

Evaluación de la calidad de las imágenes de rostros utilizadas para la identificación de las personas

Heydi Méndez-Vázquez, Leonardo Chang, Dayron Rizo-Rodríguez y
Annette Morales-González

Centro de Aplicaciones de Tecnologías de Avanzadas (CENATAV), La Habana,
Cuba

{hmendez, lchang, drizo, amorales}@cenatav.co.cu

Resumen. La evaluación automática de la calidad de las imágenes de rostros es un tema de gran importancia dentro del desarrollo de sistemas biométricos de reconocimiento de rostros. En este trabajo, se describe un esquema para evaluar si las imágenes de rostros cumplen los parámetros establecidos en la norma ISO/IEC 19794-5, que determinan si tienen valor identificativo y si pueden ser utilizadas en documentos de identificación personal. Se presentan nuevos algoritmos para el análisis de las imágenes y la clasificación de estas respecto a los parámetros que se evalúan en el esquema propuesto. El esquema propuesto se implementa en una biblioteca de vínculos dinámicos (DLL) que brinda ventajas para su uso en diferentes aplicaciones. Se evalúa la propuesta con imágenes de buena y mala calidad y se muestra de manera experimental, la factibilidad de su uso.

Palabras clave. Reconocimiento de rostros, evaluación de la calidad, estándares, procesamiento de imágenes.

Face Image Quality Evaluation for Person Identification

Abstract. Automatic evaluation of face image quality is an important topic in the development of face recognition systems. This paper describes a framework for assessing the conformity of face images with the parameter set in the standard ISO/IEC 19794-5 that determines if given images possess an identification value and can be used in personal identification documents. New algorithms for image analysis and image classification with respect to parameters evaluated in the proposed framework are presented. The proposal is implemented as a dynamic link library (DLL), offering many advantages for the use in different applications. Good and bad quality images are evaluated, and the feasibility of the use of this framework is shown experimentally.

Keywords. Face recognition, quality evaluation, standards, image processing.

1 Introducción

El rostro es la característica biométrica más comúnmente utilizada por los seres humanos. Desde hace varias décadas, las imágenes de rostros han sido utilizadas para identificar a los individuos en los documentos personales. Los avances tecnológicos han permitido que las imágenes digitales del rostro sean además la base de sistemas biométricos importantes como la verificación e identificación de personas, la video-vigilancia y el análisis de las acciones humanas. Con el aumento de los sistemas biométricos basados en las imágenes de rostros, ha surgido la necesidad de definir un formato de datos estándar, que establezca los requisitos indispensables para que las imágenes que se utilizan tengan valor identificativo y que permita la interoperabilidad entre diferentes sistemas.

La norma internacional ISO/IEC 19794-5 [20], creada por el Comité Internacional para los Estándares de Información y Tecnología, y adoptada por la Organización Internacional de la Aviación Civil (ICAO), tiene como objetivo establecer los requisitos de las imágenes de rostros para aplicaciones de reconocimiento de personas y definir un formato para el almacenamiento e intercambio de las fotografías.

Con el fin de mejorar la precisión y efectividad del reconocimiento de rostros, ya sea de manera automática o llevada a cabo por humanos, la norma incluye más de 15 requisitos, que se dividen en tres grupos generales:

- Especificaciones de escena (pose facial, expresión, uso de accesorios).
- Características fotográficas (iluminación, nitidez, distancia focal, exposición, saturación).
- Atributos propios de las imágenes digitales (resolución, nivel de compresión, formato de los archivos, formas de almacenamiento).

En la Figura 1 se muestran ejemplos negativos y positivos de los parámetros que se establecen en esta norma. Mientras más se garantice el cumplimiento de estos requisitos, mayor será el valor identificativo de las imágenes de rostros. Por otra parte, para lograr darle un uso real y eficaz a las fotos plasmadas en documentos de identificación personal, como las tarjetas de identidad, los pasaportes y las licencias de conducción, la mayoría de los gobiernos exigen cada vez con mayor fuerza el cumplimiento obligatorio de muchos de estos parámetros a la hora de tomar las fotografías. Es por ese motivo que la comprobación automática del cumplimiento de estos requisitos, se ha convertido en un tema de gran importancia dentro del desarrollo de sistemas biométricos de reconocimiento de rostros. Varias empresas importantes en el campo del reconocimiento de rostros [31], cuentan con sistemas comerciales para la captura y evaluación de las fotografías siguiendo este estándar [3, 5, 22]. Sin embargo, la mayoría de estos sistemas evalúan sólo algunos de los parámetros en dependencia del propósito específico para el que hayan sido diseñados, tienen elevados costos y no brindan información de los algoritmos utilizados para evaluar cada medida, por lo que son poco flexibles y no permiten una configuración variable de los parámetros utilizados en conjunto para determinar la calidad final de una imagen de rostro.

En este trabajo se presenta un esquema para la validación de la calidad de imágenes de rostros, basado en la evaluación de los parámetros establecidos en la norma ISO / IEC 19794-5. Se describen brevemente los algoritmos utilizados para evaluar cada uno de los parámetros tomados en cuenta en el esquema propuesto y algunos detalles de su implementación. Se proponen adicionalmente

funcionalidades para obtener, a partir de la imagen evaluada, una imagen adecuada para documentos de identidad. Finalmente se presentan los resultados experimentales obtenidos utilizando la herramienta propuesta en la evaluación de un conjunto de imágenes de rostros escogidas para este propósito.

2 Evaluación de la calidad de imágenes de rostros

Entre los requisitos que establece la norma ISO/IEC 19794-5, se seleccionaron para evaluar, los más importantes dentro de las especificaciones de la escena y las características fotográficas, que son los determinantes en el valor identificativo de la imagen de rostro [15], estos aparecen listados en la Tabla 1. Existe otro grupo de parámetros no imprescindibles para que las imágenes tengan valor identificativo, pero recomendables a la hora de utilizar las fotos en documentos de identidad, los cuales se especifican en la Tabla 2.

Entre los requisitos que establece la norma ISO/IEC 19794-5, se seleccionaron para evaluar, los más importantes dentro de las especificaciones de la escena y las características fotográficas, que son los determinantes en el valor identificativo de la imagen de rostro [15], estos aparecen listados en la Tabla 1. Existe otro grupo de parámetros no imprescindibles para que las imágenes tengan valor identificativo, pero recomendables a la hora de utilizar las fotos en documentos de identidad, los cuales se especifican en la Tabla 2.

Cada una de las medidas seleccionadas es analizada de manera individual, sin embargo basta que la imagen no cumpla con uno de los parámetros que aparecen en la Tabla 1, para determinar que no tiene valor identificativo. Por ese motivo se decidió proponer una estructura arbórea, como se muestra en la Figura 2, en la que se tienen en cuenta de manera conjunta los resultados de las evaluaciones individuales para determinar la calidad final de una imagen de rostro. En el árbol que se propone, el orden de evaluación de los distintos parámetros está determinado por la influencia que tiene cada uno de ellos en la evaluación de los restantes.

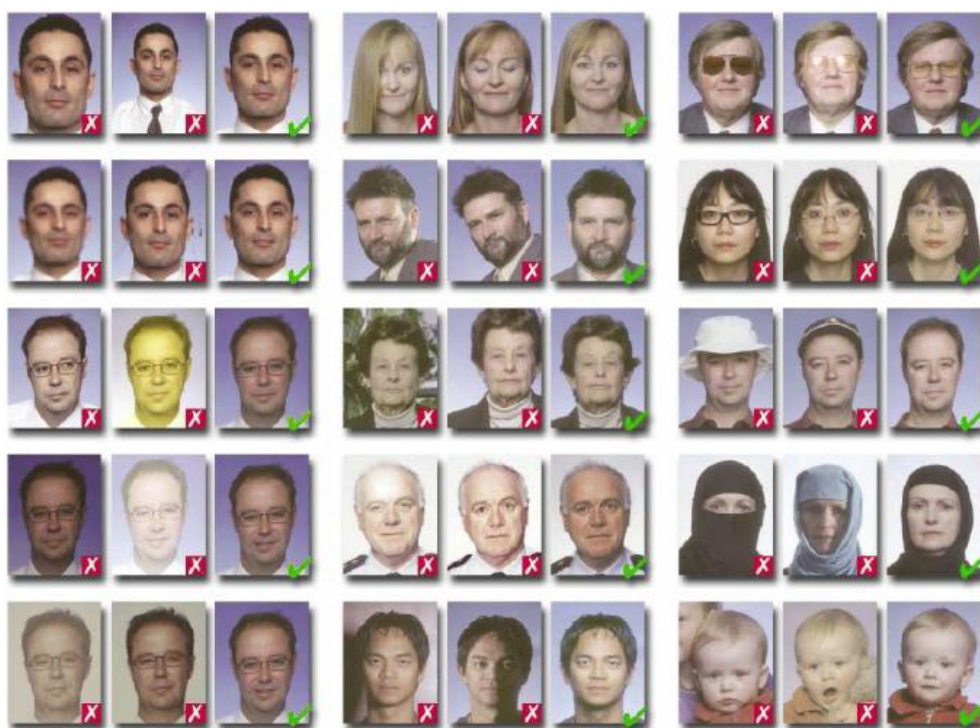


Fig. 1. Ejemplos del cumplimiento de los requisitos especificados en la norma ISO/IEC 19794-5

