



**UNIVERSIDAD AUTÓNOMA DEL ESTADO
DE HIDALGO**

INSTITUTO DE CIENCIAS BÁSICAS E INGENIERÍA

**IDENTIFICACIÓN BIOMÉTRICA DE ROSTROS
UTILIZANDO PCA EN ESPACIOS REDUCIDOS**

T E S I S

**QUE PARA OBTENER EL TÍTULO DE
INGENIERO EN ELECTRÓNICA Y TELECOMUNICACIONES**

P R E S E N T A

PABLO HERNÁNDEZ TORRES

ASESOR: DR. JOEL SUÁREZ CANSINO

PACHUCA DE SOTO, HIDALGO. MAYO DE 2006

*El presente trabajo es
dedicado a mi familia, y en
especial, a mi madre,
Josefina Torres Osorio*

Índice general

Lista de Figuras	VII
Lista de Tablas	IX
Prefacio	XI
Resumen	XIII
Parte I	Introducción
1	
Introducción	3
1.. Planteamiento del Problema	4
2.. Metodología Empleada	7
3.. Aportes Teóricos y Tecnológicos	8
2	
Estado del Arte	11
1.. Algunas Definiciones Básicas	12
2.. Análisis Crítico del Estado del Arte	26
3.. Conclusión	32
3	
Discusión y Análisis	33
1.. De las correlaciones entre imágenes	34
2.. Media por fila o columna	45
3.. Autenticación utilizando los coeficientes de correlación	48
4.. Correlaciones por pixel	49
5.. Conclusión	54
Parte II	Desarrollo

VI	
4	
PCA en una sola imagen	59
1.. PCA por filas	59
2.. PCA por columnas	63
3.. PCA mixto	63
4.. Comparando rostros en PCA filas o columnas	63
5.. PCA clásico en otra arquitectura	65
6.. El porque de PCA por filas y columnas	69
7.. Comparación de las implementaciones PCA	69
8.. Comentarios	74
5	
Implementación de los algoritmos	77
1.. GUI	77
2.. Comparación de las implementaciones	83
6	
Conclusión final del trabajo	89
Referencias	91
Apéndice: Algoritmos Empleados	94
Apéndice: Generalización	96
Apéndice: Imágenes utilizadas	98
Apéndice: Código	99

Lista de Figuras

2.1.	Componentes principales en tres dimensiones	16
2.2.	Componentes principales en tres dimensiones	20
2.3.	Valores propios: Magnitud vs Importancia	23
2.4.	Imagen media	27
2.5.	Imágenes con imagen media substraída	28
2.6.	Nube de puntos de los datos	28
2.7.	Nube de puntos rotada	29
3.1.	Imágenes de un mismo sujeto	35
3.2.	Nube de puntos para un mismo sujeto	36
3.3.	Nube de puntos para la misma fotografía	36
3.4.	Dos sujetos diferentes	37
3.5.	Nube de puntos de dos personas distintas	37
3.6.	Rostro con rotación horizontal de aproximadamente 45°	38
3.7.	Nube de puntos con una imagen rotada 45°	38
3.8.	Diversas facetas de un rostro	39
3.9.	Diversos rostros	40
3.10.	Algunas facetas de un rostro	44
3.11.	Algunas facetas de tres rostros diferentes	45
3.12.	Conjunto de fotografías de un misma persona	51
3.13.	Nube de puntos para las variables	53
3.14.	Nuevo conjunto de fotografías	53
3.15.	Nueva nube de puntos para las variables	55
4.1.	Barrido ocular humano	70
5.1.	GUI: ventana principal	78
5.2.	GUI: listado de directorios	79
5.3.	GUI: igualación de los rostros	79
5.4.	GUI: Menús	80

VIII

B.1.	Generalización de una respuesta	96
C.1.	Imágenes utilizadas como entrenamiento	98
C.2.	Imágenes adicionales de entrenamiento	98
C.3.	Imágenes utilizadas como prueba	99

Lista de Tablas

1.1.	Base de rostros del INAOE	8
3.1.	Correlaciones para un mismo sujeto	39
3.2.	Correlación sin media substraída	41
3.3.	Correlación con media substraída	42
3.4.	Diferencia de correlaciones	43
3.5.	Correlación sin media substraída	44
3.6.	Correlaciones después de sustraer la imagen media	44
3.7.	Correlaciones después - Correlaciones antes	44
3.8.	Diferencia de correlaciones para Mary	46
3.9.	Diferencia de correlaciones para Hugo	46
3.10.	Diferencia de correlaciones para Miguel	46
3.11.	Correlación con media de otro grupo restada	46
3.12.	Correlación con media de otro grupo restada	47
3.13.	Correlación con media por filas restada	47
3.14.	Correlación con media por columnas restada	48
3.15.	Muestras de nueve variables	52
3.16.	Correlación entre nueve variables	52
3.17.	Muestras de nueve variables	54
3.18.	Correlación entre nueve variables	54
4.1.	Carga de trabajo en PCA	72
4.2.	Carga de trabajo en PCA	73
5.1.	Resultados de PRUEBA I	85
5.2.	Resultados de PRUEBA II	86

Prefacio

Muchos habrían podido llegar a la sabiduría si no se hubiesen creído ya suficientemente sabios.

—Juan Luis Vives

La identificación de personas por medio de elementos clásicos tales como, por ejemplo, el nombre, un número de identificación personal o NIP, ha llegado actualmente a ser poco confiable debido a los continuos fraudes que se cometen en actividades varias como transacciones bancarias, falsificación de títulos, compras en centros comerciales, accesos restringidos a áreas de trabajo y algunos otros casos donde es necesaria la validación de identidad, etc.

Por esta razón, se hace necesario profundizar en métodos alternativos de identificación que sean más confiables. Una opción en este sentido es dada por medio del análisis de características biológicas que distinguen a un ser humano de otro. Ejemplos de estas características hay varios, entre los que usualmente se encuentran las huellas digitales, las palmas de la mano, la voz, el iris, el rostro, el ADN, la manera de caminar, la escritura etc., entre las más comunes, otras menos comunes son la dinámica de la escritura sobre un teclado, geometría de las orejas, patrones térmicos y/o vasculares, patrones de los vasos sanguíneos de la retina, dinámica de la firma e incluso el olor natural de una persona, etc.

Cada una de estas alternativas biológicas presenta, por supuesto, sus ventajas y desventajas con respecto a las otras. Entre las ventajas, podemos señalar la

que se refiere a la sencillez en la obtención de las propiedades biológicas que definen a cada una de ellas. Es decir, no es lo mismo averiguar las características asociadas al ADN que aquellas asociadas al rostro humano. En cuanto a las desventajas, podemos señalar las formas en las que es posible engañar a tales alternativas, pero ese es un tema demasiado amplio que se sale de nuestros propósitos y por tanto es evitado en este trabajo.

Basados en la característica de sencillez, en el presente trabajo se ha optado por analizar la identificación de seres humanos por medio del rostro, ya que solamente se requiere tomar una fotografía del mismo. Nótese que nos estamos refiriendo a la facilidad con que es obtenido el rasgo humano, no a la fiabilidad del método basado en la identificación por medio del rostro. Este sería un punto a discutir, el cual se sale de los objetivos de este trabajo.

No importa qué método se emplee, cada uno de ellos se analiza por medio de herramientas matemáticas que pueden incluir, entre otras cosas, el tema de espacios vectoriales. Es decir, una imagen, sea ésta una huella digital o bien un rostro, se puede considerar como un vector en un espacio vectorial multidimensional. Como bien es sabido, estos espacios vectoriales son generados por medio de bases que son conjuntos de vectores que tienen la característica de ser linealmente independientes y, en términos de los cuales pueden escribirse todos los demás elementos del espacio vectorial.

Esta característica de los espacios vectoriales se aprovecha para construir un conjunto de imágenes que sea el conjunto base; es decir, que con él se generen todas las otras imágenes miembros del espacio vectorial de imágenes. Utilizando también las propiedades métricas del espacio vectorial así definido, es posible especificar la ‘distancia’ entre dos imágenes, hecho que sugiere la posibilidad de emplear esta técnica en el reconocimiento de rostros, tal y como se analiza en este texto.

Existen, por supuesto, técnicas matemáticas que explotan lo anterior, y dos de ellas se denominan Análisis de Componentes Principales, por brevedad PCA, y Análisis de Componentes Independientes, conocido brevemente como ICA (este es un método más general). Ambas técnicas presentan algunos problemas en casos bien específicos de reconocimiento de rostros, pero sólo los que involucran a PCA se analizan en este trabajo. Esperamos entonces que estos temas aquí tratados, arrojen un poco más de luz sobre el empleo de la técnica PCA en el reconocimiento de rostros.

Resumen

A lo largo de este trabajo se presenta un método estadístico y del álgebra lineal para el análisis de datos. Este método llamado PCA (Análisis de Componentes Principales) es también utilizado para el reconocimiento de rostros y objetos por lo que no está limitado a un tipo concreto de datos.

La característica primordial de este método es la reducción de los datos de una manera tal que la estructura o relación principal de estos no se pierde y de esa forma se maneja casi la misma información pero con menos valores ya que se han desechado los que son irrelevantes y no expresan una relación importante entre ellos.

Debido a esa característica de reducción, PCA es utilizado también en la compresión de datos (como por ejemplo imágenes) con pérdida de información; en PCA clásico, si se desean comprimir imágenes implicaría utilizar más de dos para poder llevar a cabo la compresión, en los métodos propuestos de barridos horizontales y verticales una sola imagen es necesaria para comprimirla.

En este trabajo también se expone la forma no convincente o no conveniente en la que el Análisis de Componentes Principales hace uso de la media y por lo tanto, como ya lo mencionamos anteriormente, se proponen 3 formas diferentes de realizar el análisis, en los que si bien la esencia de PCA se sigue aplicando, el uso de la media implica hechos diferentes al de PCA clásico y PCA modular.

Se divide el texto en dos partes y a grandes rasgos podemos decir que la primera de ellas contempla el análisis del método PCA y el planteamiento de un problema con respecto a el que pensamos podría mejorar su desempeño. La segunda parte se enfoca a la proposición de tres nuevas formas de realizar PCA y se expone el desarrollo de las mismas y posteriormente se llevan a cabo pruebas para verificar el desempeño de los PCA's propuestos con respecto a los PCA's existentes para verificar si hubo una mejoría en el reconocimiento o si hubo un deterioro o simplemente si el desempeño permaneció inalterado.

En forma breve el contenido de los capítulos de la parte I es el siguiente:

Capítulo 1. En este capítulo se establece el problema relacionado con el Análisis de Componentes Principales y su posible solución, mostrando para ello los pasos a seguir con el fin de aceptar o rechazar la hipótesis.

Capítulo 2. En este capítulo se desglosan la teoría e investigaciones existentes relacionadas al Análisis de Componentes Principales tomando algunas veces una actitud crítica y propositiva hacia PCA.

Capítulo 3. Aquí se da solidez al problema planteado y a la conjetura que pretende resolverlo.

Y los capítulos de la parte II:

Capítulo 4. En este apartado se desarrollan los métodos propuestos y la forma de realizar el reconocimiento de rostros con ellos.

Capítulo 5. En este capítulo se describen la interfaz gráfica de usuario para implementar los métodos y las pruebas realizadas sobre ellos para la comparación de los métodos de Análisis de Componentes Principales.

Capítulo 6. Aquí se dan las conclusiones arrojadas por la investigación.

Capítulo 1

INTRODUCCIÓN

Planteamiento del Problema y Metodología de Solución

Vivir, sufrir, morir; tres cosas que no se enseñan en nuestras universidades y que sin embargo encierran toda la sabiduría necesaria al hombre.

—Auges

Introducción

La autenticación biométrica de personas es, en la actualidad, uno de los métodos que mayormente se investigan con el objeto de validar la identificación en centros de trabajo, transacciones, pesquisas forenses, etc. Una de las técnicas que más se ha investigado en este sentido es la que involucra a la autenticación de rostros humanos.

En este caso se han empleado algoritmos tales como Análisis de Componentes Principales (PCA) y Análisis de Componentes Independientes (ICA), métodos que se encuentran relacionados. Sin embargo, se ha encontrado que en algunos casos bien localizados dichas técnicas, y en especial PCA, tienen problemas al momento en que se lleva a cabo la autenticación.

Parece una razón obvia, que dicho problema se encuentre directamente relacionado con los cambios de los rostros del conjunto de imágenes de prueba con respecto al conjunto de imágenes de entrenamiento, pero también—y parece ser la razón más importante y es la que se estudia aquí—esto se debe a la interferencia que de forma intrínseca hace la técnica PCA en las correlaciones (al menos que se indique lo contrario en este trabajo nos referimos a corre-

lación lineal simplemente como correlación) de las imágenes del conjunto de entrenamiento.

Este capítulo se avoca entonces a plantear el problema en el área de autenticación biométrica de rostros que nos interesa analizar, así como el estado del arte y la metodología seguida para solucionar el problema ya indicado en forma somera.

1. PLANTEAMIENTO DEL PROBLEMA

El tamaño del conjunto de entrenamiento empleado en la autenticación biométrica por medio de rostros humanos consiste esencialmente de solamente una imagen por sujeto. Por lo cual, tal cardinalidad tiene la ventaja de presentar menor carga de trabajo para el sistema biométrico, pero indudablemente, tal sistema deberá ser capaz de acertar ante las diferencias entre las imágenes del sujeto del conjunto de prueba y las imágenes del conjunto de entrenamiento—lo cual es una característica deseable en cualquier sistema biométrico de autenticación e identificación—, es obvio que será poco frecuente que la imagen utilizada en el proceso de generalización sea exactamente igual a la imagen de entrenamiento, pero unos cambios serán más drásticos que otros y estos traerán problemas al momento de autenticar a una persona.

Los cambios posibles a partir de un rostro de expresión neutra sin rotaciones ni traslaciones son los siguientes:

1. La persona utiliza gafas oscuras o transparentes.
2. La persona cierra los ojos.
3. La persona sonrío.
4. La persona cambia de peinado o corte.
5. La persona se deja crecer o se corta el bello facial.
6. La persona se cubre con maquillaje.
7. Rotaciones y traslaciones del rostro de la persona.
8. Cambios de expresión por ira, asombro, etc.

De tales cambios, el más severo y que representa menos aciertos a la hora de autenticar es el no. 7 y, aunque en menor medida, el no. 1; el maquillaje

sólo afecta cuando introduce sombras o luces muy extremas en el área aplicada. Como ya se ha mencionado, ICA y PCA clásicos tienen la importante característica de que en ellas, y para cada persona, se emplea solamente una imagen del rostro correspondiente en el proceso que determinará el espacio de imágenes sobre el cual se hará finalmente la autenticación.

Se propone entonces tomar no solamente una imagen como normalmente se hace, sino una mayor cantidad de ellas con el objeto de eliminar la dificultad antes mencionada. Sin embargo para evitar la carga de trabajo que ello implica, se optó por otra forma de realizar PCA de tal manera que no interfiriera unas imágenes con otras afectando sus correlaciones, y así no terminar construyendo el espacio de imágenes del mismo modo a como se procede clásicamente.

Como se verá profundizado más adelante, la imagen media que se resta al conjunto de imágenes de entrenamiento y a las mismas imágenes de prueba, es la que introduce cambios en las correlaciones naturales existentes entre las imágenes. Sin embargo existe una variante de PCA—PCA modificado o PCA por módulos—que toma más de una imagen por persona con el objetivo de construir un espacio para cada individuo con sus propias imágenes, tal forma de hacer PCA no afecta de forma contraproducente las correlaciones debido a que la imagen media es de un conjunto de imágenes de la misma persona.

Esto, por supuesto, nos condujo a la pregunta acerca de la forma en la que deberíamos llevar a cabo de ahora en adelante la técnica PCA para definir el nuevo o los nuevos espacios sobre los que se haría la autenticación y claro está que la respuesta a esta pregunta fue que se debería hacer PCA sin afectar la correlación ya existente entre una imagen de rostro y otra, como con el uso de la imagen media que afecta el reconocimiento, ya sea aumentando la falsa aceptación o el falso rechazo de aquel sistema que llegase a hacer de PCA su núcleo importante.

Por tanto pensamos que, llevar a cabo PCA de manera que no se tenga que restar la imagen media del conjunto de entrenamiento (cuando este contiene imágenes de diferentes personas) mejoraría los aciertos de este método, pero además que no se ocupara más de una imagen por individuo (como ocurre en PCA por módulos). Esa forma, creemos, consiste en utilizar, en una especie de barrido horizontal o vertical, las filas o las columnas (o ambas al mismo tiempo) de la imagen única del sujeto para crear un espacio propio para esa persona, de esta forma no intervienen otras imágenes fuese cual fuese la forma de restar la media. Al igual que en PCA por módulos, en esta nueva proposición, se calcula un espacio por cada individuo, pero la diferencia en cuanto a carga de trabajo se hace evidente.

En PCA clásico, los píxeles toman la función de variables aleatorias, mientras que las imágenes forman las observaciones, de tal forma que al restar la media a cada variable—para después proseguir con los restantes pasos de PCA—estamos restando la imagen media e interfiriendo con los valores de correlación entre las imágenes, por lo cual, como se ha dicho anteriormente, creemos que se afecta el reconocimiento, es por ello que, otra propuesta consiste en tomar a cada imagen como la variable aleatoria y a sus píxeles como las observaciones o las muestras de esa variable, de esta forma al restar la media a cada variable, estaríamos restando la media propia de cada imagen en el conjunto de entrenamiento y en el conjunto de prueba sin interferir las correlaciones posibles entre imágenes.

En este trabajo consideramos a nuestras imágenes o a sus píxeles como variables aleatorias, ello por la arbitrariedad en los valores de intensidad de los píxeles de las imágenes y de las imágenes mismas, ya que no podemos saber con anticipación las imágenes que se le presentarán a PCA; lo que no nos interesa o lo que no es importante en PCA es la función de densidad de probabilidad que tienen las variables aleatorias, como en el caso del método ICA (Independent Component Analysis), un método relacionado y más general que PCA, donde sí tiene importancia la densidad de probabilidad de los datos.

Llegamos así al punto en que podemos plantear nuestros objetivos generales y específicos como sigue:

Objetivo General

Proporcionar una metodología que aumente la confiabilidad en el proceso de autenticación biométrica por medio de rostros humanos cuando se emplean técnicas basadas en análisis de componentes principales.

Objetivos Específicos

1. Elaborar una base de datos de imágenes de rostros o buscar alguna(s) en medios cibernéticos y elegir la más adecuada para los propósitos de nuestro trabajo.
2. Determinar uno o más métodos para realizar PCA sin utilizar la imagen media del conjunto de entrenamiento.
3. Elaborar sistema de cómputo en MATLAB con el cual se pueda experimentar con la base de datos de imágenes creada y con imágenes de rostros de prueba, y cuyos algoritmos deben basarse en los métodos propuestos en el objetivo específico 2.

2. METODOLOGÍA EMPLEADA

Resumiendo, nuestra hipótesis dice, que al restar la imagen media de un conjunto de imágenes de rostros—cuando estos sean de diferentes sujetos, como sucede con el conjunto entrenamiento de PCA clásico—, mermaría el desempeño del algoritmo PCA ocasionando que disminuya su confiabilidad, esto debido al cambio de las correlaciones posibles existentes entre una imagen y otra, que como se verá es más alta mientras más parecidas sean las imágenes.

Entonces para probar nuestra hipótesis debemos construir uno o más PCA's que no afecten las correlaciones que pudiesen encontrarse entre las imágenes que deseamos autenticar, esto es, PCA's que no interfieran una imagen con otra, como por ejemplo, a través de la imagen media. Posteriormente con tales PCA's encontrados llevar a cabo las pruebas pertinentes cuyos resultados nos revelarán la exactitud de la hipótesis.

Por tal motivo se escribieron, primeramente, los algoritmos de PCA clásico y PCA por módulos originales, para de esta forma verificar la solvencia de este método en sus dos formas, posteriormente, se escribieron los algoritmos nuevos propuestos, para de igual forma analizar su solvencia y registrar los resultados. Semejante información nos será de primordial importancia para vislumbrar las mejoras o inconvenientes de las propuestas de PCA. Finalmente, resta mencionar que todas las funciones que soportan al sistema de cómputo se escribieron en MATLAB 6.5 en forma modular para que resultara más práctico manejarlas, modificarlas, corregirlas, escalarlas, etc.; y se corrieron sobre una PC AuthenticAMD, con procesador AMD-K6(tm)-III, con 56.0 MB de memoria RAM y sobre una PC TEXA MATCH 4 Intel(R) Pentium(R) 4, con una CPU 2.40 GHz y con 256 MB en RAM; además, para probar los programas empleados en el sistema se utilizó la base de rostros del INAOE (ver referencia [1]).¹ Cada rostro tiene cientos de imágenes diferentes (ver tabla 1.1). Dentro de esas imágenes que se mencionan, las hay con diferentes expresiones y con ángulos de rotación de -45° y 45° con respecto a un eje que pasa por la parte superior de la cabeza y por el cuello; de las gafas, las más relevantes son las transparentes y las oscuras, otras tienen el armazón en color negro pero los vidrios son transparentes, etc.

¹La base de datos consta de 18 sujetos—17 en verdad, uno de ellos (Hugo) se repite pero con el cabello y más corto y sin barba y se trata como un sujeto diferente (Hugo.)—

Tabla 1.1 Base de rostros del INAOE

Persona	1 Mary	2 Arturo	3 Avendaño	4 Carlos	5 David	6 Edith	7 Trilce	8 Eugenio	9 Gerardo
gafas 1	18	17	18	18	17	18	18	18	18
gafas 2	18	18	18	18	18	18	18	18	18
gafas 3	17	18	18	18	18	18	18	18	18
gafas 4	18	18	18	18	17	18	18	18	18
gafas 5	17	18	18	18	18	18	18	18	18
gafas 6	18	18	18	18	18	18	12	18	18
gafas 7	18	18	18	18	18	18	18	18	18
gafas de sol	12	12	12	12	12	12	12	12	12
sin gafas	18	18	18	18	18	17	18	18	18
gafas extras	0	0	0	0	0	0	0	18	18
imágenes p/ persona	154	155	156	156	154	155	150	174	174
Persona	10 Héctor	11 Hugo	12 Hugo.1	13 Luis	14 Magoly	15 Miguel	16 Nicandi	17 Tomás	18 Ulises
gafas 1	18	18	18	18	18	18	18	16	18
gafas 2	18	0	18	18	18	18	18	17	18
gafas 3	18	18	18	18	18	18	18	18	1
gafas 4	18	18	18	18	18	18	18	18	18
gafas 5	18	18	18	18	18	18	18	18	18
gafas 6	0	0	0	18	0	18	18	18	0
gafas 7	18	18	18	18	18	18	17	18	16
gafas de sol	12	12	12	12	12	12	12	12	12
sin gafas	18	18	18	18	18	18	18	18	18
gafas extras	18	0	0	0	0	0	18	18	0
Imágenes p/ persona	156	120	138	156	138	156	173	171	136

Total 2772

^a Las "gafas extras" son las gafas que normalmente usa la persona.

3. APORTES TEÓRICOS Y TECNOLÓGICOS

El trabajo tiene aportes teóricos acerca de la forma en que influye en la calidad de la autenticación el que se considere o no la imagen media de distintos individuos en PCA. Mencionamos ya, que PCA por módulos, no la considera, pero tiene la desventaja del tiempo consumido en el entrenamiento y en la identificación cuando la base de datos es enorme, por lo que un PCA sin esta característica que utiliza más de una imagen por sujeto y que aumentase la fiabilidad de la autenticación sería por lo más necesaria.

Desde el punto de vista tecnológico, se aporta un sistema computacional que tiene la característica de contar con una interfaz gráfica, por medio de la cual el usuario puede ajustar diversos parámetros con el objeto de analizar el

comportamiento de algoritmos basados en PCA, y visualizar la influencia que sobre ellos tiene el que se considere o no la utilización de la imagen media.

Capítulo 2

ESTADO DEL ARTE

La principal sabiduría no es el profundo conocimiento de las cosas remotas, desusadas, oscuras y sutiles, sino el de aquellas que en la vida cotidiana están ante nuestros ojos.

—Milton

Introducción

En este capítulo se analizan algunas de las investigaciones previas sobre el tema realizadas por otros investigadores. La actitud que tomaremos será crítica y propositiva, enfocándonos exclusivamente sobre los resultados que se relacionan con el problema ya planteado en secciones anteriores.

Para comprender mejor los conceptos involucrados, es conveniente primero hacer una relativamente breve digresión acerca de algunos tópicos como, por ejemplo, la interpretación vectorial que se le da a una imagen, así como el concepto mismo de ésta. El tipo de imágenes en que se está interesado es precisamente el de una discreta.

Como se sabe, una imagen discreta se encuentra definida por una serie de puntos de diferente intensidad, las que pueden ser dadas en color o tonalidades de grises. Si la imagen es bidimensional, entonces estas intensidades se encuentran localizadas en puntos de un plano cartesiano. Si (x, y) son las coordenadas de un punto \boldsymbol{x} sobre este plano, entonces la intensidad en el mismo se denota por $I(\boldsymbol{x})$; es decir, es una función en dos variables.

Lo anterior sugiere entonces que una imagen discreta se puede representar por medio de una matriz \mathcal{I} cuyos índices (i, j) tienen una correspondencia con el vector \boldsymbol{x} , y cuyas entradas \mathcal{I}_{ij} denotan a la intensidad $I(\boldsymbol{x})$ en los

índices correspondientes. El tamaño de la matriz \mathcal{I} se denomina resolución de la imagen, y la calidad de esta última depende, por supuesto, del valor de la resolución. Entre mayor resolución se tenga, entonces mayor es la calidad. Con estas definiciones introductorias básicas es posible ahora profundizar sobre el asunto que nos interesa.

1. ALGUNAS DEFINICIONES BÁSICAS

Una imagen dada por la matriz \mathcal{I} , cuya resolución sea $m \times n$, puede ser transformada a un vector en el espacio Euclidiano mn -dimensional. Esta transformación se puede obtener tomando las primeras n entradas del vector a construir como aquellas de la primera fila de \mathcal{I} , las segundas n entradas como aquellas de la segunda fila, y así sucesivamente hasta las m -ésimas n entradas como aquellas de la fila m de \mathcal{I} .

Si \mathbf{X} es el vector así definido, entonces es claro que $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$, donde \mathbf{x}_i es la fila i de la matriz \mathcal{I} . Si, en general, se tienen k imágenes $\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_k$, todas ellas con la misma resolución $m \times n$, entonces se puede formar una matriz de imágenes $\mathcal{P} = [\mathbf{X}_1^T, \mathbf{X}_2^T, \dots, \mathbf{X}_k^T]$ de tamaño $mn \times k$, donde \mathbf{X}_i^T es el vector transpuesto de \mathbf{X}_i .

