

JOSÉ AUGUSTO CADENA MOREANO

Universidad Técnica de Cotopaxi
jose.cadena@utc.edu.ec

GALO ALFREDO FLORES LAGLA

Universidad Técnica de Cotopaxi
galo.flores@utc.edu.ec

MANUEL WILLIAM VILLA QUISHPE

Universidad Técnica de Cotopaxi
manuel.villa@utc.edu.ec

JUAN CARLOS CHANCÚSIG CHISAG

Universidad Técnica de Cotopaxi
juan.chancusig@utc.edu.ec

OSCAR ALEJANDRO GUAYPATÍN PICO

Universidad Técnica de Cotopaxi

Recibido: 05/02/2018

Aprobado: 06/06/2018

Resumen

En este artículo, se describen los principales métodos de reconocimiento facial en sub-espacios lineales: Eigenfaces, Fisherfaces, método Bayesiano, Análisis de componentes Independientes (ICA), y Tensorfaces; y en sub-espacios no lineales: Curvas Principales y PCA no lineales, y Kernel-PCA, así como también el método de clasificación de Máquinas de Vectores de Soporte (SVM). Además, se presentan las bases de datos de rostros construidas y empleadas para la evaluación del desempeño de los diferentes métodos de reconocimiento facial, tales como: Yale, Yale extendida B, CAS-PEAL, FERET, y ORL.

Palabras clave: reconocimiento facial, biométrico, espacios lineales, espacios no lineales, base de datos.

FACIAL RECOGNITION IN SUB-SPACES: LINEAR AND NON-LINEAR, FACE DATA BASES AND SUPPORT VECTOR MACHINE

Abstract

In this article, the main methods of facial recognition in linear sub-spaces are described: Eigenfaces, Fisherfaces, Bayesian method, and Independent Component Analysis (ICA); And in non-linear sub-spaces: Main Curves and non-linear PCA, and Kernel-PCA, as well as the classification method of Support Vector Machines (SVM). Also, the databases of faces constructed and used for the evaluation of the performance of different facial recognition methods, such as Yale, Yale extended B, CAS-PEAL, FERET, and ORL, are presented.

Keywords: face recognition, biometric, linear spaces, non-linear spaces, database.

Introducción

El reconocimiento facial es un método biométrico empleado para la identificación de individuos mediante las características de las caras humanas. Este método es aplicado en múltiples disciplinas con fines de seguridad o comerciales.

Según Chen y col (2017, p. 1767), las investigaciones del reconocimiento facial automatizado comenzaron con el análisis de patrones faciales, basándose en las características estructurales y geométricas de los rostros, esta premisa, ha sido desde entonces uno de los basamentos más significativos en la generación de algoritmos de reconocimiento facial, así mismo, estos autores, destacan que en este tema, se han realizado grandes avances en los últimos años, implementándose progresivamente un gran número de sistemas de reconocimiento y modelado facial.

Según Stan y Jain (2011, p. 1), el primer sistema automatizado de reconocimiento facial fue desarrollado en la tesis doctoral en 1973 por el Ph.D. Takeo Kanade. En 1987, Sirovich y Kirby enfocaron el estudio del reconocimiento facial en la técnica de Análisis de Componentes Principales (PCA), el cual fue el trabajo base implementado por M. A. Turk y A. P. Pentland en 1991 para el desarrollo del método Eigenfaces.

Otro de los mayores hitos en el reconocimiento facial, es el método de Fisherface propuesto por Belhumeur et al., el cual aplica el Análisis de Componentes Principales (PCA) y posteriormente el Discriminante Lineal de Fisher (FLD), lo cual maximiza la relación entre la distribución entre clases y la distribución intra-clases. “En todos los experimentos Fisherfaces reportados alcanzaron una

menor tasa de error que los experimentos Eigenfaces” (Stan & Jain, 2011, p. 2).

Moghaddam, en 1999 fue pionero en proponer la aplicación del método bayesiano en la generación de algoritmos de identificación facial, basado este en el teorema del matemático Thomas Bayes, y en el cual, se realiza una inferencia estadística usando un método PCA eficiente (Moghaddam, 1999, p. 1).

Posteriormente, se desarrollaron métodos no lineales, que conllevan unos cálculos más complejos, pero resultados más mesurados serán descritos en el presente artículo, tales como: curvas principales y PCA no lineal; y Método de Kernel-PCA y Kernel-Fisher (Deshpande & Ravishankar, 2017, p. 1174).

Además también se expone el método de clasificación de Máquinas de Vectores de Soporte (SVM), para sub-espacios lineales y no lineales.

Para evaluar el desempeño de los métodos anteriormente mencionados, se construyeron bases de datos de rostros.

Categorización

Según Crosswhite y col. (Crosswhite, y otros, 2017, p. 1), un sistema de reconocimiento facial, como un sistema biométrico, puede operar en uno o ambos de los siguientes modos:

- Verificación de rostro (o autenticación): implica una comparación uno a uno, de los parámetros de una imagen de un rostro de consulta contra una imagen de un rostro almacenado en una base de datos, el cual verifica si la identidad corresponde a la analizada.
- Identificación de rostro (o reconocimiento): implica una comparación de un rostro de consulta con múltiples rostros almacenados en la base de datos, para así asociar la identidad del rostro de

consulta a alguno de los almacenados en la base de datos.

Sub-espacio de rostros

A pesar de los significativos avances en el área de reconocimiento facial, aun representan un gran esfuerzo los análisis donde el punto de vista, iluminación, expresión, oclusión y los accesorios faciales varían considerablemente (Rajawat, Pandey, & Rajput, 2017, p. 2). Esto puede ser ilustrado a través de un sub-espacio de rostros o Manifolds (un espacio matemático abstracto que se parece a los espacios descritos por la geometría euclídea), en el cual se pueden encontrar características únicas en los individuos, de fácil identificación y comparación, que dado a esto pueden facilitar el proceso de reconocimiento facial (Stan & Jain, 2011, p. 5).

El uso de las técnicas de modelado subespacial ha representado un importante avance en la tecnología de reconocimiento facial. En la figura 1(a) se ilustra el sub-espacio de rostros en contraste con el sub-espacio sin rostro y en la figura 1(b) se ilustra un par de individuos en el sub-espacio de rostros. El sub-espacio de rostros cuenta con las aparariencias faciales mientras que el sub-espacio sin rostro abarca todos los objetos que no pertenecen al rostro. Por tanto, al analizar los sub-espacios, estos son no lineales y no convexos.

