

Modifications on Illumination, Distance function and Gabor masks for Elastic Bunch Graph Matching

Juan Carlos Gutierrez Caceres
Universidad Católica San Pablo
Universidad Nacional de San Agustín
Arequipa, Perú
Email:jcgutierrezc@gmail.com

José Rodrigo Galdos Chávez
Universidad Católica San Pablo
Arequipa, Perú
Email:jose.galdos.chavez@ucsp.edu.pe

Elian Laura Riveros
Círculo de Investigación
Universidad Nacional de San Agustín
Arequipa, Perú
Email: elian.laura.riv@gmail.com

Abstract—Face recognition is one of the most important field in computer vision, although there are many proposals and research papers still have limitations in real applications where uncontrolled conditions such as illumination, view angle, facial expressions, resolution and image quality, etc. are the main problems. To solve these issues there are a large number of methods for recognition, which can be grouped according to the approach by which the recognition process is addressed, such groups are; holistic methods and feature based.

In this paper we address a feature-based method called Elastic Bunch Graph Matching (EBGM) which is appropriate for an uncontrolled environment for its tolerance to changing background and certain variations of pose, also by having a lower sensitivity variations of illumination which is one of the weaknesses in the holistic methods, which compared to them EBGM without any modification obtained results alongside well-known methods as Principal Component Analysis (PCA).

EBGM has several parameters that can be configured and most of the works in literature uses the default settings of the original author. In this respect this paper presents modifications to these parameters as regards the number of models, illumination enhancement in the preprocessing phase, the configurations Gabor masks and modifying the similarity function. Finally we corroborate the existence of improvements on our experimental results.

Index Terms—EBGM, Illumination, Face recognition, biometrics, uncontrolled environment.

I. INTRODUCCIÓN

EL reconocimiento de rostros se ha vuelto en área de gran interés en visión computacional, y sobre todo en los últimos años donde se han presentado varios avances en materia de investigación, también se ha puesto en manifiesto su importancia para su uso en seguridad, ya sea para usos de autenticación o para seguridad y vigilancia (identificación), donde las cantidad de situaciones donde es aplicable se incrementan cada año.

Muchas de las situaciones donde se desea aplicar el reconocimiento de rostros en la actualidad, comparten la característica de ser un ambiente no controlado, esto quiere decir que no se tiene control sobre las condiciones de iluminación, ángulo de visión, oclusiones e incluso la colaboración de los sujetos a reconocer entre otros factores. Un ejemplo es la vídeo vigilancia, que es uno de los escenarios mas retadores para el reconocimiento de rostros en la actualidad.

Los métodos de reconocimiento de rostros se pueden dividir de varias formas[1] entre ellas esta la división entre métodos holísticos y basados en características [2]. Métodos muy conocidos como PCA, Linear Discriminant Analysis (LDA), Kernel Fisher Analysis (KFA) son métodos holísticos que tienen muy buenos resultados cuando se usan para situaciones particulares donde iluminación, pose, oclusión, fondo de imagen, ángulo de visión y otros factores [3] están de alguna manera controlados. Para el caso de vídeo vigilancia los factores que dificultan el reconocimiento aumentan[1]. Por ello los métodos holísticos tienen debilidades como los cambios de iluminación, la posición del rostro, ángulo de visión, el fondo de la imagen que se procesa, diferencia entre tamaños de imagen, etc.

Para abordar el reto de reconocer rostros en vídeo vigilancia elegimos un método basado en características llamando EBGM [4] que se basa en el concepto que las imágenes de los rostros reales tienen muchas características no lineales que no son abordadas por los métodos que usan un enfoque holístico, este método extrae información sobre partes específicas del rostros que son llamados puntos "fiduciales" y son considerados características biométricas. Pero EBGM presenta varios puntos débiles; no distingue que puntos extraídos del rostro pueden ser más importantes que otros, y la forma como estos puntos son elegidos es a través de modelos que se usan para formar un grafo del rostro por lo tanto es dependiente de que modelos son usados, la configuración de las máscaras de Gabor influye en la precisión para encontrar los puntos fiduciales, no existe pre-procesamiento más allá de una normalización de la información de los píxeles y ajuste de tamaño de la imagen. La variación de estos factores son tienen un efecto en la tasa de aciertos del reconocimiento.

