La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

Método para Autenticación Facial basado en SURF e implementado en dispositivos móviles.

Mendoza Martínez Cyntia, Pedraza Ortega Jesús Carlos, Sotomayor Olmedo Artemio, Rodríguez Reséndiz Juvenal, Carrillo Serrano Roberto Valentín, Aceves Fernández Marco Antonio.

Facultad de Ingeniería, Universidad Autónoma de Querétaro

Resumen

En el presente artículo se propone una metodología para el proceso de autenticación facial, el cual incluye una etapa de preprocesamiento de imágenes así como del uso del algoritmo SURF (Speeded Up Robust Features) dentro de la etapa de extracción de características. Se implementó este algoritmo robusto en la computadora y posteriormente en diferentes dispositivos móviles como Smartphones y Tablets con sistema operativo Android. La metodología propuesta consta de los las siguientes seis etapas principales: imágenes del rostro, normalización, detección del rostro, extracción de características (preprocesamiento), coincidencias y decisión, donde a partir de la definición de un umbral se podrá determinar si la autenticación fue exitosa o errónea. Se requirió del uso de una etapa de preprocesamiento mediante el uso de técnicas de ecualización de histograma con el fin de que cada imagen tenga una distribución uniforme de sus niveles de gris, y que como resultado se pueda obtener una imagen mejorada para posteriormente pasar al siguiente paso de la metodología. Dentro de la implementación se utilizaron bases de datos públicas, por ejemplo; The Extended Cohn-Kanade Dataset (CK+), Caltech Faces y FERET. Finalmente, los resultados obtenidos se implementaron en diferentes dispositivos móviles. Los resultados que se pretendan alcanzar deberán de tener un promedio de aciertos arriba del 75%.

Palabras clave: Face authentication, SURF, dispositivos móviles.

1. Introducción

La biometría se refiere a la identificación de una persona en base a sus características físicas y de comportamiento. Incluye el estudio de los métodos para reconocer únicamente los seres humanos basados en uno o más rasgos físicos. La biometría se puede dividir en tres clases:
1) Biometría fisiológica, relacionada con la forma del cuerpo (Reconocimiento facial, huella digital, escaneo del iris, escaneo de retina, escaneo de la mano).
2) Biometría del comportamiento, relacionada con el comportamiento de una persona (Escaneo de la voz, escaneo de la firma, pulsación de teclas).
3) Biometría química, ésta involucra la medición de parámetros químicos tales como el olor, la composición química de la transpiración, etc.

Dentro de la biometría fisiológica, el reconocimiento facial destaca como una de las aplicaciones más utilizadas y de mayor éxito durante los últimos años. El reconocimiento de rostros humanos, es un campo de investigación multidisciplinario donde convergen diferentes áreas de estudio como: procesamiento de imágenes, reconocimiento de patrones, redes neuronales, lógica difusa, clustering, máquina de soporte de vectores, y psicología entre otros. El reconocimiento facial puede operar de dos modos: reconocimiento o identificación y autenticación o verificación. Llevar a cabo un reconocimiento o identificación facial significa dar una imagen de la cara y se requiere que el sistema diga quién (si él o ella) es la más probable identificación. En este procedimiento se dice que la

coincidencia es de 1:K, donde K representa el número de clases, es decir, se compara la imagen de entrada con las K existentes en la base de datos para concluir si se trata de una coincidencia o no [1].

Mientras que en la autenticación o verificación facial, dada una imagen del rostro y una estimación de la identificación, se requiere que el sistema diga si es verdadera o falsa la estimación que se realizó. En este caso la coincidencia es de 1:1 ya que dada una imagen de entrada se compara con una sola clase de imágenes (la cual representa un conjunto de imágenes de la misma persona).

