

RECONOCIMIENTO DE IMÁGENES FACIALES

PATRICIA SANTANA CEMA

01-830279

Trabajo de grado presentado para optar el título de Maestría de Magister en
Ciencias-Matemática Aplicada

DIRIGIDO POR:

HERNÁN ESTRADA BUSTOS

UNIVERSIDAD NACIONAL DE COLOMBIA

FACULTAD DE CIENCIAS

DEPARTAMENTO DE MATEMÁTICAS

Bogotá, 2010

RECONOCIMIENTO DE IMÁGENES FACIALES

PATRICIA SANTANA CEMA

01-830279

Magister en Ciencias-Matemática Aplicada

DIRIGIDO POR:

HERNÁN ESTRADA BUSTOS

UNIVERSIDAD NACIONAL DE COLOMBIA

FACULTAD DE CIENCIAS

DEPARTAMENTO DE MATEMÁTICAS

Bogotá, 2010

Índice general

1. Introducción	1
1.1. Orden del trabajo	1
1.2. Métodos de identificación personal	2
1.3. Identificación de caras.	4
1.4. Aplicaciones modernas de la identificación de caras	4
2. Conceptos básicos de imágenes para la identificación de caras	6
2.1. Concepto de fotografía	6
2.2. Normalización de fotos	7
2.3. Construcción de la base de caras	7
3. Procedimiento para realizar reconstrucción de caras	11
3.1. Base ortogonal para la Base de caras	13
3.2. Reconstrucción	17
4. Identificación de rostros	21
4.1. Comparación de fotografías	21
4.2. Aplicaciones	25
5. Conclusiones	30
6. Apéndices	33
6.1. Introducción	33
6.2. Definición de SVD	34
6.3. Ejemplo de la factorización de A en la forma SVD	40
6.4. Cálculo de valores propios	42
6.4.1. Reducción de matrices a formas más simples	43
6.5. Pseudocódigo	48

Capítulo 1

Introducción

Acontecimientos como el del 11 de Septiembre de 2001 en Nueva York, el del 11 de marzo del 2004 en Madrid y recientemente los atentados en el metro en Moscú, han hecho que se incremente la preocupación por seguridad y se busquen sistemas efectivos para garantizar a los ciudadanos su bienestar. En dichos sistemas intervienen factores humanos, administrativos y técnicos, propiciando un complejo proceso que requiere de avances científicos relacionados con tecnología computacional de vanguardia. Dado que la problemática actual de seguridad mundial se aborda desde una perspectiva interdisciplinar, la matemática aplicada juega un papel protagónico y es donde queremos contribuir a este vasto tema.

1.1. Orden del trabajo

Este trabajo se ha dividido en cuatro partes. En la primera, se consideran diferentes métodos de identificación personal, en particular la identificación de caras con algunas de sus aplicaciones en diferentes países. En la segunda parte, se darán algunas definiciones básicas de imágenes y la forma de construcción de la base de caras. En la tercera parte, se establecerá el procedimiento para la reconstrucción de caras, discutiremos el fundamento de nuestro método de identificación que básicamente utiliza la teoría de las componentes principales de la matriz de caras y la representación de una imagen facial en un conjunto de vectores característicos. En la cuarta parte, se establece el procedimiento de identificación de caras mediante la proyección de fotografías externas en el espacio de la base de caras. El algoritmo de componentes principales han sido discutidos

en la literatura especializada, sus bases radican en los trabajos de Pentland y Turk [1], y se ha convertido en estándar sus conceptos y procedimientos cuando se trabaja reconocimiento de imágenes faciales por medio de componentes principales, nuestro interés es comprender el proceso y las dificultades del reconocimiento de imágenes faciales, finalmente se expondrá la identificación de NN y la identificación de retratos hablados que puede ofrecer alternativas a los métodos utilizados por las autoridades de Medicina Legal y criminalística para esclarecer casos de justicia.

1.2. Métodos de identificación personal

Durante mucho tiempo la identificación de las personas tan sólo estaba relacionada con la memoria fotográfica y ésta es subjetiva ya que tiene un alto grado de error en el reconocimiento de personas. Muchos de los métodos usados actualmente para la identificación están basados en la tecnología biométrica. Esta tecnología se define como: Medición de distintas características físicas exclusivas y determinantes de la identidad de un individuo [2].

Algunos ejemplos de identificación biométrica son las huellas dactilares, las retinas, los patrones faciales etc. Sistemas de identificación basados en la huella dactilar se remontan a la cultura China en el siglo XIV cuando los comerciantes imprimían las huellas de la palma de la mano y los pies en un papel con tinta para distinguir los individuos [3]. En 1870, el francés Alphonse Bertillon desarrolló un sistema de antropometría descriptiva como un método para identificar individuos basado en registros detallados de las medidas del cuerpo. Esta técnica fue utilizada en Francia como instrumento en las investigaciones de la policía entre 1882 y 1890 [4] Francisco Galton en su publicación *Huellas dactilares* de 1892 garantizó la invariabilidad de las huellas digitales de un individuo a lo largo de toda la vida. Además, determinó 40 características de las huellas que incluso son utilizadas hoy en día para la clasificación y propuso usarlas para la identificación personal en reemplazo del sistema de Bertillon [5]. El italiano Juan Vucetch radicado en Argentina, realizó las primeras fichas dactilares del mundo con las huellas de 23 procesados, además verificó su método con 645 reclusos de la cárcel de La Plata y en el trabajo resalto 101 rasgos de las huellas empleados para su clasificación. Mediante esta técnica en 1892 en la Argentina se detectó al primer asesino por cotejo de sus huellas dactilares. A partir de 1894 la

Policía de Buenos Aires adoptó oficialmente las huellas dactilares como sistema de identificación [6].

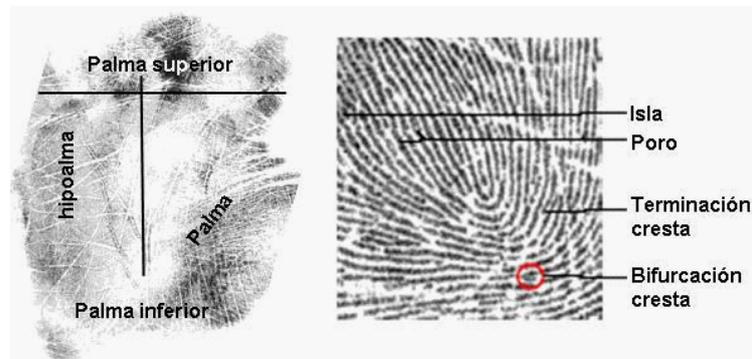


Figura 1: Partes de las huellas digitales

Los oftalmólogos Isadore Goldstein y Carleton Simon, publicaron en 1935 el trabajo de investigación *A New Scientific Method of Identification*, en el que reportan el siguiente descubrimiento: el dibujo que conforman los vasos sanguíneos de la parte posterior de los ojos son únicas en cada persona (Figura 2). Para el caso de gemelos idénticos, Paul Tower en su publicación de 1955 *The fundus Oculi in monozygotic twins : Report of six pairs of identical twins*, demostró que, en lo que menos se parecen los gemelos idénticos es en los patrones vasculares de la retina [7].

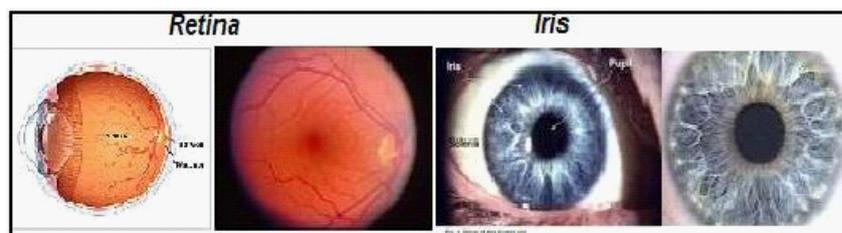


Figura 2: Gráfica de la retina e iris

Existen otros sistemas de identificación biométrica tales como el iris, la geometría de la palma de la mano, la dinámica del paso y el análisis espectral de la voz. El reconocimiento de imágenes faciales es una alternativa de identificación que, ha ganado gran interés en las últimas décadas, dado el avance en el procesamiento de imágenes y las técnicas computacionales. Este último método de identificación personal, constituye el centro de estudio de este trabajo.

1.3. Identificación de caras.

Cuando tratamos de identificar personas uno de los aspectos más relevantes es la cara ya que en ella se encuentran características típicas de la persona y se convierten en el principal foco de atención e interacción social. Desde que nacemos empezamos a aprender a identificar caras, y con el tiempo entendemos que sus facciones son una de las fuentes más importantes de información personal ofreciendo claras referencias sobre edad, estado de ánimo, género, raza, etc. Así, el más natural de los sistemas de identificación en los humanos, es el reconocimiento de caras. Si pretendemos utilizar este sistema de manera computacional nos encontramos ante una tarea bastante difícil puesto que la cara, hace parte de un ente vivo y puede adoptar diferentes expresiones ya que esta sujeta a los efectos del envejecimiento.

En 1988, una división de la policía en la ciudad de Los Ángeles usó dibujos compuestos o imágenes de video para realizar búsquedas en bases de datos de fotografías de criminales. Este procedimiento es considerado el primer sistema semiautomático de reconocimiento facial que hizo uso de la tecnología computacional [8]. El reconocimiento de caras y el procesamiento de imágenes se han convertido en un tema de bastante interés en investigación ya que es una alternativa actual para los sistemas de seguridad.

El reconocimiento facial se realiza a partir de una imagen digital en donde se analiza matemáticamente ciertas características faciales de la foto para compararlas con la información de una base de datos previamente establecida en la literatura especializada existen varios métodos para la identificación de caras como: análisis por características, red neuronal, eigenfaces, entre otros, la técnicas basadas en eigenfaces o valores principales será discutido en profundidad en los capítulos posteriores

1.4. Aplicaciones modernas de la identificación de caras

Los primeros programas de reconocimiento facial fueron instalados en Newham (Reino Unido) en 1.997. En esta ciudad se instaló un sistema de vídeo-control conectado con un programa especializado de reconocimiento de caras. Las imágenes eran transmitidas al centro de la policía y se contrastaba con la información

de una base datos de datos de delincuentes. El uso de este sistema permitió en Newham una disminución de la criminalidad en un 34 % [9].

Australia utiliza actualmente esta técnica en los aeropuertos. Allí se han instalado las denominadas "puertas inteligentes", que utilizan la identificación facial para permitir el acceso rápido de la tripulación de todas las aerolíneas. Así mismo, la aduana tiene un sistema de reconocimiento facial que compara el rostro de la persona con la imagen en el microchip del pasaporte electrónico. El objetivo es verificar que el titular del pasaporte es el propietario legítimo. Esta técnica también empezó a ser utilizada desde el año 2008 en la Terminal del aeropuerto de Manchester del Reino Unido, para ayudar a agilizar la congestión de viajeros y las preocupaciones de las autoridades policiales relacionadas con la seguridad. Las nuevas puertas de seguridad solo pueden ser usadas por ciudadanos adultos del Reino Unido y la Unión Europea que tengan pasaportes electrónicos [10]. Otro ejemplo de la utilización de identificación de rostros se da en el departamento de seguridad de EEUU que posee la base de datos más grande del mundo para el reconocimiento de caras. Dicha base está compuesta por más de 75 millones de fotografías y es la que se usa activamente para la tramitación de visas. La policía de Tampa Bay Florida utilizó en el Super Bowl XXXV un software de reconocimiento facial cuyo objetivo era la búsqueda de delincuentes que asistían al evento. Con ayuda de este procedimiento se detectaron 19 personas con órdenes de captura [9].

Vale la pena destacar que el sistema también es utilizado para evitar fraude electoral. En las elecciones presidenciales de México en el año 2000, la registraduría de ese país comparó las imágenes de los votantes con la información de la base de datos para reducir drásticamente la suplantación en la votación y recientemente se están implementando sistemas de reconocimiento de rostros para ser usados en los cajeros automáticos para evitar fraudes indebidos por la manipulación de las tarjetas de crédito [11].

Capítulo 2

Conceptos básicos de imágenes para la identificación de caras

En este capítulo consideraremos algunos elementos básicos del análisis de imágenes con el fin de definir la terminología empleada para la identificación de caras. Inicialmente daremos una definición de imagen digital, posteriormente se tratará el concepto de normalización de la imagen y finalmente se mencionará la manera como se construye la base de caras.

