

Detección Estable de los Bordos de la Oreja en Imágenes 2D

Stable Ear Border Recognition in 2D Images

Alejandro Flores Méndez y Ana Antígona Méndez Cuanalo
LIDETEA, Universidad la Salle
aflores@ci.ulsa.mx, fierro.mendez@hotmail.com

Artículo recibido en Octubre 22, 2007; aceptado en Agosto 18, 2008

Resumen

En este reporte se presenta un nuevo método para la detección estable de los bordes principales de la oreja a partir de imágenes en colores RGB. El primer paso segmenta la imagen para eliminar el cabello y cualquier rasgo que no fuera piel, esto a partir de los datos obtenidos después de transformar la imagen RGB a HSV. La imagen RGB es convertida en otra en tonos de gris y de ésta se elimina la información de aquellos píxeles que no hubieran sido seleccionados como piel. Sobre esta imagen se emplea una serie de técnicas para mejorar la detección de bordes asociados a los principales rasgos de la oreja, entre los que destacan los filtros morfológicos. Cabe señalar que el sistema tuvo un buen comportamiento aún en presencia de ruido, diferentes condiciones de iluminación para las imágenes capturadas, diversos colores de piel de las personas, etc. Este trabajo es un primer paso para la identificación y reconocimiento de personas a partir de los bordes de la oreja.

Palabras Clave: procesamiento de imágenes, reconocimiento de bordes, oreja.

Abstract

This paper contains a novel method for the stable border recognition of the main features of the ear for RGB images. For the recognition, we first do the segmentation of the image in order to eliminate the hair and any feature other than skin through the data contained in the HSV transformation. The RGB image is converted to a grayscale one, and from it the information of the pixels not related to skin is discarded. Then a series of techniques, mainly morphological filters, are used to enhance the border detection associated to the main features of the ear. The proposed system achieved good performance even in the presence of noise, different light conditions, skin colors, etc. This proposal has been thought as a first step toward human identification and recognition through ears borders.

Keywords: image processing, border recognition, ear.

1 Introducción

Identificar una persona, es establecer su individualidad a partir de aquellos rasgos o conjunto de cualidades que la distinguen de todos los demás y que hacen que sea ella misma. Cuando los rasgos a partir de los cuales se identifica a una persona son cualidades conductuales o físicas intrínsecas de ésta, se dice que se están empleando elementos biométricos, siendo el área que realiza la identificación a partir de los elementos biométricos la biometría. Si bien la biometría ha destacado en fechas recientes, no está por demás señalar que sus orígenes datan de principios del S. XIX con los trabajos de César Lombroso y Lambert A. Quetelet [9]. Dentro del mundo de la biometría existen múltiples métodos de identificación con base en ciertos aspectos del cuerpo humano como lo son: sistemas de voz, huellas dactilares, geometría de la mano, verificación de escritura y de la forma de la oreja [13]¹, entre otros.

Cabe destacar que los trabajos que emplean alguna o algunas de estas características tienen distintos tipos de avances. De hecho, aún se está evaluando la viabilidad del empleo de la oreja en la biometría, siendo una de las investigaciones en ese sentido la de los expertos de la Universidad de Leicester (Gran Bretaña) quienes están desarrollando esta tecnología a partir de fotografías de este órgano para su posterior análisis y comparación a través

¹ En 1965 se realizó en Suiza la primera identificación de un delincuente por la huella de su oreja. En España (nov. 2001) sucedió la primera sentencia condenatoria con base en evidencia parcial de este tipo.

de un software con una base de datos [1]. Sobre esta misma línea, recientemente se han realizado distintos esfuerzos para la identificación de personas a partir de las orejas, destacándose los trabajos de Yan y Bowyer [23][24][25][26][27] en donde se realiza la identificación a partir de imágenes 3D con resultados prometedores; sin embargo, éstos emplean procesos que en general son semiautomáticos y para los cuales se requiere usar marcas (*landmarks*) para la segmentación de la oreja; o bien, se hace una segmentación *ad hoc* para la cual se toman en cuenta ciertas restricciones respecto a la distancia y el ángulo con que fue capturada la imagen. Adicionalmente, en [25] los autores refieren que en general desechan la identificación a partir de imágenes bidimensionales por los problemas asociados a la segmentación de la oreja y a que los contornos que obtenían no eran estables². Otros trabajos en el área de identificación a partir de las orejas son los incluidos en [2][6][7][8][12]. Ahora bien, de éstos sólo los dos últimos se concentran en la extracción de datos a partir de imágenes bidimensionales, y sólo el artículo de Hurley, Nixon y Carter [12] ofrece una alternativa para extraer características adecuadas para la identificación; sin embargo, no es a partir de los bordes. La obtención de bordes estables para los rasgos principales de la oreja (incluidos en la Fig. 1) a partir de imágenes bidimensionales es pues importante, ya que permite emplear equipo menos especializado, además de que es computacionalmente menos exhaustivo pues procesa menos información. Por otro lado, se debe destacar que la técnica para la identificación de los bordes de la oreja aquí incluida es totalmente automatizada, contrario a lo que se encuentra en general en la literatura.

