

Aproximación de un algoritmo clasificador de edad basado en el análisis de arrugas

[Algorithm approach of a classifier based on aging wrinkle analysis]

Luis Enrique Colmenares Guillen¹⁻², Antonio Esteban Martínez Cruz¹, Francisco Javier Albores Velasco², and Maria Concepcion Perez de Celis Herrero¹

¹Facultad de Ciencias de la Computación,
Benemérita Universidad Autónoma de Puebla,
Puebla, Puebla, México

²Facultad de Ingeniería y Tecnología,
Universidad Autónoma de Tlaxcala,
Tlaxcala, Tlaxcala, México

Copyright © 2016 ISSR Journals. This is an open access article distributed under the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT: In the present investigation, the performance of a classifier based on facial analysis and detection of wrinkles is analyzed. The experimental part has been divided in two phases. The first, analyzed the proposed features, and the second, evaluated its performance on a database of faces. During the first phase, a statistical analysis of the sets analyzed has been carried out, and therefore, the correlation between the characteristics proposed for the project, have been studied and discussed. In the second phase, several tests -under different conditions- have been performed in order to determine the capability and efficiency of the system in relation to recognition tasks. In particular, the optimal size of the training set, and the performance of various measures of similarity of the two sets of proposed characteristics have been studied, also the subsets of discriminating characteristics have been evaluated.

KEYWORDS: processing, anthropometry, filter, pattern, face.

1 INTRODUCCIÓN

Dentro de la historia humana, los rostros han desempeñado una importante fuente de información en las comunicaciones sociales, ya sea que entreguen información inmediata con cierta carga emocional, o en un sentido más personalizado, con información propia sobre el individuo. Específicamente, además de expresar emociones, el rostro puede comunicar el género, etnia de origen, edad e incluso a veces experiencias de vida [1]. Es así, que el rostro, representa una invaluable fuente de información y punto común para establecer una comunicación en base a un contexto determinado, entregando una referencia entre los individuos en cuanto a su situación y cómo deben reaccionar frente a su interlocutor.

En este enfoque, cuando se tiene una persona de edad avanzada se sabe de antemano, que sus reacciones y necesidades serán diametralmente distintas, a las que podría tener un niño.

Un reto que se presenta al realizar un clasificador de edad, es que la edad de una persona resulta difícil de predecir exactamente, porque los cambios faciales son lentos y estos cambios de apariencia son dependientes de las personas. Las estructuras óseas y las arrugas, son factores que pueden ayudar a la solución de este problema.

1.1 CONCEPTO DE ARRUGAS Y CLASIFICACIÓN

Por arruga se entiende, de forma general, toda aquella depresión de la superficie de la piel que interrumpe la uniformidad de la misma y que es producida como consecuencia, principalmente, del proceso del envejecimiento.

1.1.1 CLASIFICACIÓN POR CAUSA

- Arrugas por foto envejecimiento: ligadas directamente a la acción de los rayos ultravioleta del sol.
- Arrugas producidas por el envejecimiento ligado a la edad: son aquellas que aparecen principalmente en las mejillas, el cuello o los párpados. Se originan por la pérdida de elasticidad de las estructuras de sostén de la cara.
- Arrugas de expresión: Son profundas y producidas por movimientos musculares repetitivos. Se localizan principalmente en la frente, entre las cejas, en los pliegues naso labiales y en los bordes de la boca.
- Surcos: Arrugas profundas o pliegues, permanentes, vinculados o no, al movimiento de los músculos del rostro.

1.1.2 CLASIFICACIÓN POR PROFUNDIDAD

- Tipo I: líneas muy finas y someras que no requieren tratamiento.
- Tipo II: líneas finas y arrugas muy suaves que necesitan para su corrección productos inyectables temporales o semipermanentes.
- Tipo III y IV: arrugas moderadas y profundas que se pueden tratar con inyectables temporales y semipermanentes pero que necesitan de sesiones de mantenimiento periódicas. Lo ideal para este grupo de arrugas es el tratamiento con láser.

El mejor escenario en esta área de estimación, es la colección de imágenes de las mismas personas con diferentes edades y de diferentes grupos, aunque, esto es muy difícil de obtener. Las técnicas de extracción de arrugas son recientes y están en evolución.

2 TRABAJOS RELACIONADOS

Los humanos regularmente utilizan los rostros para reconocer individuos, dada su capacidad discriminativa, y los avances en las capacidades de computación en las últimas décadas, ahora permiten reconocimientos similares en forma automática.