2.1. Concepto de fotografía

En este trabajo vamos a considerar fotografías en escala de grises comúnmente denominada en blanco y negro. Una fotografía digital en escala de grises (Figura 3) se define como un rectángulo formado por píxeles (cuadraditos), en donde cada uno de ellos tiene un determinado color en la escala de grises. Matemáticamente podemos decir que la foto digital es una función:

$$f : R \rightarrow C \quad \text{con} \quad R = ([a, b] \times [c, d]) \cap Z^2 \quad (2.1)$$

Esto es, al píxel de la posición (i, j) del rectángulo R se le asigna el color $F(i, j)$ de la escala de grises. En el lenguaje Matlab la escala de grises contiene 256 tonalidades que van desde el blanco ($f(i, j) = 255$) hasta el negro ($f(i, j) = 0$).

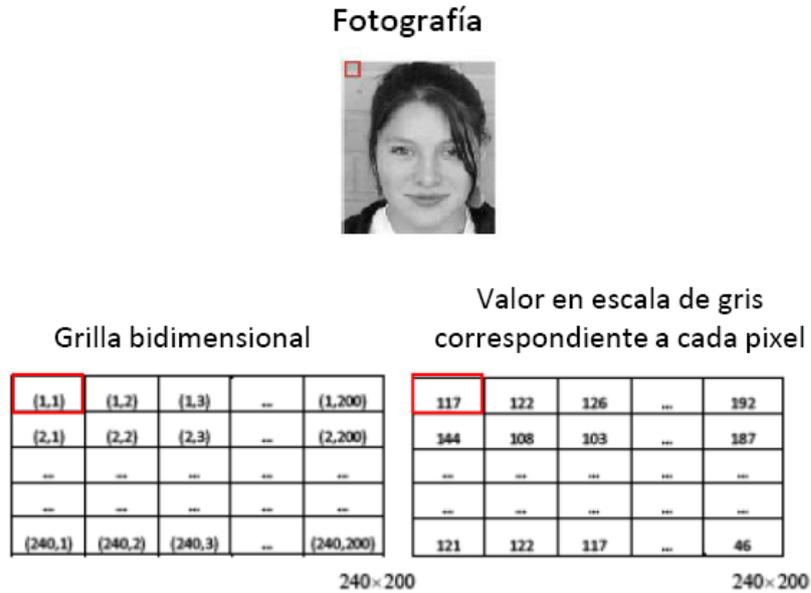


Figura 3: Definición de fotografía en escala de grises

2.2. Normalización de fotos

Para nuestro propósito, la normalización de una imagen significa en nuestro caso, considerar que todas las fotos de nuestra base tengan:

- Las dimensiones de 240×200 , esto es que contienen 48000 pixeles.
- Las fotos deben ser en blanco y negro.
- Las fotos deben representar personas en posición frontal, con diferentes expresiones faciales como risa, tristeza etc
- Las condiciones de luz deben ser similares.
- El fondo de las fotos es de color blanco.

2.3. Construcción de la base de caras

A partir del conjunto de p fotografías digitales $f(1), f(2), \dots, f(p)$, convertimos cada una de las imágenes $f(i)$ $i = 1, \dots, p$ en un vector fila de longitud $(n \times m)$ ($n = 240, m = 200$) que se obtiene al concatenar los elementos de la matriz $f(i)$. Así, para un conjunto de p fotos, se obtiene una matriz F con p vectores fila, cada uno de ellos con 48000 columnas.

Denominamos F la matriz de la base de datos de nuestro conjunto de fotos y la construimos colocando cada uno de los p vectores en las filas de dicha matriz. Por tanto, la matriz que representa la base de caras del conjunto de p fotos es un arreglo NO cuadrado de dimensión $p \times (n \times m)$:

$$F = \begin{pmatrix} f(1) \\ f(2) \\ f(3) \\ \vdots \\ f(i) \\ \vdots \\ f(p-1) \\ f(p) \end{pmatrix}_{p \times (m \times n)} \quad (2.2)$$

En este trabajo se experimento con dos bases de caras. La primera base de caras, es una base de caras disponible en internet (Figura 4) las condiciones de luz son similares y contiene 63 fotografías de 9 personas con 7 expresiones faciales diferentes y cada una de las fotos tiene resolución 240×200 , para esta base de caras la matriz F tiene dimensiones 63×48000 , es decir contiene 3024000 elementos y cada uno de ellos toma un valor entre 0 y 255. Computacionalmente nuestra matriz ocupa cerca de 2.5 Mb. Este valor se incrementa sustancialmente a medida que tomamos más imágenes con mejor calidad. Por ejemplo: Si consideramos 100000 personas y cada persona tiene 7 fotos en blanco y negro con expresiones diferentes con una resolución de $4200 \times 2800 \times 100000$ se tiene que acumular información equivalente a 970 Gb!!.

La segunda base de caras¹, las condiciones de luz son diferentes en algunas de sus fotos y contiene 80 fotografías de 16 estudiantes del Colegio Nueva Delhi con 5 expresiones faciales diferentes y cada una de las fotos tiene resolución 240×200 (Figura 5), para esta base de caras la matriz F tiene dimensiones 80×48000 , es decir contiene 3840000 elementos.

¹La base de caras del Colegio Nueva Delhi, esta almacenada en el archivo *estudiantes.mat* en el lenguaje *matlab* es una matriz $F_{80 \times 48000}$, donde cada fila representa la foto de un estudiante, con alguna pose en particular y esta disponible en psantanac@unal.edu.co

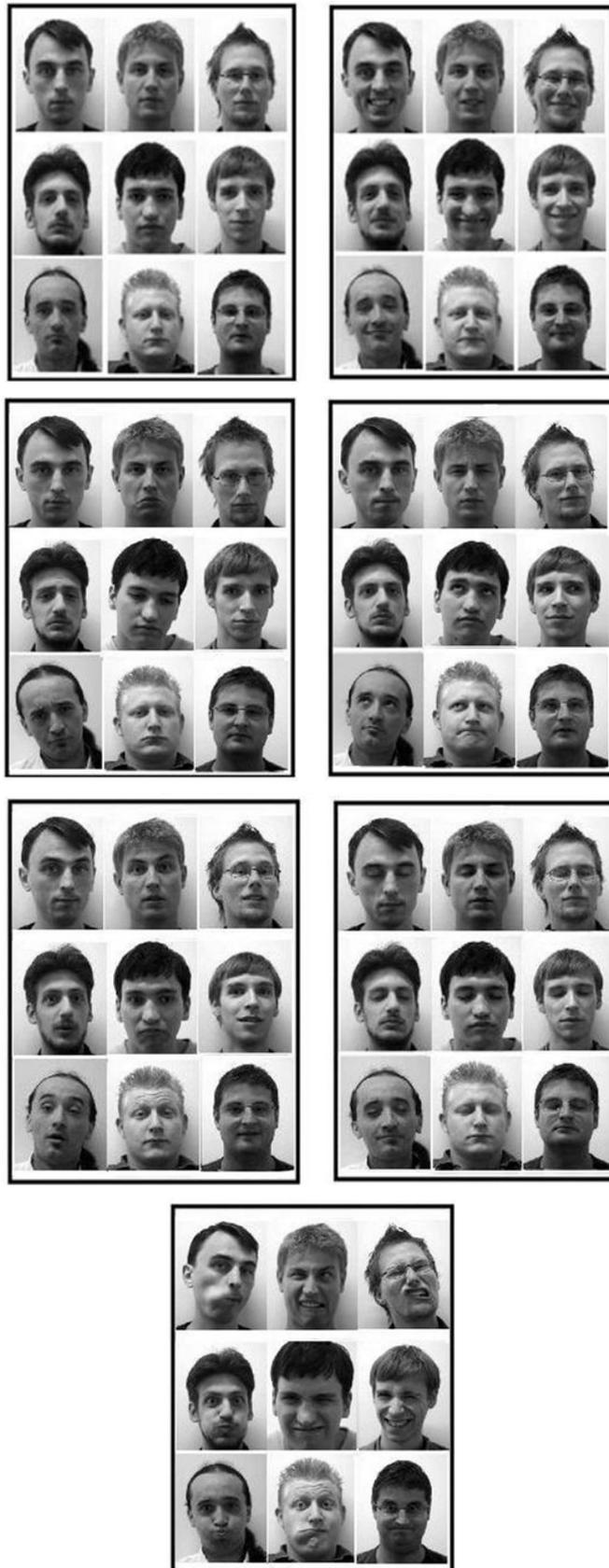


Figura 4: Primera base de caras

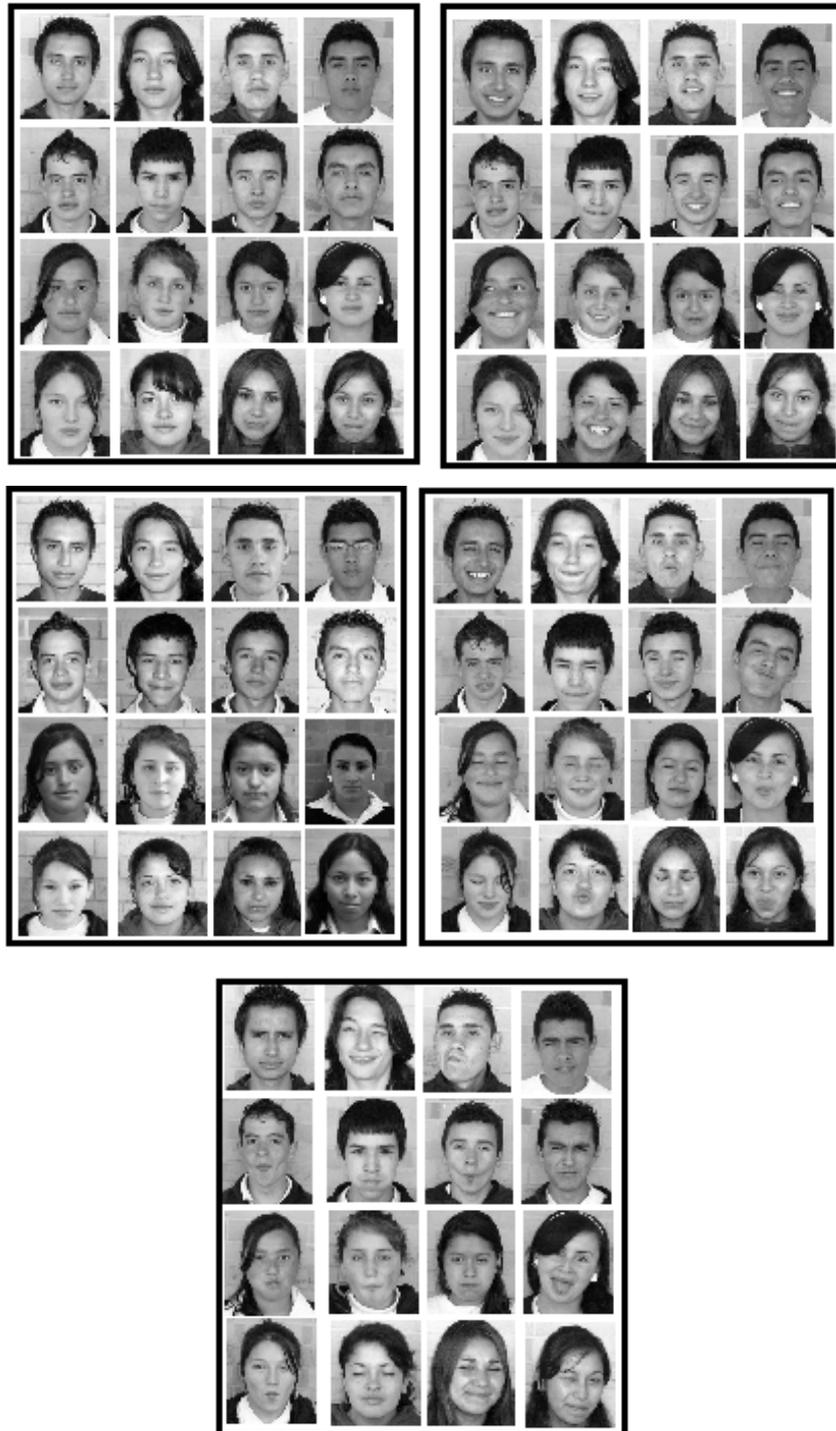


Figura 5: Segunda base de caras Colegio Nueva Delhi

Capítulo 3

Procedimiento para realizar reconstrucción de caras

Nuestro interés en este capítulo es ilustrar la técnica matemática pertinente para reconstruir una imagen que se encuentra en la base de datos F . Los dos desarrollos que se indican a continuación son fundamentales para el próximo capítulo en donde se mostrará la manera de identificar un rostro.