Todos los mecanismos de reconocimiento y autenticación que existen siguen un proceso general que incluye los siguientes pasos [2]:
1) El usuario pide acceso a un recurso (por ejemplo a un dispositivo móvil).
2) El sistema le solicita al usuario su medio de autenticación (rostro, iris, etc.).
3) El usuario entrega sus credenciales (características) de autenticación.
4) El sistema verifica las credenciales del usuario.
5) El sistema niega o proporciona al usuario el acceso al recurso.
Una vez concluido este proceso el siguiente paso es llevar a cabo la evaluación general del sistema para medir su desempeño, por lo cual diversos trabajos de investigación [3], [4], [5], toman en cuenta los siguientes parámetros;
• Verdaderos positivos (True positive, TP): El sistema reconoce las credenciales de un usuario conocido (los datos del usuario se encuentran almacenados en una base de datos) y le permite el acceso a un recurso o dispositivo.
• Falsos positivos (False positive, FP): El sistema reconoce las credenciales de un usuario no conocido (los datos del usuario no se encuentran almacenados en una base de datos) y le permite el acceso a un recurso o dispositivo.
• Verdaderos negativos (True negative, TN): El sistema no reconoce las credenciales de un usuario no conocido (los datos del usuario no se encuentran almacenados en una base de datos) y no permite el acceso a un recurso o dispositivo.
• Falsos negativos (False negative, FN): El sistema no reconoce las credenciales de un usuario conocido (los datos del usuario se encuentran almacenados en una base de datos) y no le permite el acceso a un recurso o dispositivo.

Durante la última década se han propuesto diversos algoritmos de autenticación y reconocimiento facial, los más usados en este campo son LDA, PCA, SIFT y SURF, sin embargo las principales limitantes de los sistemas que usan estos algoritmos es su dependencia a las condiciones de iluminación, posición, forma y tamaño del rostro [6].

Durante la última década se han propuesto diversos algoritmos de autenticación y reconocimiento facial, los más usados en este campo son LDA, PCA, SIFT y SURF.

En el 2011, Fedias y Saigaa, propusieron un metodo de autenticación facial basado en características estadísticas de primer orden, el cual fue implementado en computadora [7]. En el 2012, Bairagi et al. propuso un Sistema de reconocimiento facial utilizando carácterísticas de SURF y Gabor [8]. Ren et al., en el 2013 desarrolló una metodología para un Sistema de verificación facial en dispositivos móviles [9]. Y en el 2015, Stokkenes et al., propuso un Sistema multi-modal para Smart phonesutilizando reonocimiento de rostro e iris, basado en el algoritmo SURF [10][11]

En resumen, reconocimiento o autenticación facial presenta una problemática, pues el promedio de aciertos en la estimación de la identificación es bajo, en particular, se estima que se encuentra entre un 35% y un 65% de efectividad, dependiendo de las condiciones de iluminación, tamaño de la imagen, etc. por lo que el sistema de autenticación facial mediante procesamiento digital de imágenes debe realizar los ajustes necesarios para aumentar el porcentaje de reconocimiento en más de un 80%.

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

La mayoría de sistemas autenticación utilizan como medio de procesamiento de las imágenes una computadora personal y una cámara que captura el rostro y no indican que tipo de procesamiento llevan a cabo, es decir, que algoritmo se implementó. En la mayoría de los sistemas trabajan con uno o máximo dos algoritmos, sin que se pueda modificar.

2. Características Robustas y Rápidas (Spedded Up Robust Features SURF)

Características Robustas y Rápidas (SURF ó Speeded Up Robust Features), fue desarrollado por Herbert Bay et al. (Bay et al., 2006) como un detector de puntos de interés y descriptor robusto.
De forma general, SURF es un algoritmo de visión por computador, capaz de obtener una representación visual de una imagen y extraer información detallada y específica del contenido. Esta información es tratada para realizar operaciones como por ejemplo la localización y reconocimiento de determinados objetos, personas o caras, realización de escenas 3D, seguimiento de objetos y extracción de puntos de interés. Este algoritmo forma parte de la inteligencia artificial, la cual es capaz de entrenar un sistema para que interprete imágenes y determine su contenido.
Cabe mencionar que SURF está basado en su predecesor SIFT [11], aunque presenta notables diferencias. Los autores afirman que este detector y descriptor presenta principalmente dos mejoras resumidas en los siguientes conceptos [12]:
Velocidad de cálculo considerablemente superior sin ocasionar perdida del rendimiento.
Mayor robustez ante posibles transformaciones de la imagen.
Estas mejoras se consiguen mediante la reducción de la dimensionalidad y complejidad en el cálculo de los vectores de características de los puntos de interés obtenidos. SURF se compone de tres pasos consecutivos [13], los cuales son:
1) Detección de los puntos de interés.
2) Descripción de los puntos de interés.
3) Correspondencia entre puntos de interés.

