



Segmentación de rostro por color de la piel aplicado a detección de somnolencia en el conductor

Irma Josefina García Enríquez, Mariana Natalia Ibarra Bonilla, Juan Manuel Ramírez Cortés

Coordinación de Electrónica, Instituto Nacional de Astrofísica, Óptica y Electrónica
Tonantzintla, Puebla, México.

Tel/fax: (222) 234-5678, email: irmagarcia@inaoep.mx

Abstract— En este trabajo, se presenta un algoritmo para la segmentación de las regiones de piel en imágenes provenientes de una cámara web. Para la elaboración de este algoritmo, se realizó un estudio y análisis de los diferentes espacios de color. Asimismo, se propone la unión de dos planos de color el Cr del espacio YCrCb y el plano U' del espacio LUX. La finalidad es aplicar esta segmentación de piel para incorporarla a un sistema para la detección y monitoreo del estado de los ojos de un conductor. En dicho sistema, se utiliza como sensor una cámara web. Si el sistema detecta que el conductor se está durmiendo (periodos prolongados con los ojos cerrados), el sistema emite una señal auditiva de alerta.

Keywords-Segmentación de piel, Detección de ojos, Detección de somnolencia, Visión por computadora.

I. INTRODUCCIÓN

De alguna u otra manera todos dependemos ya sea en mayor o menor grado del transporte en aspectos cotidianos de nuestra vida. No obstante, al haber un incremento constante del tráfico automovilístico también se generan serios problemas sociales. Uno de los más graves, los accidentes. Cientos de personas mueren cada año en las carreteras y muchas más resultan heridas. De acuerdo con estudios recientes, la fatiga y somnolencia del conductor es una de las principales causas de accidentes de carretera, solo precedido por exceso de velocidad y estados de alcoholismo [1]. Es decir, el error humano aparece como una de las principales causas que provocan dichos contratiempos. Por tal razón, existe la necesidad de implementar sistemas de seguridad que contribuyan a aumentar la confianza en la conducción y minimizar las probabilidades de que ocurra un accidente.

Algunos fabricantes de automóviles están incorporando elementos de alta tecnología en sus modelos con la finalidad de aumentar la seguridad. Sin embargo, esto incrementa su costo y hacen que la unidad sea de difícil acceso, especialmente en nuestro país. Es decir, esta tecnología no siempre está al alcance de la mayoría de la población.

Esta es la principal motivación para el desarrollo del presente trabajo. Se han reportado diversos enfoques orientados a la solución de este problema. Algunos de estos

están basados en el análisis de señales EEG [2], la detección de la temperatura corporal, [3], y algunos signos físicos relacionados con el sueño, tales como cabeceo [4], bostezo [5], pestañeo prolongado [6,7] y combinación de estos [8,9], utilizando en muchos de los casos, técnicas de procesamiento de imágenes. La elección de la cámara web como sensor, se debe a que la cámara es un método no invasivo, a diferencia de otros métodos como los EEG por mencionar alguno. De esta manera se permite la total libertad del movimiento del conductor.

El presente documento está organizado de la siguiente manera. En la sección 2 se presenta el trabajo realizado para la detección de la piel. La sección 3 muestra la labor realizada para la ubicación de los ojos y el estado de éstos. Finalmente, en la sección 4 se exponen los resultados y en la 5 las conclusiones.

II. ELECCIÓN DEL ESPACIO DE COLOR.

La detección del color de la piel es una señal de gran utilidad para diversas aplicaciones (rastreo o detección del de personas para sistemas de seguridad, son un ejemplo). Los algoritmos existentes utilizan diversos espacios de color, tales como RGB [11], YCbCr [12], HSI [13], TSL [14], entre otros.

Detectar la piel de una imagen a través de un método computacional pudiera parecer una tarea trivial, sin embargo, implementar un método exitoso es una tarea difícil y un problema que aun no está resuelto. Una de las primeras decisiones que se deben de tomar para iniciar la implementación del método de detección de rostros, es la elección del espacio de color en el que se va a trabajar.

Para la aplicación específica presentada en este trabajo, se llevó a cabo un estudio preliminar sobre los espacios de color más adecuados, habiéndose realizado experimentación con los siguientes espacios de color: RGB, YCrCb y LUX [15]. A continuación se da un breve resumen de dichos espacios y sus características principales.

