

Localización y Seguimiento de Caras Usando Búsqueda en Rejilla

Ginés García Mateos

Dept. de Informática y Sistemas, Universidad de Murcia

30.070 Espinardo, Murcia, España

Tlf: (+34) 968 36 4608 Fax: (+34) 968 36 4151 Email: ginesgm@um.es

Cristina Vicente Chicote

Dept. Tecnología de la Información y las Comunicaciones, Universidad Politécnica de Cartagena

30.202 Cartagena, Murcia, España

Tlf: (+34) 968 32 6448 Fax: (+34) 968 32 5338 Email: cristina.vicente@upct.es

Andrés García Meroño

Consejería de Trabajo y Política Social, C.A.R.M.

Avenida de la Fama 3, 30.071 Murcia, España

Tlf: (+34) 968 36 5161 Fax: (+34) 968 36 4151 Email: andres.garcia2@carm.es

RESUMEN

En este artículo se tratan los problemas de localización y seguimiento en tiempo real de caras humanas y de ciertas componentes faciales relevantes (ojos, cejas, nariz y boca), en secuencias de imágenes en color. Haciendo uso del hecho de que el matiz de color es una característica fundamental de la piel humana, se ha realizado un estudio de los espacios de color adecuados al modelado de color de piel humana, que serán usados en el proceso de segmentación por color. Para acelerar el proceso, se ha propuesto y desarrollado una nueva técnica que permite abordar en tiempo real la segmentación de las regiones de interés, así como su localización y caracterización. Esta técnica, denominada búsqueda logarítmica en rejilla, constituye nuestra principal contribución en este trabajo. Presentamos un resumen de las pruebas llevadas a cabo, y por último se discuten las principales conclusiones y posibles vías futuras.

Palabras clave: procesamiento de imágenes, seguimiento de caras humanas, espacios de color, búsqueda logarítmica en rejilla.

1. INTRODUCCION

Dentro del área de trabajo de procesamiento de imágenes y sonido, es destacable el creciente interés por los problemas relacionados con el tratamiento de imágenes de personas. Este hecho es constatable en el elevado número de conferencias y publicaciones avaladas por los más prestigiosos grupos de investigación internacionales [1]. De acuerdo con Alex Petland [2], las aplicaciones clave dentro del procesamiento visual de personas son: la identificación a través de la cara (en general mediante las técnicas conocidas como biométricas); la supervisión, monitorización y seguimiento 3D de personas; las habitaciones inteligentes y los interfaces perceptuales.

Gran parte de esta investigación se dedica a la resolución de diversos tipos de problemas relacionados con caras humanas, tales como la detección, localización de alguna de sus partes, reconocimiento e interpretación del gesto, determinación de la situación espacial o generación de vistas sintéticas a partir de una cara. En la mayoría de los casos, el problema previo al análisis es la detección de la cara, o en el caso de secuencias, el seguimiento de la misma a través de las distintas imágenes. Sin una correcta detección, cualquier procesamiento posterior no sería posible.

Un ejemplo de aplicación de las técnicas de detección y seguimiento de caras es la animación de actores virtuales mediante captura de movimiento. En este caso, la frontera entre procesamiento y generación de imágenes se hace difusa y el resultado del análisis es la entrada directa al proceso de síntesis. Actualmente, el acercamiento usual al problema consiste en hacer uso de dispositivos físicos o marcadores claramente distinguibles, para obtener la posición 3D de ciertos puntos asociados al actor. Después, estas posiciones son usadas para ajustar el modelo de actor virtual.