Para describir el proceso mediante el cual se obtuvieron los bordes principales de las orejas a partir de imágenes en colores se ha organizado la información de la siguiente manera. En la Sección II se presentan las técnicas de procesamiento de imágenes utilizadas en este trabajo. En la Sección III se describen diversos esquemas desarrollados para su detección, tras lo cual se incluyen en la Sección IV algunos ejemplos de los resultados obtenidos con el sistema propuesto. Se finaliza el trabajo incluyendo las conclusiones y el trabajo futuro.

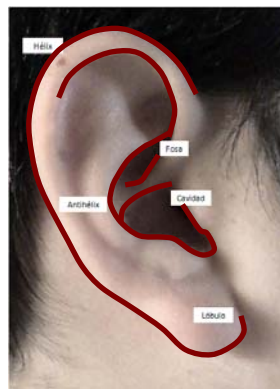


Fig. 1. Rasgos principales de la oreja. El helix, es la parte del contorno exterior de la oreja; la fosa, formada por el pliegue que une a la cara con la oreja, el antihelix es el pliegue formado entre el helix y la fosa; la cavidad es donde se encuentra ubicada la entrada del oído, y el pliegue final es llamado lóbulo

2 Procesamiento de Imágenes

2.1 Definición de imagen

Una imagen digital es un arreglo bidimensional de *pixeles*. El valor de cada pixel se representa mediante una función f , donde $f(x, y)$ representa el nivel de brillantez, color, o intensidad de la imagen en tales coordenadas. Por lo tanto, una imagen en blanco y negro puede ser representada por una matriz de dimensión $m \times n$, donde $f(x, y) \in \{0, 1\}$. Además de la representación en blanco y negro, otras opciones son las imágenes en tonos de gris, para las cuales

² Que los contornos sean estables se refiere a que los resultados para dos tomas de un mismo individuo no deben incluir variaciones importantes.

$f(x,y)$ representa un nivel de intensidad típicamente asociado a un entero entre 0 y 255 y otras más sirven para describir imágenes en color [10].

2.1.1 Imágenes en color: los espacios RGB y HSV

Existen varias representaciones de imágenes de color (RGB, CMY, HSV, HSI, entre otras). La diversidad de espacios o representaciones del color es explicada por todos los caminos que ha seguido el hombre en aras de esclarecer el proceso de percepción del color. Ahora bien, para las imágenes en el espacio RGB, se tiene que cada pixel está asociado con la intensidad de tres componentes, Rojo, Verde y Azul (o *RGB* por sus siglas en inglés, de donde reciben su nombre). Este espacio se basa en el modelo de síntesis aditiva, mostrado en la Fig. 2, aunque existen otras formas de representar imágenes en color como el modelo HSV.



Fig. 2. Síntesis aditiva empleada por el modelo RGB

El modelo de color *HSV*, creado por Alvy Ray Smith (1978) [18], define el color en términos de su tonalidad, saturación y valor, también referido como brillo, (o en inglés *Hue*, *Saturation* y *Value*). La tonalidad se asocia al tipo de color y se representa mediante un ángulo cuyos posibles valores van de 0 a π . La saturación, también llamada pureza de excitación o pureza colorimétrica, indica el nivel de saturación del color. Así tenemos que a menor saturación, mayor será la tonalidad grisácea. Por otra parte, el valor del color o brillo representa la “distancia” al blanco/negro. Este modelo es una alternativa al RGB cuando se desean detectar pixeles con tonos similares (Fig. 3), lo cual permite a la postre segmentar una imagen [11][15]. Para nuestro caso, permite distinguir la piel del cabello.

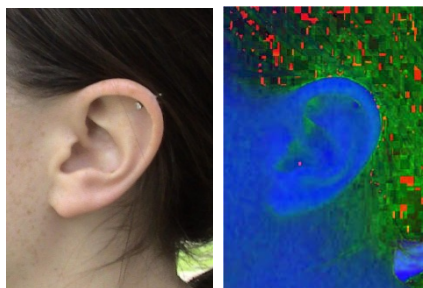


Fig. 3. Imágenes de la oreja en los espacios RGB (izq.) y HSV (der.)

Para transformar una imagen del modelo RGB al HSV [11] se emplearon las siguientes ecuaciones³:

$$D = \cos^{-1} \left(\frac{\frac{1}{2}[(R - G) + (R - B)]}{\sqrt{(R - G)^2 + (R - B)(G - B)}} \right) \quad (2.1)$$

³ A lo largo del documento en las ecuaciones se empleará la abreviatura *c.c.* para referirse al texto *en caso contrario*.

$$H = \begin{cases} D & \text{si } B \leq G \\ 360 - D & \text{c.c.} \end{cases} \quad (2.2)$$

Para el valor -de luminosidad- (V) y la saturación (S) tenemos que:

$$V = \max\{R, G, B\} \quad (2.3)$$

$$S = \begin{cases} (V - \min\{R, G, B\})/V & \text{si } V \neq 0 \\ 0 & \text{c.c.} \end{cases} \quad (2.4)$$

Otra transformación comúnmente empleada convierte una imagen del modelo RGB a otra en tonos de gris. Para llevar a cabo este proceso se empleó la siguiente ecuación:

$$Gris = (\max\{R, G, B\} + \min\{R, G, B\})/2 \quad (2.5)$$

Esta transformación es común porque en ocasiones permite manejar un menor volumen de información, sin que esto influya negativamente en la obtención de los resultados (Fig. 4).



