

Reconocimiento de rostros en tiempo real mediante análisis de componentes principales

David Loza Matovelle

Universidad Internacional del Ecuador
Quito, Ecuador

Email: davidlozaing@hotmail.com

Juan Pablo Semanate

Universidad Internacional del Ecuador
Quito, Ecuador

Email: juan.semanate@epn.edu.ec

Pedro Román

Universidad Internacional del Ecuador
Quito, Ecuador

Email: peter1roman@hotmail.com

Resumen—El presente artículo, propone un sistema de reconocimiento de rostros compuesto por tres elementos: primero, un detector de rostros que trabaja a través del algoritmo Viola-Jones el cual reconoce rostros de forma rápida. Además, se ha implementado un identificador del color de piel para robustecer el sistema de detección. Segundo, se ha normalizado la imagen del rostro eliminando el brillo y corrigiendo la posición del mismo. Con ello, se obtiene un rostro con similares características a las imágenes almacenadas en la base de conocimiento. Por último, la identificación del rostro se realizó a través de la técnica de análisis de componentes principales (PCA). Se han obtenido resultados positivos en la identificación de rostros a través del algoritmo propuesto.

Palabras claves: *Detección de Rostros, Reconocimiento de Rostros, Análisis de Componentes Principales (PCA)*

I. INTRODUCCIÓN

El reconocimiento de rostros es un tema ampliamente analizado, debido a la gran cantidad de aplicaciones basadas en seguridad, biometría, sistemas inteligentes, interacción persona-computador entre otros. Por ejemplo, en [7] y [9] se consideran elementos de seguridad como el iris, perfiles, etc.

Las investigaciones en este campo, actualmente se centran en resolver problemas generados principalmente por: baja calidad de video, imágenes de caras con poca resolución, ambientes no controlados y variación en la iluminación. En [18] se estudia la detección de rostros con baja resolución y se analiza el grado de efectividad que se tiene con técnicas estándar. También, en [1] se analiza el reconocimiento de rostros bajo condiciones de luz diversas.

El encontrar una forma rápida y confiable de reconocer rostros es un tema de activa investigación. En [16] basándose en la detección de movimiento se estima la posición del rostro, lo cual facilita los cálculos posteriores y reduce el tiempo de procesamiento. En [6] se presenta un método de detección rápida de objetos. También, el uso de tecnologías “más frecuentes” es elemento de amplios estudios. La adquisición de datos con una cámara web estándar se analiza en [12] y en [13] esta tecnología se asocia al reconocimiento de rostro en redes sociales.

Existen diferentes formas de desarrollar el tema de reconocimiento de rostros. En el enfoque estadístico se encuentran varias técnicas, entre ellas el análisis discriminante lineal (LDA *Linear Discriminant Analysis*) que se basa en encontrar

relaciones lineales entre los rostros de la base de conocimiento, que mejor discriminen el fotograma de prueba. El discriminante lineal de Fisher (FLD *Fisher Linear Discriminant*) trata de encontrar la mejor función que discrimine el fotograma y que sea una combinación lineal de los rostros de la base de conocimiento. Por último, el enfoque que se ha empleado en el presente artículo está basado en el análisis de componentes principales (PCA *Principal Components Analysis*) y se describe en la sección II-B de este artículo.

Por otro lado, existen técnicas independientes al contexto estadístico que han logrado mejores resultados pero poseen un alto costo computacional o equipos demasiado sofisticados para su implementación. Entre estas técnicas está el uso de redes neuronales que en un principio eran muy lentas y tenían un alto consumo de los recursos. Actualmente, por el avance de la tecnología el uso de estas técnicas es más común. En [10] se analiza el color de piel y labios a través de una red neuronal. Otro método de alto costo computacional es el uso de plantillas deformables las cuales permiten realizar el análisis en diferentes posiciones. En [2] se analiza el rostro de tal manera que para no perder información relevante se da ciertos pesos a zonas que en el rostro son invariantes como puede ser nariz, ojos o mejillas.

