

Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales
Departamento de Computación

Análisis de un Reconocedor de Género Basado en
Técnicas Lineales

Trabajo de cátedra:
Introducción al Procesamiento Digital de Imágenes
Departamento de Computación
Facultad de Ciencias Exactas
Universidad de Buenos Aires

Florencia S. Iglesias

Profesor: Julio César Jacobo Berllés
JTP: María Elena Buemi
Buenos Aires, 2012

Análisis de un Reconocedor de Género Basado en Técnicas Lineales

Resumen Los reconocedores de género de personas a partir de imágenes de caras, que se encuentran basados en técnicas lineales de clasificación han cobrado importancia en los últimos años debido a su simplicidad y bajos requerimientos computacionales. Nuestro trabajo consistió en analizar uno de estos reconocedores, el cual combina las técnicas de LDA, PCA y Naive Bayes. Hemos entrenando y probado el reconocedor con dos bases de caras de personas diferentes; uno de ellos constituido por fotos de sólo 8 personas y el otro compuesto por 200 individuos; ambos datasets conformados por la misma cantidad de hombres que de mujeres.

Keywords: Reconocimiento de género, PCA, LDA

1 Introducción

El reconocimiento de género de una persona a partir de una imagen de su cara es uno de los problemas de clasificación demográfica estudiados en los últimos años.

Uno de los enfoques desde el cual se ha abordado el tema se basa en extraer un conjunto de datos discriminantes de cada imagen por medio de técnicas lineales, y utilizar dicha información para realizar la clasificación. Entre las técnicas lineales más conocidas se encuentran LDA (Linear Discriminant Analysis), PCA (Principal Components Analysis) e ICA (Independent Components Analysis). Juan Bekios-Calfa, José M. Buenaposada y Luis Baumela [1] presentan varios clasificadores basados en este enfoque. En este trabajo se busca analizar la performance de uno de ellos, el cual combina LDA y PCA para realizar la extracción de características de las imágenes, que serán útiles para efectuar la clasificación.

Este trabajo está organizado de la siguiente manera: en la sección 2, se presenta una síntesis de trabajos realizados sobre este tema. En la sección 3, se describe el algoritmo utilizado para el reconocimiento de género, dando énfasis en las técnicas de LDA y PCA, y su combinación en el algoritmo principal. En la sección 4, se detallan los conjuntos de imágenes utilizados y los experimentos realizados, y en sección 5 se muestran las observaciones y resultados de los mismos. Finalmente, en la sección 6 se presentan las conclusiones obtenidas.

2 Estado del Arte

La construcción de un reconocedor de género de una persona a partir de una imagen de su cara, fue abordado desde diferentes enfoques. Moghaddam

y Yang [2] presentaron un reconocedor del género basado en Support Vector Machines (SVM) y en Radial Basis Function (RBF). Según los autores, presenta una tasa de acierto superior a las obtenidas por reconocedores de patrones tradicionales (Lineales, Cuadráticos, Fisher Linear Discriminant) y por los que utilizan técnicas más modernas, como los basados sólo en Radial Basis Function (RBF). Moghaddam y Yang experimentaron con 1775 imágenes de la base de datos FERET, y obtuvieron una tasa de acierto del 96.6 %.

En general, las técnicas utilizadas para construir los mejores clasificadores requieren poseer un sistema de alta performance. Es por ello que se está investigando la construcción de aquellos basados en técnicas que requieran menos recursos. A continuación se presentan algunos trabajos realizados desde este enfoque:

Baluja y Rowley [3] combinaron la técnica de AdaBoost con clasificadores débiles basados en comparaciones de la escala de grises de los píxeles. El mismo fue probado con un subconjunto de imágenes del dataset FERET. Éste es varias órdenes de magnitud más rápido que el clasificador basado en SVM presentado por Moghaddam y Yang [2].

Juan BekiosCalfa, Jose M. Buenaposada y Luis Baumela [1] presentaron clasificadores basados en técnicas lineales. Los mismos fueron probados utilizando varios conjuntos de imágenes, entre ellas la base de imágenes FERET utilizada por los otros trabajos mencionados. Las tasas de aciertos obtenidas por los clasificadores presentados en [1] son comparables con los resultados obtenidos con los mejores clasificadores basados en SVM [2], presentando la ventaja de necesitar menos recursos.

