

# Sistemas de reconocimiento basados en la imagen facial

## Recognition systems based on the facial image

Henry Arguello Fuentes, Ph. D. (c)

Facultad de Ingenierías Físico-Mecánicas, Universidad Industrial de Santander, Bucaramanga, Colombia  
henarfu@uis.edu.co

Recibido para revisión 27 de junio de 2011, aceptado 18 de octubre de 2011, versión final 22 de noviembre de 2011

**Resumen** – Este artículo sintetiza las principales investigaciones que se están llevando a cabo en el área de los sistemas de reconocimiento a través de la imagen facial. Se realiza la descripción de las principales líneas de trabajo en los sistemas de identificación de personas por medio de la imagen del rostro. Además, se realiza una síntesis de las últimas técnicas matemáticas para realizar la extracción de características dentro de estos sistemas de identificación.

**Palabras clave** – Imagen facial, Procesado de imágenes, Reconocimiento de rostros, Sistemas de identificación.

**Abstract** – This article summarizes the most important research projects developed in the area of recognitions and stems using facial images. It presents the description of the principal lines of study about people identification systems using face recognition. Furthermore, it contains a synthesis of recent mathematical techniques for pattern extraction in these systems.

**Keywords** – FaceImages, ImageProcessing, Facerecognition, Identificationsystems.

### I. INTRODUCCIÓN

Los sistemas de reconocimiento de identidad de una persona han evolucionado considerablemente a través de estos últimos años. En la actualidad es posible realizar esta verificación por medio de diversas señales tales como: imágenes de las huellas digitales, el iris de los ojos, la palma de la mano, el rostro, o por señales como la voz de una persona o su firma manual[20], [26]. En la actualidad cada uno de estos sistemas tiene un alto grado de eficiencia para discriminar un individuo de otro, sin embargo, ninguna de estas técnicas es 100% confiable[26].

Los sistemas basados en el rostro tienen las tasas más altas de falsa aceptación y falso rechazo, sin embargo, tienen algunas cualidades que los hacen atractivos para ser utilizados como mecanismo de autenticación, tales como la facilidad de instalación de los sensores de captura (cámara digital), y su aceptación por parte del público ya que son los sistemas de verificación menos invasivos[20]. Diferentes autores han propuesto diversas técnicas para realizar la extracción de características y reducir la dimensionalidad del problema, esto es debido a que normalmente una imagen de un rostro de sólo 50 píxeles de ancho y 50 de alto tendría una dimensionalidad de 2500. Esto hace el problema de reconocimiento de rostros altamente costoso desde el punto de vista computacional.

Dentro de las técnicas más usadas para realizar la extracción de características se encuentran: El análisis de componentes principales (PCA) o también llamada Eigenfaces[21],[34], [39], [51],[55],[56], Análisis de discriminantes lineales de Fisher o Fisherfaces (FLD)[6], [13], [54],[55], Conservación de proyecciones locales (LPP) o Laplacianfaces[10], [23],[55], Proyecciones aleatorias o Randomfaces[7], [30],[55], Redes Neuronales Artificiales (RN) [18],[22],[35],[36],[41], [42],[48], Análisis de componentes independientes (ICA)[3], [58], características locales o análisis de subregiones[40],[59],[60], correlación (CORR) y DCT[19], [22], [28] y combinaciones de dos o más técnicas antes mencionadas[28],[40], [47]. Las anteriores estrategias son basadas en imágenes fijas del individuo que se desea reconocer, sin embargo, en la actualidad una gran cantidad de investigaciones han sido realizadas alrededor de los sistemas basados en modelos 3D de la cabeza de la persona[8], [9],[25],[33], [46] y de la misma forma esquemas que trabajan con señales de vídeo y no con imágenes fijas[1], [20], [24].

Para implementar un sistema de reconocimiento de rostros se presentan 6 etapas bien definidas: Captura de la imagen, preprocesamiento, localización, escalamiento y ajuste, extracción de características y por último la clasificación y toma de la decisión[4].

Dentro de la captura de la imagen se encuentra la selección de la cámara digital y las características de iluminación, ya sean controladas o no controladas. Diferentes autores han planteado que en el desempeño de estos sistemas de reconocimiento influyen más las variaciones de iluminación en la toma de las imágenes faciales que las debidas a la escala, rotación, pose o la presencia de accesorios en el rostro. Por este motivo algunas correcciones son realizadas para tratar de compensar las variaciones de iluminación[6], [32], [36], [53].

El preprocesamiento de la imagen capturada incluye la selección del espacio de color en el que se desea trabajar o la extracción de la intensidad en el caso de que se trabaje en escala de grises[11],[27], [50].

Una vez preprocesada la imagen de entrada se determina las coordenadas donde se encuentra la posición de la cara dentro de la escena, normalmente se desea determinar las coordenadas de una subimagen formada sólo por la zona de la cara delimitada por las orejas, la frente y el mentón. Para determinar estas coordenadas normalmente se utiliza un algoritmo basado en una cascada de clasificadores básicos denominado Adaboost[52]. Otras estrategias para localizar el rostro consiste en determinar la posición de los ojos[35], [57] y en otros casos se utiliza la información del color de la piel como metodología para determinar la región que forma exclusivamente la cara[2],[32], [48], [60], o clasificadores básicos basados en redes neuronales[16].

En algunos sistemas se realiza la normalización de la cara ya sea utilizando la información de partes como los ojos, nariz u otras características[28], [54], [60] o simplemente realizando un escalado de la imagen. Los sistemas que sólo hacen hasta las operaciones ahora mencionadas se llaman sistemas de localización[16], [52].

Determinada la región que contiene el rostro se procede a extraer sus características. Para ello se puede utilizar cualquiera de las técnicas antes mencionadas o combinaciones de ellas. Normalmente esta operación produce un vector de características de baja dimensión que debe ser comparado con una base de datos de personas previamente almacenada. Si el individuo en prueba revela cual es su identidad por algún método alternativo tal como un código, tarjeta inteligente u otro medio el sistema se llama sistema de verificación o autenticación de identidad. En este caso se compara el vector de características del individuo en prueba con el almacenado en la base de datos, si esta medida es inferior a un determinado umbral la identificación es positiva en caso contrario se trata de un intruso.