Para mejorar este método de reconocimiento proponemos una modificación a las funciones de distancia que usa para determinar el reconocimiento, mejoras para el pre-procesamiento, cambios en las configuraciones de las máscaras de Gabor y en la selección de modelos para los "template" del grafo. Describiremos las pruebas realizadas a nuestra propuesta en varias bases de datos y haremos comparación con otros métodos para demostrar que la propuesta mejora el acierto de reconocimientos independientemente de las condiciones en que se presenta los rostros.

II. PROPUESTA

La propuesta es agregar técnicas y modificaciones al curso normal de EBG. Para ello usamos la implementación ofrecida por Bolme [5] la cual puede ser entendida en la figura 1: Los puntos que son modificados son los siguientes:

- Nueva selección de imágenes para los modelos de grafo con mayor variedad de rostros.
- Modificación de los tamaños del conjunto de máscaras de Gabor.
- Agregar procesos de mejora de iluminación en la fase de pre-procesamiento del algoritmo.
- Modificar la función de distancia para darle pesos a los diferentes puntos fiduciales.

A. Elastic Bunch Graph Matching (EBGM)

EBGM se basa en el concepto que las imágenes de los rostros reales tienen muchas características no lineales que no son abordadas por los métodos que usan un enfoque holístico, tales como variaciones en la iluminación, pose y expresión. Un conjunto de transformaciones de Gabor Wavelet aplicada en un punto fiducial crea una arquitectura de enlaces dinámicos que proyecta el rostro en una malla elástica.

B. Pre-procesamiento

Como parte de la propuesta incluimos la mejora de iluminación a través de equalización de histograma, transformada de logaritmo y transformada de frecuencia DCT presentada en el trabajo de Manjula [6], siendo la siguiente ecuación 1:

$$F(x, y) = c_1 * DCT + c_2 * Lg + c_3 * HE \quad (1)$$

Donde c_1, c_2, c_3 son valores tipo peso para equilibrar el efecto de las técnicas usadas, sus valores son de 0.3, 0.5 y 0.5 respectivamente.

Como parte de la implementación usada, se realiza un pre-procesamiento que consta de una transformación de perspectiva que traslada las coordenadas de los ojos a coordenadas preestablecidas (un problema que no tratamos en este artículo es el hecho que EBG necesita como información de entrada las coordenadas de los ojos, para propósito de nuestras pruebas estos datos son proveídos) y realiza un cambio de tamaño en la imagen a 128×128 , lo cual reduce el tamaño del rostro y lo acerca a tamaños que se presenta en vídeos de vigilancia reales. Además se le agrega un borde de 30 píxeles incluido dentro de los 128 píxeles de la imagen. Todas las modificaciones en pre-procesamiento se realizan en esta parte del algoritmo.

C. Nuevos modelos

Debido a que la implementación de Bolme [5] usa modelos creados a partir de imágenes incluidas en la base de datos "Ferret Gray" en donde parte de esa base de datos está declarada como desaparecida. Elegimos un nuevo conjunto de imágenes de la nueva versión conocida como "Ferret Color" e incrementamos el número de imágenes modelos de 70 en la implementación original a 150 con el fin de obtener mayor

variedad de modelos de referencia para la localización de los puntos fiduciales del rostro, y con ello aumenta la precisión de reconocimiento.

D. Variaciones en Gabor Wavelet

El Gabor Jet es un nodo en la malla elástica, el cual describe el comportamiento alrededor de un píxel. Esto es el resultado de la convolución de un píxel de la imagen con varios Gabor wavelet o máscaras de Gabor, los cuales son usados para detectar formas y extraer características.

$$W(x, y, \theta, \lambda, \phi, \sigma, \gamma) = e^{-\frac{x'^2 + \gamma y'^2}{2\sigma^2}} \cos(2\pi \frac{x'}{\lambda} + \phi) \quad (2)$$

$$x' = \cos\theta + y\sin\theta$$

$$y' = -x\sin\theta + y\cos\theta$$

Los parámetros usados para la construcción de los Gabor wavelet son los mismo que se utilizan en la implementación de Bolme[5], a continuación los explicamos brevemente:

- θ especifica la orientación del Gabor Wavelet. Siendo $\theta \in \{0, \pi/8, 2\pi/8, 3\pi/8, 4\pi/8, 5\pi/8, 6\pi/8, 7\pi/8\}$
- λ especifica el ancho de onda de la función seno, empieza con 4 píxeles y aumenta en medias octavas siendo $\lambda \in \{4, 4\sqrt{2}, 8, 8\sqrt{2}, 16\}$
- ϕ especifica la fase de la función seno, pudiendo ser par e impar que representa la parte imaginaria y la parte real del wavelet respectivamente. Siendo $\{0, \pi/2\}$
- δ especifica el radio de la Gaussiana. En este caso $\delta = \lambda$.
- γ especifica el ratio de aspecto de la Gaussiana. Este parámetro es incluido para que el Wavelet se aproxime a ciertos modelos biológicos. Siendo $\gamma = 1$.

La ecuación 2 que define el Gabor Wavelet se usa para crear máscaras donde x, y son las posiciones en la máscara. De esta manera podemos crear varios tamaños de máscaras que en la configuración original son $N \in \{25, 37, 51, 71, 101\}$ cada uno relacionado con el parámetro λ . Dando 80 configuraciones de Gabor wavelet y siendo efectivas 40 máscaras por punto fiducial, debido a que existe una máscara que extrae la parte imaginaria y otra la parte real del Wavelet.

Nuestra modificación plantea el cambio del tamaño de las máscaras de Gabor, debido a que tamaños de 71 y 101 son demasiado grande teniendo en cuenta que la imagen a convolucionar es de 128×128 y el tamaño efectivo del rostro es aun menor siendo los tamaños propuestos $N \in \{13, 19, 25, 35, 51\}$ los resultados pueden observarse en la tabla IV

E. Pesos y Función de distancia

El grafo compuesto por los nodo Gabor Jet es llamado "Face Graph" el cual es la representación del rostro a reconocer. El reconocimiento está basado en la similitud de Gabor Jet en cada nodo. La dificultad con este método es el requerimiento de marcar puntos precisos en los rostros y el hecho que no discrimina si algún punto en el rostro puede ser más significativo que otro.

$$L_{jet}(G, G') = \frac{1}{N} \sum_{i=0}^N S(J_i, J'_i) \quad (3)$$

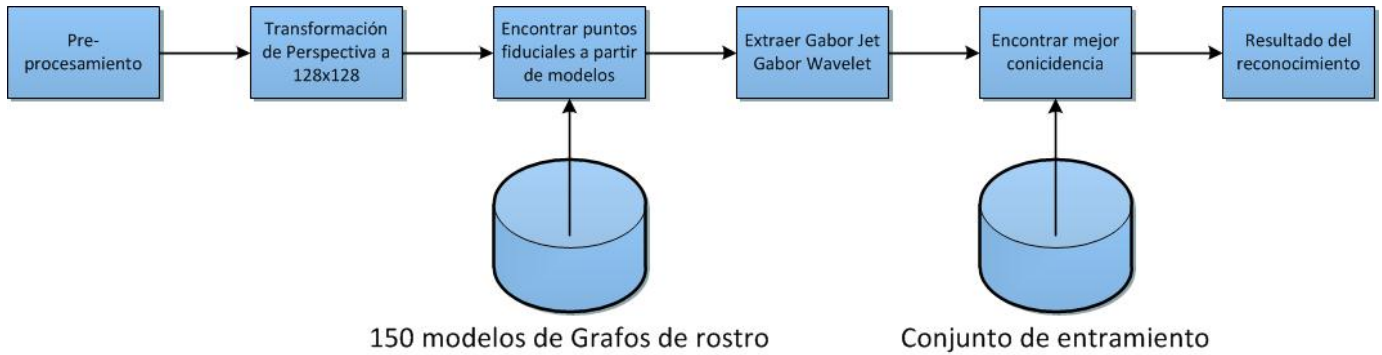


Fig. 1. Propuesta de modificación de Elastic Bunch Graph Matching (EBGM)