Por supuesto, podemos proceder de una manera diferente considerando las columnas de la matriz de imagen \mathcal{I} , en lugar de las filas de ella. En este caso, diremos que \mathbf{y}_j^T es la columna j de la matriz \mathcal{I} , y que el vector fila $\mathbf{Y} = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n)$, es un elemento del espacio Euclidiano mn -dimensional. Para el conjunto de k imágenes dado arriba, se puede formar también una matriz \mathcal{Q} de la siguiente forma $\mathcal{Q} = [\mathbf{Y}_1^T, \mathbf{Y}_2^T, \dots, \mathbf{Y}_k^T]$, la cual tiene tamaño $nm \times k$.

Un punto importante a notar aquí es que, tanto en este caso como en el previo, se puede hablar entonces de un espacio vectorial Euclidiano cuyos elementos son precisamente imágenes transformadas a vectores. Cada componente de un vector determina la intensidad en alguna entrada de la matriz que representa a la imagen.

Se han descrito entonces dos transformaciones $\mathfrak{I}_{\mathcal{P}} : \mathbb{M} \rightarrow \mathbb{R}^{mn}$ y $\mathfrak{I}_{\mathcal{Q}} : \mathbb{M} \rightarrow \mathbb{R}^{mn}$, donde \mathbb{M} es el espacio de las matrices de tamaño $m \times n$ y \mathbb{R}^{mn} es el espacio Euclidiano mn -dimensional. Estas transformaciones se definen explícitamente como sigue:

$$\mathfrak{I}_{\mathcal{P}}(i, j) = (i - 1)n + j \quad (2.1)$$

y

$$\mathfrak{T}_{\mathcal{Q}}(i, j) = (j - 1)m + i \quad (2.2)$$

donde, en ambos casos, $i = 1, 2, \dots, m$ y $j = 1, 2, \dots, n$. Por supuesto, ambas transformaciones cuentan con sus inversas definidas como sigue:

$$\mathfrak{T}_{\mathcal{P}}^{-1}(i) = ((i - 1)div(n) + 1, (i - 1)mod(n) + 1) \quad (2.3)$$

y

$$\mathfrak{T}_{\mathcal{Q}}^{-1}(i) = ((i - 1)mod(m) + 1, (i - 1)div(m) + 1) \quad (2.4)$$

donde, en ambos casos, $i = 1, 2, \dots, mn$. De esta forma, tanto la ida como el regreso de un espacio a otro se encuentran asegurados, inclusive de manera única; es decir, las transformaciones son biyectivas.

1.1 ANÁLISIS DE COMPONENTES PRINCIPALES

Desde un punto de vista de muestreo estadístico, el conjunto de k imágenes se puede considerar como un muestreo multivariable, en el cual se analiza un total de mn variables, independientemente de que se hallan tomado las filas o columnas para definir el espacio Euclidiano \mathbb{R}^{mn} . De esta forma, es válido cuestionarse acerca de la correlación existente entre cada una de las variables que definen a \mathbb{R}^{mn} , tomando como base los puntos ‘experimentales’.

Es importante resaltar en este punto el por qué podría interesar verificar la correlación existente entre las variables que definen a \mathbb{R}^{mn} . Primeramente, recordemos que, de acuerdo a las transformaciones $\mathfrak{T}_{\mathcal{P}}$ o $\mathfrak{T}_{\mathcal{Q}}$, cada variable representa un posible valor de intensidad de un pixel en una imagen.

La cuestión importante a responder es si es necesario mantener todas las variables así definidas para manejar la información contenida en las imágenes. En otras palabras, qué tantas variables habría que emplear sin que haya una pérdida sustancial de la información en las imágenes que definen a las mismas. La respuesta a esta pregunta radica en la disminución del número de variables basados en la correlación existente entre algunas de ellas. Lograr una cosa así, implicaría no tener que trabajar sobre todo el espacio \mathbb{R}^{mn} , sino sobre uno de menor dimensionalidad que aquel.

Una reducción en la dimensionalidad implica además el descifrar de una forma más simple la estructura de los datos que definen la muestra que, en este caso, se encuentra definida por las k imágenes. Es bien sabido que una forma de determinar la correlación entre dos o más variables es a través del concepto de matriz de covarianza, la cual se define en términos de varianzas y covarianzas de las variables involucradas.

Asumiendo entonces que $\mathbf{X} = (X_1, X_2, \dots, X_i, \dots, X_j, \dots, X_{mn})$ es la variable aleatoria cuya muestra es el conjunto de vectores en \mathbb{R}^{mn} , construídos por medio de las k imágenes y de acuerdo a la transformación $\mathfrak{T}_{\mathcal{P}}$ dada por ecuación 2.1, se define la covarianza entre las variables aleatorias X_i y X_j por medio de la expresión

$$Cov(X_i, X_j) = \frac{1}{k-1} (\mathfrak{X}_i - \boldsymbol{\mu}_i)(\mathfrak{X}_j - \boldsymbol{\mu}_j)^T \quad (2.5)$$

donde \mathfrak{X}_i es el vector que contiene los k valores de la muestra de la variable aleatoria X_i (a menos que se diga lo contrario, las muestras de cada variable se consideran en filas en todo este documento), valores que son dados por las intensidades asociadas a una entrada de las imágenes, y entrada que es calculada de acuerdo a la expresión 2.4. Además, $\boldsymbol{\mu}_i$ es el vector cuyas entradas son todas iguales al valor medio de los valores en \mathfrak{X}_i . Este valor medio es un escalar y se calcula de acuerdo a la expresión siguiente:

$$\mu_i = \frac{1}{k} \mathfrak{X}_i \mathbf{1}^T \quad (2.6)$$

donde $\mathbf{1}$ es un vector de tamaño $1 \times k$ cuyas entradas son todas igual a la unidad.

Para un valor i fijo, existen mn covarianzas, por lo que, tomando en cuenta cada i como fila de una matriz, se puede armar una matriz de orden mn la cual se denomina matriz de covarianza. Notando que $Cov(X_i, X_j) = Cov(X_j, X_i)$ se concluye que dicha matriz es simétrica. Adicionalmente, puesto que existen mn variables aleatorias, entonces habrá un total de mn vectores \mathfrak{X} e igual número de vectores valores medio $\boldsymbol{\mu}$, cada uno de ellos con tamaño $1 \times k$.

Sabiendo lo anterior, no es difícil ver que la matriz de imágenes \mathcal{P} se puede escribir como sigue:

$$\mathcal{P} = [\mathfrak{X}_1^T, \mathfrak{X}_2^T, \dots, \mathfrak{X}_{mn}^T]^T \quad (2.7)$$

Análogamente, se puede definir una matriz Υ en términos de los mn vectores valor medio como sigue:

$$\Upsilon = [\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T, \dots, \boldsymbol{\mu}_{mn}^T]^T \quad (2.8)$$

Bajo estas condiciones, no es difícil demostrar también que la matriz de covarianza Φ es dada entonces por la expresión

$$\Phi = (\mathcal{P} - \Upsilon)(\mathcal{P} - \Upsilon)^T \quad (2.9)$$

la cual es una matriz simétrica de orden mn . Las entradas de Φ dan las correlaciones existentes entre cada una de las variables. Si el valor de una de ellas es nulo, entonces se concluye que no existe correlación entre las variables involucradas.

En álgebra lineal se sabe que dependiendo de las bases del espacio vectorial con que se trabajen, y si la matriz de covarianza Φ se considera como una transformación lineal, entonces es posible determinar una base nueva en la que la matriz de esta transformación sea una matriz diagonal; es decir, que las correlaciones entre las variables distintas en la nueva base sean nulas. Esto elimina las correlaciones, pero todavía no reduce la dimensionalidad. Para ello, se requiere introducir un criterio adicional, como se verá más adelante.

Por supuesto, se procederá de una manera análoga cuando se emplee la transformación 2.2, en lugar de la 2.1. En este caso, la variable aleatoria en \mathbb{R}^{mn} será ahora $\mathbf{Y} = (Y_1, Y_2, \dots, Y_i, \dots, Y_j, \dots, Y_{mn})$, y la matriz \mathcal{Q} estará dada por la siguiente expresión

$$\mathcal{Q} = [\mathfrak{Y}_1^T, \mathfrak{Y}_2^T, \dots, \mathfrak{Y}_{mn}^T]^T \quad (2.10)$$

siendo los mn vectores \mathfrak{Y}_i de tamaño $1 \times k$, con entradas dadas por las intensidades de pixel de cada una de las k imágenes. De esta forma, se llega a obtener una matriz de covarianza Ψ muy similar a la dada por la expresión 2.9. A menos que se indique lo contrario, de aquí en adelante solamente se empleará la matriz de covarianza Φ , la cual se construye tomando en consideración las transformaciones 2.1 y 2.3.

Ejemplo 1

Considérese el caso simple en el que se tienen 201 ‘imágenes’, cada una con ‘resolución’ 1×3 . Las 201 ‘imágenes’ surgen al aplicar el algoritmo 2.1

localizado en el Apéndice Algoritmos Empleados. Los datos así obtenidos son aleatorios con distribución uniforme, los cuales se buscó, ex profeso, ubicar a lo largo de una recta en tres dimensiones definida por

$$z = \lambda(0,5641, 0,5895, 0,5782), \lambda \in \mathbb{R}$$

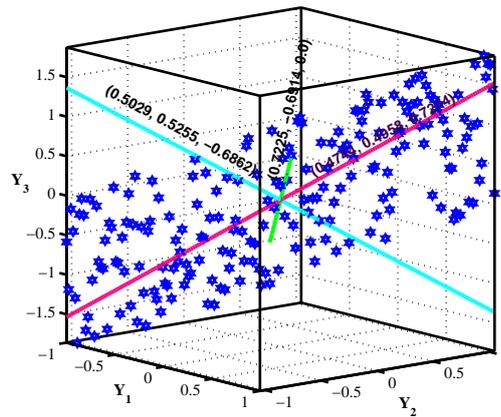


Figura 2.1 Componentes principales en tres dimensiones

Para la simulación realizada, la matriz de covarianza en dicha base es diagonal, indicando con ello que no existen correlaciones entre las variables identificadas por los nuevos vectores base encontrados

$$\beta = \{(0,7225, -0,6914, 0,0), (0,5029, 0,5255, -0,6862), (0,4744, 0,4958, 0,7274)\}$$

y que la Figura 2.1 muestra, correspondiendo al primer vector de la base encontrada un valor propio nulo. Por supuesto, estos vectores base cambian en corridas diferentes del algoritmo, debido a que en cada momento se generan datos aleatorios diferentes a las corridas previas. \square

1.2 ANÁLISIS DE IMAGEN CON PCA

Una vez descubierta la base en la que las correlaciones entre las variables son nulas, el siguiente paso consiste en definir la relación de esta nueva base con las imágenes originales a partir de las cuales se derivó. Este es un problema bastante conocido en álgebra lineal. Así, se asume que β y β' representan, respectivamente, las bases en las que se encuentran las variables aleatorias correlacionadas y las no correlacionadas.

Puesto que se asume una reducción en la dimensionalidad del espacio en que se encuentran las variables iniciales, el número de vectores en la base β es igual que el número de vectores en la base β' , aunque los elementos de esta última base son elementos con dimensión menor a la del espacio generado por la primera.

\mathbb{R}^{mn} es el espacio generado por la base β , se conviene en denotar por \mathbb{V} el espacio generado por la base β' , el cual se encuentra inmerso en el primero. La idea consiste en construir una transformación $\mathfrak{T} : \mathbb{R}^{mn} \rightarrow \mathbb{V}$ que permita pasar las imágenes a la nueva base y hacer el reconocimiento ahí. Primeramente, se debe observar que la base β de \mathbb{R}^{mn} es la base canónica, y que la base de \mathbb{V} se obtiene simple y sencillamente después de eliminar los vectores propios asociados a la matriz de covarianza, Φ , cuyos valores propios hayan sido también eliminados—con el fin de reducir la dimensionalidad—de acuerdo a ciertos criterios que se analizarán más adelante.

Es un resultado conocido en álgebra lineal que las columnas de la matriz de cambio de base (y de espacio), $\mathfrak{P} : \mathbb{V} \rightarrow \mathbb{R}^{mn}$, están dada por las coordenadas de cada uno de los elementos de la base ordenada β' del espacio \mathbb{V} , con respecto a la base β de \mathbb{R}^{mn} [2]. Puesto que esta última base es la canónica, entonces las columnas de la matriz \mathfrak{P} son simplemente los vectores propios de la matriz de covarianza, Φ , que no hayan sido eliminados. Asumiendo que la dimensionalidad de \mathbb{V} es k , entonces el tamaño de la matriz \mathfrak{P} es simplemente $mn \times k$.

Así, si $[\vec{\beta}]_{\beta'}$ son las coordenadas en columna de un vector de \mathbb{V} en base β' , entonces las coordenadas $[\vec{\beta}]_{\beta}$ de ese mismo vector en la base canónica de \mathbb{R}^{mn} son calculadas de la siguiente forma

$$[\vec{\beta}]_{\beta} = \mathfrak{P} [\vec{\beta}]_{\beta'} \quad (2.11)$$

donde $[\vec{\beta}]_{\beta'}$ es de tamaño $k \times 1$ y $[\vec{\beta}]_{\beta}$ de tamaño $mn \times 1$. Por supuesto, si se desea trabajar con coordenadas en fila, entonces la expresión a emplear es sencillamente la siguiente

$$(\vec{\beta})_{\beta} = (\vec{\beta})_{\beta'} \mathfrak{P}^T. \quad (2.12)$$

donde ahora $(\vec{\beta})_{\beta'}$ es de tamaño $1 \times k$ y $(\vec{\beta})_{\beta}$ de tamaño $1 \times mn$. Obsérvese que multiplicando la expresión 2.11 por la matriz \mathfrak{P}^T por la izquierda de ambos

miembros, o multiplicando la expresión 2.12 por la matriz \mathfrak{P} por la derecha de ambos miembros, entonces se obtienen, respectivamente, las expresiones

$$\mathfrak{P}^T [\vec{\beta}]_{\beta} = \mathfrak{P}^T \mathfrak{P} [\vec{\beta}]_{\beta'} \quad (2.13)$$

y

$$(\vec{\beta})_{\beta} \mathfrak{P} = (\vec{\beta})_{\beta'} \mathfrak{P}^T \mathfrak{P}. \quad (2.14)$$

Si en las ecuaciones 2.13 y 2.14 es posible que exista, respectivamente, la inversa $(\mathfrak{P}^T \mathfrak{P})^{-1}$, entonces es directo determinar la transformación $\mathfrak{Q} : \mathbb{R}^{mn} \rightarrow \mathbb{V}$ como el producto $(\mathfrak{P}^T \mathfrak{P})^{-1} \mathfrak{P}^T$ o $\mathfrak{P}(\mathfrak{P}^T \mathfrak{P})^{-1}$, por lo que en correspondencia se tendrá lo siguiente

$$(\mathfrak{P}^T \mathfrak{P})^{-1} \mathfrak{P}^T [\vec{\beta}]_{\beta} = [\vec{\beta}]_{\beta'} \quad (2.15)$$

y

$$(\vec{\beta})_{\beta} \mathfrak{P}(\mathfrak{P}^T \mathfrak{P})^{-1} = (\vec{\beta})_{\beta'} \quad (2.16)$$

Un punto importante a resaltar es el que se refiere a los vectores propios que generan el espacio \mathbb{V} . Se puede demostrar que ellos son linealmente independientes y, de hecho, ortogonales [3]. Esto significa que el producto escalar entre ellos es igual a cero. Sabiendo que la matriz \mathfrak{P} se construye colocando como vectores columna a los vectores propios elementos de la base β' , se concluye entonces que la matriz $\mathfrak{P}^T \mathfrak{P}$ es invertible por ser diagonal, y que la inversa, $(\mathfrak{P}^T \mathfrak{P})^{-1}$, no es otra cosa que una matriz diagonal cuyos elementos de la diagonal son el recíproco del correspondiente elemento diagonal de $\mathfrak{P}^T \mathfrak{P}$. De hecho, puesto que los vectores propios son también normales, entonces esta matriz inversa no es otra cosa que la identidad. Por lo tanto, con expresiones 2.11 y 2.15 se pueden determinar las transformaciones deseadas; a saber

$$[\vec{\beta}]_{\beta} = \mathfrak{P} [\vec{\beta}]_{\beta'} \quad (2.17)$$

y

$$[\vec{\beta}]_{\beta'} = \mathfrak{P}^T [\vec{\beta}]_{\beta} \quad (2.18)$$

Las expresiones 2.17 y 2.18 definen una arquitectura de red neuronal denominada Memoria Heteroasociativa Bidireccional, y ello da pie a relacionar el problema de reconocimiento de rostros con el tema de redes neuronales, lo cual resulta interesante abordar en un tema futuro. Sin embargo, hasta ahora solamente ha quedado claro el procedimiento a seguir para transformar las imágenes del espacio \mathbb{R} hacia el espacio \mathbb{V} y no se ha explicado qué es lo que se hace, en el proceso de autenticación, con las imágenes en \mathbb{V} .

Un punto que debe quedar claro antes de continuar, es el que se refiere a la interpretación del proceso de transformación de un espacio a otro, como el que aquí se realiza. En otras palabras, ¿Qué se ha ganado con ello?. Ya se ha visto que al menos se obtienen dos beneficios, a saber,

1. Eliminación de correlaciones,
2. Encuentro de las direcciones de máxima y mínima varianza de las variables, y
3. Reducción de dimensionalidad.

El punto no. 1 se encuentra muy relacionado con el de reducción de dimensionalidad, es el que se refiere a la obtención de una mejor definición del patrón de comportamiento de las variables en el nuevo espacio, esto debido a que se han eliminado variables que contribuyen de una manera insignificante a la especificación de dicho patrón. Las variables que se han conservado son simplemente las componentes asociadas a cada vector de la base β' la que, como ya se sabe, está definida por los vectores propios de la matriz de covarianza Φ que no han sido removidos.

Para comprender esto, recuérdese que cada variable aleatoria representa a uno de los mn píxeles de que se compone una imagen con resolución $m \times n$. Al eliminar dimensiones en el análisis de componentes principales lo que se está haciendo entonces es eliminar diferencias en las imágenes, diferencias que no contribuyen al patrón de comportamiento observado en términos de los vectores propios no removidos. De esta forma, lo que se conserva son solamente similitudes entre las imágenes, similitudes definidas precisamente por los vectores propios no removidos, donde cada uno proporciona una característica de similitud [4].

Esta es la razón principal por la cual el reconocimiento de una imagen se hace en el espacio generado por la base β' y no en el espacio original \mathbb{R}^{mn} . Por consiguiente, dada una imagen a ser autenticada, se transforma ésta al espacio \mathbb{V} , y sus correspondientes coordenadas se emplean entonces para detectar el

conjunto de vectores de la base β' que tengan características similares a la imagen dada.

Ejemplo 2

Considérese nuevamente el caso simple en el que se tienen 201 'imágenes', cada una con 'resolución' 1×3 . Las 201 'imágenes' surgen al aplicar el algoritmo 2.1 localizado en el Apéndice Algoritmos Empleados. Los datos así obtenidos son aleatorios con distribución uniforme, los cuales se buscó, ex profeso, ubicar a lo largo de una recta en tres dimensiones definida por

$$z = \lambda(0,5641, 0,5895, 0,5782), \lambda \in \mathbb{R}$$

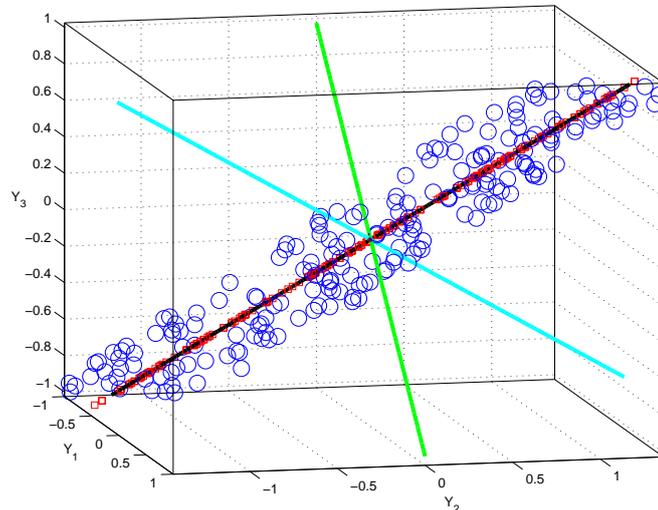


Figura 2.2 Componentes principales en tres dimensiones

Para la simulación realizada, la matriz de covarianza en dicha base es diagonal, indicando con ello que no existen correlaciones entre las variables identificadas por los nuevos vectores base encontrados

$$\beta' = \{(0,7158, 0,0000, -0,6983), (0,4413, -0,7750, 0,4524), (0,5412, 0,6320, 0,5547)\}$$

y que la Figura 2.2 muestra, correspondiendo al primer vector de la base encontrada un valor propio nulo. Por supuesto, estos vectores base cambian en corridas diferentes del algoritmo, debido a que en cada momento se generan datos aleatorios diferentes a las corridas previas.

Para ilustrar la forma en que el análisis de componentes principales categoriza en base a similitudes, para la nueva base se selecciona solamente el

tercer vector, al cual le corresponde el mayor de los valores propios (1.1203, el segundo vector propio tiene asociado un valor propio igual a 0.0525). Así, la matriz \mathfrak{B} resulta ser igual a la transpuesta del vector (0.5412, 0.6320, 0.5547). Al aplicar la transpuesta de esta matriz al conjunto de puntos que la Figura 2.2 muestra, y sobre los puntos así obtenidos se aplica ahora la matriz \mathfrak{B} , se obtienen entonces los puntos que se encuentran a lo largo de la línea definida por el único vector escogido de la base β' . La Figura 2.2 muestra también esta situación. Nótese la forma en que los puntos se alinean exactamente sobre la recta seleccionada, denotando con esto una característica de similitud entre ellos. \square

Un comentario final es el que se refiere al criterio seguido para seleccionar los vectores de la base β' . Como lo muestra el Ejemplo 2, este criterio se basa en la magnitud de los eigenvalores de la matriz de covarianza Φ ; a saber, se toman los eigenvalores con magnitud más grande. Cuántos eigenvalores se toman, depende del grado de confiabilidad deseada en el proceso de autenticación.

1.3 ANÁLISIS DE IMAGEN CON PCA POR MÓDULOS O MODIFICADO

La diferencia entre este PCA y el PCA normal es que las imágenes de rostros son del mismo sujeto, esto quiere decir que no realizamos PCA sobre un conjunto de rostros diferentes sino sobre un mismo rostro, lo que significa que para cada rostro que formará parte de la base de datos del sistema biométrico se realizará un PCA sobre sus q fotografías; si hay k rostros se realizan k PCA's clásicos de forma individual sobre cada rostro.

Todas las imágenes de los rostros pertenecen al mismo espacio \mathbb{R}^{mn} , pero sólo las pertenecientes a un mismo rostro serán ocupadas para formar su propio espacio y su base, de este modo todos los rostros que tengan $q > 1$ formarán su propio espacio. De esta manera habrá k nuevos espacios $\mathbb{V}_{1:k}$ y sus correspondientes bases $\beta_{1:k}$; estos nuevos espacios son de menor dimensionalidad. De esta forma quedan definidas k transformaciones $\mathfrak{T}_i : \mathbb{R}^{mn} \rightarrow \mathbb{V}_i$, donde i corre desde 1 hasta k . También se definen Φ_k matrices de covarianzas, de las que surgirán los vectores de las nuevas bases.

Visto así, nuestras bases \mathfrak{B}_i serán k , que como ya vimos, determinan el cambio de base $\mathfrak{B}_i : \mathbb{V}_i \rightarrow \mathbb{R}^{mn}$, es decir, desde la base nueva hasta la base canónica, que es una matriz de cambio de base más fácil de construir que aquella que hace el cambio de la base canónica al espacio nuevo. Pero debido a las razones explicadas en la sección 1.2, vimos que esta matriz de cambio de

base era ni más ni menos que la transpuesta de \mathfrak{B} , si los cambios de base los ponemos como:

$$\mathfrak{Q}_i : \mathbb{R}^{mn} \rightarrow \mathbb{V}_i \quad (2.19)$$

entonces $\mathfrak{Q}_i = \mathfrak{B}_i^T$. Note que cualquier elemento del espacio \mathbb{R}^{mn} puede ser cambiado a cualquier base \mathbb{V}_i , sin embargo, la idea es que solo las imágenes que participaron en la construcción del espacio sean movidas ahí junto con la o las imágenes con que se quieran comparar.

1.4 SELECCIÓN DE LOS VALORES PROPIOS PRINCIPALES

Existen algunos criterios que se siguen para saber qué eigenvalores principales con sus correspondientes eigenvectores deben conservarse. Ya sabemos que mientras más alto es el valor absoluto del eigenvalor más importante es su eigenvector asociado, y que los eigenvectores con eigenvalores menores tienen poca información de la relación entre los datos y por ello pueden eliminarse. Es por eso necesario saber a partir de qué magnitud (umbral) de eigenvalor podemos con toda seguridad ignorar y eliminar eigenvectores. La siguiente lista muestra algunos, quizá los más importantes, criterios de eliminación:

Gráfica de valores propios: Consiste en graficar la magnitud absoluta contra el número de eigenvalor en orden de importancia (de mayor a menor o viceversa). En el eje de las abscisas colocamos el número de eigenvalor y en el eje de las ordenadas colocamos su valor absoluto, esto produce un trazado (ver fig. 2.3) que se asemeja a una curva que va creciendo conforme el número de la componente se hace mayor.¹ El criterio consiste en identificar el punto de la gráfica donde la razón de cambio aumenta muy rápidamente, luego, se ignoran los valores propios que se encuentren hacia la izquierda de la abscisa de este punto, o lo que es lo mismo, se eligen aquellos cuyo valor absoluto sea mayor a la ordenada del punto identificado. En la gráfica de ejemplo ya mencionada, vemos que la abscisa del punto donde la pendiente empieza a crecer rápidamente es 16 y la ordenada de ese punto es 2943, por lo tanto, los eigenvalores

¹La traza en línea sólida corresponde a 18 puntos que son unidos por líneas rectas y la línea punteada muestra una aproximación a tales puntos utilizando interpolación con polinomios de Lagrange, produciendo con ello una curva suave. Note que la aproximación contiene algunos valores negativos

que sean menores a este valor serán ignorados junto con sus eigenvec- tores correspondientes. Este criterio es flexible, ya que las componentes elegidas dependen del punto de quiebre elegido por observación a lo que se le puede agregar la experiencia personal del observador.

En general, todas las gráficas de eigenvalores tendrán las mismas caracte- rísticas que la ya mencionada aunque algunas estarán menos curvadas que otras.

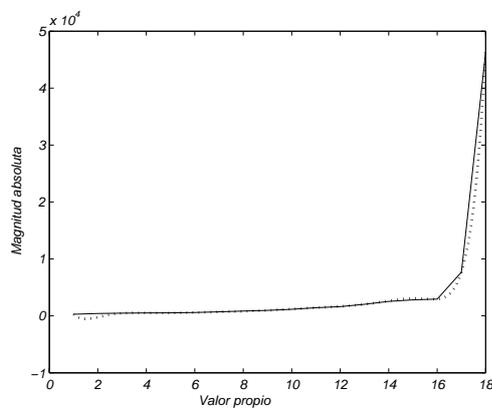


Figura 2.3 Valores propios: Magnitud vs Importancia

Promedio de valores propios: Se promedian las magnitudes absolutas de los eigenvalores y se eliminan aquellos cuya magnitud absoluta sea menor a este promedio. Este criterio es menos flexible ya que el investigador no utiliza su experiencia personal para la elección de las componentes.