El primer proceso se inicia eliminando información que no es importante o superflua ya que todas las imágenes la contienen. Formaremos la *cara promedio* \tilde{f} del conjunto de p imágenes, este es un vector que se realiza con el promedio de los valores de los píxeles de todas las p fotos que forman la base de caras.

$$\tilde{f} = \frac{1}{p} \sum_{i=1}^p f(i) \quad (3.1)$$

Tomando como ejemplo la base de caras de la figura 4, la cara promedio \tilde{f} se indica en la Figura 6



Figura 6: Cara promedio correspondiente a la primera base de caras

Ahora, para eliminar la información superflua, restamos ésta imagen \tilde{f} de cada una de las caras $f(i)$ ($i = 1, \dots, p$) que forman la matriz F y obtenemos una caricatura $a(i)$ ($i = 1, \dots, p$) de nuestra base de datos.

$$a(i) = (f(i) - \tilde{f})^T \quad (3.2)$$

Un ejemplo de caricaturas de la base utilizada en la Figura 4 es la que se indica en la Figura 7.



Figura 7: Ejemplo de caricatura de la primera base de caras.

Con las caricaturas $a(i)$ ($i = 1, \dots, p$) formamos una matriz que denominamos A , que contiene la información relevante del conjunto de personas que forman la base.

$$A_{(m \times n) \times p} = (a(1)a(2)a(3)\dots a(i)\dots a(p)) \quad (3.3)$$

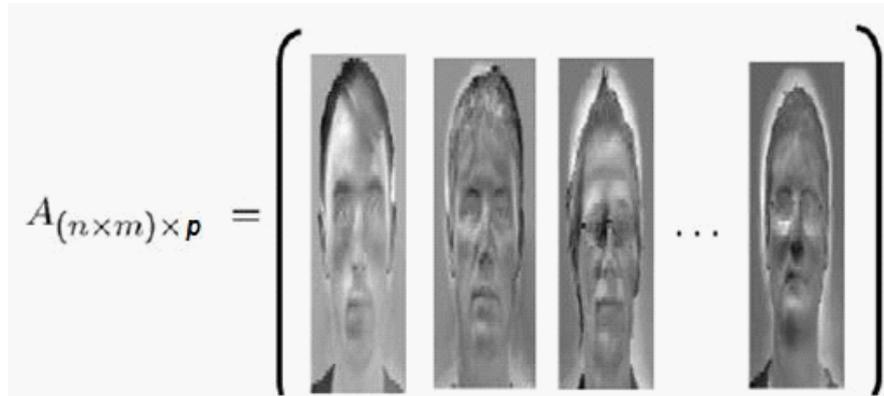


Figura 8: Construcción de la matriz de caricaturas A perteneciente a la primera base de caras

En el segundo procedimiento no determinamos ni la cara promedio ni la matriz de caricaturas, sino que tomamos $A = F^T$, es decir los valores de la matriz A son los mismos que los valores de F .

3.1. Base ortogonal para la Base de caras

Con el propósito de identificar una foto, para cualquier procedimiento debemos seleccionar una base ortogonal de tal manera que podamos expresar las fotos en función de elementos de la base, algo similar al Análisis de Fourier

A continuación mostraremos tres formas para la determinación de la base ortogonal que nos permite realizar la reconstrucción o identificación de caras.

- Definimos una matriz de Covarianza $C = AA^T$

$$C_{(m \times n) \times (m \times n)} = A_{(m \times n) \times p} A_{p \times (m \times n)}^T$$

La matriz C es una matriz cuadrada que en nuestro caso su dimensión es 48000×48000 . Con el fin de realizar el reconocimiento de caras debemos determinar sus valores y vectores propios.

$$Cu_i = \lambda_i u_i \quad (3.4)$$

$$AA^T u_i = \lambda_i u_i \quad (3.5)$$

Dado que la matriz es extremadamente grande con $48000^2 = 2,3 \times 10^9$ elementos, el procedimiento numérico requerido para encontrar sus valores propios

y sus vectores propios es una tarea costosa. Por tanto abordarlo de manera directa es bastante engorroso, sin embargo en la siguiente forma proponemos un método alternativo para llevar a cabo la búsqueda de valores y vectores propios de C .

- Para obviar la dificultad enunciada anteriormente originada por el gran tamaño de la matriz C , consideramos una matriz alternativa $L = A^T A$

$$L = A^T A, L_{p \times p} = A_{p \times (m \times n)}^T A_{(m \times n) \times p}$$

La cual tiene una dimensión $p \times p$ que es un valor razonable para realizar cálculos (por ejemplo la primera base de cara tiene 63 fotografías y por tanto la matriz L tiene 63×63 elementos). Para la determinación de los valores propios de C a partir de L procedemos de la siguiente forma:

Consideremos inicialmente la ecuación de valores propios para L

$$A^T A v_i = \mu_i v_i$$

y multipliquemos al lado izquierdo de la igualdad anterior por A :

$$A A^T A v_i = \mu_i A v_i$$

realizando un reordenamiento adecuado, obtenemos:

$$(A A^T) A v_i = \mu_i A v_i \tag{3.6}$$

se observa que el término entre paréntesis es la matriz C definida anteriormente. Por tanto es fácil reconocer que los vectores propios de C están dados por

$$u = A v_i$$

Con este procedimiento, el cálculo numérico de los vectores propios de C se reduce significativamente, pues es fácil con ayuda de un computador de escritorio y métodos numéricos convencionales realizar tales operaciones.

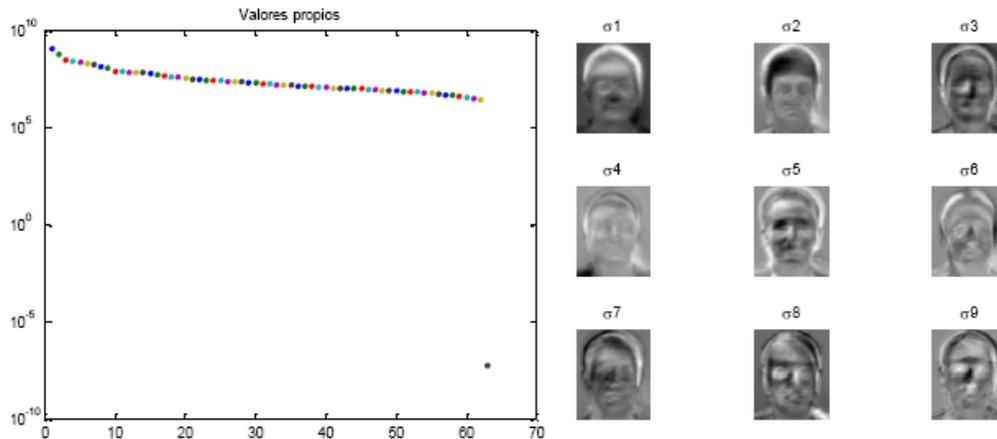


Figura 9: Valores propios λ_i y grafica de los 9 primeros vectores propios u_i de C utilizando la matriz de caricaturas A a partir del procedimiento de la ecuación 3.6

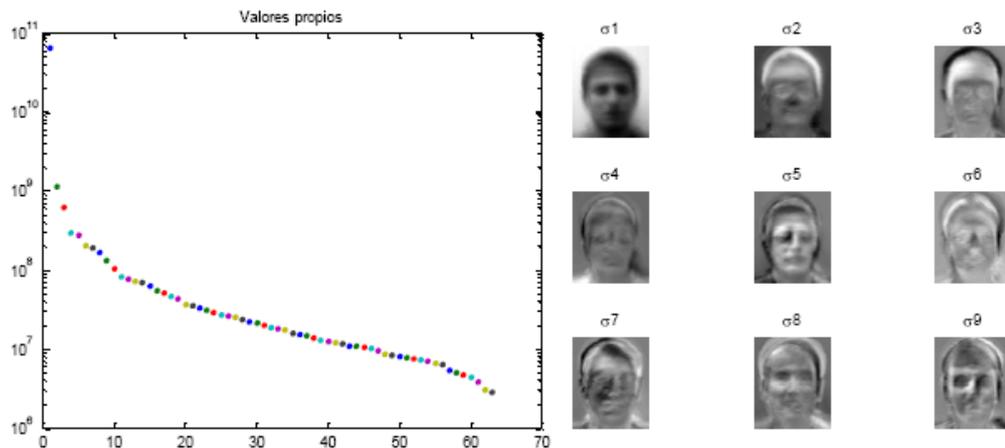


Figura 10: Valores propios λ_i y grafica de los 9 primeros vectores propios u_i de C utilizando la matriz F a partir del procedimiento de la ecuación 3.6

En las dos gráficas anteriores se observa que los valores propios decrecen cuando n aumenta, sin embargo el decrecimiento de estos valores propios es más lento en el caso en que utilizamos la caricatura f . Este comportamiento se debe a que en la matriz A se ha retirado la caricatura f , y por tanto, se ha eliminado una información superflua de la cara promedio.

- Dado que la matriz A no es cuadrada realizamos la descomposición SVD con el fin de determinar los valores singulares y la base adecuada para realizar el procedimiento de la reconstrucción e identificación de caras (ver apéndice):

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T$$

La matriz Σ es una matriz diagonal no cuadrada $m \times n$ que contiene los valores singulares descendentes de $A^T A$ o AA^T , U es una matriz ortonormal cuadrada ($U^T U = I$) y sus columnas están dadas por los vectores propios de $A^T A$ y la matriz ortogonal V ($V^T V = I$) contiene en sus columnas los vectores propios de AA^T .

De acuerdo a lo mencionado, las formas anteriormente expuestas son equivalentes a emplear el procedimiento *SVD* para la determinación de los vectores que conforman la base. Además por la complejidad impuesta por el tamaño de las matrices, concentraremos nuestros esfuerzos en la base generada por V

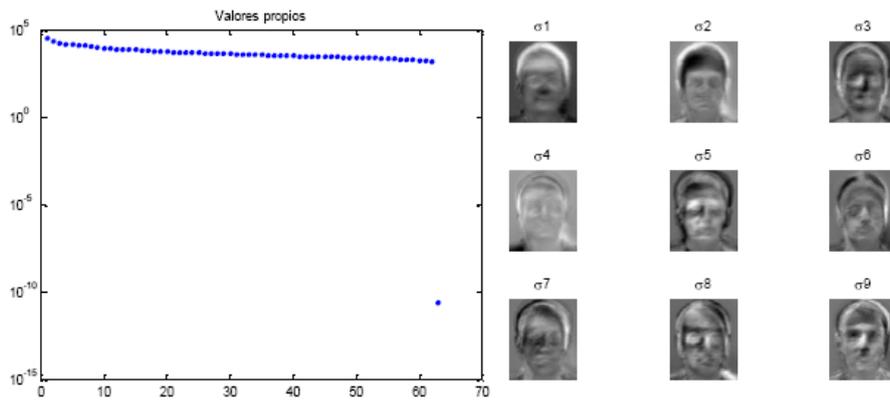


Figura 11: Valores propios y grafica de los 9 primeros vectores propios U utilizando la factorización *SVD* en la matriz de caricaturas A

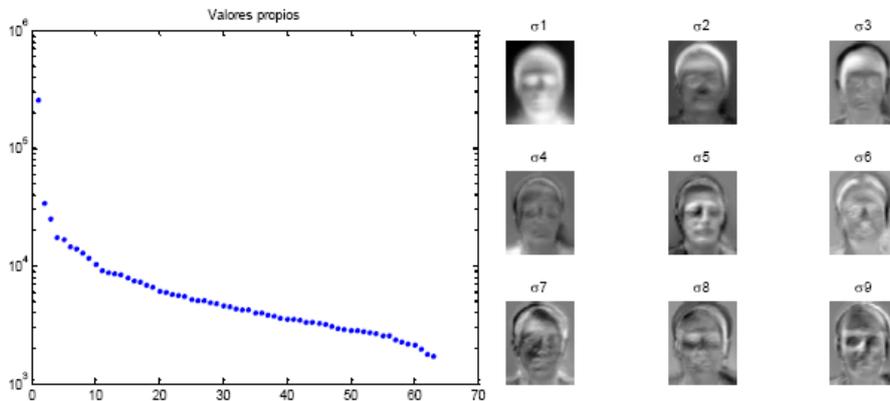


Figura 12: Valores propios y grafica de los 9 primeros vectores propios U utilizando la factorización *SVD* en la matriz F

La Figura 13 indica los últimos siete elementos de la base o sea los últimos siete vectores propios de $A^T A$, se observa que estos elementos son mucho más claros que los primeros expuestos en la Figuras 14 -15

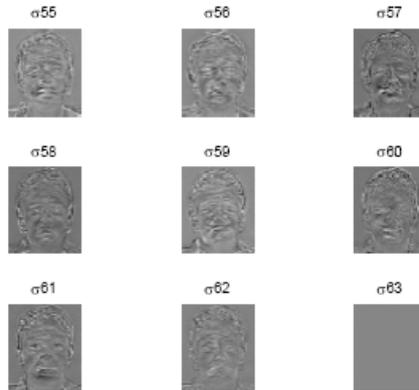


Figura 13: Grafica de los últimos 9 vectores propios $A^T A$

3.2. Reconstrucción

Obtenida la base generada por cualquier procedimiento anteriormente descrito una imagen facial $a(i)$ pueda ser expresada como una combinación lineal de los elementos de la base:

$$\tilde{a}(i) = \sum_{n=1}^r w(n)u(n) + \tilde{f} \quad (3.7)$$

Donde r es el rango de la matriz de caricaturas y $w(n)$ son los coeficientes de la expansión. Estos coeficientes los interpretamos como el *peso* con el cual el n -enésimo elemento de la base $u(n)$ contribuye a formar la foto. $a(i)$ algo similar a lo que se estudia en el análisis de Fourier (Figura 14).