Fig. 4. Imagen en tonos de gris

El *Umbralado* (thresholding), consiste en determinar si el valor (tono) de una imagen es mayor o igual a un cierto umbral (θ). Así se puede determinar una nueva imagen como:

$$th_{\theta}(x, y) = \begin{cases} 1 & \text{si } f_i(x, y) \geq \theta \\ 0 & \text{c.c.} \end{cases} \quad (2.6)$$

donde th_{θ} genera una imagen de salida para el umbral θ , mientras que f_i representa la intensidad para la imagen de entrada I . En particular esto puede utilizarse para la binarización de imágenes en tonos de gris, aunque en nuestro proceso también se emplea para la segmentación.

2.2 Segmentación de imágenes

La *segmentación* consiste en separar una imagen en elementos que correspondan a unidades significativas. Ésta se realiza con base en criterios de “homogeneidad” como lo son tono, textura, profundidad, entre otros; aunque para tal fin también es posible considerar discontinuidades, bordes, movimiento, etc. [11]. Para nuestro caso de estudio, habremos de segmentar la imagen de acuerdo a la tonalidad, para lo cual se emplea una imagen en el espacio HSV (Fig. 5) en conjunción con el umbralado.

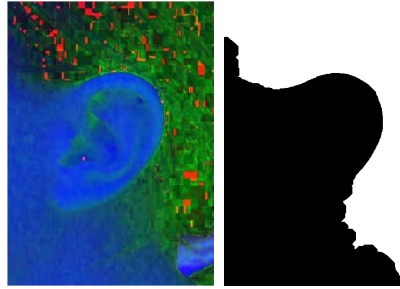


Fig. 5. Imagen HSV a segmentar (izq.) y pixeles seleccionados por f_{and} (der.)

Ahora bien, para la segmentación sobre H es conveniente definir una función de distancia. Sea $H(x,y)$ la componente H del pixel (x,y) y H_d un tono deseado. Entonces podemos definir d_H como la distancia al tono deseado como:

$$d_H = \begin{cases} |H - H_d| & \text{si } |H - H_d| \leq 0.5 \\ 1 - |H - H_d| & \text{c.c.} \end{cases} \quad (2.7)$$

De manera tal que para seleccionar (segmentar) los tonos alrededor de un tono deseado, basta sustituir en (2.6) la función f_I por d_H y elegir un valor de umbral específico.

Para la segmentación con base en V , en (2.6) se sustituye f_I por la diferencia absoluta entre $V(x,y)$ y V_d , donde V_d representa el valor deseado. El resultado de la segmentación con respecto a ambos valores (H y V) puede ser resuelto de diversas formas: como la intersección de ambas imágenes, la conjunción de éstas, etc. Sea entonces f_{and} la intersección de la segmentación a partir de H y V , y f_I una imagen de entrada, entonces se calcula la imagen de salida f_{seg} como:

$$f_{seg}(x,y) = \begin{cases} f_I(x,y) & \text{si } f_{and}(x,y) == 1 \\ 0 & \text{c.c.} \end{cases} \quad (2.8)$$

Una vez segmentada la imagen, ésta aún contiene características no deseadas, es decir “ruido”. A continuación se describe cómo es posible eliminar éste.

2.3 Filtrado

Gran parte del ruido en una imagen (f) se puede eliminar filtrándola, donde *filtrar* significa aplicar una transformación (T) para obtener una nueva imagen (g) con determinadas características acentuadas o atenuadas, lo cual se expresa formalmente como:

$$g(x,y) = T[f(x,y)] \quad (2.9)$$

Una forma comúnmente empleada para filtrar una imagen es a través de la *correlación*, la cual relaciona una imagen (f) con una mascarilla (h). Esto se denota por $f * h$, y se define como:

$$g(x,y) = (f * h)(x,y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f(x+k,y+l)h(k,l) \quad (2.10)$$

donde las coordenadas de los elementos que componen a la mascarilla se relacionan con \mathbf{Z}^2 . Así pues, normalmente para una mascarilla definida por una matriz de 3×3 se asume que sus respectivas coordenadas son:

(-1,-1)	(-1,0)	(-1,1)
(0,-1)	(0,0)	(0,1)
(1,-1)	(1,0)	(1,1)

Fig. 6. Coordenadas de una mascarilla de 3×3

La correlación es un procedimiento que se utiliza para modificar imágenes digitales a partir de la aplicación de una entrada arbitraria y una mascarilla predeterminada, la cual modifica el contenido de cada pixel de la imagen de acuerdo con el valor de los pixeles vecinos o colindantes los cuales son ponderados en términos del contenido de la mascarilla.