Si el individuo en prueba no revela previamente su identidad entonces se trata de un sistema de reconocimiento o identificación de personas[1], [2], [19], [21],[48],[58], [60]. En este caso se compara el vector de características del individuo en prueba con cada uno de los de la base de datos, se obtienen N valores de distancias, siendo N el número de individuos. Para

realizar esta comparación existen diferentes métricas tales como la distancia lineal, euclidiana, redes neuronales, Mahalanobis, coseno del ángulo entre vectores, sistemas fuzzy, coeficiente de correlación entre otras[14], [15]. De estas distancias se selecciona la más pequeña y se compara con un umbral previamente seleccionado que determina si efectivamente el individuo de prueba es una persona en la base de datos o se trata de un intruso.

De esta forma se obtiene un sistema de reconocimiento basado en la imagen facial. A través de este artículo se realiza una descripción y comparación de algunas de las técnicas mas importantes para la extracción de características y se indican las principales tendencias en las investigaciones realizadas en el área de los sistemas de reconocimiento basados en la imagen facial.

## II. EXTRACCIÓN DE CARACTERÍSTICAS

Una de las partes más importantes en un sistema de reconocimiento lo constituye la extracción de características. Como se mencionó anteriormente desde este punto de vista estos sistemas podrían ser clasificados en tres categorías: Los que trabajan con imágenes fijas, señales de vídeo normalmente de baja resolución y los que trabajan con imágenes 3D de la cabeza. Es importante resaltar que es difícil comparar cada una de estas categorías, ya que las investigaciones de cada una de ellas se realizan sobre bases de datos totalmente diferentes.

Por ejemplo, los sistemas basados en 3D son probados con imágenes tomadas con escáneres altamente costosos mediante los cuales se construyen modelos de la cabeza y debido a los costos el número de imágenes 3D es limitado, en contraste los sistemas basados en imágenes fijas cuentan con bases de datos extensas y un número grande de individuos como las que se pueden encontrar en[44]. Los sistemas basados en vídeo lógicamente que tienen el problema del excesivo tamaño de la información. Adicionalmente, cada una de las investigaciones en imágenes fijas, 3D y vídeo normalmente no realizan comparaciones entre ellas. En su lugar suministran porcentajes de desempeño basados en la categoría específica que cada investigador selecciona y realizan las comparaciones con otros trabajos que se encuentren exclusivamente dentro de la misma categoría.

Por otro lado, es fácil encontrar la comparación entre diferentes metodologías o técnicas dentro de una misma categoría. Sin embargo, existen investigaciones como la presentada en[43], donde se encontró que los sistemas basados en 3D y vídeo no aumentaban el desempeño de los sistemas de reconocimiento. Sin embargo, esto no quiere decir que esas categorías deban ser desechadas como posibles áreas de investigación, en su lugar, cada una de estas categorías se ha especializado en resolver problemas específicos del reconocimiento de rostros.

Por ejemplo, se puede observar como en las investigaciones en la categoría de 3D se presenta cierta invarianza de tales sistemas a la pose de la persona que debe ser identificada[8], [53], a su vez las dificultades debidas a la iluminación son más fáciles de corregir que en las otras categorías. Por otro lado, los sistemas basados en vídeo se han especializado en trabajar en imágenes de baja resolución para sistemas tales como la televisión, las video conferencias o los sistemas de seguimiento de individuos donde las imágenes son de baja calidad y el movimiento del individuo dentro de la escena puede suministrar información adicional que ayude a realizar la identificación[20].

De esta manera cualquiera de las 3 categorías está abierta a realizar nuevos y mejores desarrollos. Sin embargo, en general los trabajos basados en imágenes fijas presentan los mejores resultados y además el número de investigaciones y de técnicas es mucho más amplio y completo. Las otras categorías actualmente son más costosas en cuanto a su implementación y están relegadas a aplicaciones específicas o investigación.

### III. MÉTODOS BASADOS EN IMÁGENES FIJAS

Como se mencionó anteriormente los métodos basados en imágenes fijas son los más desarrollados e investigados de las 3 categorías. La técnica más utilizada y a partir de la cual se han venido construyendo las demás técnicas es la de eigenfaces o eigenvectores (PCA) inicialmente propuesta por Turk and Pentland en 1991[51]. Esta técnica reduce la dimensionalidad de las imágenes usando la técnica de análisis de componentes principales al maximizar la variación de la matriz de covarianza de todas las imágenes de la base de datos[49].

La principal ventaja de esta técnica es la facilidad de implementación y el bajo coste computacional, por lo cual permite la creación de sistemas en tiempo real. Implementaciones recientes muestran que esta técnica tan antigua sigue siendo muy utilizada[4], [39]. En algunos casos los sistemas son desarrollados con variantes de esta técnica como la desarrollada en[21],[37], [48] que usan PCA por partes CPCA o en [41] usando PCA basado en kernels específicos KPCA. Naturalmente los autores de las investigaciones que utilizan variaciones del PCA manifiestan que la eficiencia de sus sistemas son mayores que los que utilizan la técnica de eigenfaces tradicional. Sin embargo, el aumento en la eficiencia está muy relacionada con la base de datos que se utilizó en cada caso. En conclusión se puede observar como en la actualidad resulta atractivo aun desarrollar sistemas basados en PCA.

En 1997 Belhumeur, Hespanha y Kriegman[6],[17] desarrollaron la técnica de Proyecciones lineales sobre clases específicas o también llamada Fisherfaces. Esta técnica fue desarrollada principalmente para soportar variaciones de iluminación y pose de los individuos. Sin embargo, el éxito de esta técnica depende de que se disponga de varias tomas del

mismo individuo en diferentes condiciones de iluminación y pose. De esta manera la descomposición en Fisherfaces interpola o extrapola las demás condiciones de pose o iluminación que se puedan presentar. La debilidad de la técnica está en que se necesita varias tomas del mismo individuo y que estas sean representativas de las variaciones que se vayan a presentar en la aplicación real, además esta técnica es más costosa computacionalmente que la de PCA [29].