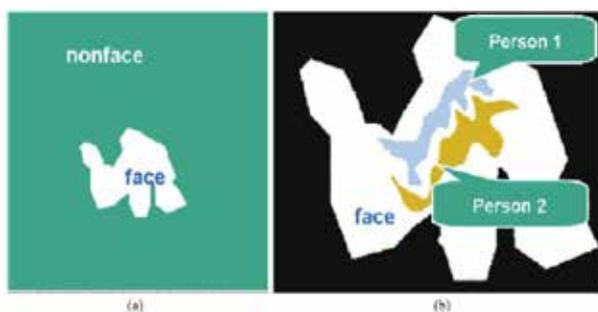


Figura 1. Sub-espacio de rostros. (a) Sub-espacio de rostros vs. Sub-espacio sin rostro. (b) Sub-espacio rostro de los diferentes individuos.

Fuente: Stan & Jain, 2011, p. 5.

De acuerdo a lo expuesto anteriormente, la detección de rostros es la acción de diferenciar entre el sub-espacio de rostros y el sub-espacio sin rostros en el espacio de la imagen de entrada, mientras que el reconocimiento de rostros se considera la acción de diferenciar entre los rostros de la variedad de individuos en el sub-espacio de los rostros.

Espacio de rostros y su dimensionalidad

Según Stan y Jain, el análisis computarizado de las imágenes de los rostros trata con una señal visual (reflejo de la luz sobre la superficie de un rostro) que es registrado digitalmente por un sensor como un matriz de valores de píxeles. Los píxeles pueden codificar color o solo intensidad, es por ello por lo que en el presente artículo se asumen imágenes en escala de grises. Para especificar una imagen determinada en el espacio de rostros, es necesario definir cada valor de los píxeles. Por tanto, la dimensionalidad del espacio está dado por la representación de los píxeles $m \times n$. Los métodos de reconocimiento que operan en esta representación sufren potenciales desventajas, la mayoría de ellos arraigados en la llamada condensación de la dimensionalidad (Stan & Jain, 2011, p. 19).

Desafíos tecnológicos

Como se expuso en la sección anterior, el problema de la detección de rostros es altamente no lineal y no convexo, aún más el de reconocimiento de rostros. A continuación, se mencionan los principales desafíos tecnológicos en el sistema automático de reconocimiento facial:

- Inmensa variedad de apariencias faciales.
- Sub-espacios complejos no lineales.
- Alta dimensionalidad y una pequeña muestra.

Sub-espacios lineales

En esta sección se describirán los métodos lineales y su generalidad, asumiendo el caso de que el sub-espacio

de los rostros es un sub-espacio lineal del espacio de la imagen.

Eigenfaces

El método de Eigenfaces, introducido por M.A. Turk y A.P. Pentland, está construido sobre la técnica de Análisis de Componentes Principales (PCA) (Stan & Jain, 2011, p. 2), en el cual se "...toma una cantidad menor de los vectores entregados por las imágenes de la base de datos pero con información necesaria para la reconstrucción de los rostros de las imágenes ingresadas". (Franco, Ospina, Cuevas, & Capacho, 2015).

Para realizar el cálculo mediante este método, se considera una imagen de rostro $I(x,y)$ de dos dimensiones N por N matricial de valores de intensidad, o un vector de dimensión N^2 . Estos vectores definen el sub-espacio de las imágenes de los rostros. Se asume que cada una de las imágenes pertenece a una de las c clases $\{X_1, X_2, \dots, X_c\}$ (Franco, Ospina, Cuevas, & Capacho, 2015, p. 146). Los nuevos vectores de características $y_k \in \mathbb{R}^m$ son definidos por la siguiente transformación lineal:

$$y_k = W^T x_k \quad k = 1, 2, \dots, N \quad (1)$$

Donde $W \in \mathbb{R}^{n \times m}$ representa una matriz ortonormal. Se define además la matriz de distribución total S_T como (Saleh, Azam, Yeo, Shanmugam, & Kannoorpatti, 2017, p. 116)":

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T \quad (2)$$

Donde $\mu \in \mathbb{R}^n$ es la medida de todas las imágenes de I . Posteriormente se aplica la transformación lineal W^T , la distribución de los vectores de características es $W^T S_T W$. Se toma aquella proyección W_{opt} que maximiza el determinante de la distribución total de la matriz de las imágenes proyectadas, obteniendo w_i donde $\{w_i / i = 1, 2, \dots, m\}$ "es el conjunto de

vectores propios n -dimensionales de S_T correspondientes a los mayores m vectores propios. Dichos vectores propios tienen la misma dimensión que las imágenes originales y se les denomina eigengaces" (Saleh, Azam, Yeo, Shanmugam, & Kannoorpatti, 2017, p. 116).

Análisis de discriminante lineal: Fisherfaces

Este método propuesto inicialmente por Belhumeur y Kriegman, contempla la aplicación inicial del Análisis de Componentes Principales (PCA) y posteriormente el Discriminante Lineal de Fisher (FLD) (Franco, Ospina, Cuevas, & Capacho, 2015, p. 148).

Según Stan y Jain (2011, p. 26), cuando hay cambios en iluminación y expresiones en el rostro, generan muchas diferencias en los datos debido a dichos cambios. Las técnicas PCA seleccionan un sub-espacio que retiene la mayor parte de esa variación, y consecuentemente la similitud en el sub-espacio de la cara no es necesariamente determinada por la identidad.

Es por ello, que para resolver el problema se aplica el método de Fisherfaces, aplicando el Discriminante Lineal de Fisher (FLD). FLD selecciona el sub-espacio lineal ϕ el cual maximiza la tasa siguiente:

$$\frac{|\phi^T S_b \phi|}{|\phi^T S_w \phi|} \quad (3)$$

Donde la matriz de dispersión entre clases es:

$$S_b = \sum_{i=1}^m N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (4)$$

La matriz de dispersión dentro de la clase viene dada por la siguiente ecuación y m es el número de sujetos (clases) en la base de datos:

$$S_w = \sum_{i=1}^m \sum_{x \in X_i} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \quad (5)$$

“El Discriminante Lineal de Fisher encuentra la proyección de los datos en los que se encuentran las clases son más linealmente separables. Se puede demostrar que la dimensión ϕ es como máximo $m-1$. Porque en la práctica S_w es usualmente singular, el algoritmo de Fisherfaces primero reduce la dimensionalidad de los datos con PCA y pueden ser computados y entonces es aplicada el FLD para reducir la dimensionalidad a $m-1$. El reconocimiento se realiza entonces mediante un clasificador NN en este subespacio final”. (Stan & Jain, 2011, p. 27)

Método Bayesiano

La formulación Bayesiana fue propuesta por Moghaddam en 1999, y la misma proyecta la tarea estándar de reconocimiento facial en un problema de clasificación de patrones binarios con Ω_I y Ω_E (1999, p. 1), al respecto, el autor postuló lo siguiente:

“Considerando un espacio característico de Δ vectores, la diferencia entre dos imágenes ($\Delta = I_j - I_k$). Se pueden definir dos clases de variaciones de la imagen facial: variaciones intrapersonales Ω_I (correspondientes, por ejemplo, a diferentes expresiones faciales e iluminaciones del mismo individuo) y variaciones extra personales Ω_E (correspondientes a variaciones entre individuos diferentes)” (Stan & Jain, 2011, p. 27).