Conservación del 98 % de la varianza: Consiste en elegir un número de eigenvectores tal que conserve al menos el 98 % de la varianza en las imágenes, como se menciona en [5]. Este criterio parece ser más complicado que los otros.

El criterio segundo tiene una magnitud límite de eigenvalor muy cercana al criterio primero (cerca del quiebre de la curva). Para comprobación se da el valor promedio de los eigenvalores mostrados en la gráfica 2.3: 4099.7. Utilizar la media de los valores es automático y computacionalmente es fácil de obtenerla y por observación de la gráfica requeriría de la intervención humana para saber el valor del umbral.

1.5 MEDIDAS DE SIMILITUD

Las medidas de similitud comparan qué tan cercanos o parecidos son dos puntos (o vectores, si se consideran las coordenadas de los puntos como sus componentes en una base canónica) en un espacio multidimensional. Existe una variedad de medidas, dependiendo de la aplicación a la cual se destina, más sin embargo las siguientes son más comunes a PCA e ICA (en especial, la medida MAH esta relacionada más a PCA por que hace uso de los valores propios). Sin embargo, no necesariamente deben ser puntos o vectores en un espacio los que se desean comparar para utilizar esta medida de similitud, si no que pueden ser cualquier grupo de valores numéricos.

Norma L1: Esta medida obtiene la suma de las diferencias absolutas entre las componentes de los coeficientes correspondientes a cada imagen:

$$L1 = \sum_{i=1}^M |a_i - a_i^{test}|$$

donde M es el número de componentes en cada grupo de coeficientes.

Norma L2: o distancia euclídea, suma las diferencias al cuadrado entre las componentes de los dos grupos de coeficientes:

$$L2 = \sum_{i=1}^M (a_i - a_i^{test})^2$$

Esta medida es el cuadrado de la distancia entre los dos puntos en el espacio euclidiano M -dimensional.

Mahalanobis: Esta distancia se utiliza en el caso que se utilice el concepto de eigenvectores y eigenvalores (como en PCA):

$$MAH = - \sum_{i=1}^M \frac{a_i \cdot a_i^{test}}{\sqrt{\lambda_i}}$$

donde λ_i son los valores propios asociados a cada vector propio de una matriz de covarianzas.

Cosenos: tratándose de vectores, esta medida no es otra cosa que el coseno del ángulo que forman los dos vectores:

$$c = \frac{\mathbf{a} \cdot \mathbf{a}_{test}}{\|\mathbf{a}\| \|\mathbf{a}_{test}\|}$$

El valor de c se encuentra entre los límites -1 y $+1$.

$L1$, $L2$, MAH y c son un número que nos dice qué tan parecidos son dos puntos, dos vectores o dos grupos de coeficientes; mientras este número más se acerque a cero en el caso de $L1$ y $L2$, a uno en el caso de c , y sea más negativo en el caso de MAH , más cercanos serán los puntos, vectores, etc.

Para un mismo par que se compara con todas de las medidas anteriores, cuando el máximo del coseno ocurra, ocurrirá el mínimo de la distancia euclidiana y el de la norma $L1$ y el valor más negativo de la distancia de Mahalanobis MAH . La medida de cosenos, es equivalente a la distancia euclidiana si antes de medir la distancia primero normalizamos la longitud de los vectores, de tal forma que cuando la medida de cosenos sea 1, la norma $L2$ será 0. Compruébelo con el siguiente desarrollo:

$$\begin{aligned} d^2(x, y) &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\mathbf{x} \cdot \mathbf{y} \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos(\alpha) \end{aligned}$$

Por lo tanto, si $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$ entonces:

$$d^2(x, y) = 2(1 - \cos(\alpha)) \quad (2.20)$$

1.6 COMPARANDO ROSTROS CON PCA CLÁSICO

La comparación es simple, una vez que hemos obtenido la nueva base β' , cambiamos a cada imagen en las columnas de \mathcal{P} a dicha base, de igual forma, las imágenes de prueba en las columnas de la matriz \mathcal{P}_{test} se cambian al mismo espacio y se comparan ahí directamente, componente contra componente, con una medida de similaridad. Ello resultará en que para un rostro de prueba habrá k números (k es el número de imágenes de entrenamiento) y de ellos el máximo o el mínimo nos dirá con cual de las de entrenamiento es más parecida la de prueba.

1.7 COMPARANDO ROSTROS CON PCA MODIFICADO

La comparación aquí es un poco diferente, primero, cada grupo de imágenes se cambia a su propia base que generó—a la cual pertenece—, después la o las imágenes de prueba se cambian a la primera base, ahí se compara(n) con las imágenes pertenecientes a esa base (recuerde que son todas del mismo sujeto), en este punto tenemos para cada imagen de prueba un grupo de coeficientes que nos dictan el parecido entre ella y cada una de las imágenes de la base, lo que se puede hacer aquí es calcular el promedio de esos coeficientes, para de esa forma obtener el parecido promedio que puede utilizarse como un coeficiente que represente la similitud entre la imagen de prueba y el sujeto de las imágenes, pero esto no nos ofreció buenos resultados en las pruebas experimentales realizadas, así que lo descartamos; otro método que muestra más aciertos consistió en elegir el coeficiente máximo o mínimo, según la medida de similitud, como el coeficiente que mide el parecido entre la imagen de prueba y el sujeto en cuestión. Una vez hecho lo anterior en cada uno de los nuevos espacios, resultará que para cada imagen de prueba habrá nuevamente un grupo de coeficientes que representarán su similitud con cada uno de los sujetos, y de esta manera se procedería a buscar el coeficiente más alto o bajo que supere el umbral o límite de semejanza establecido (si es que se sugiere uno).

Al pasar cada imagen de prueba o generalización a algún espacio, se deberá restar a esa imagen la imagen media del conjunto de imágenes del mismo rostro que generaron dicho espacio, esto refuerza nuestra idea de modificación de correlaciones, ya que según nuestra hipótesis, la correlación se modifica seriamente cuando ésta se mide entre imágenes que se les ha restado una imagen media producto de varios rostros y que si se trataba de la media producto de un conjunto de imágenes de un mismo rostro las correlaciones entre imágenes no sufrirían cambio drástico que dañase la identidad del rostro produciendo errores en el reconocimiento. Por lo que, si se pasa una imagen de un sujeto a un espacio de otro sujeto, sucederá que las correlaciones cambiarán y la identificación irá mal ya que sólo irá bien cuando se pase a un espacio del mismo sujeto debido a la no alteración severa de correlaciones.

2. ANÁLISIS CRÍTICO DEL ESTADO DEL ARTE

Es de esperarse que al modificar las correlaciones entre imágenes, la autenticación no irá bien, esto debido a la interferencia de relación entre imágenes,

que como veremos más detalladamente en el apartado de Discusión y Análisis, tienen más correlación cuando pertenecen al mismo sujeto.

Prácticamente todos los diferentes trabajos de investigación computan la media del conjunto de imágenes de la base de datos, pero solamente con el objeto de modificar a las mismas de tal forma que dicho conjunto de imágenes tenga media cero, hecho que se emplea para facilitar la construcción de la matriz de covarianza. Y esto hecho pensamos, introduce ruido y modifica correlaciones, ello en base a que la imagen media así substraída es una mezcla fantasma de todas las imágenes de las cuales se obtuvo, de tal manera que si tales imágenes no son del mismo sujeto, como podemos esperar obtener buenos resultados al restar esta imagen media—que trae información de las demás—a otra imagen con la que no está fuertemente correlacionada.

Como Figura 2.4 muestra, la imagen media es un suavizado del conjunto de imágenes. Esto no debe extrañar ya que lo que hemos utilizado es simplemente un filtro conocido como *filtro promedio* que, como bien se sabe, tiene la característica de ser un filtro pasabajos que elimina precisamente altas frecuencias y ruido [6][7]. El resultado del filtrado es entonces el mantener las características por pixel asociadas a diferentes imágenes evitando que los cambios entre éstas se den de una forma brusca permitiéndonos identificar claramente a varios sujetos y no a uno solo.



Figura 2.4 Apariencia de la imagen media

La figura 2.5 muestra el aspecto de las imágenes cuando se les ha restado la imagen media, como vemos, aún se puede distinguir los rostros de los sujetos, pero si comparamos las correlaciones entre estas imágenes, veremos que han sufrido un gran cambio.

2.1 DE LAS MEDIDAS DE SIMILITUD

El método PCA transforma nuestras imágenes originales a un nuevo espacio, espacio en el que podemos comparar rostros con mayor confiabilidad. Las medidas de similitud deberán ser capaces de tomar en cuenta los cambios



Figura 2.5 Apariencia de las imágenes después de haber restado la imagen media

acaecidos en nuestros datos al ser movidos al nuevo espacio, cambios realizados precisamente con un fin. Al utilizar una medida de distancias que no tome en cuenta tales cambios sería como ignorar el procedimiento PCA y en consecuencia nuestros resultados serían desfavorables.

Como ejemplo mire la figura 2.6 que muestra una nube de puntos en el espacio bidimensional, estos puntos representan nuestros datos originales (las coordenadas para un solo eje—digamos X —son consideradas como las observaciones de la muestra de una variable aleatoria X) antes de que se les aplicará PCA, y la nube de puntos en la figura 2.7 después de haber aplicado PCA. En esta última figura se han girado los nuevos ejes encontrados (y no se ha eliminado a ninguno) por PCA de tal forma que coincidieran con los ejes originales, así que lo único que hemos hecho es haber rotado la nube de puntos original y por lo tanto la distancia entre un punto y otro sigue siendo la misma.

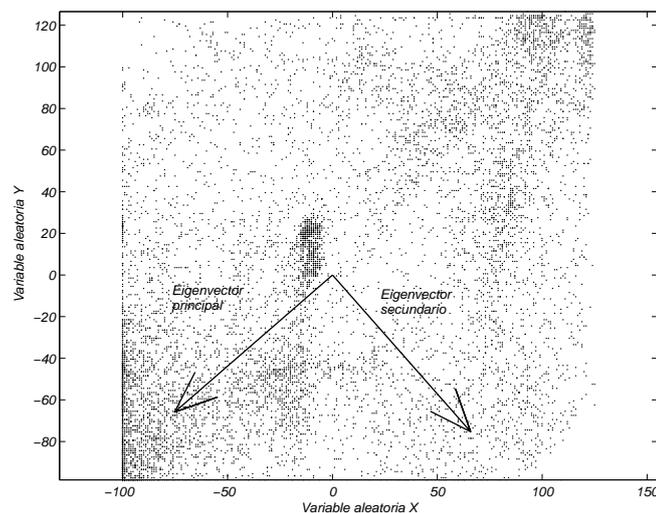


Figura 2.6 Nube de puntos de los datos originales

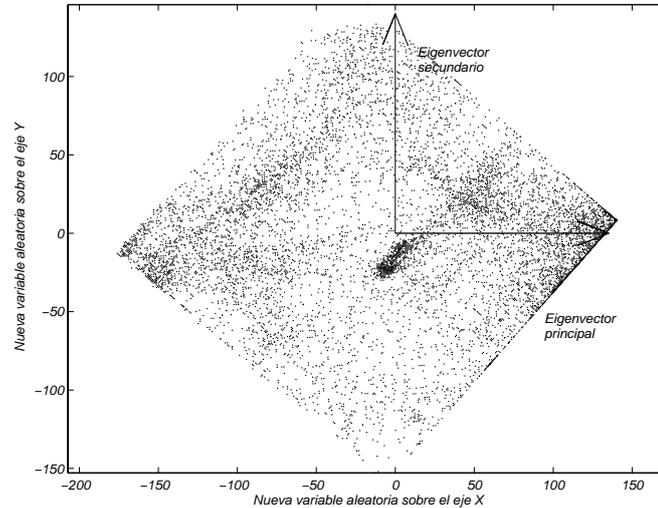


Figura 2.7 Nube de puntos rotada

Lo que ocurre con nuestros datos transformados es que ahora tienen nuevas coordenadas. Si la medida de distancias de cosenos y de norma L2 miden el coseno—valga la redundancia—y el cuadrado de la distancia entre dos puntos, respectivamente, y si ésta no ha cambiado con la sola rotación de los datos, entonces, ¿Será necesario utilizarlas?, ¿Darán mejores resultados utilizar la norma L1 o la distancia de Mahalanobis o alguna otra?

Sin embargo, si los nuevos ejes no siguen manteniendo la misma dimensionalidad ni ortogonalidad que los anteriores, no tiene sentido preguntarse si la distancia entre puntos o el ángulo entre vectores (si consideramos las coordenadas de los puntos como componentes de un vector) se conserva igual después de aplicar PCA, ya que estamos hablando de una nueva dimensión y no podemos comparar los ángulos o distancias entre dos vectores si ellos son calculados en espacios de diferente dimensión. Recuerde que cuando un vector existe en un espacio vectorial \mathbb{R}^m no puede existir también en un espacio vectorial de dimensión \mathbb{R}^n siendo $m \neq n$. Sin embargo, como la característica principal de PCA es la reducción de dimensionalidad, no tenemos que preocuparnos del dilema.

2.2 REPRESENTACIÓN DE ROSTROS

La representación de los rostros más sencilla para identificación es de aquellas basadas en plantilla, una plantilla no es más que los valores de intensidad

de pixel de las imágenes de los rostros, procesadas—normalizadas, imagen de bordes, etc.—o sin algún tipo de procesamiento; si bien es cierto que al utilizar imágenes de bordes se incrementa la eficiencia cuando de objetos se trata [8], tal vez no suceda lo mismo con rostros. También puede utilizarse más de una plantilla por rostro para comparaciones desde otros puntos de vista. Esta es la forma más sencilla de representación de un rostro, pero con altos requerimientos de memoria y de tiempo y con poca eficiencia en la certeza. Aunque podría utilizarse plantillas de menor tamaño que representen sólo ciertas características importantes de un rostro, como los ojos, la nariz, la boca, el pelo, etc.

La otra representación de un rostro es la basada en características, características geométricas como las medidas de posición y tamaño de ojos, nariz, boca, etc. o la distancia entre unas partes y otras. Por supuesto que los requerimientos de memoria serán menores y la velocidad de reconocimiento será mayor, sin embargo, la implementación para este tipo de representación es más complicada en cuanto a los algoritmos que medirán de la forma más perfecta las características mencionadas. Por su naturaleza, este tipo de representación no puede utilizarse en PCA, ya que considera a las imágenes como vectores y no toma en cuenta las características locales de los rostros si no que procesa a los rostros de forma global.

La ventaja que tiene PCA, y de hecho es una de sus características principales, es que reduce la dimensionalidad, ya que lo que utiliza es la representación de rostros basada en apariencia, esto es, una proyección de la imagen del rostro en un subespacio lineal de menores dimensiones y esto por supuesto brinda la ventaja de menor cantidad de memoria requerida y mayor velocidad en el reconocimiento, podríamos colocar a este tipo de representación de rostros como en un punto intermedio entre la representación basada en plantilla y la basada en apariencia, esto en cuanto a requerimientos de memoria, velocidad de reconocimiento y certeza.

2.3 TRUCO DE TURK Y PENTLAND

Cuando una matriz de covarianzas es muy grande, es difícil de tratarla computacionalmente debido a las limitaciones de memoria que existen, y aunque fuera posible tratarla, muchas veces es innecesario. Como ya vimos en PCA original, la dimensión de nuestra matriz de covarianzas es muy grande, de orden mn , donde mn es el número de píxeles en cada imagen. Como sabemos del álgebra lineal, de esta matriz podemos obtener un número total de mn eigenvectores, cada uno con dimensión mn , aunque no es necesario man-

tener a todos. Eso es un número muy grande si jugamos con imágenes de alta resolución y si estamos limitados por nuestra capacidad computacional.

Para evitarnos lo anterior, utilizamos el conocido truco de Turk y Pentland. Este truco nos permite tratar una matriz de covarianzas menor y nos limita a encontrar tan sólo los primeros k eigenvectores principales, donde k es el número de imágenes; la dimensión de los eigenvectores permanece inalterada.

Como conocemos, la matriz de covarianzas mide la relación entre dos variables y en PCA clásico consideramos a los píxeles de las imágenes como tales, lo que resultará en una matriz muy grande de covarianzas. Recuerde que para calcular la matriz de covarianzas se necesita restar la media a cada variable. Según el truco, hay un teorema del álgebra lineal que establece que los eigenvalores de la matriz

$$\Phi = (\mathcal{P} - \Upsilon)(\mathcal{P} - \Upsilon)^T$$

son los mismo que los eigenvalores de la matriz

$$\Phi' = (\mathcal{P} - \Upsilon)^T(\mathcal{P} - \Upsilon)$$

y que por lo tanto los eigenvectores de Φ son los mismos que los de Φ' multiplicados por la matriz de datos \mathcal{P} (ver ec. 2.21, donde \mathfrak{P}' son los eigenvectores en columna de Φ' , y si se requiere, normalizados. El segundo producto, es como si considerásemos a las imágenes como variables aleatorias, y en lugar de obtener una matriz de covarianzas de $mn \times mn$, obtendríamos una de $k \times k$, lo que resultaría en k eigenvectores de dimensión k cada uno; luego, estos multiplicados por \mathcal{P} nos darían los k eigenvectores de dimensión mn cada uno.

$$\mathfrak{P} = \mathcal{P}\mathfrak{P}' \tag{2.21}$$

Tome en cuenta que en el producto de Φ' podría parecer que se calcula las covarianzas existentes entre imágenes, pero esto no es así, dado que la forma de restar la media no es diferente de aquella del primer producto Φ . Como podemos darnos cuenta, automáticamente hemos reducido la dimensionalidad de los datos.

2.4 AUTENTICACIÓN VS IDENTIFICACIÓN

Una autenticación es una comparación uno a uno, porque de antemano se sabe el rostro de la base de datos con la que será comparada el rostro de petición (aquel que pide ser validado). En cuanto a la identificación, esta es una comparación uno a varios, es decir, la comparación de un rostro de prueba con toda o una parte de la base de datos para encontrarse o no en ella. Podríamos considerar a una identificación como numerosas autenticaciones.²

Como para cada imagen hay una medida que representa su parecido con otra imagen, es importante seleccionar un valor de umbral o límite de semejanza a tal medida, por debajo o por encima del cual pudiésemos decir que las dos imágenes pertenecen al mismo rostro.

En una identificación no necesariamente se debe establecer el umbral, ya que el resultado de una búsqueda podría tan sólo decirnos con cuál rostro de la base de datos se parece más el rostro que pide ser identificado, sin necesidad de saber si la medida de similitud entre ambos rostros sobrepasa el límite.

En este trabajo se estudia más la forma en que se comparan dos rostros, dicha forma puede utilizarse para autenticar, o para identificar una vez que se establece un umbral.

3. CONCLUSIÓN

Depende del método para autenticación utilizado, la forma en que se comparan las imágenes, en algunos métodos como PCA e ICA, se utilizan las medidas de distancia como la norma L1 y la norma L2, o medidas de ángulos como la medida de cosenos. Es importante conocer su funcionamiento para aplicarlas de forma adecuada y no influir en el desempeño del método de autenticación en el cual se aplica. Según [5], en el método ICA, el desempeño es mejor con la medida de similaridad de cosenos que con la medida de la distancia euclidiana, y que para PCA el desempeño era el mismo con cualesquiera de las dos medidas de similitud.

²A veces ambos términos son utilizados indistintamente (como en este trabajo).

Capítulo 3

DISCUSIÓN Y ANÁLISIS

El que parece sabio entre los tontos parece tonto entre los sabios.

—Quintiliano

Introducción

Un análisis cuidadoso de la forma en que se efectúa el proceso de aprendizaje empleando PCA, da lugar a suspicacias. La búsqueda de correlación entre variables aleatorias tomando como conjunto muestra de intensidades de pixel a aquellos proporcionados por cada una de las imágenes de entrenamiento de cada sujeto, deja un mal sabor de boca, ya que se presume esperar que exista poca correlación de esta forma, que considerando como conjunto muestra solamente las intensidades de pixel de imágenes para cada sujeto.

De hecho, esta suspicacia da pie a que, en este trabajo, se propongan al menos cuatro formas de emplear PCA que garanticen un índice alto de confiabilidad en el reconocimiento. Estas cuatro formas son las siguientes

1. Restar la media propia de cada imagen y adecuar los pasos posteriores de PCA para trabajar de esta manera. En este sentido cada imagen es una variable.
2. Para una imagen, cada fila define una variable. Por lo que imágenes con resolución $m \times n$ tienen m variables. En esta caso, la muestra de una variable es dada por el conjunto de intensidades en la fila correspondiente.
3. Lo mismo que en el punto 1 anterior, pero para columnas. Con resolución $m \times n$ se tienen n variables y la muestra de cada una de ellas es dada por el conjunto de intensidades en la columna.

4. Se aplica PCA tomando a cada sujeto por separado. Es decir, se crean módulos por cada conjunto de imágenes por persona, de tal forma que sobre éstos se aplica PCA. Esto equivale a contar con tantos módulos como personas haya en la base de datos.
5. Una combinación de los puntos 1 y 2, en donde en lugar de hablar de eigenrostros, se hablará de eigenfilas y eigencolumnas.

La motivación de lo anterior es debido a un análisis experimental que compara los niveles de correlación cuando se consideran la situación usual de aplicación de PCA y los puntos 1 hasta 4 anteriores. Se intuye que los valores de pixel para filas cercanas se encuentran más correlacionados que para filas lejanas.

1. DE LAS CORRELACIONES ENTRE IMÁGENES

La experiencia que se tiene manejando imágenes de rostros ha demostrado que las correlaciones entre imágenes son más altas mientras más parecidas sean las imágenes, es decir, mientras más cercanos sean los valores de intensidad de los píxeles que se corresponden. En cuanto a las imágenes de rostros, es obvio que para dos imágenes de un mismo rostro la correlación será más alta que para dos imágenes de rostros pertenecientes a sujetos diferentes, y esto bajo ciertas condiciones, dependiendo de la correspondencia rostro a rostro entre las imágenes, queriendo decir con esto que los píxeles de una parte del rostro en una imagen—poniendo como ejemplo los píxeles que esbozan la nariz—deben corresponder exactamente con los píxeles que esbozan la misma parte en la otra imagen del rostro, y así con el resto de la cara; mientras más correspondencia exista entre las mismas áreas del rostro más certeza se tendrá de que a rostros iguales más alta correlación y a rostros diferentes menor correlación. Note que tal correspondencia debe cumplirse igualmente con los puntos de vista desde los que se mira a los rostros.

La detección de rostros es utilizada cuando no existen condiciones controladas, ya que no es posible obtener el rostro específicamente donde se quiere, para ello se emplean técnicas de detección que separan el rostro del resto de la escena, cosa fácil si el fondo de la imagen estuviera controlado, pero como en ciertas aplicaciones esto no ocurre, se utilizan otros tipos de detección más sofisticados, uno de ellos es la que utiliza el detector de bordes de Canny para encontrar el contorno del rostro y dibujar una elipse alrededor de este para separar la región del rostro de la cara del resto de la imagen, sin embargo esta técnica ha sido probada con vistas frontales, con vistas desde otros ángulos esta técnica debería ser estudiada.

De la misma manera, cuando los rostros no están a la misma escala la correlación de dos imágenes de un mismo rostro no podrá ser alta, otros factores también influyen en que esto suceda, como los cambios de iluminación, y como ya se han mencionado, las expresiones faciales extremas —ya que cambian significativamente la geometría facial— y el uso de gafas oscuras, cambios en el estilo del peinado y maquillaje [9]. En el caso de las expresiones faciales, varios intentos se han hecho con gráficos de computadoras para modelarlas desde un punto de vista muscular y también métodos de transformación afines se han propuesto para normalizar expresiones.

Los rostros de las imágenes aquí utilizadas están lo suficientemente alineados unos con otros, por lo que la detección de rostros es innecesaria y no debemos preocuparnos por ello.

Para analizar la correlación entre rostros pongamos como ejemplo dos imágenes, mismas que se muestran en la fig. 3.1, cuya correlación es de 0.9367, como ya bien sabemos, esto debido a que son del mismo rostro. Para tener una idea más gráfica de esta correlación veamos una traza (fig. 3.2) en la que hemos asignado el eje X a la primer imagen y el eje Y a la segunda.



Figura 3.1 Imágenes de un mismo sujeto

La nube de puntos—puntos que representan a las entradas de las imágenes— así obtenida nos muestra cómo tales puntos se distribuyen muy cercanamente a lo largo de una recta de 45 grados (la minoría de puntos se esparcen más allá de esta recta); si asignáramos la misma imagen a ambos ejes, obtendríamos exactamente, y sin rastro de duda, una recta de puntos a 45 grados (ver fig. 3.3), y sabríamos que tal recta tendría una correlación de uno, lo cual además es obvio ya que la correlación de una variable consigo misma es la unidad.

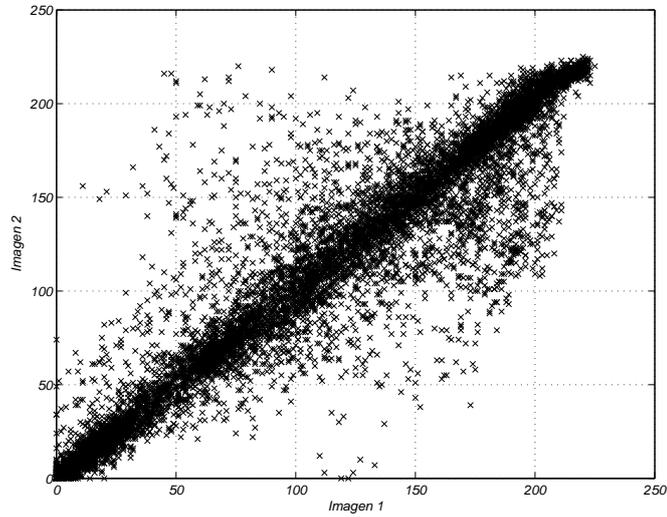


Figura 3.2 Nube de puntos para las fotografías de la figura 3.1

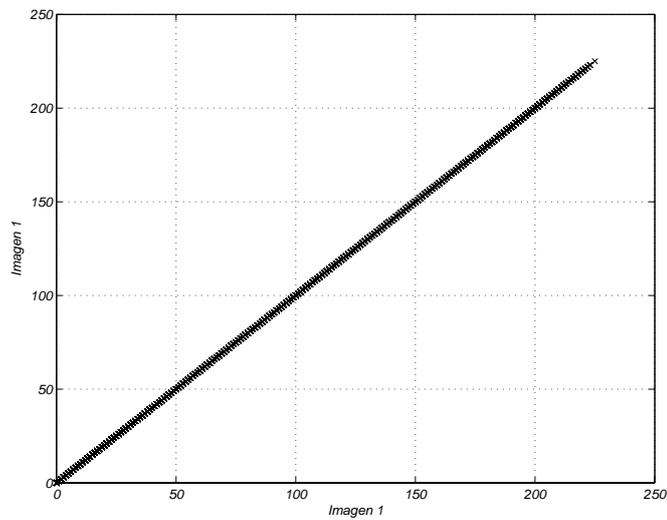


Figura 3.3 Nube de puntos para la misma fotografía

Ahora veamos la nube de puntos que forman las dos imágenes de diferentes rostros que se localizan en la figura 3.4, tal nube (fig. 3.5) se nos muestra más dispersa y no es tan visible una recta a lo largo de la cual podríamos decir se distribuyen los puntos; en tal gráfico podemos ver fuertes concentraciones de

puntos y puntos demasiado dispersos, la correlación entre estas dos imágenes es de 0.7071.