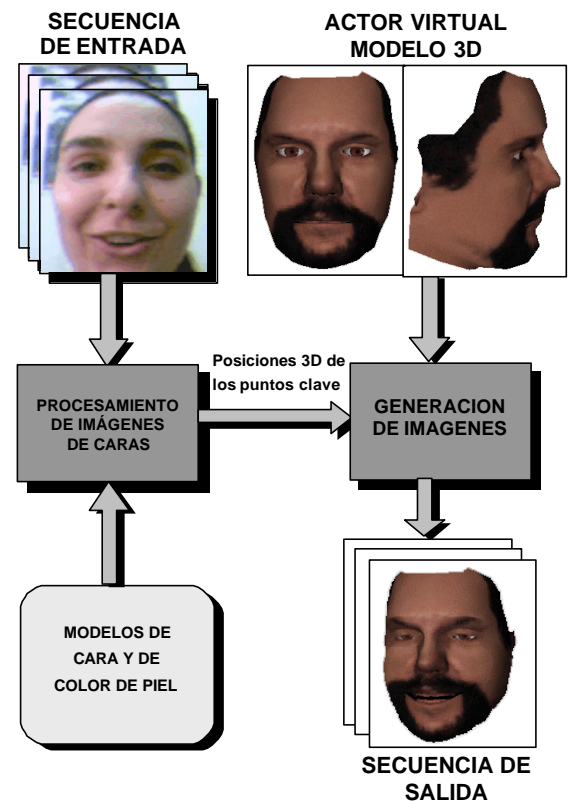


Fig. 1. Ejemplo de aplicación del procesamiento de caras: actores virtuales. El seguimiento de caras permite llevar a cabo la captura de movimiento sin marcadores artificiales. Modelo de actor virtual obtenido de <http://ligwww.epfl.ch/~fua/faces>

Como muestra la figura 1, el desarrollo de técnicas avanzadas de visión por computador permite eliminar completamente el uso de sensores o marcadores artificiales que, en la práctica, coaccionan la libertad de movimientos del actor humano. Más aun, estas técnicas tienen la capacidad de captar las sutilezas de los gestos, que de otro modo se perderían. Además, como se explica y desarrolla en [3], la creación de un modelo 3D explícito puede resultar innecesaria y la vista sintética puede conseguirse mediante la combinación lineal de vistas 2D.

El desarrollo presentado en este artículo se enmarca en una línea de trabajo sobre análisis de caras, fruto de la cual son las publicaciones [4-6], en las que se tratan diversos problemas relacionados con imágenes de caras humanas. En particular, el objetivo del trabajo aquí expuesto consiste en el seguimiento de caras y ciertos componentes faciales en secuencias de imágenes en color. En este artículo se extraen las aportaciones más novedosas realizadas, entre ellas la técnica denominada de búsqueda logarítmica en rejilla. En la sección 2 se repasan los espacios de color más extensivamente empleados para el modelado del color de piel humana. El proceso de seguimiento de caras haciendo uso de estos espacios, se desarrolla en la sección 3. En la sección 4, se presentan las pruebas realizadas y los resultados obtenidos y por último, en la sección 5, se resaltan las conclusiones alcanzadas y las posibles vías futuras de trabajo.

2. ESPACIOS DE COLOR PARA ANÁLISIS DE CARAS

Es un hecho bien conocido que la distribución de color de la piel humana tiene la propiedad de formar grupos compactos y reducidos en la mayoría de los espacios cromáticos, como comprueban, por ejemplo, Yang y Waibel en [7]. De hecho, existe una amplia variedad de trabajos en los que el color de la piel humana es usado como una característica básica para la detección, localización o seguimiento de caras, manos o piel humana en secuencias de vídeo [7, 8]. Las diferencias de color entre distintas personas son debidas a variaciones en la intensidad, más que a variaciones de matiz del color. En este punto describimos algunos de los espacios de color más frecuentemente usados en análisis de caras, que serán empleados en el proceso de seguimiento descrito en la Sección 3.

2.1. Espacios de Color Normalizados

Como ya hemos justificado, un buen espacio de representación de colores para detección de piel debería eliminar el efecto de la intensidad del color. Dos son los tipos de representaciones más usados para la descripción del color de piel: el espacio de colores cromáticos ó (r, g) normalizado [7,9], y los espacios basados en *Hue* y *Saturación* [6, 8, 10]. El espacio cromático es el resultado de normalizar los canales R y G , según la intensidad. Es decir, suponiendo que (R, G, B) es un color RGB de entrada, la salida sería $(r = R/(R+G+B), g = G/(R+G+B))$. Por otro lado, la transformación HSI puede ser obtenida según la fórmula de González y Woods [11]:

$$HUE(R, G, B) = \arcsin \frac{R - G + R - B}{2\sqrt{(R - G)^2 + (R - B)(G - B)}}, \quad (1)$$

$$SATURACION(R, G, B) = 1 - \frac{3 \cdot \min(R, G, B)}{R + G + B}. \quad (2)$$

Otros espacios parecidos son los HSV y HLS [11]. En la figura 2 se muestran dos ejemplos típicos de entrada para la localización

de caras, junto con los resultados de utilizar los dos tipos de representación antes definidos.

En la representación (r, g) la tonalidad de la piel puede ser modelada mediante una distribución de probabilidad conjunta de ambos canales, normalmente una gaussiana [7]. Ninguno de los canales aporta, a priori, más información que el otro respecto al color.