Según Stan y Jain, la medida de similitud $S(\Delta)$ puede entonces expresarse de la siguiente manera, tomando en cuenta el Teorema de Bayes (2011, p. 27).

$$S(\Delta) = P(\Omega_I|\Delta) = \frac{P(\Delta|\Omega_I)P(\Omega_I)}{P(\Delta|\Omega_I)P(\Omega_I)+P(\Delta|\Omega_E)P(\Omega_E)} \quad (6)$$

Las densidades de ambas clases son modeladas como Gaussianos de alto dimensional, empleando como base un método PCA eficiente.

$$P(\Delta|\Omega_E) = \frac{e^{-\frac{1}{2}\Delta^T \Sigma_E^{-1} \Delta}}{(2\pi)^{D/2} |\Sigma_E|^{1/2}} \quad (7)$$

$$P(\Delta|\Omega_I) = \frac{e^{-\frac{1}{2}\Delta^T \Sigma_I^{-1} \Delta}}{(2\pi)^{D/2} |\Sigma_I|^{1/2}} \quad (8)$$

Estas densidades son cero, dado que por cada $\Delta = I_j - I_i$ existe un $I_i - I_j$.

Para calcular las probabilidades $P(\Delta|\Omega_I)$ y $P(\Delta|\Omega_E)$, la base de datos de imágenes I_j son procesados con una transformación de blanqueamiento. Cada imagen es convertida y almacenada como un conjunto de dos coeficientes de sub-espacios blanqueados: $y\phi_I$ para el espacio intrapersonal y $y\phi_E$ para el espacio extrapersonal.

Luego de procesar los valores, los Gaussianos pueden ser reducidos a una simple distancia Euclidiana (los denominadores son pre-calculados) como se muestran en las siguientes ecuaciones. Estas probabilidades son examinadas y empleadas para determinar el máximo a posteriori (MAP) en la ecuación (6).

$$P(\Delta|\Omega_E) = P(I - I_j|\Omega_E) = \frac{e^{-\|y\phi_E - y^j\phi_E\|^2/2}}{(2\pi)^{k_E/2} |\Sigma_E|^{1/2}} \quad (9)$$

$$P(\Delta|\Omega_I) = P(I - I_j|\Omega_I) = \frac{e^{-\|y\phi_I - y^j\phi_I\|^2/2}}{(2\pi)^{k_I/2} |\Sigma_I|^{1/2}} \quad (10)$$

La máxima probabilidad (ML) de coincidencia de similitud es aún más simple, ya que solo se evalúa la clase intrapersonal, dando lugar a la siguiente forma modificada para la medida de similitud:

$$S'(\Delta) = P(\Delta|\Omega_I) = \frac{e^{-\|y\phi_I - y^j\phi_I\|^2/2}}{(2\pi)^{k_I/2} |\Sigma_I|^{1/2}} \quad (11)$$

En la figura 2(b), se puede observar el flujo computacional del Método Bayesiano. Los pasos de proyección son lineales mientras que los posteriores no son lineales.

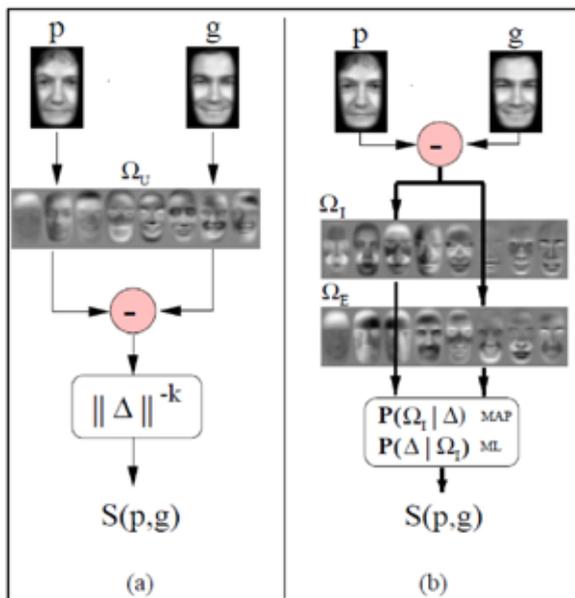


Figura 2. Diagrama de flujo para calcular la similitud g entre dos imágenes. (a) Método Eigenfaces original. (b) Similitud Bayesiana. La imagen diferente es proyectada a través de ambos conjuntos de (intra-personal/extrapersonal) eigenfaces para obtener dos probabilidades.

Fuente: Moghaddam, 1999, p. 7.

Análisis de componentes independientes (ICA)

Según Stan y Jain, el objetivo de este método es minimizar las dependencias de orden superior, los componentes encontrados por ICA están diseñados para ser no gaussianos. Mientras que PCA minimiza la covarianza de la muestra (dependencia de segundo orden) de los datos. Como las técnicas PCA, ICA produce una proyección lineal $R^n \rightarrow R^m$ pero con propiedades diferentes, aproximando la no ortogonalidad de la base A (Khedher, Illán, Górriz, Ramírez, & Brahim, 2017, p. 4).

$$x \approx Ay \quad A^T A \neq I \quad P(y) \approx \prod p(y_i) \quad (12)$$

En la figura 3, se observa un ejemplo del cálculo de un conjunto de puntos 3D. El sub-espacio 2D recuperado por la técnica ICA refleja la distribución de los datos de mejor manera en comparación a la técnica PCA.

En la figura 4, se puede observar un ejemplo básico de la técnica ICA, en la cual se muestra dos vectores IC no ortogonales, desordenados, uno de los cuales están aproximadamente alineados con el primer vector del componente principal.

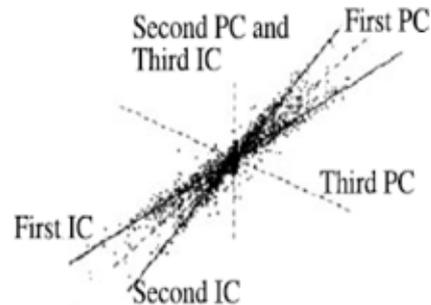


Figura 3. ICA vs. PCA descomposición de un conjunto datos 3D. Las bases de PCA (ortogonal) e ICA (no ortogonal).

Fuente: Shakhnarovich & Moghaddam, 2011, p. 30.



