

ARTÍCULO DE INVESTIGACIÓN / RESEARCH ARTICLE

Algoritmo para la obtención de parámetros antropométricos en imágenes de rostros frontales

Algorithm to obtain anthropometric parameters in frontal images of human faces

Jorge L. Camargo Díaz*

Edelmi Pérez**

Miguel Sotaquirá Gutiérrez***

Ricardo Gutiérrez De Aguas****

Universidad del Norte (Colombia)

* Jorge L. Camargo Díaz., Ingeniero Electrónico, Universidad del Norte (Barranquilla, Colombia). Forma parte de los grupos de Robótica y Sistemas Inteligentes, y Química y Biología. *jcamargo.diaz@gmail.com*.

** Ingeniero Electrónico, Universidad del Norte (Barranquilla, Colombia). *eperez@uninorte.edu.co*.

*** Ingeniero Electrónico, Ph. D. (C). Profesor de tiempo completo, Universidad del Norte (Barranquilla, Colombia). Es miembro del grupo de Telecomunicaciones y Señales actualmente Department of Bioengineering, Politecnico di Milano University. *msotaquirá@gmail.com*.

**** Licenciado en Biología y Química, Ph. D. Profesor Tiempo Completo, Universidad del Norte (Barranquilla, Colombia). Pertenece al grupo de investigación en Química y Biología. *rgutierr@uninorte.edu.co*.

Correspondencia: Km 5 vía a Puerto Colombia, tel. 0353509509, Barranquilla (Colombia).

Resumen

Este trabajo describe un algoritmo desarrollado bajo técnicas de procesamiento de imágenes para la obtención automática de los principales parámetros antropométricos del rostro en posición frontal. El algoritmo emplea un modelo *gaussiano* de la piel con el fin de detectar el rostro del individuo presente en la escena. Luego, por medio de un conjunto de ecuaciones morfológicas, se segmentan las regiones del rostro de interés (ojos, nariz y labios). Una vez detectadas estas regiones se calculan las coordenadas de los puntos de referencia faciales y, finalmente, se relacionan para obtener los parámetros antropométricos. El algoritmo es capaz de detectar automáticamente 14 de los 18 puntos de referencia definidos por Farkas, con un error promedio de ubicación menor a 3 mm. Con este conjunto de puntos se miden 23 parámetros antropométricos faciales en un tiempo promedio de 13 segundos.

Palabras clave: espacio YCbCr, dilatación, erosión, modelo *gaussiano* de la piel, antropometría facial, modelo Farkas.

Abstract

This article describes an algorithm developed under image processing techniques to automatically obtain relevant anthropometric parameters from frontal human faces. The algorithm first applies a Gaussian model of the skin to detect the individual face in the scene. Later, a set of morphological equations is applied to isolate or segment the interest regions of the face (eyes, nose and lips). Once these regions are detected, the algorithm calculates the reference point's coordinates and finally they are used to obtain the anthropometric measures of interest. The implemented algorithm is capable to automatically detect 14 of the 18 referent points defined by Farkas for frontal faces with a mean square error less than 3 mm. With this set of points 23 parameters are measured in an average computation time of 13 seconds.

Keywords: YCbCr space, Dilation, Erosion, Gaussian model of skin, Facial anthropometry, Farkas model.

Fecha de recepción: 16 de febrero de 2011
Fecha de aceptación: 18 de mayo de 2011

1. INTRODUCCIÓN

La antropometría facial es el tratado de las proporciones y medidas de la cabeza y el rostro humano [1]. Las medidas antropométricas faciales sirven para caracterizar el rostro humano y así valorar objetivamente su atractivo. El atractivo facial ayuda a entender diversas hipótesis que sostienen que esta característica puede considerarse como un certificado de salud del individuo. Además, puede ser correlacionada con la calidad seminal de los varones. También con la antropometría facial es posible estudiar los cambios morfológicos de los tejidos blandos del complejo facial para poblaciones específicas y, finalmente, lograr determinar de forma general el grado de evolución de la región facial y de la cabeza de los humanos. De allí radica su importancia biológica [2].

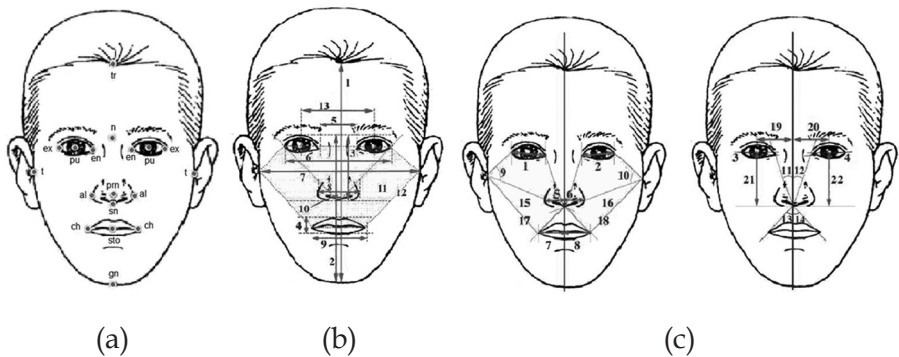


Figura 1. Parámetros antropométricos considerados: (a) puntos de referencia; (b) parámetros antropométricos del modelo del rostro frontal; (c) medidas para evaluación de la simetría facial [2]

El modelo antropométrico considerado para el rostro frontal humano consta de 32 parámetros que se calculan a partir de un conjunto de 18 puntos de referencia [1]. Según Farkas, las fronteras de los objetos del rostro definen puntos de referencia que son útiles para la caracterización del rostro. Estos puntos son: *al.*- alar, *ch.*- chelion, *en.*- endocantion, *ex.*- exocantion, *gn.*- gnation, *n.*- nasion, *prn.*- pronasal, *pu.*- pupila, *sto.*- stomion, *sn.*- subnasal, *t.*- tragon, *tr.*- trichion. Ver figura 1 (a). Teniendo en cuenta los puntos de referencia mencionados se fijan un conjunto de medidas o parámetros, que son: 1.- altura fisiognómica cara, 2.- altura media nariz, 3.- altura nariz, 4.- altura mucosa bucal, 5.- anchura interocular interna, 6.- anchura interocular

externa, 7.- anchura bi-auricular, 8.- anchura nariz, 9.- anchura boca, 10.- ángulo naso-ocular externo, 11.- área de los sentidos (cara), 12.- perímetro de los sentidos. Ver figura 1 (b). En la figura 1 (c) se encuentran: 1-2.- anchura palpebral (d-i), 3-4.- inclinación ojos (d-i), 5-6.- distancia pronasal-alagenium (d-i), 7-8.- distancia estomion-chelion (d-i), 9-10.- ángulo óculo-oto-nasal (d-i), 11-12.- ángulo separación ojos (d-i), 13-14.- ángulo naso-bucal (d-i), 15-16.- área de los sentidos (d-i), 19-20.- distancia pupila-eje cara (d-i), 21-22.- altura pupila-subnasal (d-i). La abreviatura (d-i) indica que el parámetro se mide en el lado derecho e izquierdo del rostro, teniendo en cuenta la línea imaginaria azul que se muestra en la figura 1(c). Este último conjunto de parámetros permite determinar el nivel de simetría de la cara. Esta característica es calculada con el fin de evaluar el atractivo facial del individuo, puesto que las personas en general la relacionan con la belleza del rostro [2].