Figura 3.4 Dos sujetos diferentes

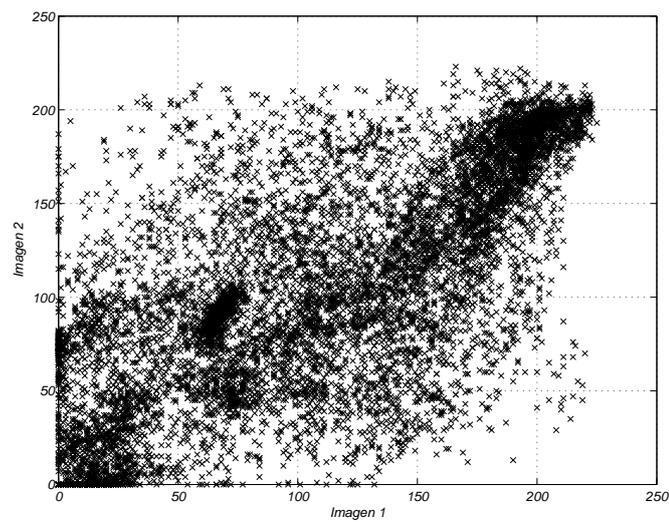


Figura 3.5 Nube de puntos para dos fotografías de diferentes sujetos

1.1 VARIACIÓN DE LA CORRELACIÓN

Hemos dicho que no siempre podemos tener la certeza de que una alta correlación corresponde a imágenes de un mismo rostro. Echemos un vistazo a las correlaciones de las imágenes de la figura 3.6.



Figura 3.6 Rostro con rotación horizontal de aproximadamente 45°

Aunque las dos fotografías pertenecen al mismo rostro, su correlación -de 0.6558- es más baja que la correlación respectiva de las imágenes en la figura 3.4. Note que el rostro de la segunda fotografía está rotado aproximadamente en 45° y por ello no existe correspondencia pixel a pixel entre las diversas zonas del rostro (la nariz no coincide con la nariz, la boca no coincide con la boca, etcetera, etc.); Para una visión más gráfica observe la nube de puntos correspondiente en la figura 3.7.

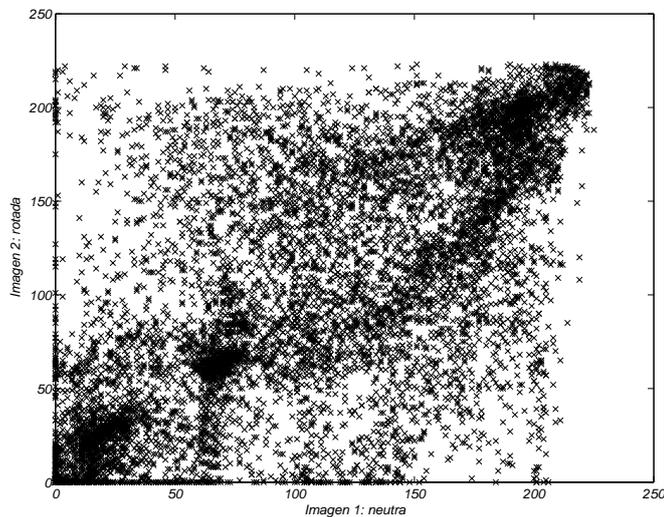


Figura 3.7 Nube de puntos correspondiente a dos fotografías del mismo rostro: una frontal y la otra rotada 45°

De forma análoga al caso de un rostro rotado, podríamos entender el porque de la baja correlación cuando los rostros contenidos en dos imágenes se encuentran desfasados en posición; para evitarnos de ello, las fotografías utilizadas en la autenticación deberán tener sus bordes en tangente con el contorno de la cara (como lo muestran las fotografías utilizadas en este contexto).

Ahora observemos las correlaciones entre las imágenes de un mismo rostro en otras situaciones, como ejemplo, vea las imágenes (enumeradas de 1 a 6 de derecha a izquierda y de arriba hacia abajo) de la figura 3.8, y compare con los respectivos coeficientes de correlación de la tabla 3.1.



Figura 3.8 Algunos de los cambios posibles más comunes en un rostro a la hora del reconocimiento

Tabla 3.1 Correlación entre diversas imágenes de un mismo sujeto

	1	2	3	4	5	6
1	1.0000	0.8235	0.7549	0.8277	0.7861	0.8196
2	0.8235	1.0000	0.6345	0.9392	0.8604	0.9542
3	0.7549	0.6345	1.0000	0.5961	0.7123	0.6666
4	0.8277	0.9392	0.5961	1.0000	0.7953	0.8993
5	0.7861	0.8604	0.7123	0.7953	1.0000	0.8643
6	0.8196	0.9542	0.6666	0.8993	0.8643	1.0000

De todas las magnitudes de correlación, la más baja es por supuesto de cuando el rostro aparece rotado (como constatamos más arriba) y cuando se utilizan gafas oscuras; y esto hablando con respecto a un rostro de expresión neutra sin rotaciones. Otros cambios menores no disminuyen significativamente

la correlación. Por tal motivo, cuando los cambios son muy drásticos entre dos imágenes, sus correlaciones serán muy bajas, y la autenticación será de poca confiabilidad en esos casos. Visualice que es lo mismo poner imágenes sin rotaciones en la etapa de entrenamiento y sí rotadas en el etapa de prueba o identificación, y viceversa.

1.2 SUSTRACCIÓN DE LA IMAGEN MEDIA

Lo que pasa al sustraer la imagen media de un conjunto de imágenes de rostros diferentes (ver fig. 3.9) es que los valores de correlación cambian súbitamente como lo muestran las tablas 3.2 y 3.3.

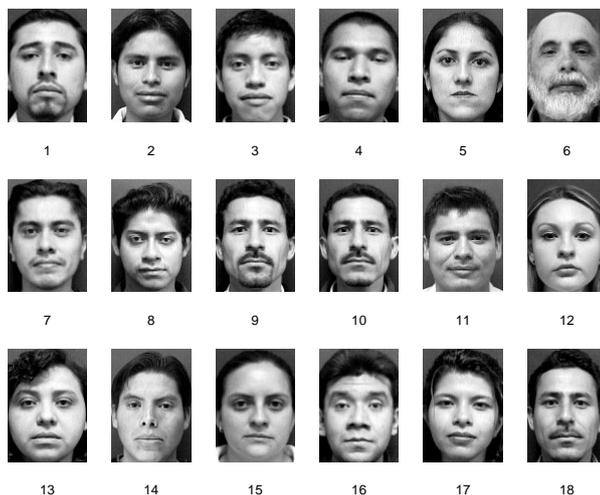


Figura 3.9 Diversos rostros

Los cambios parecen no seguir un patrón reconocible, por lo que las correlaciones, parece, no han cambiado siguiendo una regla. Para comprobación vea la tabla 3.4 que muestra la diferencia de las correlaciones entre las fotografías, antes y después de la sustracción.

¿Pero qué es lo que pasa cuando las fotografías son del mismo rostro?, la figura 3.10 muestra los rostros y las tablas 3.5 y 3.6 las correlaciones correspondientes antes y después de restar su imagen media. La tabla 3.7 muestra las diferencias entre estas correlaciones. La disminución es más constante y proporcional; verifique que la correlación más alta disminuye en menor proporción.

Tabla 3.2 Correlación antes de sustraer la imagen media

	1	2	3	4	5	6	7	8	9
1	1.0000	0.3271	0.4184	0.7071	0.5019	0.3721	0.6372	0.1790	0.6496
2	0.3271	1.0000	0.6668	0.4844	0.6675	0.0271	0.5813	0.4679	0.5440
3	0.4184	0.6668	1.0000	0.5944	0.7310	0.0769	0.7198	0.6238	0.6956
4	0.7071	0.4844	0.5944	1.0000	0.7570	0.2707	0.7722	0.3446	0.7555
5	0.5019	0.6675	0.7310	0.7570	1.0000	0.3447	0.8106	0.4623	0.7294
6	0.3721	0.0271	0.0769	0.2707	0.3447	1.0000	0.3015	0.0065	0.2399
7	0.6372	0.5813	0.7198	0.7722	0.8106	0.3015	1.0000	0.4036	0.8464
8	0.1790	0.4679	0.6238	0.3446	0.4623	0.0065	0.4036	1.0000	0.3725
9	0.6496	0.5440	0.6956	0.7555	0.7294	0.2399	0.8464	0.3725	1.0000
10	0.5663	0.5234	0.6645	0.7099	0.7059	0.2091	0.8145	0.4032	0.9201
11	0.0238	0.5923	0.5928	0.1719	0.3941	-0.0937	0.2803	0.6626	0.2865
12	0.4190	0.6018	0.7842	0.5336	0.7251	0.2093	0.7232	0.4752	0.6278
13	0.4826	0.5240	0.7781	0.6585	0.7071	0.1983	0.7149	0.5050	0.6162
14	0.2534	0.7370	0.6959	0.4493	0.7194	0.1913	0.5548	0.5382	0.5436
15	0.5330	0.5320	0.6554	0.5926	0.6523	0.1676	0.6933	0.2899	0.6266
16	0.3794	0.5792	0.8539	0.5225	0.6621	0.1476	0.6544	0.6378	0.6118
17	0.4394	0.7352	0.8504	0.6190	0.8009	0.1687	0.7556	0.6253	0.6728
18	0.5032	0.6728	0.8044	0.6977	0.8071	0.2127	0.8297	0.4997	0.8037

Tabla 3.2 (cont.) ...valores restantes.

	10	11	12	13	14	15	16	17	18
1	0.5663	0.0238	0.4190	0.4826	0.2534	0.5330	0.3794	0.4394	0.5032
2	0.5234	0.5923	0.6018	0.5240	0.7370	0.5320	0.5792	0.7352	0.6728
3	0.6645	0.5928	0.7842	0.7781	0.6959	0.6554	0.8539	0.8504	0.8044
4	0.7099	0.1719	0.5336	0.6585	0.4493	0.5926	0.5225	0.6190	0.6977
5	0.7059	0.3941	0.7251	0.7071	0.7194	0.6523	0.6621	0.8009	0.8071
6	0.2091	-0.0937	0.2093	0.1983	0.1913	0.1676	0.1476	0.1687	0.2127
7	0.8145	0.2803	0.7232	0.7149	0.5548	0.6933	0.6544	0.7556	0.8297
8	0.4032	0.6626	0.4752	0.5050	0.5382	0.2899	0.6378	0.6253	0.4997
9	0.9201	0.2865	0.6278	0.6162	0.5436	0.6266	0.6118	0.6728	0.8037
10	1.0000	0.2599	0.5969	0.5907	0.5064	0.5634	0.5796	0.6629	0.8005
11	0.2599	1.0000	0.4930	0.4638	0.6546	0.2967	0.5925	0.5632	0.4441
12	0.5969	0.4930	1.0000	0.6788	0.6616	0.6881	0.7473	0.8071	0.7597
13	0.5907	0.4638	0.6788	1.0000	0.5685	0.6164	0.7489	0.7451	0.7207
14	0.5064	0.6546	0.6616	0.5685	1.0000	0.4347	0.6863	0.7491	0.6778
15	0.5634	0.2967	0.6881	0.6164	0.4347	1.0000	0.5309	0.6217	0.6730
16	0.5796	0.5925	0.7473	0.7489	0.6863	0.5309	1.0000	0.7891	0.6730
17	0.6629	0.5632	0.8071	0.7451	0.7491	0.6217	0.7891	1.0000	0.8295
18	0.8005	0.4441	0.7597	0.7207	0.6778	0.6730	0.6730	0.8295	1.0000

Lo mismo ocurre con los siguientes tres sujetos que se muestran en la fig. 3.11, compruebe que la correlación más alta ha disminuido en menor proporción que las demás, en las tablas 3.8, 3.9 y 3.10 se muestran las diferencias antes y después de la sustracción de su imagen media.

Tabla 3.3 Correlación después de la sustracción de la imagen media

	1	2	3	4	5	6	7	8	9
1	1.0000	-0.2313	-0.3856	0.4721	-0.1664	0.3096	0.1986	-0.2775	0.2519
2	-0.2313	1.0000	-0.0107	-0.2190	-0.0030	-0.2115	-0.2542	0.0245	-0.2449
3	-0.3856	-0.0107	1.0000	-0.3524	-0.1457	-0.4928	-0.2076	0.1364	-0.1331
4	0.4721	-0.2190	-0.3524	1.0000	0.1844	0.1176	0.2539	-0.2496	0.2601
5	-0.1664	-0.0030	-0.1457	0.1844	1.0000	-0.0354	0.2090	-0.2829	0.0236
6	0.3096	-0.2115	-0.4928	0.1176	-0.0354	1.0000	-0.0634	-0.0839	-0.1470
7	0.1986	-0.2542	-0.2076	0.2539	0.2090	-0.0634	1.0000	-0.4149	0.4446
8	-0.2775	0.0245	0.1364	-0.2496	-0.2829	-0.0839	-0.4149	1.0000	-0.3827
9	0.2519	-0.2449	-0.1331	0.2601	0.0236	-0.1470	0.4446	-0.3827	1.0000
10	0.1369	-0.2100	-0.1850	0.2106	-0.0207	-0.0713	0.3636	-0.2196	0.7414
11	-0.3632	0.3199	0.1110	-0.4123	-0.3572	-0.0248	-0.5801	0.5154	-0.4617
12	-0.2269	-0.0691	0.1589	-0.3479	-0.0504	-0.0822	-0.0513	-0.1098	-0.2490
13	-0.0914	-0.2545	0.2183	0.0226	-0.0125	-0.1838	0.0064	-0.0552	-0.1946
14	-0.4115	0.3856	0.0106	-0.3603	0.0897	-0.0002	-0.4213	0.1359	-0.3166
15	0.1674	-0.0373	-0.0482	0.0524	-0.0548	0.0074	0.0876	-0.2867	-0.0252
16	-0.3184	-0.1191	0.4770	-0.3618	-0.2027	-0.2438	-0.2384	0.2392	-0.2352
17	-0.3849	0.1622	0.3217	-0.3291	0.1399	-0.4108	-0.0798	0.0920	-0.2401
18	-0.1987	-0.0268	0.0885	-0.0495	0.1342	-0.2702	0.2264	-0.2261	0.2455

Tabla 3.3 (cont.) ...valores restantes.

	10	11	12	13	14	15	16	17	18
1	0.1369	-0.3632	-0.2269	-0.0914	-0.4115	0.1674	-0.3184	-0.3849	-0.1987
2	-0.2100	0.3199	-0.0691	-0.2545	0.3856	-0.0373	-0.1191	0.1622	-0.0268
3	-0.1850	0.1110	0.1589	0.2183	0.0106	-0.0482	0.4770	0.3217	0.0885
4	0.2106	-0.4123	-0.3479	0.0226	-0.3603	0.0524	-0.3618	-0.3291	-0.0495
5	-0.0207	-0.3572	-0.0504	-0.0125	0.0897	-0.0548	-0.2027	0.1399	0.1342
6	-0.0713	-0.0248	-0.0822	-0.1838	-0.0002	0.0074	-0.2438	-0.4108	-0.2702
7	0.3636	-0.5801	-0.0513	0.0064	-0.4213	0.0876	-0.2384	-0.0798	0.2264
8	-0.2196	0.5154	-0.1098	-0.0552	0.1359	-0.2867	0.2392	0.0920	-0.2261
9	0.7414	-0.4617	-0.2490	-0.1946	-0.3166	-0.0252	-0.2352	-0.2401	0.2455
10	1.0000	-0.3805	-0.2549	-0.2191	-0.3172	-0.0993	-0.2702	-0.2321	0.2758
11	-0.3805	1.0000	0.0345	-0.0790	0.4223	-0.1370	0.1827	-0.0301	-0.2690
12	-0.2549	0.0345	1.0000	-0.0349	0.0406	0.1749	0.1724	0.2063	0.0200
13	-0.2191	-0.0790	-0.0349	1.0000	-0.2012	-0.0161	0.2307	0.0813	-0.0333
14	-0.3172	0.4223	0.0406	-0.2012	1.0000	-0.3032	0.1139	0.1376	-0.0901
15	-0.0993	-0.1370	0.1749	-0.0161	-0.3032	1.0000	-0.2427	-0.2080	-0.0285
16	-0.2702	0.1827	0.1724	0.2307	0.1139	-0.2427	1.0000	0.2152	-0.2507
17	-0.2321	-0.0301	0.2063	0.0813	0.1376	-0.2080	0.2152	1.0000	0.1834
18	0.2758	-0.2690	0.0200	-0.0333	-0.0901	-0.0285	-0.2507	0.1834	1.0000

Recordemos que a las imágenes de prueba se les resta la imagen media que se obtuvo de las imágenes del conjunto de entrenamiento, y no la del conjunto a la cual pertenece; aunque si bien es cierto que cuando la imagen de prueba es solo una, no se puede restar la imagen media. Por ello es útil indagar sobre

Tabla 3.4 Diferencia entre las correlaciones después de la sustracción de la imagen media

	1	2	3	4	5	6	7	8	9
1	0	-0.5584	-0.8040	-0.2350	-0.6683	-0.0625	-0.4387	-0.4565	-0.3977
2	-0.5584	0	-0.6775	-0.7034	-0.6705	-0.2386	-0.8355	-0.4434	-0.7890
3	-0.8040	-0.6775	0	-0.9468	-0.8767	-0.5697	-0.9274	-0.4874	-0.8287
4	-0.2350	-0.7034	-0.9468	0	-0.5726	-0.1531	-0.5183	-0.5942	-0.4954
5	-0.6683	-0.6705	-0.8767	-0.5726	0	-0.3801	-0.6016	-0.7452	-0.7058
6	-0.0625	-0.2386	-0.5697	-0.1531	-0.3801	0	-0.3649	-0.0904	-0.3869
7	-0.4387	-0.8355	-0.9274	-0.5183	-0.6016	-0.3649	0	-0.8185	-0.4018
8	-0.4565	-0.4434	-0.4874	-0.5942	-0.7452	-0.0904	-0.8185	0	-0.7552
9	-0.3977	-0.7890	-0.8287	-0.4954	-0.7058	-0.3869	-0.4018	-0.7552	0
10	-0.4293	-0.7335	-0.8496	-0.4993	-0.7265	-0.2804	-0.4509	-0.6227	-0.1787
11	-0.3869	-0.2723	-0.4818	-0.5842	-0.7513	0.0689	-0.8603	-0.1472	-0.7481
12	-0.6458	-0.6709	-0.6254	-0.8815	-0.7756	-0.2916	-0.7745	-0.5850	-0.8768
13	-0.5740	-0.7784	-0.5598	-0.6358	-0.7196	-0.3822	-0.7085	-0.5602	-0.8108
14	-0.6649	-0.3513	-0.6852	-0.8096	-0.6298	-0.1915	-0.9761	-0.4024	-0.8602
15	-0.3656	-0.5692	-0.7037	-0.5402	-0.7071	-0.1602	-0.6057	-0.5767	-0.6519
16	-0.6978	-0.6983	-0.3769	-0.8843	-0.8648	-0.3914	-0.8928	-0.3986	-0.8470
17	-0.8243	-0.5730	-0.5287	-0.9482	-0.6611	-0.5795	-0.8354	-0.5334	-0.9128
18	-0.7018	-0.6996	-0.7159	-0.7472	-0.6729	-0.4829	-0.6032	-0.7257	-0.5583

Tabla 3.4 (cont.) ...valores restantes.

	10	11	12	13	14	15	16	17	18
1	-0.4293	-0.3869	-0.6458	-0.5740	-0.6649	-0.3656	-0.6978	-0.8243	-0.7018
2	-0.7335	-0.2723	-0.6709	-0.7784	-0.3513	-0.5692	-0.6983	-0.5730	-0.6996
3	-0.8496	-0.4818	-0.6254	-0.5598	-0.6852	-0.7037	-0.3769	-0.5287	-0.7159
4	-0.4993	-0.5842	-0.8815	-0.6358	-0.8096	-0.5402	-0.8843	-0.9482	-0.7472
5	-0.7265	-0.7513	-0.7756	-0.7196	-0.6298	-0.7071	-0.8648	-0.6611	-0.6729
6	-0.2804	0.0689	-0.2916	-0.3822	-0.1915	-0.1602	-0.3914	-0.5795	-0.4829
7	-0.4509	-0.8603	-0.7745	-0.7085	-0.9761	-0.6057	-0.8928	-0.8354	-0.6032
8	-0.6227	-0.1472	-0.5850	-0.5602	-0.4024	-0.5767	-0.3986	-0.5334	-0.7257
9	-0.1787	-0.7481	-0.8768	-0.8108	-0.8602	-0.6519	-0.8470	-0.9128	-0.5583
10	0	-0.6404	-0.8518	-0.8097	-0.8236	-0.6627	-0.8497	-0.8950	-0.5247
11	-0.6404	0	-0.4585	-0.5428	-0.2323	-0.4337	-0.4098	-0.5933	-0.7131
12	-0.8518	-0.4585	0	-0.7136	-0.6210	-0.5132	-0.5749	-0.6008	-0.7396
13	-0.8097	-0.5428	-0.7136	0	-0.7697	-0.6325	-0.5182	-0.6638	-0.7540
14	-0.8236	-0.2323	-0.6210	-0.7697	0	-0.7379	-0.5724	-0.6115	-0.7678
15	-0.6627	-0.4337	-0.5132	-0.6325	-0.7379	0	-0.7736	-0.8297	-0.7016
16	-0.8497	-0.4098	-0.5749	-0.5182	-0.5724	-0.7736	0	-0.5739	-0.9237
17	-0.8950	-0.5933	-0.6008	-0.6638	-0.6115	-0.8297	-0.5739	0	-0.6460
18	-0.5247	-0.7131	-0.7396	-0.7540	-0.7678	-0.7016	-0.9237	-0.6460	0

lo que sucede cuando se resta a una imagen de un conjunto la imagen media de otro conjunto de imágenes, conjunto al que no pertenece y que por ende no participó en la obtención de su imagen media.



Figura 3.10 Algunas facetas de un rostro

Tabla 3.5 Valores de correlación antes de sustraer la imagen media

	1	2	3	4	5	6
1	1.0000	0.8298	0.5210	0.5651	0.4495	0.5502
2	0.8298	1.0000	0.6635	0.5924	0.5750	0.5883
3	0.5210	0.6635	1.0000	0.4484	0.4671	0.5498
4	0.5651	0.5924	0.4484	1.0000	0.5192	0.6339
5	0.4495	0.5750	0.4671	0.5192	1.0000	0.4132
6	0.5502	0.5883	0.5498	0.6339	0.4132	1.0000

Tabla 3.6 Correlaciones después de sustraer la imagen media

	1	2	3	4	5	6
1	1.0000	0.3842	-0.3018	-0.2360	-0.3549	-0.2716
2	0.3842	1.0000	-0.0691	-0.3858	-0.2085	-0.3893
3	-0.3018	-0.0691	1.0000	-0.3912	-0.1762	-0.1350
4	-0.2360	-0.3858	-0.3912	1.0000	-0.0943	0.0389
5	-0.3549	-0.2085	-0.1762	-0.0943	1.0000	-0.3439
6	-0.2716	-0.3893	-0.1350	0.0389	-0.3439	1.0000

Tabla 3.7 Correlaciones después - Correlaciones antes

	1	2	3	4	5	6
1	0	-0.4456	-0.8228	-0.8012	-0.8044	-0.8217
2	-0.4456	0	-0.7325	-0.9783	-0.7836	-0.9776
3	-0.8228	-0.7325	0	-0.8396	-0.6433	-0.6848
4	-0.8012	-0.9783	-0.8396	0	-0.6135	-0.5951
5	-0.8044	-0.7836	-0.6433	-0.6135	0	-0.7571
6	-0.8217	-0.9776	-0.6848	-0.5951	-0.7571	0

Las tablas 3.11 y 3.12 muestran algunas de las correlaciones de los rostros de la figura 3.9 y todas las de los rostros de la figura 3.10, respectivamente, después de que se les ha restado la imagen media recíprocamente. Verifique con las correlaciones que tenían antes (ver tablas 3.2 y 3.5) de que se les restara la imagen media.

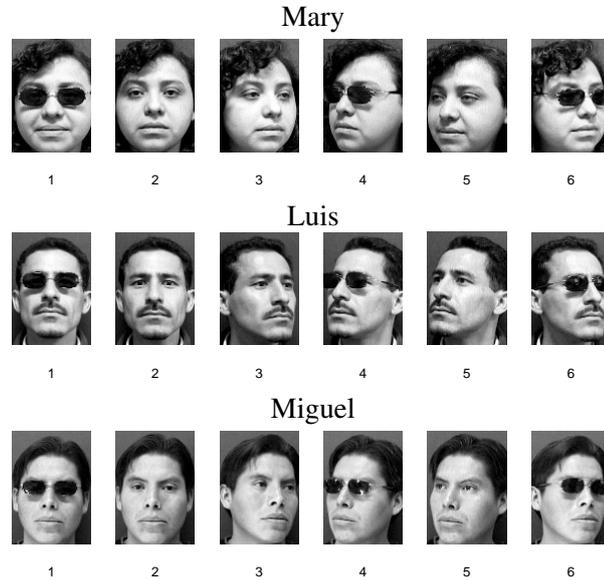


Figura 3.11 Algunas facetas de tres rostros diferentes

Como podemos observar en la tabla 3.11, algunas correlaciones han aumentado y otras disminuido, y según otras pruebas (que no mostramos aquí), esto es aproximadamente en un 50 % y 50 %. La tabla 3.12 muestra que todas las correlaciones han disminuido, sin embargo en otras pruebas con otros sujetos diferentes mostraron que estas correlaciones aumentaban, y la proporción es la misma, alrededor de 50 % por ciento aumentaba y 50 % disminuía.

2. MEDIA POR FILA O COLUMNA

Hemos mencionado lo que pasa con las relaciones entre imágenes cuando restamos alguna imagen media, ahora vamos a analizar lo que sucede con esas relaciones cuando restamos la media a cada fila o a cada columna de la imagen.

2.1 SUSTRACCIÓN DE MEDIA POR FILA

Indaguemos en la tabla 3.13 para deducir lo que acontece con las correlaciones entre las imágenes de la figura 3.9 después de haber restado la media por fila. Comparando estas correlaciones con las de la tabla 3.2, nos damos cuenta de que las correlaciones no han cambiado drásticamente como cuando se resta la imagen media.