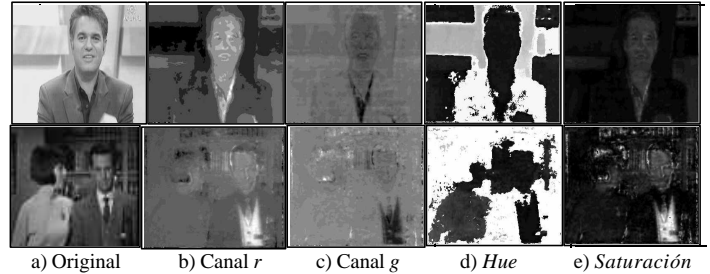


Fig. 2. Imágenes de entrada y transformación cromática. b), c) Espacio cromático o (r, g) normalizado; d), e) Espacio Hue-Saturación.

Esto no ocurre en la representación HS , donde el canal *Hue* resulta normalmente más discriminante para resolver el problema de clasificación piel/no-piel. Hemos comprobado que el color de piel se encuentra siempre en valores bajos de *Hue* (correspondientes al color negro en la figura 2d). Teniendo en cuenta la transformación RGB a HSI realizada, estos son valores próximos al rojo. El canal de *Saturación* aporta menos información, aunque es útil para distinguir el rojo del rosa-piel. No obstante, como se muestra en los ejemplos de la figura 2, ninguna de las anteriores representaciones es suficiente, por sí misma, para la detección de piel.

En los experimentos detallados en la sección 4, hemos decidido utilizar la representación HSI . Mientras que los canales H y S son útiles para la segmentación de píxeles de color de piel, el canal I resulta interesante para segmentar de los elementos faciales (ojos, cejas, nariz y boca), distinguibles por ser más oscuros que el resto de la cara.

2.2. El Espacio TSL - Tinta, Saturación, Luminosidad.

Recientemente ha sido propuesto un nuevo espacio de color perceptualmente plausible, denominado TSL [12], compuesto por los canales *Tinta*, *Saturación* y *Luminosidad*. Este espacio de color, diseñado específicamente para el modelado del color de piel, ha demostrado unos resultados muy buenos en el problema de discriminar regiones de piel humana, como se documenta en [12], a pesar de no tratarse de un modelo de color muy extendido. La transformación RGB a TSL está definida por las expresiones:

$$S(r', g') = \sqrt{\frac{9}{5}(r'^2 + g'^2)}, \quad (3)$$

$$T(r', g') = \begin{cases} \arctan(r'/g')/2\pi + 1/4 & \text{Si } g' > 0 \\ \arctan(r'/g')/2\pi + 3/4 & \text{Si } g' < 0 \\ 0 & \text{Si } g' = 0 \end{cases}. \quad (4)$$

Donde $r' = R/(R+G+B) - 1/3$ y $g' = G/(R+G+B) - 1/3$. En la figura 3 se muestra un ejemplo de transformación al espacio TSL de los ejemplos de la figura 2. Aunque, a simple vista, la mejora respecto a los anteriores espacios de color es inapreciable, la ganancia es importante en términos del rango de valores que toman los píxeles

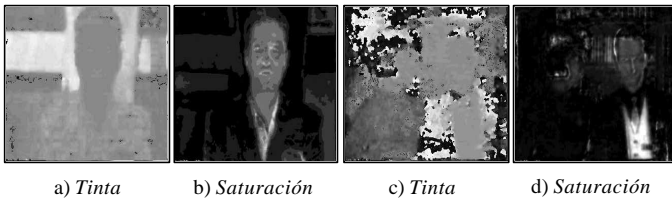


Fig. 3. Representación TSL de las imágenes de la figura 2.

de piel y, por lo tanto, en cuanto a discriminación piel/no-piel. Aunque en los experimentos documentados más adelante no se ha utilizado este modelo de color, hemos creído interesante incluirlo en esta comparativa de espacios de color, debido a los buenos resultados que ha mostrado en algunas pruebas preliminares realizadas.

3. EL PROCESO DE SEGUIMIENTO DE CARAS

En esta sección se describirán los aspectos más relevantes de nuestra propuesta para resolver los problemas de seguimiento y localización de caras o, en general, cualquier objeto distinguible por color: la técnica de búsqueda logarítmica en rejilla. Como veremos adelante, su uso permite llevar a cabo una reducción logarítmica del espacio de búsqueda de los objetos de interés, posibilitando una ejecución en tiempo real de los procesos de localización y seguimiento de caras humanas y componentes faciales. Una perspectiva global del sistema de visión donde se enmarca nuestra propuesta es mostrada en la figura 4.

