

Eigenespacios de belleza paramétricos como máquina calificadora

Ricardo Solano Monje¹, Nayeli Joaquinita Meléndez Acosta²,
Sergio Juárez Vázquez¹, Homero Vladimir Ríos Figueroa³

Universidad del Istmo, Campus Sto. Domingo Tehuantepec, Oax.,
México

² Universidad del Istmo, Campus Ixtepec, Oax.,
México

³ Universidad Veracruzana, Xalapa, Ver.,
México

rsolanomonje@gmail.com, sjuarez@bianni.unistmo.edu.mx,
nayelimelendez@gmail.com, hrios@uv.com.mx

Resumen. Este artículo presenta una máquina calificadora de belleza facial. El concepto de Máquina calificadora se debe a Eisenthal et al [4]. La máquina calificadora es implementada haciendo uso de un Eigenespacio de belleza paramétrico, tomando como ensamble vectores de distancias entre puntos de referencia de rostros. La máquina calificadora propuesta parte de la selección asistida de un conjunto de características faciales predefinidas, a partir de estas características faciales y de la posición de los puntos de referencia relativos a la imagen facial, se calcula un conjunto de distancias entre dichas características, se localizan 66 puntos de referencia, con estos puntos se construye una malla o máscara. El objetivo final de la máscara es formar un vector de distancias o longitudes entre los puntos de referencia que definen las características faciales. Estos vectores de distancias son utilizados para contruir un Eigenespacio. Las curva paramétrica acarrear la calificación de belleza de cada rostro del ensamble. La forma más simple de determinar la calificación de un rostro es proyectarla al Eigenespacio y buscar el vector más cercano. Mostramos cómo esto es suficiente para dar una calificación acertada a un rostro de entrada original.

Palabras clave: máquina calificadora, belleza facial, eigenespacios.

1. Introducción

La belleza ha captado la atención de los seres humanos desde que inicio la humanidad. La gente usa diferentes formas para que sus rostros se vean más atractivos, tales como maquillaje y cirugía cosmético. Para ejemplo, utilizamos bases para cambiar el tono de la piel, usar corrector para arrugas y pecas [1].

Esta noción del atractivo facial junto con la capacidad para evaluar de forma fiable y automáticamente la belleza facial de una imagen ha motivado este trabajo.

El componente clave de nuestro enfoque es un motor calificador de la belleza entrenado a partir de un conjunto de imágenes de rostros con el acompañamiento de una calificación de belleza proporcionado por un conjunto de evaluadores humanos.

El contenido de este artículo está organizado en cinco secciones de la siguiente manera: la segunda sección trata las bases teóricas de la aproximación de Eigenespacios. En la tercera sección se explica el proceso de embellecimiento propuesto por Leyvand et al [3], particularmente la extracción de los puntos de referencia que serán los parámetros de belleza de la máquina calificadora. La cuarta sección explica la propuesta de este artículo. Finalmente en la quinta sección se muestran las conclusiones y trabajo futuro.

2. La aproximación de eigenespacios

El reconocimiento de objetos comprende al menos tres etapas, la primera la creación del Eigenespacio a partir del ensamble de imágenes disponibles, la segunda la representación de las imágenes del ensamble en el Eigenespacio y por último la clasificación de imágenes originales, previa representación en el Eigenespacio.

2.1. La transformada K-L y reducción de dimensionalidad

El término Eigenespacio Universal se utiliza en la aproximación de apariencia para denotar el modelo que representa a un conjunto fijo de objetos. Tal modelo se genera a partir del ensamble de imágenes disponibles de los objetos a reconocer. Dicho modelo es representado por un subespacio particular llamado Eigenespacio. Para obtener el Eigenespacio a partir del ensamble, se aplica la transformada Karhunen-Loève abreviada K-L. La transformada K-L es un método bien estudiado para la compresión de información. En el caso discreto, donde la información es representada por un conjunto de vectores-imagen llamado ensamble, la transformada Karhunen-Loève minimiza el error cuadrático medio inducido al tomar solo unos términos de la expansión del total de vectores ortonormales necesarios para reconstruir el conjunto de vectores del ensamble. El conjunto de vectores ortogonales que se obtienen al aplicar la transformada K-L al ensamble es frecuentemente llamado componentes principales y la aplicación de la transformada K-L, análisis de componentes principales. Estos vectores capturan más varianza estadística que cualquier otra base. En la fase de creación del modelo –Eigenespacio- se obtienen M imágenes de $S \times R = N$ pixeles de resolución que forman el ensamble B