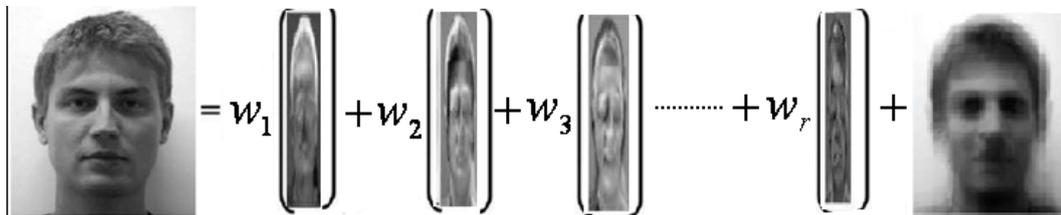


Figura14: Fotografía expresada como combinación lineal mas la cara promedio

Para observar el efecto de la contribución de las caras propias $u(n)$ en la reconstrucción de una cara $a(i)$, consideramos diferentes cantidad de elementos s para la base . La reconstrucción se mejora a medida que s aumenta como se observa en la Figura 15.

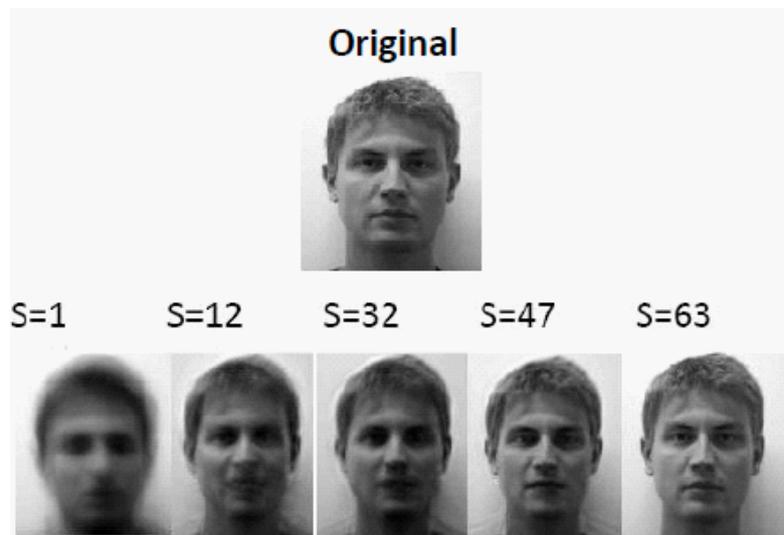


Figura 15: Reconstrucción de una imagen utilizando diferente cantidad de elementos de la base

El error de la reconstrucción se puede definir como:

$$error = \frac{\|f(i) - \tilde{a}(i)\|}{\|f(i)\|} \times 100$$

En la Figura 16 se muestra la gráfica Dimensión vs Error de la reconstrucción de la imagen correspondiente a la Figura 15, el error de la reconstrucción disminuye a medida que s aumenta.

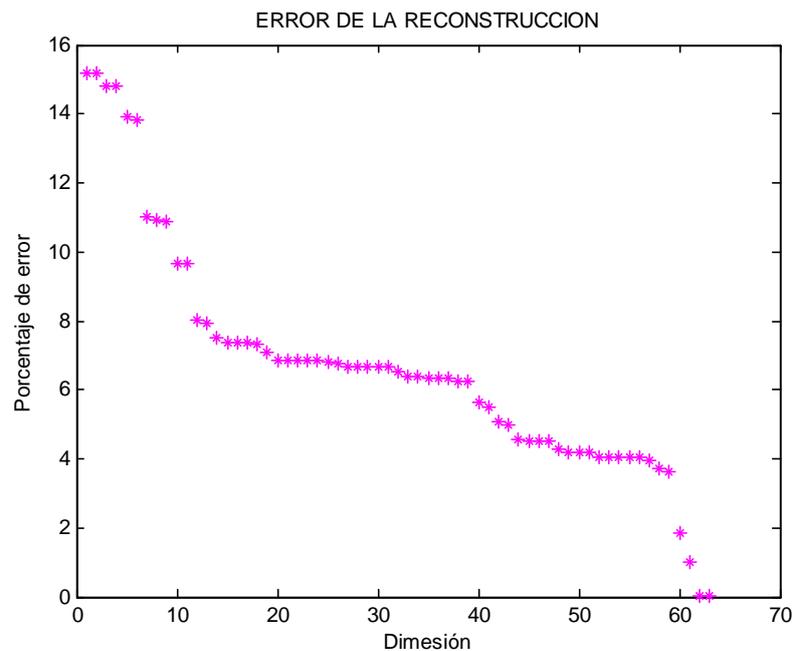


Figura 16: Dimensión vs Error de la reconstrucción

Se observa que hay un salto cuando $s = 10$ lo que nos indica que tomando solo 8 elementos de la base $u(n)$ tenemos una buena reconstrucción de la imagen $a(i)$. Este comportamiento nos sugiere que el procedimiento es rápidamente convergente, pues sólo necesitamos 8 caras propias de las 48000 caras propias. En la Figura 17 se observa un subconjunto de la base de caras de la Figura 5, vale la pena mencionar que esta base de datos tiene consideraciones de iluminación muy diferente a la considerada en la base de caras 1 o Figura 4.

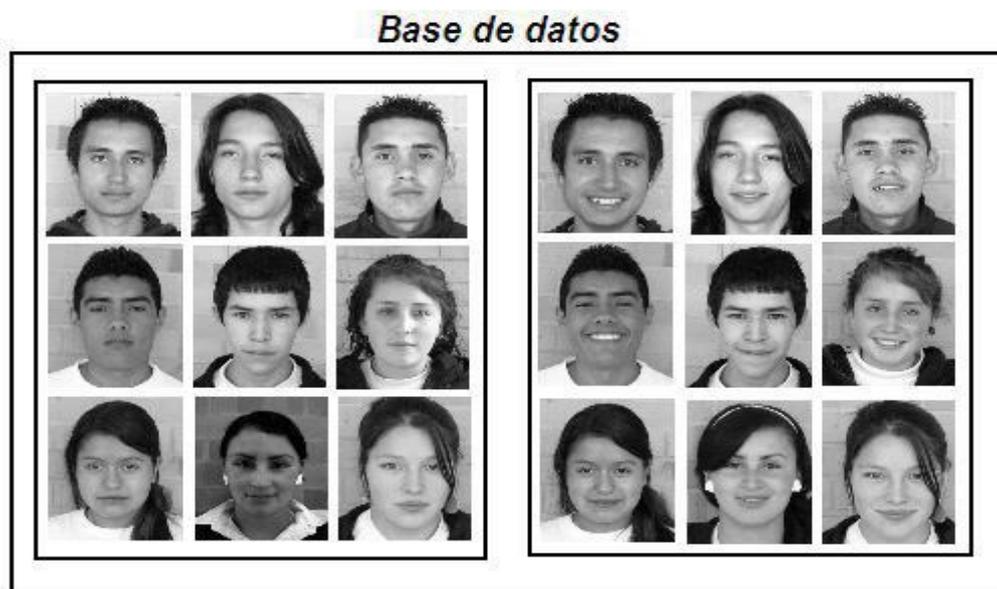


Figura 17: Subconjunto de la base de caras Colegio Nueva Delhi

A continuación se muestran dos fotografías con condiciones de iluminación diferentes (Figuras 18- y 19). Al realizar el mismo proceso de reconstrucción, utilizando como base de caras las fotos de la Figura 17 se observa que con cerca de la mitad de elementos de la base se obtiene una buena aproximación de la imagen original.

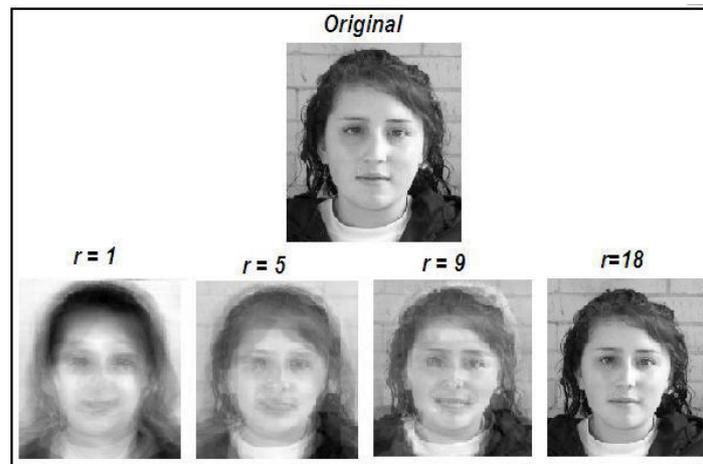


Figura 18

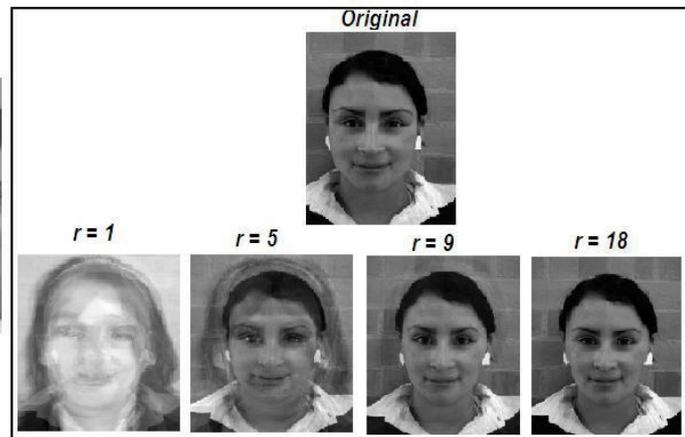


Figura 19

Capítulo 4

Identificación de rostros

En el capítulo anterior ilustramos el procedimiento matemático adecuado para reconstruir una imagen que se encuentra dentro de caras F . Así mismo calculamos el error de la reconstrucción cuando utilizamos diferente cantidad de elementos de la base empleada para la expansión. En este capítulo mostraremos las técnicas de identificación de rostros basado en el procedimiento de expansión en una base de caras. Como interés principal estudiaremos dos casos particulares:

A) Cuando la imagen de la persona que se va a identificar NO está en la base de caras pero si existen otras fotos de esa persona en la base

B) Cuando NO se tiene ninguna imagen de la persona en la base de caras. Esto corresponde al caso en que la persona no puede ser relacionada con alguien de la base y en principio los coeficientes de la expansión deben ser cero y por tanto no podemos tener similitud entre los elementos de la base y la persona que deseamos identificar.

De particular importancia para la identificación de rostros es necesario establecer un valor umbral para evitar falsas identificaciones o para evitar que se haga muy exigente el procedimiento de la identificación.