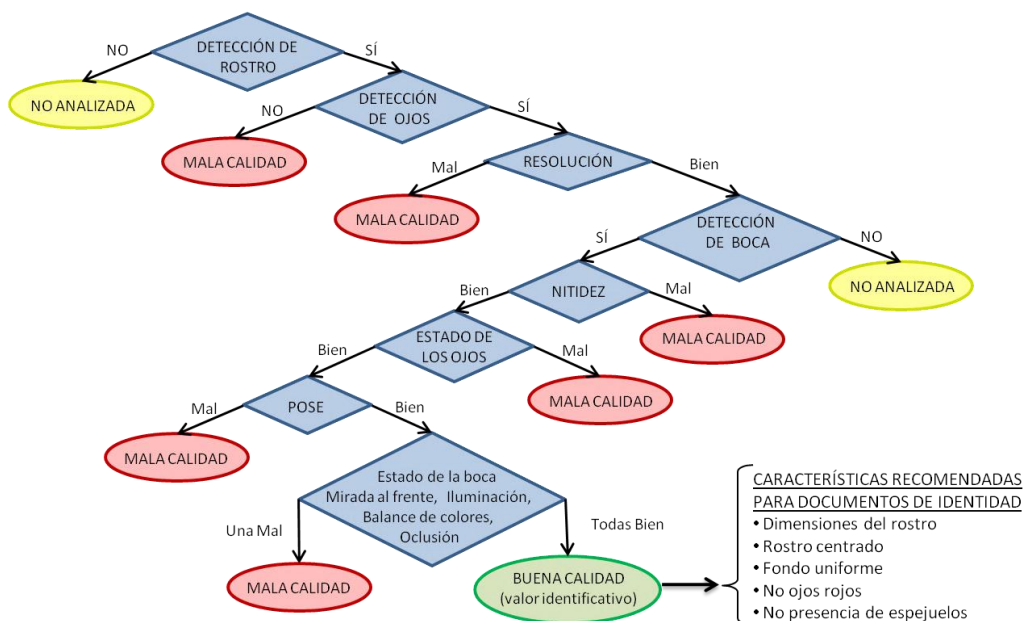


Fig. 2. Estructura arbórea para la evaluación de la calidad de una imagen de rostros

Tabla 1. Parámetros seleccionados que influyen en el valor identificativo de las imágenes de rostros

Parámetro	Especificaciones
Resolución	La distancia entre los centros de los dos ojos debe ser como mínimo de 60 píxeles.
Nitidez	Las imágenes no deben estar borrosas o desenfocadas.
Estado de los ojos	Los ojos deben estar abiertos de manera natural. El iris y la pupila deben estar claramente visibles.
Pose	La imagen debe ser lo más frontal posible.
Expresión de la boca	La boca debe estar cerrada y tener una expresión neutral.
Mirada al frente	Los ojos deben estar mirando al frente.
Iluminación	La luz debe estar distribuida uniformemente en el rostro, el cual no debe estar afectado por sombras, ni regiones brillantes.
Balance de Colores	Los colores deben ser naturales, no saturados y con un contraste adecuado.
Oclusión	Los peinados y otros accesorios no deben obstruir la imagen del rostro o los ojos.

La existencia de problemas en determinados parámetros, causa que otros no puedan ser bien evaluados. Por ejemplo, si la imagen está desenfocada o no es nítida, no hay seguridad en la evaluación con respecto al estado de los ojos y de la boca, porque los píxeles en esas regiones no estarían bien definidos. Por otra parte, esta estructura permite que se mejore la eficiencia, en cuanto al tiempo computacional necesario cuando alguna de las medidas falla y se decide, sin disminuir la eficacia, que la imagen no tiene calidad.

Si se falla en la evaluación de alguno de los parámetros del árbol, la respuesta general de la calidad de la imagen y la de ese parámetro en específico, es negativa. En ese caso, el resto de los parámetros que queden por debajo de ese nodo del árbol, no se evalúan.

Tabla 2. Parámetros adicionales seleccionados, recomendables para documentos de identidad

Parámetro	Especificaciones
Dimensiones del rostro	El largo de la cabeza debe comprender entre 70%-80% del alto de la imagen. La razón entre el ancho de la cabeza y el de la imagen debe encontrarse entre 4:7 y 1:2.
Rostro centrado	El rostro debe estar centrado de manera horizontal y la posición vertical de los ojos debe ser entre el 50 y el 70% del alto de la imagen.
Fondo uniforme	El fondo tras la silueta de la persona debe ser uniforme o tener una transición suave en una dirección.
Ojos rojos	No se admiten fotos con ojos rojos. Los colores deben ser naturales.
No presencia de espejuelos	Solo se permite el uso de espejuelos si la persona normalmente los usa. Estos no deben oscurecer los ojos y deben ser cristales claros y transparentes de manera que las pupilas y el iris queden visibles. No puede haber luces reflejadas en los espejuelos.

Como puede observarse en la Figura 2, el primer paso en el esquema que se propone, es la detección del rostro. Este es el primer paso en cualquier sistema que trabaje con imágenes de rostros, seguido de la detección de los ojos dentro de él. Este paso permite trabajar únicamente con la región del rostro, que generalmente es normalizada geométricamente a partir de las coordenadas de los ojos. La detección se realiza utilizando el algoritmo de Viola y Jones, basado en el clasificador Adaboosting [32], el cual reporta los mejores resultados en la literatura. En el caso de los rostros se utiliza el entrenamiento propuesto por Lienhart [24], mientras que en el de los ojos se utiliza el de Castrillón [9].

Como es lógico y se puede observar en la figura, si en la imagen que se intenta evaluar no

se detecta un rostro, es imposible realizar alguna evaluación. Sin embargo, si luego de detectado el rostro, no es posible detectar los ojos, esto implica que existe algún problema con ellos, es decir, que estén cerrados o sometidos a algún tipo de oclusión. Por tanto, en ese caso, se evalúa la imagen como mala y el resto de los parámetros no se evalúan.

Una vez detectado el rostro y los ojos, es posible obtener toda la información y las regiones de la imagen que son necesarias analizar para evaluar las restantes medidas. Solamente al comparar las coordenadas del centro de los ojos, es posible determinar si estas cumplen con el parámetro indispensable de resolución y en caso de no cumplirlo, se evita analizar innecesariamente el resto de los requisitos.

Antes de continuar con el resto de las evaluaciones, se realiza la detección de la boca, otra zona de referencia importante dentro de los rostros. Para esto se utiliza el mismo algoritmo con el entrenamiento de Castrillón [9], el cual tiene el inconveniente de no poder detectar las bocas que se encuentran ocultas total o parcialmente por causa de bigotes o barbas. En estos casos, no es conveniente emitir un criterio sobre la evaluación de esa región que no está claramente visible y no invalida del todo el uso de la imagen, por lo que tampoco se analiza la imagen de rostro y se deja el criterio al experto humano.

El siguiente elemento a analizar es la nitidez, la cual es decisiva en el resto de las evaluaciones, puesto que en una imagen distorsionada los píxeles están alterados y esto conlleva a errores en el análisis de las distintas regiones del rostro. En el caso de cumplir con el parámetro de nitidez, se prosigue a evaluar el estado de los ojos, el cual es igualmente decisivo al utilizarse sus centros como puntos de referencia para obtener las imágenes de las diferentes regiones del rostro.

Luego se evalúa la pose, la cual, es de los parámetros más importantes en el reconocimiento de rostros y además, si no es correcta, puede influir negativamente en el resultado de la clasificación de medidas como la iluminación, la expresión de la boca y la mirada al frente, que han sido entrenadas con imágenes de rostros frontales, por lo que una pose no frontal

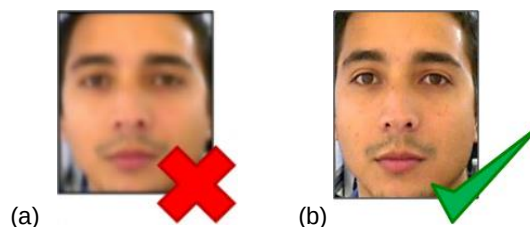


Fig. 3. Representación de la nitidez en la imagen de rostro: a) mala y b) buena

provoca que estos métodos no den una respuesta confiable.