Es importante resaltar que la técnica de Fisherfaces se basa en la técnica de PCA y por lo tanto corresponde a la reducción de dimensionalidad de las imágenes maximizando la matriz de covarianza de la base de datos disponible. En este caso existen dos matrices de covarianza, la correspondiente a las diferentes tomas de un mismo individuo o inter-clase y la matriz de covarianza debida a tomas de individuos con diferente identidad o extra-clase. De esta forma la técnica de Fisherfaces maximiza la reducción de dimensionalidad teniendo en cuenta tanto las diferencias o variaciones de un mismo individuo y de este con todos los demás de la base de datos que se esté analizando. Esta técnica no es tan conocida como la de PCA, sin embargo se pueden encontrar sistemas de reconocimiento que usan esta técnica como método de clasificación[47], [54] o variaciones recientes de la técnica como la encontrada en[58] llamada Gram-Schmidt ortogonalización por LDA (GSLDA).

En el 2002, Bartlett, Movellan y Sejnowski desarrollaron la técnica de Análisis de componentes independientes ICA[3]. Este método a diferencia de PCA aborda el problema de la representación de un rostro en un espacio de baja dimensionalidad desde un punto de vista estadístico. Existen dos versiones del algoritmo llamados ICA-I e ICA-II que abordan el problema tratando las imágenes completas o los píxeles de cada imagen como una variable aleatoria respectivamente. Es importante resaltar que ICA usa la técnica de PCA para reducir la dimensionalidad de los datos. De la misma manera como PCA proyecta las imágenes en un espacio  $n$  dimensional formado por los eigenvectores que maximizan la matriz de covarianza de la base de datos, ICA proyecta las imágenes en un espacio de baja dimensión formado por un conjunto de vectores estadísticamente independientes[5], [14], [15]. En [3] se presenta un análisis del desempeño de ICA frente a PCA, sugiriendo que este método es superior a la tradicional técnica de eigenvectores. Sin embargo, otras investigaciones han mostrado que esta superioridad de ICA depende de la base de datos bajo análisis y de las variaciones que se estén considerando en el conjunto de prueba, tales como iluminación, pose, gestos, presencia de accesorios entre otros[14], [15], sin embargo, sobresale el buen desempeño del algoritmo ICA-II sobre las otras técnicas. Otros autores consideran que los porcentajes de desempeño de los sistemas basados en ICA, PCA y LDA no se diferencian en gran medida y muestran que los porcentajes de desempeño son influenciados por el tamaño de la base de datos utilizada[55].

En el año 2005, He, Yan, Hu, Ni yogi y Zhang desarrollaron la

técnica de Laplacianfaces. Las técnicas de PCA y LDA reducen la dimensionalidad de las imágenes tratando de preservar la mayor cantidad de información global de cada una de las imágenes que forman la base de datos, por otro lado, la técnica de Laplacianfaces realiza esta reducción teniendo en cuenta las características locales de las imágenes. De esta forma esta técnica no trata de conservar la estructura total de la imagen, en lugar de esto tiene en cuenta las características parciales de la imagen lo que le permite discriminar más fácilmente dos imágenes. Esto es debido a que ciertas estructuras locales de las imágenes permiten diferenciarlas de las otras, en contraste al mirar las imágenes globalmente podrían parecer muy semejantes.

Esta técnica tiene dos etapas en la primera se realiza un mapeo de las imágenes usando la técnica de Conservación de proyecciones locales o LPP y luego realizando un análisis de PCA sobre estas imágenes mapeadas[23]. Los autores de esta técnica muestran la superioridad de esta metodología comparada con PCA o LDA. Otros estudios muestran que efectivamente la técnica de LPP es superior a las antes mencionadas, sin embargo, los resultados son dependientes del tamaño de la base de datos y del número de eigenvectores utilizado para representar las imágenes.

En el 2006, Cai, He, Han y Zhang realizaron una mejora de este método y lo llamaron Ortogonal Laplacianfaces[10]. Sin embargo, trabajos desarrollados en 2008, tal como[55], demuestran que Laplacianfaces, Randomfaces, PCA, LDA e ICA tienen porcentajes de eficiencia muy parecidos e indican que dependiendo del tamaño de la base de datos de prueba una u otra técnica resulta ser la mejor. Debido a que esta metodología es reciente no se encuentran muchas implementaciones de sistemas de reconocimiento que usen esta técnica.

#### IV. MÉTODOS BASADOS EN IMÁGENES 3D

Recientemente un gran número de trabajos han sido desarrollados en el área del reconocimiento a través de imágenes 2.5D y 3D de la cabeza de las personas. En algunos trabajos las imágenes efectivamente son tomadas con un juego de cámaras, con un escáner especializado[33] o son diferentes tomas de una persona con el mismo dispositivo[53]. En otros trabajos una estrategia de renderizado es utilizada para formar imágenes 3D a partir de una o varias imágenes 2D [9]. Para poder realizar esta construcción 3D es necesario tener un modelo general de la cabeza de las personas.

Una de las principales motivaciones de estas investigaciones es que una vez que se tiene el modelo 3D de la cabeza de las personas se pueden realizar comparaciones que son menos sensibles a factores tales como la pose, la iluminación y los gestos de las personas. En [47] se muestra que la identificación de personas utilizando métodos en 3D no supera en eficiencia las

técnicas que utilizan imágenes fijas, sin embargo, recientemente en [45] esta metodología ya tiene iguales y superiores resultados a las otras técnicas. Una de las principales áreas de trabajo en esta técnica es formar una imagen 3D de la cabeza de una persona a través de un modelo previamente calculado. Teniendo este modelo con sólo una imagen de entrada fija en 2D se puede formar su respectiva representación en 3D.