Existe otro grupo de métodos denominados 3D que solucionan algunos tipos de problemas que se encuentran en los métodos anteriores como posición del rostro e iluminación. Su principal problema es la necesidad de equipos especializados o de técnicas con mayor costo computacional que las anteriores. Basándose en estas técnicas tridimensionales también se han desarrollado métodos invasivos como son escáneres biométricos y otros sistemas similares. Algunos ejemplos de este grupo de técnicas son los artículos [8] y [3].

Para comprender los PCA, existen algunas investigaciones importantes a considerarse. Por ejemplo, en [4] se analizan las características de los PCA comparados con las FLD, se concluye que ambas técnicas son válidas y con similares porcentajes de efectividad. También, en [15] se analiza detalladamente el procesamiento de datos a través de la técnica PCA para la clasificación. El artículo presenta un sistema sencillo y rápido de reconocimiento de rostros basado en el análisis de componentes principales. Con lo cual, el algoritmo es capaz de detectar a una persona si se encuentra dentro de la base de conocimientos con una simple cámara web. El documento

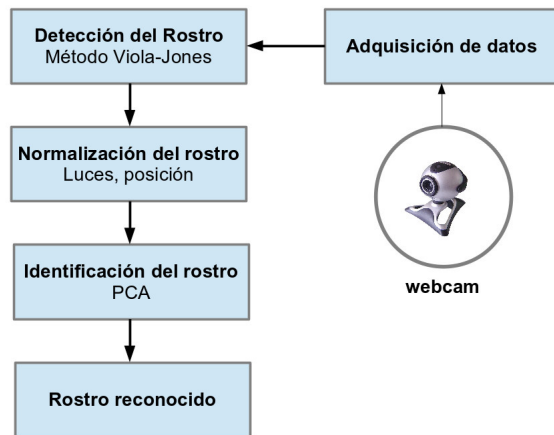


Figura 1. Proceso de reconocimiento de rostros

trata los problemas de ángulos en rostros e iluminación de manera sencilla. También, se analiza los resultados obtenidos con el sistema.

II. PROCESO DE RECONOCIMIENTO DE ROSTROS

El sistema de reconocimiento fue programado en Python. En la figura 1 se puede ver los bloques funcionales del algoritmo propuesto. En la parte superior de la figura se indica el proceso de identificación de rostro, que es adquirido a través de reconocedor Viola-Jones. Segundo, se modifica el rostro a través de la corrección del histograma y la posición del rostro. Y por último, se analiza el rostro para su clasificación a través de PCA.

A continuación, se detallan los diferentes elementos del proceso.

II-A. Localización y normalización de rostros

En la primera parte del algoritmo se localiza el rostro y se normaliza la imagen. En este proceso, se ha utilizado el método desarrollado por Paul Viola y Michael J. Jones [17] el cual fue publicado en el año 2004 y presenta un 95 % de confiabilidad. Los pasos del algoritmo son:

Primero: mediante una imagen integral se puede calcular la suma de las subregiones de la imagen rápidamente. La imagen integral es obtenida al agregar una fila y una columna ($w+1$, $h+1$) estos nuevos elementos reciben el valor de 0, cada píxel en la imagen integral estará definido como la suma de los píxeles anteriores. La imagen integral tiene diferentes aplicaciones difuminaciones, gradientes, sumatorias, promedios, etc.

A continuación, se obtienen las características de la imagen. Con este fin, se utilizó filtros con base *Haar* [11] (estos filtros toman secciones rectangulares simples de la imagen para su análisis), los cuales son calculados sobre la imagen integral. Los filtros generan características de puntos y líneas en la imagen por el contraste que existen entre las regiones del fotograma. Se pueden cambiar diferentes características dentro de los filtros: escala, orientación, tamaño y frecuencia.