De acuerdo con las técnicas de procesamiento de imágenes, el estudio antropométrico del rostro se ha abordado como un problema de detección y localización de los puntos característicos del rostro en cuanto producto del proceso de segmentación para posteriormente obtener las medidas de interés [3], [4], [5]. El algoritmo propuesto está compuesto de tres bloques generales con las siguientes funciones: 1) detección del rostro del individuo en la escena por medio del empleo de un modelo *gaussiano* del color de la piel, 2) detección de los objetos del rostro (ojos, nariz y labios) por medio de un conjunto de ecuaciones morfológicas que utilizan información del color en el espacio YCbCr, y 3) ubicación y apareamiento de los puntos de interés basándose en la información de las fronteras y elementos estructurantes característicos de cada objeto para obtener las medidas faciales. Teniendo en cuenta la complejidad de la escena, que puede presentar variaciones en el color de piel debido a la raza, condiciones de iluminación, es necesario emplear un modelo del color de la piel robusto. El entrenamiento de un modelo *gaussiano* dentro del espacio de color YCbCr ha sido ampliamente utilizado para la tarea de detectar la piel en escenas complejas [5], [6], [7], [8].

2. MODELO DE SEGMENTACIÓN DE LA PIEL

A. Espacio de color YCbCr

El espacio de color YCbCr es una representación codificada de la señal RGB. Este espacio de color está construido como una suma ponderada

de los valores RGB y por dos diferencias de color, Cr y Cb, que están formadas por la sustracción de la luminancia de las componentes roja y azul del espacio RGB. La simplicidad de la transformación y lo explícito de la separación de los componentes de la *luminancia* (brillo) con la *chrominancia* (color) hacen que este espacio sea atractivo para el modelamiento del color de la piel [9]. Estudios sobre este espacio de color han encontrado que aunque los colores de la piel de diferentes personas parezcan variar sobre un rango muy amplio, estos difieren en menor grado de la *chrominancia* que en el brillo [10], [11]. Entonces, como la luminancia puede ser removida de la representación del color en el espacio YCbCr, se prefiere este espacio para la detección del color de la piel. Además, este espacio de color también es apropiado para la detección de las regiones del rostro. Un análisis de los componentes de la *chrominancia* indicó que valores grandes de Cb y valores pequeños de Cr se encuentran alrededor de los ojos y que valores grandes de Cr son encontrados en el área de los labios. Esta caracterización permite establecer un conjunto de ecuaciones morfológicas para la segmentación de cada región de interés [10], [11].

B. Modelo gaussiano del color de la piel

Este modelo se basa en una región específica de valores de color de la piel en el sub-espacio CbCr. Asumiendo que la densidad de la piel está modelada por una *gaussiana* sencilla, la similitud de un vector de crominancia de entrada $X=[X_{Cb}, X_{Cr}]$ perteneciente a la piel está dada por la ecuación 1.

$$P(X) = \frac{1}{2\pi\sqrt{\Lambda}} \exp\left(-\frac{1}{2} \lambda^2\right) \quad (1)$$

$$\lambda^2 = (X - \mu)^T \Lambda^{-1} (X - \mu) \quad (2)$$

Donde λ es la distancia de Mahalanobis definida como indica la ecuación 2. Los vectores, $\mu = [\mu_{Cb}, \mu_{Cr}]$ vector de la media, y Λ , la matriz de la covarianza, son estimados a partir de un conjunto de entrenamiento [6], [11], [12], [13], [14]. En la figura 2 se muestra el diagrama de bloques del algoritmo para la detección de la región facial y a modo de ejemplo se muestran las imágenes que se obtienen en cada etapa del mismo.

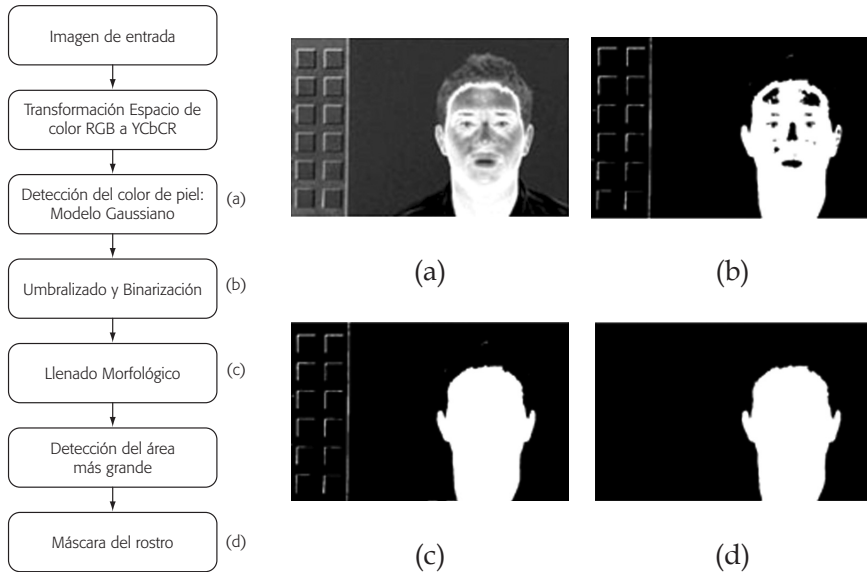


Figura 2. Diagrama de bloques del algoritmo para la detección del rostro presente en la imagen.

3. DETECCIÓN DE OBJETOS DEL ROSTRO

A. Detección de región ocular

El algoritmo de detección de los ojos utiliza las operaciones de dilatación y erosión en escala de gris con elementos estructurantes circulares. Estas operaciones son aplicadas independientemente sobre la componente de la luminancia y de la *chrominancia* para construir el mapa de los ojos. La construcción del mapa de los ojos en la componente de la luminancia se basa en la característica de que la región de los ojos contiene tanto píxeles oscuros como claros. Entonces, se combinan operadores morfológicos para enfatizar dichos píxeles [15], [16]. El mapa de los ojos en la componente de la *chrominancia* es construido a partir de la componente Cb, la inversa de Cr y la razón Cb/Cr. Los dos mapas resultantes para los ojos son combinados por medio de una operación de multiplicación lógica AND [10]. Las consideraciones anteriores pueden ser compactadas por medio de las ecuaciones 3, 4 y 5 [4].

$$\text{Mapa Ojo } C = \frac{1}{3} \left\{ C_b^2 + (C_r)^2 + \frac{C_b}{C_r} \right\} \quad (3)$$

$$\text{Mapa Ojo } L = \frac{Y(x, y) \oplus g\delta(x, y)}{Y(x, y) \ominus g\delta(x, y) + 1} \quad (4)$$

$$\text{Mapa Ojos} = (\text{MapaOjo}C) \text{ AND } (\text{MapaOjo}L) \quad (5)$$

Donde $(\tilde{C}_r)^2$, C_b^2 y C_b/C_r están normalizados dentro del rango $[0, 255]$, \tilde{C}_r es el negativo de C_r , $(255 - C_r)$. Y es la componente de la luminancia y $g\delta$ es el elemento estructurante circular [4]. Luego de la aplicación de estas ecuaciones es necesario emplear operaciones adicionales para lograr aislar del resto de la imagen los elementos de los ojos y así poder establecer la región ocular. Las operaciones que se realizan son operaciones de dilatación y binarización.