El estado de la boca, la mirada al frente, la iluminación, el balance de colores y la oclusión son medidas independientes que no influyen unas sobre las otras, por lo que se analizan todas en el mismo nivel del árbol. No obstante, éstas se evalúan ordenadas según su costo computacional, de manera que el estado de la boca, que es la que menos costo computacional requiere, es la primera y la oclusión (la más costosa) es la última.

Si como salida del árbol se determina que la imagen tiene valor identificativo, se evalúan adicionalmente las dimensiones del rostro, si este está centrado, si el fondo es uniforme, si sufre del efecto de ojos rojos y si hay presencia de espejuelos. Estas no se tienen en cuenta en la decisión final sobre la imagen, pero de manera individual se da respuesta sobre cada una de ellas para su posible uso en las fotos para documentos de identidad.

A continuación se describen brevemente los algoritmos utilizados para evaluar cada uno de los parámetros. En la mayoría de los casos son contribuciones propias a partir de métodos existentes en la literatura actualizada sobre los diferentes requisitos a evaluar.

2.1 Descripción de la evaluación de cada una de las medidas seleccionadas

Los parámetros de resolución, dimensiones del rostro y el rostro centrado, son muy sencillos de evaluar a partir de la información sobre la región de la imagen que pertenece al rostro y los ojos, es decir, la región devuelta por el algoritmo de

Viola y Jones y las coordenadas determinadas como el centro de los ojos.

Los métodos utilizados para evaluar el resto de los parámetros son más complejos y en casos como el de la iluminación y la pose, son problemas abiertos en el campo del reconocimiento de rostros.

2.1.1 Nitidez

Este parámetro evalúa el grado de distorsión de la imagen. En la Figura 3 se presenta una imagen distorsionada (a) y otra que no tiene este problema (b). Como puede observarse, cuando una imagen está borrosa o desenfocada, pierde los detalles de los bordes.

En una imagen en la que los bordes estén bien definidos existe un gran número de píxeles, pertenecientes a estos bordes, en los que sus píxeles vecinos son muy diferentes. Teniendo esto en cuenta, si se reemplaza cada píxel con el valor medio de las intensidades de los píxeles circundantes, la imagen obtenida, que es una versión suavizada de la imagen original, es muy diferente a esta. Por otra parte, si la imagen es borrosa, las intensidades de los píxeles en pequeñas vecindades es muy similar y la imagen suavizada que se obtiene es muy parecida a la original. Luego, para evaluar la nitidez se seleccionó un método que ha sido utilizado en otras aplicaciones de procesamiento digital de imágenes, el cual consiste en aplicar un filtro de media a la imagen que se está evaluando (en este trabajo se usó un filtro de 3x3 píxeles) para obtener la imagen suavizada, que luego se resta de la imagen original para saber cuánto se diferencian [11].

2.1.2 Estado de los ojos y expresión de la boca

Para el caso del estado de los ojos y la expresión de la boca, se decidió utilizar el mismo algoritmo de clasificación basado en entrenamientos diferentes. Para los ojos se utilizan imágenes pertenecientes a la región de los ojos divididas en dos clases: abiertos y cerrados. En el caso de la boca se utilizan imágenes con esta región, divididas en expresiones neutrales o no neutrales.

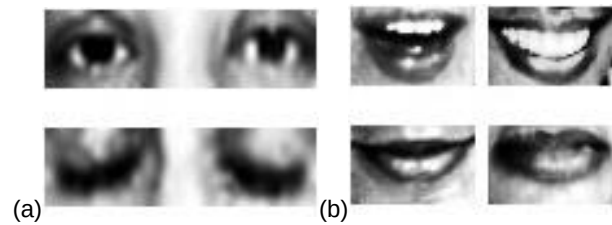


Fig. 4. Ejemplo de imágenes ecualizadas utilizadas para la clasificación de: a) ojos y b) boca

La clasificación se realiza con el método conocido como PCA, por las siglas en inglés de Análisis de Componentes Principales [19], bastante utilizado en aplicaciones de reconocimiento de patrones en imágenes por las ventajas que brinda en la reducción de la dimensionalidad. En PCA se construye una representación de baja dimensión que maximiza tanto como sea posible la varianza en los datos. Esto se logra encontrando una base lineal de los datos donde la varianza es máxima.

Las imágenes que se utilizan en la clasificación se obtienen normalizando geométricamente las imágenes de rostros originales a partir de las coordenadas de los ojos, de manera que solo contengan la región a evaluar, como se puede observar en la Figura 4.

Antes de representar estas imágenes en los subespacios correspondientes, estas son llevadas a escala de grises y preprocesadas con el método de ecualización de histogramas [18], que busca obtener una distribución más uniforme de la cantidad de píxeles asociados a un nivel de intensidad.

Las imágenes de entrenamiento ecualizadas se representan como vectores y se construye con ellos la matriz de covarianza, a partir de la cual se obtienen los valores y vectores propios [19]. Las componentes principales son los vectores propios con mayor valor propio asociado. Se representan entonces los datos en el subespacio generado por dichas componentes principales. Luego, la clasificación de las nuevas imágenes se realiza en ese subespacio, utilizando la regla del vecino más cercano (1-NN) con la distancia euclidiana.

A partir de imágenes de las bases de datos ORL [4], BioID [7] y Essex [14], se creó un

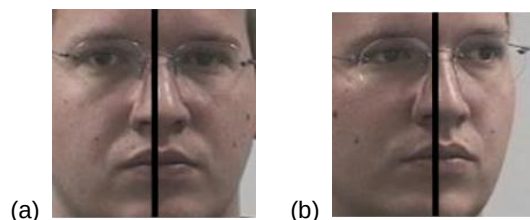


Fig. 5. Relación entre la pose del rostro y la simetría vertical: a) rostro frontal y simétrico, b) rostro no frontal y asimétrico

conjunto de 500 imágenes de ojos, 250 abiertos y 250 cerrados y uno de 500 imágenes de bocas, 250 con expresión neutral y 250 abiertas. Estas imágenes fueron clasificadas con el método 1-NN usando la representación PCA y representaciones basadas en otros métodos de reducción de dimensionalidad lineales como el Análisis Discriminante Lineal (LDA) [6] y el Escalado Multidimensional (MDS) [8] y otros no lineales como el Mapeo Isométrico de Rasgos (Isomap) [30] y el Encajado Local Lineal (LLE) [29]. En todos los casos la clasificación basada en la representación PCA fue la de mejores resultados. Para los ojos, el por ciento de imágenes clasificadas correctamente fue de 99.30% y en el caso de la boca 98.90%.

2.1.3 Pose

Este parámetro se utiliza para determinar si el rostro se encuentra en una pose frontal. Este es uno de los factores que más influye en la efectividad de los algoritmos de reconocimiento de rostros, por lo que es muy importante su evaluación. Por este motivo, llevamos a cabo un estudio detallado de los métodos de estimación de la pose existentes y se compararon varios algoritmos existentes para determinar si el rostro en una imagen tiene una posición frontal [1].

Los mejores resultados fueron los obtenidos con el método propuesto por Gao [16], en el que para evaluar la pose, se analiza la simetría del rostro. Como puede verse en la Figura 5, las desviaciones de la pose del rostro provocan afectaciones en su simetría. Para el análisis de la simetría, se utilizaron los Patrones Binarios Locales (LBP) con la medida de similitud Chi-Cuadrado para comparar las dos mitades del rostro [1].

2.1.4 Mirada frontal

Los métodos encontrados en la literatura para estimar la dirección de la vista, por lo general requieren dispositivos especializados de iluminación adicional y en otros casos necesitan adquirir imágenes de alta resolución.

El método que se propone fue presentado por autores de este trabajo en [2] y se divide en dos pasos fundamentales: la localización del iris y la estimación de la dirección de la vista.

2.1.5 Localización del iris

Dada una representación geométrica del globo ocular, donde el iris está representado por un círculo (B en la Figura 6) inscrito en una elipse que representa la esclerótica (D y E en la Figura 6), se definen β , ω y ψ como:

$$\beta = \sum_{x \in B} f(x), \quad \omega = \sum_{x \in D} f(x), \quad \psi = \sum_{x \in E} f(x).$$

Sea $g(x) = \omega + \psi - \beta$, el iris en el ojo coincide con la región circular inscrita en el cuadrado de coordenadas superiores izquierdas δ y lados de longitud d , donde:

$$\delta = x | \forall x' \neq x, g(x) \geq g(x'). \quad (1)$$

El algoritmo de localización del iris consiste en buscar en la imagen el punto δ . Para su funcionamiento, el algoritmo necesita la información del diámetro d de la región circular que se desea buscar. Este valor depende del tamaño del iris de cada individuo. A partir de un estudio realizado, como parte de este trabajo, se determinó que este valor se puede aproximar en función del tamaño de la región del ojo detectada como $d = 0.31 \times s$, donde s es el ancho de dicha región.