TABLE I
CONFIGURACIONES DE PESOS PROPUESTOS PARA LOS DIFERENTE PUNTOS FIDUCIALES

Puntos Fiduciales	1ra Config.	2da Config	3ra Config	4ta Config	5ta Config	6ta Config
ojo izquierdo	0.07	0.07	0.05	1.07	1.07	1.05
ojo derecho	0.07	0.07	0.05	1.07	1.07	1.05
punte de la nariz	0.06	0.07	0.05	1.06	1.07	1.05
pico ceja derecha	0.03	0.025	0.025	1.03	1.025	1.025
pico ceja izquierda	0.03	0.025	0.025	1.03	1.025	1.025
interior ceja derecha	0.04	0.025	0.025	1.04	1.025	1.025
interior ceja izquierda	0.04	0.025	0.025	1.04	1.025	1.025
exterior ceja derecha	0.03	0.025	0.025	1.03	0.025	1.025
exterior ceja derecha	0.03	0.025	0.025	1.03	1.025	1.025
punta de la nariz	0.05	0.08	0.05	1.05	1.08	1.05
centro base de la nariz	0.05	0.07	0.05	0.05	1.07	1.05
base derecha de la nariz	0.05	0.07	0.05	1.05	1.07	1.05
base izquierda de la nariz	0.05	0.07	0.05	1.05	1.07	1.05
parte central superior de la boca	0.05	0.0375	0.0625	1.05	1.0375	1.0625
parte central inferior de la boca	0.05	0.0375	0.0625	1.05	1.0375	1.0625
esquina izquierda de la boca	0.05	0.0375	0.0625	1.05	1.0375	1.0625
esquina derecha de la boca	0.05	0.0375	0.0625	1.05	1.0375	1.0625
parte central superior de la cabeza	0.02	0.02	0.01	1.02	1.02	1.01
parte izquierda superior de la cabeza	0.02	0.02	0.01	1.02	1.02	1.01
parte derecha,superior de la cabeza	0.02	0.02	0.01	1.02	1.02	1.01
borde izquierdo de la cara	0.02	0.02	0.01	1.02	1.02	1.01
borde derecho de la cara	0.02	0.02	0.01	1.02	1.02	1.01
centro de la barbilla	0.04	0.04	0.06	1.04	1.04	1.06
parte izquierda de la quijada	0.03	0.03	0.07	1.03	1.03	1.07
parte derecha de la quijada	0.03	0.03	0.07	1.03	1.03	1.07



Fig. 2. Puntos fiduciales que componen el grafo del rostro, los puntos son definidos en el trabajo de Wiskott [4]

$$S(J, J') = \frac{\sum_{j=1}^N a_j a'_j \cos(\phi_j - \phi'_j)}{\sqrt{\sum_{j=1}^N a_j^2 \sum_{j=1}^N a'^2_j}} \quad (4)$$

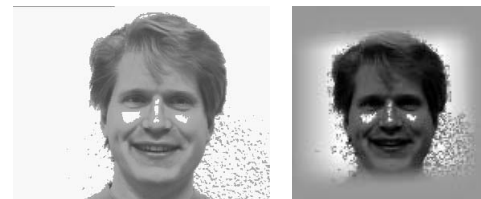


Fig. 3. Imagen antes y después del Pre-procesamiento

La ecuación 4 es la función de similitud para dos Gabor Jet, donde J y J' son los resultados de las convoluciones de las mascarar de gabor en un punto fiducial, N es la cantidad de mascarar de pares de mascarar de gabor y a y ϕ representan la parte real e imaginaria respectivamente y la ecuación 3 es la función de similitud para todo un rostro y su representación se llama Face Graph, donde G y G' son los grafos de rostros

a comparar, N es la cantidad de puntos fiduciales que contiene el grafo.

Como se puede observar en la figura 2 elegimos 25 puntos fiduciales o puntos de interés, de los cuales extraeremos información a través de las convoluciones con las máscaras de Gabor, una lista completa de estos puntos se puede ver en la tabla I. Siguiendo las recomendaciones y sugerencias presentadas en el trabajo de Bolme [5] agregamos un valor de pesos para la función de distancia nos permite darle mayor importancia a algunos puntos que otros, por ejemplo los puntos en los ojos y la nariz brindan mayor información que los puntos en el contorno de la cara. Para ello agregamos un vector de pesos que corresponde a los puntos de los rostro resultando en las ecuaciones 5 y 6. Podemos ver los valores del vector W en la tabla I

$$L_{jet}(G, G', W) = \frac{1}{N} \sum_{i=0}^N S(J_i, J'_i, w_i) \quad (5)$$

$$S(J, J', w) = w * \frac{\sum_{j=1}^N a_j a'_j \cos(\phi_j - \phi'_j)}{\sqrt{\sum_{j=1}^N a_j^2 \sum_{j=1}^N a'^2_j}} \quad (6)$$

Para la selección de los pesos se decidió agrupar los punto fiduciales del rostro en seis grupos por cercanía como se puede ver en la tabla II y asignándoles un porcentaje del total del valor de la función de similitud para que la suma de todos los valores w sea 1 en la tres primeras configuraciones, para las configuraciones 4, 5 y 6 es el mismo valor sumando uno para que para que la multiplicación de la distancia en la ecuación 6 no reduzca el valor de la operación por un valor menor a uno.

