

RECONOCIMIENTO DE ROSTROS APLICANDO SVD

Casanova, Stephanya, Hernandez, Giovanna
{stephanya.casanova,giovanna.hernandez.navarro}@correounivalle.edu.co
Universidad del Valle

Resumen—En este informe se describe el desarrollo de un sistema que permite el reconocimiento de rostros por medio de la extracción y comparación de los valores propios de un conjunto de rostros de muestra¹. Inicialmente se hace una selección de los rostros que conformarán la base de imágenes, luego se construye una matriz por cada imagen, en la que cada valor corresponde a un grado de gris en código ASCII, de esta se extraen los valores propios, y esta información se convierte en el pilar del reconocimiento del sistema, en donde cada imagen que se desee clasificar será descompuesta en sus valores propios y posteriormente se calculará la distancia entre la norma euclidiana de esta y la norma de cada una de las imágenes procesadas. En este documento se incluyen además los resultados obtenidos y el análisis de estos.

Índice de Términos- ASCII, Conjunto de Entrenamiento, Reconocimiento de Rostros, SVD, Valores Propios.

I. INTRODUCCIÓN

EN LOS ÚLTIMOS AÑOS LOS SISTEMAS BIOMÉTRICOS HAN COBRADO AUGE A NIVEL MUNDIAL EN VISTA DE SU IMPORTANCIA EN LA RESOLUCIÓN DE PROBLEMAS ENFOCADOS A LA SEGURIDAD, la cual se fundamenta en los siguientes requisitos: autenticación, autorización y auditoría. La autenticación a su vez se divide en tres tipos: algo que usted sabe, algo que usted tiene y algo que usted es. Con relación a los sistemas biométricos se encuentran clasificados en la tercera categoría, ya que permiten identificar a una persona a través de determinadas características físicas, es decir, algo que usted es.

Actualmente se han desarrollado diversos sistemas biométricos, los cuales van desde la identificación de personas a través de huellas dactilares, geometría de la mano, reconocimiento de voz, hasta la identificación a través del iris o de la retina, así como, los sistemas de reconocimiento e identificación de rostros, que son ampliamente aceptados por no ser invasivos. La clasificación por medio de SVD es un método holístico basado en subespacios, el cual busca la proyección de la imagen almacenada mas cercana a la

imagen de entrada, así como la utilización de un umbral obtenido experimentalmente para determinar durante el proceso de clasificación cuando la imagen ingresada es o no una cara, y en caso de serlo cuando corresponde a un rostro de una persona conocida o a un individuo que no se encuentra almacenado en el sistema. La medida de distancia entre vectores empleada en este trabajo por el método SVD durante el proceso de reconocimiento es la distancia euclidiana.

Este proyecto tiene como propósito desarrollar un sistema, que aplica el método SVD para el reconocimiento de rostros a partir de una base de datos de fotografías, las cuales son almacenadas bajo ciertas características y condiciones previamente establecidas.

II. TEORÍA DE LA DESCOMPOSICIÓN EN VALORES SINGULARES SVD APLICADO AL RECONOCIMIENTO DE ROSTROS

El sistema descompone imágenes de rostros en un conjunto de imágenes características llamadas “Valores propios de Rostros”, las cuales se toman como componentes principales del conjunto inicial de imágenes de entrenamiento. Estos “valores” representan las características principales de cada individuo y son los vectores propios del conjunto de imágenes; no se corresponden necesariamente con rasgos tales como los ojos, la nariz o las orejas.

El reconocimiento se lleva a cabo proyectando una nueva imagen sobre el espacio formado por los “valores propios”, denominado “Espacio de Rostros”, y clasificándola comparando su posición en ese espacio con las posiciones de los individuos ya conocidos. En términos matemáticos, el objetivo consiste en encontrar los vectores propios de la matriz de covarianzas del conjunto de imágenes, tratando a la imagen como un punto (o vector) en un espacio de dimensión enorme, donde cada coordenada corresponde al valor de un píxel de la imagen original.