Blanz y Vetter en 2003 desarrollaron esta técnica basándose en el principio de la reducción de dimensionalidad usando la técnica de análisis de componentes principales PCA. La imagen 3D que usaron es representada por un grupo de tripletas  $(X_1, Y_1, Z_1, X_2, Y_2, Z_2, \dots, X_n, Y_n, Z_n)$ . Cada uno de estos puntos es una característica de la imagen 3D que será tenida en cuenta para realizar el reconocimiento. Estas tripletas son calculadas a

partir de las duplas  $(x_1, y_1, x_2, y_2, \dots, x_n, y_n)$  de la toma fija de la imagen 2D mediante una transformación espacial[9]. Estos resultados son útiles para aplicaciones donde los individuos que se tratan de reconocer son no colaborativos o no saben de la presencia del sistema de reconocimiento. Los porcentajes de reconocimiento que estos autores muestran están por encima del 90%, sin embargo la base de datos que usan es demasiado específica y no se pueden generalizar los resultados a su contraparte en 2D.

Hu, Jiang, Yan, Zhang y Hongjiangzhang en 2004, usan la misma técnica de reconstrucción. Estos autores muestran que esta técnica es superior a su contraparte 2D cuando se usa PCA o LDA para realizar la extracción de características. Sin embargo, estos resultados son ciertos para variaciones amplias en la pose del individuo, por ejemplo de 45 grados. Es importante resaltar que esta área de investigación no requiere tener equipo especializado para su realización ya que las bases de datos de modelos 3D de personas se encuentran disponibles en línea de forma gratuita, además, se pueden obtener sistemas software gratuitos que forman la imagen 3D de una persona a partir de una simple fotografía en 2D.

#### V. MÉTODOS BASADOS EN VÍDEO

Este método comprende las aplicaciones que tienen en cuenta las relaciones entre cada uno de los cuadros de un vídeo [1]. En la mayoría de las ocasiones los vídeos que se analizan son para aplicaciones con baja resolución de vídeo, como los provenientes de sistemas de seguridad, televisión, vídeo conferencias entre otras[20].

En 2005 Gorodnichy, propone un esquema basado en la forma de asociación del cerebro humano para desarrollar una estrategia de reconocimiento a través de vídeo. En este trabajo Gorodnichy propone que su sistema debe ser capaz de reconocer un rostro dentro de una escena que tenga por lo menos 12 píxeles de separación entre los ojos de los individuos, esto motivado

por el hecho de que los seres humanos tenemos esa misma capacidad. Los resultados de este sistema alcanzan porcentajes de reconocimiento del 95% sujeto a una base de datos de videos cortos donde se presentan caras de personas y videos tomados de programas de televisión. Previamente en 2002 un informe industrial sobre eficiencia de los sistemas basados en reconocimiento facial y a través del iris de los ojos[43], encontró que los sistemas basados en video no aportaban mejores resultados que los obtenidos en trabajos equivalentes con imágenes fijas. Este mismo informe del año 2006[45], ya no menciona las técnicas basadas en video y en su defecto realza los grandes avances de las técnicas basadas en imágenes fijas y en 3D. Además, realizar la comparación entre los métodos basados en video e imágenes fijas no es posible ya que las bases de datos que utilizan estas dos metodologías son diferentes.

## VI. DESEMPEÑO DE LAS TÉCNICAS DE RECONOCIMIENTO

Como se mencionó anteriormente de los 3 métodos de reconocimiento de rostros, el de imágenes fijas es el que ha mostrado mejores resultados. Además, esta metodología tiene la mayor cantidad de investigaciones y están disponibles una amplia variedad de bases de datos de prueba. Por esta razón resulta de gran importancia determinar la eficiencia de las diferentes técnicas de este método. Existe una diversidad de bases de datos disponibles en línea y algunas son gratuitas (<http://www.face-rec.org/>).

En la actualidad una de las bases de datos más utilizada y completa para probar sistemas de reconocimiento de rostros es la base de datos FERET, que está disponible en línea en: <http://face.nist.gov/colorferet/>. La base de datos contiene 14,126 imágenes de 1,199 individuos y ocupa un espacio aproximado de 8.5GB [44]. A partir de esta base de datos se han realizado varios trabajos y estudios donde se comparan las diferentes técnicas de reconocimiento[14], [15], [43], [44], Delac, Grgic y SonjaGrgic en 2004 y 2005[14], [15], verificaron el desempeño de los algoritmos de PCA, ICA-I, ICA-II y LDA utilizando la base de datos FERET. En esta prueba se tuvieron en cuenta 3 factores en las imágenes de prueba: expresión, iluminación y toma en tiempos distintos. Se utilizaron 4 métricas para tomar la decisión final sobre la identidad de los individuos: Norma L1 o distancia lineal, Norma L2 o distancia euclidiana, coseno del ángulo entre vectores y la distancia de Mahalanobis.

De forma general ninguna de las 4 técnicas puede ser considerada como la mejor. Para condiciones del sistema de reconocimiento en que se varía principalmente la iluminación, se encontró que el método de PCA+L1 es la combinación que proporciona mejores resultados. Cuando se tiene en cuenta principalmente la variación en la expresión de los individuos, la combinación de LDA+cosenos es la que suministra un mejor desempeño, sin embargo el margen con el cual esta combinación supera a las demás es muy pequeño, por lo cual no se puede

realmente concluir que esta medida sea la más apropiada. Cuando se trata de cambios temporales en los rostros, la técnica de ICA2+coseno es la combinación que proporcionó mejores resultados. Una importante conclusión de este trabajo es que la medida L2 no mostró buenos resultados en ninguno de los 3 casos que se analizaron. Podría decirse que la combinación ICA2+coseno fue la que mostró mejores resultados, sin embargo esto no se cumple de forma general. Según lo anterior aun faltan investigaciones que permitan establecer una técnica que sea realmente superior a las demás en todas las condiciones de los individuos que forman las bases de datos. Por este motivo dependiendo de las condiciones que prevalezcan en una determinada aplicación podría escogerse la técnica a utilizar para realizar la extracción de características[15].