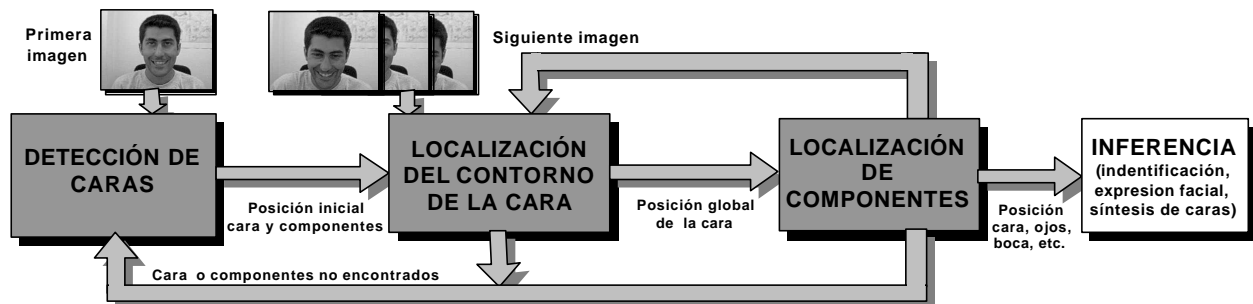


Fig. 4. Estructura global del sistema de seguimiento de caras. El resultado final del proceso es la posición de los componentes, con la que hacer algún tipo de inferencia.

En general, el seguimiento de objetos usando técnicas de visión por computador es un proceso iterativo encargado de calcular y actualizar cierta información de los objetos de interés a lo largo de una secuencia continua de imágenes. En nuestra propuesta introducimos cierta jerarquización en el proceso de seguimiento. Así, primero se localiza el contorno de la cara para posteriormente localizar los componentes faciales interiores a la misma. Ambos subproblemas hacen uso de la técnica de búsqueda en rejilla, que es mostrada en la figura 5 y será explicada en la subsección 3.2, como una manera de mejorar la eficiencia en los pasos de localización de regiones.

La inicialización del proceso de seguimiento viene dada idealmente por los resultados de un paso de detección de caras inicial. En el prototipo desarrollado esta inicialización fue llevada a cabo de forma supervisada. Por otro lado, los resultados del seguimiento serían la entrada para un módulo de inferencia, encargado de hacer uso de la información de posición obtenida en el seguimiento. Un ejemplo de una posible aplicación de

inferencia podría ser la animación de actores virtuales, comentada en la introducción. No obstante, debemos dejar claro que ambos problemas, detección e inferencia, caen fuera del alcance de este artículo.

3.1. Inicialización del Sistema

Es claro que el problema de detección de caras humanas tiene una complejidad comparable, cuando no superior, a la complejidad del problema de seguimiento. Dado que el interés del trabajo aquí presentado se centra en la parte de seguimiento y localización, se da por supuesta una detección inicial de la cara y un modelo de color de piel, que son recibidos por el prototipo a través de un fichero de configuración. Este fichero contiene el número de componentes a buscar (incluida la propia cara) y para cada uno de ellos su nombre, localización inicial (rectángulo envolvente del componente) y los umbrales (H, S, I) necesarios para su segmentación inicial. Una propuesta para la detección y localización automática de caras haciendo uso de color y textura puede consultarse en [5-6].

Con estos datos, el sistema lleva a cabo la detección de las regiones correspondientes a la cara y a cada uno de los componentes faciales en la primera imagen de la secuencia, mediante el proceso de umbralización explicado en la sección 2. De cada una de estas regiones se extraen una serie de parámetros iniciales que guiarán el resto del proceso de forma automática. Estos parámetros guía son los que describen la elipse que mejor se ajusta a cada una de las regiones detectadas: centro de la elipse, ángulo de inclinación, radio mayor y menor y área.

3.2. Localización del Contorno de la Cara

Haciendo uso de la posición actual de la cara en una imagen, el módulo de seguimiento procede a localizar la cara en la siguiente imagen de la secuencia de forma automática. Este proceso usa la técnica que hemos denominado de búsqueda logarítmica en rejilla, un método novedoso y eficiente de búsqueda de regiones que reduce la cantidad de píxeles de cada fotograma que es necesario explorar, consiguiendo así una mejora en el rendimiento global del sistema de localización y seguimiento. El esquema en pseudocódigo del algoritmo de búsqueda logarítmica en rejilla es presentado en la figura 5. Se considera que los datos a procesar están dados por la secuencia ($i_0, i_1, i_2, \dots, i_t, i_{t+1}, \dots$) de imágenes en color. El algoritmo describe el procesamiento realizado en un instante $t+1$ cualquiera.