Tabla 3.8 Diferencia de correlaciones para Mary

	1	2	3	4	5	6
1	0	-0.6304	-0.8495	-0.7753	-0.8992	-0.7230
2	-0.6304	0	-0.8434	-0.8453	-0.7773	-1.0195
3	-0.8495	-0.8434	0	-1.0170	-0.9108	-0.4766
4	-0.7753	-0.8453	-1.0170	0	-0.5621	-0.8663
5	-0.8992	-0.7773	-0.9108	-0.5621	0	-0.8645
6	-0.7230	-1.0195	-0.4766	-0.8663	-0.8645	0

Tabla 3.9 Diferencia de correlaciones para Hugo

	1	2	3	4	5	6
1	0	-0.3935	-0.9808	-0.9072	-0.9125	-0.8091
2	-0.3935	0	-0.9545	-0.9854	-0.9000	-0.9572
3	-0.9808	-0.9545	0	-1.0030	-0.9482	-0.3094
4	-0.9072	-0.9854	-1.0030	0	-0.3405	-0.9389
5	-0.9125	-0.9000	-0.9482	-0.3405	0	-0.9811
6	-0.8091	-0.9572	-0.3094	-0.9389	-0.9811	0

Tabla 3.10 Diferencia de correlaciones para Miguel

	1	2	3	4	5	6
1	0	-0.4819	-0.9897	-0.8085	-0.9541	-0.9381
2	-0.4819	0	-0.9476	-0.9934	-0.8018	-1.1284
3	-0.9897	-0.9476	0	-1.0363	-1.0052	-0.3679
4	-0.8085	-0.9934	-1.0363	0	-0.3849	-1.0013
5	-0.9541	-0.8018	-1.0052	-0.3849	0	-1.1359
6	-0.9381	-1.1284	-0.3679	-1.0013	-1.1359	0

Tabla 3.11 Correlación después de la sustracción de la imagen media de otro grupo

	1	2	3	4	5	6	7	8	9
1	1.0000	0.4284	0.5017	0.7090	0.4702	0.1612	0.6339	0.3338	0.6627
2	0.4284	1.0000	0.7551	0.5963	0.7080	-0.0085	0.6534	0.6315	0.6227
3	0.5017	0.7551	1.0000	0.6811	0.7657	0.0453	0.7693	0.7331	0.7454
4	0.7090	0.5963	0.6811	1.0000	0.7470	0.0572	0.7803	0.5108	0.7751
5	0.4702	0.7080	0.7657	0.7470	1.0000	0.0940	0.8028	0.5518	0.7336
6	0.1612	-0.0085	0.0453	0.0572	0.0940	1.0000	0.0971	0.0545	0.0925
7	0.6339	0.6534	0.7693	0.7803	0.8028	0.0971	1.0000	0.5290	0.8563
8	0.3338	0.6315	0.7331	0.5108	0.5518	0.0545	0.5290	1.0000	0.5038
9	0.6627	0.6227	0.7454	0.7751	0.7336	0.0925	0.8563	0.5038	1.0000

^a Con imagen media del conjunto de la fig. 3.10 restada.

Tabla 3.12 Correlación después de la sustracción de las imagen media de otro grupo

	1	2	3	4	5	6
1	1.0000	0.7872	0.4000	0.4778	0.2698	0.4637
2	0.7872	1.0000	0.4752	0.3907	0.2473	0.4037
3	0.4000	0.4752	1.0000	0.3071	0.2538	0.4406
4	0.4778	0.3907	0.3071	1.0000	0.3594	0.5639
5	0.2698	0.2473	0.2538	0.3594	1.0000	0.2277
6	0.4637	0.4037	0.4406	0.5639	0.2277	1.0000

^a Con imagen media del conjunto de la fig. 3.9 restada.

Tabla 3.13 Valores de correlación al restar la media por filas

	1	2	3	4	5	6	7	8	9
1	1.0000	0.2633	0.3170	0.6674	0.4240	0.3249	0.5812	0.0519	0.5989
2	0.2633	1.0000	0.6700	0.4327	0.6108	-0.1964	0.5534	0.4347	0.5100
3	0.3170	0.6700	1.0000	0.4998	0.6913	-0.0561	0.6416	0.5369	0.6225
4	0.6674	0.4327	0.4998	1.0000	0.7165	0.1733	0.7168	0.1979	0.7047
5	0.4240	0.6108	0.6913	0.7165	1.0000	0.1377	0.7696	0.3265	0.6853
6	0.3249	-0.1964	-0.0561	0.1733	0.1377	1.0000	0.1896	-0.1504	0.1373
7	0.5812	0.5534	0.6416	0.7168	0.7696	0.1896	1.0000	0.2322	0.8131
8	0.0519	0.4347	0.5369	0.1979	0.3265	-0.1504	0.2322	1.0000	0.2239
9	0.5989	0.5100	0.6225	0.7047	0.6853	0.1373	0.8131	0.2239	1.0000
10	0.5127	0.5071	0.5835	0.6521	0.6688	0.1249	0.7716	0.2469	0.9090
11	-0.0490	0.5593	0.5969	0.1025	0.3298	-0.2515	0.2206	0.6843	0.2298
12	0.3106	0.5573	0.7350	0.4157	0.6373	0.0458	0.6355	0.3450	0.5382
13	0.3909	0.5223	0.7403	0.5978	0.7012	0.1353	0.6603	0.4174	0.5497
14	0.1706	0.6972	0.7491	0.3872	0.6388	-0.1211	0.5097	0.5231	0.5122
15	0.4548	0.5203	0.5702	0.4971	0.5873	0.0664	0.6081	0.1135	0.5479
16	0.2712	0.5567	0.8198	0.4132	0.5929	0.0272	0.5565	0.5619	0.5228
17	0.3353	0.7222	0.8217	0.5203	0.7342	-0.0271	0.6753	0.5301	0.5902
18	0.4230	0.6537	0.7580	0.6263	0.7654	0.0685	0.7815	0.3650	0.7539

2.2 SUSTRACCIÓN DE MEDIA POR COLUMNA

Paralelo a lo que sucede procesando las filas, la sustracción de media por columnas tiene los mismos efectos sobre las correlaciones, es decir, la variación de ésta no es extrema. Para ello vea la tabla 3.14 y compare con la tabla 3.13 y con la tabla 3.2 de correlaciones originales para comprobar esto.

Tabla 3.13 (cont.) ...valores restantes.

	10	11	12	13	14	15	16	17	18
1	0.5127	-0.0490	0.3106	0.3909	0.1706	0.4548	0.2712	0.3353	0.4230
2	0.5071	0.5593	0.5573	0.5223	0.6972	0.5203	0.5567	0.7222	0.6537
3	0.5835	0.5969	0.7350	0.7403	0.7491	0.5702	0.8198	0.8217	0.7580
4	0.6521	0.1025	0.4157	0.5978	0.3872	0.4971	0.4132	0.5203	0.6263
5	0.6688	0.3298	0.6373	0.7012	0.6388	0.5873	0.5929	0.7342	0.7654
6	0.1249	-0.2515	0.0458	0.1353	-0.1211	0.0664	0.0272	-0.0271	0.0685
7	0.7716	0.2206	0.6355	0.6603	0.5097	0.6081	0.5565	0.6753	0.7815
8	0.2469	0.6843	0.3450	0.4174	0.5231	0.1135	0.5619	0.5301	0.3650
9	0.9090	0.2298	0.5382	0.5497	0.5122	0.5479	0.5228	0.5902	0.7539
10	1.0000	0.2140	0.5034	0.5156	0.4916	0.4649	0.4843	0.5816	0.7533
11	0.2140	1.0000	0.4662	0.4675	0.6519	0.2637	0.5922	0.5487	0.4077
12	0.5034	0.4662	1.0000	0.6232	0.6366	0.6109	0.6836	0.7462	0.6949
13	0.5156	0.4675	0.6232	1.0000	0.6373	0.5250	0.7053	0.7167	0.6835
14	0.4916	0.6519	0.6366	0.6373	1.0000	0.4144	0.7274	0.7412	0.6639
15	0.4649	0.2637	0.6109	0.5250	0.4144	1.0000	0.4083	0.5262	0.6025
16	0.4843	0.5922	0.6836	0.7053	0.7274	0.4083	1.0000	0.7414	0.5939
17	0.5816	0.5487	0.7462	0.7167	0.7412	0.5262	0.7414	1.0000	0.7796
18	0.7533	0.4077	0.6949	0.6835	0.6639	0.6025	0.5939	0.7796	1.0000

Tabla 3.14 Valores de correlación al restar la media por columnas

	1	2	3	4	5	6	7	8	9
1	1.0000	0.3181	0.3865	0.6190	0.4421	0.3471	0.5816	0.1740	0.5690
2	0.3181	1.0000	0.5524	0.4834	0.6544	0.0936	0.5284	0.4172	0.4939
3	0.3865	0.5524	1.0000	0.6020	0.7231	0.1415	0.6929	0.6043	0.6773
4	0.6190	0.4834	0.6020	1.0000	0.7374	0.2475	0.7158	0.4017	0.6827
5	0.4421	0.6544	0.7231	0.7374	1.0000	0.3508	0.7713	0.4588	0.6830
6	0.3471	0.0936	0.1415	0.2475	0.3508	1.0000	0.3230	0.0548	0.2356
7	0.5816	0.5284	0.6929	0.7158	0.7713	0.3230	1.0000	0.4385	0.7888
8	0.1740	0.4172	0.6043	0.4017	0.4588	0.0548	0.4385	1.0000	0.4160
9	0.5690	0.4939	0.6773	0.6827	0.6830	0.2356	0.7888	0.4160	1.0000
10	0.4694	0.4844	0.6539	0.6282	0.6577	0.1937	0.7552	0.4680	0.8912
11	0.0020	0.5207	0.4812	0.2301	0.4157	0.0186	0.2514	0.5600	0.2892
12	0.3738	0.5313	0.7465	0.5301	0.7053	0.2377	0.6869	0.4844	0.5713
13	0.4237	0.3841	0.6885	0.6236	0.6341	0.2218	0.6431	0.4146	0.5213
14	0.2060	0.6775	0.5891	0.4438	0.7139	0.2749	0.5014	0.4696	0.5021
15	0.5886	0.4684	0.6549	0.6850	0.6678	0.1927	0.6944	0.3689	0.6225
16	0.3677	0.4550	0.8006	0.5356	0.6490	0.2293	0.6406	0.5742	0.5906
17	0.3706	0.6784	0.8045	0.6031	0.7879	0.2126	0.7189	0.6172	0.6203
18	0.4199	0.6307	0.7881	0.6698	0.7799	0.2213	0.7805	0.5552	0.7698

3. AUTENTICACIÓN UTILIZANDO LOS COEFICIENTES DE CORRELACIÓN

Es útil mencionar que con el uso de las correlaciones existentes entre las imágenes es posible comparar dos rostros. Podemos establecer un límite al

Tabla 3.14 (cont.) ...valores restantes.

	10	11	12	13	14	15	16	17	18
1	0.4694	0.0020	0.3738	0.4237	0.2060	0.5886	0.3677	0.3706	0.4199
2	0.4844	0.5207	0.5313	0.3841	0.6775	0.4684	0.4550	0.6784	0.6307
3	0.6539	0.4812	0.7465	0.6885	0.5891	0.6549	0.8006	0.8045	0.7881
4	0.6282	0.2301	0.5301	0.6236	0.4438	0.6850	0.5356	0.6031	0.6698
5	0.6577	0.4157	0.7053	0.6341	0.7139	0.6678	0.6490	0.7879	0.7799
6	0.1937	0.0186	0.2377	0.2218	0.2749	0.1927	0.2293	0.2126	0.2213
7	0.7552	0.2514	0.6869	0.6431	0.5014	0.6944	0.6406	0.7189	0.7805
8	0.4680	0.5600	0.4844	0.4146	0.4696	0.3689	0.5742	0.6172	0.5552
9	0.8912	0.2892	0.5713	0.5213	0.5021	0.6225	0.5906	0.6203	0.7698
10	1.0000	0.2870	0.5620	0.5116	0.4690	0.5695	0.5622	0.6264	0.7772
11	0.2870	1.0000	0.4312	0.3405	0.5995	0.2402	0.4553	0.4922	0.4547
12	0.5620	0.4312	1.0000	0.5900	0.5915	0.6284	0.7241	0.7603	0.6952
13	0.5116	0.3405	0.5900	1.0000	0.4374	0.6318	0.6602	0.6322	0.6394
14	0.4690	0.5995	0.5915	0.4374	1.0000	0.3707	0.5835	0.6908	0.6459
15	0.5695	0.2402	0.6284	0.6318	0.3707	1.0000	0.5459	0.6181	0.6311
16	0.5622	0.4553	0.7241	0.6602	0.5835	0.5459	1.0000	0.7328	0.6510
17	0.6264	0.4922	0.7603	0.6322	0.6908	0.6181	0.7328	1.0000	0.7934
18	0.7772	0.4547	0.6952	0.6394	0.6459	0.6311	0.6510	0.7934	1.0000

valor de la correlación, por debajo del cual podemos decir que dos imágenes no tratan del mismo rostro, si la correlación entre dos imágenes es mayor a este limite podríamos decir lo contrario, pero no siempre con certeza, por lo ya discutido anteriormente.

De hecho las pruebas lo demuestran así, un programa sencillo que calcula la correlación de cada imagen de solicitud con cada una de las imágenes del conjunto de entrenamiento no siempre acertó en los casos que se trataba del mismo rostro en condiciones distintas, como las ya mencionadas. Tales pruebas demuestran que algunos rostros conservan alta correlación aún cuando estos aparecen en distintas situaciones en diferentes fotografías (algunos rostros no) y muestran correlación mas alta con respecto a otro rostro que consigo mismo.

4. CORRELACIONES POR PIXEL

Se observan fenómenos interesantes cuando se analizan las correlaciones por pixel. ¹ De esta manera no siempre se pueden observar muy altas correlaciones aun cuando se tengan imágenes que pertenecen a un mismo sujeto

¹Se definen un conjunto de variables aleatorias cuyas muestras por pixel son tomadas a partir de un conjunto de imágenes

ni correlaciones muy bajas cuando en el conjunto de imágenes de un mismo sujeto se encuentran fotografías que presentan el rostro con alguna rotación.

Tomando las correlaciones de esta forma se puede tener una de las siguientes situaciones:

1. Imágenes de un mismo sujeto, manteniendo el mismo ángulo fotográfico, se encuentran correlacionadas de manera aceptable, por lo menos para variables que representan píxeles vecinos.
2. Imágenes de un mismo sujeto, manteniendo el mismo ángulo fotográfico, se encuentran correlacionadas de manera inaceptable, al menos para variables que representan píxeles vecinos.
3. Imágenes de un mismo sujeto, no conservando el mismo ángulo fotográfico, se encuentran correlacionadas de manera aceptable, al menos para variables que representan píxeles vecinos.
4. Imágenes de un mismo sujeto, no conservando el mismo ángulo fotográfico, se encuentran correlacionadas de manera inaceptable, por lo menos para variables que representan píxeles vecinos.
5. Imágenes de diferentes sujetos, con diferentes gestos y ángulos fotográficos, se encuentran correlacionadas de manera aceptable, por lo menos para variables que representan píxeles vecinos.
6. Imágenes de diferentes sujetos, con diferentes gestos y ángulos fotográficos se encuentran correlacionadas de manera inaceptable, por lo menos para variables que representan píxeles vecinos.

Las situación 3 es completamente equivalente con la situación 5, mientras que la situación 4 es equivalente con la situación 6, por lo que se puede considerar solamente las primeras cuatro situaciones sin pérdida de generalidad.

Que dos variables aleatorias se encuentren aceptablemente correlacionadas quiere decir que el valor absoluto del coeficiente de correlación lineal, ρ , cumple la relación $1 - \epsilon < |\rho| < 1$, para un valor $0 < \epsilon < 1$ previamente establecido, el cual se toma de tal forma que $1 - \epsilon \approx 1$.

Dada una imagen de prueba \vec{t} , afirmamos que para garantizar una autenticación no espuria de esta imagen en el espacio \mathbb{V} , con base \mathfrak{B}' , es necesario construir una matriz de covarianza, $[\Phi]_{\mathfrak{B}}$, tal que las muestras que la definen cumplan con la situación 1 en una gran proporción y, en menos medida, con la situación 2.

No es difícil ver el origen de la aceptabilidad e inaceptabilidad de la correlación en las situaciones 3 y 4, y por tanto de las situaciones 5 y 6, respectivamente. En estos casos, los cambios bruscos en intensidad de píxeles, debido a diferentes gestos y ángulos fotográficos, marcan la diferencia.

La situación 3 es una bastante engañosa, en la que, debido a un cambio en intensidad, un par de variables no correlacionadas inicialmente pueden llegar a ser correlacionadas, aun cuando en el conjunto de imágenes de entrenamiento estén considerados sujetos diferentes. Es decir, se crea correlación pero a partir de muestras no representativas, lo cual es una situación indeseable que puede conducir a una autenticación espuria.

Asimismo, la sola presencia de condiciones similares a la situación en 3, en lo que a imágenes, o ángulos fotográficos diferentes, de sujetos diferentes se refiere, hace deseable el construir de esta forma una matriz de covarianza. Figura 3.12 muestra un conjunto de seis imágenes de un mismo sujeto bajo el mismo ángulo fotográfico, pero con diferentes gestos.

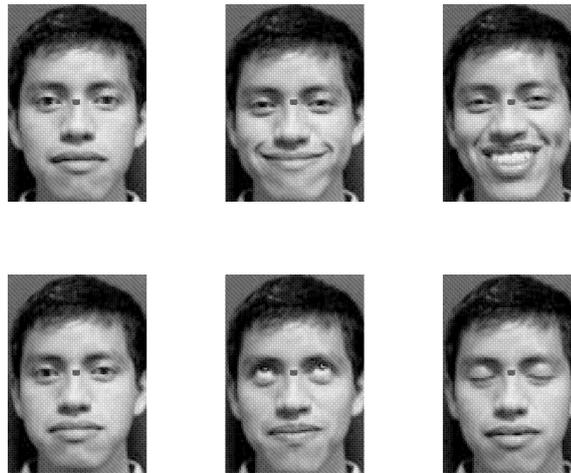


Figura 3.12 Conjunto de imágenes para definir las muestras de las variables aleatorias

En cada una de las seis imágenes, se muestra una región central sombreada que indica el lugar del cual se extraen las muestras de nueve variables aleatorias con el fin de ejemplificar lo dicho líneas arriba. Las muestras son dadas por tabla 3.15, y a las cuales se les obtiene la correlación que tabla 3.16 señala. Finalmente, figura 3.13 ilustra las situaciones 1 y 2, tomando en cuenta las muestras 1, 2 y 4 de tabla 3.15.

Tabla 3.15 Muestras de nueve variables

Pixel	Intensidad					
(56,39)	185	204	207	216	209	198
(56,40)	208	203	189	210	216	207
(56,41)	218	183	180	199	203	208
(57,39)	186	221	211	219	211	209
(57,40)	212	209	191	217	215	218
(57,41)	218	185	185	197	208	212
(58,39)	183	221	209	211	210	210
(58,40)	205	209	193	209	214	221
(58,41)	215	194	186	199	206	217

Tabla 3.16 Correlación entre nueve variables en base a las muestras de la tabla 3.15

	1	2	3	4	5	6	7	8	9
1	1.0000	0.0010	-0.5806	0.8627	-0.0603	-0.6007	0.7484	-0.0429	-0.6127
2	0.0010	1.0000	0.6819	-0.0955	0.9097	0.6520	-0.1273	0.7575	0.6635
3	-0.5806	0.6819	1.0000	-0.7188	0.6911	0.9808	-0.7274	0.5161	0.9373
4	0.8627	-0.0955	-0.7188	1.0000	-0.0271	-0.7303	0.9682	0.1222	-0.5966
5	-0.0603	0.9097	0.6911	-0.0271	1.0000	0.6376	-0.0527	0.8887	0.7487
6	-0.6007	0.6520	0.9808	-0.7303	0.6376	1.0000	-0.7001	0.5383	0.9522
7	0.7484	-0.1273	-0.7274	0.9682	-0.0527	-0.7001	1.0000	0.1945	-0.5351
8	-0.0429	0.7575	0.5161	0.1222	0.8887	0.5383	0.1945	1.0000	0.7153
9	-0.6127	0.6635	0.9373	-0.5966	0.7487	0.9522	-0.5351	0.7153	1.0000

En ella se puede ver que el punto (185,186) es el que marca la diferencia en la alta correlación entre la variable correspondiente a la primera fila y la correspondiente a la cuarta fila. Podemos decir que este punto linealiza al conjunto de puntos correspondiente a la correlación entre la variable de la primera fila y la segunda fila.

Una ligera modificación al conjunto de rostros dado por figura 3.12 es mostrado por figura 3.14, en la que la tercera imagen de la segunda fila de figura 3.12 se ha sustituido por la imagen de la misma persona con un ángulo fotográfico diferente. En este caso, la región de muestreo contiene pixeles con intensidad más oscura, tal y como muestra tabla 3.17 en su columna del extremo derecho. Este hecho es la razón por la cual se producen correlaciones altas, según lo indica tabla 3.18. La alta correlación entre las variables correspondientes a la primera y segunda fila es debido al punto (111,143), tal y como lo ilustra figura 3.15.

La forma de disminuir este tipo de problemas es garantizando de inicio que, dentro de un conjunto de ángulos fotográficos diferentes, cada ángulo cuenta

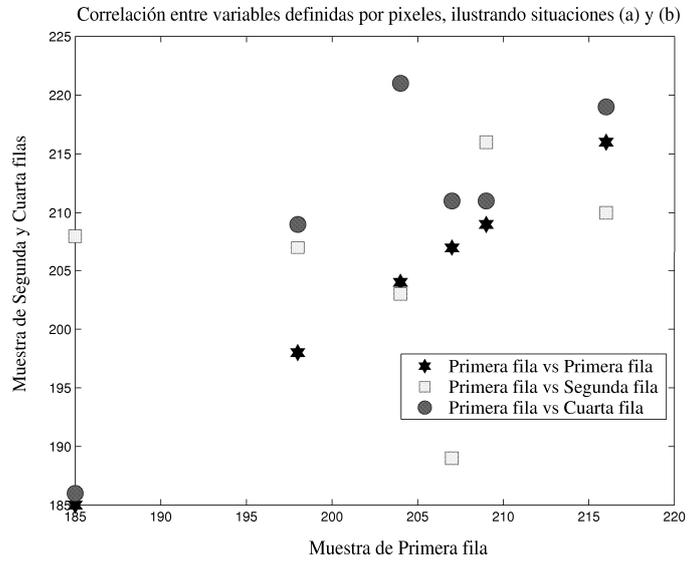


Figura 3.13 Nube de correlación entre variables indicadas en tabla 3.15

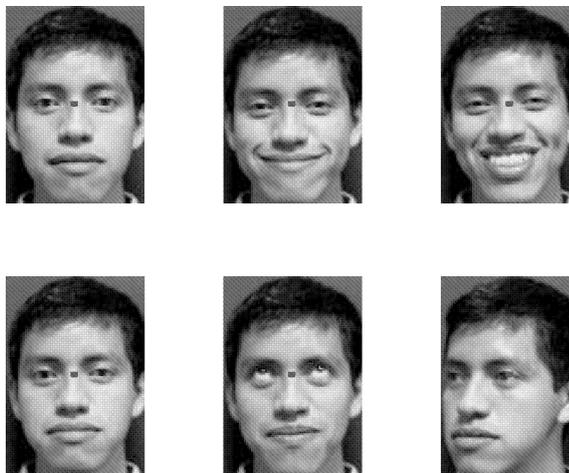


Figura 3.14 Conjunto modificado de imágenes para definir nuevas muestras de las variables aleatorias

con un conjunto de imágenes de un mismo sujeto con diferentes gestos. El problema es que el contar con una sola imagen dificulta el análisis.

Tabla 3.17 Muestras de nueve variables del conjunto de entrenamiento modificado

Pixel	Intensidad					
(56,39)	185	204	207	216	209	111
(56,40)	208	203	189	210	216	143
(56,41)	218	183	180	199	203	167
(57,39)	186	221	211	219	211	30
(57,40)	212	209	191	217	215	80
(57,41)	218	185	185	197	208	81
(58,39)	183	221	209	211	210	53
(58,40)	205	209	193	209	214	120
(58,41)	215	194	186	199	206	109

Tabla 3.18 Correlaciones de nueve variables en base a muestras en tabla 3.17

	1	2	3	4	5	6	7	8	9
1	1.0000	0.9099	0.5242	0.9901	0.9514	0.8926	0.9873	0.9535	0.8912
2	0.9099	1.0000	0.7946	0.9226	0.9829	0.9682	0.9136	0.9851	0.9752
3	0.5242	0.7946	1.0000	0.5518	0.7318	0.8264	0.5232	0.7075	0.8235
4	0.9901	0.9226	0.5518	1.0000	0.9701	0.9176	0.9987	0.9722	0.9213
5	0.9514	0.9829	0.7318	0.9701	1.0000	0.9769	0.9619	0.9969	0.9829
6	0.8926	0.9682	0.8264	0.9176	0.9769	1.0000	0.9050	0.9674	0.9977
7	0.9873	0.9136	0.5232	0.9987	0.9619	0.9050	1.0000	0.9675	0.9098
8	0.9535	0.9851	0.7075	0.9722	0.9969	0.9674	0.9675	1.0000	0.9750
9	0.8912	0.9752	0.8235	0.9213	0.9829	0.9977	0.9098	0.9750	1.0000

La observación realizada al final del párrafo previo está relacionada con el hecho de que no se puede llevar a cabo una estadística de un solo punto, y refleja una posible restricción en la forma clásica de aplicar PCA. Sugerimos que es posible solucionar esto por medio de una sola imagen.

5. CONCLUSIÓN

La correlación que hemos mencionado en este apartado, ha sido la correlación lineal, el porque de utilizar dicha correlación en lugar de otro tipo de correlación ha sido respondido de manera implícita al mostrar las nubes de puntos entre dos imágenes (del mismo y de diferentes rostros). Hemos visto cómo el cambio de correlación se ve afectado al restar la imagen media de un conjunto, mas cuando este conjunto es para un mismo rostro los cambios en los coeficientes de correlación son más proporcionales o predecibles, o cuando se resta la media sin interferir unas imágenes con otras, como por ejemplo, cuando se resta la media por fila o columna en una sola imagen.