2.1.6 Estimación de la dirección de la vista

Para estimar la dirección de la vista, una vez ubicado el iris usando la técnica anterior, es necesario comparar la suma de las intensidades de los píxeles en las regiones a ambos lados del iris. Varios criterios pueden ser utilizados para estimar el foco de atención usando las regiones de esclerótica visible. En este caso se propone combinar las informaciones de ambos ojos. El

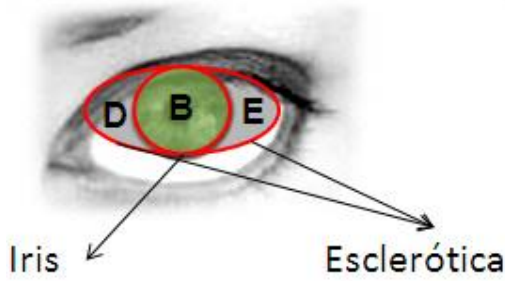


Fig. 6. El Globo Ocular dividido en 3 conjuntos de píxeles

algoritmo consiste en calcular la desviación de las razones de intensidad entre las regiones interiores y las regiones exteriores de ambos ojos. Inicialmente estas razones pueden ser calculadas a partir de las sumatorias de intensidades ω y ψ para cada ojo, de la siguiente forma:

$$r_1 = \frac{\psi_i}{\omega_d}, r_2 = \frac{\psi_d}{\omega_i} \quad (2)$$

donde r_1 y r_2 representan las razones de intensidad del ojo izquierdo y derecho respectivamente.

Sin embargo cuando uno de los ojos se encuentra afectado por una sombra, los valores r_1 y r_2 no son útiles para la estimación. Con el objetivo de hacer el algoritmo robusto a los cambios monotónicos de iluminación, se usan los valores ω y ψ normalizados con respecto a su variación interna.

Sean $\min = \min_{x \in W \cup Z} f(x)$ y $\max = \max_{x \in W \cup Z} f(x)$, las sumatorias normalizadas en el conjunto $W \cup Z$ se definen como:

$$h = \frac{\psi - \min}{n(\max - \min)}, \quad s = \frac{\omega - \min}{n(\max - \min)} \quad (3)$$

A partir de aquí podemos redefinir entonces las razones de intensidad como:

$$r_1 = \frac{h_i}{s_d}, r_2 = \frac{h_d}{s_i} \quad (4)$$

Cuando los ojos se dirigen a la izquierda los valores s_i y s_d son menores que h_d y h_i respectivamente, por tanto r_1 y r_2 se hacen ambas superiores a 1. Cuando la vista se dirige a la derecha sucede lo contrario, y las razones de intensidad se hacen menores que 1.

$$k(v): \mathbb{R}^+ \rightarrow \{\text{frontal}, \text{derecha}, \text{izquierda}\}$$

$$k(v) = \begin{cases} \text{Derecha} & \text{si } v < \mu_d \\ \text{Frontal} & \text{si } \mu_d < v < \mu_i \\ \text{Izquierda} & \text{si } v > \mu_d \end{cases} \quad (5)$$

Usando la media de las razones de intensidad (v) y mediante dos umbrales (μ_d y μ_i) se determina entonces si la mirada es frontal, derecha o izquierda, según la Ecuación 5.

Los umbrales μ_d y μ_i fueron entrenados usando un conjunto de muestras etiquetadas manualmente, aceptándose como vista frontal los valores en un rango entre 0.7 y 1.3.

2.1.7 Iluminación

La iluminación es otro de los factores que más influye en el desempeño de los algoritmos de reconocimiento de rostros. A pesar de que el tema del reconocimiento de rostros bajo condiciones variables de iluminación ha sido ampliamente abordado en la literatura, muy poco se encuentra sobre la evaluación de este parámetro en las imágenes de rostros.

Se desarrolló para este propósito un nuevo algoritmo basado en el análisis local del comportamiento de la iluminación sobre el rostro, debido a que esta afecta de manera diferente las distintas partes de la cara [28]. Primeramente, se ubica de manera automática una malla triangular sobre la imagen a analizar como puede observarse en la Figura 7, en la que cada triángulo define una región con una luminancia aproximadamente constante. De cada región, se determina ese valor promedio de luminancia y se compara con el valor esperado para una imagen de buena calidad. Este clasificador, MPQ, determina la probabilidad media de que las regiones del rostro estén bien iluminadas.

Adicionalmente, se utiliza una red neuronal, NNQ, para clasificar específicamente el comportamiento de la iluminación en las regiones correspondientes al área de la nariz, por estar identificada como la zona del rostro donde más

sobresalen las variaciones de iluminación [28]. Si el resultado de ambos clasificadores coincide, esta será la clasificación final de la imagen, de lo contrario, es necesario determinar cuál de los dos tiene la decisión correcta.

Con este objetivo se define lo que se llamará una *función de confianza* para cada uno de los dos clasificadores, la cual retornará un valor en el intervalo [0;1] que represente la seguridad de la clasificación, donde 1 significa una alta confianza y 0 muy poca seguridad en el resultado obtenido por el clasificador.

La función de confianza para el clasificador de la red neuronal, NNQ, se define de la siguiente forma:

$$NNC = |NB - NM|, NNC \in [0; 1] \quad (6)$$

donde NB es la salida de la neurona que representa a las imágenes bien iluminadas y NM es la salida de la que representa a las mal iluminadas. Esta función representa cuán bien separados y por tanto cuán confiables son los valores obtenidos por cada una de las neuronas de salida de la red neuronal.

Para MPQ, la función de confianza se define como:

$$MPC = \begin{cases} 0.99 + \frac{|\bar{P} - \alpha|}{100 \cdot \mu} & \text{si } |\bar{P} - \alpha| > \delta \\ \frac{0.99 \cdot |\bar{P} - \alpha|}{\delta} & \text{si } |\bar{P} - \alpha| \leq \delta \end{cases} \quad (7)$$

donde $\bar{P}$ es la probabilidad media de que cada uno de los triángulos de la imagen esté bien iluminado y δ representa el valor para el cual el clasificador tiene una confianza de 0.99; o sea, la seguridad de que una imagen con $\bar{P} > \alpha + \delta$ está bien iluminada y de que una imagen con $\bar{P} < \alpha + \delta$ está mal iluminada. La función está diseñada para asignarle menos confianza a los valores de $\bar{P}$ en el intervalo $[\alpha - \delta; \alpha + \delta]$, siendo menor la confianza mientras más cerca se encuentre $\bar{P}$ de α . El valor μ representa el supremo de las funciones de densidad que aproximan el comportamiento de la buena iluminación de cada región triangular. Si se analiza la ecuación 7 se puede apreciar que el primer término de la función toma valores en el

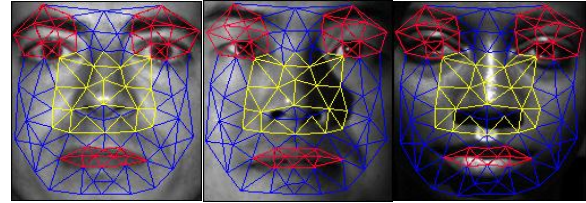


Fig. 7. Regiones del rostro utilizadas para analizar la influencia de la iluminación

intervalo [0.99 ; 1] mientras que la segunda lo hace en el intervalo [0 ; 0.99].

El valor de δ en la ecuación es determinado experimentalmente. Se escoge un valor para el cual más del 90% de los valores de $\bar{P}$ de las imágenes buenas del conjunto de entrenamiento sean mayores que $\alpha + \delta$ y que a la vez más del 90% de los valores de $\bar{P}$ de las imágenes de entrenamiento seleccionadas como malas sean menores que $\alpha - \delta$.

La calidad de una imagen respecto a la iluminación, estará finalmente determinada por la función:

$$CCQ = \begin{cases} MPQ \times NNQ & \text{si } MPQ = NNQ \\ MPQ & \text{si } MPC > NNC \\ NNQ & \text{si } MPC \leq NNC \end{cases} \quad (8)$$

Imágenes de las bases de datos de pruebas internacionales Yale B Extendida [23] y XM2VTS [26] fueron utilizadas para entrenar y evaluar el funcionamiento del método propuesto en la clasificación de imágenes bien y mal iluminadas.

