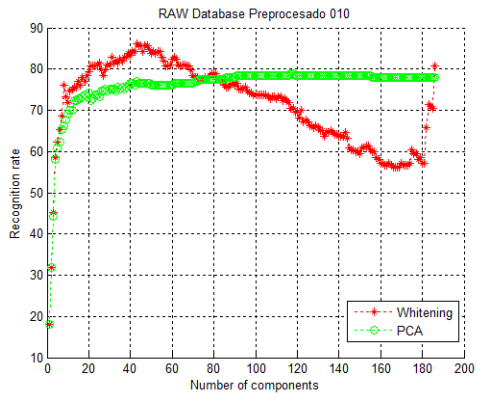
7.1. EXPERIMENTO 1

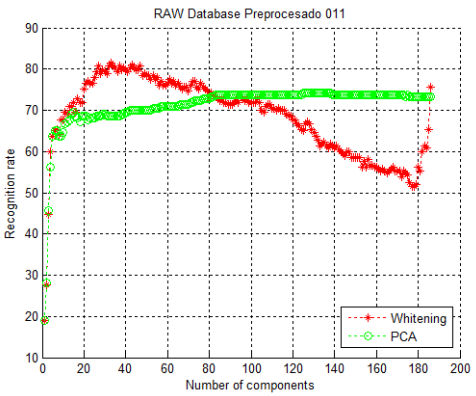
En este primer experimento se compara PCA con Whitening para todas las bases de datos y preprocesados.

7.1.1. BASE DE DATOS AR









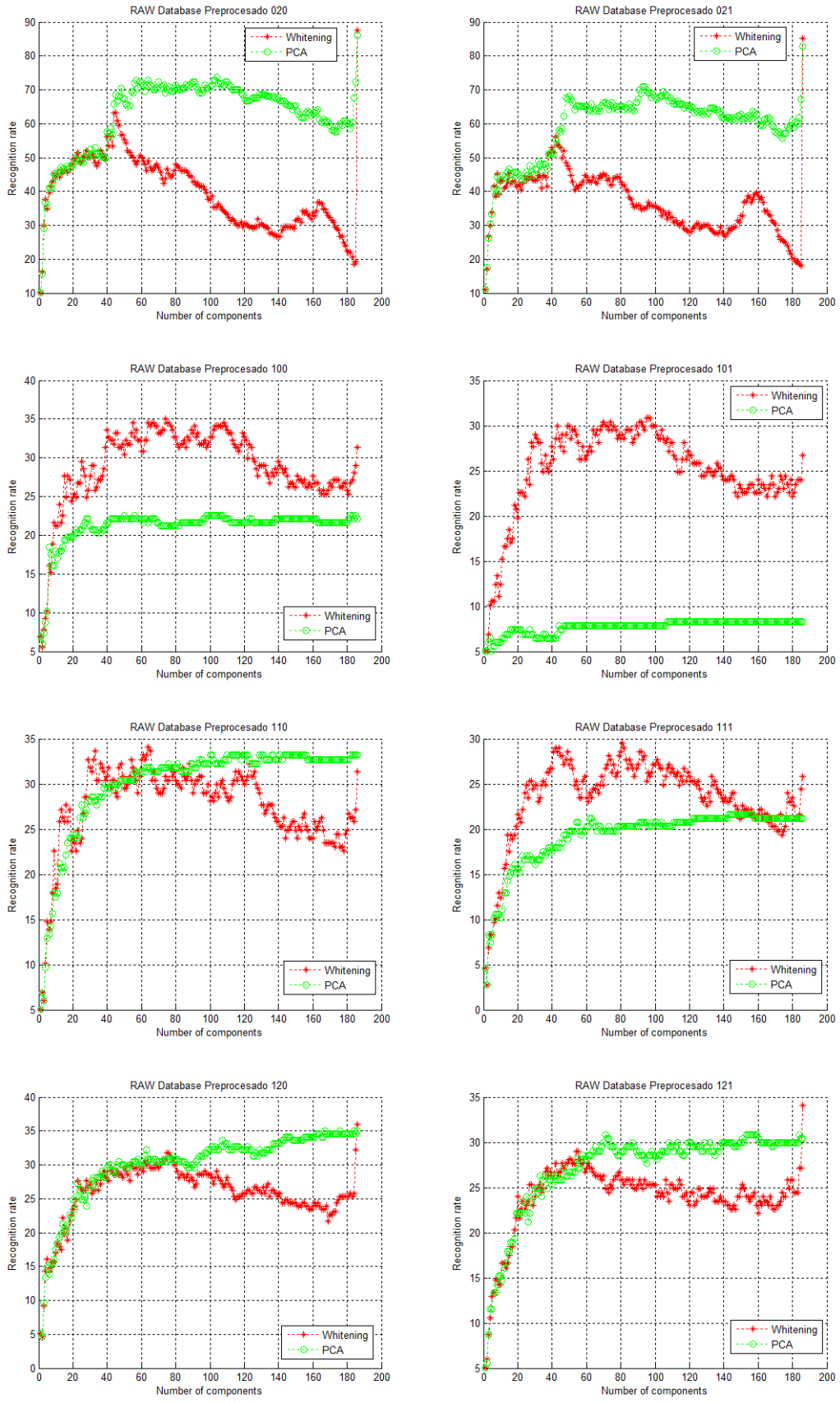


FIGURA 7.1.1.1: Gráficos de los resultados de reconocimiento para Whitening y PCA con todos los preprocesados en la BBDD AR

Preprocesado	Reconocimiento máximo (%)	Nº Componentes	Componentes (%)
000	63,60	77	41,40
001	63,13	71	38,17
010	78,80	117	62,90
011	74,19	125	67,20
020	86,17	186	100
021	82,95	186	100
100	22,58	50	26,88
101	8,29	107	57,53
110	33,18	100	53,76
111	21,66	143	76,88
120	35,02	167	89,78
121	30,88	71	38,17

TABLA 7.1.1.1: Resultados para PCA en BBDD AR. La primera columna hace referencia al preprocesado utilizado, la segunda corresponde al nivel de reconocimiento máximo alcanzado, en la tercera se muestra el número de componentes en el cual se alcanza este máximo y en la última, el porcentaje que representa este número de componentes respecto al total de componentes

Preprocesado	Reconocimiento máximo (%)	Nº Componentes	Componentes (%)
000	76,96	49	26,34
001	77,42	44	23,66
010	86,17	43	23,12
011	81,57	33	17,74
020	87,56	186	100
021	85,25	186	100
100	35,02	74	39,78
101	30,88	95	51,08
110	34,10	64	34,41
111	29,49	81	43,55
120	35,94	186	100
121	34,10	186	100

TABLA 7.1.1.2: Resultados para Whitening en BBDD AR.

Preprocesado	Mejora de Whitening sobre PCA (%)
000	121,01
001	122,63
010	109,36
011	109,94
020	101,60
021	102,78
100	155,10
101	372,22
110	102,78
111	136,17
120	102,63
121	110,45

TABLA 7.1.1.3: Comparación en el porcentaje de reconocimiento de Whitening respecto PCA en BBDD AR. Un valor por encima del 100% indica mejoría respecto al valor con el que se compara.

Preprocesado	Mejora de Whitening sobre PCA (%)
000	36,36
001	38,03
010	63,25
011	73,6
020	0
021	0
100	-48
101	11,21
110	36
111	43,36
120	-11,38
121	-161,97

TABLA 7.1.1.4: Se compara entre ambas técnicas, cuanto antes se alcanza el máximo nivel de reconocimiento por componentes.

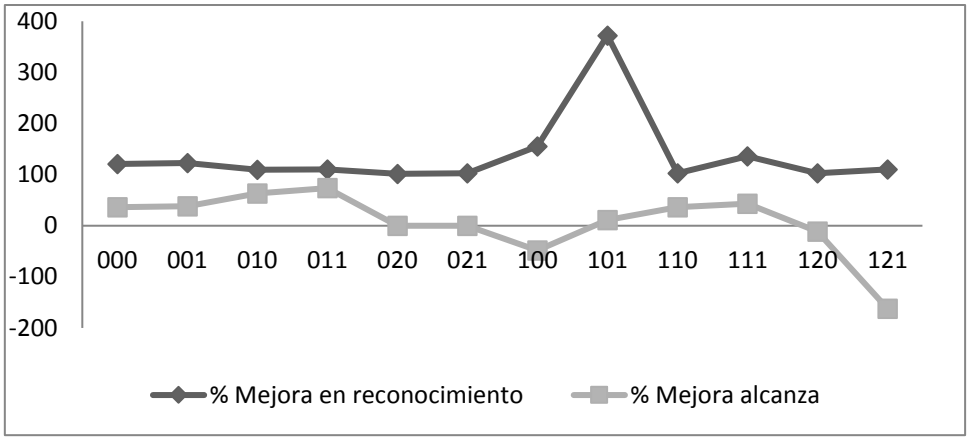


FIGURA 7.1.1.2

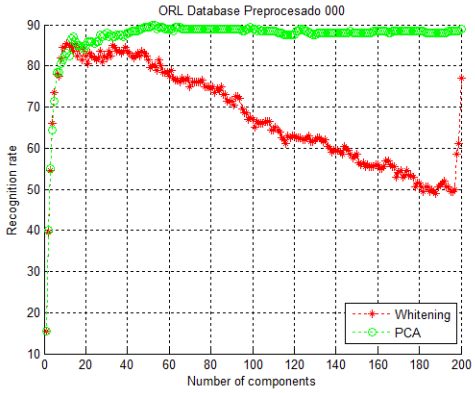
Comparando los resultados de esta primera base de datos se observa como el nivel de reconocimiento puntual es mayor tras aplicar Whitening respecto a PCA, llegado a un número de componentes determinado. Observando los gráficos se observa como en Whitening una vez se ha alcanzado la tasa máxima de reconocimiento, empeora cuantas más componentes son utilizadas. Sin embargo, con PCA, la tendencia en la tasa de reconocimiento es a mejorar conforme se utilizan más componentes.

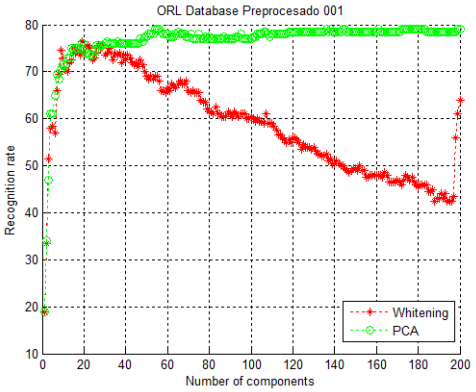
Se puede observar en la tabla 7.1.1.3, como para la base datos estudiada, AR, se mejora la tasa de reconocimiento en todos los preprocesados utilizando Whitening.

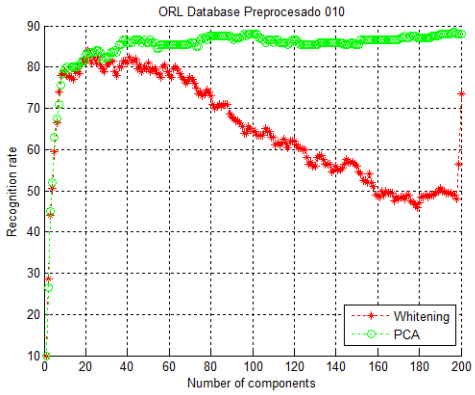
En la tabla comparativa 7.1.1.4 se observa cuanto antes se alcanza el máximo nivel de reconocimiento por preprocesado, comparándose únicamente el número de componentes utilizadas. Un valor positivo indica que se ha alcanzado antes Whitening que PCA, siendo el valor negativo el caso opuesto. Si se observa la tabla, en términos generales, se alcanza antes el porcentaje mayor de reconocimiento con Whitening que con PCA, salvo en tres ocasiones: preprocesados 100, 120 y 121; con PCA es donde se alcanza con mayor brevedad el máximo nivel de reconocimiento.

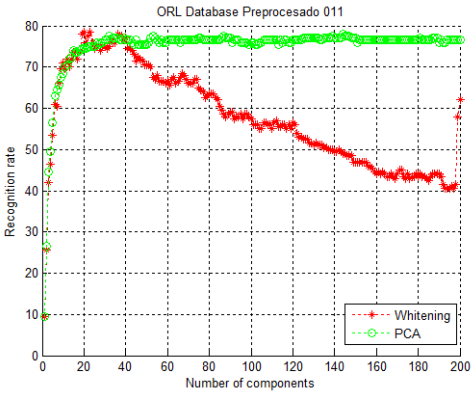
El último gráfico, figura 7.1.1.2, sirve para observar mejor las dos últimas tablas, 7.1.1.3 y 7.1.14. Se observa como con carácter general como se mejora tanto en el porcentaje de reconocimiento como en el tiempo para alcanzar este máximo, siempre y cuando se lance el reconocimiento de forma gradual respecto al número de componentes utilizadas, como se ha realizado en este estudio, es decir, se utiliza primero la componente 1, después se lanza el reconocimiento con la componente 1 y la 2, etc.

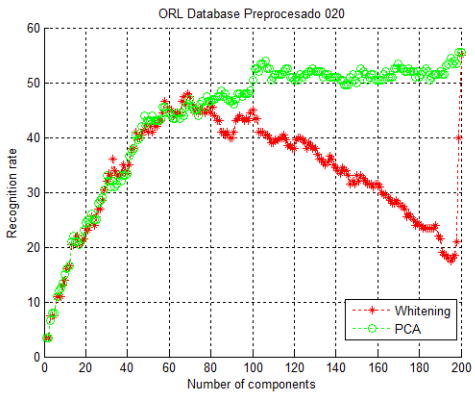
7.1.2. BASE DE DATOS ORL

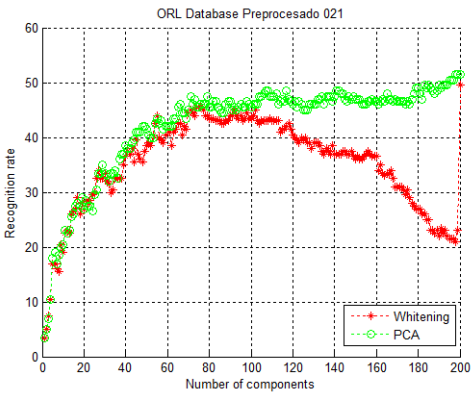


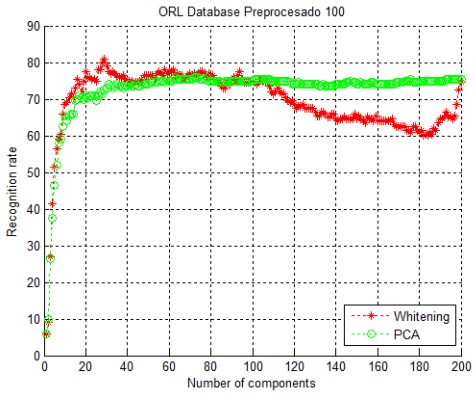


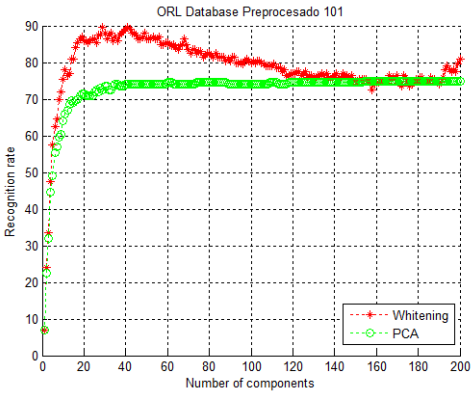












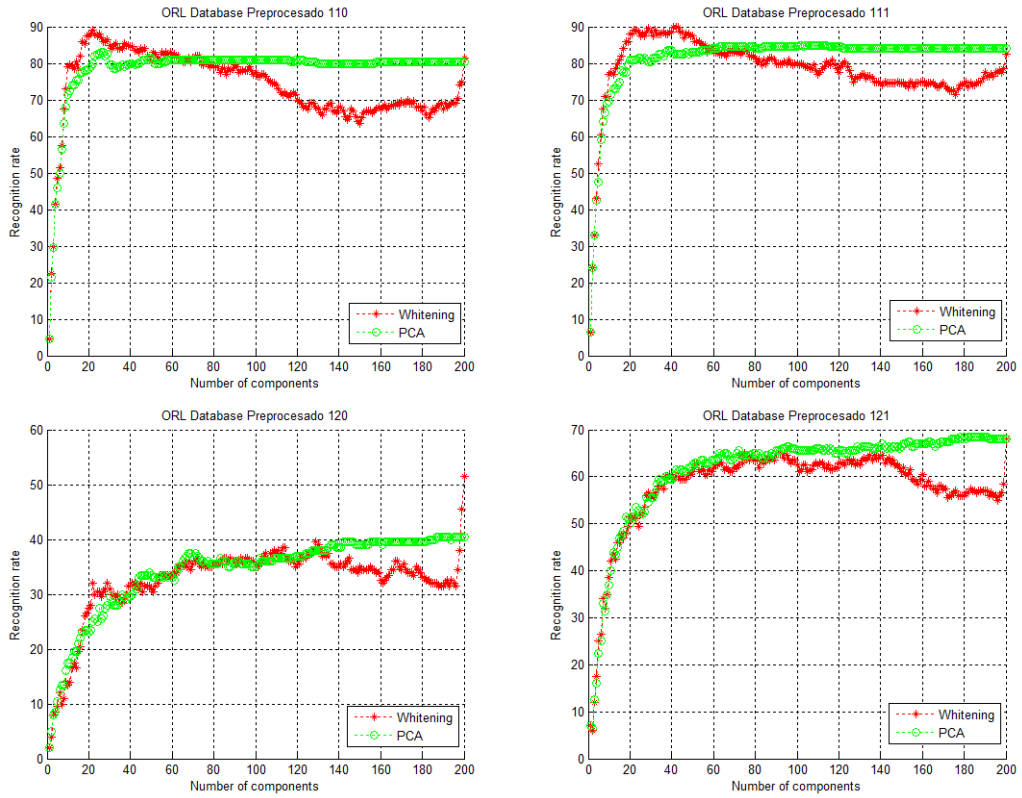


FIGURA 7.1.2.1: Gráficos de los resultados de reconocimiento para Whitening y PCA con todos los preprocesados en la BBDD ORL

Preprocesado	Reconocimiento máximo (%)	Componente	Componentes (%)
000	90	52	26
001	79	54	27
010	88,5	196	98
011	78	144	72
020	55,5	199	99,5
021	51,5	198	99
100	76	63	31,5
101	75	154	77
110	83	26	13
111	85	101	50,5
120	40,5	187	93,5
121	68,5	178	89

TABLA 7.1.2.1: Resultados para PCA en BBDD ORL

Preprocesado	Reconocimiento máximo (%)	Componente	Componentes (%)
000	85,5	11	5,5
001	76,5	19	9,5
010	84	21	10,5
011	78,5	20	10
020	55	200	100
021	49,5	200	100
100	81	29	14,5
101	89,5	29	14,5
110	89	22	11
111	90	41	20,5
120	51,5	200	100
121	68	200	100

TABLA 7.1.2.2: Resultados para WHITENING en BBDD ORL

En esta base de datos, Whitening mejora a PCA cuando se aplica el detector de caras Viola-Jones:

Preprocesado	Mejora de Whitening sobre PCA (%)
000	95
001	96,83
010	94,91
011	100,64
020	99,1
021	96,12
100	106,58
101	119,34
110	107,23
111	105,88
120	127,16
121	99,24

TABLA 7.1.2.3: Comparación en el porcentaje de reconocimiento de Whitening respecto PCA en BBDD ORL.

Preprocesado	Mejora de Whitening sobre PCA (%)
000	78,85
001	64,81
010	89,29
011	86,11
020	-0,50
021	-1,01
100	53,97
101	81,17
110	15,38
111	59,41
120	-6,95
121	-12,36

TABLA 7.1.2.4: Se compara entre ambas técnicas, cuanto antes se alcanza el máximo nivel de reconocimiento por componentes.

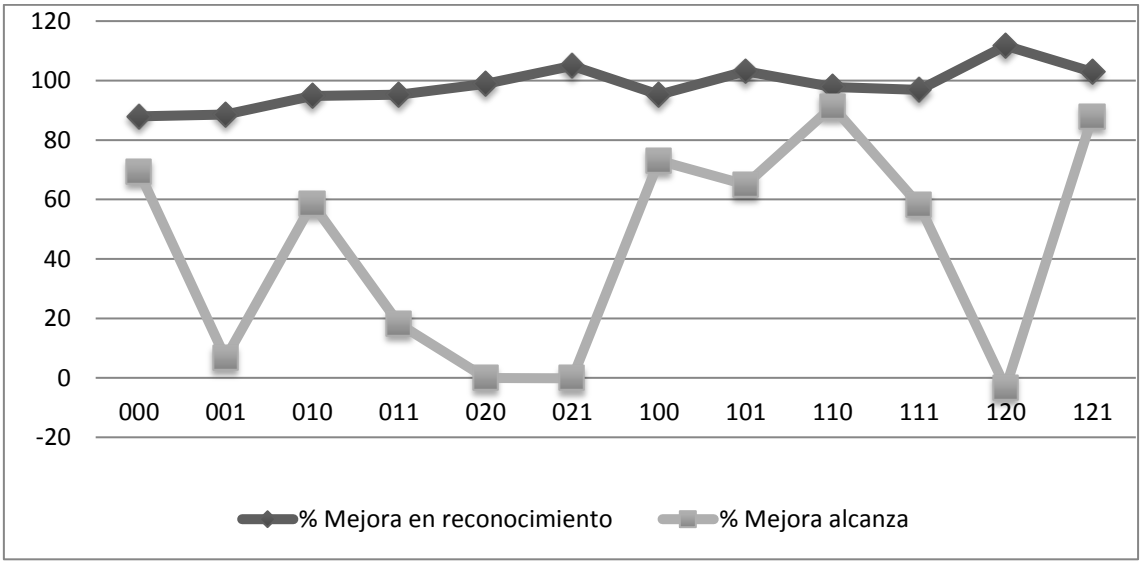


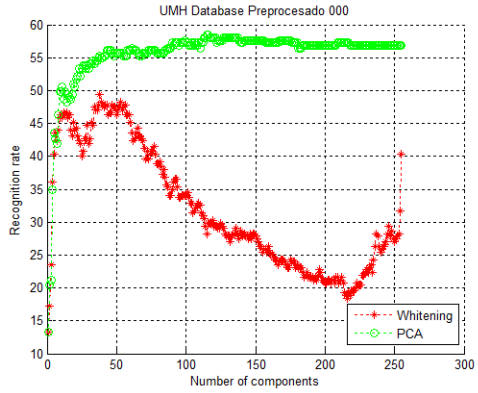
FIGURA 7.1.2.2

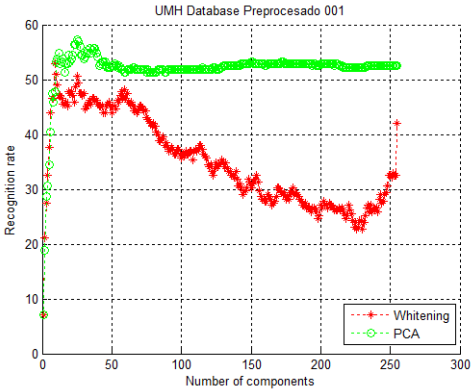
Se observa en la tabla 7.1.2.3 de la base de datos ORL, donde son comparadas ambas técnicas por lo que respecta al porcentaje de reconocimiento, como con la técnica

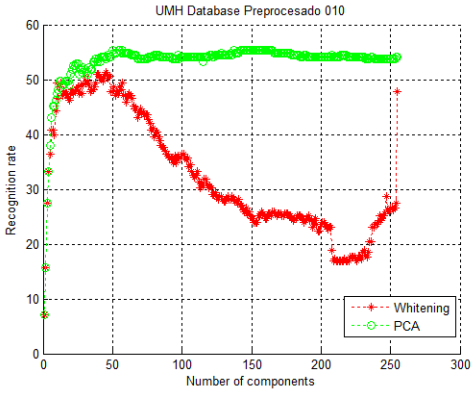
Whitening los valores de reconocimiento son mejores una vez se ha aplicado Viola-Jones, excepto para el preprocesado 121. Sin el detector de Viola-Jones el porcentaje de reconocimiento con Whitening, sin ser mejor, es muy similar a PCA.