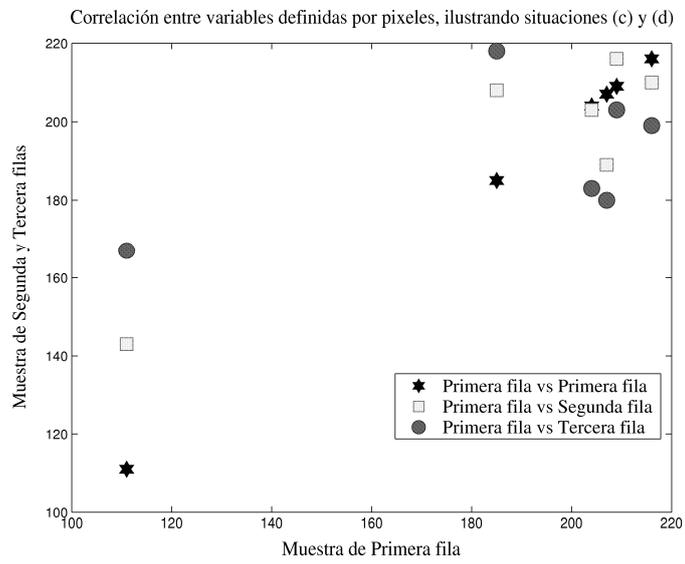


Figura 3.15 Nube de correlación entre variables indicadas en tabla 3.17

Capítulo 4

PCA EN UNA SOLA IMAGEN

Si te aprovechas de las lecciones de la sabiduría vivirás en todas partes sin disgusto y serás feliz en tu estado; la riqueza te dará placer porque tendrás mayores medios de hacer bien a muchos; la pobreza, porque te hallarás con menos inquietudes y sobresaltos; la gloria porque te verás honrado; la oscuridad porque serás menos envidiado.

—Plutarco

Introducción

Las ventajas de utilizar una imagen por persona son indispensables para cualquier sistema biométrico de autenticación de alto rendimiento, ello reduce la carga de trabajo y disminuye el tiempo total de procesamiento. Se proponen en este apartado formas diferentes de realizar PCA sin tener que utilizar más de una sola imagen y sin afectar de manera significativa las correlaciones naturales entre las imágenes de los rostros. En los PCA's que veremos a continuación no será necesario utilizar el Truco de Turk y Pentland que se utiliza en PCA clásico [8] o PCA por módulos o modificado ya que la cardinalidad del número de variables se ha reducido de ser igual a la resolución de las imágenes— $m \times n$ —a ser igual al número de filas— m —, de columnas— n —o del total de imágenes de entrenamiento— k —, donde por mucho $m, n, k \ll m \times n$.

1. PCA POR FILAS

Un punto importante a considerar cuando se toma cada fila como una muestra diferente, conjunto de muestras al que se tiene que realizar el análisis PCA, es que se están definiendo tantas variables como filas tenga la imagen y que la muestra de esta variable tiene entonces una cardinalidad igual al número de

columnas de la imagen. Desde este punto de vista, la matriz de imágenes se sustituye ahora por la matriz imagen de una sola persona y, por consiguiente, la matriz de covarianza es ahora dada por el producto de esta matriz y su transpuesta. Esto da una matriz simétrica de orden igual al número de filas de la matriz imagen.

Si se desea realizar el producto transpuesta de matriz imagen por matriz imagen, entonces se obtiene una matriz simétrica de orden igual al número de columnas de la matriz imagen. Nótese que en los dos casos anteriores se asume que a cada valor de pixel se le ha restado la media de la fila correspondiente.

Desde esta perspectiva, dada una matriz imagen \mathcal{I} de tamaño $m \times n$, la fila \mathcal{X}_i se toma ahora como una muestra de la variable aleatoria X_i correspondiente a esa fila. En este sentido, cada columna \mathcal{Y}_i de la matriz imagen representa un valor posible de la variable aleatoria $\mathbf{X} = (X_1, X_2, \dots, X_m)$.

Puesto que PCA trabaja ahora sobre la imagen de una sola persona, y no sobre un conjunto de imágenes de varias personas, la misma matriz imagen sirve para definir la matriz de covarianza, quien ahora medirá la correlación existente entre las filas de la matriz imagen \mathcal{I} .¹ Por consiguiente, la media de la fila i de la matriz imagen es simplemente dada por la expresión

$$\mu_i = \frac{1}{n} \mathcal{X}_i \mathbf{1}^T.$$

Se define entonces un vector media

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_m)$$

con el cual se define también la matriz de covarianza por fila de la siguiente forma

$$\Phi = (\mathcal{I} - \Upsilon)(\mathcal{I} - \Upsilon)^T$$

siendo $\Upsilon = \boldsymbol{\mu}^T \mathbf{1}$ una matriz $m \times n$, donde $\mathbf{1}$ es un vector fila cuyas n entradas tienen un valor unitario. Así, Φ es una matriz de orden m , por lo que el número de eigenvalores y eigenvectores es también igual a m .

Esto significa que, una vez conocidos los eigenvectores, se construye una matriz \mathfrak{B} , de orden m , cuyas columnas son precisamente los eigenvectores

¹Según pruebas en imágenes de rostros que ocupan toda el area de la imagen, la correlación es más alta mientras más cercanas se encuentren las filas o columnas

encontrados para la matriz de covarianza Φ . Este conjunto de eigenvectores forma entonces una base ordenada $\beta' = \{\vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_m\}$ para el espacio de las nuevas variables cuya correlación es nula. Haciendo una analogía con PCA original, donde a los eigenvectores de la matriz de covarianzas a menudo se les conoce como eigenrostros (por tener un número de componentes igual al número de píxeles de las imágenes y a que al ser graficados cada uno muestra la forma de un rostro), en este PCA podríamos llamarles a los eigenvectores de la matriz de covarianzas, eigencolumnas, ya que el número de componentes de cada uno es igual al de las columnas, siguiendo un razonamiento similar, en PCA por columnas a los eigenvectores los podríamos llamar eigenfilas.

Se puede demostrar además que estos eigenvectores son ortogonales. Es un resultado bastante conocido en el álgebra lineal que si $\vec{\beta}_i$ y $\vec{\beta}_j$ son eigenvectores de una matriz simétrica Φ , con eigenvalores λ_i y λ_j diferentes, entonces $\vec{\beta}_i \vec{\beta}_j^T = 0$; es decir, los eigenvectores son ortogonales. La normalización de los vectores preserva la pertenencia al conjunto de eigenvectores, así que se puede hablar de una base β' ortonormal. Además, se cumple que $\mathfrak{P}\mathfrak{P}^T$ es igual a la matriz identidad de orden m , por lo que $\mathfrak{P}^{-1} = \mathfrak{P}^T$.

La matriz \mathfrak{P} define simplemente un cambio de variables. En este sentido, las variables originales $\mathbf{X} = (X_1, X_2, \dots, X_m)$ se cambian ahora a las variables $\mathbf{Z} = (Z_1, Z_2, \dots, Z_m)$ vistas desde la nueva base β' . Esta transformación es dada por la expresión

$$[\mathbf{Z}]_{\beta'} = \mathfrak{P}^T [\mathbf{X}]_{\beta}$$

donde, en el caso que nos ocupa, β es la base canónica. Como puede observarse de esta expresión, el valor medio de \mathbf{Z} es igual a cero. Considerando las componentes Z_i y Z_j , entonces

$$Z_i = \vec{\beta}_i [\mathbf{X}]_{\beta}$$

y

$$Z_j = \vec{\beta}_j [\mathbf{X}]_{\beta}$$

cuyas muestras se obtienen por medio de los vectores muestra de la variable \mathbf{X} . Si \mathfrak{X}_i es la muestra de la variable X_i , entonces las muestras de Z_i y Z_j son dadas por las expresiones

$$\mathfrak{z}_i = \sum_{k=1}^m \beta_{ik} \mathfrak{x}_k$$

y

$$\mathfrak{z}_j = \sum_{l=1}^m \beta_{jl} \mathfrak{x}_l.$$

Por consiguiente, se encuentra que se satisface la expresión de covarianza

$$\mathfrak{z}_i \mathfrak{z}_j^T = \vec{\beta}_i \Phi \vec{\beta}_j^T$$

o, en general,

$$\mathfrak{z} \mathfrak{z}^T = \mathfrak{P}^T \Phi \mathfrak{P}$$

la cual es precisamente la matriz diagonal cuyos elementos en la diagonal son los valores propios de Φ . Es la transformación Φ vista desde la nueva base β . Puesto que los elementos fuera de la diagonal son nulos, esto significa que las nuevas variables no se encuentran correlacionadas, no al menos con correlación lineal.

Bajo el mapeo $\mathfrak{P}^T [\vec{\beta}]_{\beta}$, se obtiene un punto $[\vec{\beta}]_{\beta'}$ en el nuevo espacio generado por β' , cuyos vectores base $\{\vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_m\}$ definen cada uno de ellos variables (Y_1, Y_2, \dots, Y_m) en direcciones de máxima varianza en el espacio original, lo que permite definir proyecciones $[\vec{\beta}]_{\beta'}$, cuya normalización es útil para definir un conjunto de pesos que describen la contribución de cada elemento de la base β' .

Así pues, dada una imagen de entrenamiento, las columnas de ella representan valores posibles de la variable aleatoria \mathbf{X} . Una vez que esta variable aleatoria se centra con media cero, la variable obtenida es mapeada al nuevo espacio con el fin de obtener sus correspondientes proyecciones. Para cada columna se tendrá que definir proyecciones de esta forma. Se conviene en que el conjunto de proyecciones máximas por cada columna será el peso final empleado en la autenticación, para esa imagen de entrenamiento. Nótese que hasta aquí no hemos empleado la idea de reducción de dimensionalidad.

2. PCA POR COLUMNAS

Se realiza un proceso completamente similar cuando se conviene considerar muestras por columna en lugar de por fila, en este caso se define un número de variables aleatorias igual al número de columnas y el tamaño de la muestra es igual al número de filas de la matriz imagen. En este caso, a cada valor de pixel se le resta la media de la columna correspondiente.

3. PCA MIXTO

En este tipo de PCA utilizamos tanto PCA por filas como por columnas para llevar a cabo la autenticación; cada PCA se realiza por separado, la diferencia esta en la forma de comparar los coeficientes obtenidos para cada imagen por cada PCA. Es menester decir que el consumo de tiempo será del doble que en un PCA de filas o columnas simple, esto por causas obvias.

4. COMPARANDO ROSTROS EN PCA FILAS O COLUMNAS

Es necesario mencionar que en el espacio generado con filas o columnas es un espacio en donde no existe de forma entera una imagen si no que ésta la tenemos por partes en sus filas o columnas, por lo que la comparación es realizada a este nivel.

Se llevaron a cabo dos tipos de prueba, la primera de ellas consistió en obtener para las imágenes de generalización su propio espacio de forma similar a como se realiza con las imágenes de entrenamiento, restando su propia media por fila o columna según fuera el caso, posteriormente, las imágenes no se pasaron a los espacios creados por ellas mismas y no se realizó la comparación entre las imágenes en su nuevo espacio ya que la comparación no se haría en el mismo espacio para todas las imágenes; por lo anterior, se probó cotejar los espacios generados por las imágenes, ya que a dos imágenes iguales, dos espacios igualmente generados, los resultados no fueron muy favorables, por lo que se optó por la segunda forma, en ella, las imágenes de prueba se pasaban al espacio de la imagen de entrenamiento con la cual se quería confrontar, por tanto, primero se debería sustraer la media fila o columna de la imagen de entrenamiento en cuestión a la imagen de generalización, esto, por supuesto, refuerza la idea de la interferencia en las correlaciones, ya que, cuando la media fila o columna de la imagen de entrenamiento que se restará a la imagen de prueba es del mismo rostro no se interferirá con las correlaciones por tratarse

del mismo sujeto y la confrontación irá bien, de lo contrario, las correlaciones se verán afectadas y la comparación irá mal, reforzando así, la autenticación, note que de igual manera sucede con PCA modificado. En esta segunda forma los resultados fueron mucho mejores, por lo que es el caso que vamos a ver a continuación.

Dejemos que $\mathbf{Z}_k = [Z_{k,1}, Z_{k,i} \dots Z_{k,j}, Z_{k,p}]$ sea nuestro nuevo vector aleatorio (nuestra k -ésima imagen en el nuevo espacio generado por ella misma) y que $\mathcal{Z}_{k,i}$ es la muestra de cada variable aleatoria $Z_{k,i}$, y por lo tanto su nueva fila o columna en ese espacio. Para una imagen de prueba dada, $\mathbf{Z}_{test_k} = [Z_{test_k,1}, Z_{test_k,i} \dots Z_{test_k,j}, Z_{test_k,p}]$ es el vector aleatorio que la representa en el k -ésimo espacio generado por la k -ésima imagen de entrenamiento, espacio en el que se llevará a cabo la comparación; de forma análoga a una imagen de entrenamiento $\mathcal{Z}_{test_k,i}$ representa la muestra de cada variable aleatoria $Z_{test_k,i}$.

La comparación imagen a imagen la podemos realizar con cualesquiera de las medidas de similitud. Supongamos que estamos tratando las filas de las imágenes, el resultado de comparar una imagen con otra es un vector columna, donde cada una de sus componentes nos dice el parecido entre la fila de una imagen y la correspondiente fila de la otra imagen. Supongamos que utilizamos la medida de cosenos para comparar una imagen de prueba con la k -ésima imagen de entrenamiento, la medida de similitud entre sus i -ésimas filas es:

$$c_{i,k} = \frac{\mathcal{Z}_{i,k} \cdot \mathcal{Z}_{test_{i,k}}}{\|\mathcal{Z}_{i,k}\| \|\mathcal{Z}_{test_{i,k}}\|}$$

Si realizamos la comparación de todas las filas obtendremos el vector $\vec{\mathbf{c}}_k$ de tamaño $m \times 1$ si conservamos todos los eigenvectores o de tamaño $p \times 1$, donde $p < m$, si se conservaron sólo los p eigenvectores principales; ahora, como el vector $\vec{\mathbf{c}}_k$ nos muestra el parecido entre las filas de las imágenes, lo que podemos hacer es promediar las componentes de ese vector y obtener así un número que representa el parecido entre las imágenes correspondientes.

Si el número de imágenes de entrenamiento es k , entonces para cada imagen de prueba del sistema habrá k vectores $\vec{\mathbf{c}}$ de similitud, si promediamos el parecido entre filas para obtener el parecido general entonces para cada imagen de prueba tendremos k coeficientes, como en este caso estamos utilizando medida de cosenos el coeficiente más cercano a 1 pertenecerá a la comparación donde la imagen de entrenamiento se parece más a la imagen de prueba.

Cuando se trata de confrontación de imágenes columna a columna, el procedimiento es exactamente el mismo salvo por las dimensiones que son poco diferentes. Por ejemplo, la dimensión del vector \vec{c} es de $1 \times p$, para $p \leq n$, dependiendo del número de eigenvectores principales conservados.

Resumiendo, cada imagen de entrenamiento genera un eigenespacio con sus filas o columnas para trasladarse ahí; para comparar una imagen de prueba con una del entrenamiento deberemos primero trasladar la de prueba al espacio donde se encuentra la del entrenamiento y realizar ahí la comparación.

5. PCA CLÁSICO EN OTRA ARQUITECTURA

Ahora veremos la forma de hacer PCA sin meter ruido a las imágenes a través de la imagen media, ahora lo que haremos será restar la media propia a cada imagen y obtener la matriz de covarianza de forma diferente, le llamaremos PCA de media propia en este documento. Recuerde que asumimos que \mathbf{X} era nuestro vector aleatorio (ver página 14) con mn componentes o variables aleatorias y donde el tamaño de la muestra de cada variable aleatoria era k . Anteriormente La dimensionalidad de nuestro espacio original \mathbb{R} era de mn , en esta otra arquitectura la dimensionalidad es menor ya que corresponde a una de tamaño k (número de imágenes). Entonces se asume que ahora $\mathbf{X} = (X_1, X_2, \dots, X_i, \dots, X_j, \dots, X_k)$ y que cada componente de este vector es una variable aleatoria; una imagen ahora será considerada como una variable aleatoria. Los pasos en esta otra forma de hacer PCA son los mismos de PCA clásico, lo único que ha cambiado es la dimensionalidad, ahora la n -ésima entrada de las imágenes es un punto en el espacio y cada imagen contiene las coordenadas correspondientes a un eje de todos los puntos. De igual forma que calculamos la covarianza entre las dos variables X_i y X_j en PCA clásico calculamos la covarianza en esta otra arquitectura; note que solo cambia el término del denominador (vea la ec. 2.5):

$$Cov(X_i, X_j) = \frac{1}{mn - 1} (\mathfrak{X}_i - \boldsymbol{\mu}_i)(\mathfrak{X}_j - \boldsymbol{\mu}_j)^T \quad (4.1)$$

donde \mathfrak{X}_i es el vector fila que contiene los mn valores de la muestra de la variable aleatoria X_i , valores que son dados por todas las intensidades de pixel asociadas a una sola imagen. En este contexto $\boldsymbol{\mu}_i$ es un vector cuyas entradas son todas iguales a la media de una imagen y no la media de una entrada (o pixel) específica de las imágenes:

$$\boldsymbol{\mu}_i = [\mu_i, \mu_i, \dots, \mu_i]_{1 \times mn} \quad (4.2)$$

donde

$$\mu_i = \frac{1}{mn} \boldsymbol{x}_i \times \mathbf{1}^T \quad (4.3)$$

$\mathbf{1}$ es un vector de tamaño $1 \times mn$ cuyas entradas son todas iguales a uno. Para formar una matriz que contenga todas las covarianzas existentes entre las variables aleatorias, formamos primero la matriz de imágenes (o variables) con la ecuación:

$$\mathcal{P} = [\boldsymbol{x}_1^T, \boldsymbol{x}_2^T, \dots, \boldsymbol{x}_k^T] \quad (4.4)$$

(ver ec. 2.7 en la pág. 14) y de forma análoga construimos la matriz de valores medio Υ , formada con los vectores $\boldsymbol{\mu}_i$:

$$\Upsilon = [\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T, \dots, \boldsymbol{\mu}_k^T] \quad (4.5)$$

ecuación parecida a la ec. 2.8. Finalmente obtenemos nuestra matriz de covarianzas, que como ya sabemos es simétrica (note que $Cov(X_i, X_j) = Cov(X_j, X_i)$), cuadrada y de orden k :

$$\Phi = (\mathcal{P} - \Upsilon)^T (\mathcal{P} - \Upsilon)$$

De esta manera el tamaño de la matriz de covarianzas será de $k \times k$. A partir de Φ obtendremos los vectores (columna) propios que formarán la matriz de cambio de base \mathfrak{P} , cuyo tamaño será igualmente de $k \times k$ si se toman en cuenta todos los eigenvalores, o de tamaño $k \times p$, $p < k$, si solo se toman los p principales vectores propios.

Lo que nos resta responder es ¿Cómo identificar en el nuevo espacio generado con esta arquitectura?. Sabemos que en PCA clásico comparábamos directamente las imágenes en su nuevo espacio porque cada una de ellas era un punto y utilizábamos sus coordenadas para compararlas, pero ahora los puntos representan a los pixeles y no tiene sentido comparar las coordenadas del i -ésimo pixel con las coordenadas del j -ésimo pixel en su nuevo espacio porque no requerimos una medida de similitud entre dos pixeles sino entre dos imágenes.

Primero definamos un conjunto de imágenes que no estén correlacionadas, es decir, que su matriz de covarianzas sea una matriz diagonal. Para ello aplicamos la matriz de transformación \mathfrak{P} a la matriz de imágenes \mathcal{P} :

$$\mathcal{S} = \mathcal{P}\mathfrak{P} \quad (4.6)$$

Tales imágenes en las columnas de \mathcal{S} serán utilizadas como imágenes base. Si medimos las correlaciones entre ellas veremos que serán nulas. Esta forma fue inspirada por el método ICA, ya que en tal método, se utilizan imágenes linealmente independientes como imágenes base [5], tales imágenes son obtenidas a partir de un conjunto de imágenes conocidas como imágenes de entrenamiento, posteriormente, estas imágenes y las del conjunto de prueba son expresadas en términos de ellas y luego comparadas. Sería punto y aparte indagar sobre que imágenes base son mejores, las que son linealmente independientes o las que son no correlacionadas.

Aunque si bien es cierto que las variables independientes no están correlacionadas y que las variables no correlacionadas no son necesariamente independientes. Y debido a que la independencia si implica no correlación, muchos métodos ICA coaccionan el procedimiento de estimación de tal forma que este siempre da estimaciones no correlacionadas de las componentes independientes (imágenes base en nuestro caso).

Cabe mencionar que la ecuación 4.6 utilizada para obtener las imágenes base es similar a aquella ecuación (ver ec. 2.21) utilizada para obtener los eigenrostros con el truco de Turk y Pentland.

Las imágenes base son imágenes adecuadas para obtener otras imágenes a partir de ellas, esto significa que cualquier imagen la podemos expresar como una combinación lineal de las imágenes base:

$$\begin{aligned} \mathcal{P}_i &= a_{1,i}\mathcal{S}_1 + \dots + a_{p,i}\mathcal{S}_p \\ \mathcal{P} &= \mathcal{S}A \end{aligned} \quad (4.7)$$

donde i denota el número de columna en las matrices y p el número de imágenes base. Así cada rostro tendrá su conjunto de coeficientes y será único para el, pero mientras más parecidos sean dos rostros más similares serán sus coeficientes. Obviamente dos fotografías del mismo rostro tendrán coeficientes cercanamente parecidos (bajo ciertas circunstancias, mencionadas anteriormente). Es obvio que para comparar dos imágenes basta con comparar sus coeficientes con al-

guna medida de similaridad, y en base a estas medidas y otros criterios decidir si dos imágenes pertenecen al rostro del mismo sujeto.

Por lo tanto, la matriz A que contiene en sus columnas los coeficientes correspondientes a cada imagen es obtenida despejando \mathcal{P} de la ecuación 4.6:

$$\begin{aligned} \mathcal{S}\mathfrak{P}^{-1} &= \mathcal{P}\mathfrak{P}\mathfrak{P}^{-1} \\ \mathcal{S}\mathfrak{P}^{-1} &= \mathcal{P} \end{aligned} \quad (4.8)$$

y así obtenemos a cada una de nuestra imágenes en \mathcal{P} en términos de las imágenes base en \mathcal{U} , si comparamos esta ecuación con la ecuación 4.7, notaremos que \mathfrak{P}^{-1} es nuestra matriz A de nuestras imágenes de entrenamiento en \mathcal{P} . Para obtener los coeficientes de las imágenes de prueba consideremos la siguiente ecuación:

$$\mathcal{P}_{prueba} = \mathcal{S}A_{prueba} \quad (4.9)$$

donde ponemos a las imágenes de prueba en términos de las imágenes base obtenidas con las imágenes de entrenamiento, A_{prueba} contiene en sus columnas los coeficientes correspondientes a cada imagen de prueba, coeficientes que dictan la combinación lineal necesaria para obtener una imagen. Despejando en la ecuación anterior tenemos:

$$\mathcal{S}^{-1}\mathcal{P}_{prueba} = \mathcal{S}^{-1}\mathcal{S}A_{prueba} \quad (4.10)$$

$$\mathcal{S}^{-1}\mathcal{P}_{prueba} = A_{prueba} \quad (4.11)$$

Como ya mencionamos, si la dimensionalidad de nuestro espacio canónico original no es muy grande, podremos omitir el truco de Turk y Pentland [8] y de esta forma si así lo quisiéramos podríamos mantener la dimensionalidad del espacio canónico original; esto en base a que con el truco mencionado obtenemos solamente un número de eigenvectores igual al número de puntos, cuando podría tenerse una cantidad de eigenvectores igual a la cantidad de variables aleatorias, más como son los principales eigenvectores los que se obtienen y nuestra dimensionalidad es muy grande, con ello es suficiente. En la nueva base, las imágenes obtenidas no están correlacionadas, contrario a lo que sucede en la otra arquitectura de PCA donde los pixeles eran los que no se encontraban correlacionados y las imágenes si.

6. EL PORQUE DE PCA POR FILAS Y COLUMNAS

Entre los problemas que se pueden tener con esta forma de emplear PCA, se encuentra incluido el reconocimiento de rostros del mismo sujeto pero girados un cierto ángulo con respecto a un eje que bisecta el plano de una imagen frontal del sujeto. Entre las ventajas de emplear un método así, se encuentra la disminución en el tamaño de matrices de covarianza y de cambio de base. En ambos casos, y asumiendo el caso peor, de un problema de orden mn se pasa a uno de orden m . Adicionalmente, consideramos que el modelo holístico usualmente empleado por otros investigadores tiende a perder detalles que con el método local propuesto se incluyen de manera natural.

La interpretación local es una alternativa a la interpretación holística del reconocimiento de rostro. Generalmente, esta localidad se define en términos de regiones determinadas por puntos importantes del rostro. Los ojos, la boca, y la nariz, son puntos de referencia para definir estas regiones. Sin embargo, se pueden emplear resultados experimentales para sugerir que el reconocimiento considera también recorridos horizontales y (o) verticales del rostro, tomando cada conjunto de valores de pixel por columna y (o) fila como definiendo microregiones.

Estos resultados experimentales muestran que en el recorrido por el rostro, la visión se detiene por breves lapsos de tiempo en regiones de interés (ver fig. 4.1), como las indicadas arriba, que generalmente contienen cambios abruptos de intensidad [10]. Desde un punto de vista matemático formal, y asumiendo que el rostro es una imagen digital, esto se puede interpretar como un mecanismo de diferenciación de columnas y (o) filas. En este trabajo se propone entonces un mecanismo tal tomando como base el análisis PCA de filas o columnas de la imagen digital.

Una identificación más precisa no solamente debe contemplar PCA sobre filas y columnas, sino también sobre conjunto de imágenes de cada sujeto. Esto último se puede emplear como una aproximación gruesa a la identificación, mientras que la combinación de los dos primeros como un refinamiento en el reconocimiento.

7. COMPARACIÓN DE LAS IMPLEMENTACIONES PCA

Para juzgar correctamente los distintos PCA's desde el punto de vista de la optimización del tiempo, primero debemos considerar los intervalos necesarios



Figura 4.1 Barrido ocular humano [10].

que se necesitan para cada una de sus fases, como la del entrenamiento y la de prueba, de esta forma podemos analizar en que procesos unos aventajan a otros y en cuales se rezagan. La optimización total será la suma de los tiempos individuales para cada proceso y de esta forma veremos cual de las modos de PCA es de manera global el mejor optimizador de tiempo. El análisis de la carga de trabajo también es importante ya que el tiempo de los procesos de PCA depende no solo de los procesos que sobre los datos se tienen que hacer, si no también sobre la cantidad de datos sobre los que se tienen que hacer.

Considerando que es un poco difícil medir lo anterior, sólo consideraremos la cantidad de datos manipulados como un criterio para saber que PCA es mejor en el sentido de carga de trabajo. Para nuestro fin, la unidad para medir la cantidad de datos será un double de MATLAB que puede representar, un pixel, un elemento de la matriz de covarianza, una media, etc., en resumen, es un valor numérico.

7.1 CARGA DE TRABAJO

Rastreamos la carga de trabajo total a través de la cantidad de datos que maneja PCA por cada paso, por ejemplo, supongamos que tenemos 50 rostros con 5 fotografías por cada uno, en PCA original o en PCA filas, columnas o PCA de media propia se utiliza solamente una imagen de las 5 que tenemos disponibles, en PCA modificado se utilizan más de una, en este caso supongamos que utilizamos la totalidad de las 5 imágenes por rostro. Consideremos el caso de que cada imagen tiene resolución de 115×80 píxeles. Por lo tanto $m = 115$, $n = 80$, $k = 50$, $q = 5$.

