



**FACULTAD
DE INGENIERIA**

Universidad de Buenos Aires

75.19 – Teoría de Comunicaciones

Grupo 6

Trabajo Práctico

Paper: “Combining local face image features for identity verification”

Tutor: Dra. Araceli Proto

Alumnos:

Braithwaite, Diego Jonathan

83.446

Mercuri Guillemi, Marcos Agustín

83.624

Contenido

INTRODUCCION3

RECONOCIMIENTO DE IMÁGENES4

METODO PCA (Principal Component Analysis)5

BASE DE DATOS7

IMPLEMENTACION DE LA SOLUCION8

CONCLUSIONES.....12

BIBLIOGRAFIA13

INTRODUCCION

Durante los últimos años, el reconocimiento facial se ha convertido en una de las aplicaciones más estudiadas en campos como la biometría, el procesamiento de imagen o el reconocimiento de patrones. Una de las razones que ha llevado a este crecimiento son las necesidades cada vez mayores de aplicaciones de seguridad y vigilancia utilizadas en diferentes ámbitos.

Estos sistemas en su objetivo de determinar la identidad de una persona, pueden tener en cuenta o analizar muchas características. Estas características pueden ser tanto físicas, como sería el caso de las pupilas, la cara, el iris, etc., o se puede tratar de características psicológicas como serían los gestos.

Las aplicaciones basadas en el reconocimiento facial, presentan algunas limitaciones o inconvenientes como pueden ser:

- Orientación del rostro.
- Ruido.
- Iluminación (incluyendo interior / exterior).
- Expresión facial.
- Oclusión debido a objetos o accesorios tales como gafas de sol, sombreros...
- Vello facial.
- Envejecimiento.

Aun así algunas de estas barreras, se puede solucionar de una forma eficaz como sería el caso del envejecimiento. Para ello bastaría con una actualización periódica de la base de datos de la aplicación.

Debido a que las aplicaciones basadas en el reconocimiento facial no necesitan de la interacción del individuo, son aplicaciones de mucho interés para el entorno de video vigilancia, aeropuertos, comisarías, etc. De igual modo son aplicaciones necesarias y de utilidad en entornos controlados donde por ejemplo se pueda llevar a cabo el testeo de las aplicaciones antes de salir al mercado.

Durante este proyecto, tomando como referencia el paper seleccionado, *“Combining local face image features for identity verification”*, decidimos profundizar e investigar en las distintas técnicas de reconocimiento facial que se utilizan en la actualidad. Una de las más populares es PCA (Principal Component Analysis), en la cual nos basaremos para realizar el presente trabajo.

RECONOCIMIENTO DE IMÁGENES

Hay dos familias de técnicas de reconocimiento facial: técnicas basadas en la apariencia y técnicas basadas en modelos. Nos centraremos en los métodos basados en la apariencia.

El objetivo de este método es clasificar las imágenes en un nuevo subespacio vectorial, donde se entrena al sistema con diferentes caras, donde cada una de ellas es un punto en dicho subespacio.

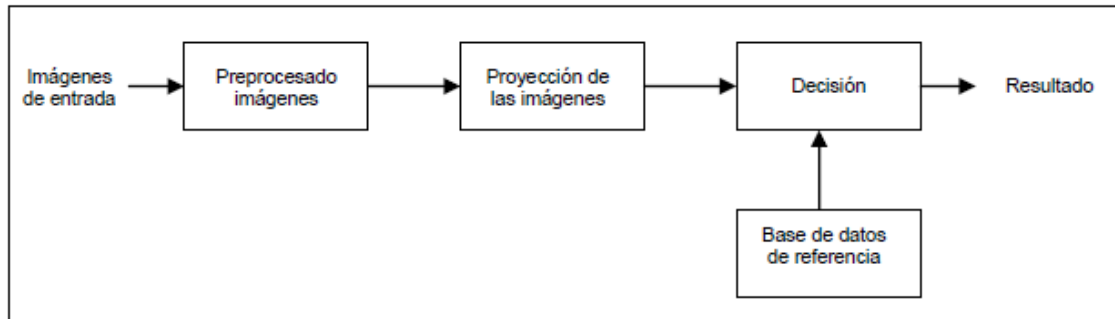
No analizaremos los sistemas basados en modelos pero podemos mencionar algunas ventajas de los basados en la apariencia frente a éstas:

- Más rápidos.
- Menor complejidad de implementación.
- No requieren de un conocimiento previo de las imágenes.
- Permiten una menor tamaño de las imágenes.