B. Detección de la región labial

Para lograr detectar esta región se parte del siguiente hecho: la región de los labios contiene valores altos de *chrominancia* roja y valores bajos de *chrominancia* azul en comparación con las demás regiones faciales dentro del espacio YCbCr. Además, los labios tienen una respuesta relativamente baja a la característica C_r/C_b pero una respuesta alta a C_r^2 [17]. Por esto, la diferencia entre C_r^2 y C_r/C_b se utiliza para enfatizar la región de los labios. Las ecuaciones 6 y 7 recopilan las consideraciones anteriores y se muestran a continuación [4]:

$$\text{MapaLabios} = C_r^2 \left[C_r^2 - \eta \left(\frac{C_r}{C_b} \right) \right]^2 \quad (6)$$

$$\eta = 0.95 \left(\frac{\frac{1}{n} \sum_{(x,y) \in FG} C_r^2(x,y)}{\frac{1}{n} \sum_{(x,y) \in FG} C_r(x,y) / C_b(x,y)} \right) \quad (7)$$

Donde C_r^2 y C_b están normalizados al rango $[0, 255]$ y n es el número de píxeles dentro de la región facial (FG). El parámetro η se estima como la relación del promedio C_r^2 al promedio C_r/C_b [4], [18], [19], [20], [21]. Luego de la aplicación de estas ecuaciones se realizan operaciones de dilatación y binarización para segmentación definitiva de la región labial y la creación del mapa correspondiente.

C. Detección de la región nasal

Una vez obtenidos los mapas de los ojos y de los labios, se utiliza la información de sus coordenadas para determinar la ubicación de la región nasal, dado que esta siempre se ubica dentro de las dos regiones mencionadas. Concretamente, primero se calcula el punto central de cada cuadro de la máscara que inscribe a cada ojo y con estas coordenadas se establecen los lados verticales del recuadro que inscribe a la nariz. Después, se toma el punto inferior derecho del cuadro inferior de la máscara de los ojos como la coordenada del lado superior del recuadro de la nariz. Por último, se utiliza la coordenada de la esquina superior derecha del mapa de los labios para determinar el lado inferior del recuadro de la nariz. El lado derecho mencionado corresponde a mitad derecha del rostro del individuo presente en la imagen.

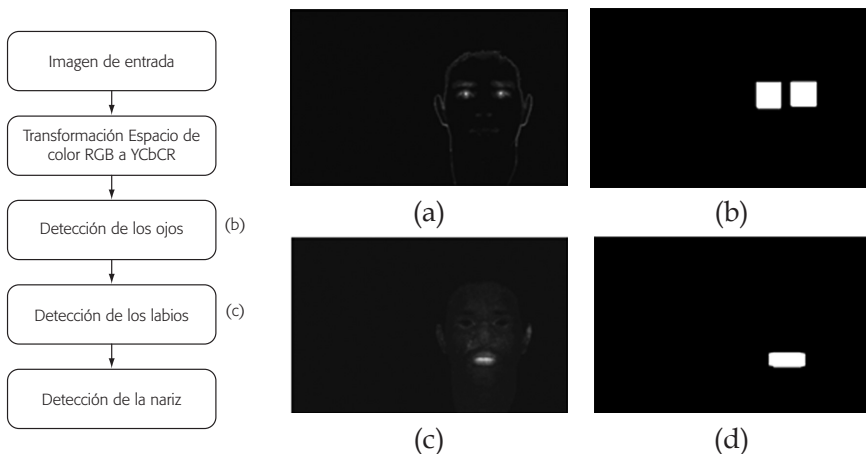


Figura 3. Diagrama de bloques del algoritmo de detección de las regiones faciales.

En la figura 3 se puede observar el diagrama de bloques del algoritmo de detección de las regiones faciales y las imágenes que resultan luego de la aplicación de las ecuaciones para el realce de los objetos faciales. Como se indicó, es necesario realizar operaciones adicionales para lograr aislar los objetos del rostro del resto de la imagen y poder construir una máscara binaria de la región correspondiente.

4. DETECCIÓN DE PUNTOS CARACTERÍSTICOS

A. Puntos de la región ocular

Como se puede ver en la figura 1 (a), para el ojo están definidos tres puntos, (endocantion, exocantion y pupila). Hasta el momento el algoritmo ha calculado una máscara que asegura que dentro de ella se encuentran los ojos. Esto quiere decir que los ojos no tienen delineadas sus fronteras en una representación binaria. Por lo tanto, se realiza el siguiente procedimiento para lograr definir una estructura correspondiente a los ojos del individuo similar a la que aparece en la figura 4 (e). Se emplea una multiplicación sobre el plano de crominancia C_b , dado que este resalta la región ocular, por una constante igual a 2.4. El valor indicado se estableció por medio de pruebas en las que se observó que los valores de C_b , correspondientes al fondo de la imagen y a los ojos consiguen ser más blancos (se enfatizan); mientras que los valores correspondientes a la piel del rostro se mantienen iguales (se atenúan). Posteriormente se realiza un umbralizado adaptativo para binarizar la región.

El procedimiento mencionado binariza toda la región ocular, por lo que en la imagen segmentada pueden estar presentes las cejas del individuo. Para eliminarlas, se calcula el centro de masa de cada región de la máscara del ojo y luego se eligen a los elementos binarios más cercanos al centro de masa respectivo. Finalmente, se calculan los puntos extremos de los elementos binarios hallados y se eligen los puntos derecho inferior e izquierdo inferior como los puntos endocantion y exocantion para cada uno de los ojos.

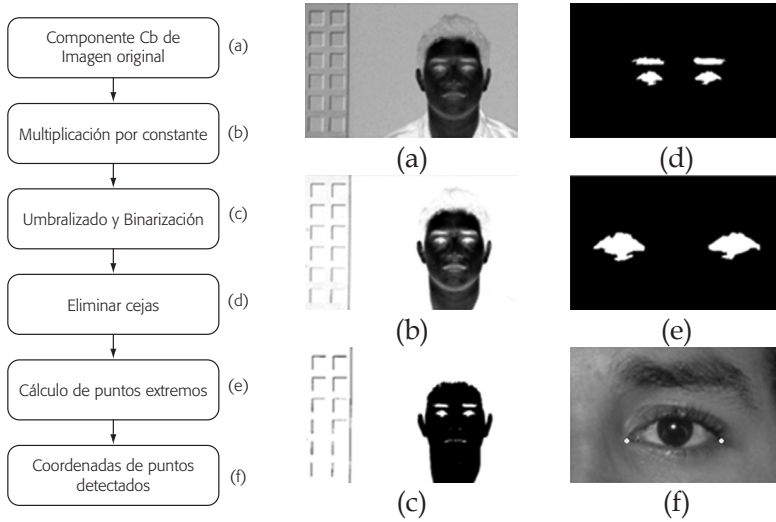


Figura 4. Diagrama de bloques del algoritmo para la detección del rostro presente en la imagen.