En la figura 6 se muestra una interpretación gráfica del mismo algoritmo. El proceso consta básicamente de dos fases: primero se calcula el rectángulo mínimo que contiene a la cara, y

Algoritmo: Búsqueda Logarítmica en Rejilla

Entrada

$(x_1^t, y_1^t), (x_2^t, y_2^t)$: Esquinas del rectángulo contenedor del objeto en el instante t .
 i_{t+1} : Nueva imagen de la secuencia, en el instante $t+1$.
 $U = (u_1, \dots, u_n)$: Umbrales para la segmentación por color.
 r : Nivel de resolución requerido.

Salida

$(x_1^{t+1}, y_1^{t+1}), (x_2^{t+1}, y_2^{t+1})$: Esquinas del rectángulo contenedor en el instante $t+1$.
 $P^{t+1} = (p_1, p_2, \dots, p_n)$: Polígono correspondiente al contorno del objeto detectado.

1. Búsqueda del valor de x_1^{t+1} , el nuevo mínimo de x en el rectángulo contenedor.

1.1. Empezar la búsqueda en el centro del rectángulo contenedor en el paso anterior:

$$xx = (x_1^t + x_2^t)/2 ; d = (x_1^t - x_2^t)/2 .$$

1.2. Mientras $d > 1$ hacer.

1.2.1. Explorar el segmento: $(xx, y_1^t)-(xx, y_2^t)$ en la imagen i_{t+1} , usando los umbrales U .

1.2.2. Si hay una intersección significativa con el objeto, entonces: $d = d/2 ; xx = xx + d$.

1.2.3. Sino: $xx = xx - d$.

1.3. $x_1^{t+1} = xx$.

2. Búsqueda del valor de x_2^{t+1} , el nuevo máximo de x en el rectángulo contenedor.

2.1, 2.2 y 2.3. Igual que 1.1, 1.2 y 1.3, pero sustituyendo $(xx = xx + d)$ por $(xx = xx - d)$, y viceversa.

3. Búsqueda del valor de y_1^{t+1} , el nuevo mínimo de y en el rectángulo contenedor.

3.1, 3.2 y 3.3. Igual que 1.1, 1.2 y 1.3, pero sustituyendo x por y , y viceversa.

4. Búsqueda del valor de y_2^{t+1} , el nuevo máximo de y en el rectángulo contenedor.

4.1, 4.2 y 4.3. Igual que 3.1, 3.2 y 3.3, pero sustituyendo $(yy = yy + d)$ por $(yy = yy - d)$, y viceversa.

5. Localizar los puntos del contorno del objeto.

5.1. Para $yy = y_1^{t+1}$ hasta y_2^{t+1} , con incrementos de tamaño $(y_2^{t+1} - y_1^{t+1})/r$, hacer:

5.1.1. Explorar el segmento: $(x_1^{t+1}, yy) - (x_2^{t+1}, yy)$ en la imagen i_{t+1} , usando los umbrales U .

5.1.2. Almacenar los 2 puntos de intersección con el objeto como partes del contorno en P^{t+1} .

5.2. Para $xx = x_1^{t+1}$ hasta x_2^{t+1} , con incrementos de tamaño $(x_2^{t+1} - x_1^{t+1})/r$, hacer:

5.2.1. Explorar el segmento: $(xx, y_1^{t+1}) - (xx, y_2^{t+1})$ en la imagen i_{t+1} , usando los umbrales U .

5.2.2. Almacenar los 2 puntos de intersección con el objeto como partes del contorno en P^{t+1} .

Fig. 5. Estructura en pseudocódigo del algoritmo de búsqueda logarítmica en rejilla, aplicado para la detección de la cara y los componentes faciales.

posteriormente se obtiene el contorno de la misma con tanta precisión como se desee, especificada en el parámetro r . La primera parte es realizada en los pasos 1-4 del algoritmo y es representada en las figuras 6a)-c). La segunda parte la forma el paso 5 y es mostrada en la figura 6d). En general, puesto que es de esperar que la nueva posición de la cara sea cercana a la antigua (trabajando con frecuencias por encima de las 15 imágenes por segundo), la localización de la cara en la imagen nueva siempre parte de la posición en la imagen anterior.

Una vez obtenida la región correspondiente a la cara, se calculan sus ejes de simetría, figura 6e), que permitirán al paso de localización de componentes faciales comprobar ciertas restricciones estructurales. Este

paso es un simple ajuste de una distribución gaussiana, y no forma parte del algoritmo de búsqueda en rejilla.