Entre las desventajas podemos citar:

- Están más afectados por cambios en la orientación o expresión de la cara
- Más dificultades frente a cambios en la iluminación

Métodos basados en la apariencia



Estudiaremos las siguientes 3 clasificaciones de imágenes para implementar el PCA.

Imágenes de entrenamiento: son utilizadas en el método PCA para conseguir las matrices de proyección.

Imágenes de referencia: son las que están almacenadas, son conocidas por el sistema y son utilizadas para saber si una imagen de entrada pertenece o no a algún sujeto registrado en el sistema.

Imágenes de test: son las que recibiremos en el sistema y serán evaluadas para saber si pertenecen a algún sujeto registrado en la base de datos.

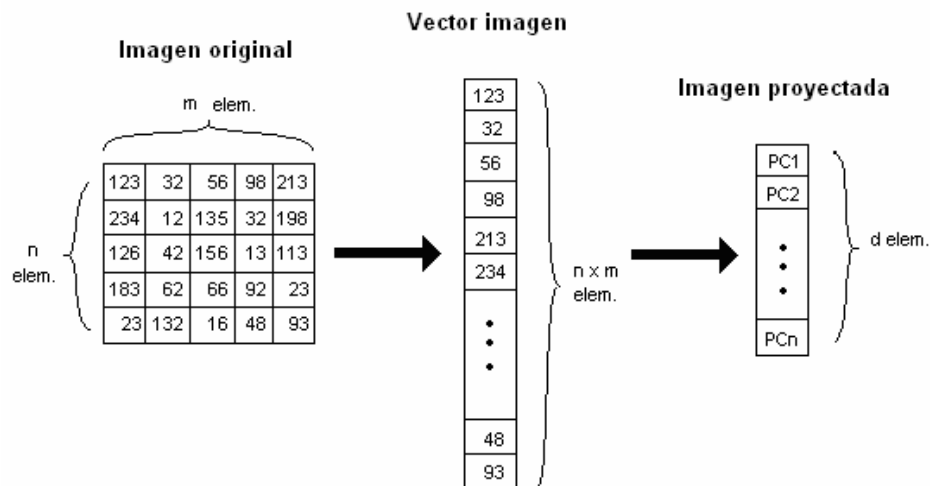
METODO PCA (Principal Component Analysis)

PCA es un algoritmo de reducción dimensional que permite encontrar los vectores que mejor representan la distribución de un grupo de imágenes. Permite representar una imagen de una cara usando una base que se ha conseguido a partir de muchas observaciones de diferentes caras.

El objetivo de este método consiste en representar una imagen en términos de un sistema de coordenadas óptimo reduciendo el número final de componentes que tendrá la imagen.

Esta base de transformación depende de las observaciones y se forma a partir de los d autovalores más significativos de la matriz de covarianza, los cuales representan las componentes principales que son más comunes en imágenes de diferentes caras.

La reducción dimensional realizada por PCA es equivalente al número de autovectores que se utilicen. Por lo tanto la imagen proyectada por PCA tendrá una dimensión de valor d , tal como se muestra en la siguiente imagen.



Estos autovectores representan las componentes principales que son más comunes en imágenes de diferentes caras, otra forma de verlo, es que son los ejes del subespacio donde proyectaremos las imágenes de caras humanas, los cuales se corresponden con las direcciones de máxima varianza de los datos.

La matriz de transformación, esta formada por los autovectores correspondientes a los d autovalores más significativos.

Otra de las cualidades de PCA, es que realiza una discriminación entre clases, es decir, todas las caras correspondientes a un mismo individuo están cerca entre si, mientras que las imágenes correspondientes a individuos diferentes están más alejadas. Esta propiedad es más destacable en las primeras componentes de las imágenes proyectadas.