Karimi and Krzyak en 2007, realizaron una comparación entre las técnicas de PCA, FLD, Laplacianfaces, PCA+filtrosGabor, FLD+filtrosGabor. En este experimento se consideró la influencia de utilizar las tres componentes de color (RGB). La base de datos que se utilizó en este experimento fue la de CVL con 114 individuos y 7 tomas por individuo y la de Georgia con 50 individuos y 15 tomas por individuo. Los resultados mostraron que Laplacianfaces es la técnica que mejores resultados presentó. Además, este trabajo indicó que realizar una combinación de los resultados de los canales RGB mejora el desempeño de los sistemas de reconocimiento. Ya sea en color o en escala de grises este trabajo muestra que la técnica de Laplacianfaces tiene un desempeño del 100% lógicamente sujeta a las bases de datos que se utilizaron [29].

Es importante resaltar que los porcentajes de reconocimiento de los diferentes trabajos varía significativamente de un trabajo a otro. Algunos trabajos de investigación muestran en sus resultados eficiencias por debajo del 90%[2], [14], [15], [21], [28],[37], [38], [42],[48], [54]mientras en otros casos cercanos al 100%[29], [35], [41], [47], [60]. Por otro lado en 2006 el reporte de vendedores de sistemas de reconocimiento[45], mostró que las empresas que se dedican a vender sistemas comerciales han logrado obtener porcentajes del 99.99% para un porcentaje de falsa aceptación de sólo 0.001%. Desafortunadamente este reporte no muestra las técnicas que fueron usadas para obtener estos desempeños ya que los algoritmos están protegidos por los derechos de autor de las diferentes empresas que participan en este concurso. Sin embargo, sus resultados son importantes ya que utilizan para probar los sistemas la base de datos FERET la cual es ampliamente conocida y suficientemente diversa. En este concurso participan empresas como Google que tiene una filial que se dedica a desarrollar este tipo de aplicaciones.

## VII. COMBINACIÓN DE CLASIFICADORES

Como se mostró anteriormente ninguna de las técnicas existentes hasta el momento garantiza ser la mejor en cualquier condición de las imágenes de la base de datos. Por este motivo

diversos investigadores han optado por realizar combinaciones de técnicas para lograr aumentar la eficiencia de sus sistemas. Algunas de las combinaciones utilizadas se indican en la Tabla 1. Se muestran los métodos, la referencia donde se pueden encontrar los resultados, el porcentaje de falsa aceptación (PFA), la base de datos que se utilizó en el experimento (En caso de que no sea una base de datos internacionalmente conocida, se indica que es local y se muestran el número de individuos utilizados y el número de tomas por cada individuo) y la eficiencia del sistema calculada como el número de veces que acertó el algoritmo dividido por el número total de imágenes

**Tabla 1.** Combinaciones utilizadas en los sistemas de reconocimiento

Métodos	Referencias	PFA (%)	Bases de datos	Eficiencia (%)
PCA+RN	[42]	4.6	Local (20 individuos con 10 tomas)	92.59
LDA+RN	[42]	6.08	Local (20 individuos con 10 tomas)	88.89
KPCA+SMV	[41]	-	ATT	97
DCT+RN	[22]	-	Local (40 individuos con 10 tomas)	88.5
PCA+Fuzzy-RN	[35]	0.42	Local (40 individuos con 24 tomas)	93.7
CPCA+RN	[48]	-	Local (100 individuos con 1 toma)	90
PCA+FiltrosGabor	[29]	-	CVL	93
LDA+FiltrosGabor	[29]	-	CVL	93
PCA+LDA+NN	[47]	-	Yale	99.5
PCA+CORR	[28]	12	Local (109 imágenes de 43 individuos)	85
DCT+CORR	[28]	12	Local (109 imágenes de 43 individuos)	84

#### VIII. SISTEMAS DE RECONOCIMIENTO EN TIEMPO REAL, IMPLEMENTACIONES EN HARDWARE

La mayoría de trabajos que se han indicado en este artículo no especifican las prestaciones de los sistemas de computo utilizados. Esto es debido a que en la mayoría de las ocasiones los autores están más interesados en obtener porcentajes de eficiencia altos y no se preocupan por la velocidad o facilidad de utilización del sistema desarrollado. Sin embargo, algunos autores han realizado diversas investigaciones alrededor de las implementaciones de estos sistemas en tiempo real. Las implementaciones se realizan en DSPs, sistemas VLSI y Chips de propósito específico. Algunas de las implementaciones que se pueden encontrar están basadas principalmente en la plataforma C6000 de la empresa Texas Instruments[4], [16], [22].

Gilbert y Yang en 1993 desarrollaron un sistema de reconocimiento mediante un chip de propósito específico que procesaba a 1000 cuadros por segundo, sin embargo este sistema que utilizaba la correlación para realizar la comparación de los individuos tenía una resolución de sólo 68x68 píxeles.

Fatemi, Kleihorst, Corporaal y Jonker en 2003, realizaron la implementación de un sistema de reconocimiento basado en un esquema multiprocesador embebido dentro de una cámara de vídeo CCD de propósito general. Este trabajo logró una

de la base de datos. Como puede observarse los resultados son muy diversos en cuanto a su porcentaje de eficiencia. La base de datos utilizada en cada experimento incide en el porcentaje de desempeño que los autores indican en sus trabajos. Una de las técnicas que sobresale en estas combinaciones es la de redes neuronales, ya que en una gran cantidad de trabajos esta técnica es conjugada con las técnicas tradicionales para dar una mayor eficiencia. Los porcentajes de falso rechazo no son suministrados en todos los trabajos, lo que hace imposible comparar los resultados.

velocidad de 4.2 milisegundos por rostro reconocido, tenía una eficiencia del 90% y una tasa de falso rechazo del 1%. La técnica de reconocimiento empleada fue las redes neuronales de base radial RBF [18].

Batur, Flinchbaugh y Hayes en 2003, realizaron la implementación de un sistema DSP TMS320C6416 de la técnica de PCA que podía reconocer un rostro en alrededor de 3.5 segundos[4].

GökselGünlü en 2005 realiza la implementación de un sistema en tiempo real en un DSP TMS320C6713, en el cual se usa la técnica de DCT+RN. Los resultados de este trabajo mostraron que en promedio es posible realizar el proceso de entrenamiento de la base de datos (400 imágenes) en 0.5 segundos[28].