$$B = \{I_1, I_2, \dots, I_M\}$$

Típicamente la imagen media \bar{x} del ensamble

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M I_i \quad (1)$$

es sustraída de cada imagen del ensamble resultando en el conjunto

$$X = \{I_1 - \bar{x}, I_2 - \bar{x}, \dots, I_M - \bar{x}\}$$

El conjunto X puede ser visto como matriz de $N \times M$. Esto nos permite crear la matriz de covarianza $\Omega = XX^T$ que caracteriza la distribución de las M imágenes en \mathfrak{R}^N . Las componentes principales del ensamble se obtienen descomponiendo la matriz de covarianza Ω en vectores y valores singulares ¹.

Sean \bar{u}_i y λ_i los eigenvectores y eigenvalores de Ω respectivamente, entonces el Eigenespacio o las componentes principales viene dado por

$$U = [\bar{u}_1, \bar{u}_2, \dots, \bar{u}_M]$$

U se utiliza como matriz de transformación de espacios, del espacio de la imagen al espacio de baja dimensión.

Una vez que la base ha sido creada cada uno de los vectores-imagen $x_i \in X$ se proyectan al subespacio de menor dimensión generado por U , por medio de

$$y_i = U^T x_i \quad i = 1 \dots M$$

La forma más simple de determinar la clase de una imagen original es proyectarla al Eigenespacio y buscar el vector más cercano.

3. Proceso de embellecimiento propuesto por Leyvand

Partiendo de la imagen facial de entrada, se determinan una variedad de características faciales predefinidas. Las características faciales -8 en total- están formadas por puntos de referencia y corresponden a 2 cejas, 2 ojos, labio superior, labio inferior, contorno de la nariz y contorno del rostro. A partir de las características faciales y de la posición de los puntos de referencia se calcula un conjunto de distancias entre dichas características. Leyvand et al en [3]. localiza 84 puntos de referencia, mientras que Eisenthal et al. en [4] localiza 37 puntos de referencia. Con estos puntos se construye una malla o máscara. El objetivo final de la máscara es formar un vector de distancias o longitudes entre los puntos de referencia que definen las características faciales.

El trabajo de Leyvand et al. obtiene 234 distancias que se almacenan en un vector. El hecho de que el grupo de Eisenthal et al. localice 37 puntos de referencia y el

¹ Obsérvese que Ω está en $\mathfrak{R}^{N \times N}$. En [2] Turk et al proponen una forma eficiente de obtener los eigenvectores de Ω a partir de $\Omega' = X^T X$. En este caso Ω' está en $\mathfrak{R}^{M \times M}$.

grupo de Leyvand et al. localice 84 tiene una explicación sencilla, de hecho, Leyvand et al. usan la máquina calificadora de Eissenthal et al., pero mientras que Eissenthal et al. persiguen en su trabajo calificar la belleza de un rostro, Leyvand et al. realiza la modificación y embellecimiento de un rostro. Ya que la modificación de la imagen de entrada se llevará a cabo con transformaciones *warping*, se requieren más puntos para realizar una transición suave hacia la imagen de salida, por ello el uso de más puntos de referencia y la construcción de la malla respectiva.

El objetivo de la máscara es la creación de un vector de distancias. Cada vector es de dimensión 234, y para hacerlo invariante a escala, se normaliza con la raíz cuadrada del área del rostro. Sin embargo, es claro que este tipo de normalización puede cambiarse según nuestros intereses, por ejemplo, podría normalizarse dividiendo cada entrada del vector entre la distancia mayor de dicho vector.

Estas distancias que se obtienen de cada rostro definen un punto o vector en un espacio multidimensional -un espacio de rostros particular-. Debe observarse que este espacio de rostros es muy simple pues es meramente un conjunto de vectores con cierto orden de prioridad. Es en este espacio donde se busca el punto más cercano al rostro de entrada, que corresponda al rostro más atractivo del conjunto de rostros atractivos. Esta búsqueda es ponderada por belleza pero aun así el esquema es demasiado simple.

El espacio de rostros está formado por los vectores de distancias -obtenidos de las máscaras- de un conjunto de 92 rostros bellos, previamente seleccionados. Estos 92 rostros bellos fueron calificados por jueces externos al sistema. Las mallas de los 92 rostros junto con su calificación forman el espacio de rostros particular que se define en [5]. En este sentido, y como puede apreciarse, no estamos frente a un espacio de rostros sofisticado.