Para la localización del punto pupila primero se ubica el iris del ojo, que por su forma circular se halla por medio de un algoritmo detector de círculos sobre toda la imagen en escala de grises. En este caso se aplica la transformada circular de Hough [22], [23], [24], [25]. Como la pupila se encuentra en el centro del iris, sus coordenadas pueden ser halladas calculando el centro del círculo detectado [26], [27], [28], [29]. El iris se establece realizando una búsqueda del centro del círculo más cercano al centro de masa de los objetos hallados para establecer los cantones [figura 4 (e)]. Este punto se elige como el punto pupila.

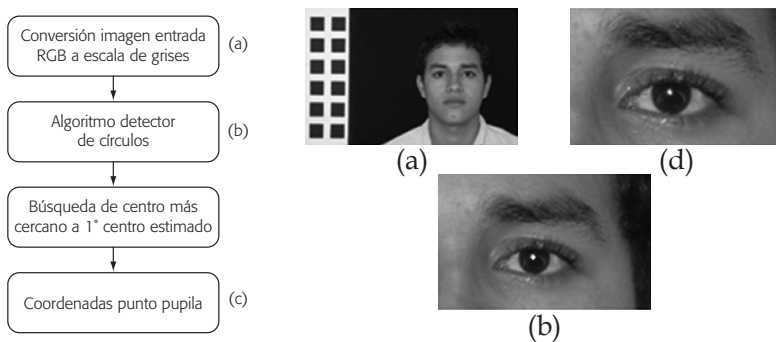


Figura 5. Diagrama de bloques del algoritmo para la detección del rostro presente en la imagen.

B. Puntos de la región bucal

En esta región primero se ubican los puntos extremos de la mucosa labial definidos como chelion, ver figura 1 (a). Para esto se parte de la observación de que los labios en su unión forman un patrón más oscuro que el resto de la región labial. Con base en esto se realiza una búsqueda de mínimos para detectar este patrón y luego se realiza una búsqueda de los puntos extremos horizontales del patrón encontrado [17], [18], [30]; ver figuras 6 (a) y (b). Para la ubicación del punto medio de los labios (stomion), se calcula la proyección vertical en esta región. Esta proyección permite estimar por medio de una recta horizontal la unión de los labios. Después, los puntos extremos hallados previamente, son proyectados sobre la recta y se calcula el punto medio. En la figura 6 (c) se nota que las coordenadas de este punto se eligen como el Stomion.

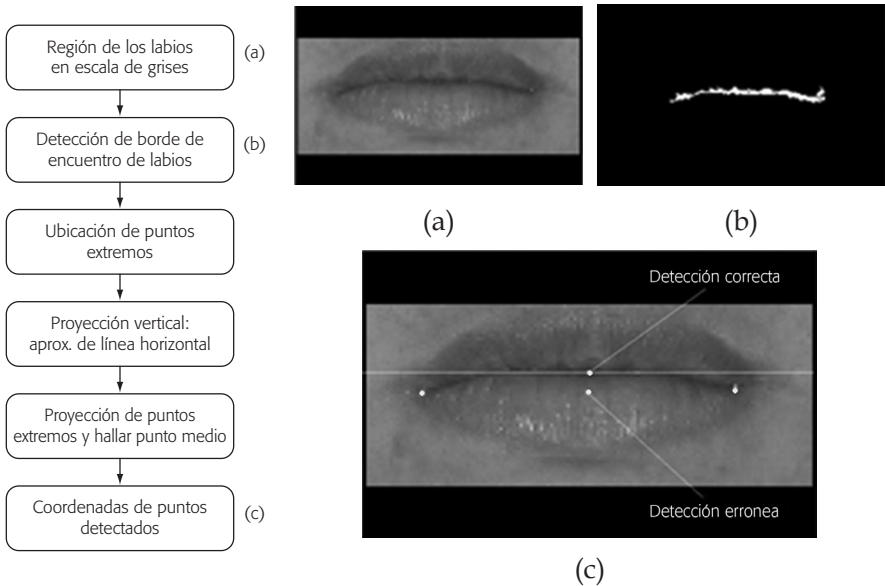


Figura 6. Diagrama de bloques del algoritmo para la detección del rostro presente en la imagen.

C. Puntos de la región nasal

Para el caso de la nariz se realiza la extracción de cuatro puntos: alar (d), alar (i), subnasal y pronasal.

Para los puntos alares primero se aplica un filtro *gaussiano* sobre la imagen en escala de grises con el fin de eliminar el ruido de alta frecuencia producido por la iluminación. Luego se aplica un detector de bordes de Canny para hallar las fronteras de la nariz y se buscan los puntos extremos (alares) sobre el eje horizontal, ver figura 7 (b).

En la figura 1(a), se puede apreciar la ubicación del punto subnasal. Se nota como característica en la ubicación de este punto que se localiza en el borde inferior y entre las ventanas de la nariz. Se define este problema como la búsqueda de ese borde inferior y para esto se debe elegir un punto cuya componente vertical esté dentro de las ventanas de la nariz; empleando la proyección integral, específicamente, la proyección vertical se logra detectar las ventanas de la nariz [31], [32], ver figura 7 (b). Con esta información se determina la componente horizontal del punto en cuestión. Para determinar la componente vertical se detectan los picos de la proyección horizontal (figura 8 (c)) y luego se determina su posición media.

Una vez que se detecta la componente vertical, se realiza una búsqueda sobre ese eje en la imagen binaria de los bordes detectados, figura 7 (b). El punto subnasal se ubica en el borde inferior sobre el eje vertical.

Sin embargo, es posible que el detector de bordes no entregue una frontera continua como la que se aprecia en la figura 7 (b). Por lo tanto, se establecen dos condiciones para disminuir el error en la ubicación cuando esto se presente: 1. el punto no puede estar muy alejado de la componente horizontal de la ventana de la nariz y si lo está, entonces se elige esta posición como el subnasal; 2. el punto debe estar por encima de la línea recta trazada por la coordenada del borde alar más inferior de la nariz.

Con el objeto de conseguir un algoritmo robusto, el punto pronasal es calculado teniendo en cuenta dos distribuciones de iluminación típicas que surgen en el momento de la adquisición. En la primera, la luz recorre la nariz de arriba hacia abajo. En la segunda, la nariz recibe la luz por el flash de la cámara digital.

Para este caso se elige trabajar por medio de la representación invertida de la imagen, porque los valores de los píxeles son constantes. Es decir, mientras que los valores de la porción iluminada de la nariz en la imagen

no invertida varían en un rango de 220 a 255, la imagen invertida establece para la porción iluminada valores de intensidad iguales a cero. Entonces, sobre el eje vertical definido previamente para hallar al punto subnasal, se realiza una búsqueda de los píxeles que tengan un valor igual a cero y se guardan en un vector. Adicionalmente, se establece un valor de umbral que permita discriminar entre los dos casos de iluminación considerados. Si el tamaño del vector que contiene los valores iguales a cero supera el umbral (primer caso), entonces, el punto pronasal es elegido como la posición del último píxel con valor de cero del vector. Si no se supera el umbral (segundo caso), se calcula la posición del píxel que corresponde a la mitad del vector.