Para este trabajo se utiliza el *Filtro Gaussiano*, el cual permite eliminar algunos rasgos indeseables en la imagen como brillos, pecas, poros de la piel, entre otros, y cuyos bordes se consideran como “ruido” para la aplicación (Fig. 7).



Fig. 7. Imagen original (izq.) e imagen filtrada (der.)

Los elementos de la mascarilla se determinan de acuerdo a la siguiente función:

$$h(x, y) = e^{-\left(\frac{x^2 + y^2}{2\sigma^2}\right)} \tag{2.11}$$

Para esto, es necesario seleccionar el tamaño de una mascarilla adecuada que se aplica a la imagen y calcular el valor del elemento (pixel) central de la mascarilla como la suma de los productos de entre los valores contenidos en la vecindad de la mascarilla y los de la imagen original [15], de acuerdo a (2.10). Por ejemplo, la mascarilla para un filtro Gaussiano asociado a una matriz de 5×5 obtenida mediante (2.11) para $\sigma = 1$ es:

$$h = \begin{bmatrix} 0.02 & 0.08 & 0.14 & 0.08 & 0.02 \\ 0.08 & 0.37 & 0.61 & 0.37 & 0.08 \\ 0.14 & 0.61 & 1.0 & 0.61 & 0.14 \\ 0.08 & 0.37 & 0.61 & 0.37 & 0.08 \\ 0.02 & 0.08 & 0.14 & 0.08 & 0.02 \end{bmatrix} \tag{2.12}$$

Una vez eliminado el ruido, se facilita la detección de *contornos*, la cual se detalla enseguida.

2.4 Detectores de contornos

Los contornos son aquellos pixeles con un tono distinto del de los que están en su vecindad. Los contornos se pueden obtener mediante la correlación en conjunto con el umbralado. Éste es el proceso empleado para las técnicas más simples como Roberts, Sobel y Prewitt; sin embargo, estas técnicas de detección de los contornos en una imagen son muy sensibles al ruido, y aún cuando en nuestro caso la imagen ya ha sido procesada mediante un filtro Gaussiano,

es conveniente utilizar técnicas de detección de contornos de segundo orden como lo son el Laplaciano y Canny [14] entre otras. De hecho, fue esta última técnica, con la que se obtuvieron los mejores resultados tras experimentar con distintos valores para los métodos mencionados sobre las imágenes ya segmentadas, convertidas a tonos de gris y filtradas (Fig. 8).



Fig. 8. De izquierda a derecha se muestran los mejores bordes obtenidos con Prewitt, Roberts, Zero-crossing, Sobel y Canny

Ahora bien, estos bordes o contornos contienen aún porciones que no son de interés para el presente trabajo, pues se desea obtener fundamentalmente los contornos del hélix, antihélix, fosa, cavidad y lóbulo. Una forma de eliminar algunos de los bordes espurios, es a través de los operadores morfológicos, los cuales se describirán a continuación.

2.5 Morfología matemática

La morfología matemática cuenta con dos operadores básicos llamados *dilatación* y *erosión* los cuales normalmente operan sobre una imagen f y una mascarilla⁴ h binarias. La dilatación, denotada por el operador \oplus , se define como:

$$(f_i \oplus h)(x, y) = (th_1(f_i * h))(x, y) \tag{2.13}$$

Por su parte, la erosión, denotada por \ominus , se define por:

$$(f_i \ominus h)(x, y) = (th_{\#_1(h)}(f_i * h))(x, y) \tag{2.14}$$

siendo $\#_1(h)$ una función que cuenta el número de 1's que aparecen en h . Así tenemos que en realidad la única diferencia entre ambas funciones es el umbral.

La erosión es un operador morfológico cuyo resultado elimina las estructuras que no “contengan” a la mascarilla. La dilatación por su parte, es la operación “dual” de la erosión, y su efecto sobre la imagen es engrosar los objetos siempre que la mascarilla se “intersecte” con alguno de los elementos que la conforman [11]. La composición de estos operadores permite generar otros que son conocidos como *clausura* y *apertura*. Para la clausura, se realiza una erosión tras la dilatación; es decir:

$$clausura_{x,y}(f_i, h) = ((f_i \oplus h) \ominus h)(x, y) \tag{2.15}$$

mientras que en la apertura se invierte este orden. Estas funciones reciben ese nombre porque su efecto es justamente el de “cerrar” o “abrir” espacios existentes entre objetos que aparecen en una imagen siempre que éstos estén lo suficientemente cercanos. En la siguiente figura (Fig. 9), se muestra el efecto del operador de clausura.

⁴ La mascarilla, dentro de morfología matemática, recibe el nombre de **elemento de estructura**.