La mayoría de las clasificaciones automáticas de la edad a través de imágenes combinan dos componentes: una representación de la imagen y un proceso de estimación de la edad, luego la diferencia se realiza en el algoritmo utilizado para conseguir dicho objetivo.

2.1 MODELOS PARA LA REALIZACIÓN DE CLASIFICADORES

Hay cinco diferentes modelos que son comúnmente usados hasta la fecha para la clasificación de la edad en un rostro humano:

1. Modelos antropométricos, su objetivo es obtener una aproximación de la edad e implementar simulaciones para representar los cambios de la apariencia con el paso del tiempo [2].
2. Modelos activos de apariencia, son ampliamente utilizados como una característica para la estimación de edad.
3. Los colectores de edad, donde la subyacente estructura de la edad facial se proyecten características de menor dimensión.
4. Modelos basados en características que describen la edad del rostro con un conjunto de texturas características discriminantes.
5. Modelos de simulación de edad.

2.2 MÉTODOS PARA LA REALIZACIÓN DE CLASIFICADORES

2.2.1 TRANSFORMADA WAVELETS LINEAL

La transformada wavelets lineal puede ser utilizada para clasificar grupos de personas jóvenes y adulta con edades desde los 16 a los 85 años, identificando cambios topológicos de la piel facial. Este método, se basa en la premisa, que la piel de la cara tiende a cambiar drásticamente con el paso del tiempo [3]. Las wavelets poseen la ventaja de entregar un análisis rápido

sin tener que utilizar máscaras más grandes que las de 5x5 píxeles en varias orientaciones. En el trabajo [4], se obtienen resultados muy precisos y que son indistintos de la pose o la rotación de la cabeza de la imagen analizada. El método desarrollado alcanza una tasa de aciertos del 90.8% en la base FG-NET, para 5 grupos de edad comprendidos entre los 16 a los 75 años [4].

2.2.2 MAPAS BINARIOS DE CAMPOS RECEPTIVOS GAUSSIANOS

Los mapas binarios de campos receptivos gaussianos utilizan una pirámide gaussiana de escala invariante de media octava. La representación de la imagen se construye usando operadores gaussianos derivativos. Se asume que estos mapas son una representación robusta de la imagen que codifica cambios faciales a través del proceso de envejecimiento [5]. Una representación tensorial se usa para codificar los cambios de forma del proceso de envejecimiento conservando la estructura tridimensional del espacio. Se utiliza Principal Component Analysis (PCA) multilinear para reducir la dimensionalidad del espacio tensorial. Y mediante Relevance Vector Machine (RVM) se realiza el método de regresión para estimar la edad. Éste método en particular entrega una aproximación de la edad del individuo con un error medio de 5.16 años [5].

2.2.3 MODELO DE RANKING

El envejecimiento puede ser visto como un problema de series de tiempo [6]. Obviamente, el incremento de edad es monótono con el paso del tiempo, además las características importantes pueden ser útiles para distinguir la edad en diferentes estados, con una determinada persona. En este trabajo, se utiliza un método denominado RankBoost para la extracción de características, que se basan en la relación ordinal entre pares de datos de la imagen. Con ello, se puede organizar fácilmente los datos con tal de obtener la información ordinal acorde a la edad de cada sujeto. Se obtiene una tasa de aciertos de un 91% considerando 3 grupos etnarios [6].

2.3 FILTRO CANNY

Uno de los métodos más importantes, para realizar una detección global de bordes sobre una imagen es el conocido, método de Canny [8]. Esta técnica, que se caracteriza por ser optimizada para la detección de bordes diferenciales, consta de tres etapas principales: filtrado, decisión inicial, e histéresis.

En 1986, Canny propuso un método para la detección de bordes, el cual se basa en tres criterios, estos son:

- Un criterio de detección expresa el hecho de evitar la eliminación de bordes importantes y no suministrar falsos bordes.
- El criterio de localización, establece que la distancia entre la posición real y la localizada del borde se debe minimizar.
- El criterio de una respuesta que integre las respuestas múltiples correspondientes a un único borde.

Uno de los métodos relacionados con la detección de bordes es el uso de la primera derivada, por que tiene el valor de cero en todas las regiones donde no varía la intensidad y tiene un valor constante en toda la transición de intensidad. Por tanto, un cambio de intensidad se manifiesta como una alteración brusca en la primera derivada, característica que es usada para detectar un borde, y además, representa el algoritmo de Canny. El algoritmo de Canny consiste en tres grandes pasos:

- Obtención del gradiente: en este paso se calcula la magnitud y orientación del vector gradiente en cada pixel.
- Supresión no máxima: en este paso se logra el adelgazamiento del ancho de los bordes, obtenidos con el gradiente, hasta lograr bordes de un pixel de ancho.
- Histéresis de umbral: en este paso se aplica una función de histéresis basada en dos umbrales; con este proceso se pretende reducir la posibilidad de aparición de contornos falsos.