4.1. Comparación de fotografías

Con el propósito de determinar si la foto digital de la cara de una persona, que denominamos a , o *imagen de prueba*, se encuentra dentro de la base de caras, procedemos de la siguiente manera: determinamos la proyección de imagen a en el espacio ortogonal formado por las *caras propias* :

$$a = \sum_{i=1}^r w(n)u(n) + \tilde{f} \quad (4.1)$$

Donde r es el rango de la matriz de caricaturas, $u(n)$ es el n -enésimo elemento de la base de la matriz de caricaturas y $w(n)$ representa el *peso* o sea la contribución del n -enésimo término de la base para formar la cara. Al conjunto de pesos $w(n)$ de la imagen digital a o *imagen de prueba* lo denominamos δ_a ;

$$\delta_a = \{w_1, w_2, w_3, \dots, w_r\} \quad i = 1 \dots n \quad (4.2)$$

Si deseamos establecer la similitud de a con alguna de las p fotos que forman la base de caras debemos comparar los pesos de a con los pesos de cada una de las caras que se encuentran en la base:

$$\delta_i = \{w_1^{(i)}, w_2^{(i)}, w_3^{(i)}, \dots, w_r^{(i)}\} \quad i = 1 \dots n \quad (4.3)$$

Con el fin de determinar la correspondencia, se comparan los δ_a con cada uno de los δ_i mediante la distancia euclidiana entre los vectores pesos:

$$\|\delta_a - \delta_i\| = l \quad (4.4)$$

donde l indica la distancia entre los vectores pesos. Definimos que hay correspondencia cuando esa distancia l es menos que un umbral previamente establecido h

$$l \leq h$$

A continuación vamos a ilustrar lo anterior. Consideremos la base de datos compuesta por 9 caras con expresiones neutras

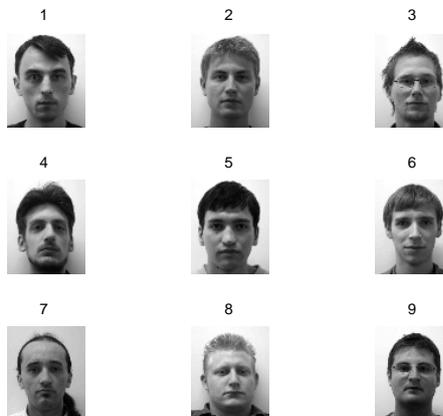


Figura 20: Base de caras de 9 personas con expresiones neutrales

Si la imagen de la persona que vamos a identificar es la indicada en la Figura 21, notéese que a simple vista esta persona si esta en la base de datos pero con expresión neutra.



Figura 21: Imagen de la cara de la persona que se desea identificar

Al correr el programa *identificacion.m* que esta escrito en el lenguaje *matlab* y cuyo pseudocódigo se encuentra en los Anexos con un umbral h podemos esperar que esta cara se identifique con la cara 2 de la base de caras ya que la imagen a es de esa persona pero con una expresión facial distinta. Vale la pena observar que si el umbral es demasiado bajo no podemos identificar la persona y si es muy alto, tenemos una correspondencia equivocada.

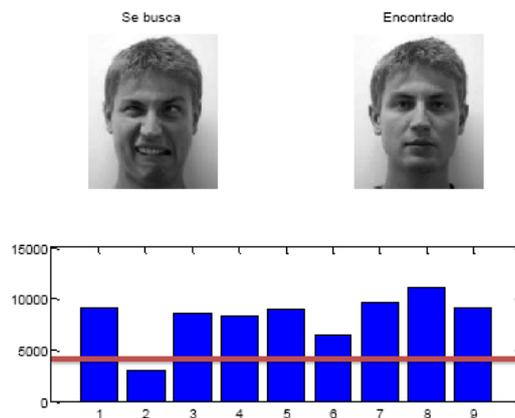


Figura 22: Proceso final al correr el programa

En el caso B en el cual la persona no esta en la base de caras como es el caso de la Figura 23 que en este caso es un estudiante del Colegio Nueva Delhi.

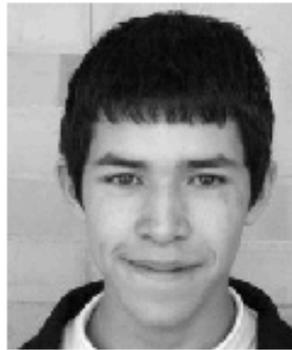


Figura 23: Imagen de la cara de la persona que se quiere identificar

Al realizar el anterior procedimiento con el mismo umbral h empleado para hacer la identificación, observamos que NO es posible asociar una persona en la base de datos con la imagen de la persona que queremos identificar y el programa debe responder como es obvio como persona no encontrada pues la cara de la persona no es parecida a ninguna de las consideradas en la base al menos para el umbral anterior, si subiéramos el umbral obtendríamos una correspondencia equivocada.

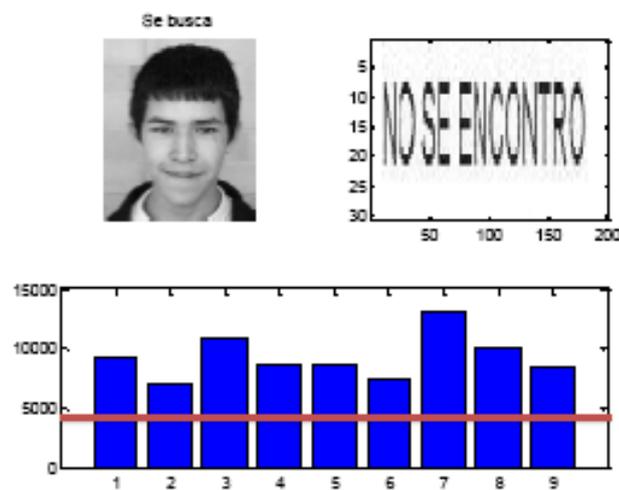


Figura 24: Proceso final al correr el programa

Aunque lo anterior se realizo para una base de caras relativamente pequeña (9 caras) y con una base de prueba de 9 estudiantes con expresion neutra, al correr el programa en todos los casos este arrojo *no se encontro*, el procedimiento es el mismo cuando tenemos una base mayor.

4.2. Aplicaciones

La teoría anteriormente expuesta nos ilustra la manera de como podemos, con ayuda del álgebra lineal, desarrollar procedimientos computacionales eficientes para identificar rostros de personas. Como aplicación final de este trabajo consideramos los casos de reconocimiento de personas NN e identificación en retratos hablados.

Identificación de NN

Esta aplicación es de suma utilidad en nuestro país dado la aparición de cadáveres NN que son producto del alto índice de inseguridad resultado del conflicto interno. Con el fin de abordar, en una etapa inicial, este problema, hemos considerado una base de datos de personas con expresiones faciales neutras parecidas a las empleadas por la Registraduría del Estado Civil en las imágenes que aparecen en la cédula de ciudadanía. Vale la pena recalcar que la exigencia de la Registraduría Nacional al tomar una foto para la cédula la persona debe tener los ojos abiertos, tener expresión neutra y la imagen se toma sobre un fondo blanco requisito que es semejante a la expresión neutra de los individuos de la Figura 25

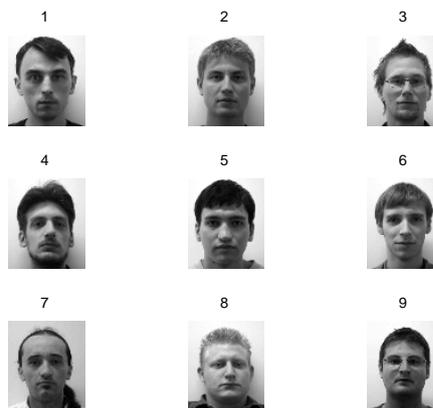


Figura 25: Simulación de una base de datos de la Registraduría

Supongamos para el caso de identificación de una persona NN que no tiene mutilaciones faciales y que tiene los ojos cerrados, hemos considerado la simulación con una imagen como la indicada en la Figura 26, vale la pena anotar

que la persona NN a identificar debe estar en la base de fotografías de la Registraduría Civil ya que el hecho de una persona colombiana no tenga documento de identidad, implica que su vida cotidiana se vea afectada en aspectos tan sencillos como la imposibilidad de cambiar un cheque en un banco, abordar un avión, hacer cualquier trámite notarial, pedir una cita médica en su EPS, sacar el RUT, tomar posesión de un cargo, etc. Las *leyes 757 de 2002* y *999 de 2005* señalan que todos los colombianos deben renovar su cédula de ciudadanía, ya que a partir del 31 de julio de 2010 el único documento de identificación válido en el país para los mayores de edad será la cédula amarilla con hologramas, documento de identidad que le permite a Colombia contar con un sistema de identificación biométrico con tecnología de punta y altos estándares de seguridad que impiden su falsificación.



Figura 26: Simulación de una persona muerta y cuya identidad es desconocida

Al realizar el procedimiento explicado en las secciones anteriores, en el programa *NN.m* el pseudocódigo es parecido al de *identificacion.m*, se diferencia en que la base de caras son solo fotografías de personas con expresiones neutras (Figura 25) y la base de prueba son fotografías de personas con ojos cerrados, encontramos que con un cierto umbral, logramos encontrar la imagen de la persona viva con ojos abiertos como se indica en la Figura 27.

Para las dos bases trabajadas se encontró concordancia entre la persona buscada y la persona encontrada por tanto consideramos que los resultados en este caso son satisfactorios y es un inicio estimulante para emprender el estudio profundo al caso de identificación de personas NN. Un reto no abordado aquí y que dejaremos para una investigación futura es el caso de personas registradas como NN con deformaciones faciales

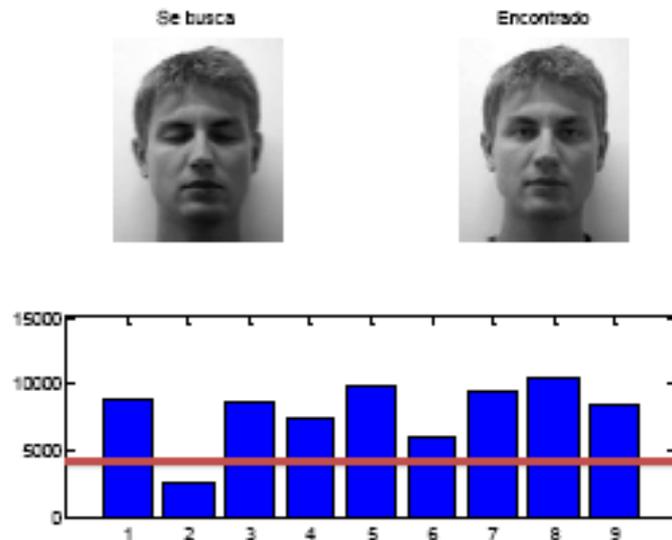


Figura 27: Proceso final al correr el programa

Identificación de retrato hablado

Otra aplicación muy interesante que se desprende de lo explicado arriba es el reconocimiento de retratos hablados". Se define el retrato hablado como: *Una disciplina artística mediante la cual se elabora el retrato o rostro de una persona extraviada o cuya identidad se ignora tomando como base los datos fisonómicos aportados por testigos e individuos que conocieron o tuvieron a la vista a quien se describe* [12].

En Colombia por el alto grado de delincuencia el retrato hablado es utilizado por las autoridades de nuestro país para iniciar la búsqueda de sospechosos que participaron en una actividad judicial por tanto constituye un factor importante en el reconocimiento de delincuentes un ejemplo de este es el caso de alias *Bochas* (Figura 28), el autor material del crimen de Jaime Garzón. Luego de su captura se hizo el cotejo y las autoridades sostuvieron que fue asombroso el parecido del retrato hablado con el individuo [13].