2.1 Detección de puntos de interés

SURF hace uso de la matriz Hessiana, más concretamente, del valor del determinante de la matriz, para la localización y la escala de puntos de interés en una imagen.
Ahora bien, el motivo para la utilización de dicha matriz Hessiana es respaldado por su rendimiento en cuanto a la velocidad de cálculo y a la precisión. Lo realmente novedoso del detector incluido en el descriptor SURF respecto de otros detectores es que no utiliza diferentes medidas para el cálculo de la posición y la escala de los puntos de interés individualmente, sino que utiliza el valor del determinante de la matriz Hessiana en ambos casos.
Por lo tanto, dado un punto p=(x,y) de la imagen I, la matriz Hessiana definida como $H(p,\sigma)$del punto p y perteneciente a la escala σ se define como se puede observar en la ecuación 1.

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{xy}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix} \quad (1)$$

Donde $L_{xx}(p,\sigma)$ representa la convolución de la derivada parcial de segundo orden de la Gaussiana $\frac{\partial^2}{\partial x^2} g(\sigma)$ con la imagen *I* en el punto *p*. De manera análoga ocurre con los términos $L_{xy}(p,\sigma)$, $L_{yy}(p,\sigma)$ de la matriz.
A pesar de que los filtros gaussianos son óptimos para el análisis del espacio-escala, se ha implementado una alternativa a los filtros gaussianos en el detector SURF debido a una serie de limitaciones de estos filtros (como la necesidad de ser discretizados, la falta de prevención total del indeseado efecto aliasing, etc.), esta alternativa son los filtros tipo caja (box-filters).
Estos nuevos filtros aproximan las derivadas parciales de segundo orden de las gaussianas y pueden ser evaluados de manera muy rápida usando imágenes integrales, independientemente del tamaño de éstas. Las imágenes integrales, cuya definición se encuentra ampliamente detallada en [14](Derpanis, 2007) y [15] (Viola y Jones, 2002) son calculadas mediante la ecuación 2.

$$Ii_{\sum}(x,y) = \sum_{i=1}^{i\leq x} \sum_{j=1}^{j\leq y} I(i,j) \quad (2)$$

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

Donde *(x,y)* representa la posición del punto en la imagen y $I_i(x,y)$ representa la intensidad de la imagen en el punto. Una vez que la imagen integral ha sido creada, se puede calcular la suma de las intensidades de una región por medio de la ecuación 3.

$$\sum I = Ii_D + Ii_A + Ii_B + Ii_C \quad (3)$$

De esta forma, el tiempo necesario para el cálculo de las operaciones de convolución es independiente del tamaño de la imagen. De este modo resulta que el espacio escala es analizado mediante la elevación del tamaño del filtro, en vez de reducir el tamaño de la imagen como es el caso del detector SIFT. Ésta diferencia se puede apreciar en la Figura 1.

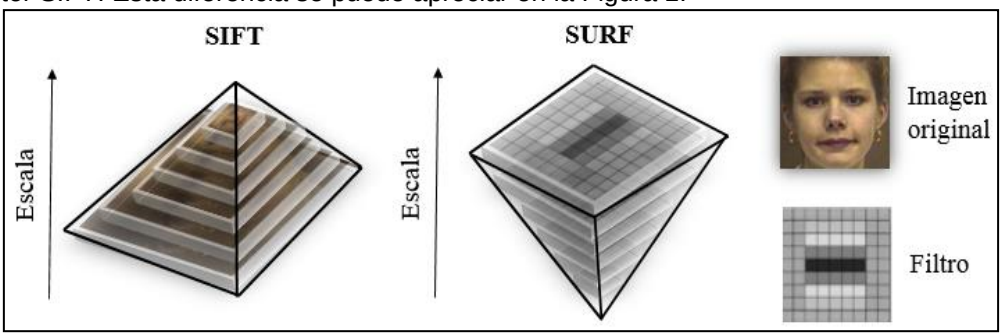

Figura 1 Representación de la intensidad de una región respecto de la imagen integral.

Las aproximaciones de las derivadas parciales se denotan como D_{xx}, D_{xy} y D_{yy}. En cuanto al determinante de la matriz Hessiana, éste queda definido por la ecuación 4.

$$\det(H_{aprox.}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (4)$$

Donde el valor de 0.9 está relacionado con la aproximación del filtro gaussiano. En la práctica este valor es constante y no tiene un impacto significante en los resultados de los experimentos [16]. En la Figura 2 se puede observar la representación de la derivada parcial de segundo orden de un filtro gaussiano discretizado y la aproximación de la derivada implementada en el caso del descriptor SURF.