3 PROPUESTA

3.1 HERRAMIENTAS

Una de las restricciones fundamentales de este trabajo es que funciona en tiempo real. Por este motivo, se desarrollara un software de procesamiento de las imágenes, el cual será íntegramente programado en el lenguaje C++, utilizando una

biblioteca denominada OpenCV (Open Computational Vision) para el filtrado de la imagen y el proceso de entrenar el clasificador.

3.2 SISTEMA DESARROLLADO

Este proyecto se centra en la aparición de arrugas faciales de individuos, una característica inevitable. El sistema desarrollado debe de tener la capacidad de pre-procesar las imágenes con el fin de obtener los rasgos faciales para su posterior análisis. Además, cada rasgo será tratado y analizado por separado utilizando distintas técnicas de tratamiento de imágenes con el objetivo de conseguir los detalles de dicho rasgo con la mayor precisión posible. No se extraen, sólo los puntos más significativos, sino que se intenta procesar el rasgo, en su totalidad.

Una vez obtenidas las imágenes de cada rasgo, se procede a la caracterización necesaria. La última etapa del sistema se ocupa de la identificación de los individuos incluidos en la base de datos, dando como resultado final la relación de densidad de los individuos. Una etapa de evaluación verifica la calidad del sistema, aquí se ocupara la misma base de datos para comparar resultados, nuevas imágenes y bases de imágenes pueden ocuparse como entrada para ser clasificadas.

3.3 BLOQUES DEL SISTEMA

El sistema consta de 4 grandes bloques mostrado en la Figura 1.

- Pre-procesado de imágenes.
- Extracción de características.
- Clasificación de resultados.
- Evaluación del sistema.

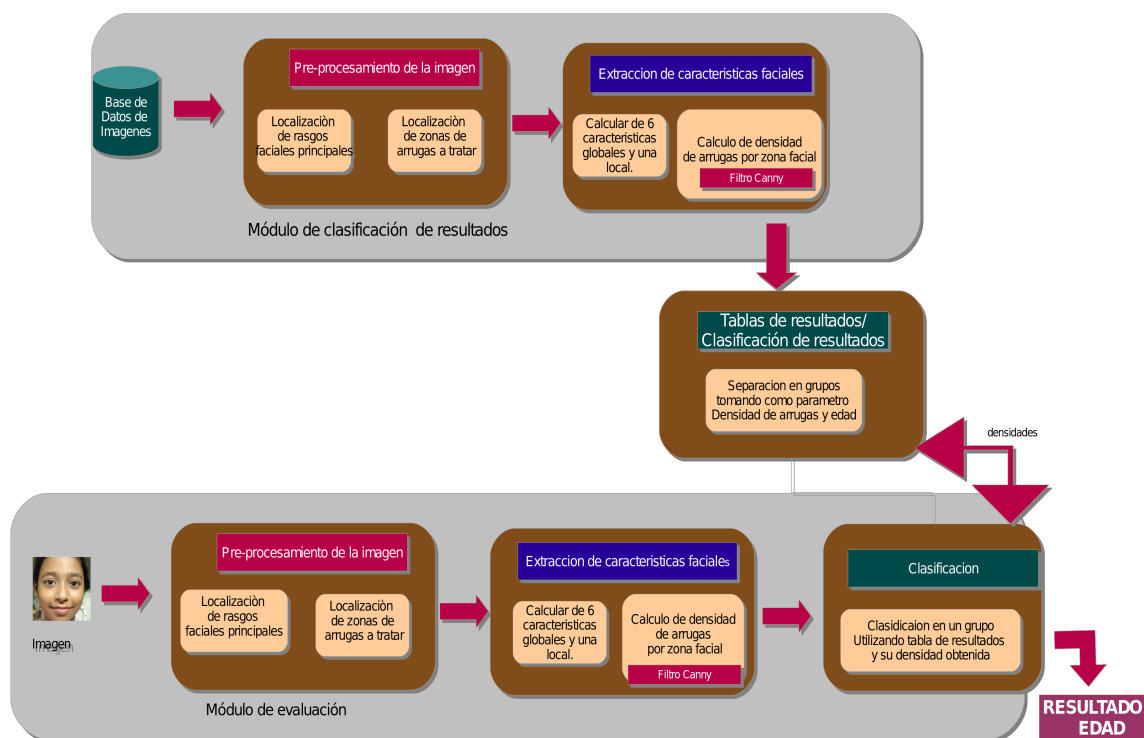


Fig. 1. Esquema general del sistema clasificador.