La imagen de cada individuo puede ser representada

mediante una combinación lineal de los “valores propios”. Cada rostro puede aproximarse utilizando únicamente los mejores valores, es decir, aquellos que tienen los mayores valores propios y por tanto, los de máxima varianza. Estas M “mejores valores propios” crean un subespacio M -dimensional de todas las imágenes posibles.

III. DESCRIPCIÓN DEL SISTEMA

3.1 Procesamiento y Clasificación de la Imagen de Rostro Humano

A continuación se presenta un diagrama que describe de manera general el funcionamiento de la aplicación, a partir de sus dos procesos principales: Entrenamiento y reconocimiento. (Ver Fig.1).

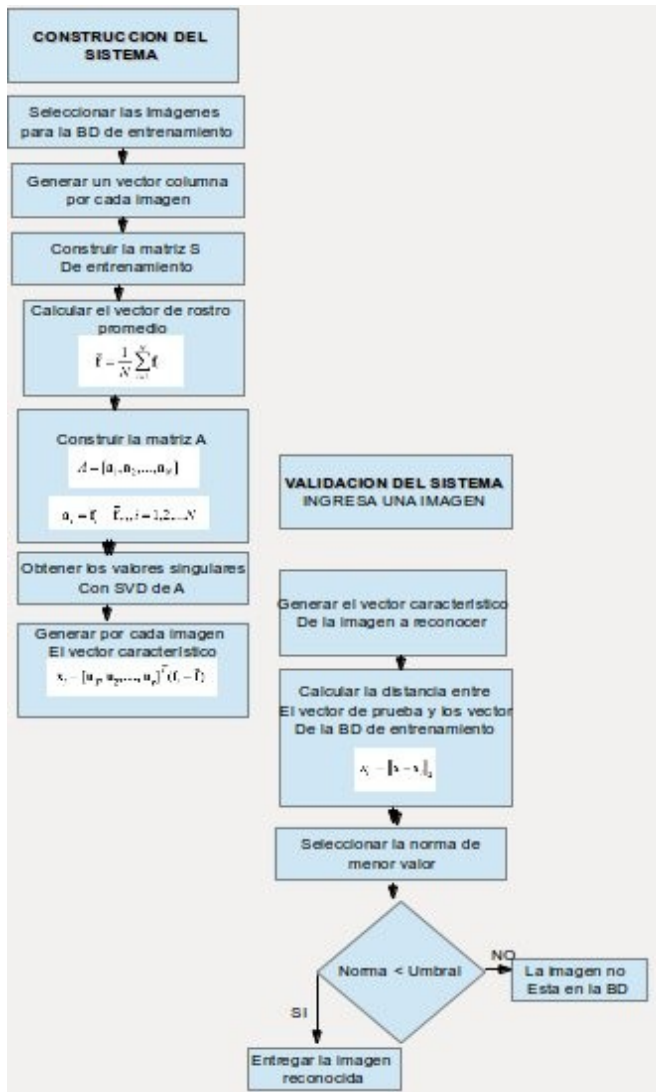


Figura 1. Diagrama de flujo para el proceso de reconocimiento de rostros humanos por medio de SVD

Fase de entrenamiento

1. Adquisición de la colección de imágenes iniciales (“BD de entrenamiento”).
2. Calcular los “Valores propios” del set de entrenamiento, teniendo en cuenta las M imágenes que corresponden a los mayores valores propios. Estas M imágenes definen el “Espacio de Rostros”.
3. Calcular la correspondiente distribución en el espacio M -dimensional para cada individuo, proyectando las imágenes en el “Espacio de Rostros”.

Fase de Reconocimiento

1. Proyectar la nueva imagen de entrada sobre el “face space” y determinar si ésta es una cara, comprobando si es próxima a este espacio.
2. Si es una cara, utilizar un clasificador para determinar la clase más próxima a la que pertenece la nueva imagen.