PROCEDIMIENTO

Los pasos para la implementación de PCA son:

1. Se entrena al sistema con un set de imágenes de entrenamiento. Con dichas imágenes se obtiene la matriz de proyección W de la siguiente forma:

- a. Calcular la media del vector:

$$\underline{\underline{\mu}} = \sum_{i=1}^N \underline{x}_i$$

- b. Estimar la matriz de covarianza:

$$\underline{\underline{S_T}} = \sum_{i=1}^N (\underline{x}_i - \underline{\mu}) \cdot (\underline{x}_i - \underline{\mu})^T$$

- c. Calcular los autovectores y autovalores de S_T y generar W de forma que sus columnas son los autovectores correspondientes a los autovalores más significativos (con mayor energía):

$$\underline{\underline{S_T}} \Rightarrow \underline{\underline{W}} = (\underline{e}_1 \ \underline{e}_2 \ \cdots \ \underline{e}_M)$$

2. El sistema necesita de un set de imágenes de referencia las cuales son procesadas y proyectadas ($y = W \cdot x$) en el nuevo subespacio para después ser utilizadas para conocer si una imagen pertenece o no a un individuo registrado en el sistema.
3. Cuando las imágenes de referencia están correctamente proyectadas, se realiza el mismo procedimiento con las imágenes de test. Estas pueden pertenecer a personas que estaban incluidas entre las imágenes de referencia pero mostrando diferentes expresiones faciales o diferentes condiciones del entorno.
4. Finalmente se dirá que ha habido reconocimiento si se cumple alguna condición, en general se considera que hay un reconocimiento positivo si se cumple una condición: $d \Rightarrow y - y < d_{th}$, donde d_{th} es una distancia threshold.

BASE DE DATOS

Existen diferentes variantes de bases de datos de imágenes para realizar estos estudios. Podemos mencionar:

- Base de datos Olivetti (ORL Database)
- Base de datos XM2VTS
- Base de datos BANCA
- Base de datos YALE

Para el presente trabajo hemos utilizado la ORL Database.

Se trata de una base de datos formada por 40 individuos, a razón de 10 imágenes por individuo tomada entre 1992 y 1994. En esta base de datos se pueden encontrar tanto imágenes de hombres como de mujeres.

Como es posible observar en la siguiente figura, en esta base de datos se pueden encontrar imágenes con diferentes posiciones y expresiones faciales (ojos abiertos/cerrados, sonriendo o no,...), lo cual será una dificultad añadida para el sistema. Del mismo modo se puede observar como en algunos sujetos, se tendrán imágenes con distintos detalles faciales (gafas o no, bigote o no,...).

Por último decir que todas estas fotos han estado tomadas bajo unas condiciones de iluminación y de fondo de imagen controladas. El tamaño de todas estas imágenes es de 92x112 píxeles, con 256 niveles en la escala de grises para cada píxel.



IMPLEMENTACION DE LA SOLUCION

TECNOLOGIA

La presente aplicación se desarrolló en Java 1.6 utilizando las siguientes librerías de apoyo para manejar el álgebra que el método requiere y el tratamiento de imágenes PGM que la base ORL nos brinda:

- jai_codec-1.1.3-alpha.jar
- jmatharray.jar
- jmathio.jar
- jmathplot.jar

Así mismo para la interfaz de usuario se utilizaron los componentes de Swing.

ENTREGABLE

Se entrega un CD que contiene:

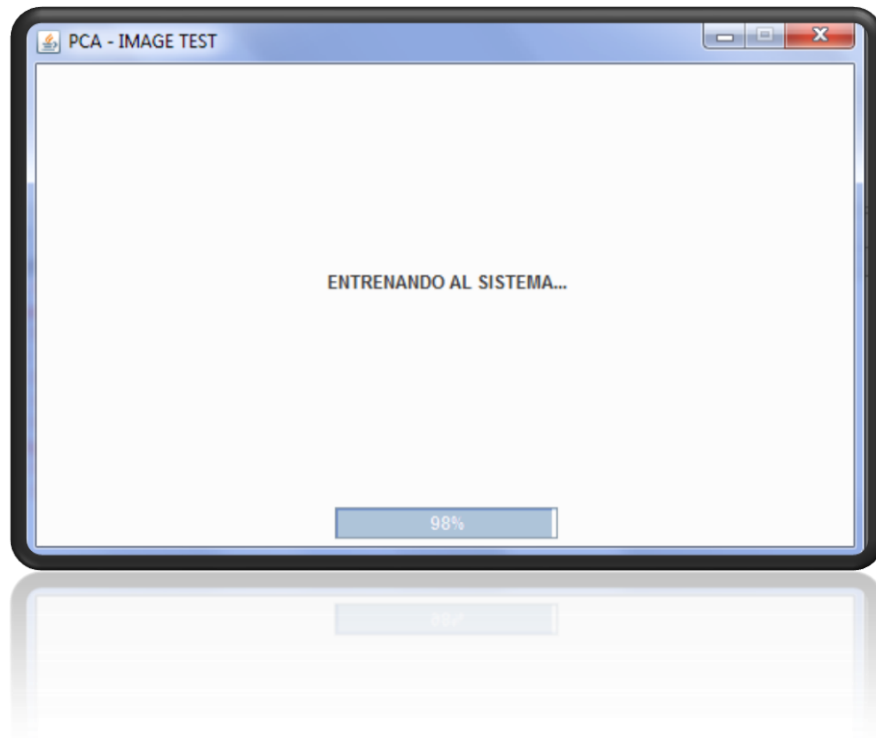
- Carpeta de imágenes (de entrenamiento, de referencia y de test)
- Archivo ejecutable: tp.bat
- Archivo jar que contiene la aplicación en sí.

EJECUCION DE LA APLICACIÓN

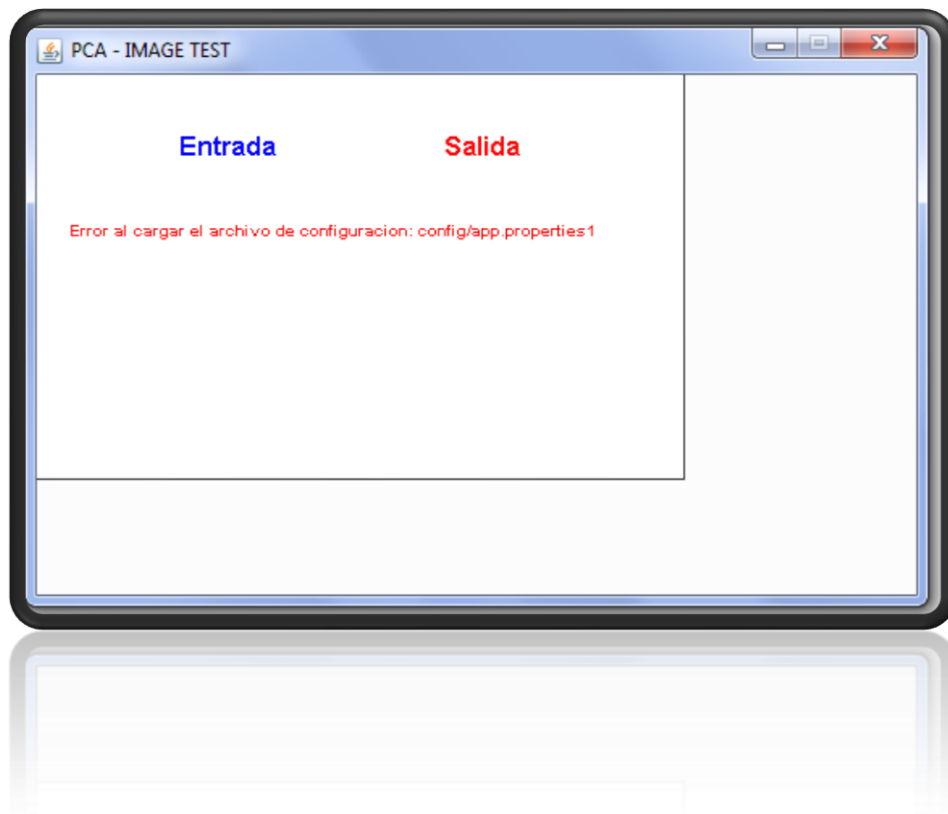
Se debe indicar en el archivo config/app.properties la ruta donde están las imágenes que requiere la aplicación.

Ejecutar el archivo tp.bat para inicializar la aplicación.

A continuación la aplicación se iniciará y en una primera instancia realizará las tareas de entrenamiento y carga en memoria de la base de datos.