Dado que el punto nasion es el punto donde se da la concavidad de la nariz entre los ojos, esta se determina mejor en imágenes del rostro de perfil. Sin embargo, se puede realizar una buena estimación para el caso de rostros en posición frontal. Para la ubicación de este punto se calcula el punto medio entre los puntos extremos superiores de los objetos binarios de los ojos detectados figura 4 (e). Del punto medio calculado solo se toma la información de su componente horizontal. La componente vertical se obtiene de la componente vertical del punto subnasal hallado previamente.

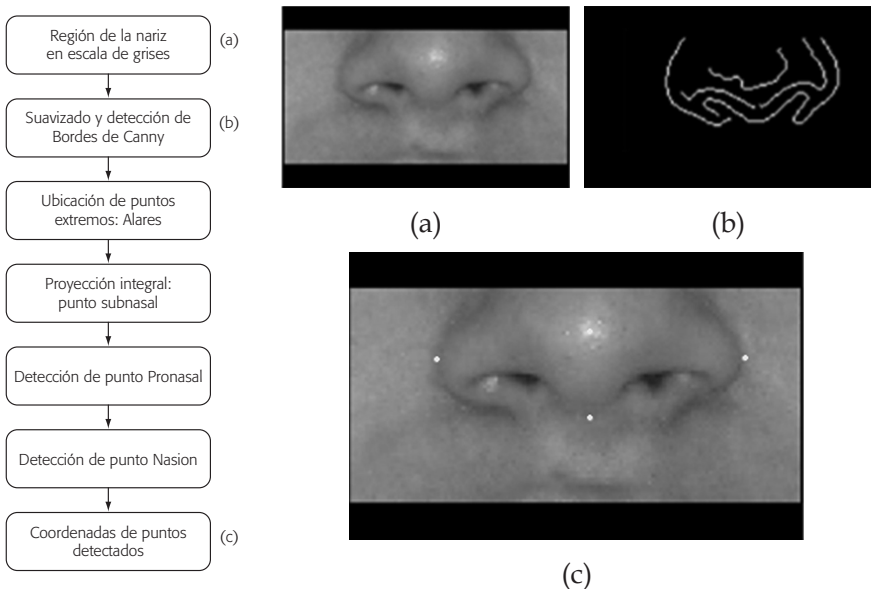


Figura 7. Diagrama de bloques para la detección de los puntos de la región nasal. Proyecciones verticales y horizontales.

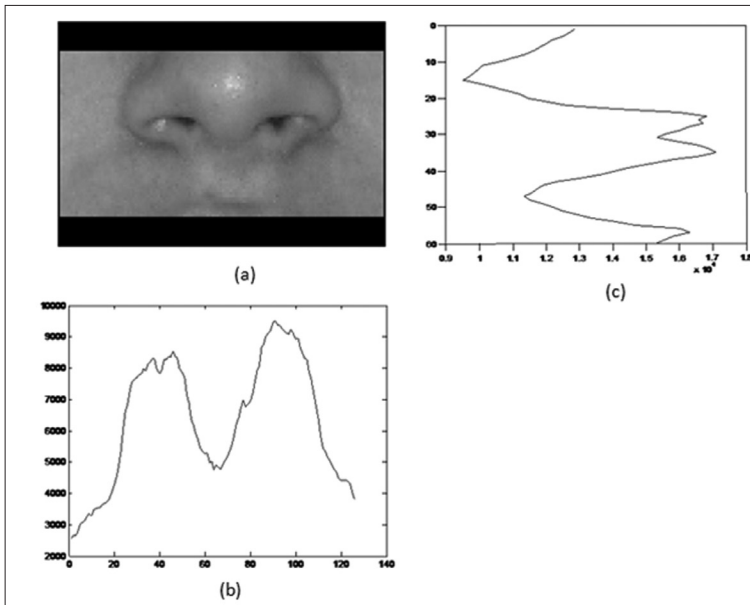


Figura 8. proyección integral: (a) Región de la nariz, (b) proyección horizontal, (c) proyección vertical.

5. PRUEBAS Y RESULTADOS

Las pruebas realizadas al algoritmo se realizaron con el fin de validar estadísticamente el error promedio de ubicación de los puntos generados por el algoritmo con respecto al punto ideal correspondiente; en específico se determinó que un desempeño óptimo generaría un error promedio de ubicación dentro del rango [0 - 3] mm. El algoritmo fue probado sobre 50 imágenes. En la muestra se tuvieron en cuenta individuos de diferente raza y sexo, y las edades estuvieron dentro del rango 18-25 años. Con respecto a las condiciones de adquisición se dispuso un fondo de escena de color negro, y en un primer plano se destaca el rostro del individuo en paralelo con una regleta graduada que facilita la conversión de las medidas obtenidas de píxeles a milímetros.

A. Desempeño de algoritmo de segmentación de la piel

Después de aplicar el modelo *gaussiano* pueden aparecer conectados elementos a la región del rostro de tamaño considerable. Para determinar el

porcentaje de acierto que realiza el algoritmo en la detección del rostro, se compara la región segmentada por el algoritmo con otra imagen que contiene la región ideal que debería ser segmentada. En la comparación se mide el grado de traslape o intercepción entre las dos imágenes y se establece una curva de desempeño del algoritmo de segmentación. Esta curva se obtiene consignando la cantidad de detecciones correctas de la región dentro de un rango de tolerancia permisible. En esta curva la tolerancia está en el eje x y a medida que aumenta indica que la detección es más estricta y que las regiones deben coincidir en mayor grado [7].

La figura 9 muestra la curva de desempeño obtenida. Se puede observar que para un porcentaje de tolerancia entre el 85 y 90% se obtuvo un porcentaje de clasificaciones correctas entre 75 y 80%. El porcentaje de tolerancia obtenido indica que la cantidad de imágenes utilizadas para el entrenamiento del modelo *gaussiano* es adecuada y que permite obtener un grado de precisión satisfactorio en la segmentación de la piel.