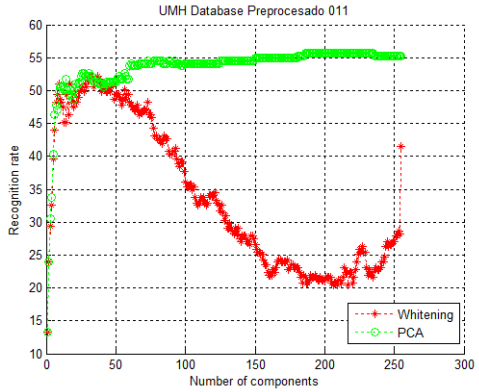
Al igual que en la anterior base de datos, con Whitening, se alcanza el nivel máximo de reconocimiento en las primeras componentes para posteriormente empeorar. Con PCA se tarda un poco más en alcanzar el nivel máximo de reconocimiento en términos generales, pero la tendencia es a quedarse en valores cercanos al máximo porcentaje de reconocimiento conforme el número de componentes aumenta, al contrario que con Whitening, donde en algunos casos empeora de forma notable.

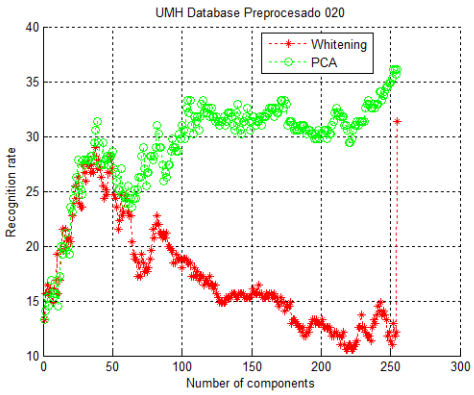
7.1.3. BASE DE DATOS UMH

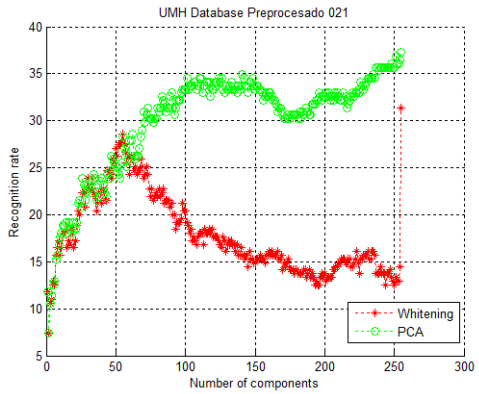


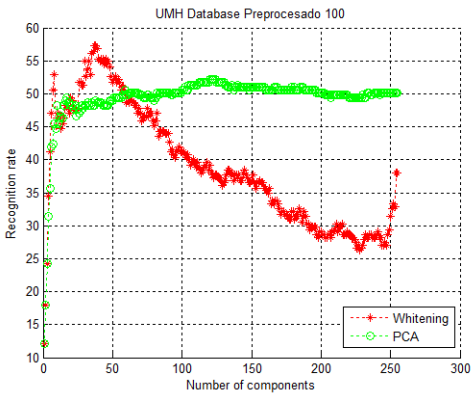


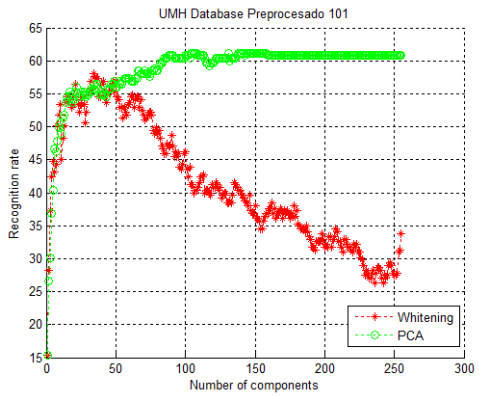












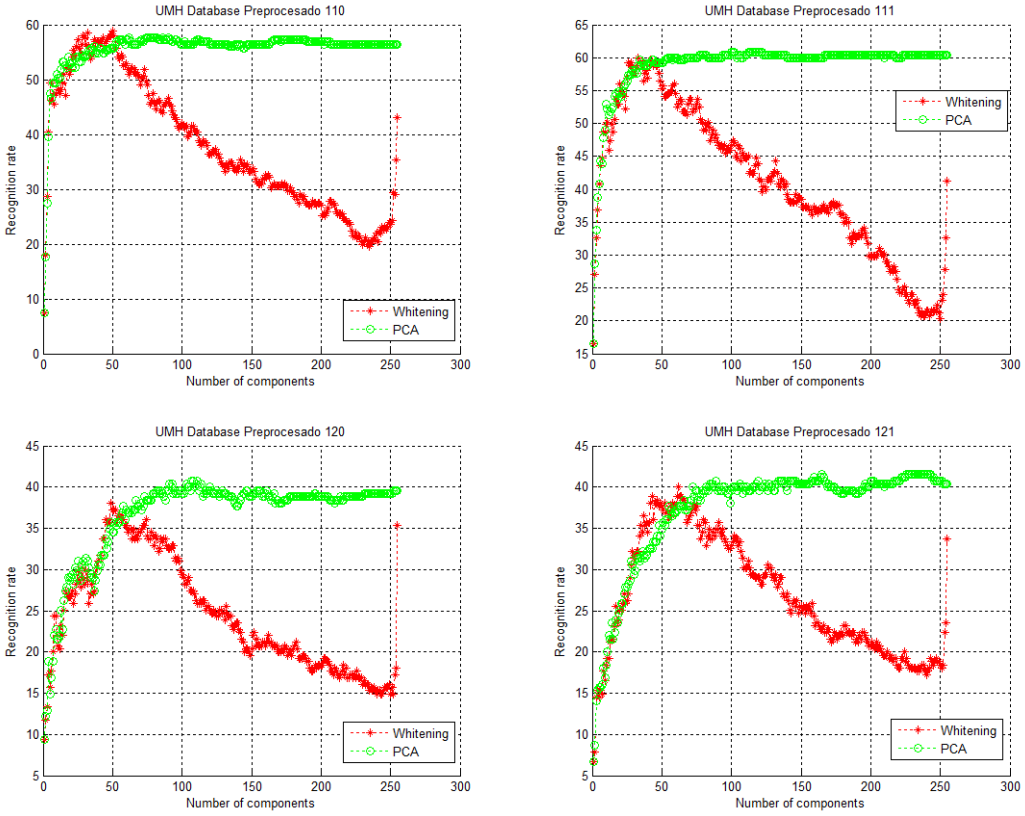


FIGURA 7.1.3.1: Gráficos de los resultados de reconocimiento para Whitening y PCA con todos los preprocesados en la BBDD UMH

Preprocesado	Reconocimiento máximo (%)	Nº Componente	Componentes (%)
000	58,43	115	45,1
001	57,25	25	9,80
010	55,29	50	19,61
011	55,69	186	72,94
020	36,08	252	98,82
021	37,25	255	100
100	52,16	120	47,06
101	61,18	104	40,78
110	57,65	62	24,31
111	60,78	100	39,22
120	40,78	107	41,96
121	41,57	165	64,71

TABLA 7.1.3.1: Resultados para PCA en BBDD UMH

Preprocesado	Reconocimiento máximo (%)	Nº Componente	Componentes (%)
000	49,41	38	14,90
001	52,94	9	3,53
010	51,37	45	17,65
011	52,16	32	12,55
020	31,37	255	100
021	31,37	255	100
100	57,25	37	14,51
101	58,04	34	13,33
110	58,82	50	19,61
111	60	33	12,94
120	38,04	49	19,22
121	40	62	24,31

TABLA 7.1.3.2: Resultados para WHITENING en BBDD UMH

Preprocesado	Mejora de Whitening sobre PCA (%)
000	84,56
001	92,47
010	92,91
011	93,66
020	86,96
021	84,21
100	109,77
101	94,87
110	102,04
111	98,72
120	93,27
121	96,23

TABLA 7.1.2.3: Comparación en el porcentaje de reconocimiento de Whitening respecto PCA en BBDD UMH

Preprocesado	Mejora de Whitening sobre PCA (%)
000	66,96
001	64
010	10
011	82,0
020	-1,19
021	0
100	69,17
101	67,31
110	19,35
111	67
120	54,21
121	62,42

TABLA 7.1.2.4: Se compara entre ambas técnicas, cuanto antes se alcanza el máximo nivel de reconocimiento por componentes.

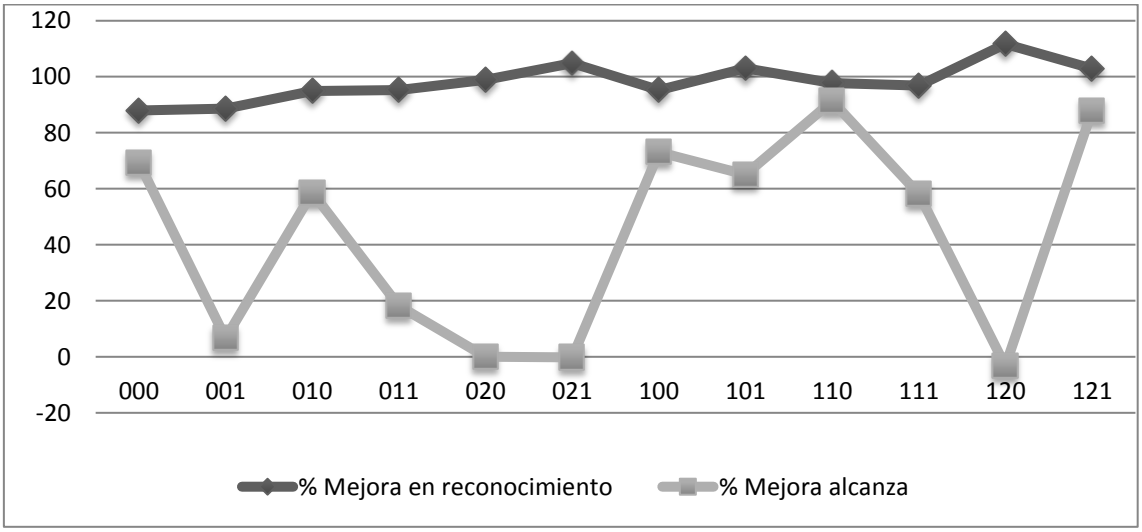
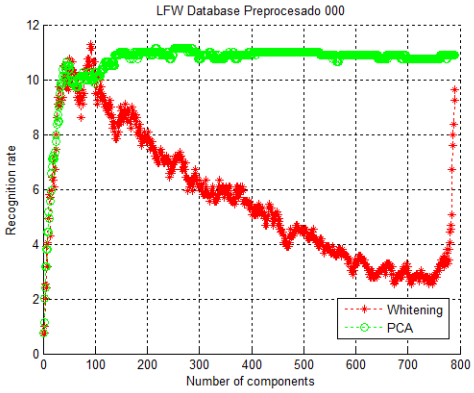


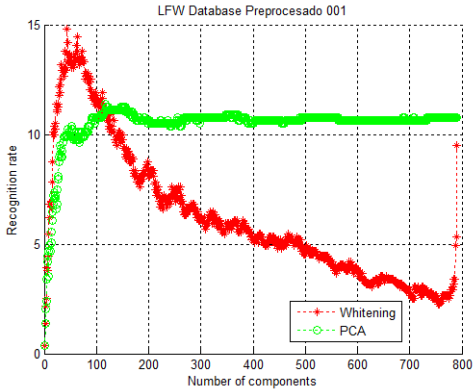
FIGURA 7.1.3.2

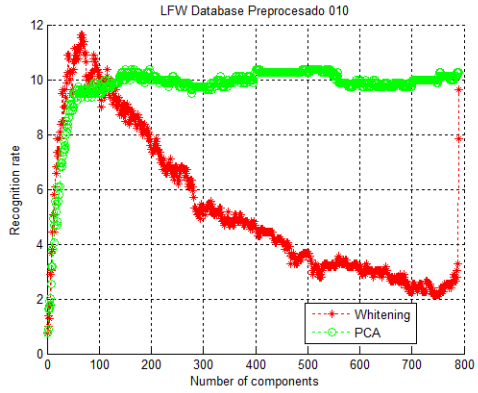
Para esta base de datos, UMH, ocurre el caso opuesto observado en las anteriores dos bases de datos estudiadas. Cuando es utilizado Whitening para el reconocimiento, empeora respecto a PCA en cuanto al reconocimiento en todos los casos, menos en los casos 100 y 110.

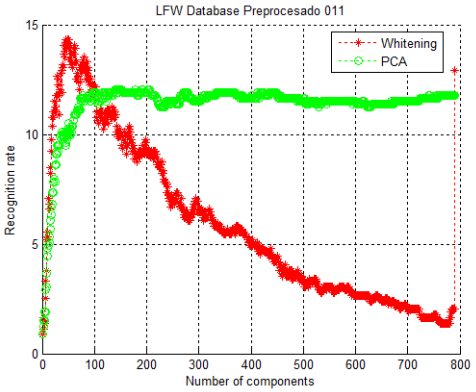
Por lo que respecta a cuanto se tarda en alcanzar este máximo porcentaje de reconocimiento. Si se compara el número de componentes que se han utilizado para llegar a ese máximo, con Whitening se mejora en todos los preprocesados el tiempo de máximo reconocimiento, menos en el caso 020. Sucede lo mismo que en los anteriores estudios, con Whitening se alcanza el nivel de reconocimiento con mayor brevedad que con PCA. Sin embargo, conforme aumenta el número de componentes tiende a empeorar, al contrario que con PCA.

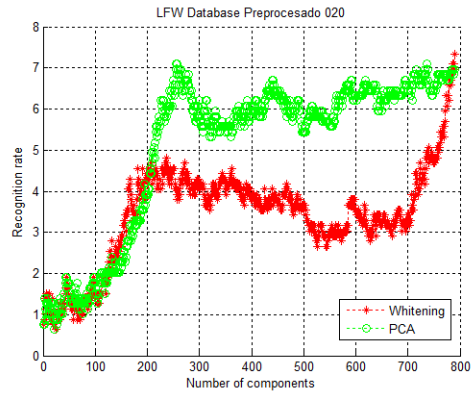
7.1.4. BASE DE DATOS LFW

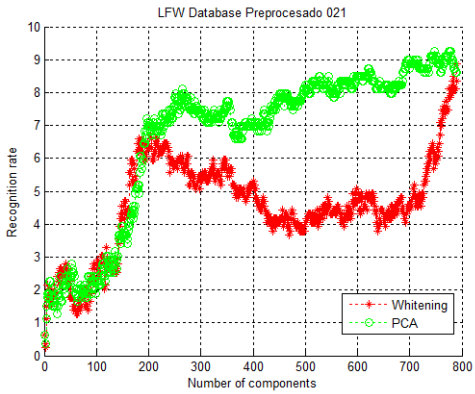


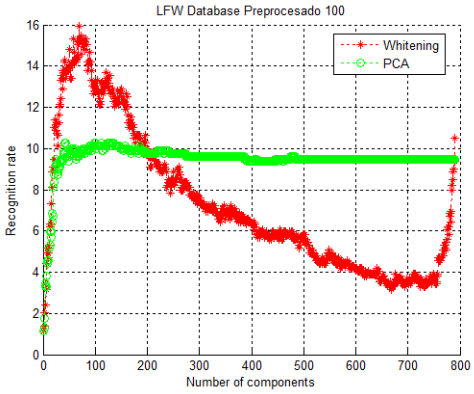


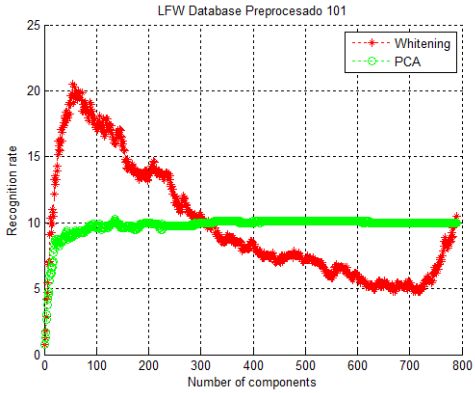












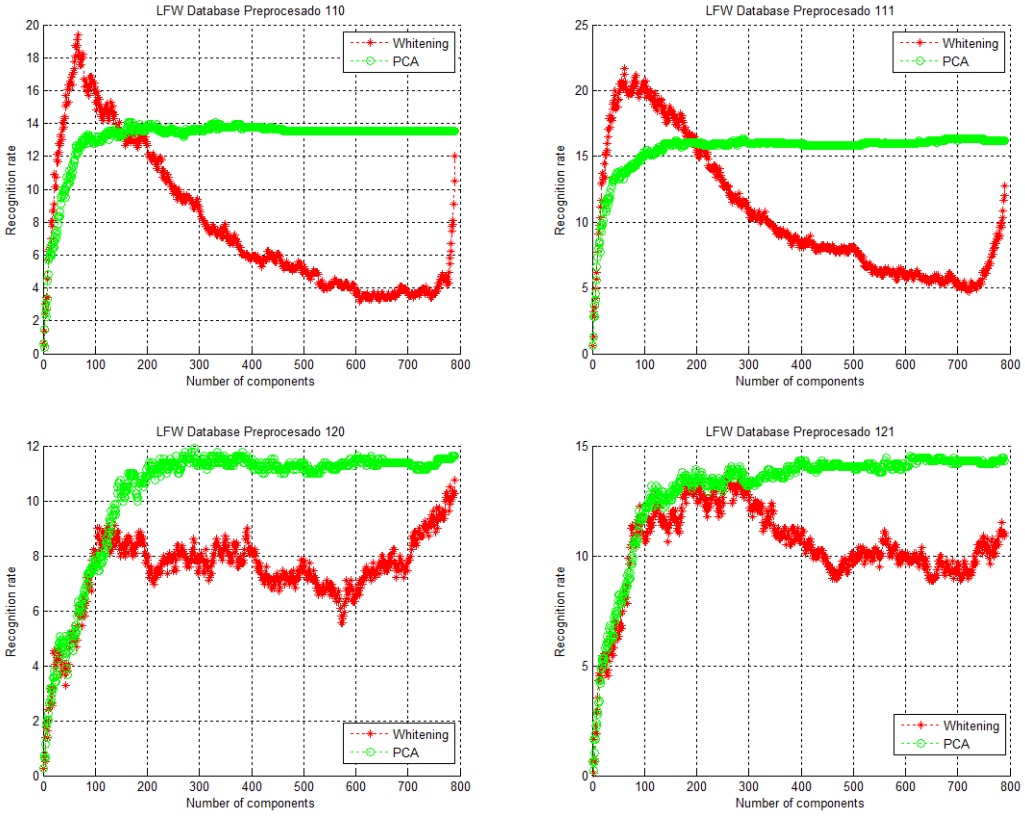


FIGURA 7.1.4.1: Gráficos de los resultados de reconocimiento para Whitening y PCA con todos los preprocesados en la BBDD LFW

Preprocesado	Reconocimiento máximo (%)	Nº Componente	Componentes (%)
000	11,14	215	27,21
001	11,39	117	14,81
010	10,38	151	19,11
011	12,02	139	17,59
020	7,09	255	32,28
021	9,24	747	94,56
100	10,25	42	5,32
101	10,25	135	17,09
110	14,05	164	20,76
111	16,33	289	36,58
120	11,90	290	36,71
121	14,56	609	77,09

TABLA 7.1.4.1: Resultados para PCA en BBDD LFW

Preprocesado	Reconocimiento máximo (%)	Nº Componente	Componentes (%)
000	11,27	92	11,65
001	14,81	44	5,57
010	11,65	65	8,23
011	14,30	46	5,82
020	7,34	789	99,87
021	8,86	790	100
100	15,95	69	8,73
101	20,51	54	6,83
110	19,37	68	8,61
111	21,65	61	7,72
120	10,76	790	100
121	13,8	261	33,04

TABLA 7.1.4.2: Resultados para WHITENING en BBDD LFW

Preprocesado	Mejora de Whitening sobre PCA (%)
000	101,14
001	130
010	112,19
011	118,95
020	103,57
021	95,89
100	155,55
101	200
110	137,84
111	132,56
120	90,42
121	94,78

TABLA 7.1.4.3: Comparación en el porcentaje de reconocimiento de Whitening respecto PCA en BBDD LFW

Preprocesado	Mejora de Whitening sobre PCA (%)
000	57,21
001	62,39
010	56,95
011	66,91
020	-209,41
021	-5,76
100	-64,29
101	60
110	58,54
111	78,89
120	-172,41
121	57,14

TABLA 7.1.4.4: Se compara entre ambas técnicas, cuanto antes se alcanza el máximo nivel de reconocimiento por componentes.

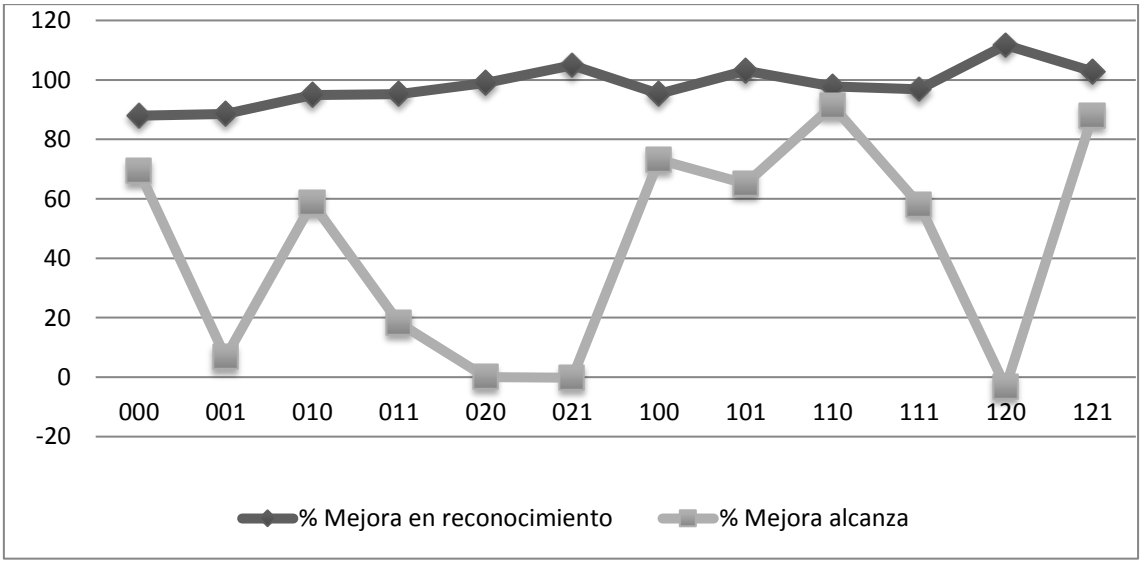


FIGURA 7.1.4.2

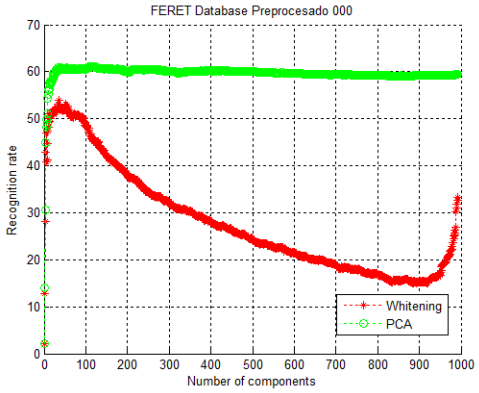
En esta base de datos, LFW, se puede observar en las tablas separadas por técnicas, 7.1.4.1 y 7.1.4.2, como los porcentajes de reconocimiento son muy bajos para ambas técnicas.