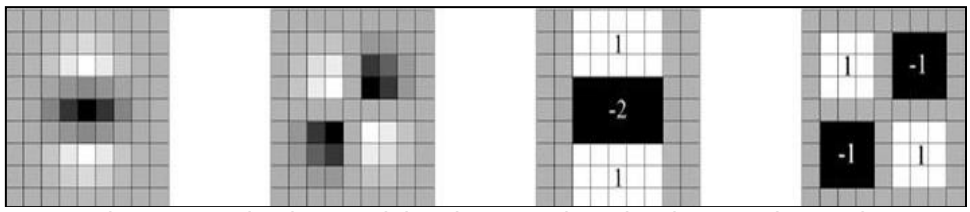

Figura 2. Derivadas parciales de segundo orden (Bay et al., 2006).

De izquierda a derecha de la Figura 4 se aprecian las derivadas parciales de segundo orden discretizadas y recortadas en las direcciones *y* así como *xy*, así como las aproximaciones de las mismas mediante los filtros tipo caja.

La imagen de salida obtenida tras la convolución de la imagen original con un filtro de dimensiones 9 x 9, que corresponde a la derivada parcial de segundo orden de una gaussiana con $\sigma = 1.2$, es considerada como la escala inicial o también como la máxima resolución espacial (s=1.2, correspondiente a una gaussiana con $\sigma = 1.2$). Las capas sucesivas se obtienen mediante la aplicación gradual de filtros de mayores dimensiones, evitando así los efectos de aliasing (curvas) en la imagen.

El espacio escala para SURF, al igual que en el caso de SIFT, está divido en octavas. Sin embargo, en SURF las octavas están compuestas por un número fijo de imágenes como resultado de la convolución de la misma imagen original con una serie de filtros tipo caja más grandes.

El incremento o paso de los filtros dentro de una misma octava es el doble respecto del paso de la octava anterior, al mismo tiempo que el primero de los filtros de cada octava es el segundo de la octava predecesora, como se muestra en la Figura 3.

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

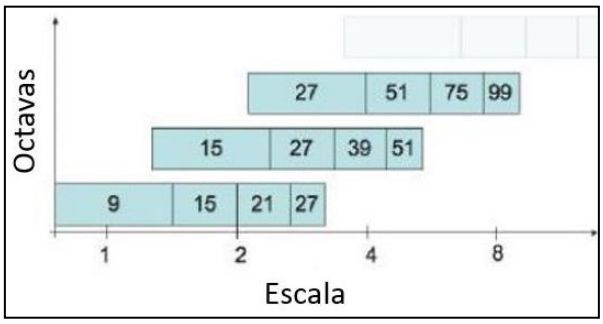

Figura 3 Representación de la longitud de los filtros de diferentes octavas [12].

De esta manera se obtienen las siguientes series de octavas con sus respectivos filtros:

- Octava inicial: $9x9 \overset{6}{\rightarrow} 15x15 \overset{6}{\rightarrow} 21x21 \overset{6}{\rightarrow} 27x27$
- Octava siguiente: $15x15 \overset{12}{\rightarrow} 27x27 \overset{12}{\rightarrow} 39x39 \overset{12}{\rightarrow} 51x51$
- Octava siguiente: $27x27 \overset{24}{\rightarrow} 51x51 \overset{24}{\rightarrow} 75x75 \overset{24}{\rightarrow} 99x99$
- Y así sucesivamente…

Finalmente para calcular la localización de todos los puntos de interés en todas las escalas, se procede mediante la eliminación de los puntos que no cumplan la condición de máximo en un vecindario de $3x3x3$. De esta manera, el máximo determinante de la matriz Hessiana es interpolado en la escala y posición de la imagen. En este punto se da por concluida la etapa de detección de los puntos de interés.

2.2 Descripción de puntos de interés

Antes de pasar a la creación del descriptor, la siguiente etapa corresponde a la asignación de la orientación de cada uno de los puntos de interés obtenidos en el paso anterior. Es en esta etapa donde se otorga al descriptor de cada punto la invarianza ante la rotación mediante la orientación del mismo.