Para el desarrollo del sistema se escogieron 20 personas y de estas se procesaron 8 diferentes gestos de cada una, lo que generaría una BD de entrenamiento de 160 caras. Cada imagen es representada como una matriz de dimensiones $w \times h = N$ píxeles, cada píxel posee un valor de intensidad diferente (valores entre 0 y 255) que se obtiene al transformar la imagen de RGB a escala de grises, las fotografías están formadas por coordenadas $I(x, y)$. Esta información de los píxeles debe ser llevada a un vector columna de $M \times 1$, donde M es la multiplicación de w y h , para el caso: $w=92$, $h=112$, y $M= 10304$, esto genera un vector f_i de 10304×1 .

Luego de obtener cada uno de los vectores de las 160 caras se construye una matriz $M \times N$: (10304×160)

$$S = [f_1, f_2, \dots, f_N]$$

Posterior a esto, se calcula el vector de cara promedio de S , así:

$$\bar{f} = \frac{1}{N} \sum_{i=1}^N f_i$$

donde f_i es el vector que representa una imagen y $N= 160$ es la cantidad de imágenes.

Seguido, se construye la matriz A para todo el conjunto de entrenamiento S .

$$A = [a_1, a_2, \dots, a_N]$$

donde, $a_i = f_i - \bar{f}, i = 1, 2, \dots, N$

Se obtienen los valores singulares de a por medio de SVD.

$$A = USV^T$$

Donde,

$$\begin{array}{c}
 \boxed{A} \\
 m \times n
 \end{array}
 =
 \begin{array}{c}
 \boxed{U} \\
 m \times n
 \end{array}
 \begin{array}{c}
 \boxed{\begin{array}{c} 0 \\ 0 \end{array}} \\
 n \times n
 \end{array}
 \begin{array}{c}
 \boxed{V^T}
 \end{array}$$

$$\begin{array}{ll}
 A = 10304 \times 160 & U = 10304 \times 160 \\
 S = 160 \times 160 & V = 160 \times 160
 \end{array}$$

A continuación, se genera por cada imagen i su vector característico x_i , así:

$$x_i = [u_1, u_2, \dots, u_r]^T (f_i - \bar{f})$$

donde r es el rango de la matriz A .

X_i : vector característico de la imagen i de las $N = 160$ imágenes, de dimensiones 160×1

Finalmente, para el ejercicio de validación del sistema se escoge una imagen de un rostro humano y se verifica la existencia de esta en el subespacio del rostro, para esto se deben generar el vector característico de esta y calcular la norma 2 de la distancia entre este y cada uno de los vectores de la BD de entrenamiento.

$$\varepsilon_i = \|x - x_i\|_2$$

Cuando se tengan todos los E_i calculados se escoge la norma de menor valor de todas, y se compara con el umbral = 3400 (este valor fue seleccionado a propósito y conforme a una serie de pruebas realizadas a 50 imágenes de rostros humanos para disminuir el error de acierto en las imágenes que no se encuentran en la BD de entrenamiento). La norma mas pequeña representar la imagen mas parecida, si la norma es mayor que 3400 se dice que la cara no existe en el conjunto de entrenamiento.

3.2 Interfaz

El programa de clasificación de rostros humanos **COLDFACE UV** fue elaborado en JAVA implementando el paquete JAMA, en su versión delgada, para la generación de los valores singulares de cada una de las imágenes. A continuación se presenta la interfaz de inicio del software:

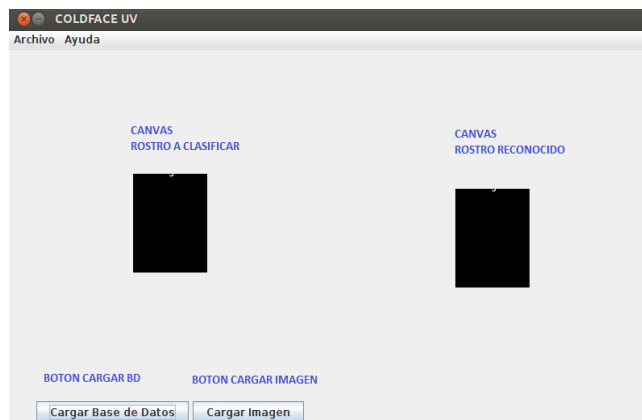


Figura 2. Interfaz del software COLDFACE UV

La interfaz se compone de:

1 Botón “Cargar Base de Datos” [ctrl B], en este se selecciona la base de datos de entrenamiento del sistema, el software esta diseñado para tomar 20 personas y de estas 8 gestos por cada una en formato .pgm.

1 Botón “Cargar Imagen” [ctrl I], aquí se selecciona la imagen que se desea clasificar, esta debe estar en formato .pgm.

2 Canvas, en donde se presentan la imagen a clasificar y la imagen que el sistema obtuvo luego del proceso de reconocimiento.

Para la ejecución del software debe escribirse en consola:
 java -cp .:Jama-1.0.3.jar -Xms1024M -Xmx1024M
 ReconocedorGUI

IV. EXPERIMENTACIÓN Y RESULTADOS

En la tabla 1 se muestran los resultados obtenidos con el sistema desarrollado, se hicieron pruebas con 24 imágenes diferentes, de las cuales 14 se encuentran en la base de datos empleada en el sistemas, otras 7 corresponden a imágenes de rostros de personas que no hacen parte de la base de datos de construcción y las 3 restantes son imágenes diversas que no son rostros humanos.

Tabla 1 Resultados obtenidos con el Software COLDFACE UV para el reconocimiento de rostros por medio de SVD

REF	Rostro	BD	Acierta Objeto	Acierta Gesto
NoCarro.pgm	NO	NO	SI	NA
no_face.pgm	NO	NO	SI	NA

no_face1.pgm	NO	NO	SI	NA
Face.pgm	SI	NO	SI	NA
Face2.pgm	SI	NO	SI	NA
cara_m.pgm	SI	SI	SI	SI
S2-5	SI	SI	SI	SI
S2-10	SI	SI	SI	NO
S5-10	SI	SI	NO	NO
S6-9	SI	SI	SI	NO
S7-6	SI	SI	SI	SI
S8-9	SI	SI	SI	NO
S9-10	SI	SI	SI	NO
S10-10	SI	SI	NO	NO
S11-10	SI	SI	SI	NO
S12-9	SI	SI	SI	NO
S18-10	SI	SI	SI	NO
S19-10	SI	SI	SI	NO
S20-9	SI	SI	SI	NO
S30-4	SI	NO	SI	NA
S32-8	SI	NO	SI	NA
S35-5	SI	NO	SI	NA
S38-6	SI	NO	NO	NA
S40-8	SI	NO	NO	NA

Según la anterior tabla se tiene un **83,33%** de acierto en las pruebas en general, un **80%** de éxito con fotografías que no pertenecen a la BD de construcción (*S30-4, S32-8, S35-5, S40-8, NoCarro, no_face, no_face1, face, face2*), un **100%** de acierto con imágenes que no corresponden a rostros humanos (*NoCarro, no_face, no_face1*), y en casos de rostros con partes faltantes como ojos y nariz (*cara_m, cara_m1*) un acierto del 100%.

En las siguientes gráficas de torta se presentan los porcentajes de acierto y no acierta en las pruebas y los porcentajes de las condiciones de las fotografías.

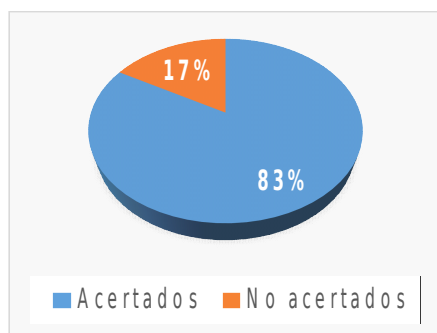


Figura 3. Diagrama de torta para % de acertados y no acertados.