3 Sistema de Reconocimiento del Género

3.1 Linear Discriminant Analysis (LDA)

Dados $x_1 \dots x_n$ datos d-dimensionales etiquetados en c clases, LDA permite obtener una proyección de los datos en un espacio de menor o igual dimensión que los datos entrantes, con el fin de que la separabilidad de las clases sea la mayor posible. La base del espacio transformado, W , se obtiene maximizando la siguiente función objetivo:

$$J(W) = \frac{|W^t * S_B * W|}{|W^t * S_W * W|} \quad (1)$$

donde S_B es la matriz de dispersión inter-clase y S_W es la matriz de dispersión intra-clase.

El W óptimo puede obtenerse resolviendo la siguiente ecuación generalizada de autovalores:

$$S_B * W = S_W * W * D \quad (2)$$

donde D es la matriz diagonal de autovalores y los autovectores son las columnas de W .

Un problema presente al intentar aplicar LDA a un conjunto de datos es que la matriz S_W puede ser singular y, por lo tanto, W no siempre puede computarse. Una solución alternativa que permite obtener resultados similares consiste reemplazar en la fórmula anterior, S_W por la matriz de covarianza de los datos de entrada [4].

3.2 Principal Components Analysis (PCA)

PCA permite obtener la proyección de una imagen sobre un subespacio de autovectores. Dado un set de imágenes $X = \{X_1, X_2, \dots, X_n\}$, la base de autovectores correspondientes a aplicar PCA sobre X se obtiene calculando los autovectores de la matriz de covarianza de X ; la misma se calcula realizando el siguiente cálculo:

$$C = A * A^t \quad (3)$$

donde A es el resultado de restar a cada elemento de X la imagen promedio de X , llamémosla I , es decir

$$A = \{\Phi_1 \dots \Phi_n\} \quad (4)$$

donde

$$\Phi_i = X_i - I \quad (5)$$

para cada $1 \leq i \leq n$

3.3 LDA + PCA en el Reconocedor de Género

Aplicar LDA a un conjunto de imágenes puede provocar problemas computacionales, debido al tamaño de las mismas. Una técnica utilizada para evitar estos problemas es aplicar PCA al conjunto de datos de entrada, manteniendo del subespacio obtenido sólo los autovectores asociados a autovalores no nulos, y luego aplicar LDA. En esta técnica la realización de PCA provoca la reducción de la dimensión de los datos que serán la entrada de LDA. Esta combinación de PCA y LDA permite obtener resultados similares a los que se obtendría si se aplicara solamente LDA [5].

El sistema de reconocimiento de género presentado por Juan BekiosCalfa, Jose M. Buenaposada y Luis Baumela [1], utiliza esta técnica de PCA+LDA para manipular el conjunto de imágenes de entrada y, con la información obtenida, realizar el entrenamiento de un clasificador Naive Bayes, según el Algoritmo 1. En el mismo, X es la matriz que tiene en cada columna una imagen vectorizada y L es un vector con el género al que corresponde cada columna de X , K es la cantidad de autovectores de la base de PCA que no se descartan y R es la tasa de acierto del mismo.

Algoritmo 1**Entrada:** X, L, K **Salida:** R Dividir X en n folds $F = \{X_1, X_2, \dots, X_n\}$ Dividir L en los n folds correspondientes $E = \{L_1, L_2, \dots, L_n\}$ **para** $i = 1$ hasta n **hacer** $X_{test} = X_i$ // Se testea con el fold i $L_{test} = L_i$ $X_{entrenamiento} = F - \{X_i\}$ // Se entrena con el resto de los folds $L_{entrenamiento} = E - \{L_i\}$ $[P_PCA, I] = PCA(X_{entrenamiento})$ // P_PCA es la base de autovectores;// I es la cara promedio vectorizadaSe le asigna a B los primeros K autovectores de P_PCA $Y = B^T * (X_{entrenamiento} - \{I, I, \dots, I\})$

// Se proyecta las imágenes de entrenamiento en el subespacio de PCA

 $P_LDA = LDA(Y)$ $Z = P_LDA * Y$ // Se proyectan los datos en el subespacio de LDA $C = entrenarClasificadorBayesiano(Z, L_{entrenamiento})$ $Z_{test} = PLDA * B^T * (X_{test} - \{I \dots I\})$ $R = clasificar(C, Z_{test}, L_{test})$ // R es la tasa de acierto**fin para**

Para seleccionar el K óptimo, el reconocedor implementado por Bekios-Calfa, Buenaposada y Baumela, realiza N iteraciones del Algoritmo 1, donde N es la cantidad de píxeles de las imágenes, y en cada iteración se prueba la performance del Algoritmo 1 con diferentes K , tomando, en la iteración i , $K = i$ (con $1 \leq i \leq N$). Ese reconocedor permite obtener el K óptimo y su performance asociada, medida en tasa de acierto.