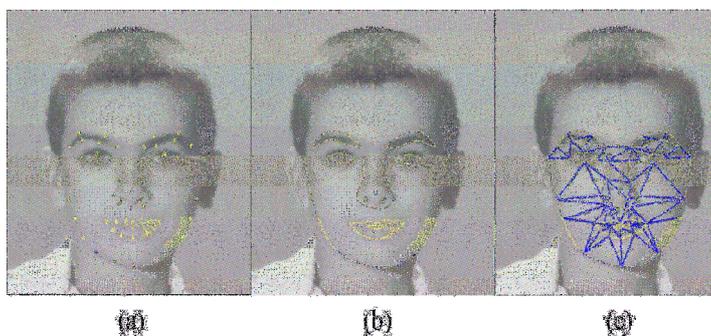


Fig. 1. (a) Los 66 puntos de referencia. (b) Las características faciales. (c) Las 89 distancias entre puntos de referencia

El componente clave del sistema de embellecimiento durante la búsqueda en el espacio de rostros bellos, es la máquina de aprendizaje que califica de forma automática la belleza facial. La máquina calificadora de Leyvand utiliza la técnica SVM Support Vector Machine.

4. Eigenespacios de belleza paramétricos

La propuesta de este artículo plantea el uso de Eigenespacios como los desarrollados por Nayar et al. en [6] para hacer una función calificadora de belleza con Eigenespacios, tomando como ensamble los vectores de puntos de referencia de rostros bellos y feos.

En nuestro caso y para fines explicativos, se considera que partiendo del rostro de entrada se extrae un total de 66 puntos de referencia de manera manual como se muestra en la Fig. 1. Al igual que en [3], los puntos de referencia son localizados en los contornos de las 8 diferentes características faciales: las dos cejas, los dos ojos, el labio inferior y superior de la boca, la nariz y límites del rostro. La posición de cada uno de los 66 puntos de referencia extraídos son usados para construir la máscara. La máscara consiste de 89 distancias y éstas forman el vector de distancias.

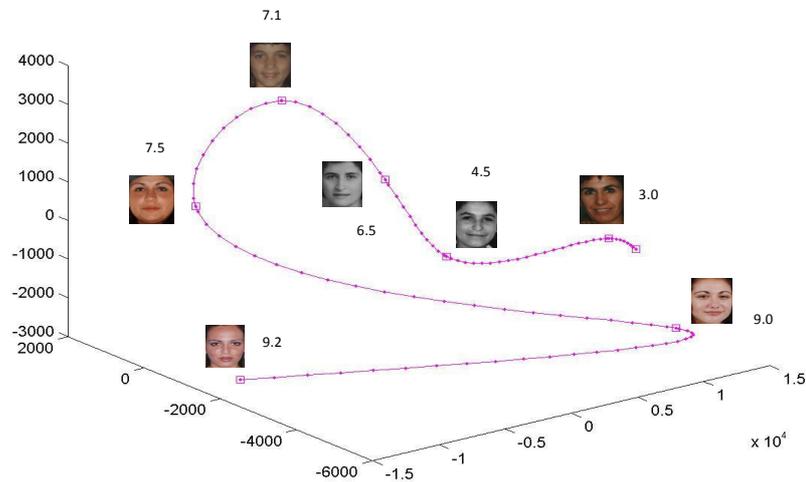


Fig. 2. Curva paramétrica en el Eigenespacio de máscaras de los rostros del ensamble. Los vectores-máscaras se interpolan, permitiendo localizar con mayor precisión la mejor calificación. Cada vector del ensamble original así como todos los puntos interpolados tienen asociado o acarrean una calificación de belleza, que permite calificar un rostro de entrada original.

Usando los vectores obtenidos de las máscaras de cada imagen del ensamble se crea un Eigenespacio y se mapean cada una de las máscaras, según se explicó la creación del Eigenespacio y el mapeo en la sección 2. Una vez hecho el mapeo de cada máscara al Eigenespacio, se interpolan los vectores mapeados y se crea una curva paramétrica entre cada uno de estos vectores que representan a las máscaras de los rostros del ensamble. La curva paramétrica acarrean la calificación de belleza. Para calificar una imagen rostro de entrada original, se obtiene la máscara de la imagen, se mapea al Eigenespacio y usando la curva paramétrica interpolada se encuentra el vector que más se le aproxima al mapeado, finalmente dicho vector similar

proporciona la calificación que acarrea consigo que le fue previamente asignada. Una representación 3D de la curva paramétrica se observa en la Fig. 2. Vale mencionar que la curva en el Eigenespacio es de dimensión 66 o menor, la mostrada en la Fig. 2 es de dimensión 3 para propósitos ilustrativos.