A partir de estas bases de datos se creó un conjunto de entrenamiento con 200 imágenes con buena iluminación y 200 imágenes afectadas por variaciones en la iluminación. Mientras que para las pruebas, se creó un subconjunto con 2400 imágenes buenas y otro que contiene 1400 imágenes malas.

Las imágenes en el conjunto de pruebas, fueron clasificadas con el método propuesto y con el único método encontrado en la literatura para evaluar la calidad de una imagen de rostro respecto a la iluminación, basado en la simetría del rostro [16]. La efectividad total clasificando tanto imágenes buenas como malas con el método propuesto fue de 96.77%, mientras que

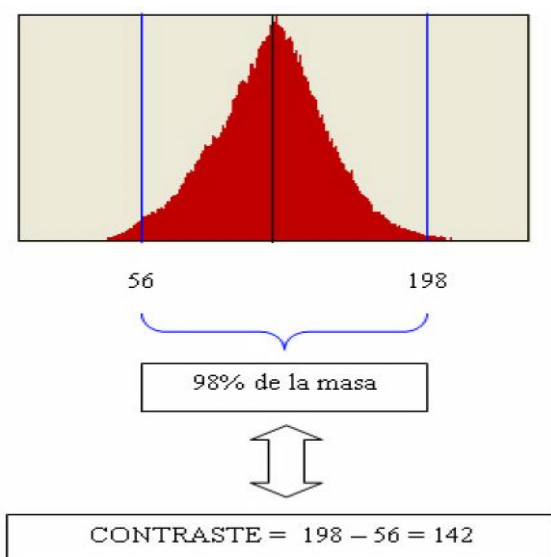


Fig. 8. Esquema del cálculo de contraste

solo se obtuvo un 63.71% con el método basado en la simetría.

2.1.8 Balance de Colores

Este parámetro se define para evaluar el balance de los colores en la imagen, por tanto está relacionado con características fotográficas de la misma. En esta evaluación se tienen en cuenta el Contraste, la Saturación y la Exposición como indicadores que se incluyen dentro de la estimación del comportamiento de los colores de la imagen.

El contraste es la diferencia relativa en intensidades, entre un punto de una imagen y sus alrededores. La medida seleccionada para evaluar el contraste se describe en [21] y se basa en el histograma de las intensidades de los píxeles de la imagen en el modelo RGB. Inicialmente se localiza el centro de masa del histograma y se calcula el valor de la amplitud del 98% de la masa partiendo de dicho centro. En este caso se considera que un 2% representa el ruido en la imagen. El contraste está dado entonces por la amplitud del 98% de la masa de los píxeles, como se ejemplifica en la Figura 8.

Para la evaluación de la saturación y la exposición se hace la conversión de la imagen

del espacio de colores RGB al espacio HSL. En el espacio HSL se trabaja con las componentes saturación y luminancia, esta última se utiliza para la evaluación de la exposición. En ambos casos se analiza el valor promedio de estos parámetros en la imagen.

Por cada uno de estos tres indicadores se definen dos umbrales, uno mínimo y uno máximo. El intervalo definido por estos umbrales en cada uno de los indicadores, representa los valores de comportamiento para imágenes de buena calidad, en este caso con colores naturales. Para determinar estos umbrales, se formó un conjunto de 1000 imágenes de bases de datos de pruebas internacionales capturadas en condiciones controladas [14, 26] y se calcularon los valores mínimo y máximo obtenidos para cada uno de los tres factores. Luego, para validar y refinar estos valores, fue necesario crear 6 subconjuntos de imágenes, representativo de cada una de las posibles afectaciones: sobre-contraste, sub-contraste, sobre-saturación, sub-saturación, sobre-exposición, sub-exposición.

Al no existir una base de datos con este propósito, se tomaron imágenes a 50 personas, con una cámara digital Logitech QuickCam Orbit AF [25], que es capaz de automáticamente ajustar la calidad de la imagen capturada en cuanto a estos factores. Denominamos a este conjunto de imágenes BSCE (buenas condiciones de saturación, contraste y exposición).

A partir de las imágenes en BSCE se crean dos subconjuntos más, uno para caracterizar las alteraciones de sobre afectación y otro para las de sub afectación, por cada una de las medidas. En el caso del contraste y la saturación, se modifican manualmente estos factores en las imágenes utilizando un software de procesamiento digital de imágenes.

En el caso de la exposición no se tienen herramientas automáticas para variar este factor de calidad, pero la cámara utilizada sí posee una opción para afectar la imagen con respecto a dicho parámetro, tanto para sobre exposición como para sub exposición. Luego, en el momento de retratar a cada individuo, además de la imagen de buena calidad, se tomó una imagen sometida a condiciones de sobre exposición y otra a sub exposición. En la Tabla 3 se muestra

Tabla 3. Ejemplos de imágenes en cada uno de los siete subconjuntos

un ejemplo de las imágenes en cada uno de los siete subconjuntos.

Para determinar finalmente la calidad respecto al balance de colores de una imagen, se evalúan estos tres factores y se emite una estimación final, donde la buena calidad está representada por la evaluación positiva de los tres, siendo de mala calidad en caso contrario.

2.1.9 Oclusión

La oclusión es un factor que influye en los algoritmos de reconocimiento de rostros, pues implica que se oculten rasgos que pueden ser decisivos para este propósito. Los métodos existentes en la literatura, la mayoría realizan análisis globales de la imagen de rostro y por lo general, confunden otras variaciones en la imagen del rostro, principalmente de iluminación y expresión, con problemas de oclusión.

En este trabajo se propone un nuevo método basado en el análisis local de las regiones del rostro para determinar grupos de regiones con un comportamiento desviado de lo normal. Para esto la imagen es dividida en $M \times N$ cuadrículas regulares y en dependencia de la cantidad de regiones detectadas como erróneas y la forma en que éstas estén agrupadas en la imagen, se decide si hay un problema de oclusión o no.

A partir de un entrenamiento realizado usando imágenes con oclusión y sin oclusión etiquetadas manualmente, se determinó el umbral para la clasificación y además se determinó que los mejores resultados eran obtenidos para regiones de 8×8 píxeles. En la Figura 9 se muestran algunos ejemplos de las regiones detectadas con problemas, en imágenes en las que hay oclusión.

2.1.10 Fondo uniforme

El algoritmo para evaluar la uniformidad del fondo, se diseñó a partir de un grupo de resultados encontrados en la literatura sobre segmentación de imágenes de rostros. El proceso seguido se resume en la Figura 10.

Primeramente, se utiliza una plantilla que tiene un área de fondo preestablecida. Cuando esta plantilla se adapta a la imagen de rostro, el área que es considerada como fondo se evalúa según la desviación estándar de la intensidad de los píxeles que la conforman. Si el resultado de esta técnica es que el fondo no es uniforme, se entra en más detalles para asegurar que la plantilla no haya incluido como fondo parte de la persona (por ejemplo, hombros o pelo). En este caso se segmenta el pelo de la persona mediante un análisis frecuencial y de color del mismo [10] y luego se segmenta la piel utilizando un modelo de color elaborado a partir de la propia imagen.



Fig. 9. Ejemplos de detección de regiones erróneas en imágenes con problemas de oclusión

La silueta total (pelo y piel) es analizada en cuanto a su convexidad y compacidad para determinar si la segmentación fue exitosa. Si la segmentación es mala, se determina que el fondo no es uniforme ya que objetos del fondo fueron segmentados junto con la persona. Si es buena, se analiza el área exterior de la silueta (fondo), para determinar si la desviación estándar del color de los píxeles se encuentra bajo el umbral designado para identificar uniformidad de color.

2.1.11 Ojos rojos

Uno de los problemas que se puede presentar al capturar imágenes de rostros es la presencia del efecto conocido como “ojos rojos”, consistente en

la aparición de pupilas rojas en fotos tomadas con flash en ambientes de poca luz y con un flash situado cerca de la lente de la cámara.

La evaluación de este parámetro parte de la posición detectada de los ojos. Teniendo una región rectangular G , que delimita el área de un ojo representada en un espacio de colores RGB (como se muestra en la Figura 11a), se calcula el grado de coloración roja presente en dicha región ($redness_G$), utilizando la ecuación propuesta en [17]:

$$redness_G = \sum_x \sum_y \frac{R(x,y)^2}{G(x,y)^2 + B(x,y)^2 + K} \quad (9)$$

donde $R(x,y)$ representa el valor de rojo del píxel en la posición (x,y) , de la misma forma $G(x,y)$ y $B(x,y)$ representan respectivamente los valores de verde y azul en dicho píxel, y K es una constante para evitar singularidades.