En este trabajo hemos implementado una variación de ese reconocedor; la misma consiste en tomar el siguiente N :

$$N = \min(\text{cantidad de píxeles, cantidad de caras}) \quad (6)$$

Se decidió realizar esta modificación debido a que la cantidad de píxeles de la imagen suele ser mayor que la cantidad de caras de la base, y por lo tanto tomando este nuevo N estamos disminuyendo la cantidad de ciclos que realiza el reconocedor.

4 Experimentos

Se realizaron experimentos para probar el reconocedor de género explicado en 3.3; el algoritmo fue entrenado con datasets de caras de personas, analizando su performance y observando como varía la misma en función de la cantidad de

autovectores del subespacio de PCA elegida. Utilizamos 2 datasets, que llamamos A y B y cuya descripción damos a continuación:

- A: dataset de caras perteneciente a la Universidad de Cambridge[6]¹. El mismo se encuentra conformado por imágenes en blanco y negro de 40 personas diferentes: 4 de ellas mujeres y 36 hombres. De cada individuo se poseen 10 imágenes. Todas ellas fueron tomadas contra un fondo oscuro y corresponden a fotos frontales de las caras de las personas, con una tolerancia de un cierto movimiento lateral. Para algunos individuos, se varió la iluminación, las expresiones faciales (con ojos abiertos u ojos cerrados, sonriendo o no) y los detalles faciales (con gafas o sin gafas). El tamaño de las mismas es de 92 x 112 píxeles. Para nuestro experimento sólo se utilizaron las 4 mujeres y 4 de los hombres, debido a que se deseaba probar el clasificador utilizando una base con la misma cantidad de individuos de cada género; se usó sólo una imagen por individuo seleccionado.
- B: dataset de caras FEI[7]². El mismo está conformado por imágenes de caras de 200 personas diferentes. La cantidad de hombres y de mujeres cuyas caras pertenecen al dataset es exactamente la misma y es igual a 100. Nosotros hemos utilizado el subconjunto de imágenes de la base FEI conformado sólo por imágenes de caras frontales alineadas. Este subconjunto está compuesto por dos caras por personas: una sonriendo y la otra no. En nuestro experimento se utilizaron las imágenes de caras sin sonrisa.

Previamente a realizar el experimento, se recortaron las imágenes de ambas bases y se redimensionaron las mismas usando OpenCV 2.3.1 face detector ³, que se encuentra basado en el algoritmo de Viola Jones[8]. En la Fig. 1. se pueden observar las 8 caras elegidas de A, luego de haber sido recortadas y redimensionadas a 67x67. La Fig. 2. es una muestra de las caras de B, recortadas y redimensionadas a 162x193.



Fig. 1. Caras de la base de imágenes de Cambridge, luego de haber sido recortadas y redimensionadas

¹ Disponible online en http://www.cl.cam.ac.uk/Research/DTG/attarchive/pub/data/att_faces.zip

² Disponible online en <http://fei.edu.br/~cet/facedatabase.htm>

³ <http://opencv.willowgarage.com>.



Fig. 2. Muestra de la base de imágenes FEI, luego de haber sido recortadas y redimensionadas

Para cada conjunto de datos, se ejecutó el algoritmo del clasificador del género 40 veces, utilizando 5 *cross validation* en todos los casos.

5 Resultados

En esta sección se presentan los resultados obtenidos al probar la performance del reconocedor de género, entrenando el mismo 40 veces con los dos conjuntos de datos explicados en la sección anterior. La mejor tasa de acierto obtenida con A es 80 % y la peor es 50 %, por lo cual las tasas de errores varían considerablemente entre las diferentes ejecuciones; las mejores tasas de acierto se obtuvieron trabajando con un subespacio de PCA de 3 o 4 autovectores. Por otro lado, al entrenar el sistema usando B, se obtuvo una tasa de acierto máxima de 86 % y una tasa de acierto mínima de 81.4 %, por lo cual las tasas de errores se encuentran entre el 20 % y el 14 %; en todas las ejecuciones las máximas tasas de acierto se obtuvieron considerando un subespacio de PCA conformado por una cantidad de autovectores entre 26 y 71. La Tabla 1. muestra las 3 mejores y 3 peores tasas de acierto obtenidas con cada dataset, junto con la dimensión del subespacio de PCA elegida en cada caso.