Figura 28: Retrato hablado de "Bochas"

En el retrato hablado solo figuran rasgos significativos y por tanto no hay precisión de la persona como en las fotos. Para llevar a cabo el experimento con el programa que desarrollamos, consideramos la imagen de un sospechoso como la imagen de la Figura 29



Figura 29: Simulación de un retrato hablado

Aquí esta imagen se obtuvo por medio del programa Corel Draw realizando un filtro de Canny para delimitar los bordes y en últimas esta imagen puede corresponder a una imagen con las características de un retrato hablado. Al realizar el reconocimiento, con el programa *retratohablado.m* el pseudocódigo es parecido al de *identificacion.m*, se diferencia en que la base de caras son solo fotografías de personas con expresiones neutras (Figura 25) y la base de prueba¹ son simulaciones de retratos hablados, con un cierto umbral, logramos encontrar la correspondencia mostrada en la Figura 30

¹La base de prueba para los retratos hablados, está almacenado en el archivo *retratos.mat* en el lenguaje *matlab* es una matriz $P_{9 \times 48000}$, donde cada fila representa el retrato hablado de un individuo y está disponible en psantanac@unal.edu.co

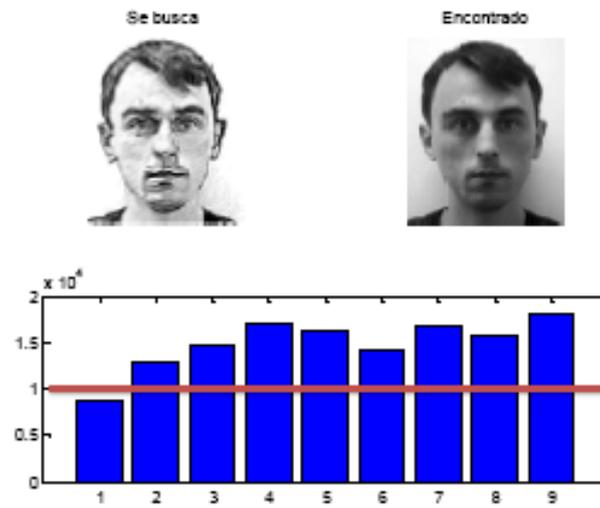


Figura 30: Proceso final al correr el programa

Aquí el valor del umbral para el reconocimiento es un poco mas difícil de definir dada las características de la imagen del retrato hablado. En la siguiente imagen se observa que para un umbral mas alto o mas bajo no se obtiene un reconocimiento adecuado y podemos decir que el método falla.

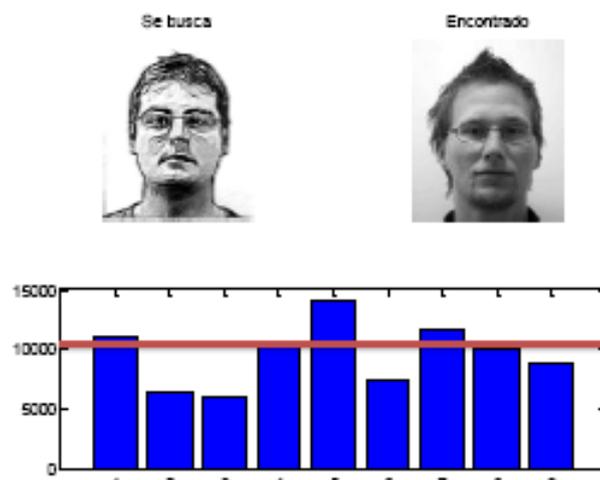


Figura 31: Proceso final al correr el programa

Capítulo 5

Conclusiones

En este trabajo se abordó el tema de la identificación de rostros utilizando el álgebra lineal y técnicas computacionales basados en el lenguaje *matlab*. Se codificaron fotografías para formar una base de datos y así construir una matriz que se le denominó *base de caras*, se trabajaron con dos bases de caras una de 63 fotografías en la cual habían 9 hombres cada uno con 7 expresiones faciales y otra de 80 fotografías de estudiantes del Colegio Nueva Delhi la cual la conformaban 16 personas cada una con 5 expresiones, a cada base de caras se le encontró un conjunto de vectores que conformaron una base ortogonales, para la formación de esta se utilizó la factorización *SVD*. Con ello se realizaron esencialmente tres cosas: Reconstrucción de cada fotografía, Identificación de rostros y Aplicaciones a la identificación de rostros.

Denominamos reconstrucción de la fotografía al poder representar una fotografía como una combinación lineal de elementos de la base, los diferentes experimentos mostraron que con una base de datos de caras de 63 fotografías que posee buenas condiciones de luz, se necesitan menos de la mitad de los elementos de la base para hacer una buena reconstrucción ya que al comparar la fotografía original con la reconstrucción el grado de error es mínimo.

Con una base de caras de 80 fotografías con diferentes condiciones de luz se necesitan cerca de la mitad de los elementos de la base para tener una buena reconstrucción. Teniendo en cuenta que la cantidad de elementos de la base concuerda con la cantidad de fotografías, esto indica que este procedimiento converge rápidamente.

En la identificación de rostros fue centro de interés dos casos: Cuando la

imagen de la persona a la que se va a identificar no está en la base de caras pero si existen otra foto de esa persona en dicha base y cuando no se tiene ninguna imagen de la persona en la base de caras.

Los experimentos para el primer caso arrojaron que con la utilización de 63 fotografías de hombres dividida en una base de caras de 9 fotografías de personas con expresión neutra y una base de prueba de 54 fotografías de las mismas personas, en la cual cada individuo tiene 6 expresiones diferentes a la expresión neutra. Al compararlas se obtuvo 5 errores correspondiente al 9,2% de error, con un umbral de 4900.

Con el uso de 80 fotografías de estudiantes del Colegio Nueva Delhi dividida en una base de caras con las 16 personas con expresión neutra y una base de prueba de 63 fotografías de las mismas personas, en la cual cada individuo tiene 4 expresiones diferentes a la expresión neutra. Al compararlas se obtuvo 8 errores correspondiente al 14% de error, manejando un umbral de 4900.

Para el segundo caso en el reconocimiento de imágenes faciales en el que no se tiene ninguna imagen de la persona en la base de caras, manejando el mismo umbral, con una base de caras de 9 hombres con expresiones neutras y una base de prueba de 9 estudiantes también con expresiones neutras al realizar el proceso de identificación en todos los casos el programa arroja *no se encuentra*

En este trabajo se trato en una etapa inicial dos aplicaciones en el reconocimiento de imágenes faciales: La identificación de NN e Identificación de retratos hablados, temas que no se encuentran en la literatura especializada. Para la identificación de NN se uso una base de caras 9 fotografías con expresiones neutras se comparo con una base prueba de 9 personas con ojos cerrados que simularon personas NN. Al comparar dichas bases todas las fotografías concor-daron, utilizando un umbral de 4900. En el caso del retrato hablado se utilizo la misma base de de caras y la base de prueba que representaba un retratos habla-dos se simulo con el programa Corel Draw con un umbral de 10000 se obtuvo 1 error que correspondiente al 11%

Con base de caras pequeñas el reconocimiento de imágenes faciales con las que se experimentaron en este trabajo y utilizando el procedimiento *SVD*, el programa no gasta mucho tiempo en su ejecución y se obtuvo una tasa de rechazo pequeña, teniendo en cuenta que posee limitantes como la baja resolución en todas las fotografías, ya que estas tienen dimensiones 240×200 .

Falta experimentar la eficiencia del procedimiento *SVD* en el reconocimieto

facial, donde la imagen que se quiere reconocer no está en condiciones óptimas, es decir puede que la imagen este borrosa, muy oscura o con mucha luz.

Con respecto a las aplicaciones este trabajo, se sugiere experimentar con fotos de personas NN reales y con retratos hablados pintados por un dibujante para obtener resultados más fieles a la realidad respecto al reconocimiento.

Capítulo 6

Apéndices

6.1. Introducción

Para resolver un sistema lineal de ecuaciones

$$Av = \lambda v$$

es necesario factorizar la matriz A , y para ello existen varias formas de hacerlo:

- Método LU que consiste en escribir A como

$$A = PLU$$

en donde P es una matriz de permutación, L es una matriz triangular inferior y U es una matriz triangular superior.

- Método QR : la matriz A se escribe como

$$A = QR$$

donde Q es una matriz unitaria, R es una matriz triangular superior.

- Método SVD en donde

$$A = U\Sigma V^T$$

U y V son matrices unitarias y Σ es una matriz diagonal.

En este apéndice consideraremos la factorización SVD y su relación con los valores y vectores singulares. Este procedimiento es útil para nuestro propósito ya que tenemos una matriz de base de fotografías F o una matriz de caricaturas A , ambas NO cuadradas

6.2. Definición de SVD

Dada una matriz, no necesariamente cuadrada, $A_{m \times n}$, la descomposición *SVD* es una factorización que se escribe como:

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T \quad (6.1)$$

Esta factorización *SVD*, presenta las siguientes características:

- La matriz Σ es una matriz diagonal en donde los elementos de la diagonal principal son los valores singulares descendentes¹ de $A^T A$ o AA^T .
- La matriz U es ortonormal $U^T U = I$ y esta formada por los vectores propios de $A^T A$.
- La matriz V es ortonormal $V^T V = I$ y esta formada por los vectores propios de AA^T .

La matriz $A \in \mathbb{R}^{m \times n}$, es una matriz NO cuadrada y puede ser vista como una transformación lineal

$$A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

En donde el vector x definido en \mathbb{R}^n le corresponde la imagen Ax que pertenece a \mathbb{R}^m . Dos espacios fundamentales asociados con la transformación lineal son la imagen de A y el espacio *nulo*.

- La imagen de A es un subespacio de \mathbb{R}^m que esta definida como

$$\text{Imagen}(A) = \{Ax \in \mathbb{R}^m \mid x \in \mathbb{R}^n\}$$

- El *rango* es la dimensión de la imagen de A y se define como

$$\rho(A) = \dim(\text{Imagen}(A))$$

el rango también es definido como la dimensión del espacio de las columnas de A .

¹Los valores singulares son raíces cuadradas de los valores propios

- El *espacio nulo* es un subespacio de \mathbb{R}^m y se define como

$$N(A) = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

- La *nulidad* es la dimensión del espacio nulo de A y se define como

$$v(A) = \dim(N(A))$$

Una relación entre la nulidad y el rango de una matriz $A_{m \times n}$ esta dada mediante la siguiente relación

$$n = \rho(A) + v(A)$$

es decir que el *rango de A* mas la *nulidad de A* es igual al número de columnas de A .

La definición de *SVD* se fundamenta en las matrices $(A^T A)_{n \times n}$ o $(AA^T)_{m \times m}$. Podemos observar las siguientes propiedades de estas matrices y su relación con A y A^T :

Lema: $A^T A$ y AA^T son simétricas²

Demostración $(A^T A)^T = A^T A$ y $(AA^T)^T = (AA^T)$

Lema: $N(A^T A) = N(A)$

Demostración: Demostrar que $N(A) \subseteq N(A^T A)$ si $x \in N(A)$ entonces $x \in \mathbb{R}^n$ con $Ax = 0$, si multiplicamos por A^T a la izquierda se tiene que $A^T Ax = 0$ luego $x \in N(A^T A)$. Demostraremos que $N(A^T A) \subseteq N(A)$. Suponemos que $x \in N(A^T A)$. Entonces $A^T Ax = 0$ recordemos que el producto interno estándar es $(w, z) = z^T w$. Podemos observar que $(A^T Ax, x) = (Ax, Ax) = 0$. Por lo tanto $Ax = 0$. Esto significa que $x \in N(A)$.

Lema: $Rango(A^T A) = Rango(A) = Rango(A^T) = Rango(AA^T)$

Demostración: $Rango(A^T A) = m - \dim(N(A^T A)) = m - \dim(N(A)) = Rango(A)$

Lema: Si v es un vector propio de $A^T A$ asociado con un valor propio diferente de cero λ . Entonces Av es un vector propio de AA^T , asociado con el mismo valor propio.

Demostración $AA^T(Av) = A(A^T A)v = A(\lambda v) = \lambda(Av)$

²Una matriz es simétrica si $A^T = A$

Teorema: El rango de una matriz $A \in \mathbb{R}^{m \times n}$ que tiene r valores singulares no nulos $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ con $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$ entonces el rango de A es r

Demostración: Sean $\{v_1, v_2, \dots, v_n\}$ una base de \mathbb{R}^n formada por los vectores propios asociados a $A^T A$, de tal forma que los valores propios correspondientes a $A^T A$ satisfacen que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ entonces:

$$\begin{aligned} Av_i Av_j &= (Av_i)^T Av_j = v_i (A^T Av_j) = v_i \lambda_j v_j \\ Av_i Av_j &= \sigma_j^2 (v_i v_j) = \begin{cases} 0 & \text{si } i \neq j \\ \sigma_j^2 & \text{si } i = j \end{cases} \end{aligned}$$

Luego $[Av_1, Av_2, \dots, Av_n]$ es ortogonal. Sea r el número de valores propios no nulos de A , esto es r es el número de valores propios no nulos de $A^T A$. Se tiene que $Av_i \neq 0$ si y solo si $1 \leq i \leq r$ entonces $[Av_1, Av_2, \dots, Av_r]$ son linealmente independientes lo cuales pertenecen al espacio de las columnas de A denotado $\text{col}(A)$. Supongamos que $y = Ax$ donde x se puede expresar como: $x = c_1 v_1 + c_2 v_2 + \dots + c_n v_n$ y

$$\begin{aligned} y &= Ax = c_1 Av_1 + c_2 Av_2 + \dots + c_r Av_r + c_{r+1} Av_{r+1} + \dots + c_n Av_n \\ y &= Ax = c_1 Av_1 + c_2 Av_2 + \dots + c_r Av_r + 0 + \dots + 0 \end{aligned}$$

Entonces y esta en el espacio generado por $[Av_1, Av_2, \dots, Av_r]$ que es ortogonal para $\text{col}(A)$ por tanto $\rho(A) = r$

Teorema: SVD :Sea $A \in \mathbb{R}^{n \times m}$, con rango r . Entonces existen números reales $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$ y una base ortonormal $v_1, v_2, \dots, v_m \in \mathbb{R}^m$ y una base ortonormal $u_1, u_2, \dots, u_m \in \mathbb{R}^n$ tal que:

$$Av_i = \sigma_i u_i \quad \forall i = 1 \dots r \quad A^T u_i = \sigma_i v_i \quad \forall i = 1 \dots r \quad (6.2)$$

$$Av_i = 0 \quad \forall i = r + 1 \dots n \quad A^T u_i = 0 \quad \forall i \text{ desde } r + 1 \text{ hasta } n \quad (6.3)$$

Las ecuaciones anteriores implican que v_1, v_2, \dots, v_m son valores propios de $A^T A$, y u_1, u_2, \dots, u_m son valores propios de AA^T , además $\sigma_1, \sigma_2, \dots, \sigma_r$ son los valores propios diferentes de cero para $A^T A$ y AA^T .