Este modo de determinar el predominio de color rojo para emitir un criterio sobre la presencia o no de ojos rojos, falla cuando son otros motivos los que contribuyen a la coloración roja de la región, como pueden ser coloración

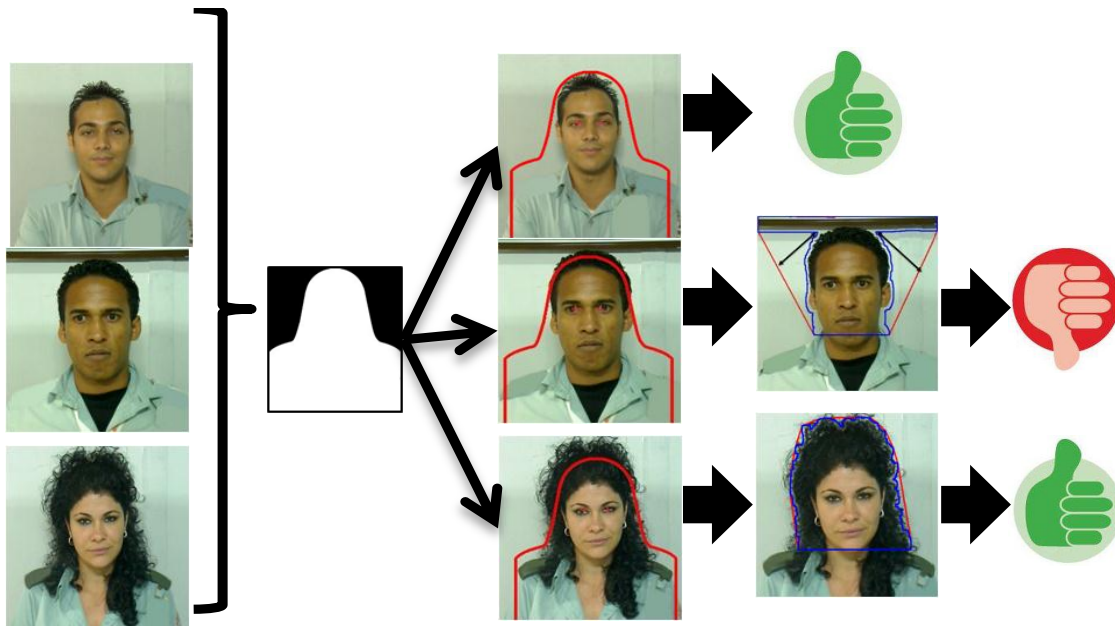


Fig. 10. Pasos del proceso para evaluar la uniformidad del fondo

roja de la piel que circunda los ojos, enrojecimiento de la esclerótica o presencia de maquillaje rojo, como se observa en la Figura 11b. Para corregir este problema, se dividió la región G en dos regiones mostradas en la Figura 11c: una externa a la región del iris (G1) y otra que delimita el área de iris y pupila (G2). Se calcula el grado de coloración roja en cada una de las regiones por separado y lo que se utiliza para determinar la presencia o no de ojos rojos es la relación entre los valores de ambas regiones ($relative_redness_G$), que da una medida del contraste de color rojo existente entre ellas. Finalmente se compara el valor $relative_redness_G$ obtenido con un umbral que determina si en la región hay predominio del color rojo.

Para calcular el umbral que determinará la presencia o no de ojos rojos se tomaron dos conjuntos de imágenes, el primero donde todas presentaban ojos rojos y el segundo donde todas tenían una coloración de iris y pupila normal. Se calculó para todas las imágenes el valor de $relative_redness_G$ y con un listado ordenado de estos valores fue posible determinar el valor de $relative_redness_G$ que corresponde a cada error de falsa aceptación (clasificar una imagen con presencia de ojos rojos cuando en realidad no los tiene) y verdadero rechazo (clasificar una imagen como normal cuando en realidad presenta ojos rojos). El valor de $relative_redness_G$ escogido como umbral es el corresponde a una tasa de errores de falsa aceptación igual a la de verdadero rechazo.

2.1.12 Presencia de espejuelos

Para determinar si el individuo representado en la imagen porta espejuelos o no, se desarrolló un algoritmo basado en la propuesta de Park [27] para eliminar de forma automática los espejuelos en imágenes de rostros. Para esto se utiliza el método PCA, entrenando con imágenes de rostros representativas de las regiones de los ojos sin espejuelos. Luego, cuando se evalúa cada imagen en particular, se halla el error de reconstrucción de la imagen a partir del subespacio entrenado, el que será mayor para las imágenes con espejuelos, ya que estas no están representadas en la muestra de entrenamiento.

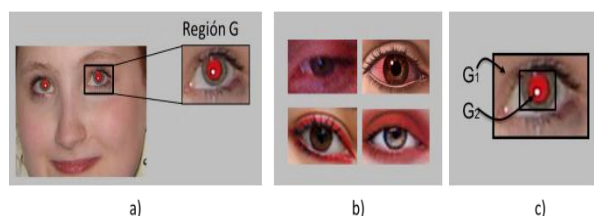


Fig. 11. a) Región G para la evaluación de ojos rojos. b) Causas que aumentan la coloración roja en G. c) Regiones G1 y G2 en que se divide G para hallar el contraste de color rojo

2.2 Algunos detalles de implementación

El proceso de evaluación de la calidad de una imagen de rostro descrito en la sección anterior, se implementa en una biblioteca de vínculos dinámicos o DLL, escrita en lenguaje C++, utilizando funcionalidades de la Biblioteca de Visión por Computadoras de Código Abierto (OpenCV) [3]. El tiempo promedio que demora el análisis de una imagen de 800x600 píxeles en la que se evalúan todos los parámetros, en una PC con procesador Intel Core 2 Duo a 2.4 GHz y con 2 Gb de memoria RAM, es de 850 milisegundos. Esto permite que pueda ser utilizada en cualquier aplicación de reconocimiento de rostros.

La DLL brinda una serie de facilidades para la evaluación de la calidad de imágenes de rostros. Primeramente presenta una estructura de implementación en la que es posible, de manera fácil y rápida, añadir métodos nuevos o modificar los existentes para evaluar cada parámetro. Los usuarios pueden modificar los valores de los umbrales utilizados por cada una de las medidas, permitiendo que sean ajustados de una forma más adecuada a los distintos escenarios donde se requiera su aplicación.

A estos se le brinda la información de la respuesta de cada una de las medidas evaluadas, además de la calidad general de la imagen. Por otro lado, se brinda la posibilidad de activar o desactivar las diferentes medidas que se deseen evaluar, respetando siempre el orden estructural en forma de árbol que se ha explicado anteriormente. Esto es otra opción para lograr una adaptabilidad del componente a las necesidades particulares de cada usuario.

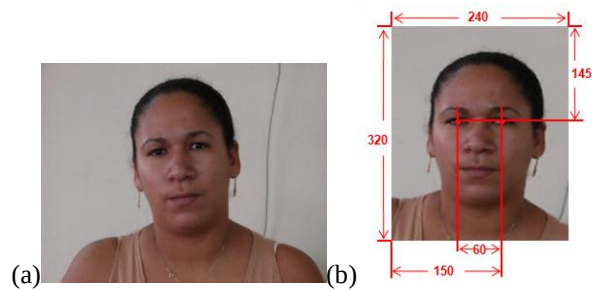


Fig. 12. Ejemplo de imagen: a) capturada y b) normalizada a una plantilla predefinida

La imagen, para ser evaluada por la DLL, debe estar en escala de colores. La DLL puede recibir lo mismo la dirección en disco de la imagen, que la información de la misma en memoria. Además de los resultados de la evaluación, el componente devuelve las coordenadas de los ojos, situados en el rostro detectado en la imagen, las cuales pueden ser corregidas por los usuarios para una evaluación más precisa y utilizadas como punto de partida para cualquier aplicación de reconocimiento de rostros.

3 Obtención de imágenes para documentos de identidad

Una de las principales aplicaciones que tiene la DLL que se propone es la obtención de imágenes adecuadas para documentos de identidad. Para utilizar una imagen de rostro evaluada como buena, como fotografía de un documento de identidad, generalmente se requiere que esta tenga un tamaño predefinido. Por otra parte, es deseable que todas las imágenes tengan de fondo un color único y constante, preferiblemente blanco, ya sea para ahorrar tinta en la impresión o para homogeneizar los fondos de las imágenes tomadas en distintas condiciones. Con el fin de responder a estos objetivos, de manera auxiliar hemos desarrollado e insertado funcionalidades a nuestra DLL con estos propósitos.