Figura 4. ICA vs. PCA descomposición de un conjunto de datos 3D. Izquierda: la proyección de los datos sobre los dos principales componentes (PCA). Derecha: la proyección de datos sobre los dos componentes independientes (ICA).

Fuente: Shakhnarovich & Moghaddam, 2011, p. 30.

Multilinear SVD: Tensorfaces

De acuerdo a Stan y Jain, los métodos de análisis lineales expuestos en las secciones anteriores han demostrado ser adecuados cuando la pose, la iluminación o la expresión son fijadas a través de la base de datos de los rostros. Cuando algunos de estos aspectos varían, la representación lineal del subespacio no captura esta

variación de la mejor forma (Li & Jain, 2015, p. 907). Vasilescu y Terzopoulos (2001, p. 3), propusieron una alternativa multilínea, llamado Tensorfaces. El álgebra multilínea, ofrece un marco matemático para analizar el multi-factor estructural del conjunto de imágenes y para abordar el complicado problema de los factores constituyentes. Esta técnica emplea un tensor extensión de la descomposición del valor singular de la matriz convencional (SDV), conocida como “N-mode SDV”, la cual descompone un tensor n-dimensional A como se muestra en la ecuación 13.

$$A = L \times_1 U_1^T \times_2 U_2^T \dots \times_N U_N^T \quad (13)$$

El rol del tensor central L en esta descomposición es similar al rol en la matriz de valor singular (SDV).

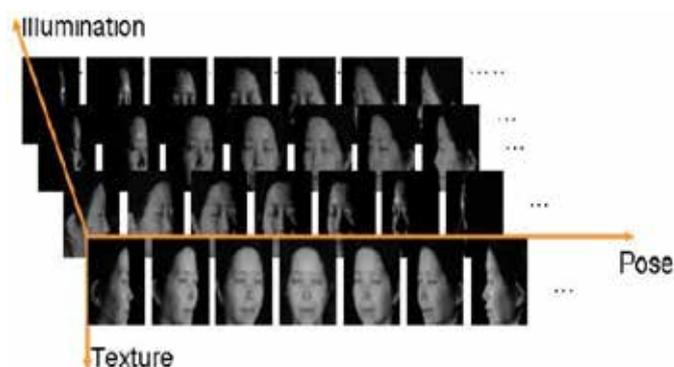


Figura 5. Ejemplo de Tensorfaces.

Fuente: Xu Q., 2017.

El tensor noción se aplica en el conjunto de imágenes de los rostros, considerando un conjunto de imágenes de N_p rostros de personas, cada fotografiado en N_v puntos de vista, con N_i iluminaciones y N_e expresiones faciales. Esto se ejemplifica, en la figura 5, solo con un tensor de 4 dimensiones (Stan & Jain, 2011, p. 32).

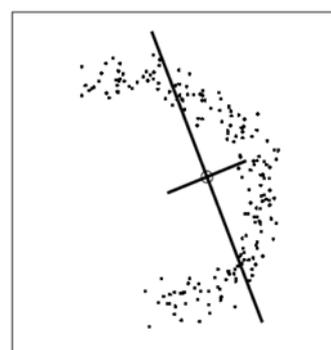
Sub-espacios no lineales

En esta sección se describirán las técnicas que no asumen que el espacio principal es lineal.

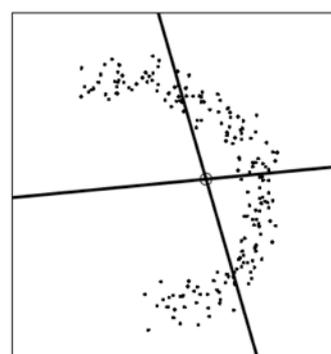
Curvas principales y PCA no lineales

De acuerdo a Stan y Jain, la propiedad definitoria de los sub-espacios principales no lineales o curvas principales, es que la imagen inversa del sub-espacio en el espacio original R^N es una superficie dimensional baja no lineal (curva) que pasa por el medio de los datos, lo cual disminuye la sumatoria total de la distancia entre los puntos de datos y sus proyecciones sobre esa superficie (Howard, Apley, & Runger, 2017, p. 2).

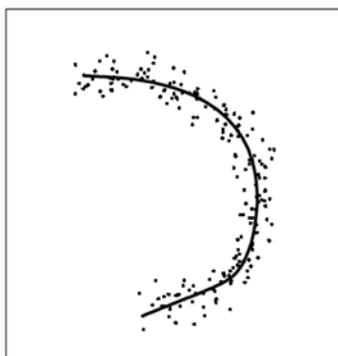
En la figura 6(c), se puede observar un ejemplo de una curva principal.



(a)



(b)



(c)

Figura 6. (a) PCA básico (lineal, ordenado y ortogonal). (b) ICA básico (lineal, desordenado y no ortogonal). (c) Curva principal (sub-espacio no lineal parametrizado).

Fuente: Moghaddam, 1999, p. 6

Uno de los métodos más simples para el cálculo de los sub-espacios es el PCA no lineal de redes neuronales multicapa como se muestra en la figura 7.

Según Shanmuganathan (2016, p. 4) las Redes Neuronales Artificiales: “en términos simples es un modelo computacional biológicamente inspirado, que consiste en procesar elementos (llamados neuronas), y conexiones entre ellos con coeficientes (pesos) enlazados a las conexiones. Estas conexiones constituyen la estructura neuronal y se unen a esta estructura son los algoritmos de entrenamiento y recuperación. Las redes neuronales se denominan modelos conexionistas debido a las conexiones encontradas entre las neuronas” (2004, p. 6).

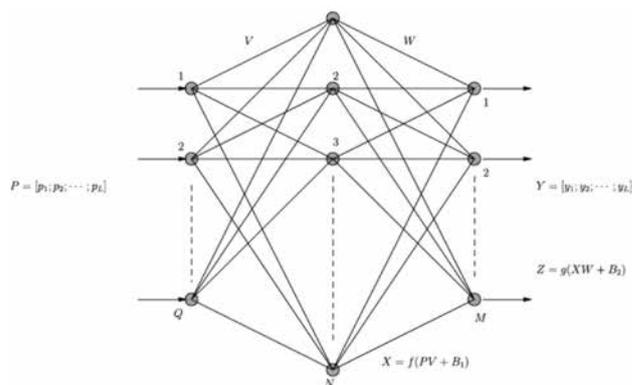


Figura 7. Estructura de una Red Neuronal con modelo para cálculo de sub-espacios principales $\in \mathbb{R}^k$ en el espacio de entrada $x \in \mathbb{R}^N$.

Fuente: Shanmuganathan, 2016, p. 73).