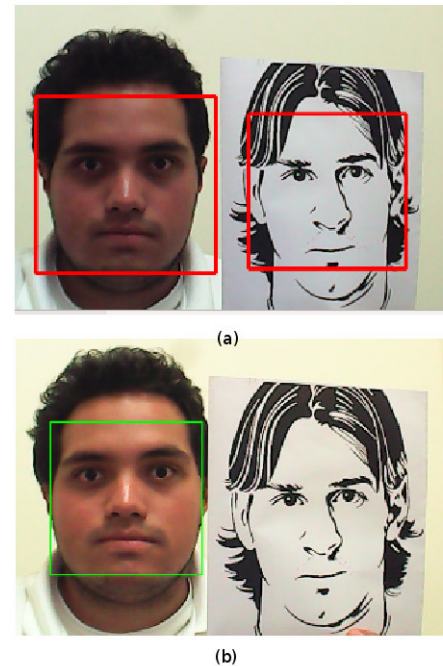


Figura 2. Localización de rostros y detección de color de piel (a) Sin el detector de piel (b) Con el detector de piel.

En este punto, se procede a la detección en función de un modelo inducido durante el entrenamiento (el entrenamiento es basándose en ejemplos positivos y negativos). Para clasificar se utiliza un método denominado *boosting* que utiliza varios clasificadores sencillos que en conjunto forman uno más robusto, de ahí el nombre de clasificador en cascada. En el método propuesto se utiliza el clasificador *AdaBoost* [5]. Las ventajas de este método son su rapidez y gran eficiencia.

Las desventajas de este método son que los filtros no reconocen color ni se puede detectar otro modelo que no se haya inducido en el entrenamiento, tampoco detecta cambios o rotaciones en el rostro. Se pensó en un principio el agregar varios clasificadores para diferentes posiciones del rostro, pero en las pruebas se rechazó la idea debido a que la clasificación era muy lenta.

Como refuerzo al sistema de detección se emplea el detector de color de piel para descartar falsos positivos (rostros que no lo son) ya que el método Viola-Jones puede ser engañado con una cara que no sea real (fotos, dibujos). El detector de piel convierte la imagen a un formato YCbCr para que no afecte en demasía la iluminación y se pueda hacer más notoria la presencia de piel en el fotograma. Se analiza la imagen buscando un determinado rango de colores, entre los cuales se encuentran los diferentes tonos de piel. La figura 2 presenta los resultados de la técnica propuesta.

Después de la extracción de rostros se necesita preprocesar la imagen para que tenga el mismo tamaño que los ejemplos de la base de conocimiento y normalizada su intensidad. El rostro se escala a una dimensión de 200×200 píxeles. Al



Figura 3. Preprocesamiento de imágenes

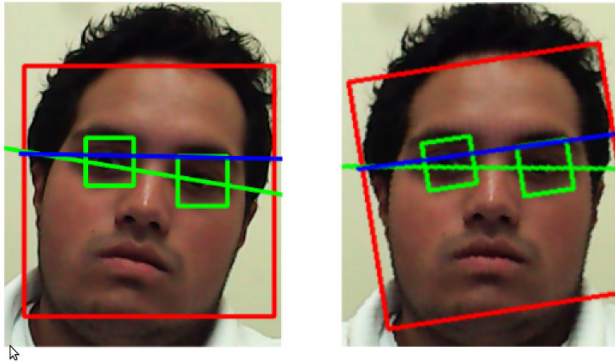


Figura 4. Corrección del ángulo en el fotograma

normalizar la intensidad, se reduce los cambios de luz que afectan sustancialmente los resultados. De acuerdo con [14] se tienen varias formas de eliminar los efectos producidos por las luces. Algunas opciones son: pasar a escala de grises, ecualizar el histograma, corrección gamma, filtros, etc. Para el proyecto se han empleado el transformar la imagen a escala de grises, ecualizar el histograma y suavizar el fotograma. La figura 3 muestra el resultado del pre-proceso.