3.3.1 PRE-PROCESADO DE LAS IMÁGENES

Los dos primeros sub-bloques del esquema del sistema desarrollado forman parte del pre-procesado de las imágenes necesario para seguir avanzando con el análisis. La entrada al primer bloque son las imágenes de cada sujeto y los ficheros con las coordenadas de los puntos marcados. Los dos primeros bloques corresponden al pre-procesado común para todas las imágenes y se procede al pre-procesado individual de cada rasgo.

Diferentes regiones de la cara como el parejo entre los ojos (entrecejo), punta de la nariz y mentón son detectadas, mediante filtros como se muestra en la Figura 2.

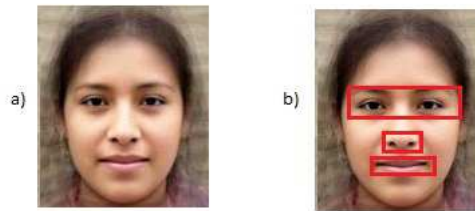


Fig. 2. Localización de rasgos principales. a) Rostro detectado. b) Rasgos detectados después de aplicaciones filtros.

Enseguida los rasgos de las arrugas son calculados usando la porción de la frente. La porción superior de las mejillas, las regiones de los párpados y las esquinas de los ojos. Una vez que la posición media del ojo derecho (x_r, y_r) y del ojo izquierdo (x_l, y_l) son determinados, la distancia entre los dos ojos es d y es calculado por la formula (1).

$$d = \sqrt{(x_r - x_l)^2 + (y_r - y_l)^2} \quad (1)$$

Basado en la estructura de la geometría de la Figura 3(a), se realizan los cálculos siguientes:

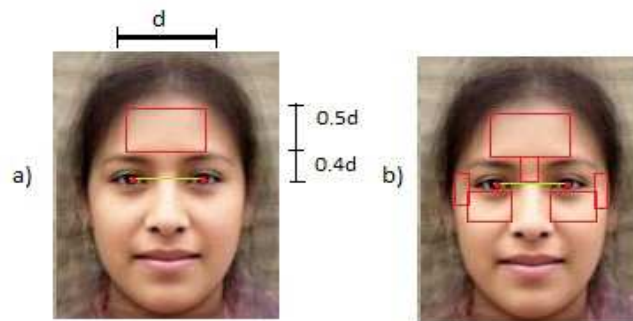


Fig. 3. Detección de rasgos de estudio. a) Imagen con distancia ocular calculada. b) Imagen con área de arrugas localizadas

- La distancia vertical entre los ojos y los globos oculares pueden estar colocada a $0.4 \cdot d$. Así se propone la porción de la frente sobre la región de los globos oculares con un rectángulo de tamaño $d \cdot 0.5d$.
- La región rectangular en el rostro que va debajo de los ojos hasta la punta de la de la nariz es considerada la porción de los párpados.
- La distancia entre las dos cejas es seleccionado con respecto a la línea principal.
- La región cerca a las esquinas de los ojos es seleccionada hasta la altura de la ceja.

3.3.2 EXTRACCIÓN DE CARACTERÍSTICAS FACIALES

Como se mencionó en el apartado anterior, es necesario un pre-procesado individual de cada rasgo para obtener todas las características morfológicas con más precisión, representado en la Figura 4.

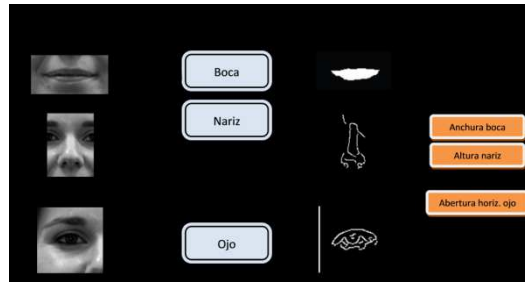


Fig. 4. Características morfológicas localizadas en la imagen de entrada.

Obtención de características globales: Las características globales representan las razones antropométricas, estas incluyen varias razones de distancias de todos los objetos fáciles, como el globo ocular derecho y el globo ocular izquierdo, la nariz, el mentón, los labios y la frente [7]. Las características locales principalmente usados, son las arrugas de las porciones de la frente, esquina de los ojos, los párpados y el entrecejo [9].