Analizando la complejidad del algoritmo, encontramos que la mayor parte de la carga computacional es originada por la exploración de segmentos que aparece en los pasos 1.2.1, 2.2.1, 3.2.1, 4.2.1, 5.1.1 y 5.2.1. Esta exploración consiste básicamente

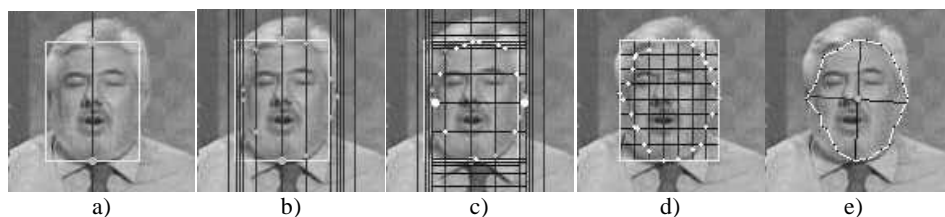


Fig. 6. Descripción del proceso de localización de caras mediante el método de búsqueda logarítmica en rejilla a)-d) y extracción de los ejes de simetría e).

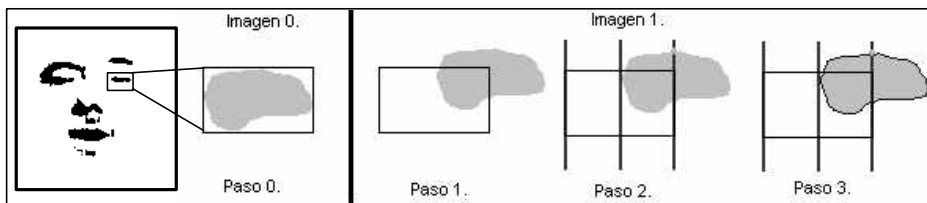


Fig. 7. Localización de los componentes faciales. A la izquierda, la cara encontrada en la figura 6e), umbralizada en intensidad. Inicialmente (imagen 0) el rectángulo envuelve al objeto ojo. En la siguiente imagen de la secuencia, el objeto se desplaza. Se exploran 3 segmentos verticales, obteniendo los puntos de colisión (paso 2). Finalmente, a partir del punto de colisión se busca el contorno (paso 3).

en buscar los trozos del segmento que son de color de piel, de acuerdo con el modelo de color U . Suponiendo que el rectángulo contenedor de la cara en el instante t es de tamaño $n \times n$, cada segmento explorado tiene n píxeles de tamaño y es fácil ver que el algoritmo de búsqueda en rejilla tiene un orden de complejidad de $O(r \cdot n + n \cdot \log_2 n)$. En concreto, el número promedio de píxeles explorados será: $2r \cdot n + 4n \cdot \log_2 n$. Comparado con una exploración exhaustiva, que requeriría un $O(n^2)$, vemos que la mejora depende del parámetro de precisión r , que indica el número de segmentos de la rejilla en la figura 6d). Un valor menor implica menos tiempo de ejecución, aunque una peor resolución de la región encontrada. En cualquier caso, la máxima eficiencia que se puede conseguir es de $O(n \cdot \log_2 n)$.

3.3. Localización de los Componentes Faciales

El proceso de búsqueda de los componentes faciales de interés (cara, ojos y nariz) es mostrado en la figura 7. Como ya se ha justificado, estos elementos destacan por ser más oscuros que el resto de la cara, por lo que podemos hacer una segmentación por intensidad. En particular, la región de búsqueda de los componentes es determinada mediante un proceso similar a la búsqueda de la cara. En este caso, la búsqueda en rejilla se realiza solamente sobre tres segmentos verticales, colocados según la posición anterior del componente.

Es posible aprovechar la información calculada sobre el movimiento global de la cara, para buscar los componentes faciales de interés. Así, se podrían estimar sus nuevas posiciones suponiendo que éstos se han movido solidariamente con la cara y sin variar su posición respecto a ella. No obstante, esta suposición no siempre es válida ya que un cambio en el gesto de la persona, la oclusión de alguno de los componentes, o más aun, una rotación de la cara respecto al plano de la imagen, harán que el uso de esta información dirija la búsqueda de manera errónea. Sin embargo, dado que el sistema trabaja en tiempo real, es factible asumir que entre dos imágenes consecutivas de la secuencia no habrá cambios bruscos.

4. PRUEBAS Y RESULTADOS

La técnica de localización y seguimiento expuesta en este artículo ha sido implementada en un prototipo utilizando el lenguaje de programación Delphi 5.0, de Borland. Una imagen del aspecto de este prototipo es mostrada en la figura 8. Para realizar las pruebas se han empleado 13 secuencias de vídeo en formato AVI y resolución 320x240, seis de las cuales han sido adquiridas utilizando una cámara de videoconferencia QuickCam de Logitech de bajo coste y las siete restantes han sido grabadas de la televisión usando una tarjeta sintonizadora y capturadora ATI All-In-Wonder 128 PRO. En las secuencias seleccionadas se encuentran tanto ejemplos sencillos, como casos más complejos

en los que aparecen oclusiones de algunos componentes faciales, rotaciones de la cara, cambios de iluminación, etc.