En la normalización del rostro se consideró la posición del mismo. Para detectar la inclinación se analizó las posibilidades que dispone el detector Viola-Jones de encontrar subelementos como por ejemplo ojos. Con esta idea, se puede calcular el ángulo que existe entre los ojos y realizar una transformación lineal de la imagen para obtener un rostro en posición frontal. En la figura 4 se puede ver una aplicación de la técnica propuesta.

II-B. Reconocimiento de Rostros

A continuación, se analiza el PCA y los cálculos realizados para la aplicación en el proyecto.

Análisis de componentes principales PCA: PCA es una técnica estadística que permite ordenar datos según su importancia, así como, reducir el conjunto de datos eliminando los que son poco relevantes. Esto permite crear modelos representativos del comportamiento de sus variables. Su estudio principalmente permite:

- Reducir la dimensionalidad de un grupo de datos. Elimina características que contienen poca información y conserva las más relevantes.
- Guarda los componentes principales que contribuyen más a su varianza. Estos serán una combinación de las variables originales.
- Se aprecia correlaciones entre las variables que otros métodos no suelen considerar. Por ejemplo existen sistemas de reconocimiento facial que toman en cuenta detalles como distancias entre ojos o longitud de la boca, con la técnica PCA se toman relaciones como detalles específicos en la piel, lunares, cicatrices, etc., que antes no eran consideradas.

A continuación se detalla la formulación matemática de la teoría general de PCA. Para luego pasar a la aplicación en el reconocimiento de rostros.

Primero, los datos son transformados a un vector de dimensión $1 \times m$:

$$v_i = (x_1, x_2, \dots, x_m) \quad (1)$$

Luego, se establece una matriz de dimensiones $n \times m$ que representan n vectores de longitud m . La matriz tiene la siguiente forma:

$$X = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{pmatrix} \quad (2)$$

Se obtiene un vector promedio que es la suma de todos los vectores divididos para n y se expresa como:

$$\Psi = \frac{1}{n} \sum_{i=1}^n v_i \quad (3)$$

Se procede a normalizar la matriz restando a cada vector el vector promedio (Ψ), con lo cual los vectores estarán centrados alrededor de cero.

$$\Phi_i = v_i - \Psi \quad (4)$$

Para obtener la matriz de covarianza se multiplica la matriz X por su transpuesta, esta es una matriz de dimensiones $n \times n$ y luego se divide para la cantidad de vectores menos 1. El resultado es una matriz diagonal positiva la cual tiene n valores propios los que si son distintos están asociados a un vector propio ortonormal.

$$\text{cov}(X) = \frac{XX^T}{n-1} \quad (5)$$

De $cov(X)$ obtenemos sus valores y vectores propios, los que deben estar ordenados en forma decreciente respecto a los valores propios.

$$Det(cov(X) - \lambda I) = 0 \quad (6)$$

$$cov(X) - \lambda I = 0 \quad (7)$$

Una vez obtenidos estos resultados sabemos que es la base generadora en la cual podemos proyectar cualquier vector. Si el resultado es una combinación lineal, pertenece al sub espacio vectorial de la base antes mencionada. La matriz de covarianza por los vectores propios u_p , será igual a los valores propios λ_p por u_p .

$$cov(X)u_p = \lambda_p u_p \quad (8)$$

Finalmente, al multiplicar un vector X por los vectores propios calculados, nos dará como resultado un vector p , el cual representa las proyecciones de X en u_p , al ser comparado con los valores propios podremos saber la varianza.

$$p = Xu_p \quad (9)$$

La técnica de PCA es útil pero tiene ciertas limitaciones. A continuación se menciona cuando este modelo no sería efectivo.

- Solo se puede manejar datos cuantitativos, los datos no deben tener ninguna jerarquía o importancia entre sí.
- Si las variables no tienen correlación, el análisis no es aplicable.
- Se asume que los datos tienen un comportamiento lineal.