Primero hay que realizar el cálculo de la respuesta de Haar en ambas direcciones tanto en x como en y, esto se lleva a cabo mediante las funciones representadas en la Figura 4, donde el color negro tiene el peso de -1 y el color blanco tiene el peso de +1. Además, El área de interés para el cálculo de las respuestas de Haar es el área circular centrada en el punto de interés y de radio *6s*, siendo *s* (donde *s* ≥*1*) la escala en la que el punto de interés ha sido detectado. De la misma manera, la etapa de muestreo depende de la escala, tomándose como valor a s. Respecto de las funciones onduladas de Haar, se toma el valor 4s, por tanto dependiente también de la escala, como referencia, donde a mayor valor de escala mayor es la dimensión de las funciones onduladas.

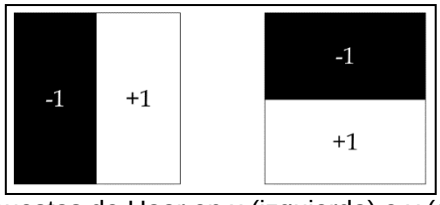

Figura 4 Respuestas de Haar en x (izquierda) e y (derecha) [12].

Tras haber realizado todos estos cálculos, se utilizan imágenes integrales nuevamente para proceder al filtrado mediante las máscaras de Haar y obtener así las respuestas en ambas direcciones. Asimismo, son necesarias únicamente seis operaciones para obtener la respuesta en la dirección x e y. Una vez que las respuestas onduladas han sido calculadas, son ponderadas por una gaussiana de valor σ=2.5s centrada en el punto de interés.

Las respuestas son representadas como vectores en el espacio colocando la respuesta horizontal y vertical en el eje de abscisas y ordenadas respectivamente. Después, se obtiene una orientación dominante por cada sector mediante la suma de todas las respuestas dentro de una ventana de orientación móvil cubriendo un ángulo de π/3 siguiendo las especificaciones recomendadas por Bay [16]. La orientación final del punto de interés será finalmente aquella cuyo vector sea el más grande

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

dentro de los seis sectores en los que ha sido dividida el área circular alrededor del punto de interés. Esta orientación se puede observar en la Figura 5.

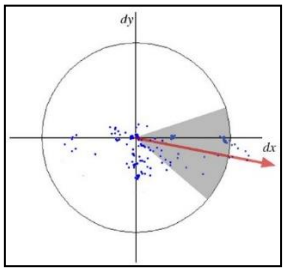

Figura 5 Asignación de la orientación de un punto de interés [12].

Ahora bien, se procede con la creación del descriptor SURF para cada punto de interés. Como primer paso se construye una región cuadrada de tamaño *20s* alrededor del punto de interés con la orientación calculada en la etapa anterior. Esta región es a su vez dividida en 4x4 sub-regiones dentro de cada una de las cuales se calculan las respuestas de Haar de los puntos con una separación de muestreo de *5x5* en ambas direcciones. Por simplicidad, se consideran d_xy $d_y$las respuestas de Haar en las direcciones horizontal y vertical respectivamente relativas a la orientación del punto de interés. En la Figura 6 están representadas tanto las respuestas de Haar en cada una de las sub-regiones alrededor del punto de interés así como las componentes d_xy $d_y$de uno de los vectores.

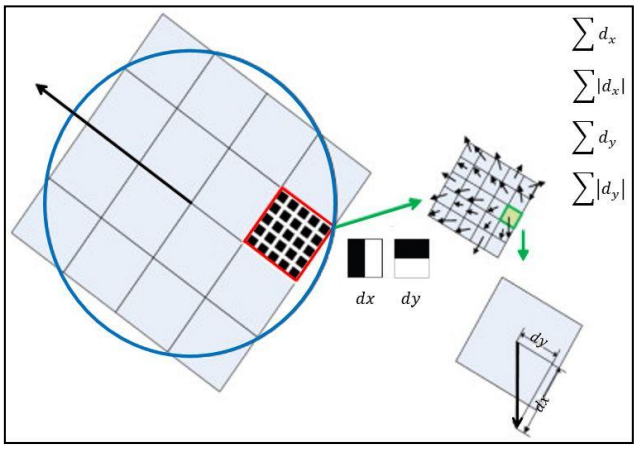

Figura 6 Respuestas de Haar en las sub-regiones del punto de interés [12].

Para dotar a las respuestas d_x y $d_y$de una mayor robustez ante deformaciones geométricas y errores de posición, éstas son ponderadas por una gaussiana de valor $\sigma = 3.3s$ centrada en el punto de interés. En cada una de las sub-regiones se suman las respuestas d_x y $d_y$obteniendo así un valor de d_x y $d_y$representativo por cada una de las sub-regiones.