En el primer caso, normalizar de manera automática todas las imágenes a un tamaño fijo en el que todas las caras queden en la misma posición, es un proceso sencillo que solamente

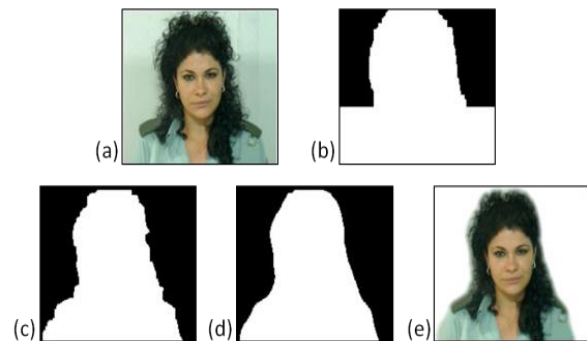


Fig. 13. Pasos para eliminar el fondo en una imagen de rostro

requiere realizar un conjunto de transformaciones afines de rotación, traslación y escala partiendo de las coordenadas de los ojos detectadas y predefiniendo la posición de estos en la plantilla a utilizar, como puede observarse en la Figura 12.

Por otra parte, para eliminar el fondo, las imágenes ya deben haber sido evaluadas positivamente en cuanto a la presencia de fondo uniforme. Para la segmentación de la silueta se utiliza el mismo esquema descrito en la Sección 2.1.8 para segmentar el pelo y la piel. Una vez obtenida la silueta que comprende pelo y piel (ver Figura 13b) se debe segmentar la parte de los hombros de la persona, pues esta no está incluida en la silueta inicial detectada. Encontrar la región de los hombros supone varios problemas, ya que los vestuarios de las personas pueden variar mucho, tanto en estilos como en colores y texturas, por lo que es difícil crear un modelo que sea capaz de describir toda esta área. Además, se le suma el problema de que la pose de la persona puede no ser correcta y no se pueden hacer supuestos sobre la posición o simetría de los hombros en la imagen. Se decidió entonces modelar el color del fondo obtenido a partir de la primera segmentación y utilizando este modelo de color se segmenta el fondo. Con este paso se obtiene una nueva silueta que ya incluye los hombros, como se muestra en la Figura 13c. Esta nueva silueta se suaviza (Figura 13d) y se elimina de la imagen original la parte de fondo que esta delimita (Figura 13e). En la DLL se brindan además facilidades para establecer el color de fondo que se desee.

Tabla 4. Resultados en la evaluación de cada medida de manera individual

Parámetro	Cantidad Imágenes	
	Buenas	Malas
Resolución	5082	78
Nitidez	4235	925
Estado de los ojos	4796	391
Pose	3057	2103
Expresión de la boca	2368	2792
Mirada al frente	4877	283
Iluminación	4078	1082
Balance de Colores	4525	635
Oclusión	4991	169
Dimensiones del rostro	4844	316
Rostro centrado	4778	382
Fondo uniforme	2710	2450
Ojos rojos	65	65
No presencia de espejuelos	5062	98

4 Resultados experimentales

La experimentación para comprobar el funcionamiento de la propuesta consta de dos fases. En una primera fase, se comprueba la eficacia de cada una de las medidas de manera individual, con imágenes clasificadas manualmente como buenas y malas por cada parámetro. Luego, se comprueba de manera general el funcionamiento de la evaluación de las imágenes divididas en dos grupos, uno con imágenes en la que se cumplen todos los parámetros y otro con imágenes que presentan problemas en al menos un parámetro. Se utilizaron para este propósito un conjunto de 5160 imágenes obtenidas mediante un sistema de captura en vivo en explotación. Adicionalmente se usaron 65 imágenes propias de los autores, con ojos rojos, para la clasificación de esta medida, ya que al no utilizarse una cámara con flash en el sistema de captura en vivo, ninguna imagen está afectada por este factor. En la Tabla 4 se muestra la distribución de la cantidad de

imágenes clasificadas manualmente por expertos en cada uno de los parámetros a evaluar.

Los resultados obtenidos en la primera fase de experimentación, pueden ser observados en la Tabla 5, donde se presentan además los resultados obtenidos con dos sistemas comerciales a los que llamaremos “A” y “B”. Por cada una de las medidas se especifica el por ciento de error en la clasificación de imágenes malas (EM) y el por ciento de error en la clasificación de imágenes buenas (EB).

Como puede observarse en la Tabla 5, el sistema “B” sólo evalúa aquellos parámetros relacionados con la dimensión y proporciones del rostro en la imagen y ni siquiera con esos pocos parámetros logra ser más preciso que el nuestro. Por su parte, el sistema “A”, es bastante similar a la propuesta en cuanto a los parámetros que se evalúan, pero nuestros métodos obtienen mejores resultados, muy superiores en la mayoría de los casos. Solamente al evaluar la expresión de la boca obtenemos un mayor porcentaje de error en la clasificación de imágenes buenas y esto se debe fundamentalmente a errores en la detección precisa cuando la persona tiene bigotes o los labios muy gruesos. Puede apreciarse que en nuestra propuesta, el error en la clasificación de imágenes buenas es mayor que el de imágenes malas. Además de la selección de los umbrales para minimizar el error de permitir pasar imágenes con problemas, muchas de las imágenes utilizadas para realizar los experimentos están bien con respecto al parámetro evaluado pero mal respecto a otro.

Como se mencionó anteriormente, un parámetro puede influir negativamente en la decisión de otro, por ejemplo, en una imagen con una pose no frontal, el pelo puede estar ubicado en una región que normalmente corresponda a la cara y esto puede provocar que se determine que hay oclusión cuando en realidad no la hay. Este problema se resuelve de cierta forma con la utilización de la estructura jerárquica definida, en la que todas las medidas son tomadas en cuenta pero analizadas según el orden de influencia de unas sobre otras.

Para comprobar la eficacia en la evaluación tomando en cuenta la estructura de árbol definida, de las 5160 imágenes, se utilizaron 3600 imágenes con problemas en uno o varios

Tabla 5. Resultados en la evaluación de cada medida de manera individual

Parámetro	Propuesta		Sistema A		Sistema B	
	EM (%)	EB (%)	EM (%)	EB (%)	EM (%)	EB (%)
Resolución	0.0	0.0	0.0	0.0	0.0	0.0
Nitidez	12.4	14.65	43.06	24.62	-	-
Estado de los ojos	3.90	13.40	3.84	17.84	-	-
Pose	9.03	19.80	8.70	20.67	-	-
Expresión de la boca	2.75	22.00	2.93	13.20	-	-
Mirada al frente	3.70	11.90	0.87	57.65	-	-
Iluminación	2.94	6.71	21.92	58.46	-	-
Balance de Colores	0.95	3.75	3.75	77.33	-	-
Oclusión	7.22	13.70	-	-	-	-
Dimensiones del rostro	0.0	0.0	0.0	0.0	0.0	0.0
Rostro centrado	0.01	0.03	36.12	27.14	3.02	3.85
Fondo uniforme	4.9	12.2	20.93	20.23	-	-
Ojos rojos	5.00	16.67	6.67	18.33	-	-
No presencia de espejuelos	1.94	2.48	9.30	5.15	-	-
Resolución	0.0	0.0	0.0	0.0	0.0	0.0

parámetros y 885 sin ningún tipo de afectación. Los por cientos de errores cometidos en determinar si estas imágenes poseen o no valor identificativo, se muestran en la Tabla 6.

Como puede apreciarse, en general disminuyen los por cientos de ambos errores, el de clasificar erróneamente imágenes buenas y el de permitir imágenes malas. Se puede apreciar que se mantiene menor el error de permitir

Tabla 6. Resultados en la evaluación general

Calidad de la imagen	EM (%)	EB (%)
con valor identificativo	4.11	11.29

imágenes malas, lo cual es lo más deseable por el posterior uso de estas imágenes en sistemas de identificación. Al analizar el 4.11% de error en la clasificación de imágenes malas, pudimos comprobar que correspondía fundamentalmente a requisitos que no son decisivos en el valor identificativo de la foto. De 148 imágenes malas clasificadas como buenas, 102 solo tenían el problema de que el rostro no estaba centrado, 7 tenían espejuelos, pero de marcos finos y sin reflejar la luz y 58 no tenían el fondo uniforme. Solamente 10 imágenes (0.28%) tenían problemas realmente en cuanto a la pose y la expresión.

5 Conclusiones

En este trabajo, se describe un esquema que permite realizar la evaluación automática de la calidad de una imagen de rostro. Este paso preliminar es de gran importancia para la mayoría de los sistemas biométricos basados en el análisis del rostro.