Método de Kernel-PCA

Según Stan y Jain, el análisis de componentes principales no lineal ha resurgido gracias a Schölkopf, por el método de eigenvalor de kernel. La metodología básica del método de Kernel-PCA es aplicar un mapeo no lineal de entrada $\varphi(x): \mathbb{R}^N \rightarrow \mathbb{R}^L$, y posteriormente solucionarlo a través de la técnica PCA lineal obteniendo un espacio característico \mathbb{R}^L , donde L es mucho mayor que N y posiblemente infinito. El mapeo $\varphi(x)$ está hecho a través de la función de kernel (2011, p. 35).

$$k(x, y) = [\varphi(x) \cdot \varphi(y)] \quad (14)$$

Donde $k(x, y)$ en el espacio de entrada corresponde al producto punto en el espacio característico dimensional alto. Asumiendo que la proyección de los datos en un espacio característico es centrada, entonces la covarianza se puede expresar como:

$$\Sigma K = [\varphi(x) \cdot \varphi(y)^T] \quad (15)$$

Resultando la ecuación del eigenvector $\lambda V = \Sigma K V$, por lo cual, existe un coeficiente w_i .

$$V = \sum_{i=1}^T w_i \varphi(x) \quad (16)$$

Usando la definición de la ecuación 14, y la ecuación 15, se define el resultado como una matriz $K K_{ij} = [\varphi(x) \varphi(y)]$ con dimensión T por T conduciendo al problema del eigenvalor equivalente formulado en términos de Kernels en el espacio de entrada.

$$T \lambda w = K w \quad (17)$$

Donde $w = (w_1, w_2, \dots, w_T)$ es el vector de coeficientes de expansión de un determinado eigenvector V como se define en la ecuación 16.

“Posteriormente, los componentes principales de KPCA de cualquier vector de entrada pueden ser calculados eficientemente con evaluaciones de kernel simples contra el conjunto de datos”. (Stan & Jain, 2011, p. 36).

El enésimo componente principal y_n de x está dado por la ecuación 18.

$$y_n = (V_n \cdot \varphi(x)) = \sum_{i=1}^T w_i^n k(x, x_i) \quad (18)$$

Donde V_n es el enésimo eigenvector del espacio característico definido por φ .

En la figura 8, se puede observar gráficamente la idea básica del método de Kernel-PCA. “En algún espacio característico dimensional alto F (Abajo a la derecha) se obtiene un PCA lineal, como un PCA en espacio de entrada (arriba). Sin embargo, F no está linealmente relacionado al espacio de entrada (vía Φ), las líneas de contorno de la proyección constante sobre el eigenvector principal (dibujado como flecha) van a ser no lineales en el espacio de entrada. Nótese que no se puede dibujar una pre-imagen del eigenvector en el espacio de entrada, porque puede que no exista” (Schölkopf, Smola, & Müller, 1998, p. 1.306).

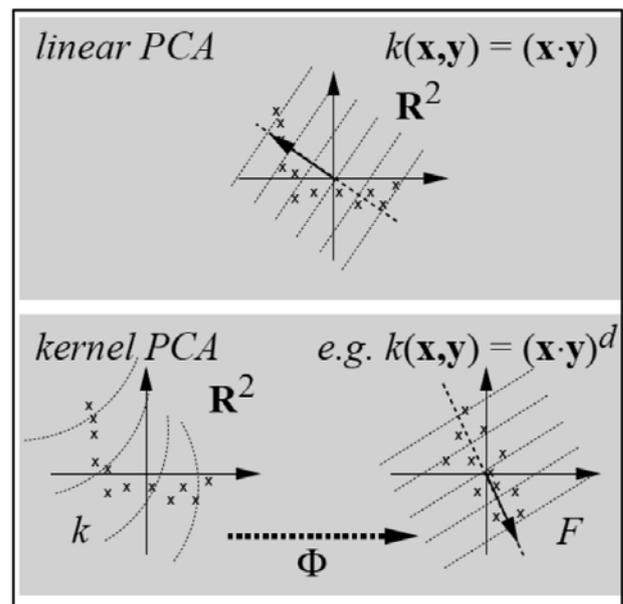


Figura 8. La idea básica del método de Kernel-PCA.

Fuente: Schölkopf, Smola, & Müller, 1998, p. 1.306.

Principales bases de datos de reconocimiento de rostros

En la presente sección, se expondrán las principales bases de datos de rostros desarrolladas, y las consideraciones técnicas tomadas durante la captura de las imágenes.

Base de datos Yale

Belhumeur, Hespanha y Kriegman en 1997, construyeron la base de datos de rostros Yale, la cual está conformada por 165 imágenes en escala de grises de 15 individuos (11 imágenes por individuo), los cuales contemplan diferentes expresiones faciales, iluminación y uso de lentes. En la figura 9, se muestra la sesión de imágenes completa para un individuo de estudio de la base de datos Yale (Wang, y otros, 2014, p. 5).

Base de datos Yale extendida B

En 2001, Georhiades, Belhumeur y Kriegman (2001, p. 644) ampliaron el trabajo previo de la base de datos Yale. La base de datos de rostros Yale Extendida B, contiene 16.128 imágenes de 640 X 480 en escala de

grises de 28 individuos bajo 9 poses y 64 condiciones de iluminación. Además, se incluyó en la base de datos un conjunto de imágenes recortadas con únicamente el rostro de los individuos como se muestra en la figura 10.



Figura 9. Ejemplo de sesión completa para un individuo de la base de datos Yale.

Fuente: Wang, Lu, Wang, Li, Yan, & Hu, 2014, p. 6.



Figura 10. Ejemplo de base de datos de rostros de Yale Extendida B.

Fuente: Chan, Tahir, Kittler, & Pietikäinen, 2013, p. 1.170.

Base de datos CAS-PEAL

En 2004, Cao y Shan construyeron una base de datos de rostros que contiene 99.594 imágenes de 1040 individuos, de los cuales 595 son hombres y 445 son mujeres, en donde se tomaron las imágenes con diferentes poses, expresiones faciales, iluminación y accesorios (Gao, Ma, Zhao, & Yuille, 2017, p. 10).

Se instalaron nueve cámaras en semicírculo, con el fin de capturar simultáneamente en pose a cada individuo. Se tomaron 18 imágenes con el rostro hacia arriba y hacia abajo, además se consideraron 5 diferentes expresiones faciales, 6 accesorios (3 lentes y 3 gorras) y 15 orientaciones de iluminación. En la

figura 11, se muestran ejemplos de la base de datos CAS-PEAL.



Figura 11. Ejemplo de imágenes capturadas en la base de datos de rostros CAS-PEAL.

Fuente: Gao, Ma, Zhao, & Yuille, 2017, p. 11.