PCA aplicado al reconocimiento de rostros: Para aplicar PCA al reconocimiento de rostros se realizaron los siguientes pasos:

Vectorización de una imagen, para el estudio presentado se utiliza las imágenes de dimensión 200×200 píxeles que contiene los datos más importantes de los rostros como son los ojos, nariz y boca. El fotograma se convierte en un vector de una fila con 40000 columnas. Las fotografías son consideradas como los vectores v_i .

Se calcula el vector medio, el cual representa un rostro medio (Ψ) de la base de conocimiento. Como ejemplo, tenemos la figura 5. Se termina el cálculo del PCA como se describió en la sección anterior, con la obtención de autovectores y autovalores. En la figura 6 se presenta los resultado de proyectar las fotografías de la base de conocimiento.

Por último, se mantienen los autovectores k correspondientes a los mayores autovalores asociados. La elección del número de autovalores es un punto crítico en el cálculo, un valor alto de autovalores ralentizará el proceso y en cambio un valor bajo tendrá poca eficiencia.

Proyección de cada imagen de entrenamiento en el subespacio PCA: Para el entrenamiento se utilizaron cinco fotos de cada sujeto ligeramente diferentes. A continuación, lo que realizamos es proyectar cada imagen de entrenamiento en el subespacio. Cada rostro Φ del conjunto de entrenamiento

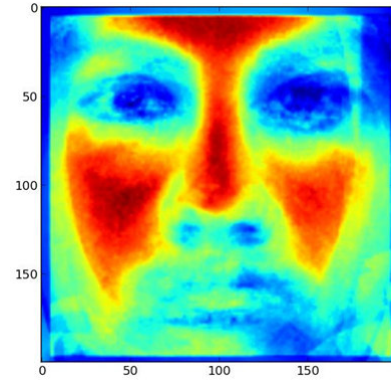


Figura 5. Rostro medio

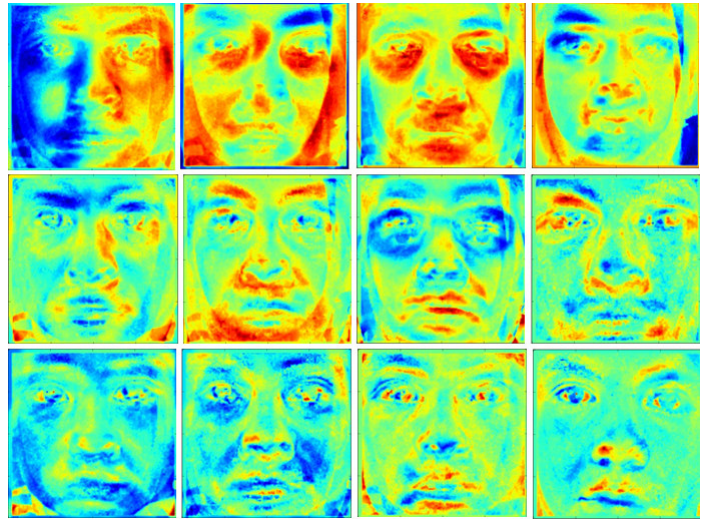


Figura 6. Rostros de la base de conocimiento proyectados

(menos la media) puede ser representado como una combinación lineal de los autovectores más representativos.

$$\tilde{\Phi}_i = \sum_{j=1}^k w_j u_j \quad (w_j = u_j^T \Phi_i) \quad (10)$$

Cada rostro de entrenamiento normalizado es representado en esta base como un vector.

$$\Omega_i = \begin{bmatrix} w_1^i \\ w_2^i \\ \vdots \\ w_k^i \end{bmatrix} \quad i = 1, 2, \dots, M \quad (11)$$