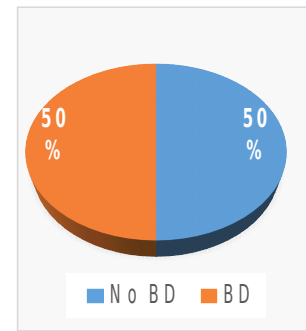


Figura 4. Diagrama de torta para No aciertos, en donde 2 casos corresponden a imágenes que se encuentran en la BD de entrenamiento y 2 no se encuentran.

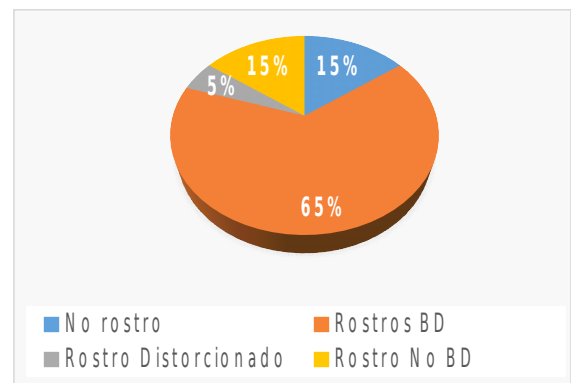


Figura 5. Diagrama de torta para Aciertos, en donde 13 casos corresponden a imágenes que se encuentran en la BD de entrenamiento, 3 no se encuentran, 3 no son rostros humanos y 1 es un rostro humano distorsionado.

A continuación se presentan 6 pruebas en donde se evidencian los aciertos y fallas del software COLDFACE UV.

Caso 0. Imagen de No Rostro Humano



Figura 6. Caso 0, donde la imagen es diferente a un rostro humano

Aquí se logra una clasificación acertada de la imagen, ya que esta no hace parte de la BD de entrenamiento. Cabe mencionar que la distancia para imágenes que no son de rostros humanos son valores muy altos superando las unidades de mil.

Caso 1. Imagen Rostro Humano No BD de entrenamiento



Figura 7. Caso 1, donde la imagen es una rostro humano pero no hace parte de la BD de entrenamiento

En este ejemplo se observa que el programa responde adecuadamente ante una imagen de un rostro que no hace parte de la BD de entrenamiento.

Caso 2. Imagen Rostro Humano en BD de entrenamiento con elemento del rostro distorsionado.



Figura 8. Caso 2, donde la imagen es un rostro humano pero con distorsión

De nuevo aquí se comprueba que el sistema es capaz de cubrir las distorsiones de un rostro que se encuentre en la BD de entrenamiento.