Base de datos FERET

Entre 1993 y 1996, P.J. Phillips, H. Moon, P. Rauss y S.A. Rivisi (1997, p. 141) construyeron la base de datos FERET (Face Recognition Technology), la cual consiste en 1196 individuos y cuatro clases de conjunto de pruebas: fafb (imágenes con diferentes expresiones faciales), fafc (imágenes con diferentes orientaciones de iluminación), dupI (imágenes tomadas en menos de 18 meses) y dupII (imágenes tomadas en los 18 meses siguientes). El programa (FERET) es administrado por la Agencia DARPA (Defense Advanced Research Projects Agency) y (NIST) (National Institute of Standards and Technology).

En la figura 12, se muestra un pequeño ejemplo de las imágenes capturadas en la base de datos FERET.



Figura 12. Ejemplo de imágenes de la base de datos de rostros FERET.

Fuente: Lu, Tan, & Wang, 2013, p. 45.

Base de datos ORL

En 1994, un grupo de trabajo de los laboratorios AT&T, conformaron la base de datos ORL conformada por 10 imágenes diferentes de 40 individuos. Las imágenes fueron capturadas en distintos periodos de tiempo, ángulos de iluminación, expresiones faciales y acceso-

rios. Las dimensiones de las imágenes son 92 x 112 píxeles cada una, en escala de grises. En la figura 13, se observa un ejemplo de las imágenes capturadas para un individuo (Xu, y otros, 2014, p. 1954).



Figura 13. Ejemplo de imágenes de la base de datos de rostros ORL.

Fuente: Xu, y otros, 2014, p. 1.955.

Máquinas de Vectores de Soporte (SVM)

Según Jia y Martínez, SVM (Support Vector Machines) son una de las técnicas más útiles en clasificación de problemas y un claro ejemplo es el reconocimiento facial (Ma & Guo, 2014, p. 97).

En principio, se expondrá la idea básica del método SVM para clasificación binaria y luego se explicará cómo esta técnica puede ser aplicada a un problema de clasificación multi-clases.

Clasificación binaria

Las Máquinas Virtuales de Soporte (SVM) pertenecen a la clase de los clasificadores de máximo margen. Realizan el reconocimiento de patrones entre dos clases encontrando una superficie de decisión que tiene la distancia máxima a los puntos más cercanos en el conjunto de entrenamiento que se denominan vectores de soporte (Hamedi, Salleh, & Noor, 2015, p. 84).

Tomando un conjunto de entrenamiento de puntos x_i

en \mathbb{R}^n , donde $i = 1, 2, \dots, N$, y cada x_i pertenece a una de las dos clases identificadas por la etiqueta $y_i \in \{-1, 1\}$. Se asume que son datos linealmente separables, la meta de clasificación de margen máximo es separar las dos clases por un hiper-plano llamado Hiper-plano Óptimo de Separación (OSH), tal que, la distancia a los vectores de soporte es maximizada (Xiang, Yu, & Qu, 2014, p. 832). El Hiper-plano Óptimo de Separación es calculado a través de la ecuación 19.

$$f(x) = \sum_{i=1}^L \alpha_i y_i X_i \cdot X + b \quad (19)$$

Los coeficientes α_i y b , son soluciones al problema de programación cuadrática. Por consiguiente, se empleará la ecuación (20) para realizar la clasificación multi-clases.

$$d(x) = \frac{\sum_{i=1}^L \alpha_i y_i X_i \cdot X + b}{\|\sum_{i=1}^L \alpha_i y_i X_i\|} \quad (20)$$

Donde $|d|$ es la distancia desde x hasta el hiper-plano y el signo de d es el resultado de la clasificación por x .

Según Xiang y Qu, lo anteriormente expuesto, puede ser empleado para casos de superficies separadas no lineales. Cada punto x en el espacio de entrada es mapeado a un punto $z = \varphi(x)$ de un espacio característico dimensional alto, en el cual los datos son separados por un hiper-plano. El espacio característico puede ser expresado como una función de kernel $K(x, y)$, por lo cual la superficie de decisión se calcula a través de la ecuación 21 (Xiang, Yu, & Qu, 2014, p. 833):

$$f(x) = \sum_{i=1}^L \alpha_i y_i K(x, x_i) + b \quad (21)$$

Donde, los coeficientes α_i y b , son soluciones al problema de programación cuadrática, y $f(x)$ no depende de la dimensionalidad del espacio característico. La familia de funciones de kernel $K(x, y)$ es el polinomio de kernel, donde d es el grado del polinomio.

$$K(x, y) = (1 + x \cdot y)^d \quad (22)$$

Clasificación multi-clases

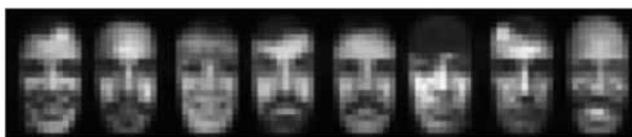
De acuerdo con Zhang y col. (2015, pp. 1.181-1.182), existen dos estrategias básicas para resolver problemas de q -clases a través del Método de Máquinas de Vectores de Soporte (SVM):

- i. En el enfoque uno-vs-todo q Máquinas de Vectores de Soporte (SVM) son entrenadas. Cada uno de los SVM separa una sola clase de todas las clases restantes.
- ii. En el enfoque de pares q $(q-1)/2$ máquinas son entrenadas. Cada SVM separa un par de clases. Los pares de clasificadores están dispuestos en árboles, donde cada nodo de árbol representa un SVM.

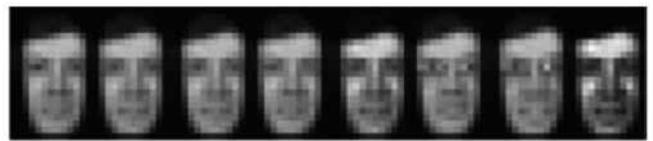
Comparación empírica de los métodos de sub-espacio

Reconocimiento basado en PCA

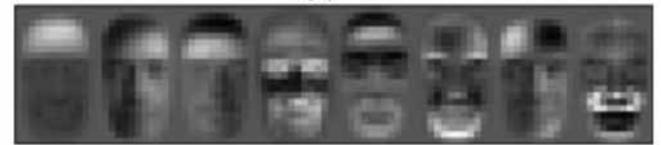
El algoritmo base para este experimento de reconocimiento facial fue una correspondencia estándar PCA (Eigenface). Los primeros ocho eigenvectores principales fueron calculados como se muestra en la figura 14(c), a partir de una partición simple. En la tabla 1, se muestra la proyección de la prueba realizada sobre el sub-espacio de lineal 20-dimensional calculado con PCA, obteniendo una tasa media de reconocimiento de 77,31%, alcanzado una tasa máxima de 79,62%. En la figura 15, se observa una tasa de reconocimiento del experimento de 86,46% (línea punteada). Claramente, el desempeño es degradado por la dimensionalidad $252 \rightarrow 20$ (Stan & Jain, 2011, p. 38).