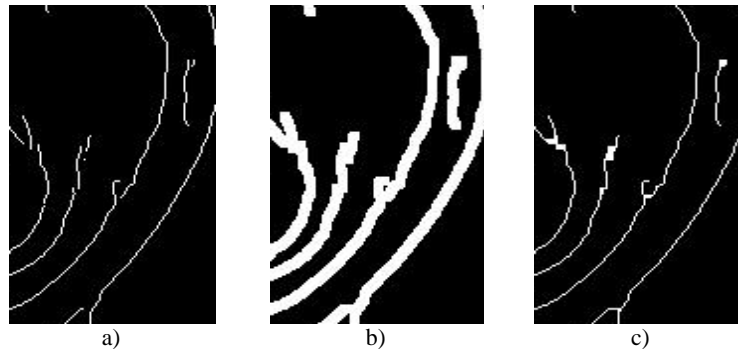


Fig. 9. a) Detalle de los bordes, b) dilatación y c) clausura (erosión después de la dilatación)

Un ejemplo más de una función morfológica que se apoya en los operadores básicos es la *contracción*, la cual elimina elementos en la frontera de un objeto, preservando su característica de Euler [14]. Así pues, podemos definir la contracción como:

$$\text{contracción}_{x,y}(f_1, g) = (\text{th}_2(f_1 * g))(x, y) \& f_1(x, y) \quad (2.16)$$

Las técnicas de esta sección se combinaron para detectar los contornos deseados como se indica a continuación.

3 Descripción de la propuesta para la Detección de Contornos

A diferencia de los trabajos realizados por Yan y Bowyer [4][5][6][7][8] y los de otros autores [9][10][11], concentrados principalmente en la identificación de una persona a partir de imágenes 3D que contienen orejas a partir de procesos semiautomáticos, el presente trabajo incluye una alternativa para obtener elementos que sirvan a la identificación a partir de imágenes bidimensionales, siendo otros trabajos en este sentido [7][8][12][13], aunque ninguno de éstos ofrece un proceso para la obtención de bordes estables de los rasgos principales de las orejas.

Los principales bordes de una oreja, mostrados en la Fig. 1, son elementos a través de los cuales es posible identificar o reconocer a un individuo. Para su detección, se consideró que dadas las condiciones de captura, la luz puede provocar sombras y/o brillos que nos lleven a detectar contornos que no se desean. Adicionalmente, el cabello también se desea eliminar, pues los contornos asociados al mismo se considerarían ruido al no describir la geometría de la oreja. Así pues, para la detección de los contornos se realizaron una serie de pasos que se describen en el siguiente pseudocódigo, el cual recibe una imagen RGB como argumento de entrada y genera una imagen en blanco y negro con los contornos deseados:

```
Ear-edges (im)
1. [h,s,v] ← rgb2hsv (im)
2. im_seg ← segment (im, h, v, thH, thV)
3. im_gray ← rgb2gray (im_seg)
4. im_gauss ← fGauss (im_gray)
5. im_canny ← canny (im_gauss)
6. im_cerr ← imclose (im_canny, nClose)
7. im_res ← del_small (im_cerr, thSmall)
8. return im_res
```

La función `rgb2hsv` convierte una imagen RGB a otra HSV. Para ello se utilizan las ecuaciones (2.1) a (2.4). Los resultados que se obtienen a partir de tal proceso se presentan en la Fig. 3.

A partir de la imagen en el espacio HSV se segmenta la piel como aquella región para la que se tienen coeficientes similares de tono y valor. Los valores deseados para éstos se determinan de manera dinámica con base

en las características existentes en cada imagen como sigue. Para la segmentación sobre H se calcula la moda y se considera este valor para H_d en (2.6), mientras que para V_d se emplea un valor igual al 80% de la media de V . Sus respectivos valores de umbral son en este caso $thH = thV = 0.2$. Este proceso de segmentación (Fig. 5), realizó en términos generales correctamente la división entre la piel y el cabello de la imagen. El propósito de haber obtenido la moda y la media para la segmentación se debe a que esto permite que el sistema se adapte mejor a las diversas características de las imágenes capturadas relativas a la luz, el tamaño, color de piel y cabello, etc. Así pues, el resultado de `segment` son todos aquellos pixeles de `im` que hubieran sido seleccionados por el proceso descrito en la sección 0.

La imagen resultante del proceso de segmentación (`im_seg`), la cual emplea el modelo RGB, es transformada en otra en tonos de gris vía `rgb2gray` por medio de la ecuación (2.5). Esto se hace necesario para poder eventualmente calcular los bordes de la imagen, aunque la experimentación nos mostró que era preciso realizar primero un filtrado Gaussiano, ya que aún se detectaban bordes debidos a “imperfecciones de la piel” como poros, pecas, etc. En este caso se utilizó (2.12) como mascarilla para el filtro Gaussiano. Tras el filtrado Gaussiano (Fig. 7), muchos de los rasgos que no nos interesan se han eliminado en la imagen filtrada (`im_gauss`), dejando sólo aquellos rasgos más representativos que distinguen la oreja.