Morizet, Amiel, Hamed y Ea en 2007 realizaron la comparación de implementar un sistema de reconocimiento en tiempo real en un PC, un DSP y un FPGA. La técnica de reconocimiento usada fue la de PCA. Los resultados mostraron que el DSP y el PC (3GHz) tienen ordenes de procesamiento de 5 milisegundos, mientras las tarjetas FPGA del orden de los 60 milisegundos. Aunque las tres soluciones tienen tiempos de procesamiento adecuados para una implementación en tiempo real, es importante resaltar que el PC es una solución muy costosa y ocupa un gran espacio físico, la solución basada en

FPGA es la más compacta pero la más compleja de implementar y la solución DSP parece ser la que tiene la mejor relación entre facilidad de implementación, desempeño y espacio requerido[39].

Otra de las tendencias de las implementaciones hardware es realizar sistemas basados en clúster de computadores para examinar bases de datos de gran tamaño. Chunhong, Cuangda y Xiaodong en 2003 realizaron la implementación de un clúster basado en equipos Pentium III logrando reducir la velocidad de procesamiento alrededor de 4 veces de lo que haría un solo computador[12].

Ming, Ma, Zhang y Xiong en 2006 realizaron el mismo experimento de[12], sin embargo esta vez los equipos eran Pentium IV de 3.5GHz, en este trabajo los autores obtienen un tiempo de 3.8 segundos por reconocimiento con un solo nodo del clúster y 0.29 segundos usando 20 lo que les proporciona una reducción de procesamiento de alrededor de 13 veces[38].

En 2005, Meng, Guang-da su, Li, Fu y Zhou implementaron un clúster basado en 6 PC Xeon de 2.4GHz logrando un tiempo de 1.094 segundos para encontrar un rostro dentro de una base de datos de 2.560.000 individuos[37]. Finalmente en 2006, Lahdenojat, Maunu, Laiho y Paasio implementaron un sistema basado en un vector de procesadores que logra comparar un patrón de referencia con 100 imágenes en alrededor de 1.5 milisegundos[31].

## IX. CONCLUSIONES

Una situación que dificulta comparar los resultados es que los diferentes trabajos de investigación muestran los porcentajes de rostros que clasifican correctamente y no discriminan entre los porcentajes de falso rechazo y falsa aceptación.

Diversos trabajos muestran que el preprocesamiento aumenta considerablemente el desempeño de estos sistemas de reconocimiento. En una gran cantidad de investigaciones los buenos resultados se deben más a la corrección de iluminación, alineación, escalado que a la misma técnica que se use para realizar la selección de patrones. Existe alguna discusión aún sobre si el color puede ser un factor que ayude a aumentar el porcentaje de reconocimiento, algunos autores aseveran que este no tiene incidencia y otros mediante experimentos demuestran lo contrario.

En la actualidad no existe una técnica de reconocimiento que en realidad pueda ser considerada como la mejor, en lugar de ello dependiendo de las condiciones específicas donde va a operar el sistema de reconocimiento podría seleccionarse la técnica más apropiada.

Es importante que cuando se realice cualquier investigación en esta área del conocimiento se utilicen bases de datos internacionales como la FERET que permiten comparar verdaderamente los resultados con otros trabajos similares.

Los porcentajes de reconocimiento actuales de los trabajos de investigación que usan una sola técnica de extracción de características están alrededor del 90%. Las implementaciones industriales que utilizan la base de datos FERET para ser probados tienen un desempeño aproximado del 99.99% con una tasa de falsa aceptación de tan solo 0.001%. Estos valores deben ser tenidos en cuenta cuando se desee realizar cualquier implementación a futuro.

De este trabajo de investigación se observó que en la mayoría de implementaciones de estos sistemas los investigadores prefieren realizar combinaciones entre las diferentes técnicas existentes y que esta metodología efectivamente aumenta el porcentaje de reconocimiento de los sistemas. A si mismo, se encontró que las técnicas mas utilizadas en las investigaciones son PCA y cualquiera de sus variaciones CPCA o KPCA y redes neuronales artificiales RN. Dentro de esta última técnica los investigadores usan principalmente las máquinas de soporte vectorial SVM, las funciones de base radial RBF y en algunos casos el perceptrón multicapa MLP.

Se encontró además que los DSPs son uno de los hardware preferidos a la hora de realizar implementaciones en tiempo real y que existe una diversidad de trabajos basados en la arquitectura C6000 de la empresa Texas Instruments.

No se encontraron implementaciones en tiempo real de aplicaciones basadas en modelos 3D de la cabeza, esta puede ser un área de investigación para futuros trabajos.

## AGRADECIMIENTOS

Principalmente a Dios fuente de toda inspiración. Este trabajo es financiado por la Vicerrectoría de Investigación y Extensión de la Universidad Industrial de Santander, mediante el proyecto de investigación con código interno 5537 y titulado: Sistema en tiempo real de verificación de identidad a través de la imagen facial.

## REFERENCIAS

- [1] Aggarwal G.; Roy Ch. A. K. y Chellappa R., 2004. A system identification approach for video-based face recognition. En: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, Vol. 4, pp. 175- 178.
- [2] Amnuaykanjanasin P.; Aramvith S. y Chalidabhongse T.H., 2006. Real-time face identification using two cooperative active cameras. En: Proceedings of the 9th International Conference on Control, Automation, Robotics and Vision (ICARCV '06), 5-8, pp. 1-6.
- [3] Bartlett M. S.; Movellan J. R. y Sejnowski T. J., 2002. Face recognition by independent component analysis. En: IEEE Transactions on Neural Networks, Vol. 13, No. 6, pp. 1450-1464.
- [4] Batur A. U.; Flinchbaugh B. E. y Hayes III M. H., 2003. A DSP-based approach for the implementation of face recognition algorithms. En: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Proceedings, (ICASSP