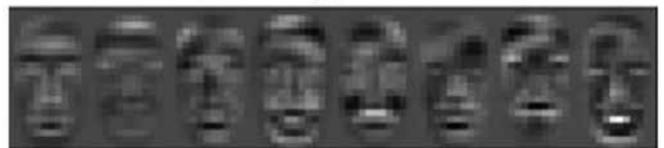
(a)



(b)



(c)



(d)

Figura 14. Experimento sobre la Base de datos FERET. (a) Gran cantidad de rostros de la galería. (b) Múltiples pruebas para un individuo, con diferentes expresiones faciales, lentes, luz ambiental variables, y contraste en la imagen. (c) Eigenfaces. (d) Imágenes bases de ICA.

Fuente: Stan & Jain, 2011, p. 37.

Reconocimiento basado en ICA

Según Stain y Jain, el reconocimiento ICA para este experimento, está basado en dos algoritmos: el algoritmo JADE y el algoritmo Hyvärinen y Oja. En ambos algoritmos se contempla un paso de blanqueo con la técnica PCA. La proyección del sub-espacio ICA del conjunto de pruebas fue obtenida usando $y = A^{-1}x$. En la tabla 1, se muestra la tasa media de reconocimiento de 77,30%, alcanzando una tasa máxima de 82,90% (2011, p. 39).

Reconocimiento basado en Kernel-PCA

Para el reconocimiento facial basado en el método Kernel-PCA, los parámetros de los núcleos gaussiano, polinomial y sigmoidal fueron ajustados primero para obtener el mejor rendimiento con un conjunto de validación de partición 50/50 resultando ser los mejores

kernels para este conjunto de datos. En la tabla 1, se muestra la tasa media de reconocimiento alcanzada es de 87,34%, y obteniendo una tasa máxima de 92,37%. En la figura 16, se puede observar que este método fue ligeramente superior a los métodos PCA e ICA (Stan & Jain, 2011, p. 39).

Reconocimiento basado en el método bayesiano

Para la correspondencia a través del método bayesiano, se formó adecuadamente Δ_s para las dos clases Ω_i (Fig. 15b) y Ω_E (Fig. 15a), donde se usaron una densidad dual estimada basada en PCA $P(\Delta | \Omega_i)$ y $P(\Delta | \Omega_E)$, en la cual ambos son modelados como un simple gaussiano con dimensiones sub-espaciales de k_i y k_E respectivamente, donde k fue dividido en dos densidades, dado por $k_i = k_E = k/2 = 10$ para el modelado.

Usando el cálculo del MAP, con el método bayesiano se obtuvo una tasa media de reconocimiento alcanzada de 94,83%, y obteniendo una tasa máxima de 97,87%, como se muestra en la tabla 1. En la figura 16, se puede observar que la desviación estándar para este algoritmo fue realmente baja.

Tabla 1. Precisiones de reconocimiento con $k = 20$ proyecciones sub-espaciales utilizando la validación cruzada de cinco veces. Los resultados son en porcentajes.

Partition	PCA	ICA	KPCA	Bayes
1	78.00	82.90	83.26	95.46
2	79.62	77.29	92.37	97.87
3	78.59	79.19	88.52	94.49
4	76.39	82.84	85.96	92.90
5	73.96	64.29	86.57	93.45
Mean	77.31	77.30	87.34	94.83
SD	2.21	7.66	3.39	1.96

Fuente: Stan y Jain, 2011, p. 39.

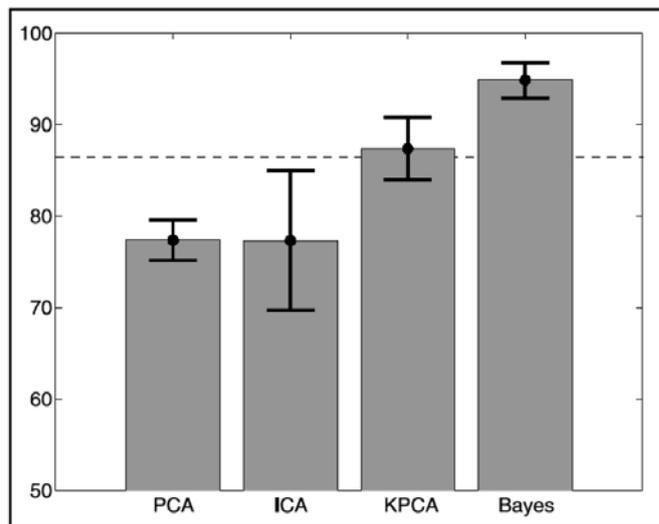


Figura 15. Desempeño de reconocimiento de los métodos PCA, ICA, y KPCA versus el método bayesiano (MAP) coincidiendo con un subespacio dimensional $k = 20$. La línea discontinua indica el rendimiento del más cercano-vecino que coincide con los vectores de imagen.

Fuente: Stan & Jain, 2011, p. 37.

Conclusión

El reconocimiento facial como método biométrico, ha conllevado el desarrollo de un conjunto de métodos para la elaboración de los sistemas automatizados, en los cuales se han considerado los sub-espacios lineales y los sub-espacios no lineales como metodologías matemáticas para considerar en el análisis de las imágenes aspectos como: poses, ángulos de iluminación, expresiones faciales y accesorios de los individuos sujetos a estudio.

Para la evaluación de dichos métodos, una gran cantidad de investigadores se han dado la tarea de construir bases de datos de rostros de diversas dimensiones en cuanto a los sujetos de estudio, tomando en consideración los aspectos mencionados anteriormente en las imágenes de entrada.

Para que la investigación en el área de reconocimiento facial continúe avanzando, se deben prolongar mejoras de la base de los métodos, y el incremento y progreso de las bases de datos, es por ello que son desarrollados en el presente artículo ambos temas.