Obtención de características locales: Las características locales representan las zonas de las arrugas. Utilizando las zonas de las arrugas en la imagen pre-procesada, se obtiene la característica W (*wrinkles*) que es una parte importante de la propuesta del algoritmo. La detección de bordes es extensamente usada para la detección de discontinuidades en una imagen. La característica W se calcula, usando un filtro Canny, de la siguiente manera:

- La imagen de entrada es convertida a escala de grises. El filtro Canny es aplicado y resulta una imagen binaria con las arrugas. (Figura 5.a).
- Los píxeles blancos del área de arrugas proporciona información sobre las arrugas en la imagen facial.
- En la imagen binaria el valor 1 es usado para el pixel blanco y 0 para el pixel negro. Entonces la suma de los pixeles 0 en la zona de arrugas en una imagen binaria, que es proporcional a las arrugas que se encuentran en el rostro, esto se muestra en la Figura 5.b.

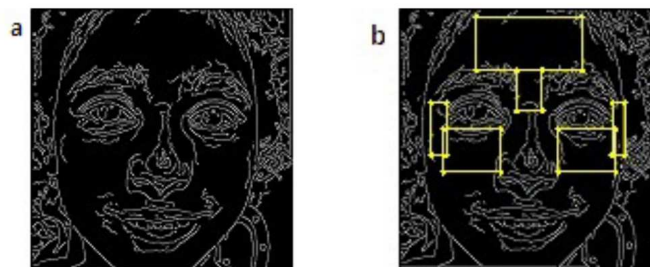


Fig. 5. Obtención de densidad de arrugas. a) Aplicación de filtro Canny. b) zonas de arrugas detectadas.

- Característica W = (suma de pixeles 0 en el area de la frente/número total de pixeles en el area de la frente) + (suma de pixeles 0 en el parpado izquierdo / número total de pixeles izquierdo) + (suma de pixeles 0 en el parpado derecho / número total de pixeles derecho) + (suma de pixeles 0 en el área de la esquina del ojo izquierdo/número de pixeles en el área de la esquina del ojo izquierdo) + (suma de pixeles 0 en área de la esquina del ojo derecho/número de pixeles en el área de la esquina del ojo derecho).

3.3.3 CLASIFICACIÓN DE RESULTADOS

En esta etapa se registrará la densidad de arrugas de cada individuo en un conjunto de datos que será dividido en tres rangos de edades: niños, edad media y edad avanzada. Para después calcular que rangos de densidades corresponden a cada grupo de edad.

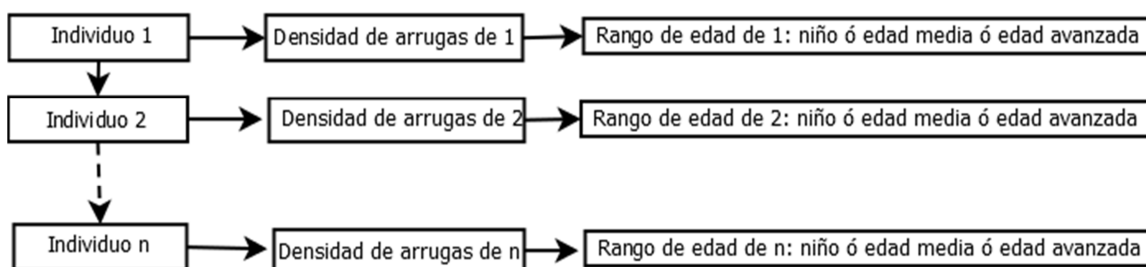


Fig. 6. Clasificación de individuos por densidad de arrugas y rango de edad.

El conjunto de datos que se observa en la Figura 6, se obtendrán los resultados en la etapa de clasificación, después se calculará un promedio de densidades para cada rango de edad ver la Figura 7.

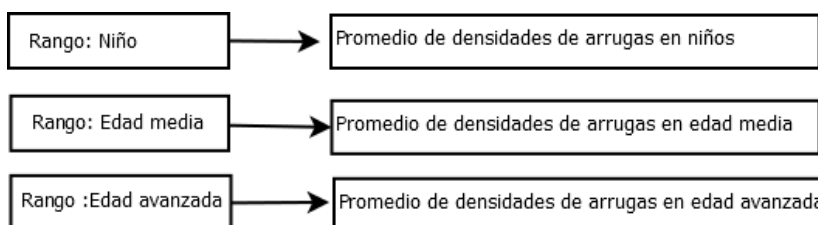


Fig. 7. Asignación de densidades a cada rango de edad

3.3.4 EVALUACIÓN DEL SISTEMA

En este bloque se utilizará el sistema, teniendo como entrada la misma base de datos de imágenes. Se realizará la comparación de la densidad de arrugas de cada imagen y posteriormente se clasifica en un rango de edades que genera los resultados de la tabla 2. El bloque se muestra en la Figura 8.