Las pruebas de tiempo han sido realizadas en un PC de gama media, con procesador AMD K6 a 350 Mhz, con 32 Mbytes de memoria RAM. Hemos llevado a cabo una comparación entre el tiempo obtenido por un algoritmo de seguimiento realizando una búsqueda exhaustiva de la imagen y el algoritmo de búsqueda logarítmica en rejilla,

utilizando distintos valores de precisión (parámetro r en el algoritmo de la figura 5). Los resultados obtenidos son mostrados en la tabla 1.

Algoritmo	Búsqueda exhaustiva	Búsqueda Logarítmica en Rejilla			
		$r = 4$	$r = 8$	$r = 16$	$r = 32$
Tiempo por imagen (ms)	36,41	14,51	16,42	20,23	27,85
Tiempo normalizado	100%	39,6%	45,1%	55,5%	76,5%
Frecuencia (imágenes/s)	27,5	68,9	60,9	49,4	35,9

Tabla 1. Tiempos del algoritmo de detección para distintos tamaños de rejilla

Para resolución $r = 32$ segmentos, la ganancia resulta poco significativa. Sin embargo, para resoluciones menores la mejora es considerable, llegando la reducción hasta el 40% para $r = 4$. Como dato más interesante, vemos que para $r = 8$ la velocidad media conseguida es de 16.4 milisegundos por imagen, sin contar los tiempos de entrada/salida. Esto permite una frecuencia de procesamiento máximo algo superior a las 60 imágenes/segundo, con lo cual el objetivo de tiempo real es claramente alcanzado.

En la figura 9 se muestran algunos ejemplos de segmentación con tres de las secuencias de prueba. En las 13 secuencias usadas, la posición de la cara es obtenida con una precisión similar a los resultados mostrados en la figura 9. En general, el prototipo diseñado ha mostrado una buena robustez frente a condiciones de iluminación, expresión facial y rotación de la cara. En la mayoría de los casos donde el cuello es visible, este es incluido en la región de cara, como se puede ver en la figura 9a). Por otro lado,



Fig. 8. Ventana de trabajo del prototipo implementado.

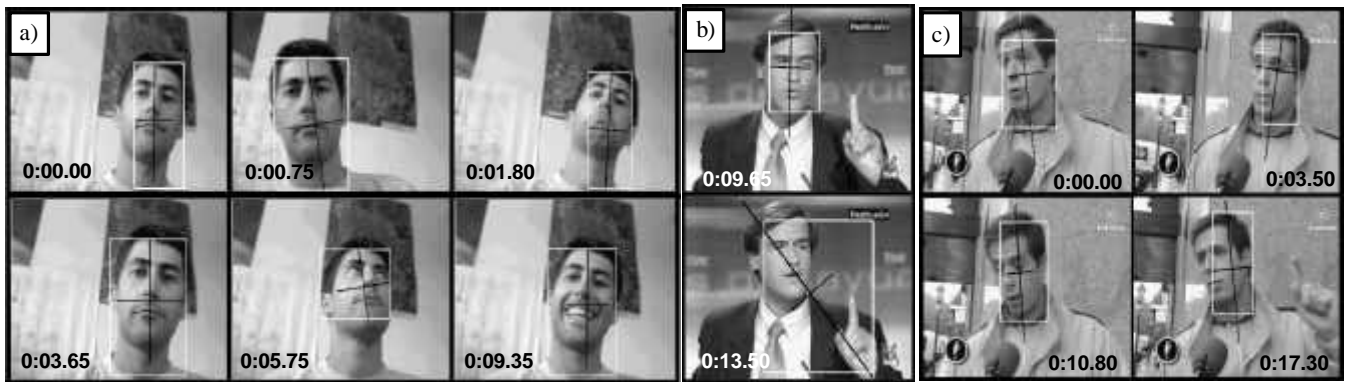


Fig. 9. Ejemplos de seguimiento de la cara para tres secuencias. Por claridad, se muestra sólo el rectángulo contenedor y el eje principal de la cara.

en algunos ejemplos donde aparece una mano, como el de la figura 9b), esta es introducida erróneamente en el rectángulo contenedor de la cara. No se ha realizado ninguna medida sobre la exactitud y precisión del proceso de seguimiento, por lo que no nos es posible comparar con otros métodos.