Una vez creada la curva paramétrica entre cada uno de los vectores que representan a las máscaras de los rostros del ensamble y ya que la curva paramétrica acarrea la calificación de belleza es necesario realizar una segunda interpolación lineal para proporcionar una calificación a los vectores de las máscaras interpolados.

Teniendo la curva paramétrica la de las máscaras de rostros y la interpolación lineal de sus calificaciones, dada una imagen de entrada se busca la máscara del rostro que más se le aproxime, localizada esta máscara obtenemos su calificación y esta calificación es proporcionada al rostro de entrada.

5. Resultados

Nuestros resultados están basados en un ensamble de 50 de imágenes de rostros de mujeres con su calificación de belleza respectiva. En la Fig. 3 se muestran algunas de las imágenes usadas en el ensamble. Todo el desarrollo está basado en la base de imágenes Fg-Net [10].

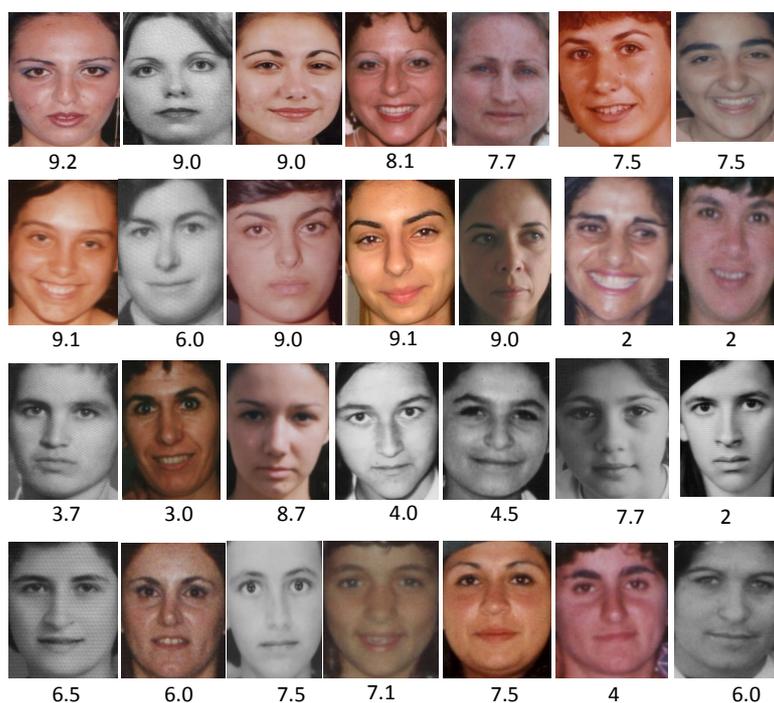


Fig. 3. 28 imágenes usadas en el ensamble

El espacio de rostros está formado por los vectores de distancias -obtenidos de las máscaras- de un conjunto de 50 imágenes de rostros diferentes, previamente previo marcado de la máscara.

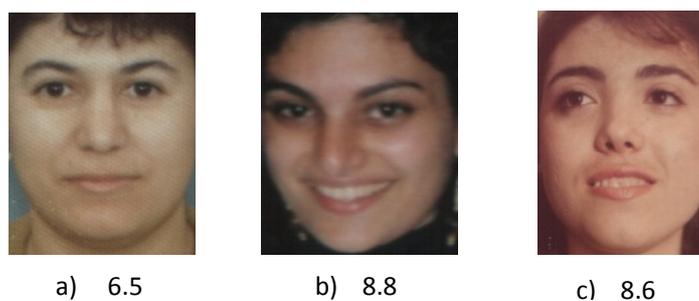


Fig. 4. Algunos resultados obtenidos de utilizar el método propuesto en este artículo.

La Fig. 5 muestra los resultados obtenidos sobre la curva paramétrica, podemos observar que la imagen a) de la Fig. 4 es la única que es equivalente a una máscara de los rostros de ensamble, las otras son semejantes a máscaras creadas por interpolación.