Para la distribución de los valores vistos en la tabla II se toma en consideración que el borde de la cabeza no ofrece tanta información como el centro del rostro e intentamos darle mayor importancia a los ojos y la nariz.

TABLE II

DISTRIBUCIÓN DEL VALOR DE LOS PESOS PARA LA MODIFICACIÓN DE LA FUNCIÓN DE SIMILITUD

Grupo	Nro Puntos	Conf. 1	Conf. 2	Conf. 3
Ojos y puente de la nariz	3	20%	50%	15%
Cejas	6	20%	15%	15%
Nariz	4	20%	20%	20%
Boca	4	20%	15%	25%
Contorno cabeza	5	10%	10%	5%
Quijada y mandíbula	3	10%	10%	20%

III. RESULTADOS

Las pruebas y modificaciones presentadas en este artículo se basa en la implementación realizada por David Bolme [5].

Para la realización de las pruebas en este artículo cada base de datos fue dividida en grupos de 3 a 4 imágenes por sujeto, dependiendo de la base de datos y cada grupo fue usado como entrenamiento mientras el resto es usado como prueba. Este proceso se repite hasta que todos los grupos hayan sido usados como entrenamiento, siendo la cifra final el promedio de los resultados de cada grupo.

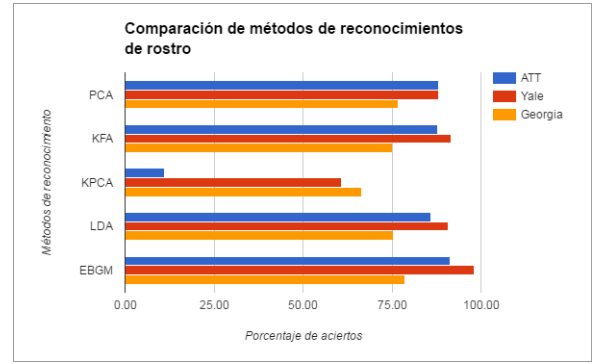


Fig. 4. Comparación de EBGM con otros algoritmos en las base de datos AT&T, Yale y Georgia

A. EBGM en comparación de otros métodos

Se realizó una comparación con otros algoritmos usado en reconocimiento de rostros como PCA[7], KFA[8], Kernel Principal Component Analysis (KPCA)[8], LDA[9] implementados en matlab[10] y varias bases de datos como ORL/ATT, YaleA, Georgia Tech para demostrar el rendimiento de estos algoritmos junto a EBGM en varias situaciones.

TABLE III

COMPARACIÓN CON BASES DE DATOS ATT, YALE A Y GEORGIA

	ATT/ORL	YALE A	Georgia
PCA	88.03%	88.18%	76.84%
KFA	87.78%	91.72%	75.03%
KPCA	10.95%	60.71%	66.43%
LDA	86.02%	98.21%	75.50%
EBGM	91.43%	98.21%	78.74%

La tabla IV muestra todas las modificaciones planteadas a EBGM en lo que respecta a porcentaje de aciertos.

Como podemos ver existe una mejora cuando se usa los nuevos parámetros de las máscaras de Gabor debido a los tamaños mas pequeños, extrae información mas precisa sobre el punto fiducial. La modificación a la ecuación de similitud 3 y 4 agregando pesos también muestran una mejora en el porcentaje de aciertos, sobretodo las configuraciones 4, 5, 6. Y también la mejora usando las técnicas de pre-procesamiento propuestas por Manjula[6] mejora su rendimiento sobretodo en la base de datos Yale.