A. RGB

Es el espacio de color que surge de la combinación de tres colores (rojo, verde y azul) y es uno de los más ampliamente utilizado para la transformación y el almacenamiento digital de los datos de la imagen. Los algoritmos que se utilizan en este trabajo se obtienen a partir de este espacio de color.

B. YCbCr

YCrCb es una señal no lineal codificada del RGB, usada por estudios europeos de televisión y para la compresión de imágenes. El color es representado por luma (que es la luminancia), calculada del RGB, construido como una suma ponderada de los valores del RGB, y dos valores diferentes de color Cr y Cb que se forman restando del luma los componentes del RGB rojos y azules.

$$Y = 0.299R + 0.587G + 0.114B \quad (1)$$

$$C_r = R - Y \quad (2)$$

$$C_b = B - Y \quad (3)$$

C. Lux

LUX: por Logarithmic hUe eXtension. Es un espacio de color no lineal basado en una transformación logarítmica del RGB.[5],[6].

$$L = (R + 1)^{0.3}(G + 1)^{0.5}(B + 1)^{0.1} - 1 \quad (4)$$

$$H = \begin{cases} \frac{M}{2} \left(\frac{R+1}{L+1} \right) & \text{si } R < L \\ M - \frac{M}{2} \left(\frac{L+1}{R+1} \right) & \text{en otro caso} \end{cases} \quad (5)$$

$$X = \begin{cases} \frac{M}{2} \left(\frac{B+1}{L+1} \right) & \text{si } B < L \\ M - \frac{M}{2} \left(\frac{L+1}{B+1} \right) & \text{en otro caso} \end{cases} \quad (6)$$

D. Análisis de los Espacios de color

De acuerdo con las referencias consultadas, uno de los espacios de color más utilizados para la detección de piel, es el YCbCr. Sin embargo, ante condiciones de iluminación muy fuertes, se presentan deficiencias. Como una alternativa interesante, se plantea utilizar el nuevo y poco conocido espacio de color LUX [15], dadas las características de separabilidad obtenidas en el plano U. Los espacios de color en los cuales se realizaron las pruebas pertinentes fueron el RGB, YCbCr, HSI y LUX. Las pruebas hechas consistieron en formar 4 imágenes de fondo en 4 colores diferentes: Rojo, Verde, Azul y Amarillo. Sobre cada imagen se colocó un rectángulo con una gama de color que se consideró empíricamente semejante a distintos tonos de piel y que

MATLAB proporciona bajo la denominación „copper“. Esta imagen compuesta se utilizó en pruebas de separabilidad en los espacios de color seleccionados, con los resultados indicados en la figura 1.

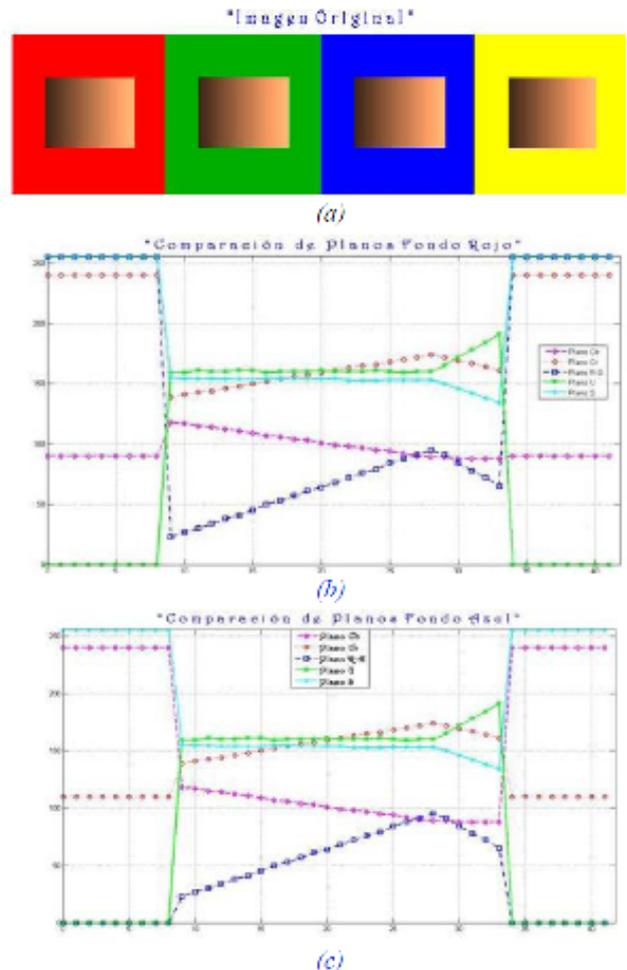


Fig. 1. (a) Imagen prueba sobre fondo rojo, azul, verde y amarillo. (b) Resultados fondo rojo, (c) fondo azul.