5. CONCLUSIONES

En este artículo hemos presentado una solución a los problemas de localización y seguimiento de caras humanas en secuencias de imágenes en color. La técnica usada se basa fundamentalmente en segmentación por color en el espacio *HSI*. Los resultados preliminares han sido muy positivos, demostrando que un acercamiento relativamente simple es viable. Para evitar la elevada carga computacional que una transformación *HSI* completa requeriría, hemos propuesto una técnica de reducción del espacio de búsqueda mediante una rejilla logarítmica. De esta forma, se permite llevar a cabo el seguimiento de objetos en una secuencia en tiempo real.

Está claro, sin embargo, que todavía queda mucho trabajo por desarrollar para conseguir un proceso de seguimiento robusto y preciso en situaciones no controladas. Por ejemplo, resultaría interesante que el modelo de color no sea fijo, sino que se adapte según la evolución de la iluminación en la escena, al estilo de las técnicas propuestas en [8]. Por otro lado, para hacer más robusto el seguimiento, podría hacerse uso de las técnicas de filtrado de Kalman, de manera que una pérdida temporal de la cara o un error en la posición, no provocaran la pérdida de la consistencia del sistema. Finalmente, creemos que es necesario profundizar en otro tipo de criterios para verificar que el objeto encontrado es realmente el deseado. En el caso de caras, un posible criterio podría ser la existencia de simetría a lo largo del eje vertical. De hecho, nuestro prototipo incluye la posibilidad de comprobar esta simetría, aunque la sobrecarga que introduce en el tiempo de ejecución no permite su uso en tiempo real.

6. AGRADECIMIENTOS

Agradecemos a los doctores Alberto Ruiz García e Isidro Verdú Conesa, por su colaboración e interés en el trabajo aquí desarrollado.

Este trabajo ha sido financiado por el proyecto MCYT, del Ministerio Español de Ciencia y Tecnología, número DPI-2001-0469-C03-01.

7. REFERENCIAS

- [1] A. Pujol, J.L. Alba, J.J. Villanueva, "Supervised Hausdorff-based Measures for Face Recognition", in the Proc. SNRFAT' 2001, Castellón, 16-18 May 2001
- [2] A. Pentland, "Looking at People: Sensing for Ubiquitous and Wearable Computing", IEEE Transactions on PAMI, Vol. 22, No. 1, January 2000
- [3] S. Ullman, "High-level Vision: Object Recognition and Visual Cognition", The MIT Press, 1996
- [4] G. García Mateos, C. Vicente Chicote, "A New Model and Process Architecture for Facial Expression Recognition", LNCS 1876, Springer, in Proc. of Joint IAPR Int. Workshops S+SSPR'2000, Alicante, Spain, September 2000
- [5] G. García Mateos, C. Vicente Chicote, "A Unified Approach to Face Detection, Segmentation and Location Using HIT Maps", in the Proc. of the IX Spanish Symposium SNRFAT' 2001, Benicasim, Castellón, Spain, 16-18 May 2001
- [6] G. García Mateos, C. Vicente Chicote, "Face Detection on Still Images Using HIT Maps", in the LNCS, Springer, 3rd Int. Conf. on Audio- and Video-based Biometric Person Authentication, Halmstad, Sweden, June 6-8, 2001
- [7] J. Yang, A. Waibel, "A Real-Time Face Tracker", in Proceedings of the Third IEEE Workshop on Applications of Computer Vision, pp. 142-147, 1996
- [8] L. Sigal et al., "Estimation and Prediction of Evolving Color Distributions for Skin Segmentation Under Varying Illumination", IEEE Conference on Computer Vision and Pattern Recognition, Barcelona, 2000
- [9] N. Oliver, A. P. Pentland, F. Berard, "Lafter: Lips and Face Real Time Tracker", IEEE Conf. on Computer Vision and Pattern Recognition, 1997
- [10] K. Sobottka, I. Pitas, "Looking for Faces and Facial Features in Color Images", PRIA: Advances in Mathematical Theory and Applications, vol. 7, no. 1, 1997
- [11] A. Ford, "Computer Graphics Colour Spaces", http://www.zenger.informatik.tu-muenchen.de/lehre/vorlesungen/graphik/info/csc/COL_20.htm
- [12] J.C. Terrillon, S. Akamatsu, "Comparative Performance of Different Chrominance Spaces for Color Segmentation and Detection of Human Faces in Complex Scene", Images Vision Interface'99, Trois-Rivieres, Canada, pp.1821
- [13] M.H. Yang, N. Ahuja, D. Kriegman, "A Survey on Face Detection Methods", IEEE Trans. on Pattern Analysis and Machine Intelligence, to appear in 2002