Consideremos la fase de entrenamiento, en PCA original, al restar la imagen media y al calcular la matriz de covarianza utilizamos 9200×50 valores de pixel en total, el resultado es una matriz de covarianza de 50×50 . Como utilizamos el truco de Turk y Pentland, en el proceso de calcular los eigenvectores y los eigenvalores se utiliza esa matriz de covarianza junto con todas las imágenes, el resultado es un conjunto de 50 eigenvectores de tamaño 9200×1 cada uno y 50 eigenvalores, hasta aquí la cantidad de datos no ha variado mucho ya que sólo se le aumentan los 50 datos de los eigenvalores, el total de componentes debidos a los eigenvectores es igual al total de píxeles que estamos manejando, esto es 460,000 datos. Automáticamente, al utilizar el truco ya mencionado hemos conservado únicamente los k primeros eigenvectores y desechado los restantes.

Ahora para PCA por filas, en su fase de entrenamiento, al calcular la matriz de covarianza de cada imagen tenemos 9200×50 píxeles en total, la misma cantidad que en PCA original, para cada imagen resulta una matriz de covarianza de 115×115 , esto multiplicado por el total de imágenes resulta en un total de 661,250 valores, una cantidad mayor que en PCA original. Al momento de calcular los vectores y valores propios de cada imagen, obtenemos para cada una de ellas 115 vectores propios con 115 componentes cada uno y su correspondientes valores propios, esto nos da un total de 661,250 valores que maneja PCA filas hasta este instante del entrenamiento. Como a cada imagen le corresponden 115 valores propios, entonces todos ellos suman un total de $115 \times 50 = 5750$. Todo esto sin haber eliminado aún ningún eigenvector. Hasta este punto nos damos cuenta de que PCA por filas maneja más datos que el PCA original, lo que significa más carga de trabajo.

Si analizamos por columnas a las imágenes, vemos que la matriz de covarianzas ha cambiado a una de tamaño 80×80 por imagen, lo que en total da 320,000 datos, algo mucho menor que en PCA por filas, lo cual resulta obvio ya que el número de columnas es menor que el número de filas. Los eigenvectores

serán 80 por cada imagen, con 80 componentes cada uno, lo que resulta en un total de 320,000 componentes; nuevamente, esto es algo menor que PCA por filas. Como a cada eigenvector se le asocia un eigenvalor, a cada imagen le corresponden 80 eigenvalores, que por 50 imágenes, es un total de 3600 datos más que procesar.

Haciendo cuentas, PCA original, PCA filas y PCA columnas, todos ellos llegan con la misma cantidad de datos (los píxeles de las imágenes), al momento de calcular la media y la matriz de covarianza podemos decir que todos están en iguales condiciones de trabajo, después de calcularla, PCA original se quedó con una matriz de covarianzas de 50×50 , PCA filas con 50 de 115×115 cada una y PCA columnas con 50 de 80×80 cada una. Posteriormente al calcular los vectores y valores propios sobre cada una de estas matrices de covarianza los PCA's se quedaron como lo muestra la tabla 4.1.

Tabla 4.1 Carga de trabajo para cada uno de los PCA's propuestos

	PCA original	PCA filas	PCA columnas
Total de componentes	460,000	661,250	320,000
Total de eigenvalores	50	5750	3600

¿Quién tiene mayor carga de trabajo?, pues PCA filas, en este caso más que PCA columnas por obvias razones, lo contrario sucedería si el número de columnas de las imágenes fuera mayor que el número de filas. Note que la reducción de dimensionalidad ha ayudado mucho a PCA original, lo mismo haría con los demás PCA's.

Ahora, con PCA modificado, por su propia naturaleza, utilizamos—en este ejemplo—5 veces más imágenes que en los demás PCA's. Como se utilizan sólo 5 imágenes por rostro entonces para cada rostro se obtiene una matriz de covarianzas de 5×5 , y luego por 50 imágenes; esto es el doble que PCA original. De igual forma el número de eigenvectores y eigenvalores es 5 veces mayor con respecto a PCA original. Con PCA de media propia hasta el cálculo de los valores y vectores propios difiere en PCA clásico en el tamaño de su eigenvectores ya que sólo tienen 50 componentes cada uno, pero recordemos que maneja también 50 imágenes base del mismo tamaño que las imágenes de entrenamiento.

Finalmente, el total de valores numéricos para las imágenes de entrenamiento que se manejarán en los nuevos espacios generados (excepto en PCA de media propia que utiliza coeficientes que representan una combinación lineal

de imágenes base) para la identificación son como se muestran en al penúltima fila de la tabla 4.2, misma que resume lo discutido aquí.

Tabla 4.2 Carga de trabajo para las implementaciones PCA

	PCA original	PCA filas	PCA columnas	PCA modificado	PCA media propia
a)	$115 \times 80 \times 50$	$115 \times 80 \times 50$	$115 \times 80 \times 50$	$115 \times 80 \times 5 \times 50$	$115 \times 80 \times 50$
b)	50×50	$115 \times 115 \times 50$	$80 \times 80 \times 50$	$5 \times 5 \times 50$	50×50
c)	50×9200	$115 \times 115 \times 50$	$80 \times 80 \times 50$	$5 \times 50 \times 9200$	50×50
d)	50	115×50	80×50	5×50	50
e)	50×50	$115 \times 80 \times 50$	$115 \times 80 \times 50$	$5 \times 5 \times 50$	50×50
f)	—	—	—	—	$115 \times 80 \times 50$

^a Cada celda representa el total de valores numéricos que se procesan para cada caso.

La lista siguiente muestra el significado de cada fila en la tabla 4.2:

- a): Total de datos en las imágenes iniciales.
- b): Datos en las matrices de covarianzas.
- c): Datos en los eigenvalores.
- d): Eigenvalores
- e): Total de datos por las imágenes en el nuevo espacio.
- f): Total de datos debidos a las imágenes base.

La interpretación de la tabla es simple, por ejemplo, para la fila de eigenvalores, las columnas representan el total de componentes que se procesan por cada PCA. Las cantidades que aparecen como factores dicen por sí mismos lo que representan.

Como lo hemos dicho, lo anterior sólo representa carga de trabajo, sería interesante analizar el efecto que el tamaño de los datos tiene sobre el desempeño de los procesos que PCA aplica sobre ellos y si hay procesos específicos para cada PCA, como por ejemplo en PCA de media propia, donde se tienen que calcular las imágenes base, que es un paso análogo pero no igual al de utilizar el truco de Turk y Pentland para calcular los eigenvalores de una matriz de covarianzas muy grande.

8. COMENTARIOS

Una línea de investigación adicional es la que se refiere a la incorporación de datos de entrada al análisis de componentes principales que sean perceptual o cognitivamente más realistas que los valores de pixel mismos, por ejemplo como salidas de filtros de Gabor o wavelets [11].

Dentro de la comunidad que investiga la autenticación de rostros, se ha suscitado incluso una discusión en relación a la escala en la cual opera de manera satisfactoria el análisis de componentes principales. Se afirma que las características asociadas a valores propios altos contienen información relativa a la orientación, además de la especificación de la clase a la cual pertenece. Por otro lado, se dice que para los eigenvalores con valor pequeño se tiene información de identidad del rostro [11].

Una pregunta más general e incluso filosófica, que no tiene que ver necesariamente con el análisis de componentes principales, es la que se refiere al proceso de reconocimiento o autenticación, cuestionándose si este proceso es holístico o parcial [12][13][14]. La mayor parte de los trabajos sobre reconocimiento de rostros emplean las imágenes completas para realizar con ellas el análisis de componentes principales; sin embargo, aunque también existen trabajos relacionados con el análisis del reconocimiento de imágenes a través de particiones de éstas en regiones de píxeles [15], ninguno de ellos realiza un estudio serio acerca de la influencia del tamaño de estas regiones en el proceso de reconocimiento, estudio que, pensamos, arrojaría luz acerca de las escalas en las que opera de manera adecuada el análisis de componentes principales.

Generalmente, las regiones seleccionadas en las imágenes comprenden áreas en las que se encuentran localizados (diríamos hasta ubicados completamente) rasgos importantes en el reconocimiento como, por ejemplo, boca, nariz y ojos, aunque tales regiones no cubren necesariamente todo el rostro. Algunos de los trabajos encontrados en la literatura muestran un funcionamiento exitoso del reconocimiento, aunque queda todavía la duda acerca del tamaño mínimo de las regiones involucradas.

Queda incluso la duda de su buen funcionamiento cuando los sujetos no empleados en el entrenamiento se encuentran con lentes, o su rostro ha sufrido algunas de las modificaciones ya señaladas arriba. Suena razonable esperar que la técnica de dividir una imagen en sub-imágenes tenga éxito aun en estas condiciones; sin embargo, resulta también razonable cuestionarse en este caso acerca de la granularidad de la partición para que esto siga siendo cierto. Resulta interesante conocer también el papel que juegan los valores propios

cuya magnitud se ha tomado como pequeña, cosa sobre la cual se dice muy poco.

Por lo tanto, las preguntas que resultan interesante atacar desde un punto de vista del reconocimiento automático de rostros son las siguientes

1. ¿A qué nivel de granularidad mínimo deben especificarse las regiones de una imagen para que sea aplicable la técnica de análisis de componentes principales?
2. ¿Resulta importante considerar los eigenvalores con valor propio tomado como pequeño, en el caso de granularidades en las que no funcionen adecuadamente los valores propios definidos como grandes?
3. ¿Tienen alguna contribución ambos grupos de eigenvalores cuando la imagen a reconocer presenta alguna modificación en relación con las imágenes empleadas como entrenamiento?

Capítulo 5

IMPLEMENTACIÓN DE LOS ALGORITMOS

Lo malo de hacer sugerencias inteligentes es que uno corre el riesgo de que se le asigne para llevarlas a cabo.

—Anónimo

Introducción

En la computación de nuestros tiempos es imperiosa la implementación y utilización de sistemas en donde las interfaces gráficas de usuario nos permiten interactuar con el mundo de las computadoras de una manera más simple a través del uso de controles gráficos de sencillo funcionamiento que sustituyen la ardua tarea de tener que comunicarse con nuestro programa en modo texto a través de la consola, sin embargo, algunas veces no es tan fácil construir interfaces gráficas. MATLAB es una potente herramienta para implementar y correr algoritmos que requieren de potentes funciones que implementen muchas de las funciones matemáticas y no matemáticas que existen en nuestra días, además nos permite construir interfaces gráficas, a un modo de prototipo, de una manera muy sencilla, aunque tanto los programas como las GUI creadas con las funciones que provee no son independientes y sólo corren con el MATLAB en ejecución.

1. GUI

Para poder probar los métodos PCA originales existentes con los nuevos propuestos, fue necesario construir un pequeño programa de cómputo que no sólo implementará lo algoritmos de los PCA's, si no que también implementará una interfaz gráfica para utilizarlos, realizar entrenamientos de las imágenes ya sea

con uno o con otro algoritmo o realizar alguna prueba con cualquiera de esos algoritmos, agregar imágenes, borrarlas, insertar nuevos sujetos, etc.

El sistema permite agregar varias imágenes por sujeto para el entrenamiento y para la prueba también pero sin clasificarlas por sujeto. De tal forma que podemos entrenar algunas imágenes con algún PCA y otras imágenes con otro, no necesariamente todas las imágenes tienen que ser entrenadas y además varias imágenes de un mismo sujeto pueden estar entrenadas con el mismo tipo de PCA. El que se permita tener varias imágenes por sujeto es debido a que este sistema permite entrenar con PCA modificado. En cuanto a las imágenes de prueba, se pueden agregar al sistema tantas como se quiera, pero sólo serán autenticadas aquellas que se marquen para tal efecto. Serán autenticadas con el método indicado (uno a la vez).

A continuación se muestran las ventanas que constituyen toda la interfaz gráfica de usuario de nuestro sistema.

La ventana principal es la que contiene todas las opciones disponibles y la que nos permite explorar las imágenes de entrenamiento y prueba (vea la fig. 5.1).

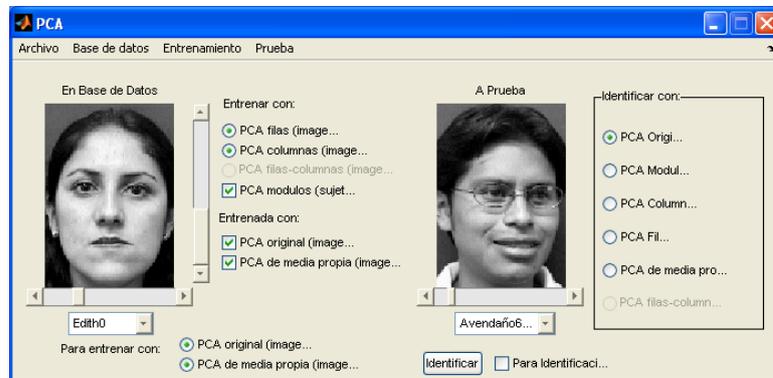


Figura 5.1 Ventana principal

La ventana de la figura 5.2 nos sirve, como parece ya obvio, para explorar los directorios en búsqueda de las imágenes que deseemos agregar a un sujeto ya existente o para agregar un nuevo sujeto.

La última de las ventanas la utilizamos para mostrar las relaciones encontradas por el sistema entre las imágenes de prueba y entrenamiento (ver fig. 5.3).

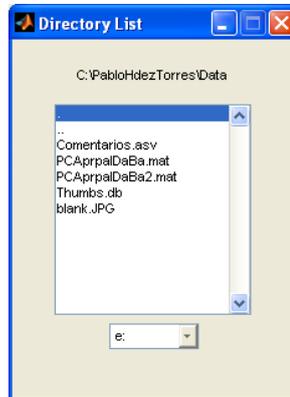


Figura 5.2 Listado de los directorios y sus contenidos



Figura 5.3 Imágenes de prueba vs Imágenes de entrenamiento

Así mismo, como pudimos observar, la ventana principal contiene una barra con tres menús, además, algunas de esas opciones también aparecen en menús contextuales que se despliegan cuando el puntero del mouse esta sobre los ejes que muestran las fotografías. Las opciones para los menús “Base de Datos”, “Entrenamiento” y “Prueba” se muestran en la figura 5.4.

1.1 MENÚS

1.1.1 Menú Archivo. Los ítems para este menú son:

Guardar imágenes de entrenamiento: Si se realizó algún cambio al contenido de las imágenes de entrenamiento, podemos guardarlo, ya que de otra manera esto se hará hasta el final, cuando se salga de la interfaz gráfica, por lo que si previamente ocurre algún suceso que nos impida salir correctamente de la GUI los cambios podrían no guardarse.



Figura 5.4 Menús y sus opciones

Guardar imágenes de prueba: La misma situación explicada para las imágenes de entrenamiento puede pasar para las imágenes de prueba

Salir: Salimos de la interfaz gráfica guardando los cambios en los contenidos de las imágenes de prueba y entrenamiento.

1.1.2 Menú Base de Datos. Las opciones de este menú están encaminadas a administrar las imágenes que se utilizarán como entrenamiento.

Agregar imagen: Agrega una nueva imagen al registro del sujeto actual, el sujeto actual es aquel que se muestra en el eje correspondiente para las imágenes de entrenamiento.

Agregar sujeto: Agrega un nuevo registro y la primera imagen de un sujeto.

Borrar imagen: Borra la imagen actual que se muestra en los ejes respectivos. Si esa imagen es la única imagen del sujeto, entonces se borra todo el registro del sujeto.

Borrar sujeto: Borra el registro del sujeto actual.

Borrar a todos los sujetos...: Limpia por completo la estructura que contiene los registros de los sujetos de entrenamiento, dejándola en una estructura vacía.

1.1.3 Menú Entrenamiento.

PCA original: Entrenamos las imágenes marcadas con PCA original.

PCA de media propia: Entrenamos las imágenes PCA de media propia.

Para PCA original, si no esta marcada más de una imagen, la opción correspondiente del menú estará desactivada, lo mismo sucederá con PCA de media propia.

1.1.4 Menú Prueba. Las opciones de este menú son para el manejo de las imágenes de prueba, cabe mencionar que al igual que las imágenes de entrenamiento éstas se almacenan en su propia estructura y cuando esta estructura esta vacía (porque no se han agregado imágenes) ciertas opciones del menú estarán desactivadas. En la estructura mencionada, cada imagen se encuentra en un registro y además las imágenes no se manejan por sujetos.

Agregar imagen: Agregamos una nueva imagen en un nuevo registro.

Borrar imagen: Borra la imagen actual que se muestra en el eje correspondiente de las imágenes de prueba.

Borrar todo: Pone la estructura que contiene las imágenes de prueba a una estructura vacía.

1.2 CONTROLES DE LA VENTANA PRINCIPAL

Todos los controles, a excepción de los menús desplegables, nos sirven para controlar el entrenamiento y la identificación de las imágenes guardadas en nuestras estructuras.

Controles bajo “Entrenar con:”. Los radio botones y el checkbox que se encuentran bajo este recuadro tienen la doble función de entrenar e indicar imágenes entrenadas con PCA filas, con PCA columnas o PCA módulos. Cada vez que se da un click a estos controles cambiarán su estado, si el cambio es de ‘apagado’ a ‘encendido’ significa que la imagen (caso filas y columnas) o el sujeto (caso módulos) actual no estaba entrenado con el método respectivo y por lo tanto se entrenará; si el cambio es de ‘encendido’ a ‘apagado’ significa que la imagen ya estaba entrenada y por lo tanto se borrará del entrenamiento, de tal forma que siempre el estado del control nos dirá si una imagen está o no entrenada con el método que representa el control. El estado del control se guarda junto con la imagen o el sujeto para que cada vez que la imagen (y por lo tanto un sujeto) aparezca en los ejes de entrenamiento se recuperen los estados de los controles y se muestre el valor correcto.

Controles de “Para entrenar con:”. Cada imagen en la base de datos puede ser marcada, ya sea para que sea entrenada con PCA original o con PCA de media propia, los radio botones de esta parte de la ventana tienen esa función. Cuando se elija una opción del menú ‘Entrenamiento’ sólo las imágenes marcadas con la opción correspondiente se entrenarán. La razón de por qué primero se marcan las imágenes y luego con un control diferente se entrenan o se quitan del entrenamiento es debido a que en los mencionados métodos para agregar una imagen en el entrenamiento se necesitan re-entrenar a todas las demás imágenes ya entrenadas junto con la nueva imagen, ahora imagínese re-entrenar el sistema cada vez que se quiere agregar una nueva imagen, eso sería mucha pérdida de tiempo computacional, por lo que primero se deberán elegir las imágenes que quieran ser entrenadas con tales métodos y luego se entrenarán todas juntas.

Controles bajo “Entrenada con:”. Estos checkboxes muestran si la imagen de entrenamiento presente está o no entrenada con PCA original o PCA media propia.

Deslizadores. El deslizador horizontal nos permite movernos entre los sujetos y el deslizador vertical nos permite movernos entre las imágenes del sujeto.

Controles de prueba. Con el checkbox “Para identificación” marcamos a las imágenes de prueba que queremos identificar una vez que presionemos el botón correspondiente. El método que se utilizará para la identificación depende del radio botón marcado en el cuadro “Identificar con:”.

2. COMPARACIÓN DE LAS IMPLEMENTACIONES

En esta sección analizamos y comparamos a cada uno de los algoritmos desde el punto de vista de acierto y error, es decir, verificaremos cual de ellos tiene más identificaciones correctas y menos aciertos que los demás, o si hay algoritmos cuyo desempeño es similar y se analizarán las imágenes donde los métodos fallan y donde no lo hacen, esto con el fin de averiguar si hay alguna característica presente en las imágenes que provoca que los métodos fallen. En las pruebas realizadas a los algoritmos se utilizaron los rostros de la base de datos del INAOE, base de datos ya mencionada en la sección “Metodología Empleada”. Las características de las fotografías de dicha base parecen apropiadas para cotejar los métodos PCA, ya que como criterios de comparación de los algoritmos tomamos en cuenta algunos rasgos propios del rostro que provocan más fallos en la identificación ¹ y no tomamos cuestiones que tuvieran que ver con la calidad de las fotografías obtenidas. ²

2.1 PRUEBAS

Las pruebas consisten en presentar a cada algoritmo diversos conjuntos de rostros (los conjuntos de prueba), cada conjunto con ciertas características, para después verificar con una medida de similaridad la cantidad de aciertos y errores que el algoritmo comete y posteriormente comparar con otro algoritmo al que se le ha presentado el mismo conjunto y de esta forma verificar cual ha sido el mejor. Se realizaron dos tipos de pruebas, etiquetadas como PRUEBA I y PRUEBA II, cada una de ellas divididas en dos sub-pruebas o fases, la fase A y la fase B. En las pruebas no se ha establecido un límite a la medida de similaridad por lo que aquí no podemos decir si una imagen esta o no en la base de datos sino que se va a encontrar aquella imagen que es la más parecida. El valor del límite de similaridad debe ser tal que podamos estar seguros que dos imágenes son del mismo sujeto si ese límite es sobrepasado.

El conjunto de imágenes de entrenamiento consistió de los 18 rostros de la base de datos del INAOE, todos sin rotaciones, con expresión neutra y sin accesorios. Como PCA modular debe utilizar más de una imagen (de preferencia más de 2) se utilizaron para este tipo de PCA tres imágenes adicionales con las siguientes características cada una:

¹Características como las ya mencionadas anteriormente.

²Calidad dependiente de la iluminación, del aparato fotográfico, etc.

- rotación de 45° sin accesorios y expresión neutra.
- rotación de -45° sin accesorios y expresión neutra.
- frontal con expresión neutra y con gafas oscuras.

La razón por la que elegimos imágenes con esas cualidades es porque representan situaciones de un rostro que nos provocan de la máxima cantidad de fallos en el PCA original, aunque también se podrían haber elegido imágenes que muestran la o las situaciones más probables del rostro a la hora de estar frente al sistema y tomarse la fotografía para ser identificado para de esta forma cubrir las facetas del rostro en el sentido más amplio posible .

En PRUEBA I no se eliminan eigenvectores, exceptuando PCA original y PCA modular que al utilizar ambos el truco de Turk & Pentland automáticamente limitan el número de eigenvectores a una cantidad igual al número de imágenes de entrenamiento; los eigenvalores correspondientes a tales eigenvectores son los de mayor magnitud absoluta.

En PRUEBA II se eliminan algunos eigenvectores siguiendo el criterio de la media de eigenvalores. Para métodos como PCA original y PCA modular se eliminan adicionalmente más eigenvectores.

La fase A de las pruebas consiste en utilizar la norma L1 como medida de similaridad y en la fase B utilizar la norma L2 con el mismo fin. En PCA la medida de similaridad de cosenos ofrece los mismos resultados que la norma L2 [5] por lo que no la utilizaremos en una tercera fase. El hecho de utilizar más de una medida de similaridad es con el fin de medir su influencia sobre el reconocimiento.

Los conjuntos de prueba son los siguientes:

- A. Rostros con expresión neutra, frontales y con gafas oscuras.
 - A.1 Mismas que A pero sin expresión neutra: sonrisa mostrando dentadura.
- B. Rostros con expresión neutra, rotados 45° y sin ningún accesorio.
 - B.1 Mismas que B pero sin expresión neutra: sonrisa mostrando dentadura, y con

accesorios:
lentes transparentes.

C. Rostros con expresión neutra, rotados -45° y sin ningún accesorio.

C.1 Mismas que C pero sin expresión neutra: sonrisa mostrando dentadura, y con accesorios: lentes transparentes.

D. Rostros con sonrisa mostrando dientes, de frente y sin accesorios.

2.1.1 Resultados. Los resultados mostrando los aciertos y las fallas de cada PCA se muestran en la tabla 5.1 y en la tabla 5.2.

Tabla 5.1 Resultados de PRUEBA I

Tipo de imagen de prueba	PCA original	PCA filas	PCA columnas	PCA de media propia	PCA por módulos
Fase A					
A	16	16	16	17	18
A.1	15	16	14	16	12
B	12	9	12	11	18
B.1	9	9	8	10	2
D	18	17	18	18	13
Fase B					
A	17	17	17	17	18
A.1	16	16	16	16	12
B	12	12	12	12	18
B.1	10	10	10	11	3
D	18	18	18	18	14

El apéndice B.1 muestra las imágenes utilizadas como imágenes de entrenamiento y algunas de las imágenes utilizadas en las pruebas. Dentro de esas imágenes podemos observar que Ulises es muy parecido a Hugo, en las pruebas pudo notarse tal hecho ya que se relacionaba frecuentemente a Hugo, Hugo_ y a Ulises con la misma persona y también relacionaba a Ulises con Hugo y Hugo_. No es de extrañarse que frecuentemente se relacionara a Hugo con Hugo_, pero también en algunos casos podía distinguir a Hugo de Hugo_.

Tabla 5.2 Resultados de PRUEBA II

Tipo de imagen de prueba	PCA original	PCA filas	PCA columnas	PCA de media propia	PCA por módulos
Fase A					
A	14	15	13	7	18
A.1	13	16	14	6	2
B	8	10	10	2	18
B.1	8	10	10	2	1
D	16	17	17	9	1
Fase B					
A	14	16	14	7	18
A.1	12	14	15	6	1
B	8	10	10	3	18
B.1	8	10	10	3	1
D	16	18	17	9	2

Un error muy frecuente ocurría con Hugo y Carlos, observando sus respectivas imágenes de prueba y comparándolas con sus respectivas imágenes de entrenamiento se verifico que sus rostros en las imágenes de prueba aparecían desplazados un poco hacia la izquierda a una distancia tal que su oreja derecha no era visible. De esta forma podemos ver que un desplazamiento muy grande en las imágenes puede afectar severamente la identificación.

Los cero errores de PCA módulos que se encuentran en las prueba I y II, fases A y B, con rostros de prueba tipo A y B, se deben a que las imágenes de entrenamiento se utilizaron también como imágenes de prueba. En la práctica es casi imposible que dos fotografías de un rostro salgan exactamente iguales. Este método es el que nos ofreció los peores resultados.

Observando los resultados de las tablas podemos darnos cuenta que al eliminar eigenvectores disminuimos de forma no considerable los aciertos de los métodos PCA filas, PCA columnas y PCA original mientras que para PCA de media propia y PCA modular la tasa de errores aumento considerablemente.

Las imágenes de prueba con las que más se cometen errores son las más distintas con respecto a las imágenes de entrenamiento, como cuando se usan gafas o cuando el rostro esta rotado, en cambio cuando la imagen parece casi inalterada los errores son los menos, por ejemplo, cuando el rostro solo esboza una expresión leve como una sonrisa o tiene los ojos cerrados.

La influencia de la medida de similaridad también se hace notar, ya que cuando no se eliminan vectores propios, los aciertos son más cuando se utiliza la norma L2 que la norma L1, mientras que, cuando se eliminan vectores propios la diferencia no es muy notable.