TABLA I
COMPARACIÓN DE LOS PLANOS DE COLORES BAJO DIFERENTES COLORES DE FONDO

Plano	Fondo Rojo		Fondo Verde		Fondo Azul		Fondo Amarillo	
	PC	PO	PC	PO	PC	PO	PC	PO
R-G	✓	✓	•	✓	•	✓	•	✓
Cr	✓	✓	✓	✓	✓	✓	×	•
Cb	•	×	✓	✓	✓	✓	✓	✓
S	✓	✓	✓	✓	✓	✓	✓	✓
U	✓	✓	✓	✓	✓	✓	✓	✓

✓ Buena segmentación PC = Piel clara
 × Mala Segmentación PO = Piel oscura
 • Segmentación Regular

Se realizaron pruebas aplicando el algoritmo a diferentes imágenes, incluyendo fotografías de la base de Datos VALID [16] y se analizaron sus valores de umbral en 3 diferentes planos: Cb y Cr (espacio YCbCr,) R-G (espacio RGB), U'' (espacio LUX) y S (espacio HSI). En la figura 2 se muestran los resultados obtenidos para algunas de las fotografías utilizadas.

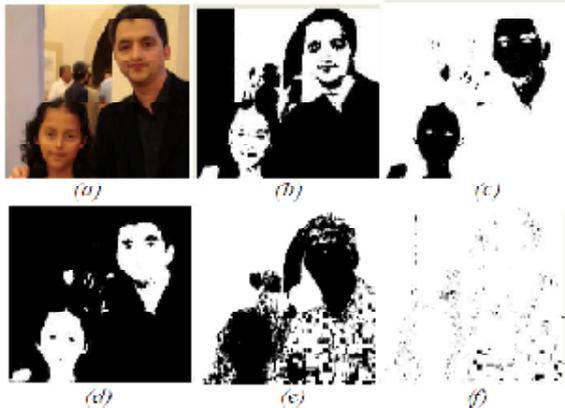


Figura 2. Pruebas en fotografía: (a) Imagen Original, (b) Plano Cr, (c) Plano Cb, (d) Plano R-G, (e) Plano U'' y (f) Plano S

En base a los experimentos, se observó que los planos Cb y Cr dan buenos resultados para la segmentación de las regiones de piel. El plano U y no se ve afectado por las variaciones de luz. Por lo tanto se ha concluido que con la combinación de estos planos se puede obtener un algoritmo más robusto a los cambios de iluminación.

I. SEGMENTACIÓN DE LAS REGIONES DE PIEL

Para realizar las pruebas con el plano U'', se desarrolló un algoritmo en MatLab para obtener el espacio LUX y de este tomar únicamente el plano U'' que es el que ofrece información de las regiones de piel. Se utilizaron para este fin las ecuaciones 2. Posteriormente, para el algoritmo del plano U'', se utilizó la ecuación siguiente para realizar la programación:

$$U' = \begin{cases} 256 \times \frac{G}{R} & \text{si } R > G \\ 255 & \text{en otro caso} \end{cases} \quad \text{Ec. 7}$$

En la figura 3, se puede observar el diagrama de flujo correspondiente a la programación del algoritmo U'. Posteriormente, en la figura 4, se aprecia el resultado de aplicar dicho algoritmo en imágenes de rostros de personas.