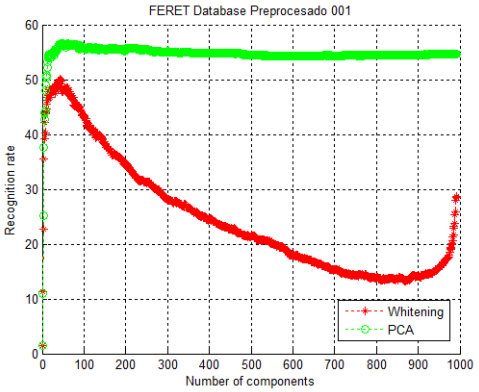
Si se comparan los resultados obtenidos entre ambas técnicas, el porcentaje de reconocimiento es siempre mayor con la técnica de Whitening, menos en los casos 021, 120 y 121.

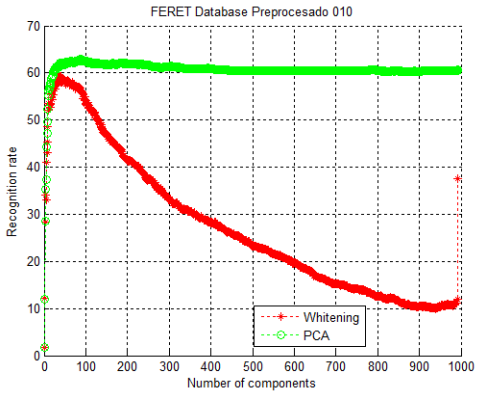
En esta base de datos, una vez se ha utilizado Viola-Jones el porcentaje de reconocimiento aumenta notablemente con la técnica Whitening respecto a PCA, aunque el ratio del mismo continua siendo bajo comparado con el de las otras bases de datos estudiadas.

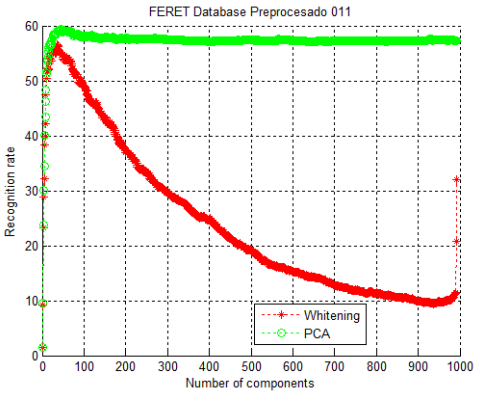
Por lo que respecta al tiempo en alcanzar el máximo nivel de reconocimiento entre ambas técnicas, este es el mismo que en los otros casos estudiados, con Whitening se alcanza en las primeras componentes para descender seguidamente y aumentar al final. Para la técnica PCA, el nivel de reconocimiento aumenta rápidamente en las primeras componentes y se mantiene prácticamente constante hasta las últimas componentes.

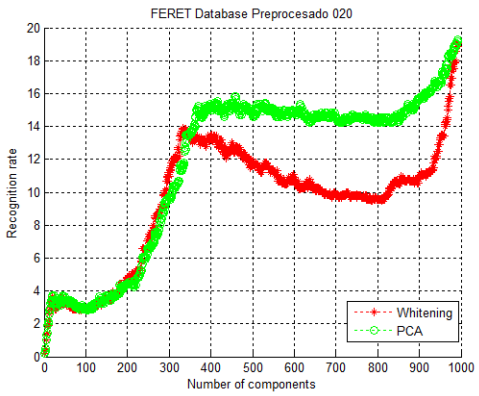
7.1.5. BASE DE DATOS FERET

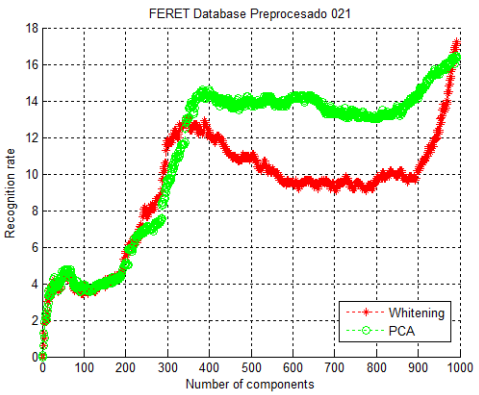


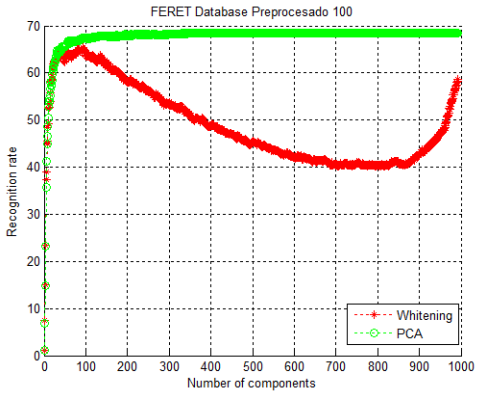


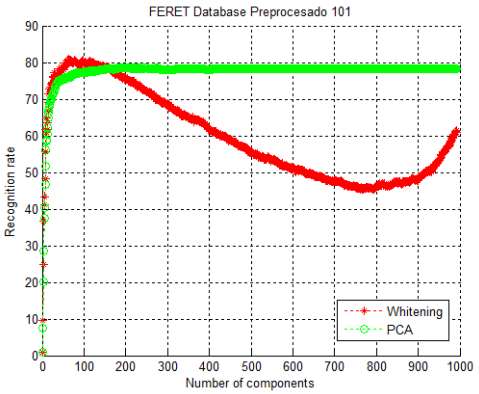












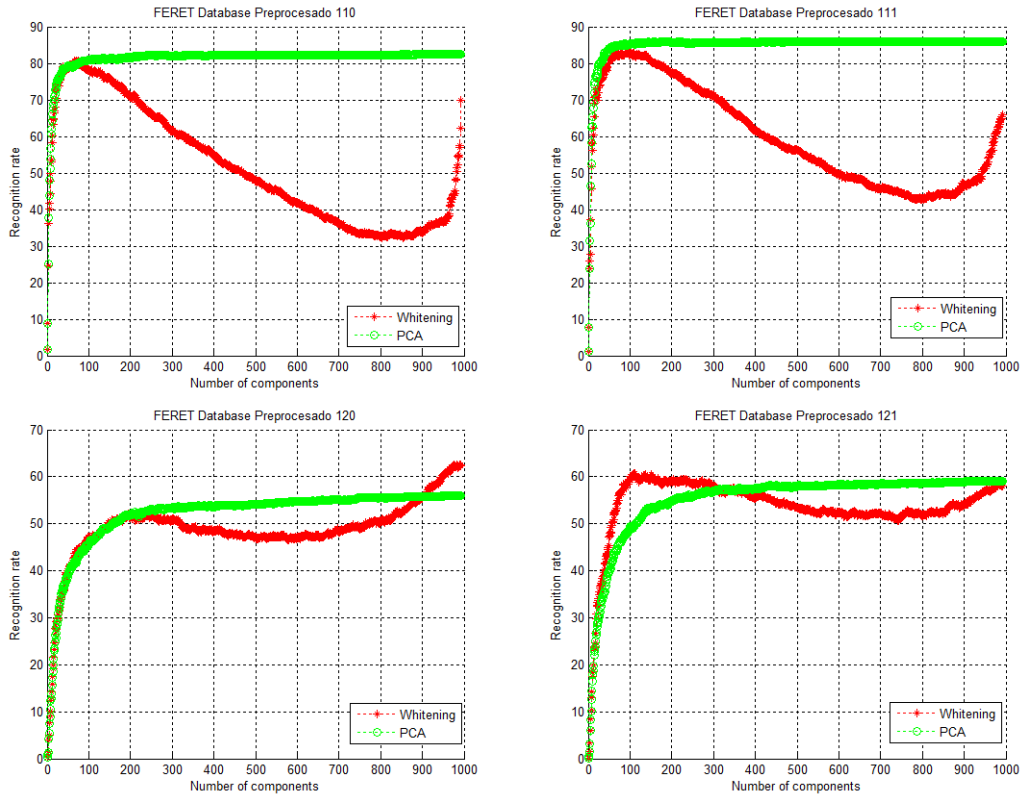


FIGURA 7.1.4.1: Gráficos de los resultados de reconocimiento para Whitening y PCA con todos los preprocesados en la BBDD FERET

Preprocesado	Reconocimiento máximo (%)	Nº Componente	Componentes (%)
000	61,19	114	11,49
001	56,55	45	4,54
010	62,90	87	8,77
011	59,37	44	4,43
020	19,25	991	99,90
021	16,43	990	99,80
100	68,55	346	34,88
101	78,53	183	18,45
110	82,46	850	85,68
111	85,99	217	21,87
120	55,95	945	95,26
121	59,07	908	91,53

TABLA 7.1.4.1: Resultados para PCA en BBDD FERET

Preprocesado	Reconocimiento máximo (%)	Nº Componente	Componentes (%)
000	53,83	35	3,53
001	50,10	42	4,23
010	59,68	36	3,63
011	56,55	36	3,63
020	19,05	991	99,9
021	17,24	992	100
100	65,32	93	9,37
101	80,95	64	6,45
110	80,64	74	7,46
111	83,27	91	9,17
120	62,5	976	98,39
121	60,79	110	11,09

TABLA 7.1.4.2: Resultados para WHITENING en BBDD FERET

Preprocesado	Mejora de Whitening sobre PCA (%)
000	87,97
001	88,59
010	94,87
011	95,25
020	98,95
021	104,91
100	95,29
101	103,08
110	97,8
111	96,83
120	111,71
121	102,90

TABLA 7.1.4.3: Comparación en el porcentaje de reconocimiento de Whitening respecto PCA en BBDD FERET

Preprocesado	Mejora de Whitening sobre PCA (%)
000	69,3
001	6,67
010	58,62
011	18,18
020	0
021	-0,2
100	73,12
101	65,03
110	91,29
111	58,06
120	-3,28
121	87,88

TABLA 7.1.4.4: Se compara entre ambas técnicas, cuanto antes se alcanza el máximo nivel de reconocimiento por componentes.

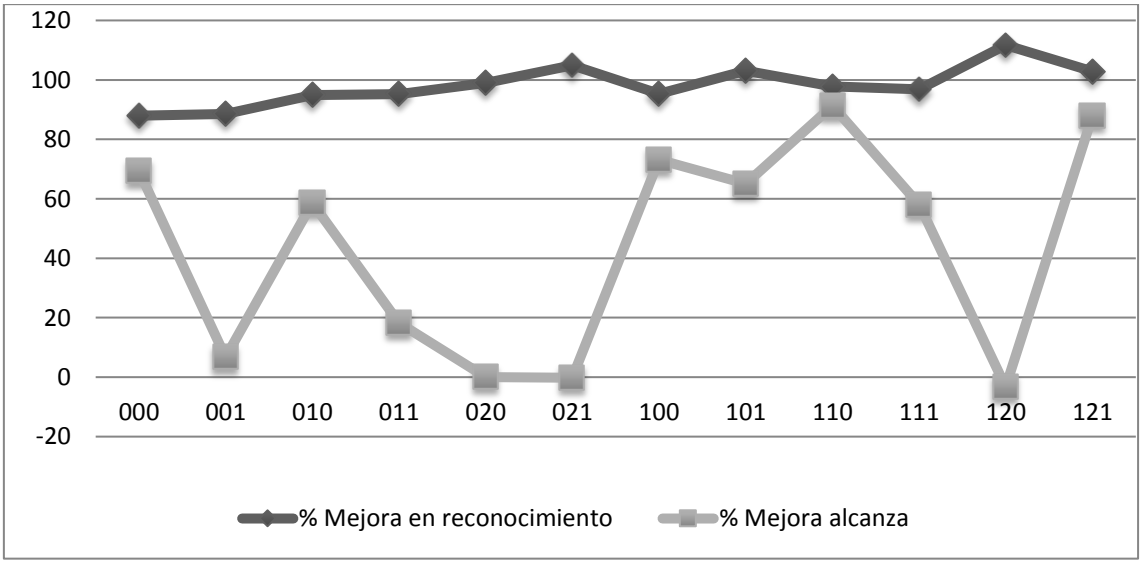


FIGURA 7.1.4.2

En la última base de datos del presente estudio, FERET, si se observan los datos de las tablas y gráficos obtenidos, se aprecia como con la técnica de PCA se alcanzan mejores resultados de porcentaje de reconocimiento respecto a Whitening, siendo superior en la mayor parte de los preprocesados salvo para los preprocesados 021, 101, 120 y 121.

Tras ser aplicado Viola-Jones, en ambas técnicas, se observa como el porcentaje de reconocimiento aumenta notablemente, especialmente para los preprocesados 120 y 121.

En los gráficos se aprecia como antes de ser aplicado Viola- Jones para PCA, el nivel máximo de reconocimiento se alcanza en las primeras componentes, disminuyendo de manera muy paulatina, pero no de una manera drástica, conforme aumentan este número de componentes. Ocurre el caso contrario cuando se ha aplicado Viola- Jones, se alcanza el nivel de reconocimiento máximo cuantas más componentes se utilizan, llegando a valores cercanos al máximo porcentaje de reconocimiento en las primeras componentes, para ir aumentando paulatinamente conforme se utilizan más componentes, al contrario que el caso anterior descrito. Para el caso de Whitening, el comportamiento es similar al resto de bases utilizadas, se alcanza el nivel de reconocimiento en las primeras componentes para después descender rápidamente a medida que son utilizadas más componentes. Cabe destacar como tanto para PCA como para Whitening, antes y después de Viola-Jones, la curva que se observa en los gráficos de los niveles de reconocimiento antes de llegar a los máximos y mínimos, tiene un carácter más abrupto para el caso donde no se ha aplicado Viola-Jones y más suave cuando se ha utilizado, es decir, el número de componentes cerca del nivel máximo de reconocimiento es menor, sin ser aplicado Viola-Jones que cuando se ha aplicado Viola-Jones.

7.1.6. RESUMEN EXPERIMENTO 1

Tras los resultados de este primer experimento, si se hace un resumen de cuando la técnica Whitening ha mejorado a PCA por lo que respecta a la tasa máxima de reconocimiento, se obtiene:

- Para base de datos, AR, se mejora en 12 de los 12 posibles casos
- Para base de datos, ORL, se mejora en 6 de los 12 posibles casos
- Para base de datos, UMH, se mejora en 2 de los 12 posibles casos
- Para base de datos, LFW, se mejora en 9 de los 12 posibles casos
- Para base de datos, FERET, se mejora en 4 de los 12 posibles casos

Por base de datos existe un total de 12 casos, las bases de datos donde se realiza el estudio son 5, sumando un total de 60 casos. Como se observa en los datos anteriores resumidos, Whitening, mejora en 32 de los 60 posibles casos, dejando para PCA los restantes 28 de 60 posibles casos. Estos 32 representan el 53.3% del total.

Si son analizados cuando se alcanzan estos máximos, separándolos por las dos técnicas estudiadas, PCA y Whitening, los resultados que se obtienen son:

- En el caso de PCA, si se observan los resultados de los porcentajes de cuando sucede el máximo nivel de reconocimiento, no se puede establecer de forma precisa cual puede ser el número óptimo de componentes a utilizar, ya que las componentes donde se alcanzan los mayores porcentajes de reconocimiento cambian de una base de datos a otra.

- Para Whitening poder establecer el número de componentes óptimo es más viable que en el caso de PCA. El nivel de reconocimiento máximo para esta técnica tiene un comportamiento prácticamente exponencial en las primeras componentes, por lo tanto, se puede establecer que el número de componentes óptimo a utilizar está alrededor del 51% del total de las componentes de cada base de datos. Con este porcentaje de las componentes disponibles, se puede obtener el 100% de las tasas de reconocimiento

máximas en todas las bases de datos estudiadas. Este porcentaje, puede incluso ser menor, si se excluye la base de datos AR, dejando el nivel óptimo de componentes a utilizar en un 21% del total. Con estas conclusiones, se puede indicar que con Whitening el tiempo computacional, así como los datos a manejar disminuyen considerablemente respecto a PCA.

En las anteriores conclusiones no se han tenido en cuenta los resultados para los preprocesados 020, 021, 120 y 121, ya que como se ha ido observando en cada base de datos, los máximos porcentajes de reconocimiento para estos preprocesados se alcanzan cuando se ha utilizado el mayor número de componentes disponibles, incluso la totalidad de las componentes disponibles por cada DB.

7.2. EXPERIMENTO 2

7.2.1. Base de datos AR

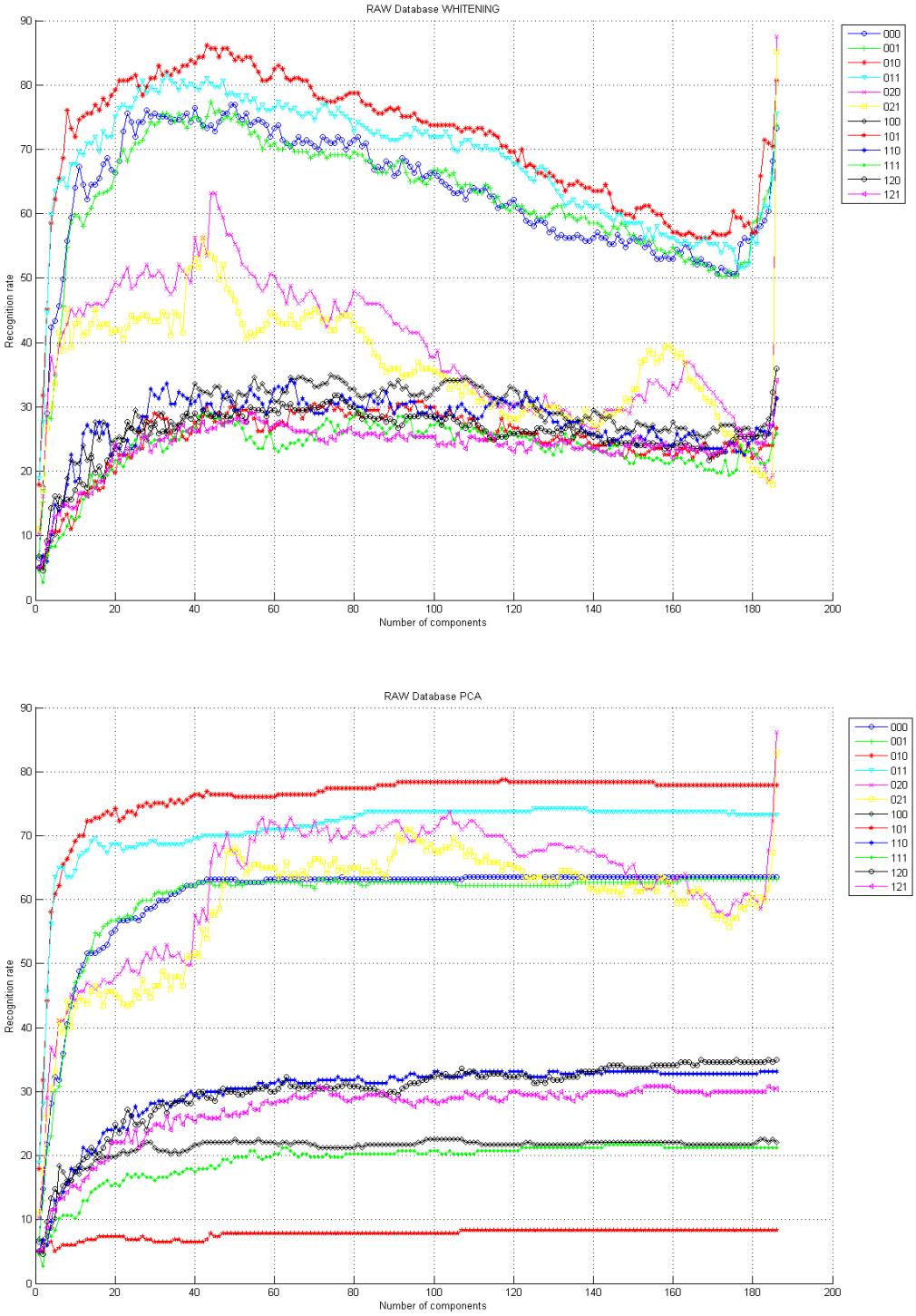


FIGURA 7.2.1.1: Gráficos de reconocimiento para Whitening y PCA con todos los preprocesados en BBDD AR

Posición	Preprocesado	Reconocimiento (%)
1	020	87,56
2	010	86,17
3	021	85,25
4	011	81,57
5	001	77,42
6	000	76,96
7	120	35,94
8	100	35,02
9	110	34,10
10	121	34,10
11	101	30,88
12	111	29,49

TABLA 7.2.1.1: Clasificación de mejor a peor preprocesado en base de datos AR con Whitening en porcentaje de reconocimiento

Posición	Preprocesado	Reconocimiento (%)
1	020	86,17
2	021	82,95
3	010	78,80
4	011	74,19
5	000	63,59
6	001	63,13
7	120	35,02
8	110	33,18
9	121	30,88
10	100	22,58
11	111	21,66
12	101	8,29

TABLA 7.2.1.2: Clasificación de mejor a peor preprocesado en base de datos AR con PCA en porcentaje de reconocimiento

Se han agrupados todos los resultados de reconocimiento para una misma base de datos y separados por técnicas, Whitening y PCA.

En esta primera base de datos de este experimento se puede observar como los preprocesados con un ratio de reconocimiento mayor son para el 010 y el 011 para Whitening, aunque en los preprocesados 020 y 021 se alcanza el mayor porcentaje de reconocimiento de una forma puntual cuando se han utilizado todas las componentes, 87.56% y 85.25% respectivamente.

En el caso de PCA, los preprocesados donde se consigue una mayor tasa de reconocimiento son los mismos que para Whitening más el 020 y el 021. En este caso también se alcanza el máximo nivel de reconocimiento cuando se han utilizadas todas las componentes, con porcentajes de reconocimiento de 86.17% y 82.95% respectivamente.

Se observa como en ambas técnicas tras ser aplicado el detector de Viola-Jones los resultados de reconocimiento empeoran notablemente respecto a los casos sin él.