Clasificación de una imagen nueva.: Para la clasificación de una nueva imagen se sigue el siguiente procedimiento. Se normaliza la imagen nueva:

$$\Phi = v - \Psi \quad (12)$$

Se proyecta sobre el espacio propio

$$\tilde{\Phi} = \sum_{i=1}^k w_i u_i \quad (w_i = u_i^T \Phi) \quad (13)$$

Por último, se representa a Φ como:

$$\Omega = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} \quad (14)$$

Se define un error $e_r = \min_i \|\Omega - \Omega_i\|$, siendo la distancia euclidiana en el espacio de rostros. En caso de que e_r se ha menor a una constante (esta constante fue determinada en las pruebas), se lo considera dentro de la base conocimiento, caso contrario se lo marca como desconocida.

III. ANÁLISIS DE RESULTADOS

Para el estudio de resultados se analizaron fotos de 12 personas. Se preprocesaron de diferentes formas y se concluyó que los mejores resultados se obtuvieron de los métodos explicados en este documento. Como primera alternativa se convirtió a escala de grises y se equalizó el histograma, como segunda opción, se convirtió a escala grises con corrección gamma. Se obtuvo un 20% de mejores resultados con la primera configuración. Esto se explica debido a que la corrección gamma solo toma en cuenta características principales como son: ojos, nariz, boca y el contorno de la cara, más no los detalles específicos los cuales, no son valorados adecuadamente por lo cual se pierde información.

La figura 7 muestra los resultados que se producen de la proyección de los rostros en el subespacio. Los círculos de diferentes colores representan grupos de datos proyectados de un mismo rostro. En este espacio se proyectó la imagen localizada por el detector Viola-Jones. Se analizó las distancias entre el punto proyectado y cada grupo de datos por individuo. Concluyendo el proceso, se tomó la distancia mínima como el posible candidato, si la distancia excede una constante se asume que el rostro no se encuentra dentro de la base.

El algoritmo con el grupo de muestra utilizado en forma regular obtuvo un 60% de efectividad bajo ambientes controlados.

La propuesta presenta problemas con respecto al alto contraste y no se adapta al cambio de apariencia, como puede ser la edad de los sujetos, el uso de lentes, barba, etc. Otro problema que se tiene en el sistema es el número de individuos que se puede registrar. Esto se debe ya que a mayor cantidad de usuarios, mayor es el número de subespacios y se corre el riesgo que exista puntos comunes entre los subespacios, con lo cual el reconocimiento no sería de lo más óptimo.

IV. CONCLUSIONES Y TRABAJOS FUTUROS

El presente artículo presenta un algoritmo sencillo de identificación de rostros. Se reconoce el rostro con el algoritmo Viola-Jones con un identificador del color de piel, el sistema funciona bien bajo condiciones normales. Además, se ha desarrollado un sistema que normaliza el rostro en el cual

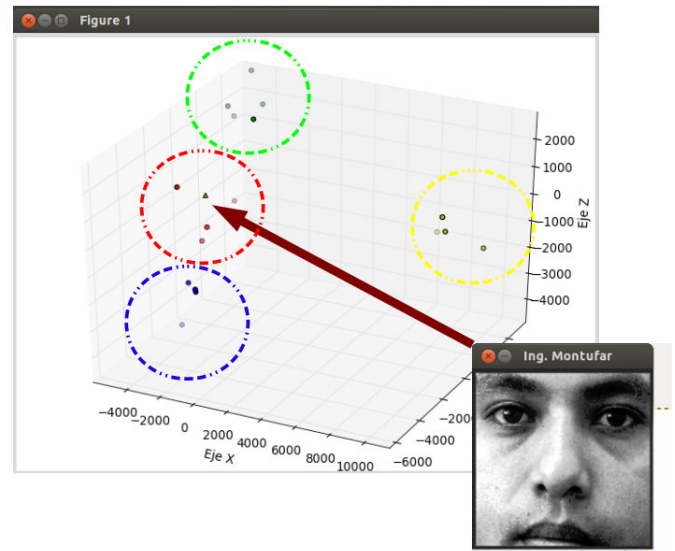


Figura 7. Resultados de la clasificación

se corrige el ángulo e iluminación, para obtener mejores prestaciones. La tasa de reconocimiento bajo circunstancias de luz normales es de un 60%.