Tabla 1. Tasas de acierto, junto con la cantidad óptima de autovectores del subespacio de PCA elegidos, entre ()

A	B
80.0 % (3)	86.0 % (71)
71.0 % (4)	85.4 % (26)
70.2 % (3)	84.5 % (56)
60.0 % (5)	82.7 % (24)
56.7 % (6)	81.6 % (16)
50.4 % (4)	81.4 % (34)

En cada ejecución del algoritmo, al ir aumentando la dimensión del subespacio de PCA elegida antes de calcular LDA, también varían las performances intermedias obtenidas:

Al utilizar la base de datos de Cambridge, en cada ejecución, las tasas de acierto máxima son marcadamente superior al resto de las tasas. Por otro lado, si se toman dos o más ejecuciones del algoritmo donde los porcentajes de acierto máximos presenten diferencias, al comparar los resultados intermedios obtenidos tomando la misma cantidad de autovectores de la base de PCA, se pueden observar diferencias considerables en las performances intermedias. En la Fig. 3., pueden observarse los resultados obtenidos en dos pruebas realizadas con este conjunto.

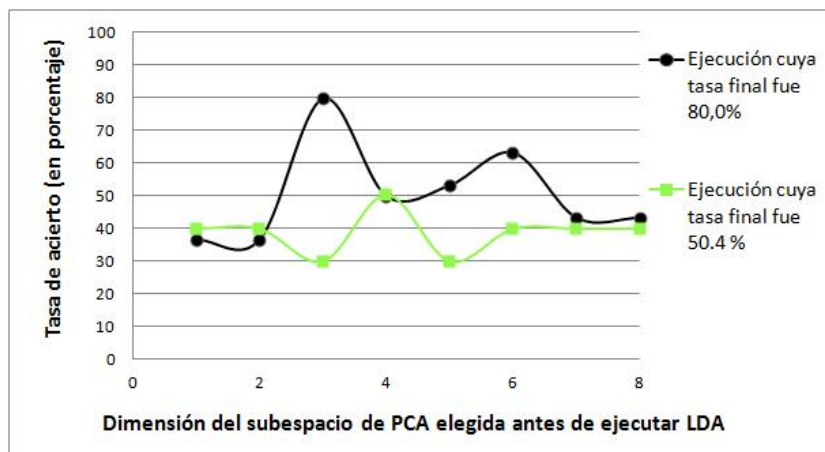


Fig. 3. Performance intermedias en dos ejecuciones del algoritmo, utilizando el dataset A. Las ejecuciones graficadas corresponden a las que brindaron la mejor y la peor tasa final.

Al utilizar el otro conjunto de imágenes, la diferencia entre la tasa de acierto máxima y la mínima obtenidas en cada ejecución, se encuentra entre el 20 % el 30 %. Sin embargo, para una ejecución del algoritmo de reconocimiento de género, al considerar la tasa de acierto calculada para una cierta dimensión del subespacio de PCA y compararla con la obtenida con un subespacio de PCA formado por un autovector menos (o uno más) puede verse poca diferencia entre los valores. Además la tendencia que presentan las tasas de acierto intermedias, es la misma para todas las ejecuciones, como puede verse al comprar las Fig. 4. y Fig. 5.: tomando un subespacio de PCA de tamaño 1 se obtiene la tasa intermedia mínima (alrededor del 60 %), sigue una fase de variaciones pequeñas

8

de las tasas (entre las cuales se encuentra la tasa máxima), y finalmente tomando una dimensión de la base de PCA igual o mayor a 175, el valor del porcentaje de acierto se estabiliza tomando un valor fijo, entre el 70% y el 80%.

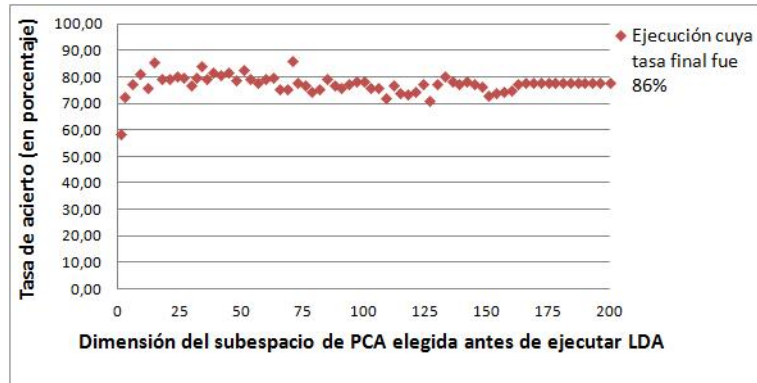


Fig. 4. Performance intermedias en la ejecución del algoritmo que brinda la mejor tasa de acierto, utilizando el dataset B.