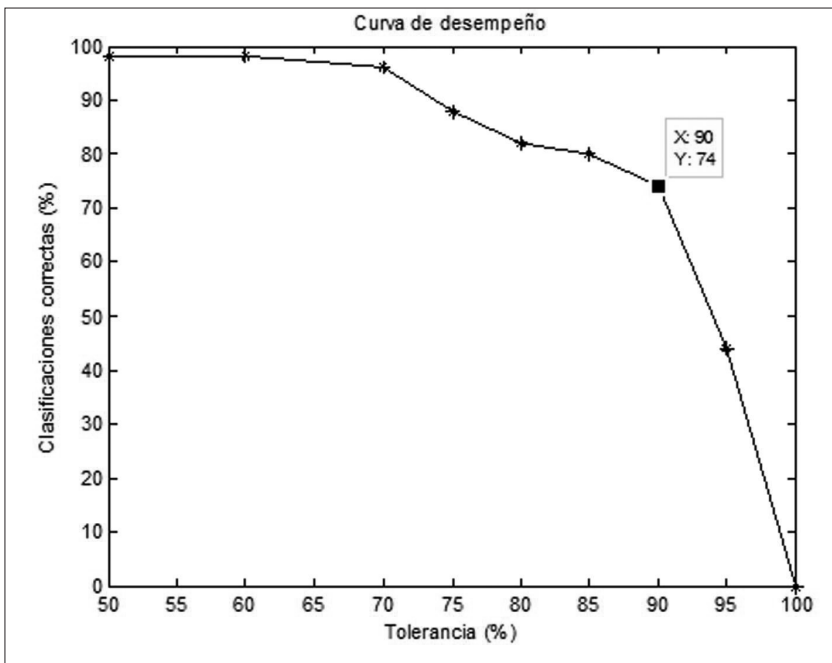


Figura 9. Curva de desempeño del algoritmo de detección de la piel del rostro

Los errores generados en esta etapa del algoritmo se deben a la información del color empleado por el modelo. Por medio del análisis de las regiones que fueron clasificadas erróneamente como piel se pudo notar que estas correspondían a regiones de la ropa y/o accesorios que presentaban una distribución de valores de color similar a la de la piel. Específicamente, los colores que el algoritmo confunde con la piel es el color amarillo en su tonalidad pastel (este caso ocurrió en las tres ocasiones en las que el individuo viste ropa con el color mencionado) y la tonalidad castaño claro presente en algunas regiones del cabello (este caso se repitió en todos los diez casos en los que el individuo tiene este color en el cabello).

B. Porcentaje de detección correcta de las regiones del rostro

Se considera como una detección correcta cuando el algoritmo es capaz de generar la máscara completa para cada uno de los objetos del rostro en la misma imagen. A cada detección correcta se le asignó un valor igual a uno, y un valor igual a cero para las detecciones erróneas. Los resultados de esta prueba se pueden observar en la tabla 1.

Tabla 1. Porcentaje de detecciones correcta de las regiones del rostro

N° Imágenes	Ojos	Nariz	Labios
50	90,77	84,62	92,31

De la tabla 1 se puede asegurar que el algoritmo alcanzó una alta tasa de detección para las regiones ocular y labial: 90.77% y 92.32%, respectivamente. Los casos que influyeron en la detección errónea de las regiones faciales se deben a elementos detectados erróneamente como piel en la etapa anterior.

En detalle, las detecciones erróneas para el caso de la máscara de los ojos se deben a que previamente el algoritmo de segmentación de la piel había detectado porciones de ropa del individuo de color blanca (la cantidad de individuos con ropa de este color es igual a 13 y la detección incorrecta corresponde solo a un caso particular). También se dio el caso para un sujeto de la base de datos en el que los aretes de color metálico reflejan la luz hacia la cámara y exhiben un color blanco (nuevamente se aclara que

la cantidad de individuos con accesorios parecidos a los señalados es igual a 6 y en un solo caso se notó este error).

De manera similar sucede para los casos de detección errónea de la región de los labios. Se observó que el algoritmo identificaba erróneamente al pabellón externo de la oreja como la mucosa labial cuando la cabeza del individuo está un poco inclinada hacia la cámara y recibe más iluminación en la región señalada que en el resto del área facial, este error se presentó en dos muestras. Además, en uno de los dos casos mencionados se notó que el individuo no delimitaba bien la frontera de color característica entre la mucosa labial y el resto del rostro.

Según los datos en la tabla 1, el porcentaje de detecciones correctas de la región nasal es igual a 84.62%. Sin embargo, es bajo si se compara con los resultados obtenidos para las regiones de los ojos y labios. Pero lo cierto es que este porcentaje pone de manifiesto que la atinada detección de la región nasal depende de la estimación correcta de las regiones de ojos y de la boca en la misma imagen.

C. Error promedio de ubicación automática de los puntos de referencia

En esta prueba se mide la distancia euclidiana de separación para cada uno de los puntos generados automáticamente con respecto a su posición ideal en cada una de las 50 imágenes de la base de datos empleada. Específicamente, a fin de conocer el error promedio, se realizó una consolidación de todos los datos de la distancia para cada punto; posteriormente se calcula su valor promedio. Asimismo se obtuvo información sobre la varianza de los datos y su desviación estándar.

Los resultados obtenidos se pueden ver en la tabla 2. Con base en los datos se puede afirmar que el valor promedio del error cometido en la ubicación automática de los puntos de referencia, se encuentra dentro de un rango aceptable (menor o igual a 3 milímetros) y que los datos cuentan con una desviación o dispersión respecto de la media aritmética aceptable según lo indica la columna de la desviación estándar. También se pueden notar en los datos las variaciones en las condiciones de iluminación sobre el rostro de los individuos durante las sesiones de adquisición. Se puede decir que la mitad derecha del rostro del individuo recibió en promedio una mejor

iluminación que la mitad izquierda. Sin embargo, esta desviación en las condiciones de iluminación no es drástica en la medida en que no ubica al error por fuera del rango de error establecido [33].

Tabla 2. Error promedio en la detección automática de puntos de referencia

Región facial	Puntos de referencia	Error promedio [mm]	Varianza	Desviación estándar [mm]
Región ocular	Exocantion (d)	2,05	1,34	1,16
	Pupila (d)	0,88	1,43	1,20
	Endocantion (d)	2,31	3,06	1,75
	Endocantion (i)	2,59	2,66	1,63
	Pupila (i)	0,73	0,22	0,46
	Exocantion (i)	2,04	1,41	1,19
Región nasal	Alar (d)	2,26	3,07	1,75
	Pronasal	1,58	2,63	1,62
	Alar (i)	1,96	8,62	2,94
	Nasion	2,7	6,25	2,50
	Subnasal	1,32	2,48	1,58
Región bucal	Chelion (d)	1,45	2,08	1,44
	Stomion	1,6	1,07	1,03
	Chelion (i)	1,45	8,09	2,84

Un análisis detallado de los valores consignados en la tabla 2 permite comparar los resultados obtenidos para los puntos de la región ocular: exocantion, endocantion y pupila. El error particular que generó la inconsistencia entre los valores para el caso del punto exocantion se debió al color de las pestañas, aunque los resultados en promedio indican que esta diferencia no afecta de manera significativa. El peor de los casos obtenidos en la ubicación de los puntos exocantion derecho e izquierdo fueron iguales a 6.1 mm y 6.25 mm, respectivamente.

Para el caso de los puntos endocantion se puede ver en la tabla 2 que el algoritmo ubica con menor error promedio al punto endocantion derecho (2,31 mm) que el endocantion izquierdo (2,59 mm). Para el punto de la

pupila derecha se obtiene un valor de ubicación promedio de 0.88 mm y para la pupila izquierda de 0.73 mm, valores menores que un milímetro. Nuevamente se destaca que la diferencia entre estos valores se debe a la iluminación. El peor caso para los puntos pupila es igual a 7.7 mm para pupila derecha e igual a 2.8 mm para el punto pupila izquierda. Además de la variación de la iluminación sobre el rostro, otra fuente de error en la ubicación de este punto se da cuando el círculo más cercano al centro de masa del objeto binario de los ojos, hallado por la transformada circular de Hough, no corresponde a la circunferencia del iris.