Para futuros desarrollos se desea implementar el uso de varias cámaras, con lo cual se podría corregir los efectos de la posición en rostros. También, se piensa emplear otras técnicas como son *FisherFaces* y realizar un estudio comparativo. Incluso se puede desarrollar técnicas híbridas a partir de las desarrolladas y obtener un sistema más robusto.

AGRADECIMIENTOS

Agradecemos a las personas que muy gentilmente colaboraron con las imágenes para el reconocimiento. También, extendemos un agradecimiento a los compañeros docentes de la Universidad Internacional del Ecuador que contribuyeron con sus comentarios y puntos de vista sobre el trabajo.

REFERENCIAS

- [1] M. S. Aliakbarian, F. S. Saleh, F. Aliakbarian, and F. S. Saleh. Robust face detection and recognition under ambient light and head pose. In *7th International Conference on Computer, Information Technology and Digital Media*, 2013.
- [2] P. Bagchi, D. Bhattacharjee, M. Nasipuri, and D. K. Basu. A novel approach for nose tip detection using smoothing by weighted median filtering applied to 3d face images in variant poses. *CoRR*, abs/1309.4573, 2013.
- [3] P. Bagchi, D. Bhattacharjee, M. Nasipuri, and D. K. Basu. A novel approach in detecting pose orientation of a 3d face required for face. *Computer Science: Computer Vision and Pattern Recognition*, 2013.
- [4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):711–720, July 1997.
- [5] Y. Freund and R. E. Schapire. A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14:771–780, 1999.
- [6] B. Heisele, T. Serre, S. Prentice, and T. Poggio. Hierarchical classification and feature reduction for fast face detection with support vector machines. *Pattern Recognition*, 36:2007–2017, 2003.
- [7] M. Krishnaswamy and G. H. Kumar. Enhancing template security of face biometrics by using edge detection and hashing. *CoRR*, abs/1401.5632, 2014.

- [8] Y. Liu and S. Wang. A virtual teleconferencing system based on face detection and 3d animation in a low bandwidth environment. *International Journal of Imaging Systems and Technology*, 20:323–332, 2010.
- [9] M. D. Marsico, C. Galdi, M. Nappi, and D. Riccio. Firme: Face iris recognition for mobile engagement. *Image and Vision Computing*, 2014.
- [10] Y. Mitsukura, E. Fukumi, and N. Akamatsu. A design of face detection system by using lip detection neural network and skin distinction neural network. In *In proceeding of: Systems, Man, and Cybernetics, 2000 IEEE International Conference on, V: 4*, 2000.
- [11] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. *MIT*, 1998.
- [12] P. Pattanasethanon and C. Savithi. Human face detection and recognition using web-cam. *Journal University of Jordan*, pages 1585–1593, 2012.
- [13] Z. Stone, T. Zickler, and T. Darrell. Toward large-scale face recognition using social network context. *IEEE Xplore*, 98-8:1408–1415, 2010.
- [14] X. Tan and B. Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, 19:1635–1650, 2010.
- [15] M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. *Journal of the Royal Statistical Society, Series B*, 61:611–622, 1999.
- [16] O. Vermeulen, A. Manzanera, and L. Lacassagne. Ultra fast grey scale face detection using vector (simd) programming. In *3rd International Conference on Signal-Image Technology and Internet-based Systems (SITIS 07)*, Shanghai, China, dec. 2007. IEEE.
- [17] P. Viola and M. J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57:136–154, 2004.
- [18] J. Zheng, G. A. Ramirez, and O. Fuentes. Face detection in low-resolution color images. In *Image Analysis and Recognition, 7th International Conference, ICIAR*, 2010.