Demostración: Sea v_1, v_2, \dots, v_m es una base ortonormal de \mathbb{R}^m consistente en los autovectores de $A^T A$ y $\lambda_1, \lambda_2, \dots, \lambda_m$ los valores propios asociados. Como $A^T A$ es semidefinida positiva, todos sus valores propios son positivos. Asumimos que v_1, v_2, \dots, v_m son ordenados tal que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$. Como $r = \text{rango}(A) = \text{rango}(A^T A)$ debe pasar que $\lambda_r > 0$ y $\lambda_{r+1} = \lambda_{r+2} = \dots = \lambda_m = 0$ para toda $i = 1 \dots n$. Definimos σ_i y u_i como

$$\sigma_i = \|Av_i\|_2 \quad \text{y} \quad u_i = \frac{1}{\sigma_i} Av_i \quad (6.4)$$

Esta definición implica que $Av_i = \sigma_i u_i$ y $\|u_i\| = 1$ para toda $i = 1, \dots, r$. Tenemos que u_1, u_2, \dots, u_r son ortonormales, ya que $\sigma_i^2 = \|Av_i\|_2^2 = (Av_i, Av_i) = (A^T Av_i, v_i) = (\lambda_i v_i, v_i) = \lambda_i$

$$A^T u_i = (1/\sigma_i) A^T Av_i = (\lambda_i/\sigma_i) v_i = \sigma_i v_i$$

Como $r < n$ Los vectores u_1, u_2, \dots, u_r son vectores propios de $AA^T \in \mathbb{R}^{n \times n}$ y el $\text{rango}(AA^T) = r$. la dimensión del espacio nulo, la nulidad $v(A) = n - r$. Sean $u_{r+1}, u_{r+2}, \dots, u_n$ una base ortonormal de $N(AA^T)$ que constituyen los vectores propios de AA^T asociados con el valor propio 0. como $u_{r+1}, u_{r+2}, \dots, u_m$ son ortonormales a u_1, u_2, \dots, u_r . Así u_1, u_2, \dots, u_r es una base ortonormal de \mathbb{R}^n . Como $N(AA^T) = N(A^T)$ nosotros tenemos que $A^T u_i = 0$ para $i = r + 1 \dots n$

Es de aclarar que:

- $\sigma_1, \sigma_2, \dots, \sigma_r$ se le denominan valores singulares
- v_1, v_2, \dots, v_m se le denominan vectores singulares a derecha
- u_1, u_2, \dots, u_r se le denominan vectores singulares a izquierda

El teorema anterior se puede ilustrar mediante el siguiente gráfico.

Pueden ser combinación de matrices singulares

$$A [v_1, \dots, v_r | v_{r+1}, \dots, v_n] = [u_1, \dots, u_r | u_{r+1}, \dots, u_n] \begin{bmatrix} \sigma_1 & & \mathbf{0} & | & \\ & \ddots & & & \\ \mathbf{0} & & \sigma_r & | & \mathbf{0} \\ \hline & & \mathbf{0} & | & \mathbf{0} \end{bmatrix}$$

Significa que $AV = U\Sigma$. Como V es unitaria o sea $VV^T = I$ se tiene luego que: $A = U\Sigma V^T$

Como consecuencia del anterior Teorema se tiene el siguiente Lema

Lema: Sea $A \in \mathbb{R}^{n \times m}$ con rango r , entonces existe $\hat{U} \in \mathbb{R}^{n \times r}$, $\hat{\Sigma} \in \mathbb{R}^{r \times r}$ y $\hat{V} \in \mathbb{R}^{m \times r}$, tal que \hat{U} y \hat{V} son unitarias, $\hat{\Sigma}$ es una matriz diagonal en donde los elementos de la diagonal principal son $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$ y

$$A = \hat{U}\hat{\Sigma}\hat{V}^T \tag{6.7}$$

El Lema anterior se puede ilustrar en la siguiente Figura 35, donde se compara $A = U\Sigma V^T = \hat{U}\hat{\Sigma}\hat{V}^T$

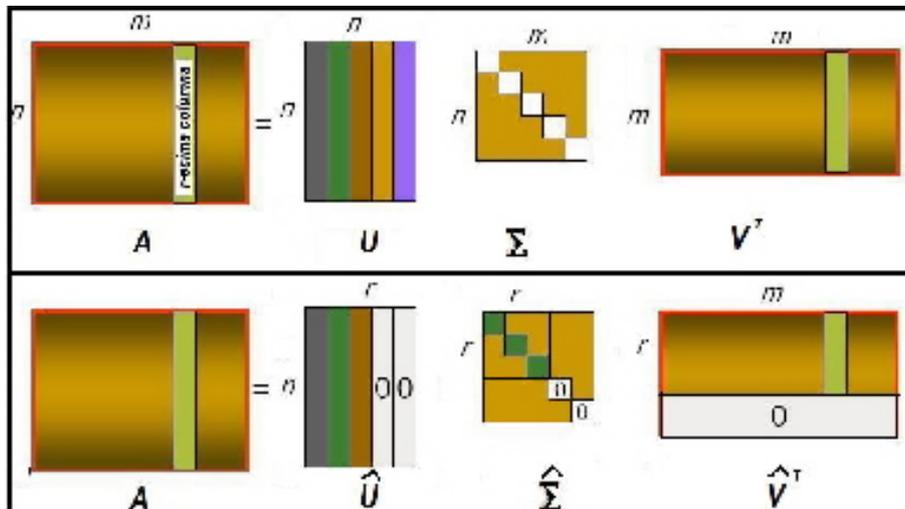


Figura 35: Descomposición SVD

Teorema: Sea $A \in \mathbb{R}^{n \times m}$ con rango r , Sean $\sigma_1, \sigma_2, \dots, \sigma_r$ los valores singulares de A asociados con sus vectores singulares a derecha e izquierda v_1, v_2, \dots, v_m y u_1, u_2, \dots, u_r respectivamente. Entonces

$$A = \sum_{i=1}^r \sigma_i u_i v_i^T \quad (6.8)$$

Demostración: Sea $B = \sum_{i=1}^r \sigma_i u_i v_i^T \in \mathbb{R}^{n \times m}$. Mostraremos que $A = B$, es suficiente mostrar $Av_j = Bv_j$, para $i = 1, \dots, m$. Como v_1, v_2, \dots, v_m es una base para \mathbb{R}^m . Si $i \leq r$, nosotros tenemos que $Av_i = \sigma_i u_i$ y $Bv_j = \sum_{i=1}^r \sigma_i u_i (v_i^T v_j)$. Como v_1, v_2, \dots, v_m es ortonormal, entonces $v_i^T v_j = 0$ si $i \neq j$ y $v_i^T v_j = 1$ si $i = j$. Por lo tanto todos los términos en la sumatoria son cero, excepto el i -ésimo término, por lo tanto $Bv_i = \sigma_i u_i$. Si $i > r$ entonces $Av_j = 0$ y $Bv_j = \sum_{i=1}^r \sigma_i u_i (v_i^T v_j) = \sum_{i=1}^r \sigma_i u_i (0) = 0$

Teorema . Sea $A \in \mathbb{R}^{n \times m}$ con rango r .

Las primeras r columnas de de U forman una base ortonormal para el espacio de columnas de A

Las primeras r columnas de de V forman una base ortonormal para el espacio de filas de A

6.3. Ejemplo de la factorización de A en la forma SVD

El proceso para factorizar una matriz A en la forma $A = U\Sigma V^T = UV^T$, lo describiremos a continuación mediante un ejemplo. Dada la matriz A no cuadrada 4×2

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Para calcular la representación de A como SVD, debemos encontrar los valores propios y vectores propios de AA^T y $A^T A$. Los vectores propios de AA^T forman las columnas de U y los vectores propios de $A^T A$ forman las columnas de V . Así mismo los valores singulares en Σ son las raíces cuadradas de los valores propios de AA^T o $A^T A$. Estos valores singulares son los elementos de la diagonal de Σ que se colocan en orden descendente.

Para nuestro caso.

$$\begin{aligned}
 AA^T &= \begin{pmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{pmatrix} \\
 &= \begin{pmatrix} 20 & 14 & 0 & 0 \\ 14 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = W
 \end{aligned}$$

Valores propios de la matriz AA^T son 29,883, 0,117, 0, 0. Con los vectores propios de W formamos la matriz U

$$U = \begin{pmatrix} 0,82 & -0,58 & 0 & 0 \\ 0,58 & 0,82 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Similarmente a partir de $A^T A$ determinamos la matriz V

$$AA^T = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 11 \\ 11 & 25 \end{pmatrix}$$

Por lo tanto

$$V = \begin{pmatrix} 0,40 & -0,91 \\ 0,91 & 0,40 \end{pmatrix}$$

Finalmente la matriz Σ

$$\Sigma = \begin{pmatrix} 5,47 & 0 \\ 0 & 0,37 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Así

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 3 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} = U\Sigma V^T = \begin{pmatrix} 0,82 & -0,58 & 0 & 0 \\ 0,58 & 0,82 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5,47 & 0 \\ 0 & 0,37 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0,40 & 0,91 \\ -0,91 & 0,40 \end{pmatrix}$$

$$A = \begin{pmatrix} 4 & 0 \\ 3 & -5 \end{pmatrix} \quad (6.9)$$

6.4. Cálculo de valores propios

La factorización SVD consiste esencialmente en encontrar los valores propios y vectores propios de las matrices simétricas $A^T A$ o AA^T , y eso es, encontrar los ceros del polinomio característico

$$|A^T A - \lambda I| = 0$$

Usualmente, para nuestro caso de identificación de caras, las matrices tiene una dimensión alta y es necesario hacer uso de algún método numérico eficiente para encontrar los ceros de dicho polinomio. La estrategia establecida divide el problema en dos partes, primero se reduce la matriz original $A^T A$ a una matriz B de estructura más simple (en el sentido que tiene mayor cantidad de ceros) que puede ser una matriz de Hessenberg³, una matriz tridiagonal⁴ o matriz bidiagonal⁵ (Figura 36) a las que es más sencillo determinar sus valores propios.



Figura 36: Matrices Hessenberg, Tridiagonal, Bidiagonal

³ $A = a_{ij}$ es matriz Hessenberg superior si $a_{ij} = 0$ para todo $i > j + 1$

⁴ $A = a_{ij}$ es matriz Tridiagonal si a_{ii} , $a_{i,i+1}$ y $a_{i+1,i}$ son diferentes de cero

⁵ $A = a_{ij}$ es matriz Bidiagonal superior si a_{ii} , $a_{i,i+1}$ son diferentes de cero

6.4.1. Reducción de matrices a formas más simples

Estos métodos se basan en la transformación de similitud o semejanza

Definición: Dos matrices A y B se dicen semejantes si existe una matriz $T_{n \times n}$ no singular⁶ tal que:

$$B = T^{-1}AT \quad (6.10)$$

Lema: La transformación de similitud tiene de preservar los valores propios, vectores propios y poseer el mismo polinomio característico

Demostración: Preservar los valores y vectores propios.