7.2.2. Base de datos ORL

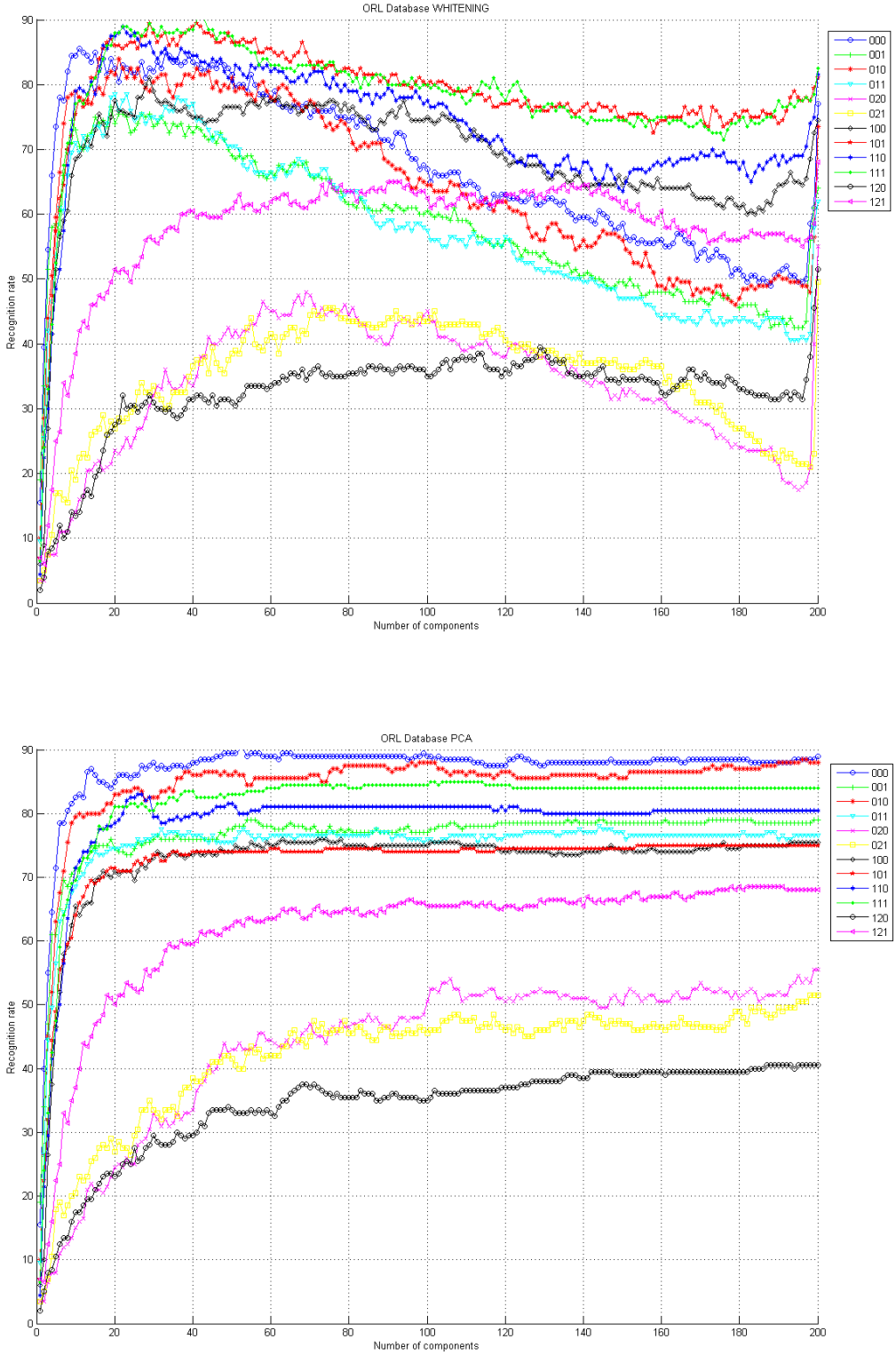


FIGURA 7.2.2.1: Gráficos de reconocimiento para Whitening y PCA con todos los preprocesados en BBDD ORL

Posición	Preprocesado	Reconocimiento (%)
1	111	90
2	101	89,5
3	110	89
4	000	85,5
5	010	84
6	100	81
7	011	78,5
8	001	76,5
9	121	68
10	020	55
11	120	51,5
12	021	49,5

TABLA 7.2.2.1: Clasificación de mejor a peor preprocesado en base de datos AR con Whitening en porcentaje de reconocimiento

Posición	Preprocesado	Reconocimiento (%)
1	000	90
2	010	88,5
3	111	85
4	110	83
5	001	79
6	011	78
7	100	76
8	101	75
9	121	68,5
10	020	55,5
11	021	51,5
12	120	40,5

TABLA 7.2.2.2: Clasificación de mejor a peor preprocesado en base de datos AR con PCA en porcentaje de reconocimiento

Para la base de datos ORL los mayores porcentajes de reconocimiento con la técnica Whitening se obtienen en los preprocesados 111, 101 y 110, donde los dos primeros tienen resultados muy similares conforme aumenta el número de componentes. Los peores resultados se obtienen con los resultados 120, 021, 020 seguido por el 121 para Whitening. En los resultados de PCA los mejores preprocesados son los 000, 010 y 111. Los peores resultados también se obtienen donde se ha aplicado el preprocesado del ecualizado Tan-Triggs 120, 021, 020 seguido del 121.

7.2.3. Base de datos UMH

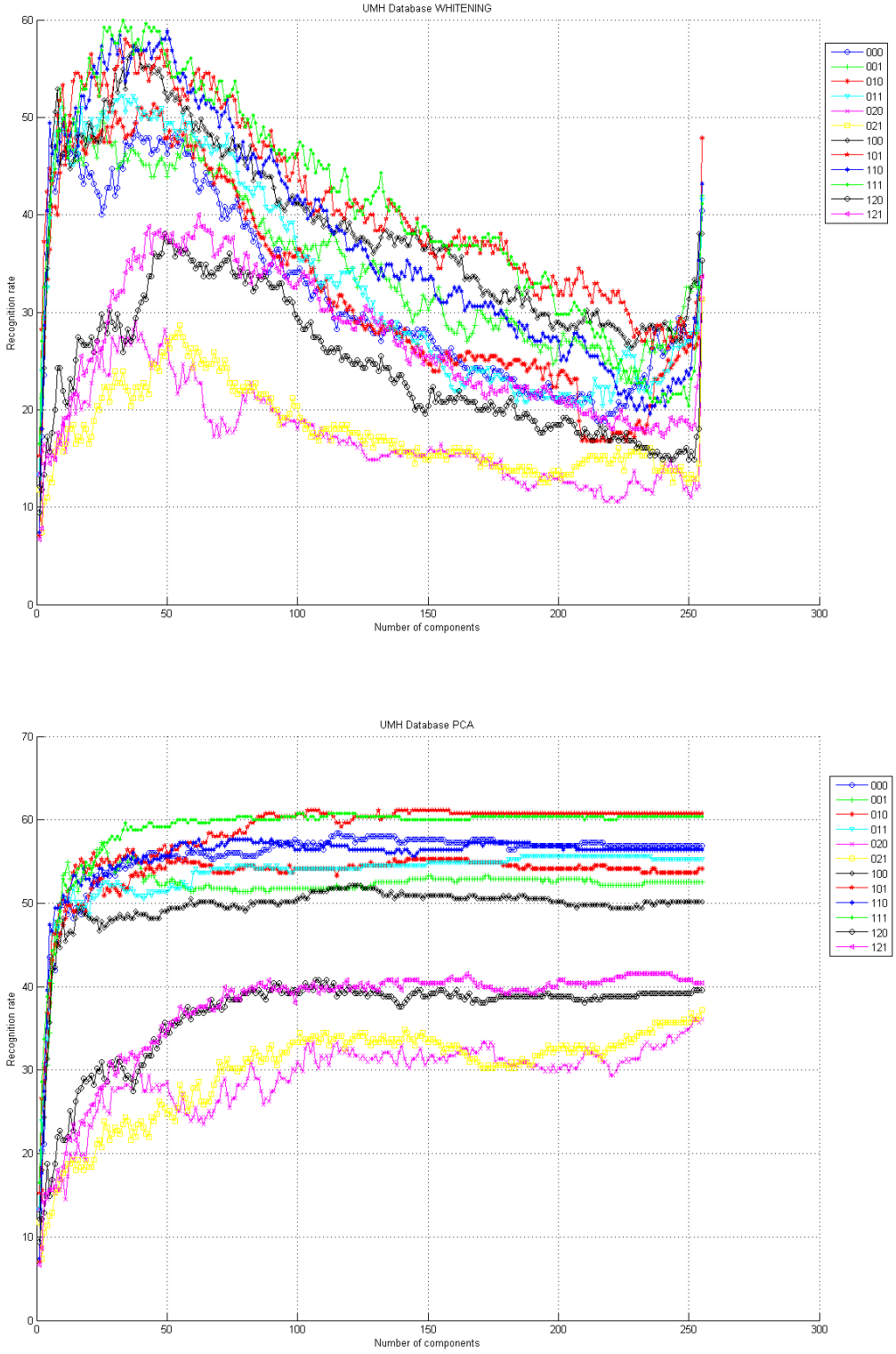


FIGURA 7.2.3.1: Gráficos de reconocimiento para Whitening y PCA con todos los preprocesados en BBDD UMH

Posición	Preprocesado	Reconocimiento (%)
1	111	60
2	110	58,82
3	101	58,04
4	100	57,25
5	001	52,94
6	011	52,16
7	010	51,37
8	000	49,41
9	121	40
10	120	38,04
11	020	31,37
12	021	31,37

TABLA 7.2.3.1: Clasificación de mejor a peor preprocesado en base de datos AR con Whitening en porcentaje de reconocimiento

Posición	Preprocesado	Reconocimiento (%)
1	101	61,18
2	111	60,78
3	000	58,43
4	110	57,65
5	001	57,25
6	011	55,69
7	010	55,29
8	100	52,16
9	121	41,57
10	120	40,78
11	021	37,25
12	020	36,08

TABLA 7.2.3.2: Clasificación de mejor a peor preprocesado en base de datos AR con PCA en porcentaje de reconocimiento

Si se observan los resultados que se obtienen aunándose todos los preprocesados bajo una misma técnica, se aprecia como para la base de datos UMH los mejores resultados bajo la técnica de Whitening se obtienen cuando se ha aplicado Viola-Jones, 111, 110, 101 y 100, de mejor a menor respectivamente. Las peores tasas de reconocimiento se obtienen cuando se ha aplicado ecualizado Tan-Triggs, 021, 020, 120 y 121, de menor a mayor tasa, respectivamente.

Para el caso PCA, los mejores resultados se obtienen también tras ser aplicado Viola-Jones, 101, 111 seguido del 000, estando muy parejo con el 110, pero teniendo la tasa puntual de reconocimiento más alta en las imágenes sin preprocesar. En este caso, los peores resultados también se obtienen cuando se ha aplicado el ecualizado Tan-Triggs, 020, 021, 120 y 121 respectivamente.

Se puede apreciar como los peores resultados se obtienen, para ambas técnicas, cuando se ha aplicado el ecualizado Tan-Triggs más el ecualizado de histograma.

7.2.4. Base de datos LFW

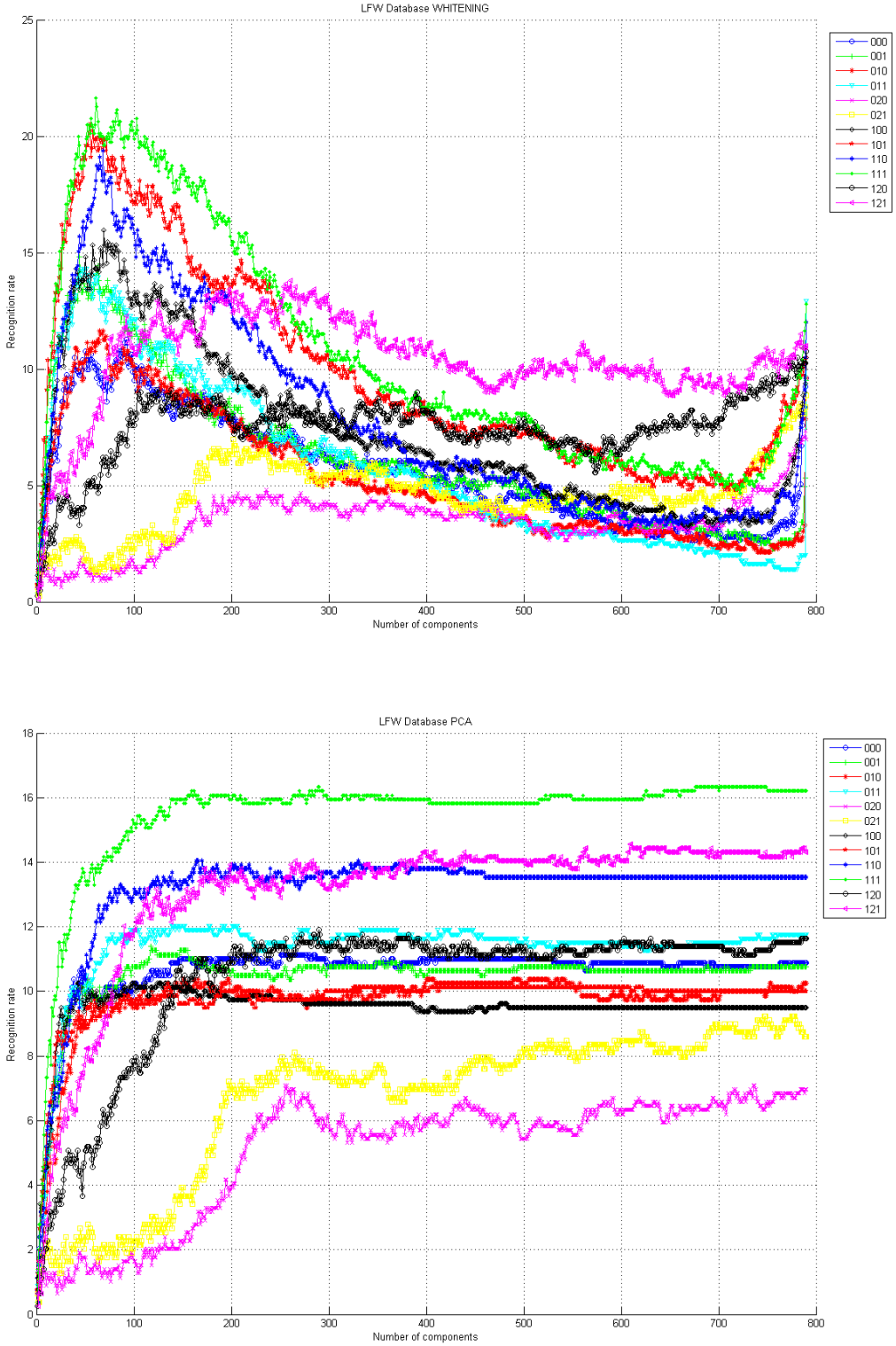


FIGURA 7.2.1.1: Gráficos de reconocimiento para Whitening y PCA con todos los preprocesados en BBDD LFW

Posición	Preprocesado	Reconocimiento (%)
1	111	21,65
2	101	20,51
3	110	19,37
4	100	15,95
5	001	14,81
6	011	14,30
7	121	13,8
8	010	11,65
9	000	11,27
10	120	10,76
11	021	8,86
12	020	7,34

TABLA 7.2.4.1: Clasificación de mejor a peor preprocesado en base de datos AR con Whitening en porcentaje de reconocimiento

Posición	Preprocesado	Reconocimiento (%)
1	111	16,33
2	121	14,53
3	110	14,05
4	011	12,02
5	120	11,9
6	001	11,39
7	000	11,14
8	010	10,38
9	100	10,25
10	101	10,25
11	021	9,24
12	020	7,09

TABLA 7.2.4.2: Clasificación de mejor a peor preprocesado en base de datos AR con PCA en porcentaje de reconocimiento

LFW es la es la base de datos donde se obtienen los peores resultados de porcentaje de reconocimiento.

En esta BBDD se observa como los mejores preprocesados para Whitening se dan cuando se ha aplicado Viola-Jones, 111, 101 y 110, respectivamente. Los peores resultados se obtienen con los preprocesados 020, 021, 120 y 121, respectivamente.

En la técnica de PCA los mejores resultados se obtienen, también, cuando se ha aplicado Viola-Jones, 111, 110 y 121, respectivamente. Para PCA los peores resultados también se obtienen con el ecualizado TAN-TRIGGS sin el detector Viola-Jones, 020, 021 y 111.

7.2.5. Base de datos FERET

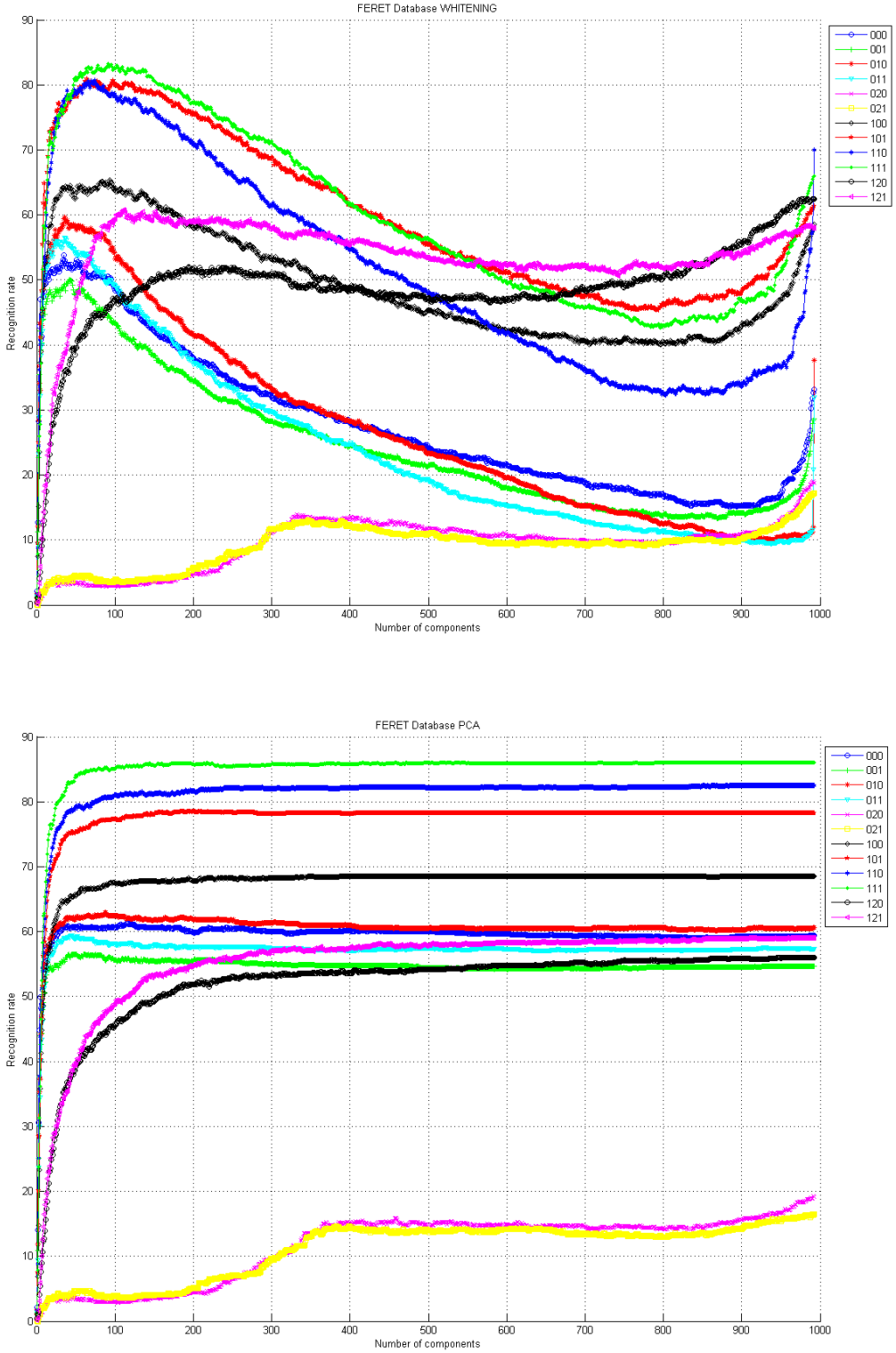


FIGURA 7.2.5.1: Gráficos de reconocimiento para Whitening y PCA con todos los preprocesados en BBDD FERET

Posición	Preprocesado	Reconocimiento (%)
1	111	83,266129
2	101	80,947581
3	110	80,645161
4	100	65,322581
5	120	62,5
6	121	60,78629
7	010	59,677419
8	011	56,552419
9	000	53,830645
10	001	50,100806
11	020	19,052419
12	021	17,237903

TABLA 7.2.5.1: Clasificación de mejor a peor preprocesado en base de datos AR con Whitening en porcentaje de reconocimiento

Posición	Preprocesado	Reconocimiento (%)
1	111	85,99
2	110	82,46
3	101	78,53
4	100	68,55
5	010	62,9
6	000	61,19
7	011	59,37
8	121	59,07
9	001	56,55
10	120	55,95
11	020	19,25
12	021	16,43

TABLA 7.2.5.2: Clasificación de mejor a peor preprocesado en base de datos AR con PCA en porcentaje de reconocimiento

Para la última base de datos del estudio, FERET, los preprocesados, tanto para Whitening como para PCA, que arrojan los mejores resultados también se alcanzan cuando se ha aplicado Viola-Jones.

Con la técnica Whitening los mejores resultados se obtienen con los preprocesados 111, 101 y 110, donde los peores resultados, para esta base de datos, también son obtenidos cuando se ha aplicado el ecualizado Tan-Triggs, 021, 020 y 001, respectivamente.

Para PCA, los mejores resultados se han obtenido con los preprocesados 111, 110 y 101, siendo los peores los preprocesados 021, 020 y 120, respectivamente.