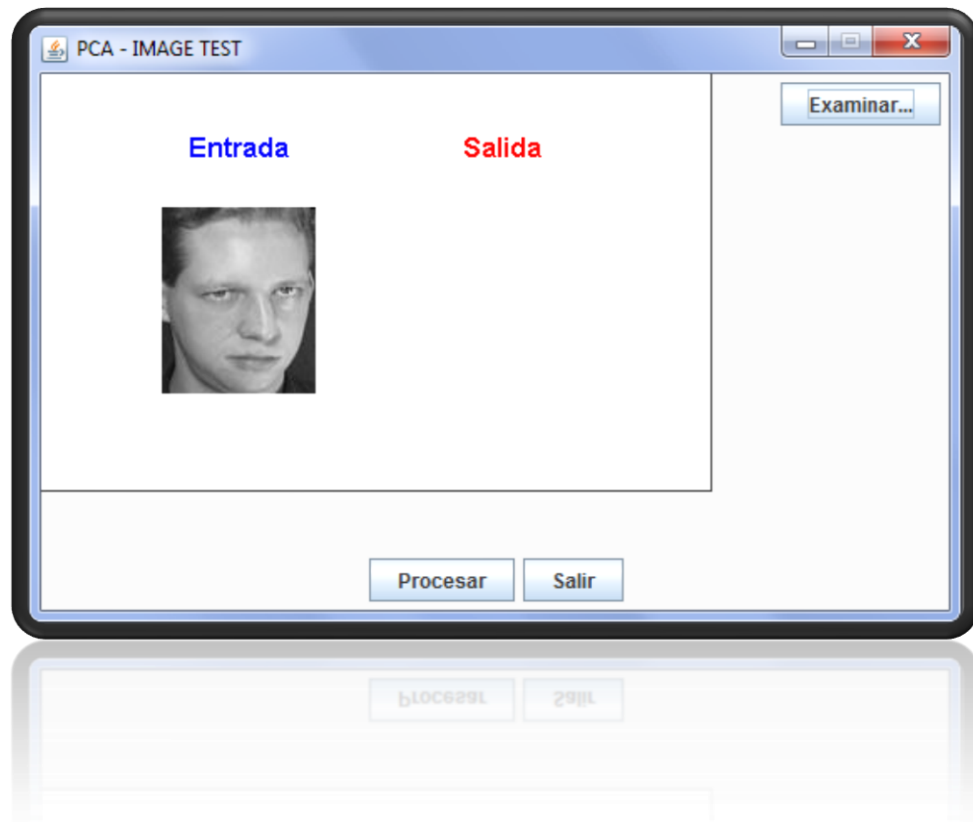
En caso que haya algún inconveniente con el inicio de la aplicación, como por ejemplo, que no encuentre la ruta configurada en config/app.properties (donde están las imágenes); o que no encuentre el archivo de configuración mencionado, se mostrará un mensaje de error y no se podrá seguir ejecutando la misma.