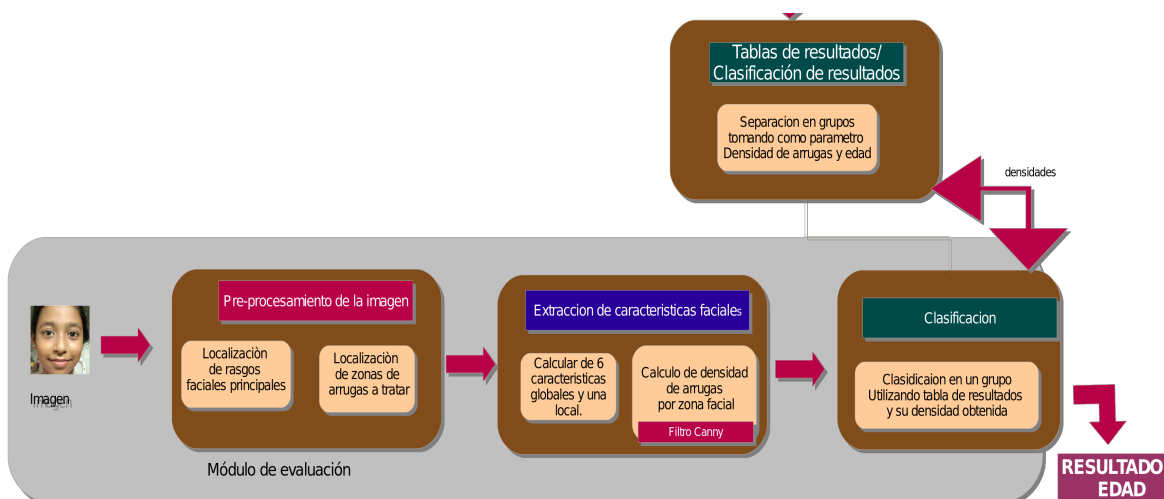


Fig. 8. Bloque/módulo de prueba del sistema

4 ANÁLISIS DE RESULTADOS

Las implementaciones realizadas y que servirán para las fases posteriores a la localización de rasgos globales y locales son:

- Ecuilización de histograma.
- Localización de ojos y pupilas.
- Filtro para mejoramiento de localización de ojos.
 - Filtro de threshold.
 - Operador de dilatación.
- Filtro Canny.

4.1 ECUALIZACIÓN DE HISTOGRAMA

La ecualización se utiliza para estandarizar el contraste y brillo de la imagen, esto es, para que distintas condiciones de iluminación no afecten la detección de los rasgos en la imagen, ver la Figura 9.



Fig. 9. Resultados de ecualización de histograma canny a) Imagen original. b) Imagen ecualizada.

4.2 LOCALIZACIÓN DE OJOS

En esta etapa, se procesan los rostros de las imágenes, previamente ecualizadas, para hallar la posición de los ojos. Se implementó el uso de tres clasificadores que representa el archivo xml, con toda la información que OpenCV necesita para poder identificar objetos y así seleccionar el más conveniente para la propuesta del algoritmo.

Se utilizan las siguientes tres bases de datos de fotografías, en ambientes totalmente controlados.

- Iranian women (Set de 150 imágenes).
- Alumnos Facultad de Ciencias de la Computación BUAP (Set 53 imágenes).
- Nottingham_scans (Set 100 imágenes).

Los clasificadores utilizados fueron:

- haarcascade_eye.
- haarcascade_eye_tree_eyeglasses.
- haarcascade_mcs_lefteye/ mcs_right eye.

Los resultados de esta fase fueron calculados en términos de porcentajes de aciertos, es decir, donde ambos ojos fueron localizados correctamente, como se observa en la Tabla 1.