El esquema propuesto permite comprobar el cumplimiento de nueve de los requisitos definidos en la ISO/IEC 19794-5 para garantizar el valor identificativo de las imágenes de rostros y adicionalmente, cinco parámetros para determinar si las fotos pueden ser utilizadas para documentos de identidad. La mayoría de los métodos utilizados para evaluar cada uno de los parámetros, son contribuciones propias a partir de métodos existentes en la literatura actualizada sobre los diferentes requisitos a evaluar.

Adicionalmente, las fotos usadas en documentos de identidad se pueden ajustar a una plantilla predefinida y eliminarles el fondo, a partir de las propuestas realizadas.

La DLL implementada a partir del esquema propuesto, permite que pueda ser utilizado en cualquier aplicación de reconocimiento de

rostros, activando cada uno de los parámetros que se desee evaluar. Tiene además una estructura en la que de manera fácil y rápida es posible añadir nuevos métodos o modificar los existentes.

Los resultados experimentales mostraron que la propuesta puede ser utilizada para aplicaciones reales, ya que tarda menos de un segundo en evaluar cada imagen y los porcentajes de error obtenidos son aceptables, aunque aún puede trabajarse en la mejora de cada uno de los métodos individuales.

Referencias

1. **Alvarez, J. & Chang, L. (2009).** Análisis de la simetría facial como medida de estimación de la pose del rostro. *Memorias del evento COMPUMAT*, ISSN 1728-6042, Noviembre 2009.
2. **Alvarez, J., Chang, L., & Silva, F. (2009).** Método de estimación de la dirección de la vista basado en rasgos oculares. *Memorias del evento COMPUMAT*, ISSN 1728-6042, Noviembre 2009.
3. **Arévalo, V., González, J., & Ambrosio, G. (2005).** OpenCV: La Librería Open Source de Visión Artificial, *Linux Free Magazine*, 10, 141–147.
4. **AT&T Laboratories Cambridge (s.f.).** The ORL Database of Faces. Retrieved from <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.
5. **ICAO Pack (s.f.).** Retrieved from <http://www.aware.com/biometrics/icaopack.htm>.
6. **Belhumeur, P.N., Hespanha, J.P., & Kriegman, D.J. (1997).** Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 711–720.
7. **BioID Face Database (s.f.).** Retrieved from <http://www.bioid.com/downloads/software/bioid-face-database.html>.
8. **Borg, I. & Groenen, P.J.F. (2005).** *Modern Multidimensional Scaling: Theory and Applications*. New York: Springer.
9. **Castrillón-Santana, M., Lorenzo-Navarro, J., Déniz-Suárez, O., Isern-González, J., & Falcón-Martel, A. (2005).** Multiple face detection at different resolutions for perceptual user interfaces. *Pattern Recognition and Image Analysis. Lecture Notes in Computer Science*, 3522, 767–808.
10. **Rousset, C. & Coulon, P.Y. (2008).** Frequential and color analysis for hair mask segmentation. *15th IEEE International Conference on Image Processing (ICIP 2008)*, San Diego, CA, USA, 2276–2279.
11. **Chang, L., Rodés, I., Méndez, H., & Del Toro, E. (2008).** Best-Shot Selection for Video Face Recognition Using FPGA. *Progress in Pattern Recognition, Image Analysis and Applications. Lecture Notes in Computer Science*, 5197, 543–550.
12. **Chen, V.C. (2004).** Evaluation of Bayes, ICA, PCA and SVM Methods for Classification. *RTO SET Symposium on Target Identification and Recognition Using RF Systems*, Oslo, Norway, 37–1–37-12.
13. **FaceVACS - PortraitAcquisition:** Acquiring and assessing digital portraits for photo-id documents (s.f.). Retrieved from <http://www.cognitec-systems.de/FaceVACS-PortraitAcquisition.15.0.html>.
14. **Spacek, L. (2008).** Computer Vision Science Research Projects. Retrieved from <http://cswww.essex.ac.uk/mv/allfaces/index.html>.
15. **Ferrara, M., Franco, A., & Maltoni, D. (2008).** Evaluating systems assessing face-image compliance with ICAO/ISO standards. *Biometrics and Identity Management. Lecture Notes in Computer Science*, 5372, 191–199.
16. **Gao, X., Li, S.Z., Liu, R., & Zhang, P. (2007).** Standardization of Face Image Sample Quality. *Advances in Biometrics. Lecture Notes in Computer Science*, 4642, 242–251.
17. **Gaubatz, M. & Ulichney, R. (2002).** Automatic Red-Eye Detection and Correction. *2002 International Conference on Image Processing (ICIP 2002)*, Rochester, New York, USA, 1-804–1-807.
18. **Gonzalez, R.C. & Woods, R.E. (2008).** Intensity Transformations and Spatial Filtering. *Digital Image Processing 3rd Edition*, (104–198), Upper Saddle River, N.J.: Prentice Hall.
19. **Hotelling, H. (1933).** Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441.

20. **INCITS, "Face Recognition Format Data Interchange"**, InterNational Committee for Information Technology Standards (INCITS) Secretariat. Information Technology Industry Council, 2004.
21. **Ke, Y., Tang, X., & Jing, F. (2006)**. The Design of High-Level Features for Photo Quality Assessment. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, New York, USA, 419–426.
22. **Kee Square, "Morpheus ICAO"**. Retrieved from http://www.keesquare.com/morpheus_icao_en.html
23. **Lee, K.C., Ho, J., & Kriegman, D.J. (2005)**. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5), 684–698.
24. **Lienhart, R., Kuranov, A., & Pisarevsky, V. (2003)**. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. *Pattern Recognition. Lecture Notes in Computer Science*, 2781, 297–304.
25. **Logitech QuickCam® Orbit AF (s.f.)**. Retrieved from <http://www.logitech.com/en-us/38/3480>.
26. **Messer, K., Matas, J., Kittler, J., Luetlin, J., & Maitre, G. (1999)**. XM2VTSDB: The extended M2VTS database. *2nd International Conference on Audio and Video-based Biometric Person Authentication (AVBPA'99)*, Washington, D.C., USA, 72–77.
27. **Park, J.S., Oh, Y.H., Ahn, S.C., & Lee, S.W. (2005)**. Glasses removal from facial image using recursive error compensation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(5), 805–811.
28. **Rizo-Rodríguez, D., Méndez-Vázquez, H., & García-Reyes, E. (2010)**. An Illumination Quality Measure for Face Recognition. *20th International Conference on Pattern Recognition (ICPR 2010)*, Istanbul, Turkey, 1477–1480.
29. **Roweis, S.T. & Saul, L.K. (2000)**. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*, 290(5500), 2323–2326.
30. **Tenenbaum, J.B. (1997)**. Mapping a Manifold of Perceptual Observations. *1997 Conference on Advances in Neural Information Processing*

Systems, Denver, CO, USA, 10, 682–688.

31. **Face Recognition Homepage: Vendors (s.f.)**. Retrieved from <http://www.face-rec.org/vendors/>.

32. **Viola, P. & Jones, M. (2001)**. Rapid object detection using a boosted cascade of simple features. *2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii, USA, I-511–I-518.



Heydi Méndez-Vázquez es graduada de Ingeniería Informática (con Honores) en el Instituto Superior Politécnico José Antonio Echevarría de Cuba, en el año 2005. Desde entonces se ha desempeñado como Investigador en el CENATAV. En el 2010 culminó su doctorado en Automática y Computación, en el tema de Reconocimiento de Rostros. Sus áreas de investigación incluyen la Biometría, el Reconocimiento de Rostros y el Procesamiento Digital de Imágenes.



Leonardo Chang se graduó de Ingeniero Informático (con Honores) en el Instituto Superior Politécnico José Antonio Echevarría de Cuba, en el año 2007. Desde entonces se ha desempeñado como Investigador en el CENATAV, Cuba. En el año 2010 se graduó de MsC. en el Área de Ciencias Computacionales en el INAOE, donde actualmente es estudiante de Doctorado en Ciencias Computacionales. Sus áreas de investigación son el reconocimiento de objetos en imágenes y su aceleración usando FPGAs y GPUs.



Dayron Rizo-Rodríguez se graduó de Ciencias de la Computación en la Universidad de la Habana, Cuba, en el año 2008. Desde entonces se ha desempeñado como Investigador en el Centro de Aplicaciones de Tecnologías de Avanzada (CENATAV), Cuba. Sus intereses de investigación incluyen el Reconocimiento de Rostros.



Annette Morales-González es graduada de Ingeniería Informática en el Instituto Superior Politécnico José Antonio Echevarría de Cuba, en el año 2005. Actualmente es investigadora y estudiante de doctorado en el CENATAV. Sus intereses de investigación incluyen

la segmentación y clasificación de imágenes, la anotación automática y la recuperación de imágenes por contenido.

Artículo recibido el 17/02/2011; aceptado el 21/10/2011.