El siguiente paso consiste en detectar los bordes principales de la imagen. Como se mencionó anteriormente, se experimentó con diferentes técnicas para obtener éstos, siendo Canny el método con el que se obtuvieron los mejores resultados (Fig. 8). Ahora bien, para resaltar la utilidad de los pasos previos, en la Fig. 10 se incluyen dos imágenes con los bordes generados por Canny sobre la imagen original, y el resultado tras la segmentación y el filtrado Gaussiano.



Fig. 10. Detección de bordes con Canny sobre la imagen original y sobre la imagen segmentada y tratada con el filtro Gaussiano

Tras obtener los bordes se aplica un proceso de clausura (`imclose`), el cual considera como imagen de entrada el resultado de Canny (`im_canny`). Este proceso utiliza una mascarilla con dimensiones `nClose`×`nClose`, el valor de `nClose` se fijó experimentalmente en 7. El objetivo de este proceso es reducir el número de objetos que aparecen en `im_canny`.

Por último `thSmall`, elimina los objetos más pequeños de una imagen a través de la aplicación reiterada del operador de contracción. En nuestro caso se reitera este operador 15 veces.

Como producto de este proceso, se logró obtener en la mayoría de los casos los bordes característicos de la oreja que son el hélix (borde), antihélix, fosa y cavidad con una cantidad de ruido aceptable (Fig. 11).



Fig. 11. Imagen original y resultado final

4 Experimentación y Resultados

Para el presente trabajo se capturaron las imágenes empleando una cámara tipo Handycam de Sony 800x Modelo DCR-SR60. Las imágenes fueron capturadas sin establecer previamente la distancia de captura por lo que éstas tienen diferentes dimensiones que oscilan entre los 470 y 560 píxeles de ancho por 650 a 950 de alto. Las diferencias en las dimensiones fueron el resultado de seleccionar manualmente el área de interés (la cual sencillamente es el área en que está contenida la oreja). La única fuente de iluminación en el momento de la captura era la propia del sitio en que se obtuvo la imagen.

La implantación del método descrito se llevó a cabo empleando MATLAB[®] junto con el Toolbox de procesamiento de imágenes sobre una máquina con SO Windows XP/SP2 y procesador Intel Centrino de 1.73GHz con 512 MB en RAM. Para cada una de las pruebas en promedio se requirieron de aprox. 8 segs. En la siguiente figura se presentan los resultados de los distintos pasos que permitieron la detección de los bordes de la oreja para una imagen con poco ruido; es decir, bien enfocada, sin cabello sobre la oreja y un área bien delimitada entre piel y cabello y sin aretes (Fig. 12).



Fig. 12. Secuencia de imágenes resultantes mediante el algoritmo Ear-edges para la detección de los bordes de la oreja. La imagen en la extrema izquierda representa la entrada, y las siguientes imágenes corresponden a las salidas para los pasos 1 al 7

A continuación se presentan algunas imágenes de orejas en distintas condiciones de iluminación y distancia. La primera de éstas figuras (Fig. 13), presenta los resultados obtenidos para personas de color, una de las cuales tenía un peinado tipo afro el cual dificultaba la detección de los bordes de la oreja, pues la piel y el cabello eran más difíciles de segmentar. Esta figura permite mostrar cómo gracias a la determinación dinámica de los tonos, se puede hacer una segmentación efectiva del área correspondiente al cabello y a la piel.



Fig. 13. En las columnas de izquierda a derecha aparecen las imágenes correspondientes a personas de color, el resultado de la segmentación, y el resultado final del sistema

Ahora bien, dado que el interés del presente trabajo es detectar los bordes más importantes de la oreja, es fundamental que éstos sean aproximadamente los mismos para una misma persona en dos tomas distintas. En la Fig. 14, se presentan dos imágenes capturadas de dos personas junto con los bordes de las orejas detectados en cada una. En estas imágenes la diferencia de los contornos es mínima, este fue el caso para la mayoría de las imágenes estudiadas, por lo que se puede considerar que los bordes obtenidos son estables. Esta estabilidad, podría presentar la posibilidad de llevar a cabo la identificación de personas mediante imágenes en 2D, pues el procesamiento con tales imágenes es computacionalmente menos exhaustivo que el necesario para procesar las imágenes 3D utilizadas en los artículos ya citados. De hecho, vale la pena mencionar que los primeros pasos en ese sentido ya se están dando, y que actualmente para un grupo de aprox. 30 personas los resultados preliminares apuntan a que se tienen valores comparables con los reportados en [5].



Fig. 14. Dos imágenes diferentes de dos personas junto con los resultados obtenidos por el sistema

Por último, y cómo se mencionó con anterioridad, se realizaron distintas pruebas con imágenes con dimensiones diferentes, con las dimensiones de las orejas también distintas, pobres condiciones de iluminación, imágenes mal enfocadas, personas con algunas manchas en la piel e incluso imágenes a las que se les adicionó intencionalmente ruido. Ejemplos de los resultados obtenidos para estos casos se incluyen en la siguiente figura (Fig. 15).