Referencias

- Belhumeur, J., Hespanha, J., & Kriegman, D. (1997). Eigen faces vs. fisherfaces: recognition using class specific linear projection. *IEEE*, 711-720.
- Cabello, E. (2004). *Técnicos de reconocimiento facial mediante redes neuronales*. Madrid: UPM.
- Chan, C., Tahir, M., Kittler, J., & Pietikäinen, M. (2013). Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors. *IEEE Transactions on pattern analysis and machine intelligence*, 35(5), 1164-1177.
- Chen, Z., Huang, W., & Lv, Z. (2017). Towards a face recognition method based on uncorrelated discriminant sparse preserving projection. *Multimed Tools Appl*, 76(17), 17669-17683.
- Crosswhite, N., Byrne, J., Stauffer, C., Parkhi, O., Cao, Q., & Zisserman, A. (2017). Template adaptation for face verification and identification. *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)* (págs. 1-8). Washington, DC: IEEE.
- Deshpande, N., & Ravishankar, S. (2017). Face detection and recognition using viola-jones algorithm and fusion of PCA and ANN. *Advances in Computational Sciences and Technology*, 10(5), 1173-1189.
- Franco, C., Ospina, C., Cuevas, E., & Capacho, D. (2015). Reconocimiento facial basado en EIGENFACES, LBHP y FISHERFACES en la BEAGLEBOARD-XM. *Revista Colombiana de Tecnologías de Avanzada*, 2(26), 145-152.
- Gao, Y., Ma, J., Zhao, M., & Yuille, A. (2017). Semi-Supervised Sparse Representation Based Classification for Face Recognition with Insufficient Labeled Samples. *IEEE Transactions on image processing*, 26(5), 1-16.
- Georhiades, A., Belhumeur, P., & Kriegman, D. (2001). From few to many: illumination, cone models for face recognition under variable lighting and pose. *IEEE*, 643-660.
- Hamed, M., Salleh, S.-H., & Noor, A. (2015). Facial neuromuscular signal classification by means of least squares support vector machine for MuCI. *Applied Soft Computing*, 30, 83-93.
- Howard, P., Apley, D., & Runger, G. (2017). Distinct Variation Pattern Discovery Using Alternating Nonlinear Principal Component Analysis. *IEEE Transactions on neural networks and learning systems*, 99, 1-11.
- Khedher, L., Illán, I., Górriz, J., Ramírez, J., & Brahim, A. (2017). Independent component analysis-support vector machine-based computer-aided diagnosis system for Alzheimer's with visual support. *International Journal of Neural Systems*, 1-18.
- Li, S., & Jain, A. (2015). *Encyclopedia of biometrics* (Segunda ed.). East Lansing: Springer.
- Lu, J., Tan, Y.-P., & Wang, G. (2013). Discriminative multi-manifold analysis for face recognition from a single training sample per person. *IEEE Transactions on pattern analysis and machine intelligence*, 35(1), 39-51.
- Ma, Y., & Guo, G. (2014). *Support vector machines applications*. Switzerland: Springer.
- Moghaddam, B. (1999). *Principal manifolds and bayesian subspaces for visual recognition*. Cambridge: Mitsubishi Electric Research Laboratory.
- Phillips, P., Moon, H., Rauss, P., & Rivsi, S. (1997). The feret evaluation methodology for face recognition algorithms. *IEEE*, 137-143.
- Rajawat, A., Pandey, M., & Rajput, S. (2017). Low resolution face recognition techniques: a survey. *3rd International Conference on Computational Intelligence & Communication Technology (CICT)*, (págs. 1-4). Ghaziabad.
- Saleh, S., Azam, S., Yeo, K., Shanmugam, B., & Kannoopatti, K. (2017). An improved face recognition method using local binary pattern method. *11th International Conference on Intelligent Systems and Control (ISCO)*, (pp. 112-118). Coimbatore.
- Samaria, F., & Harter, A. (1994). Parameterisation of a stochastic model for human face identification. *2nd IEEE Workshop on applications of computer vision*.
- Schölkopf, B., Smola, A., & Müller, K.-R. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput*, 1299-1319.
- Shakhnarovich, G., & Moghaddam, B. (2011). Face recognition in subspaces. En S. Li, & A. Jain, *Handbook of face recognition* (p. 716). London: Springer.
- Shanmuganathan, S. (2016). *Artificial neural network modeling*. Vol. 628. New Zealand: Springer.
- Stan, L., & Jain, A. (2011). *Handbook of facial recognition*. Londres: Springer.
- Vasilescu, M., & Terzopoulos, D. (2001). Multilinear analysis of image ensembles: tensorfaces. *Proceeding of European Conference on Computer Vision*, 447-460.
- Wang, J., Lu, C., Wang, M., Li, P., Yan, S., & Hu, X. (2014). Robust face recognition via adaptive sparse representation. *IEEE Transactions on cybernetics*, 1-11.
- Xiang, C., Yu, Z., & Qu, X. (2014). Support vector machine optimized by improved genetic algorithm. *Telkomnika Indonesian Journal of Electrical Engineering*, 12(1), 831-836.
- Xu, Q. (2017). *Estatistical modeling for multiple modes facial images using GND-PCA*. Jinan: repositorio de la Universidad de Shandong.
- Xu, Y., Fang, X., Li, X., Yang, J., You, J., Liu, H., & Teng, S. (2014). Data uncertainty in face recognition. *IEEE Transactions on cybernetics*, 44(10), 1950-1961.
- Zhang, W., Meng, X., Li, Z., Lu, Q., & Tan, S. (2015). Emotion recognition in speech using multi-classification SVM. *2015. IEEE 12th Intl Conf. on ubiquitous intelligence and computing and 2015. IEEE 12th Intl Conf. on autonomic and trusted computing and 2015. IEEE 15th Intl Conf. on scalable computing and communications and its. Associated Workshops (UIC-ATC-ScalCom)* (p9. 1.181-1.186). Beijing: IEEE.

Apéndice A. Análisis de Componentes Principales (PCA)

Según Stan y Jain, el método de Análisis de Componentes Principales (PCA) es una técnica de reducción de dimensionalidad basado en la extracción del número deseado de componentes principales de los datos multidimensionales (2011, p. 21).

El Análisis de Componentes Principales fue inicialmente desarrollado dentro del campo de la estadística. Más tarde fue reformulado en el campo de redes neuronales.

El método está relacionado a la Transformada de Karhunen-Loève(KLT). Este método asume que un conjunto dado de N imágenes de muestra $\{x_1, x_2, \dots, x_N\} \in \mathbb{R}^n$. Cada imagen pertenece a una de las m clases $\{C_1, C_2, \dots, C_m\}$. Se define la matriz de dispersión total S_T como.

$$S_T = \sum_{k=1}^N (x_k - \mu)(x_k - \mu)^T$$

Donde μ es la media de los datos. El Análisis de Componentes Principales determina la proyección ortogonal φ como.

$$y_k = \varphi^T x_k \text{ donde } k = 1, \dots, N$$

Esto maximiza el determinante del total de la matriz de dispersión de la muestra proyectada y_1, \dots, y_N :

$$\varphi_{opt} = \arg \max_{\varphi} |\varphi^T S_T \varphi|$$

Las propiedades principales del PCA son resumidas por las siguientes ecuaciones:

$$x \approx \varphi_k y, \quad \varphi_k^T \varphi_k = I, \quad E\{y_i y_j\}_{i \neq j} = 0$$

Reconstrucción aproximada, orto normalidad de la base φ_k , y componentes principales correlacionados $y_i = \varphi_k^T x$ respectivamente (Samaria & Harter, 1994).