Tabla 1. Tabla de resultados de los distintos clasificadores utilizados

Clasificador	Base de datos de imágenes		
	Iranian	FCC BUAP	Nottingham_scans
Eye	70.00%	86.79%	87.02%
eye_glasses	29.20%	77.35%	70.20%
mcs_lefteye/righteye	32.14%	81.13%	79.00%

4.3 FILTRO PARA MEJORAMIENTO DE LOCALIZACIÓN DE OJOS Y PUPILAS

La ubicación de ojos y pupilas es una parte crítica del algoritmo, ya que la distancia entre las dos pupilas, representa una medida antropométrica que determinará las características W , que corresponden a las zonas de arrugas.

Por ello, se propone un filtro, que mejore el rendimiento de los clasificadores, y que ayude a cubrir o eliminar zonas que sean semejantes a las características de los ojos y evitar generar falsos positivos, como en las zonas siguientes: la punta de la nariz y la parte superior de los labios (Figura 10).



Fig. 10. Ejemplos de patrones semejantes en el rostro.

4.3.1 FILTRO THRESHOLD

El filtro threshold separa las regiones de una imagen correspondientes a los objetos que se quiere analizar. Esta separación se basa en la variación de la intensidad entre los píxeles de objetos y los píxeles de fondo, ver figura 11.

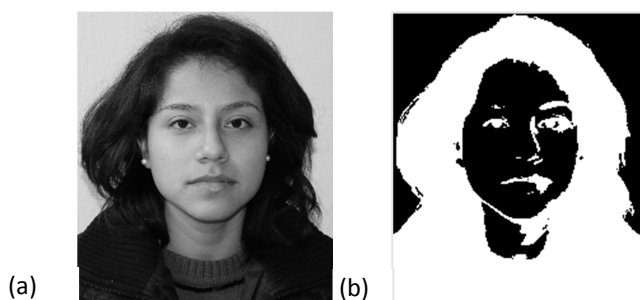


Fig. 11. Resultados de threshold a) Imagen escala de grises b) Filtro threshold binario.

4.3.2 OPERADOR DILATACIÓN

Esta operación morfológica agranda la zona blanca de la máscara y de esta forma los orificios de nariz y labios quedan invertidos de oscuros a claros, se observa en la Figura 12.

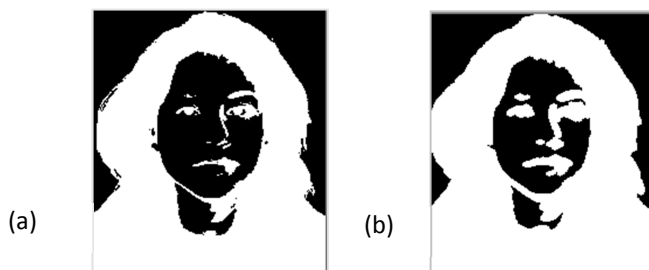


Fig. 12. Resultados de dilatación a) Imagen con threshold. b) Operación dilatación.

Se observa en la zona de los ojos que queda igualmente invertida, para evitar esta situación, se propone modificar únicamente la segunda mitad de la imagen ya que la nariz y la boca quedan definidas en esta parte de la imagen, ver la Figura 13. Observe, que los patrones de nariz y labios quedan invertidos y por ello no causarán falsos positivos.

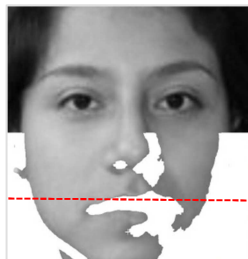


Fig. 13. Resultados de eliminación de patrones conflictivos.

Al procesar la imagen resultante de la Figura 13. Nuevamente se utilizaron los clasificadores, y los resultados que se obtuvieron fueron altamente mejorados, como se puede apreciar en la Tabla 2. Se concluye que el mejor clasificador es el mcs_lefteye/righteye, que se obtiene hasta un 96.93 % de precisión.

Tabla 2. Resultado comparativos imagen original vs. imagen con patrones conflictivo eliminados

Clasificador	Base de datos de imágenes					
	Iranian		FCC BUAP		Nottingham_scam	
	Original	Mejorado	Original	Mejorado	Original	Mejorado
Eye	70.00%	72.85%	86.79%	87.79%	87.02%	90.01%
eye_glasses	29.20%	70.00%	77.35%	77.35%	70.20%	71.42%
mcs_lefteye/righteye	32.14%	82.85%	81.13%	84.20%	79.00%	96.93%

4.4 LOCALIZACIÓN DE PUPILAS

Para la localización de pupilas, se aplica el filtro Canny y se busca en la zona de los ojos utilizando el clasificador, el mayor conjunto de píxeles oscuros agrupados en la imagen binarizada que se obtuvieron con el filtro Canny, de esta manera las pupilas son ubicadas, ver la Figura 14 (b).