- '03), Apr. 6-10, Vol. 2, pp. II - 253-6.
- [5] Becker B. C. y Ortiz E. G., 2008. Evaluation of face recognition techniques for application to facebook. En: IEEE International Conference on Automatic Face and Gesture Recognition, pp. 1-6.
- [6] Belhumeur P. N.; Hespanha J. P.; Kriegman D. J., 1997. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. En: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, pp. 711-720.
- [7] Bingham E. y Mannila H., 2001. Random projection in dimensionality reduction: Applications to image and text data. En: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 245-250.
- [8] Blanz V. y Vetter T., 2003. Face recognition based on fitting a 3D morphable model. En: Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 9, pp. 1063-1074.
- [9] Blanz V.; Grother P.; Phillips P. J. y Vetter T., 2005. Face recognition based on frontal views generated from non-frontal images. En: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005, June 20-25, Vol. 2, pp. 454- 461.
- [10] Cai D.; He X.; Han J. y Zhang H.-J., 2006. Orthogonal laplacianfaces for face recognition. En: IEEE Transactions on Image Processing, Vol. 15, Issue 11, pp. 3608-3614.
- [11] Chang T. C.; Huang T. S. y Novak C., 1994. Facial feature extraction from color images. En: Proceedings of the 12th IAPR International Conference on Pattern Recognition, Oct. 9-13, Vol. 2, pp. 39-43.
- [12] Chunhong J.; Cuangda S. y Xiaodong L., 2003. A distributed parallel system for face recognition. En: Proceedings of the Fourth International Conference on Parallel and Distributed Computing, Applications and Technologies, PDCAT'2003, Aug. 27-29, pp. 797- 800.
- [13] Dai G.; Yeung D.-Y. y Qian Y.-T., 2007. Face recognition using a kernel fractional-step discriminant analysis algorithm. En: IEEE Transactions on Pattern Recognition, Vol. 40, Issue 1, pp. 229-243.
- [14] Delac K.; Grgic M. y Grgic S., 2005. A comparative study of PCA, ICA and LDA. En: Proceedings of the 5th EURASIP Conference focuses on Speech and Image Processing, Multimedia Communications and Services, EC-SIP-M.
- [15] Delac K.; Grgic M. y Grgic S., 2005. Statistics in face recognition: analyzing probability distributions of PCA, ICA and LDA performance results. En: Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, pp. 289-294.
- [16] Delongi P.; Polikarpov' B. y Krumnikli M., 2007. Face detection by DSP using directly connected camera. En: Proceedings of the 17th International Conference Radioelektronika, Apr. 24-25, pp. 1-3.
- [17] Etemad K. y Chellappa R., 1997. Discriminant analysis for recognition of human face images. En: J. Opt. Soc. Am., Vol. 14, No. 8, pp. 1724-1733.
- [18] Fatemi H.; Kleihorst R.; Corporaal H. y Jonker P., 2003. Real-time face recognition on a smart camera. En: Proceedings of Advanced Concepts for Intelligent Vision Systems, Ghent, Belgium, Sep. 2-5, pp. 222-227.
- [19] Gilbert M. y Yang W., 1993. A real-time face recognition system using custom VLSI hardware. En: Proceedings Computer Architectures for Machine Perception, Dec. 15-17, pp. 58-66.
- [20] Gorodnichy D. O., 2005. Video-based framework for face recognition in video. En: Second Workshop on Face Processing in Video (FPiV'05), Proceedings of Second Canadian Conference on Computer and Robot Vision (CRV'05), pp. 330-338.
- [21] Gottumukkal R. y Asari V. K., 2003. System level design of real time face recognition architecture based on composite PCA. En: Great Lakes Symposium on VLSI, Proceedings of the 13th ACM Great Lakes symposium on VLSI. Washington, D. C., USA. VLSI design, pp. 157-160.
- [22] Günlü G., 2005. Improving DSP performance for artificial neural networks based face recognition. En: Proceedings of the IEEE 13th Conference on Signal Processing and Communications Applications, pp. 551-556.
- [23] He X.; Yan S.; Hu Y.; Niyogi P. y Zhang H.-J., 2005. Face recognition using laplacianfaces. En: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 3, pp. 328-340.
- [24] Huang K. S. y Trivedi M. M., 2002. Streaming face recognition using multicamera video arrays. En: Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02), Vol. 4, pp. 213-216.
- [25] Husken M.; Brauckmann M.; Gehlen S. y Von der Malsburg C., 2005. Strategies and benefits of fusion of 2D and 3D face recognition. En: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 3, pp. 174.
- [26] Jain A. K. y Pankanti S., 2008. Beyond fingerprinting: Is biometrics the best bet for fighting identity theft?. En: Scientific American Magazine.
- [27] Jones III C. F. y Abbott A. L., 2004. Optimization of color conversion for face recognition. En: EURASIP Journal on Applied Signal Processing, pp. 522-529.
- [28] Kar S.; Hiremath S.; Joshi D. G.; Chadda V. K. y Bajpai A., 2006. A multi-algorithmic face recognition system. En: International Conference on Advanced Computing and Communications, ADCOM 2006, pp. 321-326.
- [29] Karimi B. y Krzyzak A., 2007. A study on significance of color in face recognition using several eigenface algorithms. En: Canadian Conference on Electrical and Computer Engineering, CCECE 2007, Apr. 22-26, pp. 1309-1312.
- [30] Kaski S., 1998. Dimensionality reduction by random mapping. En: Proceedings of the IEEE International Joint Conference on Neural Networks, Vol. 1, pp. 413-418.
- [31] Lahdenojat O.; Maunu J.; Laiho M. y Paasio A., 2006. A massively parallel algorithm for local binary pattern based face recognition. En: Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '06), Kos, Greece, 2006, 4 P.
- [32] Lai J.; Yuen P. C.; Chen W.; Lao S. y Kawade M., 2001. Robust facial feature point detection under nonlinear illuminations. En: Proceedings of the IEEE Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems (RATFG-RTS'01), 168 P.
- [33] Li C. y Barreto A., 2006. An integrated 3D face-expression recognition approach. En: Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006, May. 14-19, Vol. 3.
- [34] Li J.; Poulton G.; Guo Y. y Qiao R.-Y., 2003. Face recognition based on multiple region features. En: Proc. VIIth Digital Image Computing: Techniques and Applications, Sydney, pp. 69-78.
- [35] Lu J.; Yuan X. y Yahagi T., 2007. A method of face recognition based on fuzzy c-means clustering and associated sub-nns. En: IEEE Transactions on Neural Networks, Vol. 18, No. 1, pp. 150-160.