7.2.6. RESUMEN EXPERIMENTO 2

Si se realiza una tabla de clasificación donde se tiene en cuenta el número de veces que un preprocesado ha obtenido un mayor ratio de reconocimiento que el resto, sumándose el resultado en todas las bases de datos, los preprocesados quedan clasificados de la siguiente forma:

Posición	Preprocesado	Nº Casos	Total	Técnica
1	111	44 42	86 de 110	Whitening PCA
2	110	40 39	79 de 110	Whitening PCA
3	101	40 26	66 de 110	Whitening PCA
3	010	31 35	66 de 110	Whitening PCA
4	000	24 38	62 de 110	Whitening PCA
4	011	29 33	62 de 110	Whitening PCA
5	001	27 29	56 de 110	Whitening PCA
5	100	34 22	56 de 110	Whitening PCA
6	121	19 23	42 de 110	Whitening PCA
7	120	17 16	33 de 110	Whitening PCA
8	020	15 14	29 de 110	Whitening PCA
9	021	10 13	23 de 110	Whitening PCA

TABLA 7.2.6.1: Clasificación de preprocesados

Separando los resultados por técnicas, la clasificación de preprocesados es la siguiente:

Posición	Preprocesado	Nº Casos
1	111	44
2	101	40
2	110	40
3	100	34
4	010	31
5	011	29
6	001	27
7	000	24
8	121	19
9	120	17
10	020	15
11	021	10

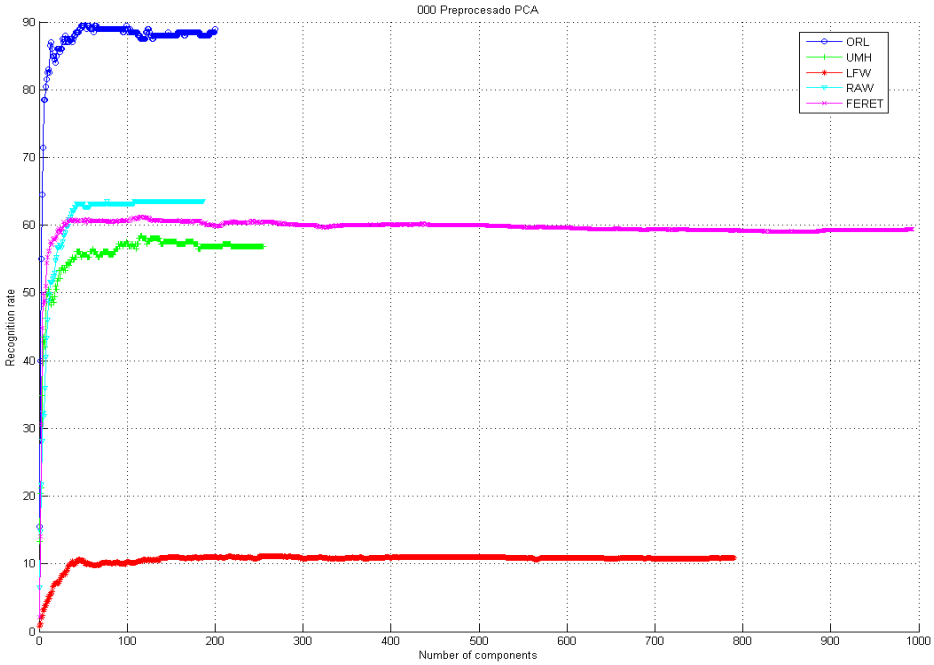
TABLA 7.2.6.2: Clasificación de preprocesados para la técnica Whitening

Posición	Preprocesado	Nº Casos
1	111	42
2	110	39
3	000	38
4	010	35
5	011	33
6	001	29
7	101	26
8	121	23
9	100	22
10	120	16
11	020	14
12	021	13

TABLA 7.2.6.3: Clasificación de preprocesados para la técnica PCA

7.3. EXPERIMENTO 3

7.3.1. PREPROCESADO 000



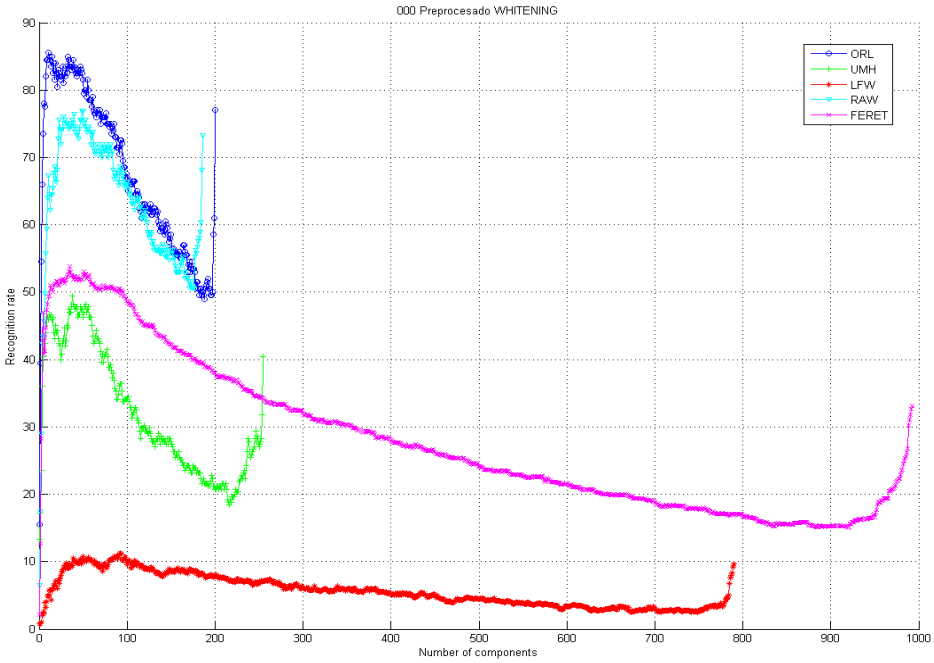


FIGURA 7.3.1: Resultados aunados para todas las BBDD en el preprocesado 001 por técnica, PCA o Whitening

En este tercer experimiento, se utilizan los resultados del preprocesado 000 en cada base de datos, es decir, las imágenes sin preprocesar, como los resultados de referencia para los otros preprocesados en esa misma base de datos, para poder observar si el preprocesado aplicado mejora o empeora el resultado respecto a las imágenes originales.

7.3.2. PREPROCESADO 001

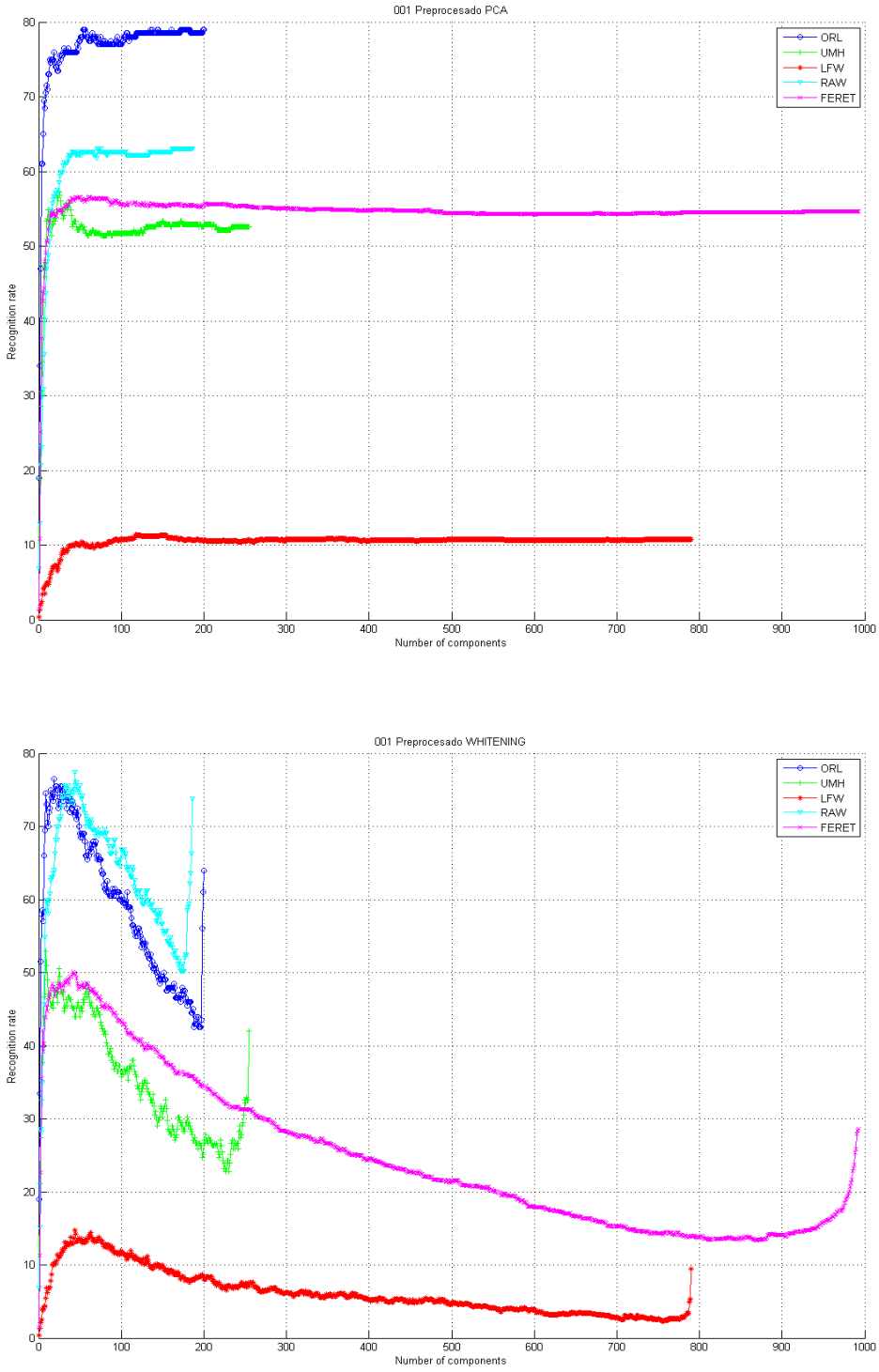


FIGURA 7.3.2 Resultados aunados para todas las BBDD en el preprocesado 001 por técnica, PCA o Whitening

		000	001	Mejora (%)
AR	PCA	63,59	63,13	-0,72
	Whitening	76,96	**77,42**	0,6
ORL	PCA	90	79	-12,22
	Whitening	85,5	76,5	-10,53
UMH	PCA	58,43	57,25	-2,01
	Whitening	49,41	**52,94**	7,14
LFW	PCA	11,14	**11,39**	2,27
	Whitening	11,27	**14,81**	31,46
FERET	PCA	61,19	56,55	-7,58
	Whitening	53,83	50,1	-6,93

TABLA 7.3.1: Comparación de los mejores resultados entre preprocesado 000 y 001 para las técnicas PCA y Whitening.

Como se puede observar en la tabla adjunta, se comprueba como el preprocesado 001, donde se ha aplicado la máscara elipsoidal, mejora el porcentaje de reconocimiento para PCA únicamente en una ocasión, en la base de datos LFW.

Para Whitening mejora en las mismas bases de datos, más la base de datos AR, con lo que podemos decir que mejora el resultado en 3 de 5 bases de datos, mientras que para PCA se mejora en 1 de 5.

7.3.3. PREPROCESADO 010

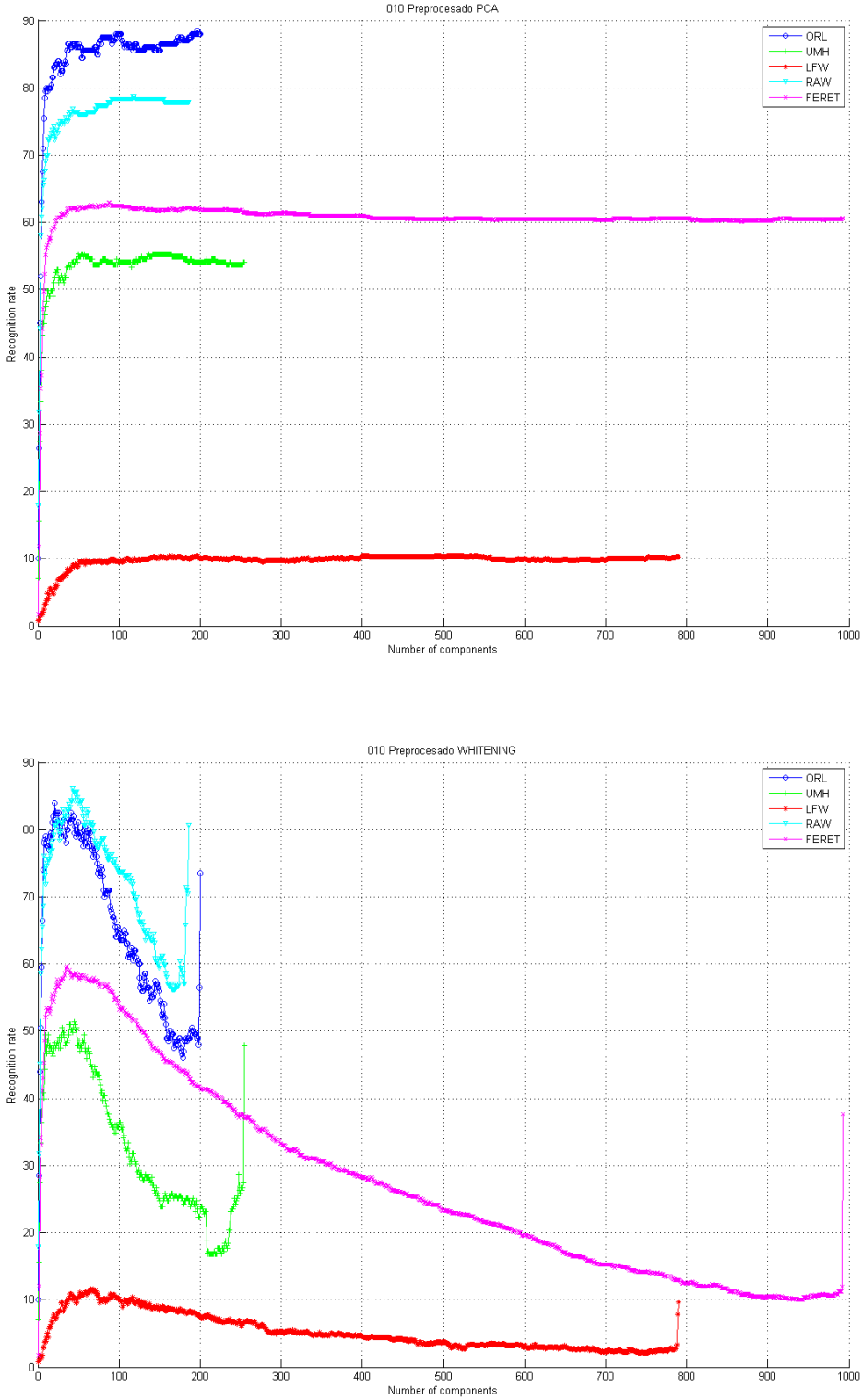


FIGURA 7.3.3: Resultados aunados para todas las BBDD en el preprocesado 010 por técnica, PCA o Whitening

		000	010	Mejora (%)
AR	PCA	63,59	**78,8**	23,91
	Whitening	76,95	**86,17**	11,98
ORL	PCA	90	88,5	-1,67
	Whitening	85,5	84	-1,75
UMH	PCA	58,43	55,29	-5,37
	Whitening	49,41	**51,37**	3,97
LFW	PCA	11,14	10,38	-5,37
	Whitening	11,27	**11,65**	3,97
FERET	PCA	61,19	**62,9**	2,8
	Whitening	53,83	**59,68**	10,86

TABLA 7.3.2: Comparación de los mejores resultados entre preprocesado 000 y 010 para las técnicas PCA y Whitening

Con el ecualizado del histograma, preprocesado 010, se obtiene una mejora en PCA para las bases de datos AR y FERET, mejorando en 2 de 5 de las bases de datos. Por el contrario, para Whitening la mejora se eleva a 4 de 5 bases de datos, mejorando para AR, UMH, LFW y FERET.

7.3.4. PREPROCESADO 011

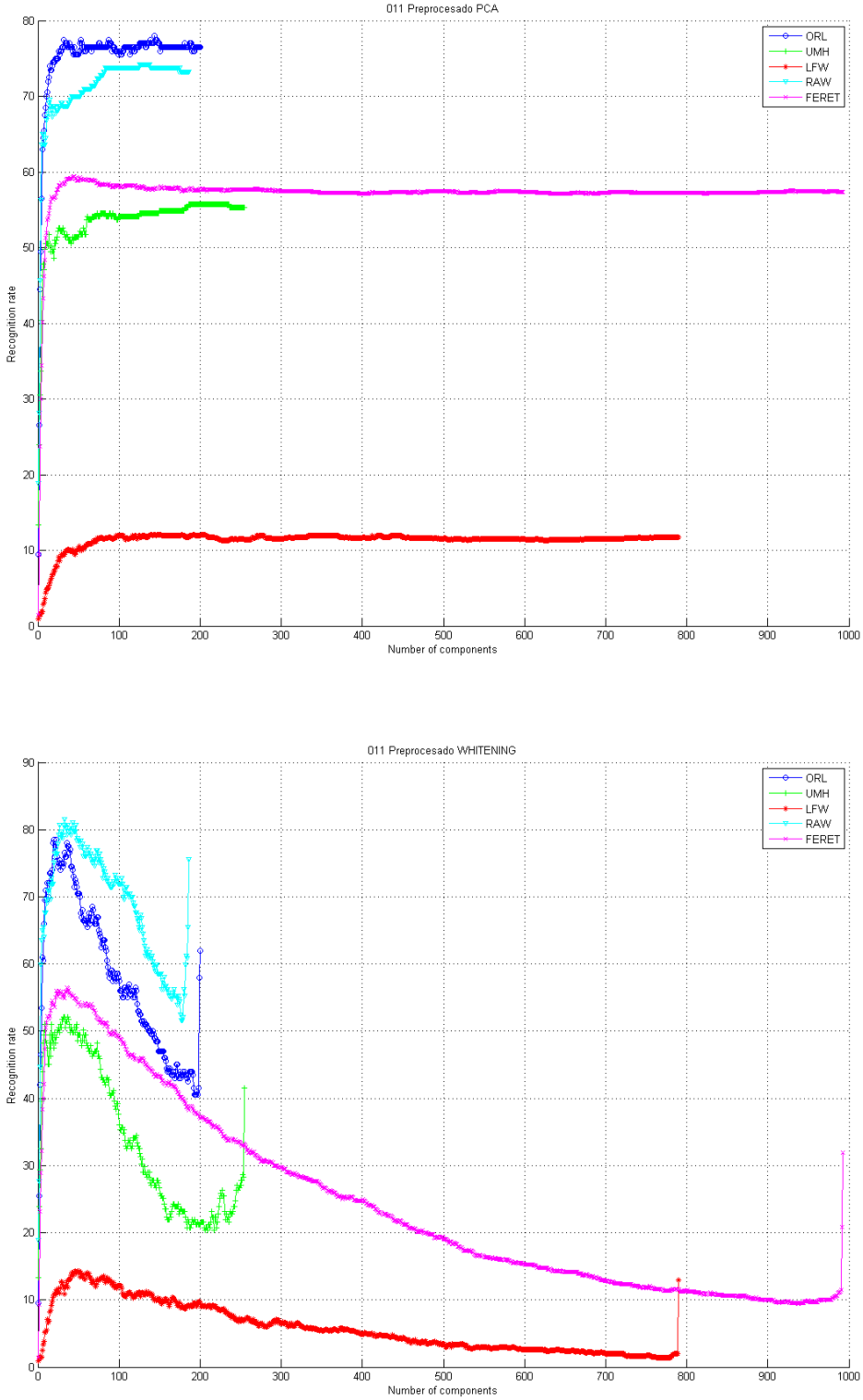


FIGURA 7.3.4: Resultados aunados para todas las BBDD en el preprocesado 011 por técnica, PCA o Whitening

		000	011	Mejora (%)
AR	PCA	63,59	**74,2**	16,67
	Whitening	76,95	**81,57**	5,99
ORL	PCA	90	78	-13,33
	Whitening	85,5	78,5	-8,19
UMH	PCA	58,43	55,69	-4,7
	Whitening	49,41	**52,16**	5,55
LFW	PCA	11,14	**12,02**	7,95
	Whitening	11,27	**14,3**	26,97
FERET	PCA	61,19	59,37	-2,96
	Whitening	53,83	**56,55**	5,06

TABLA 7.3.3: Comparación de los mejores resultados entre preprocesado 000 y 011 para las técnicas PCA y Whitening

En el preprocesado donde se realiza el ecualizado de histograma más la máscara elipsoidal, 011, mejora el porcentaje de reconocimiento respecto al preprocesado 000 en las bases de datos AR y LFW, esto es un total de mejora 2 de 5 en las BBDD. Para la técnica Whitening, este preprocesado mejora en las bases de datos AR, UMH, LFW y FERET, es decir, 4 de 5.

7.3.5. PREPROCESADO 020

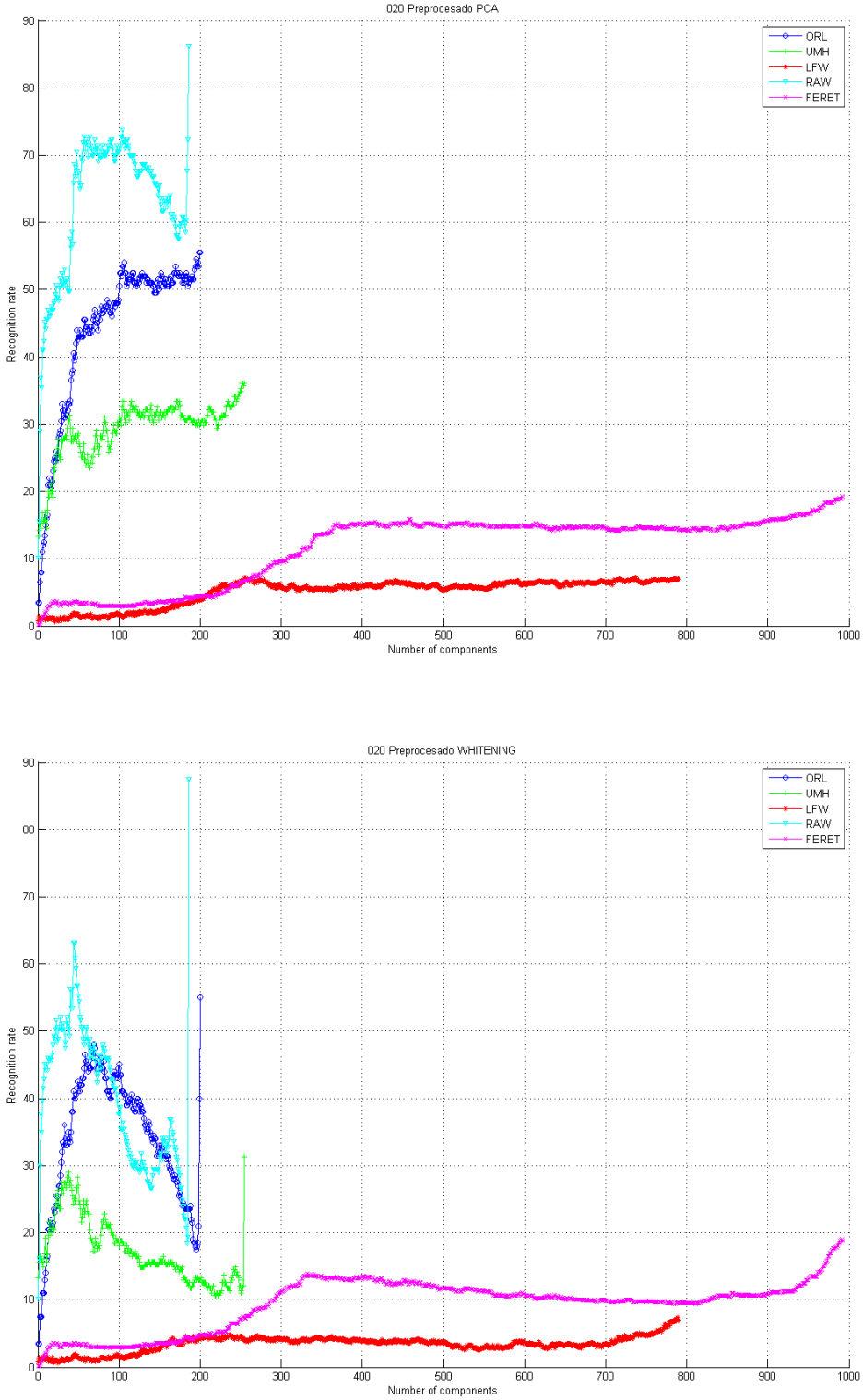


FIGURA 7.3.5: Resultados aunados para todas las BBDD en el preprocesado 020 por técnica, PCA o Whitening

		000	020	Mejora (%)
AR	PCA	63,59	**86,17**	35,51
	Whitening	76,96	**87,56**	13,77
ORL	PCA	90	55,5	-38,33
	Whitening	85,5	55	-35,67
UMH	PCA	58,43	36,08	-38,25
	Whitening	49,41	31,37	-36,51
LFW	PCA	11,14	7,09	-36,36
	Whitening	11,27	8,86	-21,35
FERET	PCA	61,19	19,25	-36,36
	Whitening	53,83	19,05	-21,35

TABLA 7.3.4: Comparación de los mejores resultados entre preprocesado 000 y 020 para las técnicas PCA y Whitening

En este preprocesado donde se ha aplicado el ecualizado Tan-Triggs, es uno de los que peores resultados ha obtenido. Los resultados obtenidos solo muestran una mejora tanto para PCA como para Whitening en la base de datos AR, con lo que solo se mejora en 1 de 5 en las BBDD.