Caso 3. Imagen Rostro Humano en BD de entrenamiento

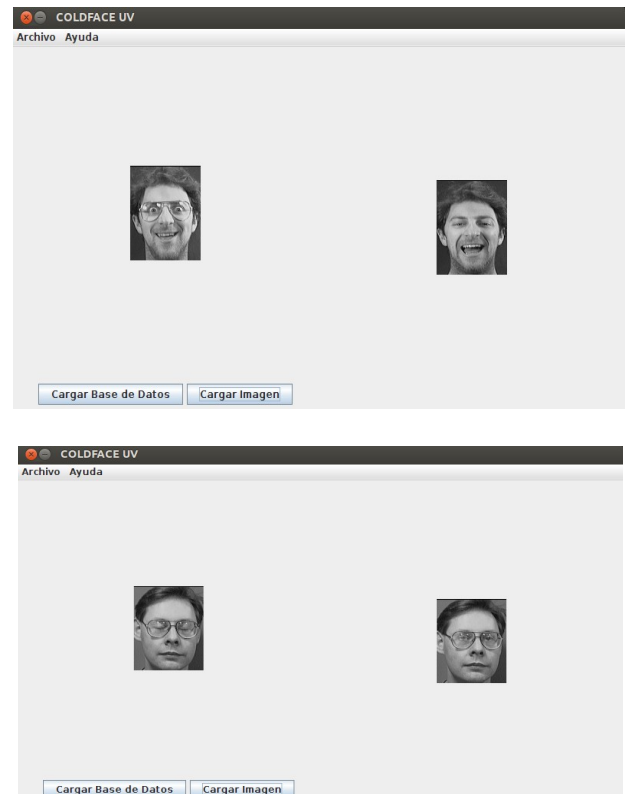


Figura 9 Caso 3, Imágenes que si hacen parte de la Bdde entrenamiento

Efectivamente, se observa que al ingresar un rostro de una persona que hace parte de la Bdde entrenamiento con diferentes gesto al software, este es capaz de clasificarlo correctamente. Se puede observar además que el software es capaz de sortear la presencia de objetos adicionales en un rostro como el caso del uso de gafas.

Caso 4. Imagen Rostro Humano en BD de entrenamiento con Error de clasificación



Figura 10. Caso 4, donde el software reconoce erróneamente una imagen de la BD

Aquí se puede evidenciar un error en el software en la clasificación de una imagen que se encuentra en la base de entrenamiento, esto se debe a que se estableció un umbral con un valor experimental de 3400 lo que hace que el software al no encontrar un valor de distancia menor muestre la imagen de error, anteriormente con un valor de umbral de 5000 el sistema reconocía a otra persona, esto es un error, sin embargo cabe mencionar que el procesamiento con SVD desconoce los elementos del rostro y trabaja estrictamente con niveles de gris.

Caso 5. Imagen Rostro Humano NO BD de entrenamiento con Error de clasificación



Figura 11. Caso 5, donde el software reconoce erróneamente una imagen que no está en la BD

Aquí se observa que el software relaciona erróneamente 2 rostros de personas distintas, esto se debe a que la distancia entre ambas imágenes está por debajo del umbral propuesto, y al observarse ambas imágenes se puede decir que estas 2 personas comparten características como lo son: la barba, las gafas, la forma del rostro y el bigote, estos elementos contribuyen en los niveles de gris de la imagen lo que hace que los valores propios de estos sean similares.

En términos generales se puede concluir que el reconocimiento de rostros por medio de SVD ofrece un alto grado de éxito, lo que lo hace una herramienta adecuada para actividades de clasificación de rostros humanos.

V. CONCLUSIONES

El empleo de SVD para el reconocimiento de rostros es un método sencillo y eficiente. Este hace uso de la vectorización de la imagen en niveles de gris desconociendo los componentes de un rostro como la nariz, los ojos, la boca entre otros, esto puede llevar al algoritmo a no reconocer apropiadamente un rostro ya que es muy factible que existan varias imágenes con la misma distribución de gris.

Para complementar el algoritmo de SVD se deberían incluir otras técnicas de extracción de características como FFT o

implementar algoritmos que se especialicen en una parte del rostro, con esto se tendría un sistema mucho más robusto y confiable.

La selección del umbral debe ser cuidadosa ya que si es demasiado bajo puede que no encuentre una imagen en la BD de entrenamiento, y en el caso contrario, para un valor alto, puede que reconozca erróneamente la imagen.

REFERENCIAS

- [1] GUOLIANG ZENG, "FACE RECOGNITION WITH SINGULAR VALUE DECOMPOSITION.", CISSE PROCEEDING, 2006
 - [2] LARS ELDÉN, "NUMERICAL LINEAR ALGEBRA AND APPLICATIONS IN DATA MINING", LINK PING UNIVERSITY, 2005.
- WEB
(mayo 28 de 2014)
- [1] http://es.wikipedia.org/wiki/Descomposici%C3%B3n_en_valores_singulares
 - [2] <http://www.math.cuhk.edu.hk/~lmlui/CaoSVDintro.pdf>