Sea $Ax = \lambda x$

Si multiplicamos por T^{-1} tenemos $T^{-1}Ax = \lambda T^{-1}x$

Si definimos $T^{-1}x = y$ tenemos que $x = Ty$

Obteniendo $T^{-1}ATy = \lambda T^{-1}Ty = \lambda y$

Es decir $By = \lambda y$

Preservar el mismo polinomio característico

$$p_b(\lambda) = |B - \lambda I| = |T^{-1}AT - \lambda T^{-1}T| = |T^{-1}(A - \lambda I)T| = |T^{-1}| |A - \lambda I| |T| = |A - \lambda I| = p_a(\lambda).$$

La clase de similitud es muy amplia. Dentro de ella está incluida un tipo de transformación mucho más restrictiva, la transformación ortogonal, en este caso la matriz de transformación es ortogonal lo que significa que su transpuesta es su inversa

$$T^{-1} = T^T$$

La estrategia para calcular los valores propios de una matriz A consiste en utilizar una sucesión de transformaciones de semejanza

$$\begin{aligned} A^0 &= A \\ A^i &= T_i^{-1}A^{(i-1)}T_i \text{ Para } i = 1, 2, \dots, m \end{aligned} \quad (6.11)$$

Para gradualmente transformar la matriz A en una matriz B de forma más simple

$$B = A^m = T^{-1}AT \text{ Donde } T = T_1T_2\dots T_m \quad (6.12)$$

Los métodos de factorización se basan en que una matriz A puede ser factorizada en una matriz de la forma

$$A = F_L F_R \quad (6.13)$$

⁶Una matriz T es singular si posee inversa

donde F_L es una matriz a izquierda F_R es una matriz a derecha Si invertimos el orden de los factores obtenemos una matriz $\hat{A} = F_R F_L = F_L^{-1} A F_L$ que es semejante a la matriz A por medio de la transformación F_L . Así cualquier factorización de dos matrices al ser multiplicadas en orden inverso constituyen una transformación de semejanza.

Reducción Householder

El algoritmo de Householder produce ceros a partir de transformaciones ortogonales y cada transformación está diseñada para formar ceros una parte de toda una columna y su correspondiente fila si la matriz es simétrica, este algoritmo hace uso de una matriz ortogonal simétrica P , denominada la matriz de Householder

$$P = I - 2ww^t$$

En donde w es un vector real con $\|w\|_2^2 = 1$, P tiene la propiedad de ser simétrica

$$\begin{aligned} P^t &= (I - 2ww^t)^t \\ P^t &= I - 2(ww^t)^t \\ P^t &= I - 2(w^t)^t w^t \\ P^t &= I - 2ww^t \\ P^t &= P \end{aligned}$$

P tiene la propiedad de ser ortogonal

$$\begin{aligned} P^2 &= (I - 2ww^t)(I - 2ww^t) \\ P^2 &= I - 4ww^t + 4w(w^t w)w^t \\ P^2 &= I \end{aligned}$$

Así $P^{-1} = P$, pero como $P^t = P$ se tiene que $P^{-1} = P^t$ con lo que se concluye la prueba de la ortogonalidad.

El vector w se escoge de tal forma que para un vector x dado se cumpla que:

$$Px = ke$$

Donde e es el vector $[1, 0, 0, \dots, 0]^t$ y k es la norma euclidiana de x . Por tanto la matriz P actúa sobre un vector dado x llevando a cero todas sus componentes excepto el primer término.

Para reducir una matriz simétrica A en la forma tridiagonal se construye una matriz de transformación T_1 como una matriz particionada que tiene 1 en la posición (1,1), una matriz de Householder P_1 de orden $(n-1) \times (n-1)$ y ceros en el resto de posiciones.

$$\mathbf{T}_1 = \left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] \begin{array}{c} \\ \\ \\ \mathbf{P}_1 \\ \end{array}$$

La matriz T_1 es ortogonal y simétrica, entonces $T^{-1} = T_1^t = T$, se elige el vector x a los últimos $n-1$ elementos de la primera columna de A , $x = (a_{21}, a_{31}, \dots, a_{n1})$, en la multiplicación $T_1 A$ se obtiene:

$$\begin{aligned} \mathbf{T}_1 \mathbf{A} &= \left[\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] \left[\begin{array}{c|cccc} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \hline a_{21} & & & & \\ a_{31} & & & & \\ \vdots & & & & \\ a_{n1} & & & & \end{array} \right] \\ &= \left[\begin{array}{c|cccc} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \hline k & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] \begin{array}{c} \\ \\ \text{irrelevante} \\ \\ \end{array} \end{aligned}$$

La transformación ortogonal completa es:

$$\mathbf{T}_1 \mathbf{A} \mathbf{T}_1 = \left[\begin{array}{c|cccc} a_{11} & k & 0 & \cdots & 0 \\ \hline k & & & & \\ 0 & & & & \\ \vdots & & & & \\ 0 & & & & \end{array} \right] \begin{array}{c} \\ \\ \text{irrelevante} \\ \\ \end{array}$$

Por tanto se han producido ceros en la primera columna y fila que son necesarios para formar la diagonal, posteriormente para colocar ceros en la segunda

columna y fila, se construye la matriz de transformación T_2 :

$$\mathbf{T}_2 = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \mathbf{P}_2 & \\ 0 & 0 & & & \end{array} \right]$$

Donde P_2 es la matriz de Householder de tamaño $(n - 2) \times (n - 2)$ que actúa sobre un vector x que tiene componentes los $n - 2$ elementos inferiores de la segunda columna de A . El bloque identidad en la esquina superior derecha asegura que la tridiagonalización alcanzada el primer paso no sea destruida, P_2 crea una nueva fila y columna de la matriz tridiagonal, una sucesión de $(n - 2)$ transformaciones reduce a la matriz A a la forma tridiagonal, el trabajo realizado por este algoritmo contiene $2n^3/3$ multiplicaciones, esta estrategia puede ser implementada para reducir matrices no simétricas a la forma tridiagonal con un costo de $5n^3/3$ multiplicaciones

Polinomio característico de una matriz tridiagonal

Sea A una matriz tridiagonal simétrica

$$\mathbf{A} = \begin{bmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_n & \alpha_n & \\ & & & \beta_n & \alpha_n \end{bmatrix}$$

Sin pérdida de la generalidad se asume que A es una matriz tridiagonal *irreducible*, esto es que $\beta_i \neq 0$ para todo i . De lo contrario A puede escribirse como una matriz triangular por bloques cuyos bloques son matrices tridiagonales irreducibles $A^{(i)}, i = 1, \dots, k$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^{(1)} & & & \\ & \mathbf{A}^{(2)} & & \\ & & \ddots & \\ & & & \mathbf{A}^{(k)} \end{bmatrix}$$

El polinomio característico de A es el producto de los polinomios característicos de los bloques $A^{(i)}, i = 1, \dots, k$, entonces es suficiente considerar matrices

$$p_i(\lambda) = \det(\mathbf{A}_i - \lambda \mathbf{I}), \quad \mathbf{A}_i = \begin{bmatrix} \alpha_1 & \beta_2 & & \\ \beta_2 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_i \\ & & \beta_i & \alpha_i \end{bmatrix}$$

Calculando

$$\begin{aligned} p_0(\lambda) &= 1 \\ p_1(\lambda) &= (\alpha_1 - \lambda) \\ p_i(\lambda) &= (\alpha_i - \lambda)p_{i-1}(\lambda) - \beta_i^2 p_{i-2}(\lambda), \quad i = 2, 3, \dots, n \end{aligned}$$

finalmente $p(\lambda) = p_n(\lambda)$

Para un determinado valor λ este algoritmo es una forma para evaluar $p(\lambda)$. Se pueden utilizar métodos numéricos clásicos para determinar las raíces de un polinomio como método del dicótomo, el de la secante o el método de Newton.

6.5. Pseudocódigo

Para el reconocimiento facial se desarrollo un programa en el lenguaje *matlab* denominado *identificacion.m* el cual tiene el siguiente pseudocódigo:

- Carga la matriz de la base de caras y muestra las fotos que se encuentran en ella.
- Determina la cara promedio, muestra la imagen de la cara promedio y construye una matriz de caricaturas.
- Realiza la factorización *SVD* para la matriz de caricaturas
- Grafica los valores propios y normaliza los vectores propios
- Se le pide al usuario la dimensión del espacio de las caras (cantidad de elementos de la base que se desea utilizar)
- Realiza la construcción de la matriz de pesos para cada fotografía de la base de caras.

- Reconstruye cada fotografía como una combinación lineal más la cara promedio y se determina el error de la reconstrucción
- Muestra las fotografía original de la base de datos, la fotografía reconstruida y el error correspondiente
- Carga la base de prueba que son las fotografías que se quieren identificar y muestra dichas fotografías.
- Se le determina la caricatura para las fotografías que se desean identificar y sus correspondientes pesos.
- Se compara la fotografía que se desea identificar con cada fotografía de la base de caras, mediante la Distancia Euclidiana, almacenando la menor distancia y la posición que ocupa esta fotografía en la base de caras
- Se establece un umbral determinado y si la distancia euclidiana es menor que dicho umbral, se muestra una cuadro en donde en la parte superior aparecen la fotografía de la persona que se desea identificar con la correspondiente fotografía en la base de datos que ocupa la posición correspondiente a la menor distancia y en la parte inferior aparece una grafica de barras donde se muestran la distancia euclidiana de los pesos de la fotografía que se desea identificar con los pesos de cada una de las fotografías de la base de caras.
- En el caso que la distancia euclidiana es mayor que dicho umbral, se muestra una figura en donde en la parte superior izquierda aparecen la fotografía de la persona que se desea identificar y en la parte superior derecha aparece las palabras *no se encontro*.

Bibliografía

- [1] Turk y Pentlan, Eigenfaces for recognition. Journal of Cognitive Neuroscience. (1999)
- [2] De la contraseña a la identificación biométrica: <http://www.tecnipublicaciones.com/teleinf/actualidad/seguridad.asp?opcion=informe>
- [3] John Vacca, Biometric technologies and verification systems, Butterworth-Heinemann, 2007
- [4] Simon A Cole, Suspect Identities: A history of fingerprinting and criminal identification, First Harvard University Press paperback edition, 2002
- [5] Davide Maltoni Dario Maio y Anil K Jain, Handbook of Fingerprint Recognition, Springer 2009
- [6] Aric W Dutelle, An Introduction to Crime Science Investigation, Jones and Bartlett Publishers, 2008
- [7] Tower Paul, The fundus Oculi in monozygotic twins :Report of six pairs of identical twins, 1955
- [8] Introducción a la biometría: <http://www.monografias.com/trabajos43/biometria/biometria3.shtml>
- [9] Sistema de reconocimiento facial: <http://www.faq-mac.com/noticias/9001/sistemas-reconocimiento-facial-raul-sanchez-vitores#>
- [10] Biometría facial: <http://www.idnoticias.com/2008/08/20/biometria-facial-a-prueba-en-aeropuertos-del-reino-unido>
- [11] Nuevas Tecnologías: http://www.euskadi.net/botoelek/otros_países/sim1_c.htm

-
- [12] Antonio Lecha Marzo, El retrato hablado ó descripción verbal de la fisonomía humana, Universidad de Michigan, 2009
- [13] <http://www.colombia.com/especiales/2002/criminalistica/tecnologia.asp>
- [14] Fornberg Bengt y Herbst Beng, Modeling in Applied Mathematics, University of Colorado
- [15] Tom M Apostol, Calculus Volumen 2, Reverté, 1998
- [16] Stanley L Grossman, Algebra Lineal, McGraw Hill, 1999
- [17] David S Watkins, Fundamentals of Matrix Computations, John Wiley and Sons, 2002
- [18] WinterAlthaus y Conde Lazaro, Metodos y algoritmos básicos del álgebra numérica,Reverte.
- [19] Skiba Yuri, Métodos y esquemas numéricos un análisis computacional,Universidad Autonoma de Mexico, 2005
- [20] Rodríguez Liceso, Conde Cristina, Comparison of Novel Dimension Reduction Method in Faces Verification, Universida Rey Juan Carlos
- [21] Alkiviadis G,Gennadi Malaschonok, Applications of Singular Value Decomposition, 2002
- [22] Ilker Atalay,Thesis faces Recognition,1996.

En la anterior bibliografía se encuentran algunos de los libros, artículos y website consultados, pero es de aclarar que este tema tiene una amplia bibliografía., tan solo en el buscador de google se encuentran aproximadamente 1.510.000 website con las palabras *thesis faces recognition*