7.3.6. PREPROCESADO 021

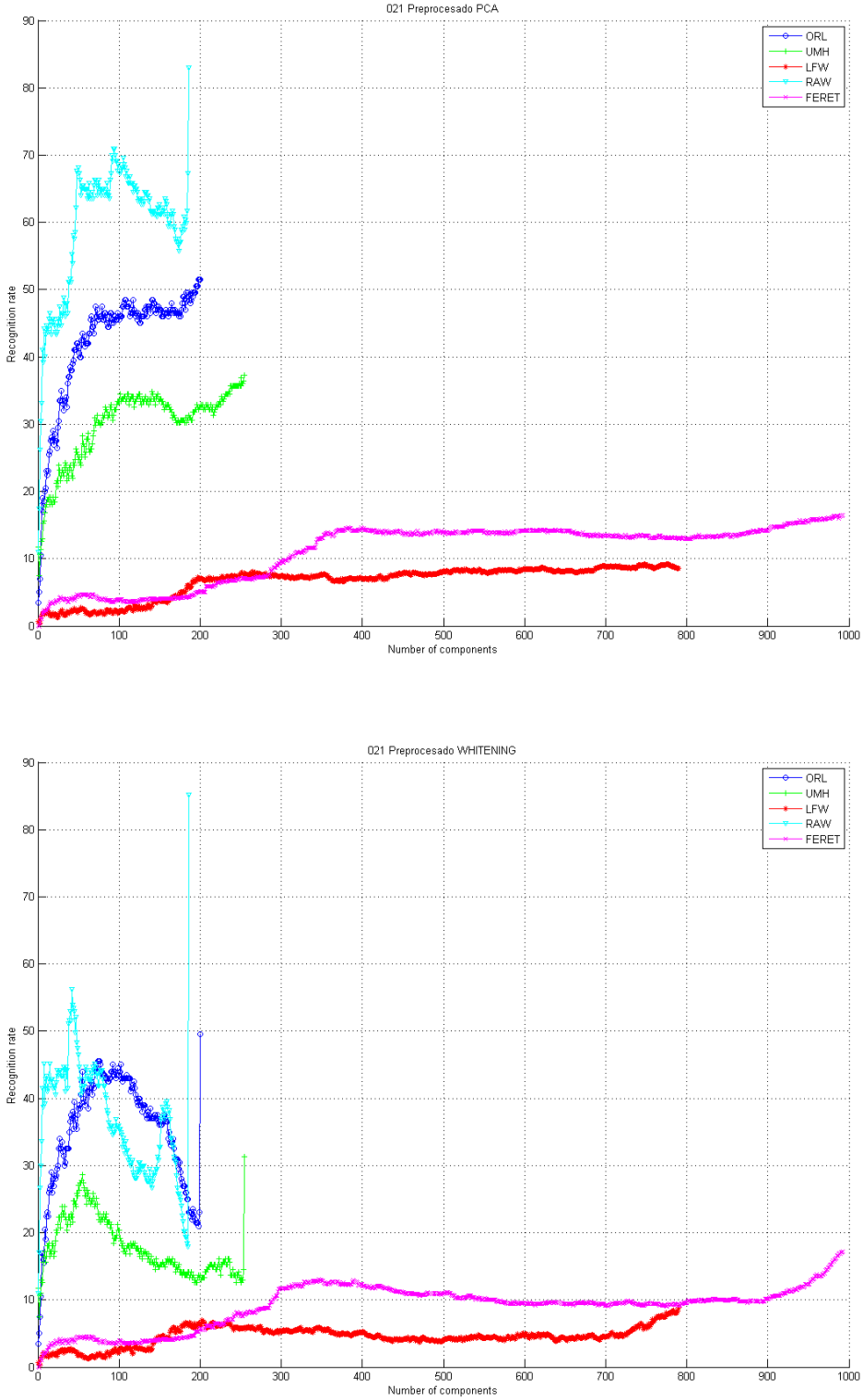


FIGURA 7.3.6: Resultados aunados para todas las BBDD en el preprocesado 021 por técnica, PCA o Whitening

		000	021	Mejora (%)
AR	PCA	63,59	**82,95**	30,43
	Whitening	76,96	**85,25**	10,78
ORL	PCA	90	51,5	-42,78
	Whitening	85,5	49,5	-42,1
UMH	PCA	58,43	37,25	-36,24
	Whitening	49,41	31,37	-36,51
LFW	PCA	11,14	9,24	-17,04
	Whitening	11,27	8,86	-21,35
FERET	PCA	61,19	16,43	-73,15
	Whitening	53,83	17,24	-67,98

TABLA 7.3.5: Comparación de los mejores resultados entre preprocesado 000 y 021 para las técnicas PCA y Whitening

Observando los resultados mostrados para el preprocesado donde se ha aplicado el ecualizado Tan-Triggs más la máscara elipsoidal, se observa, como los resultados que se obtienen son muy parecidos al caso anterior, donde solo existe mejora respecto al original en la base de datos AR, tanto para la técnica de PCA como para Whitening.

7.3.7. PREPROCESADO 100

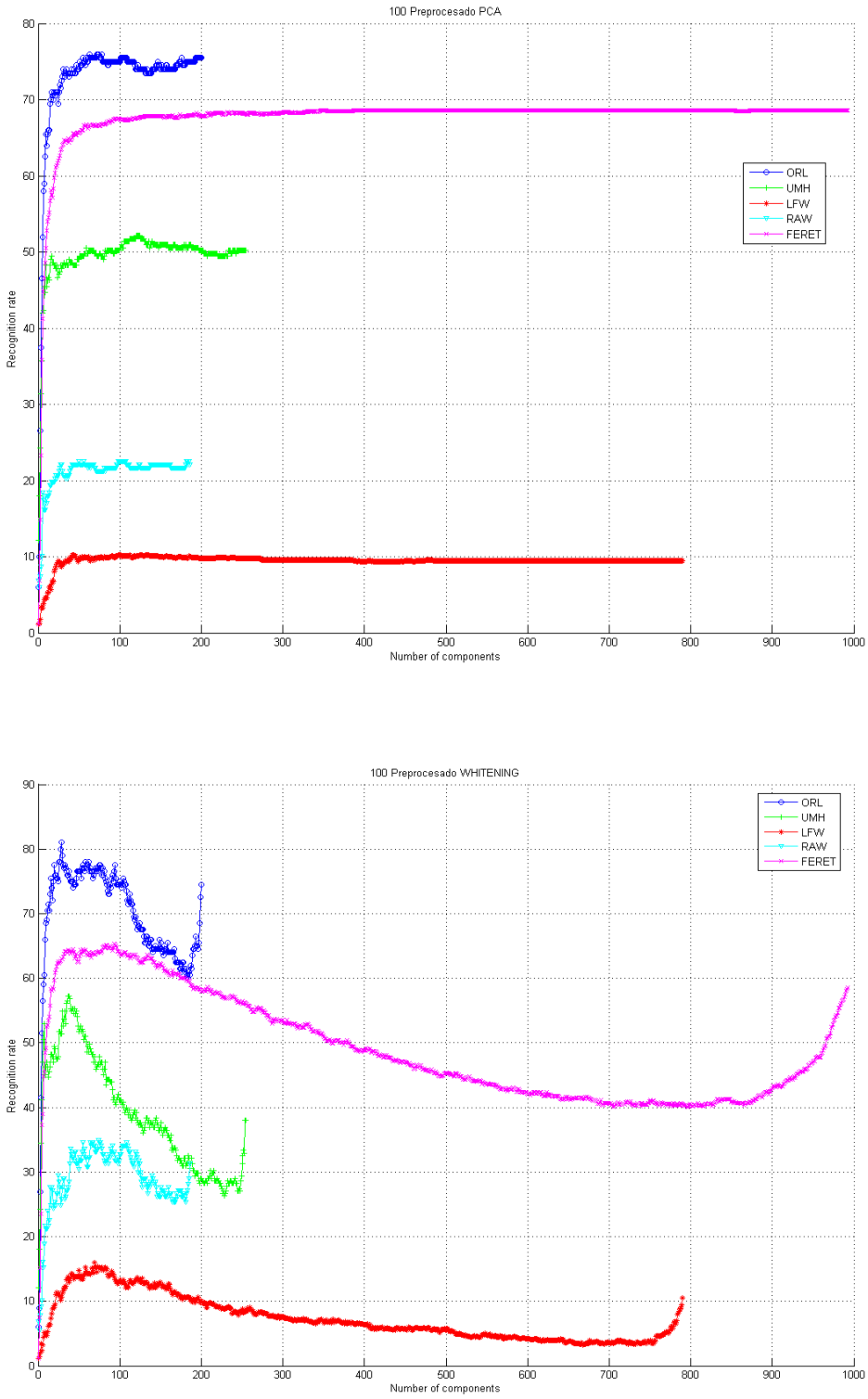


FIGURA 7.3.7: Resultados aunados para todas las BBDD en el preprocesado 100 por técnica, PCA o Whitening

		000	100	Mejora (%)
AR	PCA	63,59	22,58	-64,49
	Whitening	76,96	35,02	-54,49
ORL	PCA	90	76	-15,56
	Whitening	85,5	81	-5,26
UMH	PCA	58,43	52,16	-10,74
	Whitening	49,41	**57,25**	15,87
LFW	PCA	11,14	10,25	-7,95
	Whitening	11,27	**15,95**	41,57
FERET	PCA	61,19	**68,55**	12,03
	Whitening	53,83	**65,32**	21,35

TABLA 7.3.6: Comparación de los mejores resultados entre preprocesado 000 y 021 para las técnicas PCA y Whitening

Analizando los datos obtenidos para el preprocesado donde únicamente se ha aplicado Viola-Jones, se observa como para PCA únicamente existe una mejora en la base de datos FERET, haciendo que la mejora sea en 1 de 5 respecto del total de las BBDD. Para Whitening, se observa como mejora en 2 más respecto a PCA, sumándose la base de datos LFW y UMH a la de PCA, haciendo que la mejora global respecto al total de las BBDD sea de 3 de 5.

7.3.8. PREPROCESADO 101

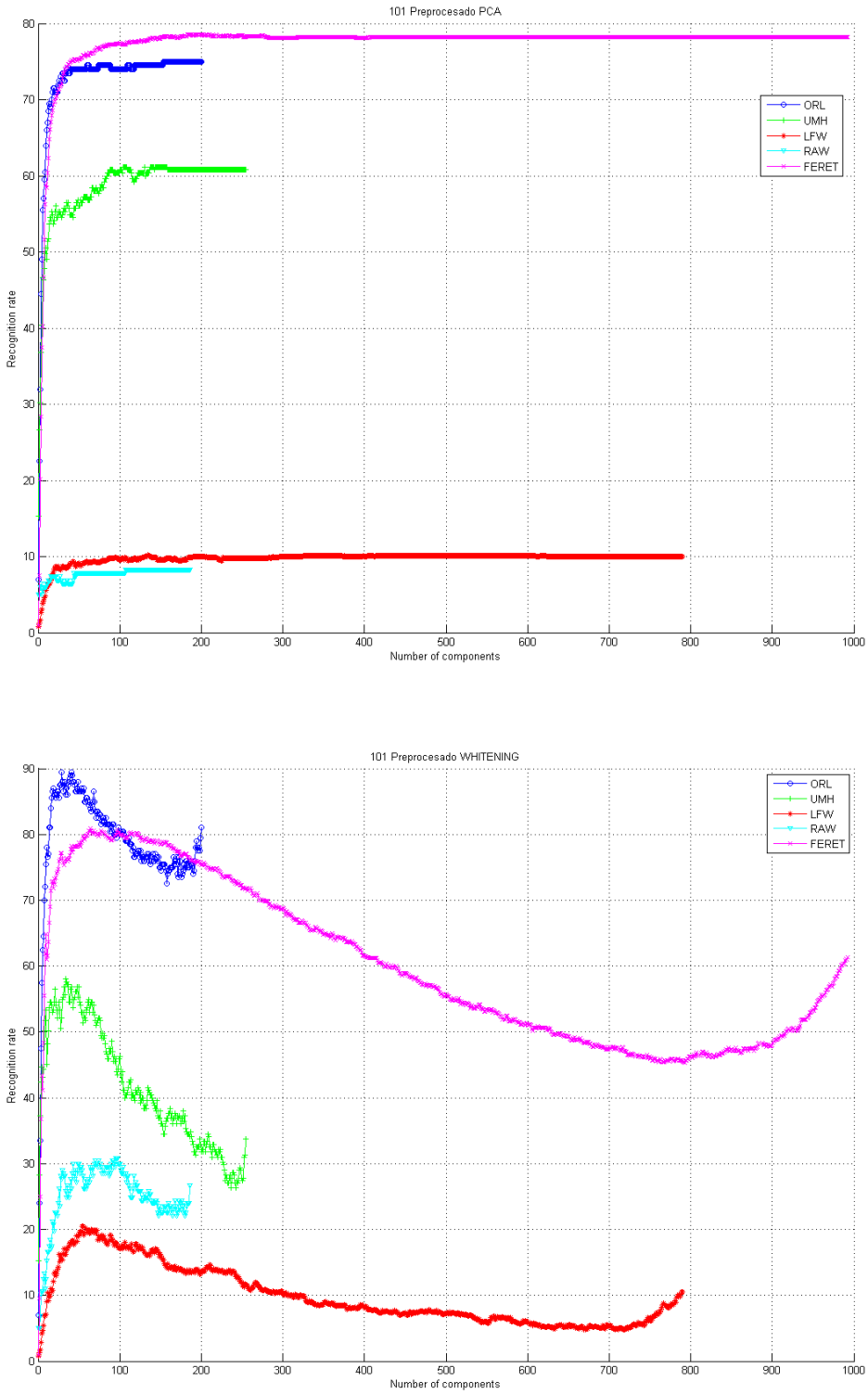


FIGURA 7.3.8: Resultados aunados para todas las BBDD en el preprocesado 101 por técnica, PCA o Whitening

		000	101	Mejora (%)
AR	PCA	63,59	8,29	-86,96
	Whitening	76,96	30,88	-59,88
ORL	PCA	90	75	-16,67
	Whitening	85,5	**89,5**	4,68
UMH	PCA	58,43	**61,18**	4,7
	Whitening	49,41	**58,04**	17,46
LFW	PCA	11,14	10,25	-7,95
	Whitening	11,27	**20,51**	82,02
FERET	PCA	61,19	**78,53**	28,34
	Whitening	53,83	**80,95**	50,37

TABLA 7.3.7: Comparación de los mejores resultados entre preprocesado 000 y 101 para las técnicas PCA y Whitening

Para el preprocesado 101, donde se ha aplicado Viola-Jones más la máscara elipsoidal, se observa como los resultados mejoran con respecto al anterior preprocesado, donde solo se aplicaba el detector Viola-Jones. En este caso para PCA, se observan mejoras en las bases de datos UMH y FERET, por lo que mejora en 2 de las 5 bases de datos.

Para el caso de Whitening, la mejora existe en las bases de datos ORL, UMH, LFW y FERET, mejorando en 4 de las 5 bases de datos posibles.

7.3.9. PREPROCESADO 110

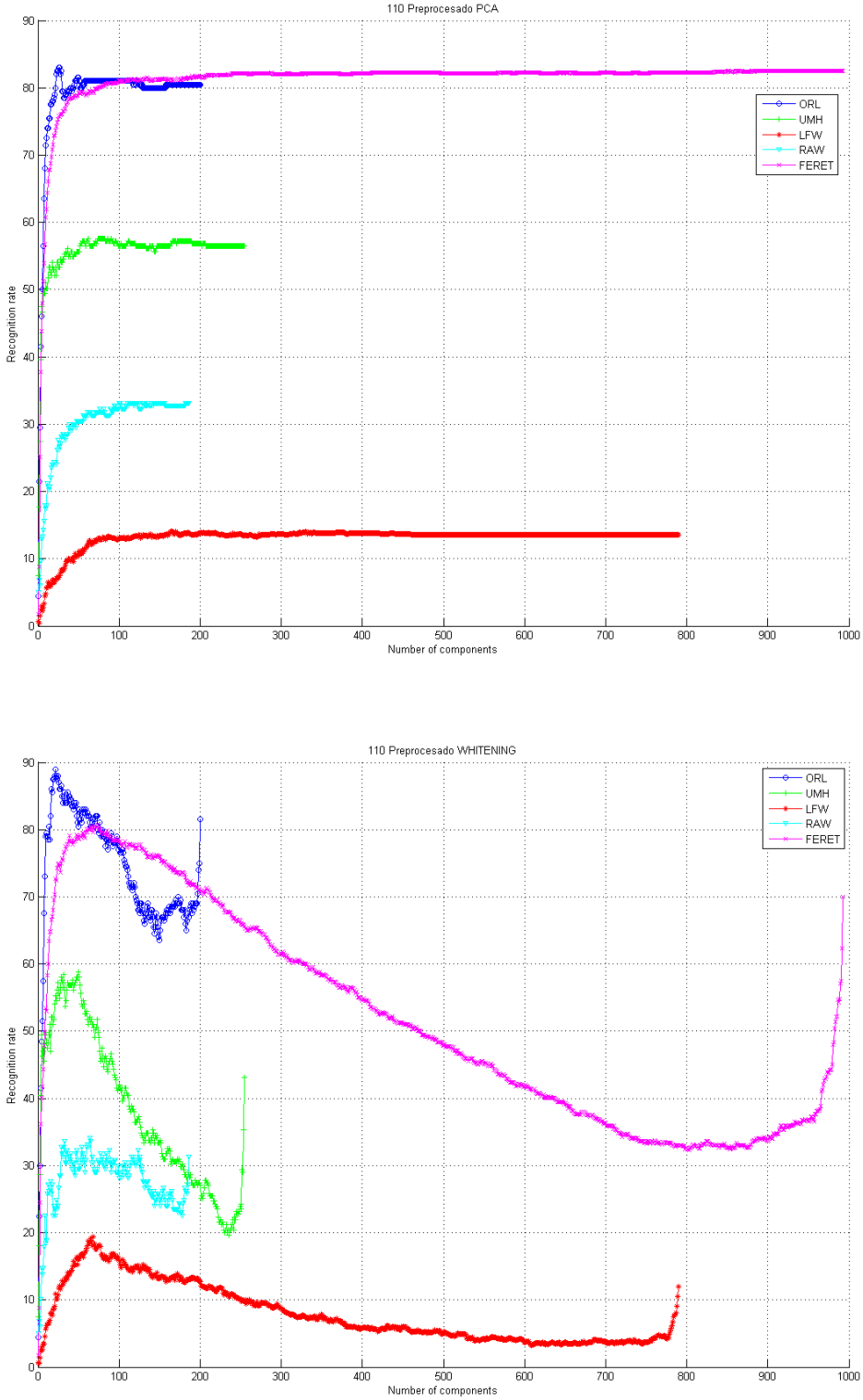


FIGURA 7.3.9: Resultados aunados para todas las BBDD en el preprocesado 110 por técnica, PCA o Whitening

		000	110	Mejora (%)
AR	PCA	63,59	33,18	-47,83
	Whitening	76,96	34,1	-55,69
ORL	PCA	90	83	-7,78
	Whitening	85,5	**89**	4,09
UMH	PCA	58,43	57,65	-1,342
	Whitening	49,41	**58,82**	19,05
LFW	PCA	11,14	**14,05**	26,14
	Whitening	11,27	**19,37**	71,91
FERET	PCA	61,19	**82,46**	34,76
	Whitening	53,83	**80,64**	49,81

TABLA 7.3.8: Comparación de los mejores resultados entre preprocesado 000 y 110 para las técnicas PCA y Whitening

En el caso del preprocesado 110, donde aparte de Viola-Jones se aplica el ecualizado de histograma, se observan unos resultados similares al del preprocesado 101.

En el caso de PCA, la mejora en el porcentaje de reconocimiento, comparándola con las imágenes originales, existe en las bases de datos LFW y FERET (2 de 5). Ahora, si se comparan para el caso Whitening, la mejora es la misma que en el preprocesado 101, mejorando en las mismas bases de datos, ORL, UMH, LFW y FERET (4 de 5).

7.3.10. PREPROCESADO 111

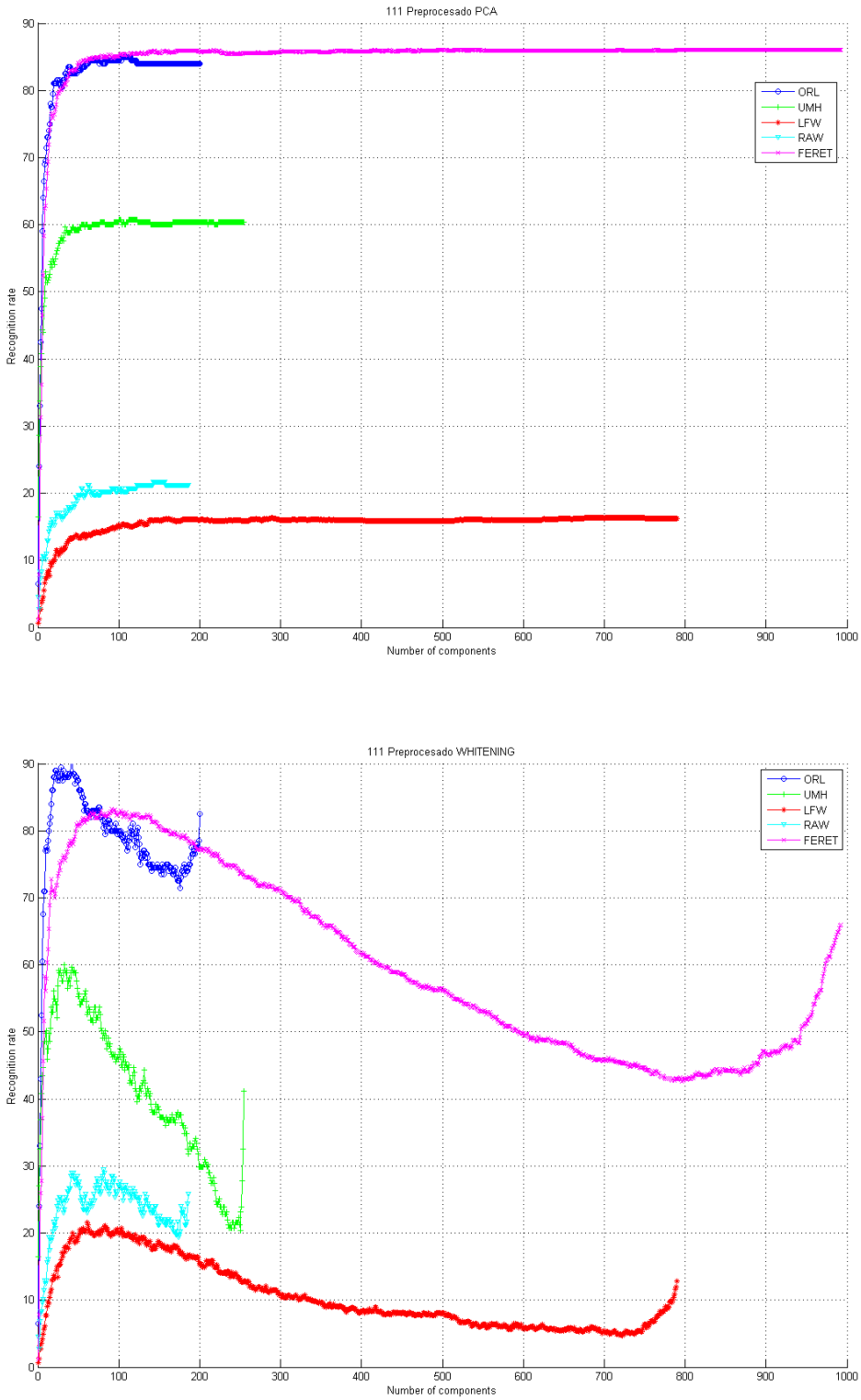


FIGURA 7.3.10: Resultados aunados para todas las BBDD en el preprocesado 111 por técnica, PCA o Whitening

		000	111	Mejora (%)
AR	PCA	63,59	21,66	-65,94
	Whitening	76,96	29,49	-61,68
ORL	PCA	90	85	-5,56
	Whitening	85,5	**90**	5,26
UMH	PCA	58,43	**60,78**	4,03
	Whitening	49,41	**60**	21,43
LFW	PCA	11,14	**16,33**	46,59
	Whitening	11,27	**21,65**	92,13
FERET	PCA	61,19	**85,99**	40,53
	Whitening	53,83	**83,27**	54,68

TABLA 7.3.9: Comparación de los mejores resultados entre preprocesado 000 y 111 para las técnicas PCA y Whitening

Los resultados que se obtienen en el preprocesado 111 son, como se podrá leer posteriormente, los mejores en cuanto a aumento de porcentaje de reconocimiento respecto a las imágenes originales sin preprocesar.