TABLE IV

MODIFICACIONES AL EBGM

Propuestas	AT&T	YALE	Georgia
Algoritmo sin cambios	91.43%	98.21%	78.73%
Primera configuración de pesos	91.67%	97.58%	80.43%
Segunda configuración de pesos	91.15%	97.86%	79.40%
Tercera configuración de pesos	90.16%	96.71%	76.70%
Cuarta configuración de pesos	92.78%	98.21%	78.90%
Quinta configuración de pesos	92.78%	98.21%	78.87%
Sexta configuración de pesos	92.66%	98.21%	78.90%
Nuevas máscaras de Gabor	93.18%	97.58%	79.43%
Con mejora en iluminación	92.82%	98.5%	78.93%

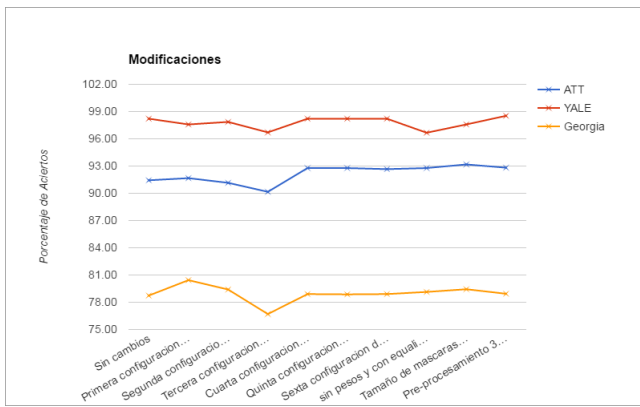


Fig. 5. Variación de los resultados a través de las diferentes modificaciones propuestas

IV. CONCLUSIÓN

Podemos ver que en la tabla III y en la figura 4 que EBGMM que es un algoritmo muy robusto, que mantiene porcentajes de aciertos muy buenos independiente de la base de datos en la que se realiza la prueba y que en ciertas situaciones es mejor que PCA, y los otros métodos holísticos.

Hemos logrado mejoras con las modificaciones propuestas, podemos ver las modificaciones en las máscaras de Gabor mejoran el rendimiento del reconocimiento y en el peor de los casos no lo afecta el rendimiento original, como se puede observar en la tabla IV. El método de pre-procesamiento propuesto por Manjula[6] mejora el reconocimiento en las tres bases de datos probadas siendo sus mejores resultados de 92.82%, 98.5% y 78.93% en AT&T, Yale y Georgia respectivamente.

Algunas configuraciones de pesos afecta el reconocimiento pero siempre hay una configuración que mejora el reconocimiento, esta variación se puede apreciar en la figura 5 donde se muestra los resultados a través de las diferentes opciones expuestas en este artículo, siendo el resultado más relevante el de la segunda configuración donde la base de datos Georgia que es la que posee la mayor cantidad de imágenes obtiene un porcentaje de acierto de 80.43%.

Estos resultados nos acercan a un método de reconocimiento que es robusto para cualquier situación y que puede ser apto para el reconocimiento de rostros en video vigilancia.

V. AGRADECIMIENTOS

El presente trabajo es respaldado por el financiamiento 009-2014-FONDECYT del Consejo Nacional de Ciencia, Tecnología e Innovación Tecnológica (CONCYTEC-PERU) y por la Universidad Católica San Pablo de Arequipa.

REFERENCIAS

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *Acm Computing Surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003.

[2] S. Tseng, "Comparison of holistic and feature based approaches to face recognition," Ph.D. dissertation, School of Computer Science and Information Technology, Faculty of Applied Science, Royal Melbourne Institute of Technology, 2003.

[3] R. Gross, J. Shi, and J. F. Cohn, *Quo vadis face recognition?* Carnegie Mellon University, The Robotics Institute, 2001.

[4] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. Von Der Malsburg, "Face recognition by elastic bunch graph matching," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 775–779, 1997.

[5] D. S. Bolme, "Elastic bunch graph matching," Ph.D. dissertation, Colorado State University, 2003.

[6] V. Manjula, "Image normalization robust using histogram equalization and logarithm transform frequency dct coefficients for illumination in facial images."

[7] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[8] M.-H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods." in *fgr*, vol. 2, 2002, p. 215.

[9] W. Zhao, R. Chellappa, and P. J. Phillips, *Subspace linear discriminant analysis for face recognition*. Citeseer, 1999.

[10] V. Struc, "The phd toolbox: Pretty helpful development functions for face recognition," 2012.