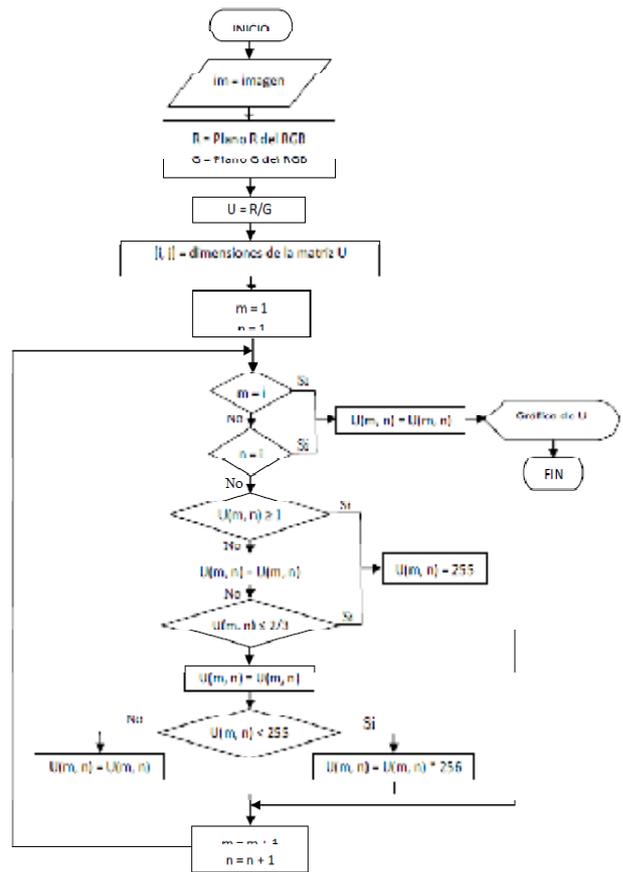


Figura 3. Diagrama de flujo del algoritmo del plano U'



Figura 4. Prueba del algoritmo U': (a) Imagen original, (b) Imagen en el plano U', (c) Imagen del plano U' binarizada

A. Unión de los planos Cr y U'

Se elabora el algoritmo de unión de los planos Cr y U'. Partiendo de la imagen original, se utiliza la ecuación 3 para obtener el plano U'' y el comando de Matlab para obtener el plano Cr. En seguida, se binarizan cada uno de estos planos de acuerdo al siguiente criterio:

$$I_{Cr}(x, y) = \begin{cases} 1, & \text{si } I(x, y) \in Th_{Cr} \\ 0 & \text{en otro caso} \end{cases} \quad (8)$$

$$I_{U'}(x, y) = \begin{cases} 1, & \text{si } I(x, y) \in Th_{U'} \\ 0 & \text{en otro caso} \end{cases} \quad (9)$$

Donde los umbrales son:

$$105 < Th_{Cr} < 140$$

$$0 < Th_{U'} < 255$$

Posteriormente se realiza la unión por medio del producto punto entre I_{Cr} e $I_{U'}$, es decir:

$$I_{skin}(x,y) = \begin{cases} 1, & \text{si } I_{Cr} \cap I_{U'} = 1 \\ 0, & \text{en otro caso} \end{cases} \quad (10)$$

El algoritmo se programó en Matlab. En las imágenes de la figura 5, se muestra la aplicación del algoritmo a algunas de las fotografías. Asimismo, el algoritmo se aplicó a las fotografías de la base de datos de VALID [22]. En la figura 6 se muestran los resultados.

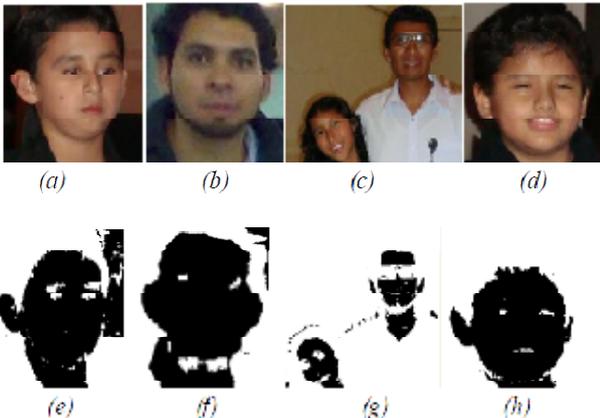


Figura 5. Aplicación del algoritmo propuesto: (a),(b),(c),(d) Imágenes originales, (e),(f),(g),(h) Imágenes después del algoritmo.

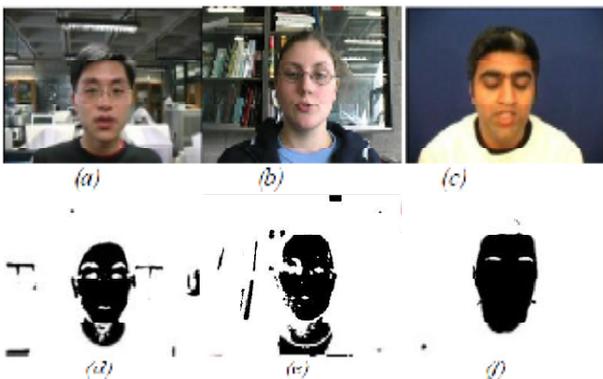


Figura 6. Fotografías a las que se les aplicó el algoritmo propuesto: (a),(c) Imágenes originales, (b),(d), Imágenes después del algoritmo.