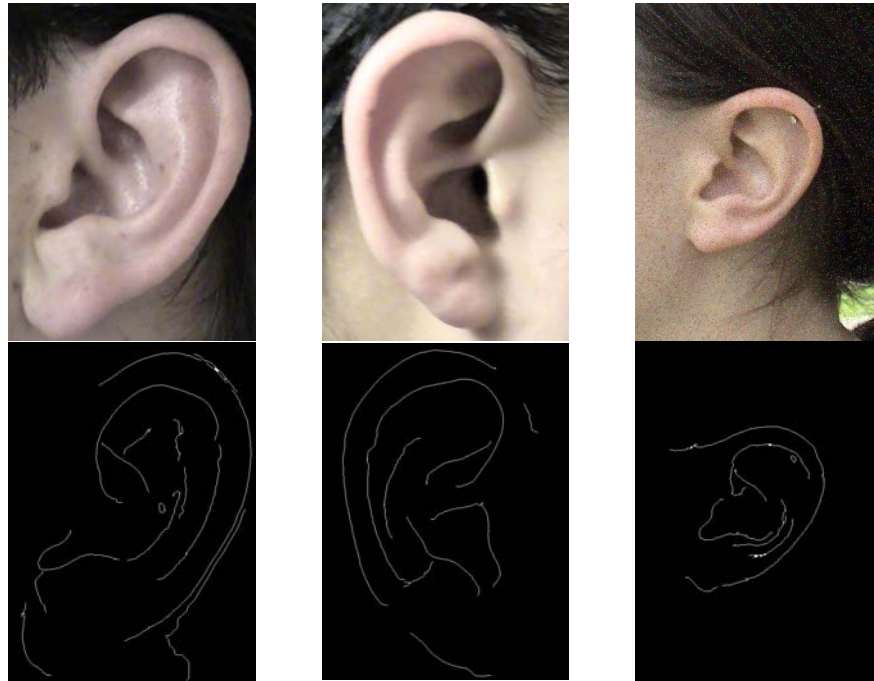


Fig. 15. Ejemplos de los resultados obtenidos (de izquierda a derecha) para imágenes con dimensiones distintas a aquellas con las que se realizaron las pruebas, mal enfocadas y con ruido sal y pimienta al 5%

5 Conclusiones y Trabajo futuro

El método propuesto permite la detección de contornos característicos de las orejas, y aunque aún se incluye cierta cantidad de ruido, este es un paso importante para identificar a las personas a través de los bordes de sus orejas de manera automática ya que la técnica ofrece una posibilidad para la obtención de bordes estables de una imagen que contiene una oreja en colores. Más aún, en este sentido ya se han realizado las primeras pruebas para la identificación de un individuo sobre un conjunto reducido (30 personas), y aunque estos datos son preliminares actualmente se tiene una identificación de rango 1 de algo más del 70%; aunque algunas ideas nuevas podrían llevarnos a mejorar este porcentaje, siendo ésta la mejor prueba de si nuestro método ofrece o no bordes estables.

El procedimiento descrito para la detección de bordes es robusto a los cambios de iluminación y es funcional para diferentes colores de piel. Además consigue detectar los bordes aún cuando las dimensiones de las orejas y de las propias imágenes se encuentren fuera de los rangos estipulados en el documento (aunque claro, dentro de ciertos límites), esto podría ayudar a mejorar las actuales tasas de identificación a partir de imágenes en 2D, la cual es una de las líneas que se buscarán estudiar como trabajo futuro. Así mismo otra línea sería la experimentación con imágenes en 3D, pues parece razonable emplear algunos de los elementos propuestos aquí para la detección de bordes en estas imágenes.

También como trabajo futuro, cabe destacar que los individuos involucrados fueron todas personas de entre 20 y 35 años; sin embargo, sería experimentalmente importante para la parte de identificación validar los resultados con

personas de la tercera edad y con niños menores de cinco años, atendiendo en particular los cambios que se dice sufre la oreja en su forma conforme avanza el tiempo.

Otra de las líneas en que se debe trabajar para mejorar las imágenes resultantes de los bordes de la oreja sería restringir las condiciones de captura, en particular en lo referente a la iluminación y distancia con que cada imagen es capturada.

Por último, se considera importante establecer un banco de imágenes de orejas, para eventualmente emplear estos resultados en la identificación/reconocimiento de personas.