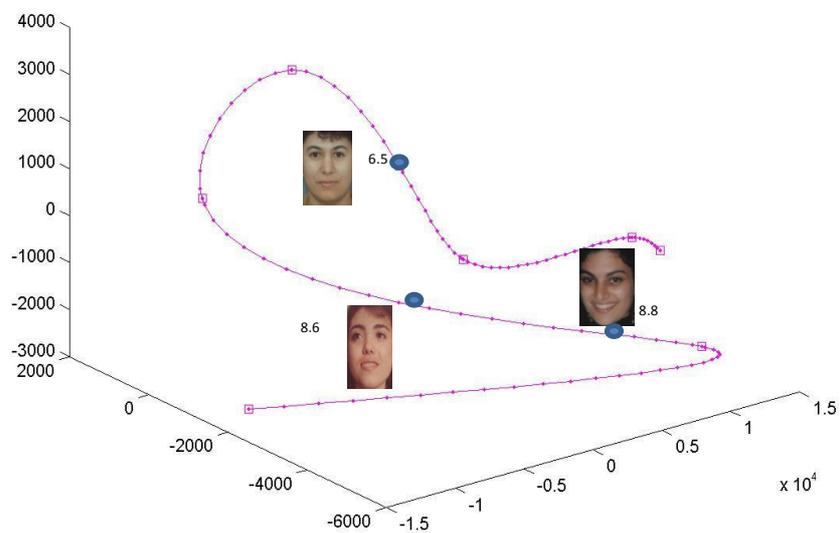


Fig. 5. Resultados sobre la curva paramétrica.

6. Conclusiones y trabajo futuro

Hemos mostrado como nuestra propuesta es factible. Nuestra aproximación tiene aspectos que son deseables en aplicaciones de respuesta casi inmediata. Los submodelos pueden agruparse en clusters según vayan siendo construidos o conforme las imágenes de los rostros vayan siendo adquiridas. Es evidente que en nuestra aproximación se pueden agregar nuevos rostros o eliminar existentes en un tiempo razonable.

La función de evaluación de belleza puede realizarse también con una red neuronal. De hecho cualquier aproximación de aprendizaje supervisado permite aproximar dicha función.

Referencias

1. Zhao, Y., Huang, X., Gao, J., Tokuta, A., Zhang, A., Yang, R.: Video Face Beatification (2013)
2. Turk, M., Pentland, A.: Face recognition using eigenfaces, Proc. Of IEEE Conf. On Comp. Vision and Pattern Recog., 586–591 (1991)
3. Leyvand, T., Cohen Or, D., Dror, G., Lischinski, D.: Data Driven Enhancement of Facial Attractiveness. ACM Trans. Graph., 27(3):9 (2008)
4. Eishenthal, Y., Dror, G., Ruppin, E.: Facial attractiveness: Beauty and the machine. Neural Computation, 18(1):119–142 (2006)
5. Bar, M., Neta, M., Linz, H.: Very first impressions. Emotion by the American Psychological Association, 6(2):269–278 (2006)
6. Nene, S., Nayar, S., Murase, H.: Software library for appearance matching. Technical Report CUCS-019-94, Department of Computer Science, Columbia University (1994)
7. Sirovich, L., Kirby, M.: Application of the Karhunen-Loeve procedure for the characterization of human faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(1) (1990)
8. Altamirano, L.C., Altamirano, L., Alvarado, M.: Non-Uniform sampling for Improve Appearance-based Models. Pattern Recognition Letters (2002)
9. Chandrasekaran, S., Manjunath, B., Wang, Y., Zhang, H.: An Eigenspace Update Algorithm for Image Analysis, Graphical Models and Image Processing, 59(5):321–332 (1997)
10. Murase, H., Nayar, S.K.: Parametric Eigenspace Representation for Visual Learning and Recognition. In: Proceedings of The International Society for Optical Engineering (SPIE), volume 2031, pp. 378–391 (1993)
11. Nayar, S.K., Murase, H., Nene, S.A.: Parametric Appearance Representation. In: Early Visual Learning, pp. 131–160 (1996)
12. Nayar, S.K., Baker, S., Murase, H.: Parametric Feature Detection. In: DARPA Image Understanding Workshop (IUW), pp. 1425–1430 (1997)
13. Cootes, T.: The Fg-Net Aging Database, <http://sting.cycollege.ac.cy/~alanitis/fgnetaging/index.htm>