Ya que el algoritmo ofrece buenos resultados para imágenes estáticas, se desarrolló un programa en MatLab que permitiese su aplicación en imágenes en movimiento provenientes de una cámara web montada en el auto. La figura 7 muestra una imagen típica.

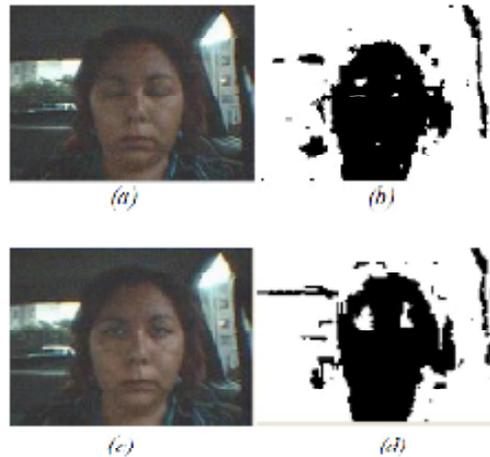


Figura 7. Imagen obtenida de la cámara a la que se le aplicó nuestro algoritmo: (a),(c) Imagen original, (b),(d) Imagen después del algoritmo.

III. DETECCIÓN DE CARA Y OJOS

Como se observa en la figura 7, en las imágenes procedentes de la cámara web, se distingue mayor ruido que en las imágenes estáticas. Es por lo anterior que el primer paso es eliminar dicho ruido. Se realiza un escaneo de la imagen y se etiquetan las regiones de piel, al mismo tiempo que se calcula su área. Finalmente la región que tiene mayor área se considera que es el rostro conservándose y eliminando lo demás. (Fig.8c).

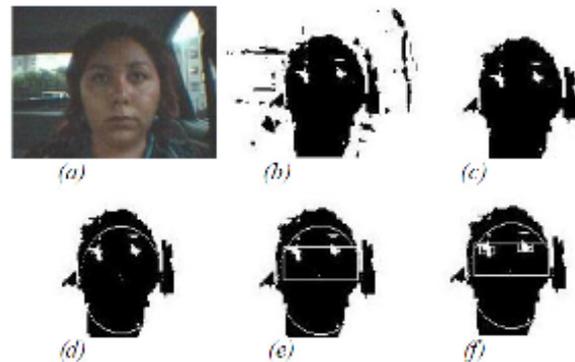


Figura 8. Secuencia del procesamiento de la imagen. (a) Imagen Original, (b) Segmentación de piel, (c) Eliminación de ruido, (d) Marcado de la elipse, (e) Región de los ojos, (f) Ubicación de los ojos

El rostro de una persona generalmente es similar a una elipse, por lo que se utiliza ésta forma para adaptarla a la cara del conductor. Para ello se deben de obtener los parámetros característicos de la elipse, es decir, el centro de la misma, su eje mayor y eje menor. El centro (X_c, Y_c)

se obtiene haciendo un promedio de todos los valores de la posición de los píxeles en la región de la piel:

$$Xc = \frac{1}{|A_s|} \sum_{(x,y) \in A_s} x \quad \text{Ec. 12}$$

$$Yc = \frac{1}{|A_s|} \sum_{(x,y) \in A_s} y \quad \text{Ec. 13}$$

donde A_s es la región de piel y $|A_s|$ su tamaño. Para calcular los tamaños del eje mayor y eje menor, lo primero es localizar las coordenadas del extremo superior del eje mayor. Posteriormente la distancia del centro a éste punto multiplicada por dos, es el tamaño del eje mayor.

Para determinar el tamaño del eje menor, basta con situar las coordenadas de los extremos derecho e izquierdo y la distancia entre estos dos es el tamaño del eje menor. Posteriormente, se dibuja la elipse alrededor de la cara utilizando la ecuación de la elipse. (Fig. 8d).

Para llevar a cabo la ubicación de los ojos, se marca un rectángulo que parte del centro con una altura igual a $0.6 \cdot (\text{longitud del eje mayor}/2)$. Fig. 8e. En seguida, para determinar el estado de los ojos, (cerrados o abiertos), el procedimiento es el siguiente.