Observando ahora en qué bases se ha realizado la mejora respecto a las distintas técnicas aplicadas, se observa para PCA que existe una mejora en 3 de las 5 bases de datos estudiadas, específicamente en la base de datos UMH, LFW y FERET. Para Whitening, la mejora existe en 4 de las 5 bases estudiadas, ORL, UMH, LFW y FERET.

7.3.11. PREPROCESADO 120

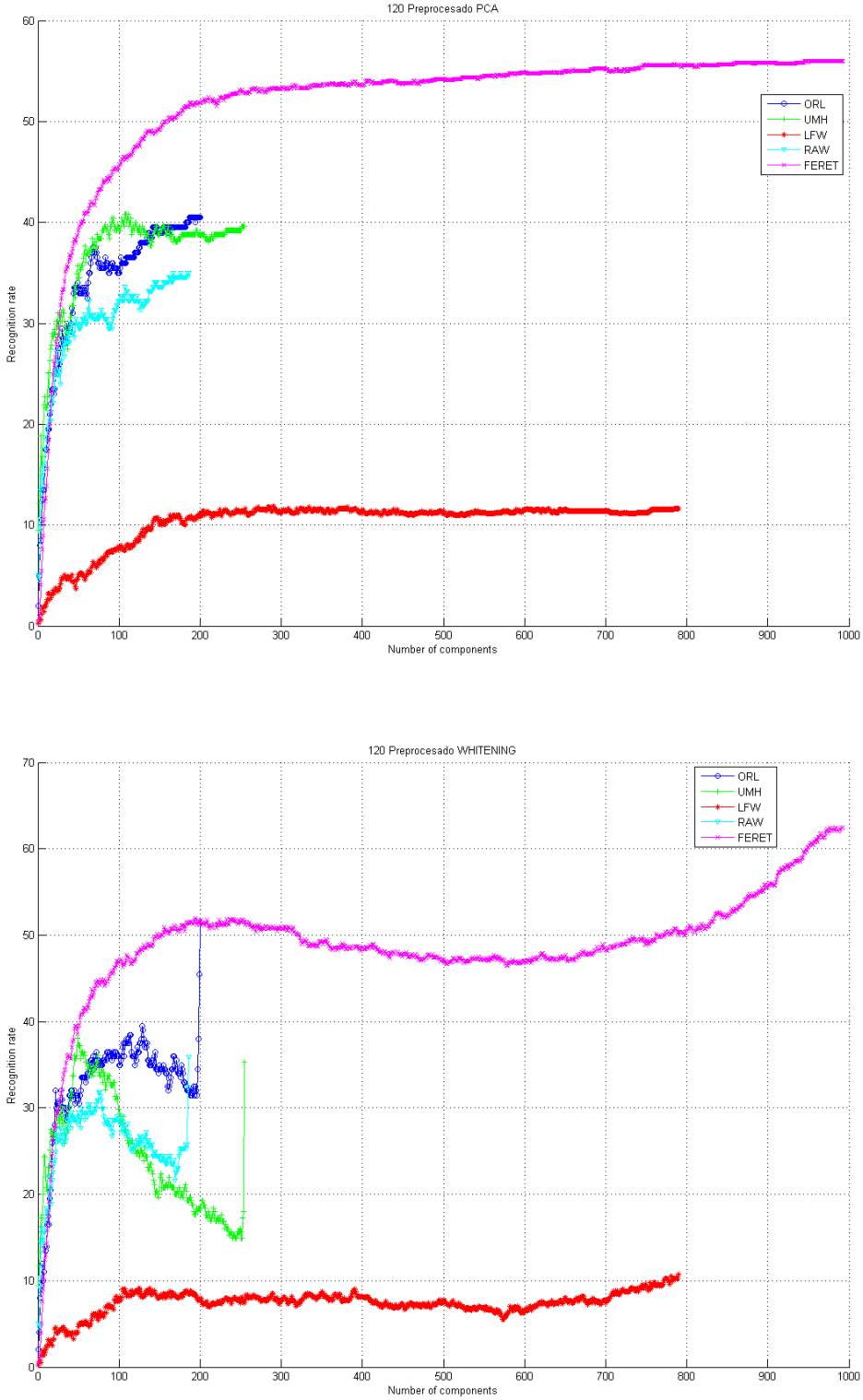


FIGURA 7.3.11: Resultados aunados para todas las BBDD en el preprocesado 120 por técnica, PCA o Whitening

		000	120	Mejora (%)
AR	PCA	63,59	35,02	-44,93
	Whitening	76,96	35,94	-53,29
ORL	PCA	90	40,5	-55
	Whitening	85,5	51,5	-39,77
UMH	PCA	58,43	40,78	-30,2
	Whitening	49,41	38,04	-23,02
LFW	PCA	11,14	**11,9**	6,82
	Whitening	11,27	10,76	-4,49
FERET	PCA	61,19	55,95	-8,57
	Whitening	53,83	**62,5**	16,1

TABLA 7.3.10: Comparación de los mejores resultados entre preprocesado 000 y 120 para las técnicas PCA y Whitening

Cuando al detector Viola-Jones se suman ambos ecualizados, histograma y Tan-Triggs, los resultados que se obtienen cuando se comparan respecto a los originales, en el caso de PCA, es de no mejoría, sino un empeoramiento en los porcentajes de reconocimiento, exceptuando para la base de datos LFW (1 de 5) donde existe una leve mejoría respecto al original. Para Whitening, la mejora solo existe en la base de datos FERET (1 de 5).

7.3.12. PREPROCESADO 121

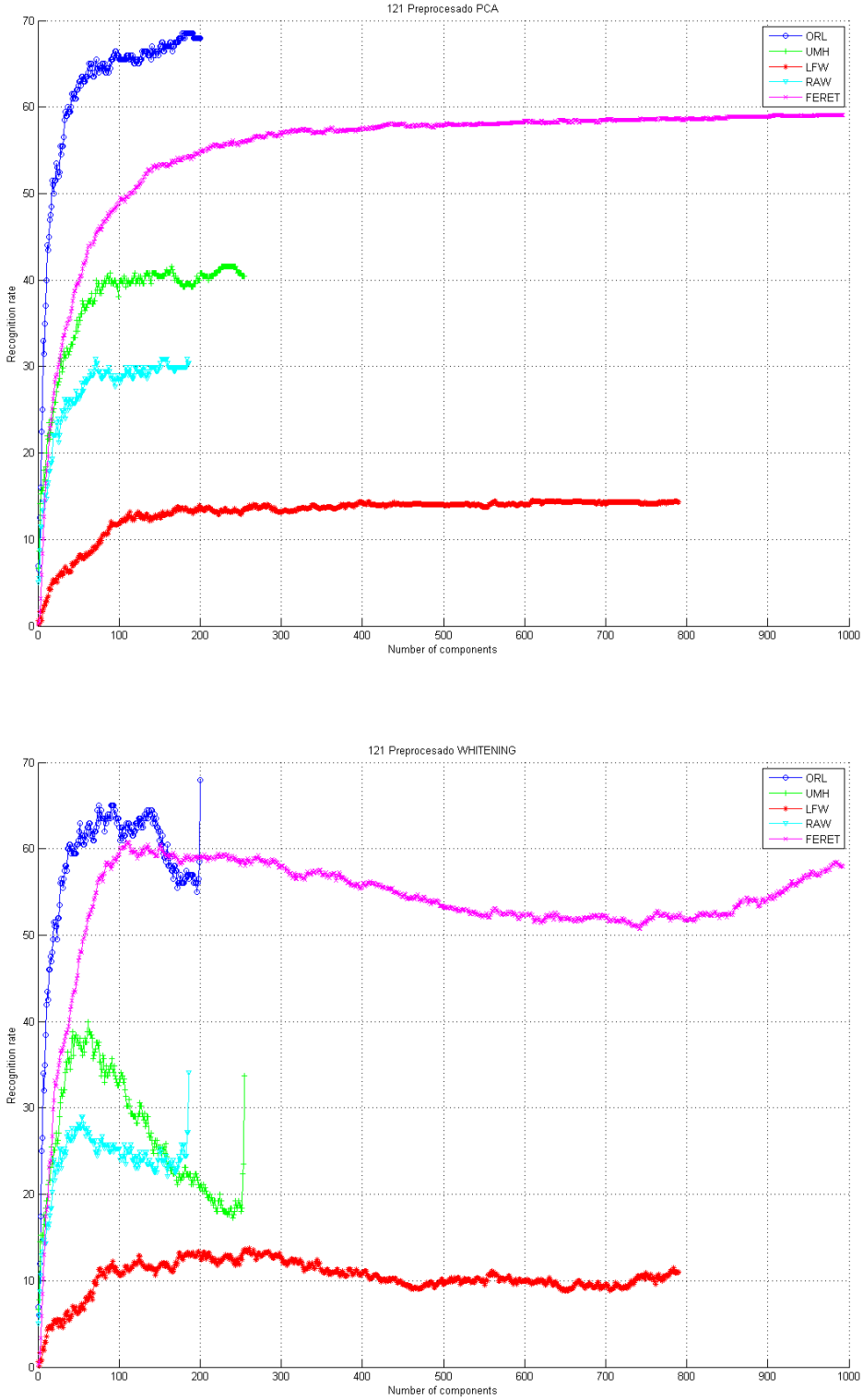


FIGURA 7.3.12: Resultados aunados para todas las BBDD en el preprocesado 121 por técnica, PCA o Whitening

		000	121	Mejora (%)
AR	PCA	63,59	30,88	-51,45
	Whitening	76,96	34,1	-55,69
ORL	PCA	90	68,5	-23,89
	Whitening	85,5	68	-20,47
UMH	PCA	58,43	41,57	-28,86
	Whitening	49,41	40	-19,05
LFW	PCA	11,14	**14,56**	30,68
	Whitening	11,27	**13,8**	22,47
FERET	PCA	61,19	59,07	-3,46
	Whitening	53,83	**60,79**	12,92

TABLA 7.3.11: Comparación de los mejores resultados entre preprocesado 000 y 121 para las técnicas PCA y Whitening

Los resultados para PCA en el último preprocesado aplicado, mejoran únicamente en una base de datos LFW (1 de 5). En el caso de Whitening, se mejora en una BBDD más respecto a PCA, sumando la base de datos FERET a la de LFW (2 de 5)

7.3.13. RESUMEN EXPERIMENTO 3

RESUMEN RESULTADOS PARA PCA

Preprocesados	Mejora	Empeora
000	--	--
001	LFW	UMH, FERET, ORL, AR
010	AR y FERET	ORL, UMH y LFW
011	AR y LFW	ORL, UMH y FERET
020	AR	AR, ORL, UMH y LFW
021	AR	ORL, UMH, LFW y FERET
100	FERET	AR, ORL, UMH y LFW
101	UMH y FERET	AR, ORL, LFW
110	LFW y FERET	AR, ORL y UMH
111	UMH, LFW y FERET	AR y ORL
120	LFW	UMH, FERET, ORL, AR
121	LFW	UMH, FERET, ORL, AR

TABLA 7.4.1: Resultados de todos los preprocesados en todas las BBDD agrupadas en dos grupos, mejora o empeora la base de datos X en el preprocesado Y para la técnica PCA.

RESUMEN RESULTADOS PARA WHITENING

Preprocesados	Mejora	Empeora
000	--	--
001	AR, UMH y LFW	ORL y FERET
010	AR, UMH, LFW y FERET	ORL
011	AR, UMH, LFW y FERET	ORL
020	AR	LFW, UMH, ORL, FERET
021	AR	LFW, UMH, ORL, FERET
100	FERET, LFW y UMH	ORL, AR
101	ORL, UMH, LFW y FERET	AR
110	ORL, UMH, LFW y FERET	AR
111	ORL, UMH, LFW y FERET	AR
120	FERET	AR, ORL, UMH y LFW
121	LFW, FERET	UMH, ORL, AR

TABLA 7.4.2: Resultados de todos los preprocesados en todas las BBDD agrupadas en dos grupos, mejora o empeora la base de datos X en el preprocesado Y para la técnica WHITENING.

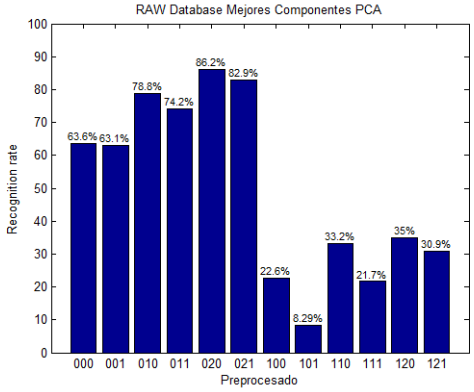
Observando la tabla donde se resumen las mejoras y no mejoras respecto a las imágenes sin realizar ningún preprocesado, se concluye que el mejor preprocesado para PCA es el preprocesado 111, ya que mejora en 3 de las 5 bases de datos estudiadas.

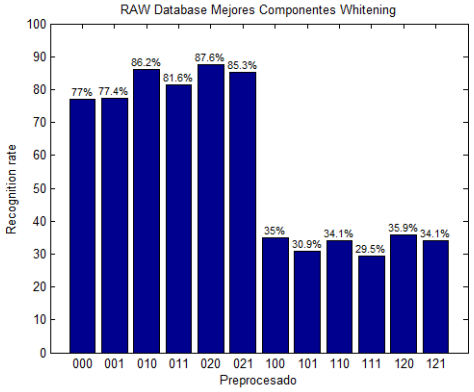
Para Whitening, está conclusión no es tan sencilla de formular, ya que existen 5 preprocesados donde se mejora, en 4 de las 5 bases de datos, el porcentaje de reconocimiento: 010, 011, 101, 110 y 111. Para ello, se debe observar los resultados obtenidos en los anteriores experimentos.

7.4. EXPERIMENTO 4

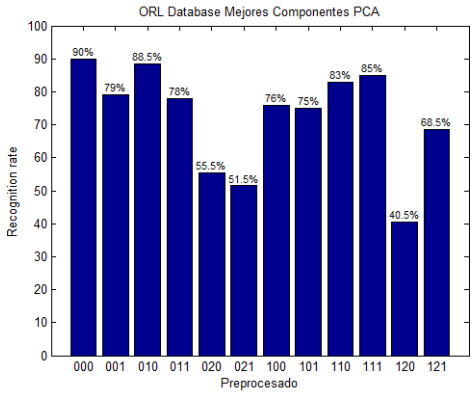
7.4.1. RATIO DE RECONOCIMIENTO POR BASE DE DATOS Y PREPROCESADO

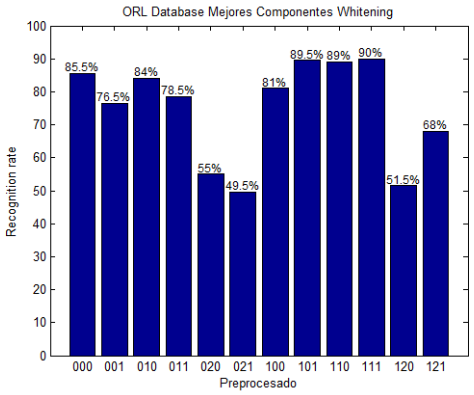
- Base de datos AR:



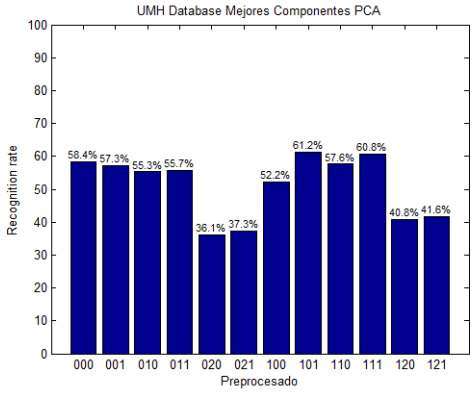


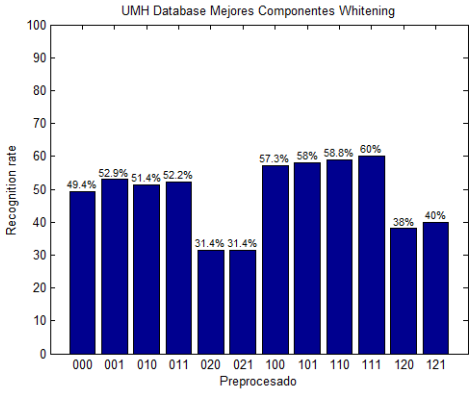
- Base de datos ORL:



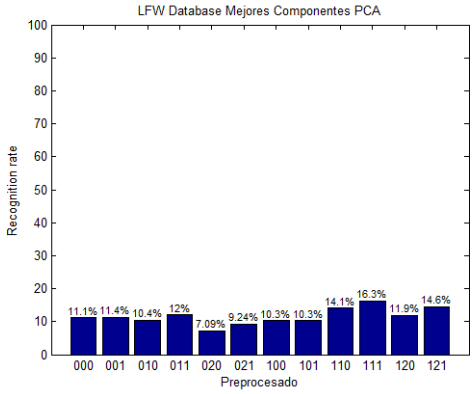


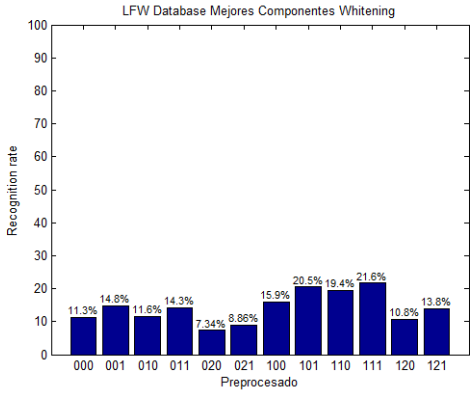
- Base de datos UMH:



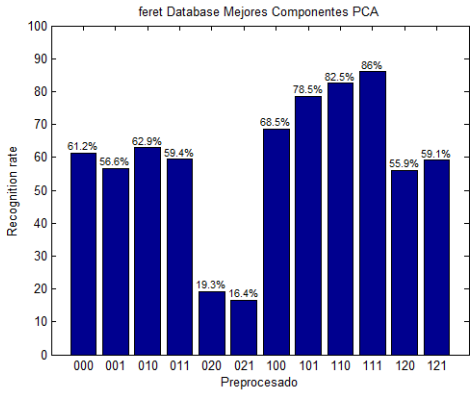


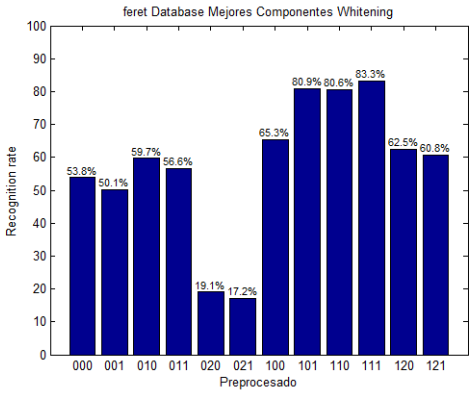
- Base de datos LFW:





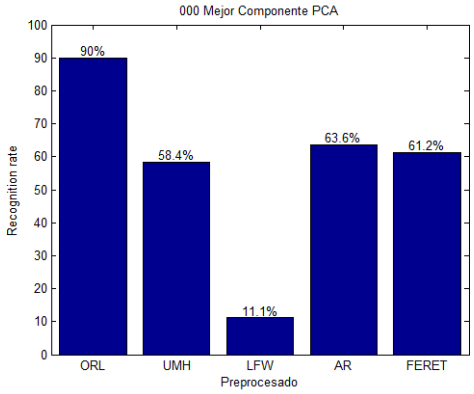
- Base de datos FERET:

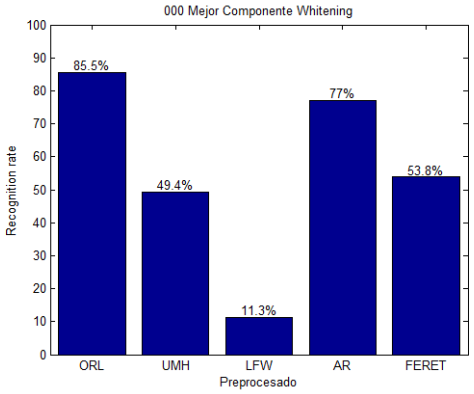




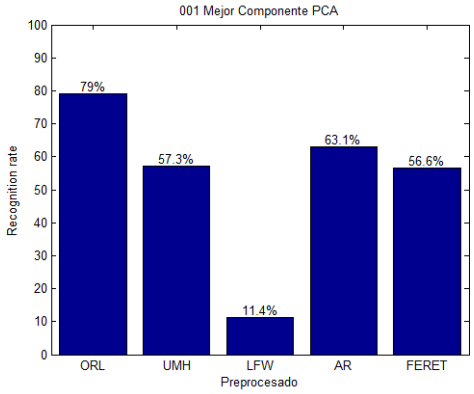
7.4.2. COMPARATIVA DE PREPROCESADOS POR BASE DE DATOS

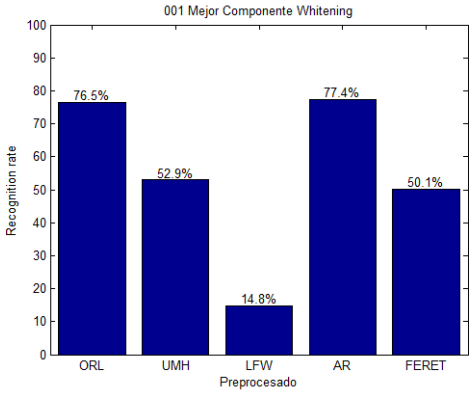
- Preprocesado 000:



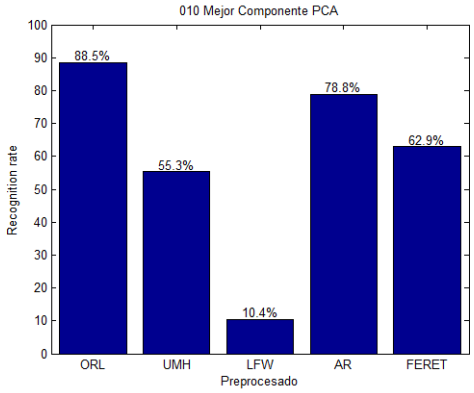


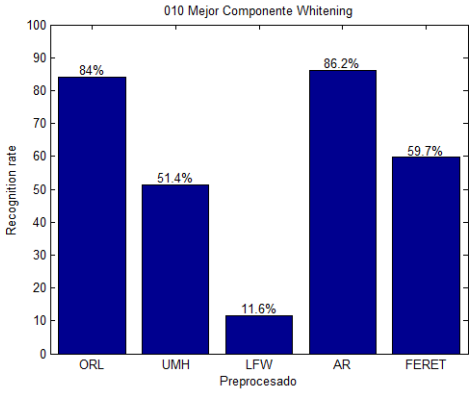
- Preprocesado 001:



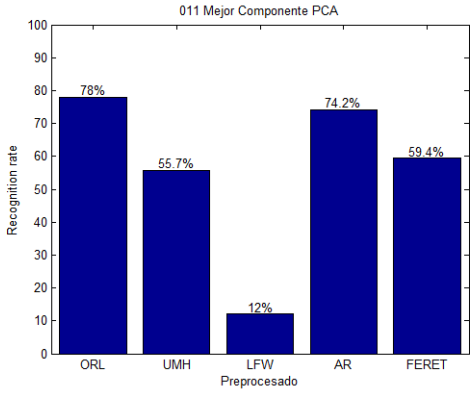


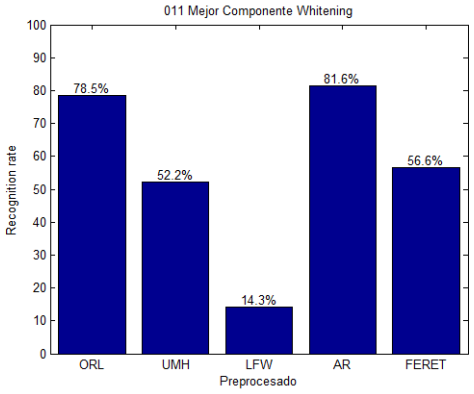
- Preprocesado 010:



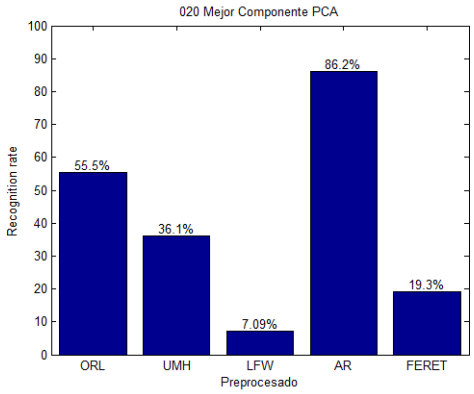


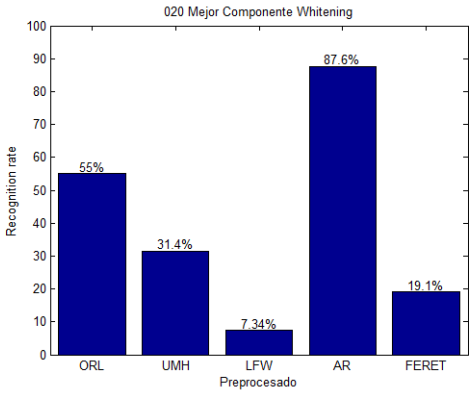
- Preprocesado 011:



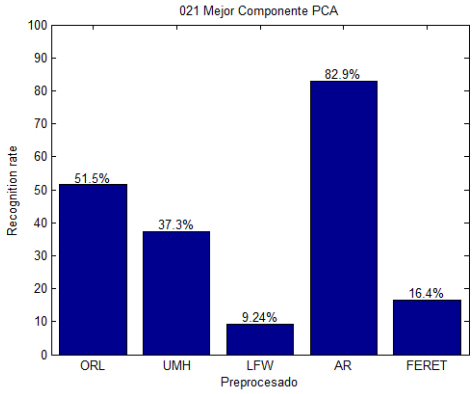


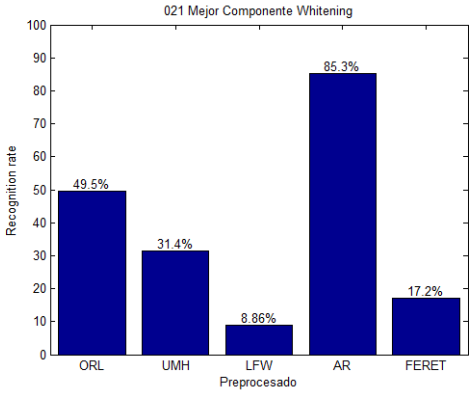
- Preprocesado 020:



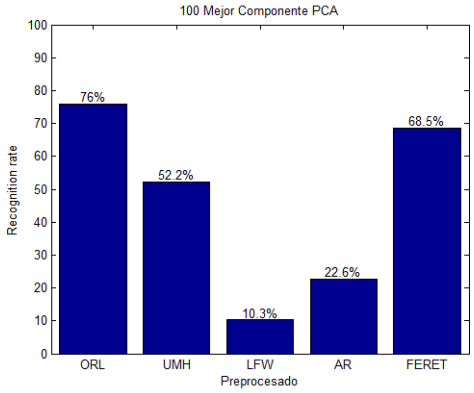


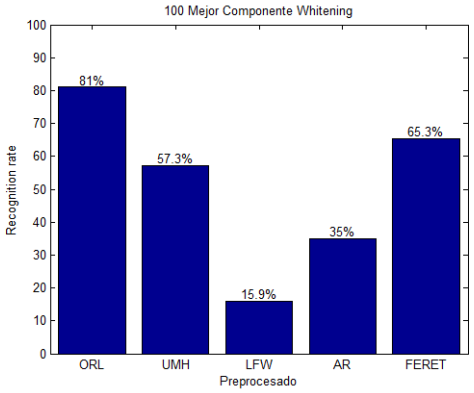
- Preprocesado 021:



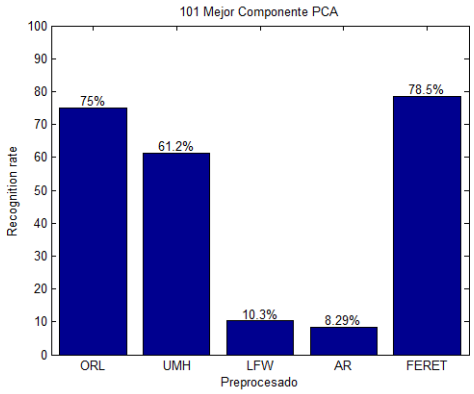


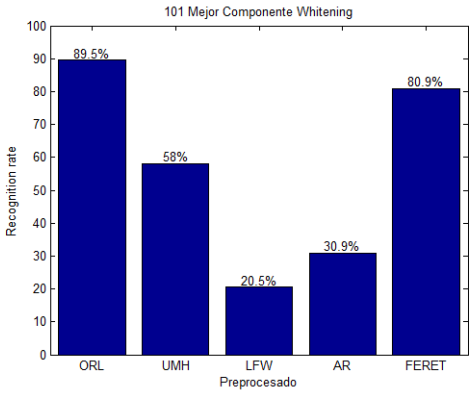
- Preprocesado 100:



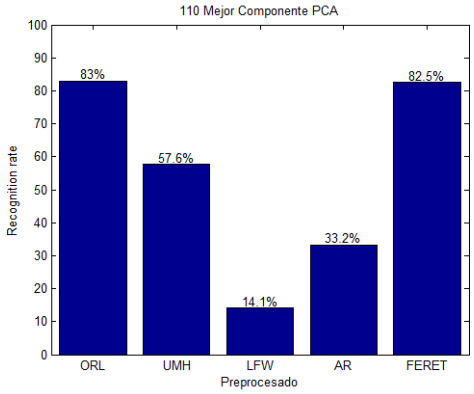


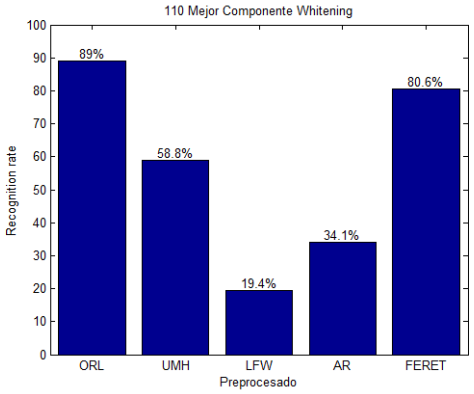
- Preprocesado 101:



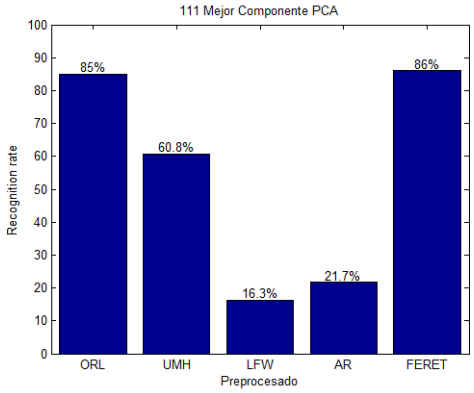


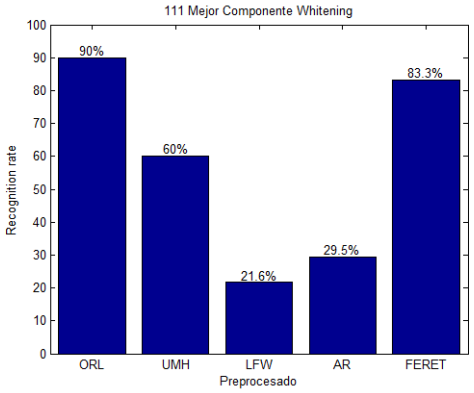
- Preprocesado 110:



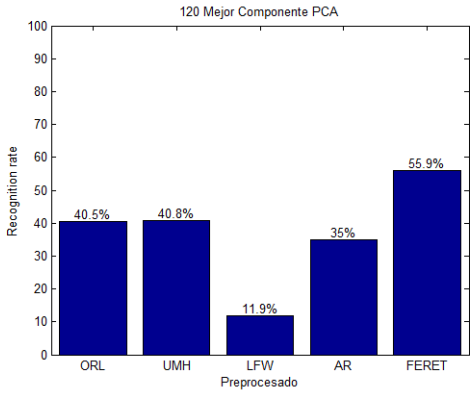


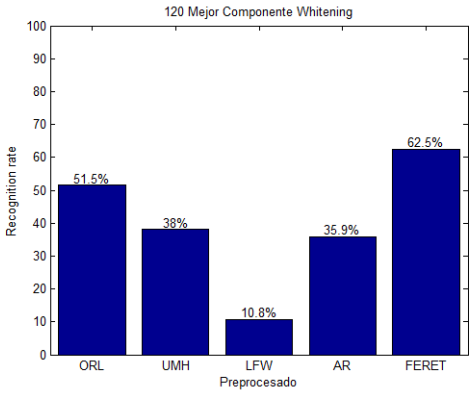
- Preprocesado 111:



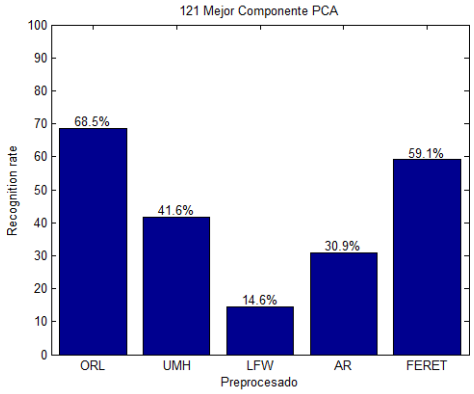


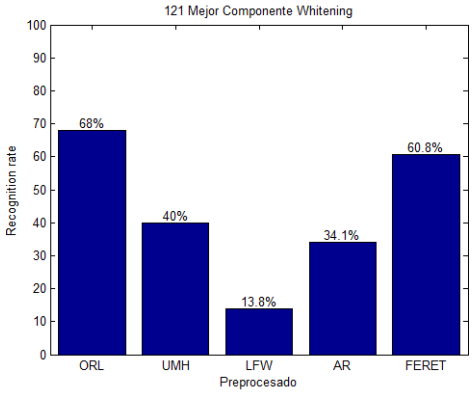
- Peprocesado 120:



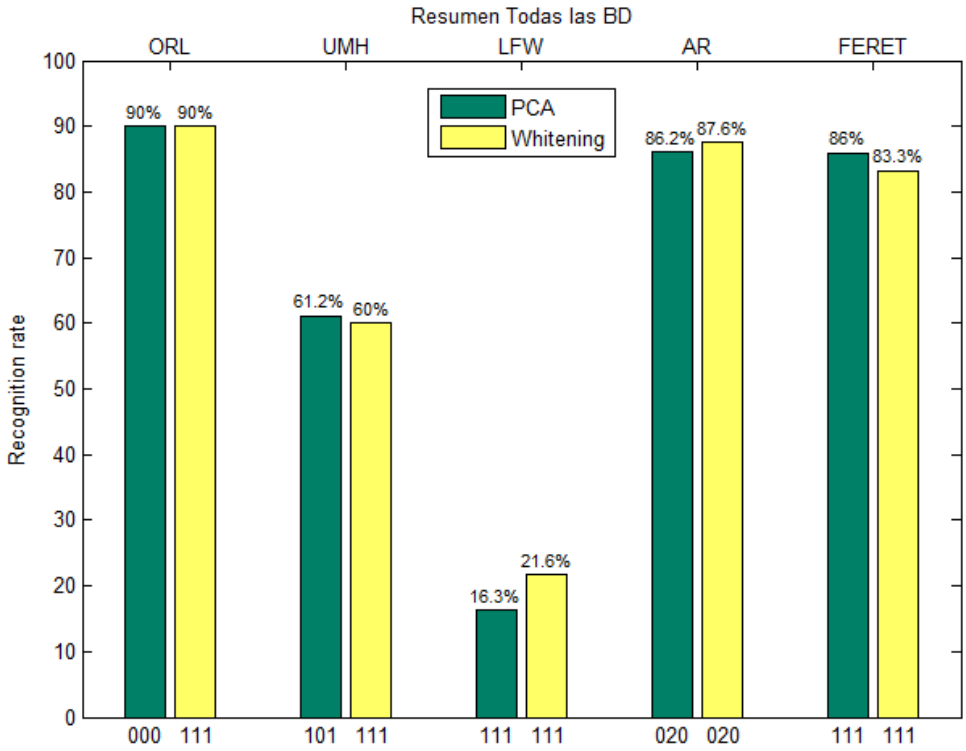


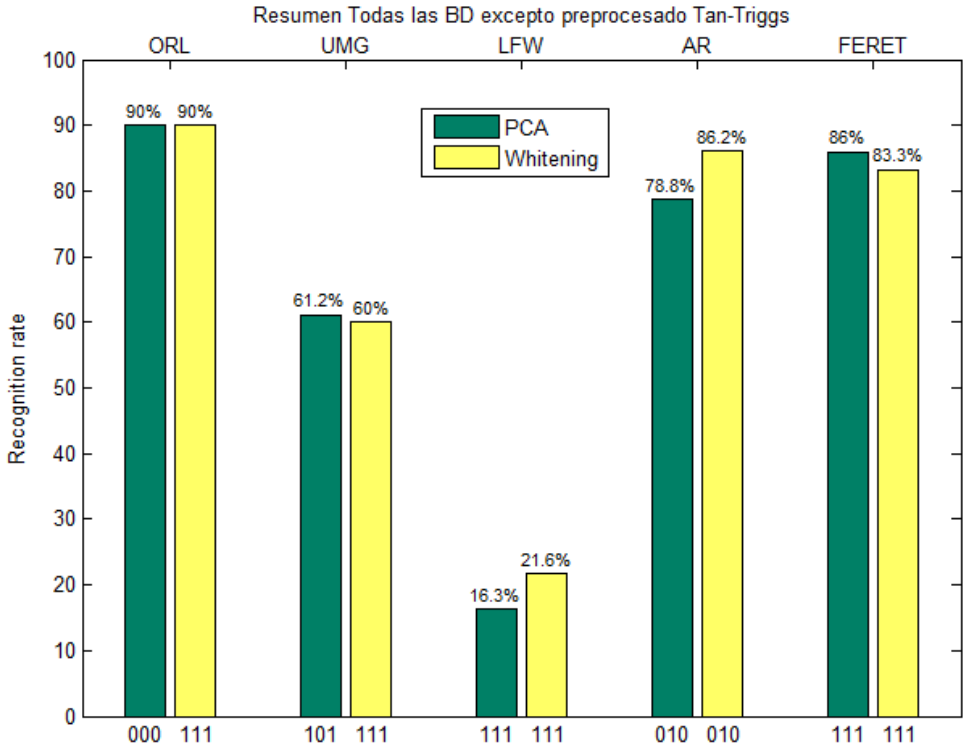
- Preprocesado 121:





7.4.3. MEJOR PREPROCESADO POR BASE DE DATOS





8. CONCLUSIONES

La finalidad de este proyecto es la de poder indicar cuál de las dos técnicas aplicadas en las bases de datos estudiadas es la mejor dependiendo de los resultados obtenidos, así como también saber cuál es el preprocesado idóneo a utilizar, independientemente de la base de datos utilizada.

Una vez analizados los experimentos, se pueden obtener las conclusiones finales que son el motivo final del presente estudio.

Si se entra en detalle a estudiar cuál de las dos técnicas es la idónea a utilizar sin aplicar ningún tipo de preprocesado, observando la tabla 8.1, se aprecia como con la técnica de PCA se obtienen mejores resultados, respecto de Whitening, en 3 de las 5 bases de datos estudiadas. En la base de datos AR, la técnica PCA comparándose con Whitening es un 17.34% peor, en ORL es un 5.26% mejor, para UMH un 18.25% mejor, en LFW es un 1.12% peor, finalizando con la BBDD FERET con un porcentaje de reconocimiento de un 13.67% mejor.

BBDD	TÉCNICA	RECONOCIMIENTO (%)
AR	PCA	63,59
	Whitening	76,96
ORL	PCA	90
	Whitening	85,5
UMH	PCA	58,43
	Whitening	49,41
LFW	PCA	11,14
	Whitening	11,27
FERET	PCA	61,19
	Whitening	53,83

TABLA 8.1: Resultados de PCA y WHITENING para todas las BBDD sin preprocesados.

Partiendo de los resultados obtenidos en el Experimento 1 donde se muestran todos los resultados en ambas técnicas por bases de datos y preprocesados, se observa como Whitening mejora en 32 de los 60 casos posibles respecto a los resultados de PCA. De estos 32 casos de mejora, 12 de ellos son sobre una misma base de datos, BBDD AR, esto es el 37.5% del total de casos mejorados en Whitening. Para PCA, en la BBDD UMH, es mejor en 10 de los 12 casos, que representa un 35.71% de los 28 casos en los que esta técnica obtiene mejores resultados respecto de la otra técnica estudiada.

Se puede concluir en esta primera fase, donde únicamente se analiza cual es la mejor técnica a preprocesar, que, si no se va a aplicar ningún tipo de preprocesado sobre las imágenes, la mejor técnica es PCA, pero si se va a utilizar imágenes preprocesadas, como los preprocesados estudiados, la mejor técnica es Whitening.

	Técnica recomendada
Sin imágenes preprocesadas	PCA
Con imágenes preprocesadas	WHITENING

Tabla 8.2: Recomendación de técnicas a utilizar.

Por lo que respecta al mejor preprocesado a aplicar sobre las imágenes, si se observan los resultados obtenidos, los preprocesados que han obtenido la tasa de reconocimiento puntual más elevada, independientemente de la técnica empleada, han sido el 000 y el 111 con un 90% de tasa de reconocimiento en ambos. El preprocesado 000 son las imágenes sin preprocesar y el 111 es el preprocesado donde se ha aplicado Viola-Jones junto con un ecualizado del histograma y una máscara elipsoidal. En sí mismo el preprocesado 000 no es ningún preprocesado, por lo tanto, se puede afirmar que el preprocesado 111 es el mejor en cuanto a resultado de nivel máximo de reconocimiento.

Si se comprueba si el preprocesado 111 es el que mejor comportamiento tiene para todas las bases de datos, se ha visto en el experimento 3 como para PCA el porcentaje de reconocimiento es mayor respecto a las imágenes originales en el preprocesado 111 en mayor medida que los otros preprocesados, realizando una mejora, como ya se ha

dicho anteriormente, en 3 de las 5 bases de datos. En el caso de Whitening en el mismo experimento 3, quedaba pendiente revisar otros datos debido a que cinco preprocesados, 010, 011, 101, 110 y 111, obtienen mejores resultados en 4 de las 5 BBDD estudiadas respecto del preprocesado 000.

Si son estudiadas las medias de la mejoría de los resultados del experimento 3 para un mismo preprocesado en todas las bases de datos respececto al preprocesado 000, se obtiene:

Técnica	**Preprocesado**	**Media (%)**
PCA	001	3,75
Whitening		4,35
PCA	010	3,1
Whitening		5,8
PCA	011	6,71
Whitening		7,08
PCA	020	-19,76
Whitening		-20,22
PCA	021	-17,4
Whitening		-31,43
PCA	100	-11,43
Whitening		3,8
PCA	101	-4,97
Whitening		18,93
PCA	110	8,22
Whitening		17,83
PCA	111	14,25
Whitening		22,37
PCA	120	-25,56
Whitening		-20,89
PCA	121	-10,21
Whitening		-11,96

Tabla 8.3: Media de reconocimiento por preprocesado y técnica

En la tabla mostrada arriba, se aprecia como para PCA y para Whitening, el preprocesado 111 es el que tiene una media mayor de reconocimiento respecto a los demás preprocesados.

En el experimento 2 se ha realizado una tabla, enumerando de mejor a peor preprocesado por base de datos. Si comparamos los preprocesados mencionados, se obtienen estos resultados:

Posición	010	011	101	110	111	Técnica
1					4	Whitening
			1		2	PCA
2	1		3	1		Whitening
	1			1	1	PCA
3			1	3		Whitening
	1		1	1	1	PCA
4		1				Whitening
		2		2		PCA
5	1					Whitening
	1					PCA
6		2				Whitening
		2				PCA
7	2	1				Whitening
	1	1				PCA
8	1	1				Whitening
	1	0	1	1		PCA
9				1		Whitening
						PCA
10						Whitening
			1			PCA
11			1			Whitening
					1	PCA
12					1	Whitening
			1			PCA

Tabla 8.4: Resumen posiciones mejores preprocesados

En la tabla anterior se puede observar como el preprocesado 111, ha sido en 6 de los 10 posibles casos, clasificado como el que mejores porcentajes de reconocimiento ha obtenido, de los cuales 4 son en la técnica Whitening y 2 en PCA.

Como en la tabla anterior no se aprecia de forma muy clara cuál puede ser el segundo, tercer, cuarto, etc… mejor preprocesado, se han comparado entre ellos y se ha realizado una tabla donde quedan clasificados por el número de veces que han obtenido una mejor posición en una base de datos respecto a los demás preprocesados en esta misma base de datos, sumando el resultado de todas las bases de datos. Esto es lo mismo que clasificar los proprocesados por el número de veces que han obtenido una mayor tasa de reconocimiento.

Posición	Preprocesado	Nº Casos	Total	Técnica
1	111	16	31 de 40	Whitening
		15		PCA
2	110	11	23 de 40	Whitening
		12		PCA
3	101	12	18 de 40	Whitening
		6		PCA
4	010	6	16 de 40	Whitening
		10		PCA
5	011	5	13 de 40	Whitening
		7		PCA

Tabla 8.5: Clasificación de los 5 mejores preprocesados

Las posiciones y preprocesados de la tabla anterior, coinciden con los de la tabla representada en el experimento 2, *Tabla 7.2.5.3*, donde estaban todos los preprocesados clasificados de esta misma forma, aunque en ella se tenía en cuenta el preprocesado 000, que como ya se ha dicho no es ningún preprocesado en sí mismo. En ese mismo experimento 2, se separaban los resultados por técnicas, las *tablas 7.5.2.4* y *7.5.2.5*, Whitening y PCA respectivamente, que confirma que para PCA el mejor preprocesado es el 111 seguido del 110, quedando las imágenes originales en el tercer lugar, aunque la mayor tasa puntual de reconocimiento se consigue con las imágenes sin preprocesar. Quedaba pendiente dilucidar para Whitening si realmente el preprocesado 111 también es el que mejor se comporta con carácter general, aparte de tener la tasa de reconocimiento puntual más elevada, así que con los datos de las tablas anteriores más los de la tabla del experimento 3, 7.5.2.4, queda confirmado que es el mejor preprocesado a aplicar con esta técnica con carácter general, seguido del preprocesado 101 y 110.

Se puede concluir el estudio indicando que, el preprocesado 111 es también el que mejores resultados ha obtenido en las bases de datos estudiadas con carácter general. También es el que mejores resultados ha obtenido en ambas técnicas estudiadas, PCA y Whitening, aunque para la técnica PCA el que mayor tasa puntual ha obtenido ha sido las imágenes originales. Por lo tanto, se puede resumir en la siguiente tabla:

	MEJOR PREPROCESADO
PCA	111
WHITENING	111

Tabla 8.6: Preprocesado óptimo a utilizar por técnicas