Referencias

- 1 **Abbas, A., and G. Ruty.** "The Role of the Human Ear in the process of Forensic Identification." *BMedSci Thesis*. Universidad de Leicester. División de Patología Forense, Departamento de Patología, May 2003.
- 2 **Alvarez, L., E. González, and L. Mazorra.** "Fitting ear contour using an ovoid model." Departamento de Informática y Sistemas. Universidad de Las Palmas de Gran Canaria, May 27, 2005.
- 3 **Burge, M., and W. Burger.** "Ear Recognition." *Biometrics: Personal Identification in Networked Society*. Edited by R. Bolle and S. Pankanti. Kluwer Academic Publishing, 1999. 273-286.
- 4 **Canny, J.** "A Computational Approach to Edge Detection." *IEEE Trans. on PAMI*, 1986: 679-698.
- 5 **Chang, K., K. Bowyer, and P. Flynn.** "Face recognition using 2D and 3D facial data." *Workshop on Multimodal user Authentication*. 2003. 25-32.
- 6 **Chang, K., K. W. Bowyer, S. Sarkar, and V. Barnabas.** "Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, September 2003: 1160-1165.
- 7 **Chen, H., and B. Bhanu.** "Contour Matching for 3D Ear Recognition." *In Seventh IEEE Workshops on Application of Computer Vision*, 2005: 123-128.
- 8 —. "Human Ear Recognition in 3D." *IEEE Trans. on PAMI*, April 2007: 738-742.
- 9 **Curriel López de Arcuate, A. M., and J. Granell Navarro.** "La huella de oreja como método de identificación." *Acta Otorrinolaringol Esp* 57 (2006): 329-332.
- 10 **Flores-Méndez, A.** "Visión por Computadora. Apuntes de Maestría en Ciencias, Área Cibernética. ULSA." Mexico, D.F., 2007.
- 11 **Gonzalez, Rafael, and Richard Woods.** *Digital Image Processing*. 3. New Jersey: Pearson Prentice Hall, 2008.
- 12 **Hurley, David J., Mark S. Nixon, and John N. Carter.** "Force field feature extraction for ear biometrics." *Computer Vision and Image Understanding*. Department of Electronics and Computer Science, University of Southampton, November 10, 2004.
- 13 **Jain, A. K., A. Ross, and S. Pankanti.** "Biometrics: A Tool for Information Security." *IEEE Trans. on Information Forensics and Security*, June 2006: 125-143.
- 14 **Kong, T.Y., and A. Rosenferg .** *Topological Algorithms for Digital Image Processing*. Elsevier, 1996.
- 15 **Nixon, M., and A. Aguado.** *Feature Extraction and Image Processing*. 2. Academic Press, 2008.
- 16 **Ortiz-Zamora, F., and F. Torres-Medina.** "Procesamiento morfológico de imágenes en color. Aplicación a la reconstrucción geodésica." Universidad de Alicante. Departamento de física, Ingeniería de sistemas y Teoría de la señal, May 2002.
- 17 **Ratha, N., A. Senior, and R. Bolle.** "Automated Biometrics." *Proceedings of International Conference on Advances in Pattern Recognition*. Río de Janeiro, Brasil, March 2001.
- 18 **Ray Smith, Alvy.** "Color gamut transform pairs." *ACM SIGGRAPH Computer Graphics*, August 1978: 12-19.
- 19 **Sossa Azuela, H.** *Rasgos descriptores para el reconocimiento de objetos*. México: Ciencia de la Computación. CIC, IPN, 2006.
- 20 **Sucar, Luis Enrique.** "Visión Computacional Apuntes Dr. Luis Enrique Sucar." Cuernavaca, Edo. Mexico: Departamento de Computación. ITESM, July 2007.
- 21 **Vizcaya, P., and L. Gerhardt.** "A Nonlinear Orientation Model for Global Description of Fingerprints." *Pattern Recognition*, 1996: 1221-1231.

- 22 **Wayman, J., A. Jain, D. Maltoni, and D. Maio.** *Biometrics Systems: Technology, Design and Performance Evaluation*. Springer, 2005.
- 23 **Yan, P., and K.W. Bowyer.** "An automatic 3D Ear Recognition System." *Proc. Third Int'l Symp. 3D Data Processing, Visualization, and Transmission*. 2006. 213-218.
- 24 **Yan, P., and K.W. Bowyer.** "Ear Biometrics using 2D and 3D images." *In Advance 3D Imaging for Safety and Security 3* (June 2005): 121-121.
- 25 —. "Empirical Evaluation of Advanced Ear Biometrics." *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshop Empirical Evaluation Methods in Computer Vision*. June 2005. 41 - 41.
- 26 —. "Biometric Recognition Using 3D Ear Shape." *IEEE Trans. on PAMI*, August 2007: 1297-1308.
- 27 —. "Multi-Biometrics 2D and 3D ear recognition." *LNCS: In Audio- and Video-based Biometric Person Authentication*. 2005. 503-512.



Alejandro Flores Méndez es Ing. en Cibernética y Sist. Computacionales (1991-1996), Universidad la Salle (ULSA); M. en C., Especialidad Computación, CINVESTAV-IPN (1996-1998). Actualmente se encuentra haciendo sus estudios de Doctorado en el Depto. de Matemáticas del CINVESTAV-IPN (2003-). Áreas de interés: Computación Inteligente, Visión por Computadora, Optimización Combinatoria. Cuenta con más de 20 artículos en congresos, 8 artículos en revistas con arbitraje, conferencista principal en 7 ocasiones. Actualmente es investigador de la ULSA y miembro de LIDETEA.



Ana Antígona Méndez Cuanalo Egresada de la Carrera de Ingeniería en Cibernética y Ciencias Computacionales en la Universidad La Salle (1986-1991), México. Estudiante de la Maestría en Ciencias – Área Cibernética de esta misma Universidad (2006-2008).