Se realiza un escaneo solamente dentro del área de la región de los ojos con la finalidad de localizar el centro de los ojos. Un primer escaneo es del extremo izquierdo al centro y el segundo del centro al extremo derecho. Al igual que se hizo para encontrar el centro de la cara se promedian los valores de la posición de los píxeles en los ojos:

$$X_{left_eye} = \frac{1}{|R_{eye}|} \sum_{(x,y) \in R_{eye}} x \quad \text{Ec. 14}$$

$$Y_{left_eye} = \frac{1}{|R_{eye}|} \sum_{(x,y) \in R_{eye}} y \quad \text{Ec. 15}$$

Donde R_{eye} es el ojo, R_{eye} su tamaño y R_{eyes} el rectángulo de la región de piel del extremo izquierdo al centro. El procedimiento se repite para el ojo derecho.

Una vez que se tienen las coordenadas del centro de los ojos, se marca un rectángulo alrededor de ellos como se aprecia en la figura 8f. Finalmente para determinar si los ojos están abiertos o cerrados se calcula el área de cada rectángulo alrededor de los ojos y el área que ocupan los ojos. Si el área de alguno de los dos ojos es mayor al 50% del área del rectángulo de los ojos, se determina que el ojo está abierto en caso contrario, el ojo está cerrado:

$$\text{Edo ojos} \begin{cases} \text{Abiertos si } A_{ey_1} \cup A_{ey_2} \geq 50\% A_{r_{eye}} \\ \text{Cerrados si } A_{ey_1} \cup A_{ey_2} < 50\% A_{r_{eye}} \end{cases} \quad \text{Ec. 16}$$

Se programó un GUIDE en matlab para realizar las pruebas. En la figura 9 se muestra esta interfaz.



Figura 9. Interfaz de Matlab para las pruebas de ubicación y estado de los ojos.

Finalmente, en la figura 10, se puede apreciar al diagrama de flujo del sistema completo.

III. RESULTADOS

Los resultados de las pruebas hechas se muestran en las tablas I y II. El sistema de detección y rastreo de ojos está limitado a ambientes iluminados. El nivel de prueba que queda por realizar, es la integración de todas sus partes para la realización de pruebas simuladas, y pruebas reales sobre un prototipo de vehículo controlado por visión para movimiento autónomo, actualmente en desarrollo.

TABLA I
RESULTADOS DE LAS PRUEBAS DE DETECCIÓN DE ESTADO DE LOS OJOS

	OK	FP	FN
Individuo 1	9	1	0
Individuo 2	8	1	1
Individuo 3	10	0	0
Individuo 4	9	1	0
Individuo 5	10	0	0
Individuo 6	7	1	2
Individuo 7	9	0	1
Individuo 8	8	1	1
Individuo 9	8	2	0
Individuo 10	7	3	1
Total 100	85%	9%	6%

OK: Estado de los ojos correcto

FP: Falsos positivos (El sistema indica ojos cerrados falsamente)

FN: Falsos negativos (El sistema indica ojos abiertos falsamente)

V. AGRADECIMIENTOS

El primer autor agradece al CONACYT por el apoyo otorgado a través de la beca para estudios de Maestría con número 212442.

VI. REFERENCIAS

[1] J. McCall, M.Trivedi, "Driver behavior and situation aware brake assistance for intelligent vehicles", Proceedings of the IEEE, Vol. 95, No. 2, February 2007.

[2] B. Jap, S. Lal, P. Fischer, E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue", Expert Systems with Applications, 36 (2), pp. 2352-2359, March 2009.

[3] Rajinda Senaratne, David Hardy, Bill Vanderaa, and Saman Halgamuge. "Driver Fatigue Detection by Fusing Multiple Cues", 4th international Symposium on Neural Networks, pp. 801 - 809, Nanjing, China, June 3-7, 2007.

[4] J.Vandenberg, "Sleepiness and head movements", Industrial Health, Vol. 44, pp. 564-576, 2006.

[5] X. Fan, B. Yin, Y. Sun, "Yawning detection for monitoring driver fatigue" Sixth International Conference on Machine Learning and Cybernetics, Hong Kong, 19-22 August 2007.

[6] Jung-Ming Wang, et.al, "Detecting driver's eyes during driving", 18th IPPR Conference on Computer Vision, Graphics and Image Processing. Taipei, China, 2005.