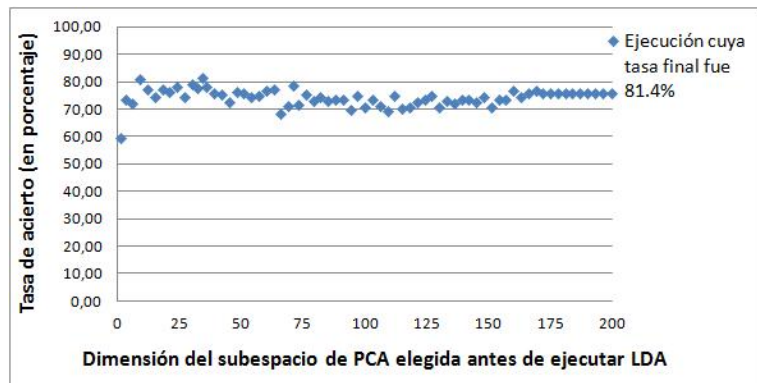


Fig. 5. Performance intermedias en la ejecución del algoritmo que brinda la peor tasa de acierto, utilizando el dataset B.

6 Conclusiones

Hemos reducido la cantidad de ciclos en el reconocedor de género propuesto por BekiosCalfa, Buenaposada y Baumela [1], que utiliza PCA y LDA para extraer

datos discriminantes de las imágenes y posteriormente utilizar dicha información para realizar la clasificación de las caras.

Utilizando el conjunto de datos pequeño (8 imágenes), se alcanzó una tasa de acierto máxima de 80 %, lo que implica que se puede obtener un clasificador del género de una persona con una alta performance entrenando el mismo con un dataset con pocas imágenes de caras. Al utilizar el conjunto de mayor cantidad de imágenes, la tasa de acierto aumentó a valores cercanos al 86 %, tomando un subespacio de PCA de dimensión 71; la obtención de un porcentaje de error cercano al 14 % lograda con este dataset, permite observar la buena performance que puede alcanzar el clasificador.

Al entrenar el reconocedor utilizando el conjunto con mayor cantidad de caras, puede observarse que independientemente de la cantidad de veces que se realiza el entrenamiento, las tasas intermedias (obtenidas al ir aumentando la cantidad de autovectores que conforman la base del subespacio de PCA elegida antes de aplicar LDA) presentan siempre la misma tendencia: con un subespacio de PCA de dimensión 1 se obtiene la tasa mínima; luego sigue una fase de variaciones pequeñas de las tasas, entre las cuales se encuentra la tasa máxima; y finalmente, tomando una dimensión de la base de PCA igual o mayor a 175, el porcentaje de acierto se mantiene fijo en un valor que se encuentra entre el 70% y el 80%. Por otro lado, al entrenar y probar la clasificador con el dataset pequeño, se obtuvo gran variabilidad en las tasas y en la cantidad de autovectores de la base del subespacio de PCA que se debe tomar.

Referencias

1. José M. Buenaposada y Luis Baumela Juan BekiosCalfa, "Revisiting linear discriminant techniques in gender recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 858–864, Apr. 2011.
2. B. Moghaddam y M.H. Yang, "Learning gender with support faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 707–711, May 2002.
3. S. Baluja and H.A. Rowley, "Boosting sex identification performance," *Int'l J. Computer Vision*, vol. 71, pp. 111–119, Jan. 2007.
4. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, 1990.
5. J. Yang y J. y. Yang, "Why can lda be performed in pca transformed space?," *Pattern Recognition*, vol. 36, pp. 563–566, 2003.
6. F. Samaria y A. Harter, "Parameterisation of a stochastic model for human face identification," *2nd IEEE Workshop on Applications of Computer Vision*, Dec. 1994, Sarasota (Florida).
7. C. E. Thomaz and G. A. Giraldi, "A new ranking method for principal components analysis and its application to face image analysis," *Image and Vision Computing*, vol. 28, no. 6, pp. 902–913, June 2010.
8. Paul Viola and Michael J. Jones, "Robust real-time face detection," *Int. J. Comput. Vision*, vol. 57, no. 2, pp. 137–154, May 2004.