El desplazamiento producido por el ruido que se filtra en la región nasal se considera como una primera fuente de error para esta región. Esta fuente de error junto con la iluminación hace que para algunas imágenes la frontera de la nariz esté desplazada. Además, se dan casos en los que el ruido no es eliminado completamente y se generen errores de ubicación particulares con valores iguales a 7.4 mm y de 1.5 cm. La diferencia en los errores de detección de estos puntos refleja la presencia de ruido remanente.

El error obtenido para el punto subnasal se debe a que el filtro empleado para suavizar la región nasal no elimina completamente el ruido y este es elegido por el algoritmo como si fuese el borde inferior de la nariz. Otra causa de error se debe al borde discontinuo arrojado por el detector de bordes implementado, que no arrojó un borde continuo sobre la coordenada vertical de búsqueda. No obstante, el resultado obtenido para este punto de 1.32 mm indica que se obtiene una ubicación buena según estimación a este punto de referencia.

El error cometido en la ubicación del punto pronasal se debe principalmente a que no se encuentra la porción iluminada de la nariz sobre la coordenada vertical que entrega el punto subnasal. Se identificó que esto se debe a una ligera inclinación del rostro del individuo hacia un lado. Pero este caso solo se presentó sobre un único individuo de toda la base de datos de imágenes y por esto se obtuvo un error bajo en la ubicación promedio para este punto (1.58 mm). De la misma manera, la inclinación de la cabeza hace que la coordenada vertical del punto subnasal utilizada para calcular el nasion corresponda a una detección errónea de este último. También se debe considerar que este punto es determinado de mejor manera en imágenes del rostro en posición lateral; sin embargo, se logra un error promedio

bajo (2.7 mm) a pesar de que en este caso solo se emplearon imágenes de rostro en posición frontal.

Además de los puntos de referencia ubicados en los labios, se identificó que los errores son generados por el principio de ejecución de este algoritmo, que realiza una búsqueda de los valores mínimos cuando determina el patrón oscuro característico de la unión de los labios. Con esto en mente, es deducible que una fuente de error puede ocurrir ante la presencia de vello facial con un color más oscuro que el patrón mencionado. Para este caso el valor de error alcanza un valor de 2.35 cm. Este error no es frecuente (únicamente se presenta para un solo individuo de dos que presentan vello facial) y, por esto, se alcanzan valores de error promedio iguales a 1.45 mm para el punto chelion derecho y de 1.66 para el punto chelion izquierdo. Para el caso del punto stomion se nota que el método utilizado para ubicarlo realiza una buena aproximación. Además, el procedimiento que ejecuta el algoritmo para ubicar este punto permite una detección correcta, aun cuando esta haya sido errónea para los puntos extremos de los labios; esto se debe a que los últimos son proyectados a una línea recta horizontal de unión, estimada a partir de la proyección vertical. Sin embargo, cuando uno de los puntos extremos es detectado en una posición muy alejada de su posición ideal, el error en la ubicación del punto stomion genera un error grande e igual a 5.3 mm.

Como parte de los resultados del algoritmo se muestra la figura 9. En esta se pueden observar las regiones del rostro detectadas figura 10 (a); además, en la figura 10 (b) se pueden ver la totalidad de los puntos de referencia detectados automáticamente.

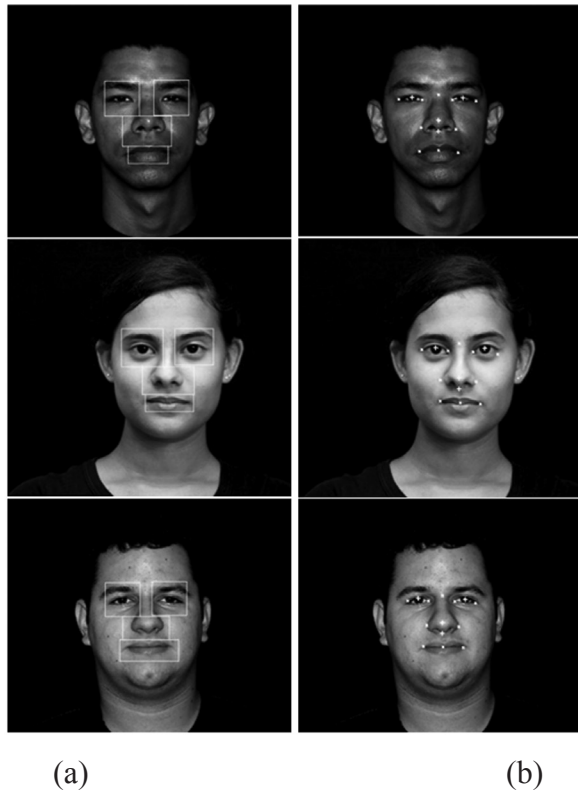


Figura 10. Resultados del proceso de segmentación. (a) Objetos del rostro detectados a partir de las máscaras de las regiones detectadas. (b) Puntos de referencia obtenidos de manera automática por el algoritmo desarrollado.

6. CONCLUSIONES

Con base en los resultados experimentales obtenidos, se puede asegurar que el algoritmo implementado calcula automáticamente 14 de los 18 puntos de referencia del modelo en un tiempo promedio de 13 segundos con un error de ubicación con respecto a la posición ideal en promedio menor a 3 mm.

Además, se destaca que en el algoritmo desarrollado se pudieron ubicar de forma automática los 4 puntos de referencia pertenecientes a la región nasal (alar (d), alar (i), pronasal y subnasal), el punto nasion y el stomion de la boca. La obtención de los seis puntos mencionados, junto con la detección

de la región nasal, representa una contribución significativa al proyecto nacional conocido [7], en el que solo consideran 8 puntos de referencia y las medidas antropométricas correspondientes a los objetos de los ojos y de los labios.

Se detectó que la iluminación de la escena es el factor que determina una correcta ubicación de los puntos de referencia y, en consecuencia, un cálculo acertado del conjunto de medidas antropométricas respectivas. La deficiencia en este parámetro de adquisición permite explicar, por ejemplo, por qué varían los valores del error promedio en la detección automática de los puntos extremos de la nariz y de los ojos. En concreto, se logró desarrollar un algoritmo de detección del rostro con una tasa aceptable de detecciones correctas (entre 75 y 80%). Lo anterior se logró incluyendo un total de 198400 píxeles de entrenamiento pertenecientes a diferentes tonos del color de la piel al modelo *gaussiano*. El efecto consiguiente que con esto se logró permitió generar un algoritmo capaz de procesar imágenes de individuos de diferente raza, y que le permite usarse como herramienta para el estudio y caracterización antropométrica de cualquier grupo étnico.

Por otro lado, se destaca la realización de una interfaz gráfica para ejecutar el algoritmo. Esta herramienta adicional permite al usuario opciones de interactividad con la cual es posible: observar una lista de resumen con las medidas de los parámetros antropométricos calculados, modificar manualmente los puntos hallados y ubicar los puntos adicionales sobre el complejo facial (ej. trignon, gnation, trichion (d), trichion (i), labial superior y labial inferior). Estas opciones adicionales incluidas en la interfaz, posibilita la obtención de la totalidad de los puntos de referencia y la totalidad de mediciones antropométricas definidas para el rostro en posición frontal.