[7] H.Veeraraghavan, N. Bird, S. Atev, N. Papanikolopoulos, "Classifiers for driver activity monitoring", Transportation Research Part C: Emerging Technologies, Vol.15, pp. 51-67, February 2007.

[8] A. Hussain, B. Bais, S. Samad, S. Hendi, "Novel data fusion approach for drowsiness detection", Information Technology Journal, Vol. 7, No. 1, pp. 48-55, 2008.

[9] Martin Golz, David Sommer, "Feature fusion for the detection of microsleeep events", Journal of VLSI Signal Processing, Vol. 49, No. 2, pp. 329-342, 2007.

[10] Liévin, M and Luthon, F., "Non linear color space and spatiotemporal MRF for hierarchical segmentation of face features in video". IEEE Transactions on Image Processing, Vol.13, No.1, January 2004

[11] Albiol, a., Torres, I., And delp, e. J.. "Optimum color spaces for skin detection". In Proceedings of the International Conference on Image Processing, vol. 1, 122-124. 2001

[12] Dadgostar Farhad, Sarrafzadeh Abdolhossein. "An adaptive real-time skin detector based on Hue thresholding: A comparison on two motion tracking methods", Pattern Recognition Letters, Vol. 27, No. 2, pp. 1342-1352, September, 2006.

[13] Vezhnevets Vladimir, et.al, "A Survey on pixel-based skin color detection techniques", International Conference on Computer Graphics and Vision, GraphyCon, Moscow, Russia, 2003.

[14] Rein-Lien Hsu, M Abdel-Mottaleb, AK Jain "Face detection in color images". IEEE Transactions on PAMI, Vol. 24, No. 5, pp. 696-706, 2002.

[15] N.A. Fox, B. A. O'Mullane, R.B. Reilly, "The realistic multi-modal VALID database and visual speaker identification comparison experiments," 5th International Conference of Audio and Video-Based Person Authentication, New York, July 2005.

[16] G. De Lannoy, A. De Decker, and M. Verleysen. "A supervised wavelet transform algorithm for R spike detection in noisy ECGs", Int. Joint Conf. in Biomed. Eng. Syst. and Tech., Madeira, Portugal, pp. 256-264, January, 2008.

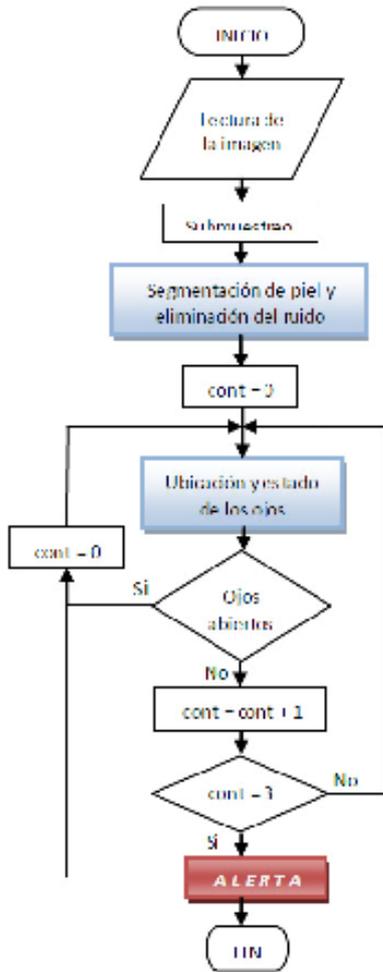


Figura 10. Diagrama de flujo del sistema

IV. CONCLUSIONES

Se presentó un algoritmo para la segmentación de regiones de piel en imágenes. Este algoritmo se aplicó a un sistema para la detección de la somnolencia del conductor usando como sensor una cámara que detecta el pestañeo prolongado a través del monitoreo de los ojos. El sistema funciona en tiempo real en ambientes iluminados, no así en lugares con muy poca o nula luz. Actualmente se está trabajando en la incorporación de un sensor más para lograr mayor robustez. Este sensor es un acelerómetro inalámbrico que se colocará en la cabeza del conductor para detectar el cabeceo. Además, se propone mejorar el algoritmo de rastreo de ojos para que funcione en ambientes de baja luminosidad, o bien, dejar la opción para que en situaciones de oscuridad total, funcione solo el acelerómetro para la detección únicamente del cabeceo del conductor. En este trabajo se presenta una de las muchas aplicaciones que puede darse a la segmentación de piel.