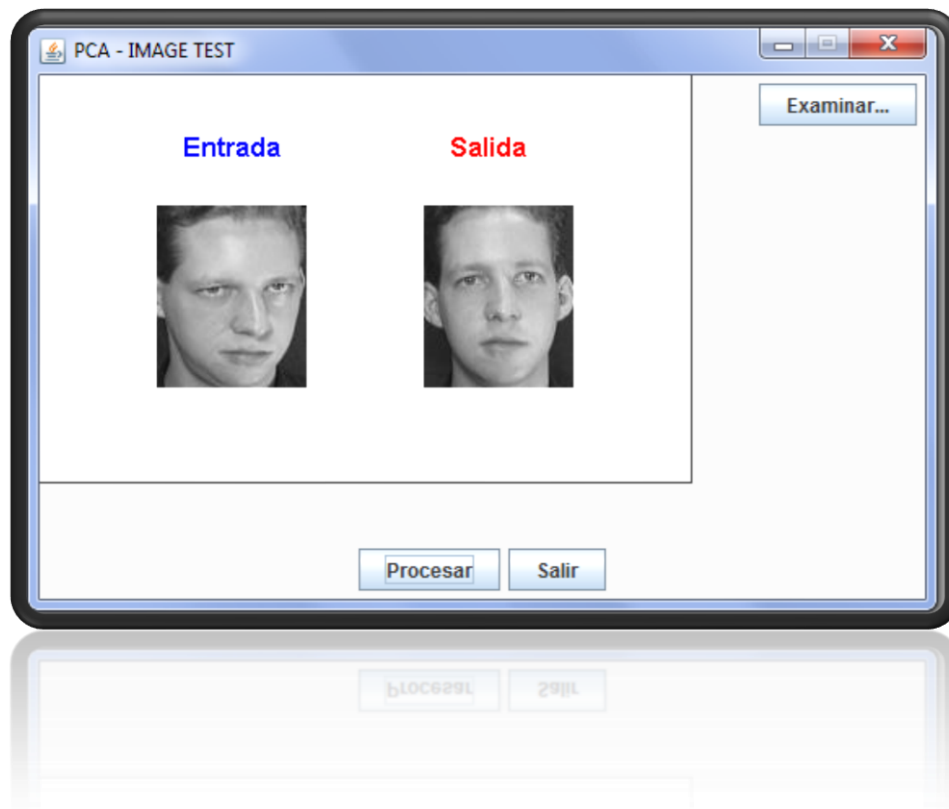
Una vez iniciada la aplicación hay que seleccionar una de las imágenes de la base de datos como imagen de test desde el botón Examinar. Para que la prueba tenga sentido, lo ideal es que la imagen de test no exista entre las imágenes de entrenamiento y en las de referencia.

Para esta entrega, hemos tomado la imagen 10 de cada individuo para realizar el test correspondiente. Las mismas se encuentran en la carpeta images/test y fueron eliminadas de las imágenes de entrenamiento y referencia.

Para la primera prueba, seleccionamos la imagen s1_10.pgm.



A continuación seleccionamos Procesar para que realice la búsqueda y arroje los resultados. En caso que no encuentre al individuo, se mostrará un mensaje descriptivo.



Para realizar la prueba de un individuo que no existe, eliminamos al s11 de las imágenes de entrenamiento y de referencia, dejando la imagen s11_10.pgm en la carpeta de test.



CONCLUSIONES

El motivo por el cual seleccionamos el paper de este trabajo práctico, fue porque era un tema no muy conocido por ambos y nos interesó aprender cómo funcionan estos sistemas.

Como experiencia podemos decir que a la hora de investigar sobre el tema de reconocimiento facial de imágenes, nos hemos encontrado con muchísima información en internet. Muchos métodos son los que se mencionan, como PCA (Principal Component Analysis), LPP (Locality Preserving Projections), LDA (Linear Discriminant Analysis), DCT (Discrete Cosine Transform). Seleccionamos PCA debido a que, si bien no es el más efectivo, es el más popular.

A lo largo del análisis y desarrollo nos encontramos con dificultades extras como el manejo de álgebra para resolver el método. Creemos que también nos ayudó a refrescar muchos conceptos que vimos en el inicio de la carrera y que es bueno saber que no deben quedar en unos apuntes, sino más bien, que van más allá, ya que estos métodos se basan exclusivamente en ellos.

Podemos afirmar que la técnica PCA funciona en óptimas condiciones para la base de datos ORL, ya que todas las pruebas que ejecutamos dieron los resultados esperados.

En resumen, podemos decir que de todos los métodos de reconocimiento e identificación facial, el PCA es el más popular y el menos complejo de implementar. En nuestro caso, a pesar de las complicaciones extraprogramáticas (el álgebra del método), pudimos implementarlo en java satisfactoriamente. Y las pruebas que se realizaron con la base ORL fueron todas según lo esperado.

BIBLIOGRAFIA

Mauricio Delbracio, Matias Mateu, "Trabajo Final de Reconocimiento de Patrones: Identificación utilizando PCA, ICA y LDA.

Algebra. 2004-2005. Ingenieros Industriales. Departamento de Matemática Aplicada II. Universidad de Sevilla.

Roger Gimeno Hernández, "ESTUDIO DE TÉCNICAS DE RECONOCIMIENTO FACIAL"

Lindsay I Smith, "A tutorial on Principal Components Analysis".

Wikipedia, "Principal component analysis".