Finalmente, se destaca que este proyecto de investigación representa un avance significativo al problema general de la extracción automática de todo el conjunto de parámetros antropométricos del rostro en posición frontal y un aporte más hacia la consolidación de la relación interdisciplinar entre la Ingeniería Electrónica con la Biología.

Referencias

- [1] L. Farkas, *Anthropometry of the Head and Face*, 2da. ed., New York: Raven Press, 1994, p. 395.
- [2] R. Gutiérrez, "Estudio de la relación entre características antropométricas, atractivo facial y calidad seminal", Tesis doctoral, Dpto. Biología Funcional y Antropología Física, Universitat de València, Valencia, España, 2004.
- [3] V. Vezhnevets, S. Soldatov, A. Degtiareva, and I. K. Park, "Automation extraction of frontal facial features", *Proceeding Sixth Asian Conference on Computer Vision (ACCV04)*, Jeju, Korea, vol. 2, pp. 1020-1025, 2004.
- [4] I. Mejía, "Extracción automática de características faciales para el estudio antropométrico en niños entre 5 y 10 años de la ciudad de Manizales", Tesis de pregrado, Dpto. Electricidad, Electrónica y Computación, Universidad Nacional de Colombia sede Manizales, Manizales, Caldas, 2004.
- [5] M. H. Yang, and N. Ahuja, "Detecting human faces in color images", 1998 International Conference on Image Processing ICIP 98, Chicago, Illinois, vol. 1, pp. 127-130, 1998.
- [6] J. Y. Lee, and S. I. Yoo, "An Elliptical Boundary Model for Skin Color Detection", *International Conference on Imaging Science, Systems, and Technology CISST'02*, Las Vegas, Nevada, pp. 24-27, 2002.
- [7] J. Min, M. Powell, and K. Bowyer, "Automatic Description of Complex Buildings with Multiple Images". *Fifth IEEE Workshop on Applications of Computer Vision WACV 2000*, Palm Springs, California, pp. 155-162, 2000.
- [8] A. Sayeed, M. Sohail, and P. Bhattacharya, "Detection of Facial Feature Points Using Anthropometric Face Model", *Multimedia system and applications*, vol. 1, pp. 189-200, 2008.
- [9] V. Vezhnevets, V. Sazonov, and A. Andreeva. "A survey on pixel-based skin color detection techniques", *Proc. Graphicon*, Moscow, Russia, 2003.
- [10] J. Ren, and L. Guo, "Applying mixed Gaussian skin models to the automatic face detection", *Proceedings 14th IEEE Signal Processing Society Workshop*, Sedona, Arizona , pp. 599-608, 2004.
- [11] J. Zhang, K. Zhou, L. Velho, B. Guo, and H. Y. Shum, "Synthesis of progressively-variant textures on arbitrary surfaces", *ACM Transactions on Graphics*, vol. 22, pp. 295-302, 2003.
- [12] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color and edge information", *IEEE ISSPA*, vol. 1, pp. 525-528, 2003.
- [13] M. J. Jones, and J. M. Rehg. "Statistical color models with application to skin detection", *International Journal of Computer Vision*, vol. 46, pp. 274-280, 1999.
- [14] J. B. Martinkauppi, and M. Pietikainen, "Facial Skin Color Modeling", in *Handbook of face recognition*, S. Z. Lee, and A. K. Jain, Ed., New York: Springer, 2005, pp. 113-135.

- [15] P. Dwayne, *Image Processing in C: Analyzing and Enhancing Digital Images*, 2nd. Ed. Lawrence, Kansas: R & D Publications, 2000.
- [16] G. X. Ritter, and J. N. Wilson, *Handbook of Computer Vision Algorithms in Image Algebra*, Toronto: CRC Press, 2004.
- [17] H. Rein-Lien, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 696-706, 2002.
- [18] W. N. Lie, and H. C. Hsieh, "Lips Detection by Morphological Image Processing", *Fourth International Conference on Signal Processing Proceedings ICSP '98*, Beijing, China, vol. 2, 1084-1087, 1998.
- [19] A. Caplier, "Lip Detection and Tracking", *11th International Conference on Image Analysis and Processing ICIAP'01*, Palermo, Italy, pp. 8, 2001.
- [20] N. Eveno, A. Caplier, and P. Y. Coulon, "A new color transformation for lips segmentation", *Fourth Workshop on Multimedia Signal Processing*, pp. 3-8, 2001.
- [21] P. Gacon, P. Y. Coulon, and G. Bailly, "Statistical active model for mouth components segmentation", *Computer and Information Science*, vol. 2, pp. 1021-1024, 2005.
- [22] R. Bruneli, and T. Poggio, "Face recognition: Features versus templates", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1042- 1062, 1993.
- [23] S. Sirohey, A. Rosenfeld, and Z. Duric, "A method of detecting and tracking irises and eyelids in video", *Pattern Recognition*, vol. 35, pp. 1389-1401, 2002.
- [24] X. Cao, and F. Deravi, "An Efficient Method for Multiple-Circle Detection", *Proceeding Third International Computer Vision*, Osaka, Japan, pp. 744-747, 1990.
- [25] B. Lamiroy, O. Gaucher, and L. Fritz, "Robust Circle Detection", *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Curitiba, Paraná, vol. 1, pp. 526-530, 2007.
- [26] R. K. K. Yip, D. N. K. Leung, and S. O. Harrold, "Line segment patterns hough transform for circles detection using a 2-dimensional array", *Proceeding of IECON '93*, Maui, Hawaii, vol. 3, pp. 1361-1365, 1993.
- [27] Y. Wei, "A Fast Finding and Fitting Algorithm to Detect Circles", *Geoscience and Remote Sensing Symposium Proceeding*, Seattle, Washington, vol. 2, pp. 1137-1139, 1998.
- [28] P. F. Zhang, and Q. M. Li, "Research on iris image preprocessing algorithm", *Proceedings of 2005 International Conference on Machine Learning and Cybernetics*, Guangzhou, China, vol. 8, pp. 5220-5224, 2005.
- [29] W. Yuan, Z. Lin, and L. Xu, "A Novel and Fast Iris Location Algorithm Based on the Structure of Human Eyes", *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Dalian, China, pp. 10388-10392, 2006.

- [30] J. A. Dargham, and A. Chekima, "Lips Detection in the Normalised RGB Colour Scheme", *IEEE International Conference on Information & Communication Technologies: From Theory to Applications ICTTA '06*, Damascus, Syria, pp. 1546-1551, 2006.
- [31] L. Moran, and R. Pinto, "Descripción de dos modelos estadísticos aplicados a la extracción de rasgos faciales: modelo de proyecciones integrales y modelo de forma activa", *IX Reunión Regional de Ramas IEEE*, Acapulco, México, 2006.
- [32] L. Yin, and A. Basu, "Nose shape estimation and tracking for model based coding", *Acoustics, Speech and Signal Processing*, vol. 3, pp. 1477-1480, 2001.
- [33] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 721-732, 1997.