La distancia entre las dos pupilas es una medida antropométrica que ayuda a la localización de área de arrugas en las siguientes fases del proyecto.

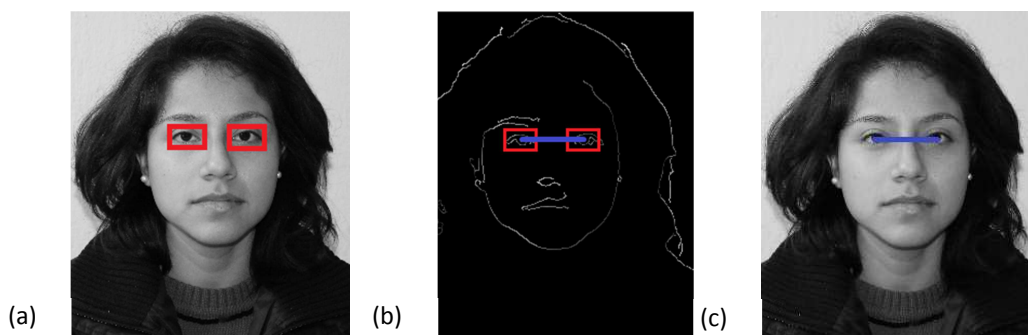


Fig. 14. Localización de pupilas a) Resultado de clasificador. b) Filtro Canny para hallar pupilas. c) Distancia entre pupilas

Los resultados que se obtuvieron servirán de soporte, para las fases siguientes del proyecto que son el cálculo de densidad de arrugas y del entrenamiento del sistema.

5 CONCLUSIÓN

En este trabajo de investigación, se muestran los resultados obtenidos por los filtros siguientes: threshold, binarización, Canny, localización de ojos y pupilas, que son necesarios para obtener la densidad de las arrugas que se propone en el algoritmo. Estos filtros serán necesarios para las siguientes fases de implementación del algoritmo, por lo que, la base de este algoritmo esta finalizado.

Se ha comprobado que el desempeño del sistema depende del buen uso de cada uno de estos filtros, es decir su calibración. Ya que de no hacerlo correctamente, se genera demasiado ruido y genera falsos positivos.

Se concluye, que tanto las áreas de las arrugas como otras características faciales relevantes, están ligadas a las propiedades antropométricas y especialmente de las distancias entre los ojos, característica que fue útil y que fue la base para este trabajo de investigación.

Mediante el entrenamiento se utilizaran bases de datos de imágenes con mayor volumen y obtener mejores aproximaciones entre las densidades de las arrugas y la edad más próxima, y clasificarlas en grupos con edades de 5 a 10 de años de margen. Se pretende utilizar el filtro de detección de bordes Sobe, ya que las arrugas son bordes no cerrados, y así poder realizar la comparación de los resultados con el filtro Canny.

REFERENCIAS

- [1] J. C. Clint and A. K. Jaén. "A case study en unconstrained facial recognition using the boston marathon bombings suspects", Technical Report, pp.4-5, 2013.
- [2] A. Ross A.K. Jain, P. Flynn. Handbook of biometrics. Springer, M. Chandra Moha, V. Vijaya Kumar, V. Venkata Krishna: "Novel Method of Adult Age Classification Using Linear Wavelet Transforms." *IJCSNS, International Journal of Computer Science and Network Security*, vol10, no.3, 2007.
- [3] Next generation identification. [Online], Available: us/cjis/fingerprints_biometrics/ngi/ngi2.
- [4] Belhumeur N, Hespanha JP, Kriegman D. "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *Computer Vision — ECCV '96*, Volume 1064 of the series Lecture Notes in Computer Science pp 43-58.
- [5] Zait D, Super BJ, Quek FKH., "Comparison of five color models in skin pixel classification" *RATFG-RTS '99 Proceedings of the International Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*. pp 58.
- [6] Peng Yang, Lin Zhong, Dimitris Metaxas, "Ranking Model for Facial Age Estimation", *International Conference on Pattern Recognition*, 2010.
- [7] Pantic,M. and Rothkrantz, JM, "Automatic analysis of facial expressions: The state of the art", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp12, 2000.
- [8] Thomas B. Moeslund. "Image and Video Processing". August 2008.
- [9] Contet-Audonneau JL, Jeanmaire C, Pauly G. "A histological study of human wrinkle structures: comparison between sun-exposed areas of the face, with or without wrinkles, and sun-protected areas", *British Journal of Dermatology*, vol. 140, pp. 1038-1047, 2000.