En la prueba I, y de manera total, el porcentaje de error por cada método fue el siguiente:

- PCA original: 20.1 %
- PCA filas: 22.2 %
- PCA columnas: 21.7 %
- PCA media propia: 19.4 %
- PCA módulos: 28.9 %

En la prueba II, y de manera también total, los porcentajes de error fueron:

- PCA original: 35 %
- PCA filas: 24.4 %
- PCA columnas: 27.8 %
- PCA media propia: 70 %
- PCA módulos: 55.6 %

Por lo que en la prueba II, el mejor método fue PCA por filas.

Por último, las variaciones en la prueba I no son tan fluctuantes como en la prueba II y los métodos propuestos son por muy poco porcentaje mejores que los PCA existentes, salvo en la prueba II, en donde PCA de media propia resultó ser un caos.

Capítulo 6

CONCLUSIÓN FINAL DEL TRABAJO

Durante este trabajo se propusieron formas diferentes de realizar Análisis de Componentes Principales sobre las imágenes con la posibilidad de mejorar el desempeño de este método. Tres formas de llevar a cabo el análisis fueron propuestas en este trabajo y fueron comparadas con dos PCA's existentes: PCA original o clásico y PCA modular.

De las pruebas realizadas sobre los métodos de Análisis de Componentes Principales se desprende que la medida de similaridad conocida como la norma L2 ofrece mejores resultados que la norma L1 aunque dichos resultados no son tan ventajosos como para despreciar a la norma L1.

Con reducción de dimensionalidad el desempeño de los algoritmos se ve afectado, pero solo de forma crítica en los métodos PCA de media propia y PCA modular y no de forma tan severa en PCA clásico pero si mucho más que PCA filas y columnas. El que la reducción de la dimensionalidad haya afectado más a los PCA's ya mencionados podría deberse a la cantidad de valores propios que ellos manejan, ya que en los PCA's por barridos verticales y horizontales la cantidad de valores propios es mayor que en los otros métodos, además recordemos que hay una reducción de dimensionalidad adicional en PCA clásico y modular cuando se utiliza el truco de Turk y Pentland.

Sobre la cuestión de cuál es el mejor método y basándonos en las pruebas concluimos que la norma L2 ofrece mejores resultados que la norma L1, que la eliminación de vectores propios afecta el desempeño pero no de forma crítica, a excepción de los métodos PCA modular y PCA de media propia y en menor medida en PCA clásico, y que el método que ofrece los resultados más pobres es PCA por módulos a excepción de cuando se eliminan vectores propios ya que en ese caso el que ofrece los peores resultados es el método propuesto PCA de media propia.

En los métodos clásico y modular no se deberían eliminar más vectores propios de los que ya existen cuando se utiliza el truco ya discutido, esto

debido a que con tal truco el número de vectores propios es proporcional al número de imágenes en el entrenamiento, y estas son una cantidad pequeña en comparación con la cantidad de píxeles que contiene cada imagen. En general, no se recomienda que el número de imágenes utilizadas sea muy cercano al número de píxeles en una imagen. En el método de media propia, de manera implícita, y sin utilizar el truco de Turk la cantidad de vectores propios es igual a la cantidad de imágenes de entrenamiento, por tal razón podría ser poco conveniente eliminar vectores propios, ya que al hacerlo estaríamos eliminando imágenes base.

Sin la eliminación adicional que se hace a PCA clásico, y sin la eliminación de vectores propios en PCA de media propia, el desempeño de todos los métodos, exceptuando a PCA modular, es similar y ni uno de los métodos es tan ventajoso como para insinuar que se trata del ideal y mejor.

Si bien los métodos propuestos no resultaron tener mejor desempeño con respecto a los existentes, estos si lograron un desempeño similar y no inferior. Sin embargo, los métodos propuestos ofrecen algunas ventajas, y por supuesto, también algunas desventajas, dentro de las ventajas podemos señalar que para PCA por filas o columnas no es necesario re-entrenar a todas las imágenes del conjunto entrenamiento cada vez que se agrega o elimina un sujeto como es el caso de PCA clásico ya que cada imagen tiene su propio espacio y no esta relacionado con los espacios de las demás imágenes; dentro de las desventajas, existe la del tiempo mayor requerido cuando se desea reconocer a un sujeto, esto debido a la existencia de más de un espacio de autenticación, lo que requiere el cambio de la imagen de prueba a la base de cada uno de ellos, además de que la comparación dentro de esos espacios es fila a fila o columna a columna; en contraste, en PCA original existe el cambio a una sola base y la comparación es imagen a imagen.

De las ventajas que PCA de media propia ofrece podemos mencionar que de manera implícita este tipo de PCA maneja una matriz de covarianzas y vectores propios de menor dimensión con respecto a los manejados en PCA original.

Cabe mencionar que actualmente alumnos de la Maestría en Ciencias de la Computación se encuentran desarrollando un Sistema Evaluador Adaptable (CAT, por sus siglas en inglés), al cual, una vez desarrollado, se le implementará reconocimiento de rostros aplicando lo estudiado en este trabajo. Lo anterior con la finalidad de que la validación de identidad para los usuarios del CAT sea de manera biométrica y, por tanto, más segura, poniendo en un segundo término el uso de contraseñas.

Referencias

- [1] Base de rostros inaoe, 2004. Departamento de Ciencias Computacionales.
- [2] Howard Anton. *Introducción al álgebra lineal*. Limusa, 2da. edition, 1999.
- [3] Raymond A. Barnett and Michael R. Ziegler. *Linear Algebra, An Introduction with Applications*. MACMILLAN, 1987.
- [4] Lindsey I. Smith. A tutorial on principal component analysis, 2002. February 26.
- [5] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski. Face recognition by independent component analysis. *IEEE Transactions on neural networks*, 13(6), November 2002.
- [6] Sanjit K. Mitra. *Digital Signal Processing, A Computer-Based Approach*. McGraw-Hill, International Edition, 2002.
- [7] Delores M. Etter. *Solución de Problemas de Ingeniería con MATLAB*. Prentice-Hall, Pearson Educación, 1997.
- [8] M. A. Vicente, O. Reinoso, C. Pérez, J.A. Sabater, and J.A. Azorín. Reconocimiento de objetos 3d mediante análisis pca. Technical report, Universidad Miguel Hernández, División Ingeniería de Sistemas y Automática, Departamento de Ingeniería, Campus de Elche Avda. del Ferrocarril s/n 03202 Elche (Alicante) España.
- [9] Wei Sun. 308-766a shape analysis in computer vision final project report: Face recognition. Technical report, Department of Electrical Engineering McGill University, Diciembre 21 1998. E-mail: wsun@cim.mcgill.ca.

- [10] David H. Hubel. *Eye, Brain and Vision*. Scientific American Library, 1995.
- [11] Hervé Abdi, Dominique Valentin, and Betty G. Edelman. Eigenfeatures as intermediate level representations: The case for pca models. *Brain and Behavioural Sciences*, pages 175–177, 1997.
- [12] Garrison W. Cottrell, MatthewÑ. Dailey, Curtis Padgett, and Ralph Adolphs. Is all face processing holistic? the view from ucsd. In M. Wenger and J. Townsend, editors, *Computational, Geometric, and Process Perspectives on Facial Cognition: Contexts and Challenges*. Erlbaum, 2000.
- [13] Peter J. B. Hancock, Vicki Bruce, and A. Mike Burton. Testing principal components representations for faces. 1997.
- [14] James C. Bartlett, Jean H. Searcy, and Hervé Abdi. What are the routes to face recognition? In M. Peterson and G. Rhodes, editors, *Perception of faces, objects, and scenes: analytic and holistic processes*, pages 21–52. Oxford University Press, Oxford, 2003.
- [15] Jiaming Li, Geoff Poulton, Ying Guo, and Rong-Yu Qiao. Face recognition based on multiple region features. In Talbot H. Sun C., Ourselin S., and Adriaansen T., editors, *Proc. VIIth Digital Image Computing: Techniques and Applications*, NATO-ASI Series C, pages 10–12. 2003.

GLOSARIO

GUI Interfaz Gráfica de Usuario.

holístico Análisis en el que se estudia el todo de un sistema a través de todas sus partes y las relaciones existentes entre ellas, relaciones visibles o no, para comprender el comportamiento del mismo.

ICA Análisis de Componentes Independientes.

generalización Proceso de extrapolación y/o interpolación para la predicción de la respuesta de un sistema.

PCA Análisis de Componentes Principales.

PCA modificado o por módulos PCA realizado por grupo de imágenes, donde las imágenes de cada grupo son de una misma persona.

Apéndice: Algoritmos Empleados**Algoritmo 2.1**

*/*Define el intervalo de valores de la primera variable.*/*

$$\mathbf{x} \in [-1, 1];$$

*/*Define las imágenes con resolución 1×3 , las cuales se colocan por columna. Los parámetros ϵ son vectores aleatorios entre -1 y 1 y de tamaño igual al tamaño de \mathbf{x} .*/*

$$X_1 = \mathbf{x};$$

$$X_2 = 0,5895/0,5641 * \mathbf{x} + 0,0 * \epsilon_2;$$

$$X_3 = 0,5782/0,5641 * \mathbf{x} + 1,0 * \epsilon_3;$$

*/*Define el valor medio de cada pixel; es decir, el valor medio por fila de \mathbf{X} .*/*

$$\boldsymbol{\mu}^T = \text{promedio}(\mathbf{X}^T);$$

*/*Define una nueva matriz de imágenes al restar el valor medio a cada imagen en \mathbf{X} . El vector fila 1 tiene unos en*

sus entradas y es de igual tamaño que el vector

\mathbf{x} .

$$\mathbf{Y} = \mathbf{X} - \boldsymbol{\mu}^T * \mathbf{1};$$

/*Calcula la matriz de covarianza de \mathbf{Y}^T , puesto que cada imagen tiene resolución 1×3 , entonces el orden de la matriz de covarianza es igual a 3.*/

$$\mathbf{C} = \text{cov}(\mathbf{Y}^T);$$

/*Determina los eigenvectores (\mathbf{V}) y eigenvalores (\mathbf{D}) de la matriz de covarianza. El eigenvector en posición $\mathbf{D}(1, 1)$ corresponde al eigenvector $\mathbf{V}(:, 1)$, y así sucesivamente.*/

$$[\mathbf{V}, \mathbf{D}] = \text{eig}(\mathbf{C});$$

/*Grafica los ejes representados por cada eigenvector.*/

$\text{grafica3D}(\mathbf{x}, \mathbf{V}(2, 1)/\mathbf{V}(1, 1) * \mathbf{x}, \mathbf{V}(3, 1)/\mathbf{V}(1, 1) * \mathbf{x});$

$\text{grafica3D}(\mathbf{x}, \mathbf{V}(2, 2)/\mathbf{V}(1, 2) * \mathbf{x}, \mathbf{V}(3, 2)/\mathbf{V}(1, 2) * \mathbf{x});$

$\text{grafica3D}(\mathbf{x}, \mathbf{V}(2, 3)/\mathbf{V}(1, 3) * \mathbf{x}, \mathbf{V}(3, 3)/\mathbf{V}(1, 3) * \mathbf{x});$

Apéndice: Generalización

A nuestras fotografías de prueba, que son aquellas imágenes del rostro de la persona que solicita ser autenticada o identificada, se les llama algunas veces, imágenes de generalización. A continuación explicamos el proceso de generalización.

Para un sistema dado, generalmente se desea una salida, entonces lo que se hace es, entrenarlo con entradas y con la respuestas que se quieren para dichas entradas, pero muchas veces es imposible entrenarlo con todas las entradas posibles, por lo que frecuentemente se hace es elegir solamente algunas de las entradas y sus respuestas correspondientes, esperando que el sistema responda a las otras entradas con las salida deseadas, esto es, que con las repuestas a unas entradas, el sistema conozca la respuesta a cualquier entrada, es decir, que generalice su respuesta. Las entradas y sus salidas utilizadas como entrenamiento, deben ser tales que caractericen el comportamiento del sistema.

Veamos un ejemplo con una gráfica. La gráfica B.1 muestra la respuesta deseada de un sistema en tiempo continuo, como podemos deducir, la cantidad de entradas posibles es infinita, por lo que se utilizan solo algunas entradas para preparar al sistema, de tal forma que para las otras entradas la respuesta sea la deseada.

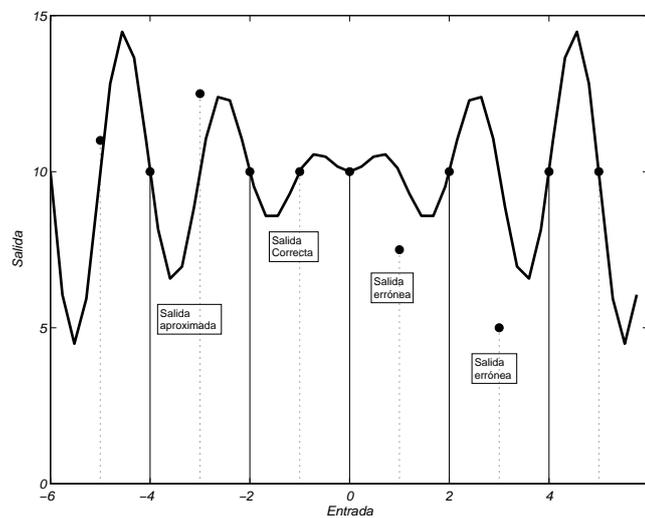


Figura B.1 Generalización de una respuesta

En cuanto a las imágenes es algo similar; no es posible entrenar un sistema biométrico de autenticación con todos los rostros posibles que se le pueden

presentar, sino que sólo se utilizan aquellos de los cuales se quiere una respuesta afirmativa por parte del sistema, por lo que para el resto de los rostros se desea una respuesta negativa. Hasta este punto, a las imágenes de prueba se les llama de generalización porque, seguramente, no serán exactamente las mismas con las que se entrenó al sistema y a las cuales el sistema deberá responder conforme al diseño de su salida.

Apéndice: Imágenes utilizadas

La figura C.1 muestra las imágenes de entrenamiento en las pruebas que se realizaron sobre los métodos PCA.

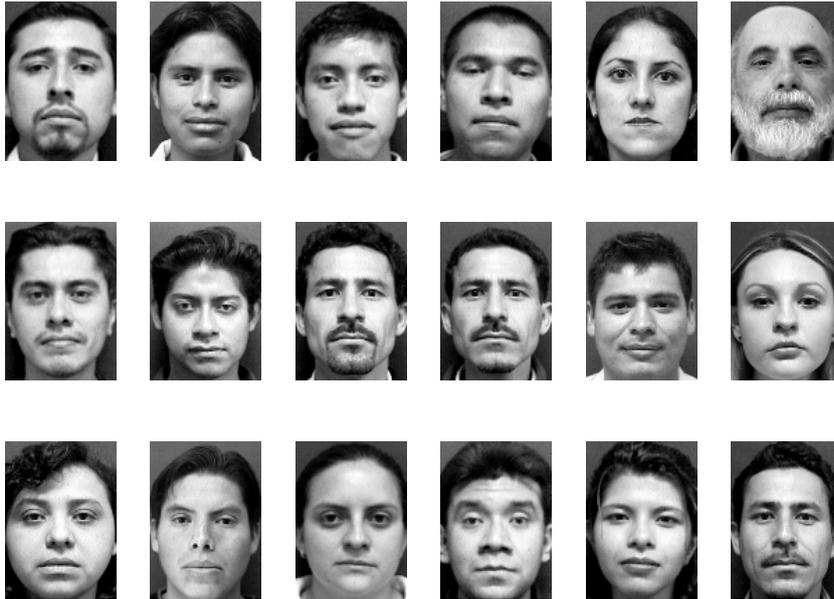


Figura C.1 Imágenes utilizadas como entrenamiento

Adicionalmente, para PCA modular se utilizaron tres imágenes más por cada sujeto, imágenes con las características de las imágenes de la figura C.2.



Figura C.2 Características de las imágenes adicionales para el entrenamiento de PCA módulos

Las figuras restantes muestran algunas imágenes utilizadas como prueba.

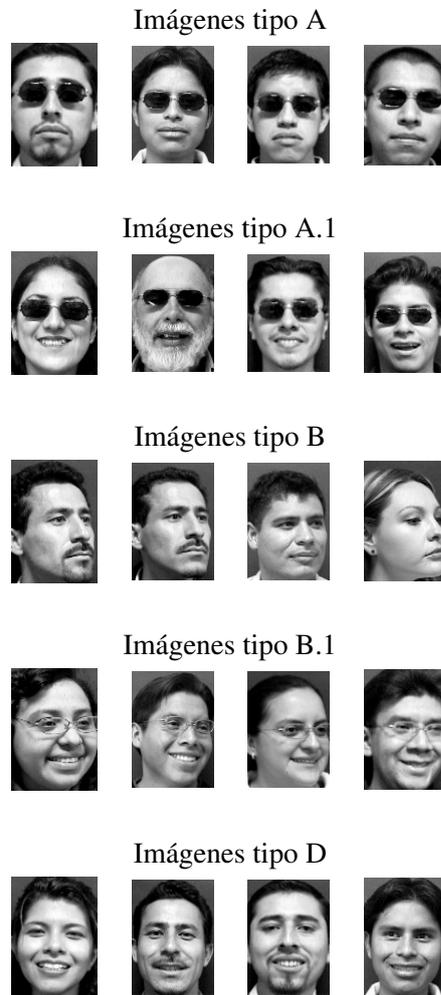


Figura C.3 Imágenes utilizadas como prueba

Apéndice: Código

Se incluye el código de las funciones que implementan a cada una de las variantes de PCA tratadas en este trabajo, funciones que son llamadas por la Interfaz Gráfica de Usuario para el entrenamiento de las imágenes.

PCA original.

```

function orig(imagenes,nombres)
%ORIG  Entrena imagenes con PCA clasico.
% ORIG(imagenes,nombres) entrena las "imagenes"
% con PCA clasico y guarda las imagenes en su
% nueva base junto con sus nombres, guarda además
% la base obtenida y la imagen media. Utiliza el
% truco de Turk & Pentland para el calculo de los
% eigenvectores.

load PCAprpalDaBa2 orig

% si el numero de imágenes no es suficiente se
% guarda una estructura vacía.
if length(nombres) > 2,
    % tamaño de matriz imágenes, re-formar matriz
    % imágenes.
    [r,c,p] = size(imagenes);
    imagenes = reshape(imagenes,[r*c p 1]);

    % calculamos la matriz de covarianzas.
    [imagenes,orig(1).imamed] = mdia(imagenes,2);
    mcov = (imagenes'*imagenes)/(r*c);

    % calculamos los pseudo-eigenvectores y
    % eigenvalores.
    [peigvecs,eigvals] = eig(mcov);
    eigvals = diag(eigvals);

    % conservamos los eigenvalores que son mayores
    % o iguales al promedio junto con sus
    % correspondientes pseudo-eigenvectores.
    k = find(abs(eigvals) >= mean(abs(eigvals)));
    eigvals = eigvals(k);
    peigvecs = peigvecs(:,k);

    % ordenamos los eigenvalores de menor a mayor
    % junto con sus correspondientes
    % pseudo-eigenvectores.
    [orig(1).eigvals,1] = sort(abs(eigvals));
    orig(1).eigvals = eigvals(1);

```

```

peigvecs = peigvecs(:,1);

% obtenemos los eigenvectores reales.
orig(1).base = normat(imagenes*peigvecs);

% cambiamos a cada imagen al nuevo espacio y
% la guardamos junto con sus nombres.
orig(1).imagenes = orig.base' * imagenes;
orig(1).nombres = nombres;
else
    orig = struct([]);
end

% guardamos la estructura a disco.
save PCAprpalDaBa2 orig -APPEND

```

PCA filas.

```

function eigfil(Im,nombre)
% EIGFIL Entrena imagen por filas.
% EIGFIL(Im,nombre) recibe la imagen en "Im" del
% rostro que va a entrenar y también recibe su
% "nombre" de tal forma que se guarda la nueva
% base generada con "Im", los valores propios y
% la imagen en el nuevo espacio con la etiqueta
% "nombre" en la estructura fils.

% cargamos la estructura "fils"
load PCAprpalDaBa2 fils

% el siguiente if es porque el comando "end" no
% funciona con arreglos vacíos.
if isempty(fils)
    fin = 0;
else
    fin = length(fils);
end

% calculamos la matriz de covarianzas.
[r,c] = size(Im);
fils(fin+1,1).filmed = mean(Im,1);

```

102

```
Im = Im-ones(r,1)*fils(end,1).filmed;
mcov = (Im'*Im)/r;

% obtenemos los eigenvectores y eigenvalores.
[eigfils,eigvals] = eig( mcov );
eigvals = diag(eigvals);

% conservamos los eigenvalores que son mayores o
% iguales al promedio junto con sus correspondientes
% eigenvectores.
k = find(abs(eigvals) >= mean(abs(eigvals)));
eigvals = eigvals(k);
eigfils = eigfils(:,k);

% ordenamos los eigenvalores de menor a mayor junto
% con sus correspondientes eigenvectores.
[fils(end,1).eigvals,1] = sort(abs(eigvals));
fils(end,1).eigvals = eigvals(1);
fils(end,1).base = eigfils(:,1);

% proyectamos la imagen y guardamos su nombre.
fils(end,1).imagen = Im*fils(end,1).base;
fils(end,1).nombre = nombre;

% guardamos la estructura a disco.
save PCAprpalDaBa2 fils -APPEND
```

PCA columnas.

```
function eigcol(Im,nombre)
% EIGCOL Entrena imagen por columnas.
% EIGCOL(Im,nombre) recibe la imagen de un rostro
% en "Im" y su "nombre" respectivo para entrenarla
% con PCA columnas y guardar en la estructura
% "cols" la nueva base formada, los valores propios
% correspondientes a la base y la imagen en su
% nueva base, todo, con la etiqueta "nombre".

% cargamos la estructura "cols"
load PCAprpalDaBa2 cols
```

```

% el siguiente if es porque "end" no funciona con
% arreglos vacíos.
if isempty(cols)
    fin = 0;
else
    fin = length(cols);
end

% calculamos la matriz de covarianzas.
[r,c] = size(Im);
cols(fin+1,1).colmed = mean(Im,2);
Im = Im - cols(end,1).colmed*ones(1,c);
mcov = (Im*Im')/c;

% obtenemos los eigenvectores y los eigenvalores.
[eigcols,eigvals] = eig( mcov );
eigvals = diag(eigvals);

% conservamos los eigenvalores que son mayores o
% iguales al valor promedio junto con sus
% correspondientes eigenvectores.
k = find(abs(eigvals) >= mean(abs(eigvals)));
eigvals = eigvals(k);
eigcols = eigcols(:,k);

% ordenamos los eigenvalores de menor a mayor
% junto con sus correspondientes eigenvectores.
[cols(end,1).eigvals,1] = sort(abs(eigvals));
cols(end,1).eigvals = eigvals(1);
cols(end,1).base = eigcols(:,1);

% proyectamos la imagen y guardamos su nombre.
cols(end,1).imagen = cols(end).base'*Im;
cols(end,1).nombre = nombre;

% guardamos la estructura a disco.
save PCAprpalDaBa2 cols -APPEND

```

PCA media propia.

```
function prop(imagenes, nombres)
```

```

% PROP  Entrena imágenes con PCA de media propia.
% PROP(imagenes, nombres) recibe un conjunto de
% "imagenes" y obtiene otro conjunto de imágenes
% llamadas imágenes base y obtiene los coeficientes
% de las "imagenes". Las imágenes base, los
% coeficientes y los nombres correspondientes se
% guardan en la estructura "prop".

load PCAprpalDaBa2 prop

% si el numero de imágenes no es suficiente guardamos
% una estructura vacía.
if length(nombres) > 2,
    % dimensiones, re-formar matriz imágenes.
    [r,c,p] = size(imagenes);
    imagenes = reshape(imagenes,[r*c p 1]);
    medims = mean(imagenes,1);

    % calculamos matriz de covarianzas.
    imagenes = imagenes - ones(r*c,1)*medims;
    mcov = (imagenes' * imagenes)/(r*c);

    % obtenemos los eigenvectores y eigenvalores.
    [eigvecs,eigvals] = eig( mcov );
    eigvals = diag(eigvals);

    % conservamos los eigenvalores que son mayores
    % o iguales al promedio junto con sus
    % correspondientes eigenvectores.
    k = find(abs(eigvals) >= mean(abs(eigvals)));
    eigvals = eigvals(k);
    eigvecs = eigvecs(:,k);

    % ordenamos los eigenvalores de menor a mayor
    % junto con sus correspondientes eigenvectores.
    [prop(1).eigvals,1] = sort(abs(eigvals));
    prop(1).eigvals = eigvals(1);
    eigvecs = eigvecs(:,1);

    % obtenemos imágenes base y coeficientes de las
    % imágenes de entrenamiento.
    prop(1).base = imagenes*eigvecs;

```

```

prop(1).coeficientes = pinv(eigvecs);

% guardamos los nombres de las imágenes
prop(1).nombres = nombres;

else
    prop = struct([]);
end

% guardamos la estructura a disco.
save PCAprpalDaBa2 prop -APPEND

```

PCA módulos.

```

function modu(imagenes,nombre)
%MODU Entrenamiento por PCA módulos.
% MODU(imagenes,nombre) entrena las imágenes
% contenidas en "imagenes" con PCA por módulos
% y las guarda con la etiqueta "nombre" en la
% estructura "mods", además guarda la base
% obtenida, la imagen media y los eigenvalores
% correspondientes. Utiliza el truco de Turk &
% Pentland para el calculo de los eigenvectores.

load PCAprpalDaBa2 mods

% el siguiente if es porque "end" no funciona con
% estructuras vacías.
if isempty(mods),
    fin = 0;
else
    fin = length(mods);
end

% tamaño de matriz imágenes, re-formar matriz
% imágenes.
[r,c,p] = size(imagenes);
imagenes = reshape(imagenes,[r*c p 1]);

% calculamos la matriz de covarianzas con truco de
% Turk & Pentland

```

106

```
[imagenes,mods(fin+1,1).imamed] = mdia(imagenes,2);
mcov = (imagenes'*imagenes)/(r*c);

% calculamos pseudo-eigenvectores y eigenvalores.
[peigvecs,eigvals] = eig(mcov);
eigvals = diag(eigvals);

% conservamos los eigenvalores que son mayores o
% iguales al promedio junto con sus correspondientes
% pseudo-eigenvectores.
k = find(abs(eigvals) >= mean(abs(eigvals)));
eigvals = eigvals(k);
eigvecs = peigvecs(:,k);

% ordenamos los eigenvalores de menor a mayor junto
% con sus correspondientes pseudo-eigenvectores.
[mods(end,1).eigvals,1] = sort(abs(eigvals));
mods(end,1).eigvals = eigvals(1);
peigvecs = peigvecs(:,1);

% obtenemos los eigenvectores reales.
mods(end,1).base = normat(imagenes*peigvecs);

% cambiamos a cada imagen al nuevo espacio y las
% guardamos junto con sus nombres.
mods(end,1).imagenes = mods(end,1).base' * imagenes;
mods(end,1).nombre = nombre;

% guardamos la estructura a disco.
save PCAprpalDaBa2 mods -APPEND
```