- [36] Luo B.; Zhang Y. y Pan Y., 2005. Face recognition based on wavelet transform and SVM. En: Proceedings of the 2005 IEEE International Conference on Information Acquisition, June 27 - July 3, Hong Kong and Macau, China, pp. 373-377.
- [37] Meng K.; Su G.; Li C.; Fu B. y Zhou J., 2005. A high performance face recognition system based on a huge face database. En: Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, Aug. 18-21, Vol.8, pp. 5159-5164.
- [38] Ming A.; Ma H.; Zhang H. y Xiong W., 2005. A grid-based face recognition approach using general template matching. En: Proceedings of the First International Conference on Semantics, Knowledge, and Grid (SKG 2005), Nov. 27-29.
- [39] Morizet N.; Amiel F.; Hamed I. D. y Ea T., 2007. A comparative implementation of PCA face recognition algorithm. En: 14th IEEE International Conference on Electronics, Circuits and Systems (ICECS '07), Dec. 11-14, pp. 865-868.
- [40] Nastar C. y Mitschke M., 1998. Real-time face recognition using feature combination. En: Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, April 14-16, Nara, Japan, pp. 312-317.
- [41] Naval Jr. P. C., 2006. Recognizing faces using kernel eigenfaces and support vector machines. En: Philippine Computing Journal, Vol. 1, March.
- [42] Nazeer S. A.; Omar N. y M. Khalid, 2007. Face recognition system using artificial neural networks approach. En: Proceedings of ICSCN 2007: International Conference on Signal Processing Communications and Networking, Feb. 22-24, Chennai, India, pp. 420-425.
- [43] Phillips P. J.; Grother P.; Micheals R.; Blackburn D. M.; Tabassi E. y Bone M., 2003. Face recognition vendor test 2002. En: Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'03), pp. 44.
- [44] Phillips P. J.; Moon H.; Rauss P. y Rizvi S. A., 2000. The FERET evaluation methodology for face-recognition algorithms. En: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, Oct., pp. 1090-1104.
- [45] Phillips P. J.; Scruggs W. T.; O'Toole A. J.; Flynn P. J.; Bowyer K. W.; Schott C. L. y Sharpe M., 2006. FRVT 2006 and ICE 2006 large-scale results. En: Face recognition vendor test 2006, [En línea] Disponible en: <http://www.frvt.org/FRVT2006/Results.aspx>
- [46] Pighin F.; Hecker J.; Lischinski D.; Szeliski R. y Salesin D. H., 1998. synthesizing realistic facial expressions from photographs. En: Proceedings of the 25th annual conference on Computer Graphics and Interactive Techniques, pp. 75-84.
- [47] Sahoozadeh A. H.; Heidari B. Z. y Dehghani C. H., 2008. A new face recognition method using PCA, LDA and neural network. En: Proceedings of World Academy of Science, Engineering and Technology, Vol. 31, pp. 218-223.
- [48] Seow M. J.; Gottumukkal R.; Valaparla D. y Asari K.V., 2004. A robust face recognition system for real time surveillance. En: Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'04), Vol. 1, pp. 631-635.
- [49] Smith L. I., 2002. A tutorial on Principal Components Analysis. [En línea]. Disponible en: [www.cs.otago.ac.nz/cosc453/student\\_tutorials/](http://www.cs.otago.ac.nz/cosc453/student_tutorials/)
- [50] Torres L. y Reutter Lorente Y., 1999. The importance of the color information in face recognition. En: Proceedings of International Conference on Image Processing, ICIP 99, Vol. 3, pp. 627-631.
- [51] Turk M. y Pentland A. P., 1991. Face recognition using eigenfaces. En: IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, pp. 586-591.
- [52] Viola P. y Jones M. J., 2004. Robust real-time face detection. En: International Journal of Computer Vision, 57(2), pp. 137-154.
- [53] Wang S.; Zhang L. y Samaras D., 2005. Face reconstruction across different poses and arbitrary illumination conditions. En: AVBPA 2005: Audio- and video-based biometric person authentication, No. 5, Hilton Rye Town NY, ETATS-UNIS, Vol. 3546, pp. 91-101.
- [54] Xie B.; Boulton T.; Ramesh V. y Zhu Y., 2006. Multi-camera face recognition by reliability-based selection. En: Proceedings of the 2006 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, Oct. 16-17, pp. 18-23.
- [55] Yang A. Y.; Wright J.; Ma Yi y Sastry S. S., 2007. Feature selection in face recognition: a sparse representation perspective. En: Manuscript submitted to IEEE Transactions on PAMI.
- [56] Yang M.-H.; Ahuja N. y Kriegman D., 2000. Face recognition using kernel eigenfaces. En: Proceedings of International Conference on Image Processing, Vol. 1, pp. 37-40.
- [57] Yazdi H. S.; Shargh A. E.; Gelyan S. N.; Deldari H.; Nasiri J. A. y Moulavi M. A., 2008. An efficient parallel eye detection algorithm on color facial images. En: Proceedings of the Ninth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, pp. 706-711.
- [58] Zheng W.; Zou C. y Zhao L., 2004. real-time face recognition using Gram-Schmidt orthogonalization for LDA. En: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04), Vol. 2, pp. 403 - 406.
- [59] Zou J.; Ji Q. y Nagy G., 2007. A comparative study of local matching approach for face recognition. En: IEEE Transactions on Image Processing, Vol. 16, No. 10, Oct., pp. 2617-2628.
- [60] Zuo F. y de With P. H. N., 2005. Real-time face recognition for smart home applications. En: International Conference on Consumer Electronics, (ICCE 2005), Jan. 8-12, pp. 35- 36.

**125** AÑOS  
1887 - 2012  
**Ingeniería con Trabajo y Rectitud**



UNIVERSIDAD  
**NACIONAL**  
DE COLOMBIA  

---

SEDE MEDELLÍN  
FACULTAD DE MINAS