Al mismo tiempo se realiza la suma de los valores absolutos de las respuestas $|d_x|$ y $|d_y|$ en cada una de las sub-regiones, obteniendo de esta manera, información de la polaridad sobre los cambios de intensidad. En resumen, cada una de las sub-regiones queda representada por un vector v de componentes d_x,d_y, $|d_x|$ y $|d_y|$, la representación matemática de este vector se puede apreciar en la ecuación 5.

$$v = \left(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|\right) \quad (5)$$

Por lo tanto, englobando las 4x4 sub-regiones, resulta un descriptor SURF con una longitud de 64 valores para cada uno de los puntos de interés identificados.

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

2.3 Correspondencia entre puntos de interés

A este último paso del algoritmo SURF también se le conoce como *matching* (coincidencia), ya que tiene como finalidad el cálculo de un valor que represente el grado de similitud entre dos imágenes, y que a continuación se puedan establecer las diferentes conclusiones. El cálculo de este valor, representado como distancia y conocido también como score, se realiza mediante la aplicación de una métrica o fórmula de la distancia entre ambas imágenes. Previo al cálculo del score, es necesario establecer las correspondencias entre los puntos clave. Dicha correspondencia se lleva a cabo mediante el cálculo de la distancia euclidiana entre los vectores de características pertenecientes a diferentes puntos de interés. Este cálculo genera a su vez otro valor que será utilizado para determinar cuál de los puntos de la imagen comparada se corresponde con su homólogo, en el caso de existir, de la primera de las imágenes.

Suponiendo que se quiere realizar el *matching* de puntos entre dos imágenes representadas por I_1 e I_2. Para cada uno de los puntos clave pertenecientes a I_1, se seleccionan los dos mejores candidatos de entre todos los puntos clave pertenecientes a I_2 mediante el criterio de máxima similitud.

Este criterio establece que los mejores candidatos para realizar el *matching* con el punto clave I_1 perteneciente a I_1 cuyo vector de características es v_1, son los puntos clave p'_1 y p'_2pertenecientes a I_2 cuyos vectores de características v'_1 y v'_2representan las distancias euclidianas mínimas d_1 y $d_2$respectivamente, en relación con v_1. Si la relación d_1/d_2entre las distancias mencionadas es suficientemente pequeña, entonces se establece el *matching* entre los puntos p_1 y p'_1pertenecientes a cada una de las imágenes. De acuerdo con Bay [16], se establece un umbral de 0.7 para el ratio d_1/d_2. Esta estrategia de *matching* recibe el nombre de "el vecino más próximo". Finalmente la puntuación o score entre las dos imágenes se obtiene mediante una relación que tiene en cuenta el número total de puntos correspondientes entre ambas imágenes. La Figura 7 muestra un ejemplo de *matching* entre I_1 e I_2.

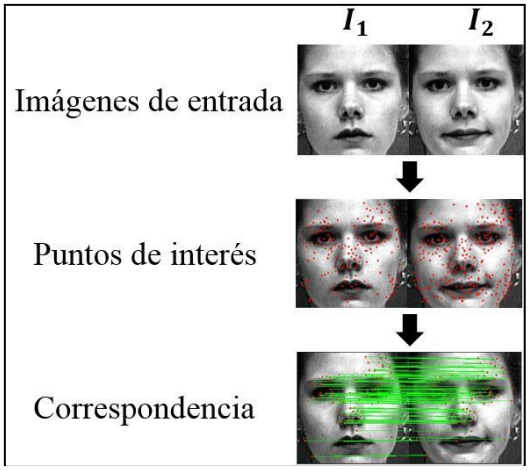

Figura 7 Correspondencia (*matching*) entre dos imágenes.

A pesar del buen desempeño del algoritmo SURF en diferentes aplicaciones como las realizadas por [17] a [22], en este trabajo se propone la implementación de una etapa adicional de preprocesamiento antes de emplear el algoritmo SURF, todo ello con el fin de obtener mejores resultados dentro de un sistema de autenticación facial.

3. Metodología propuesta

La incorporación de algoritmos de autenticación facial, en particular en dispositivos móviles ha ido desempeñando un papel muy relevante en los últimos 10 años. Esto ha sido posible gracias a la continua investigación y desarrollo de nuevos algoritmos que permiten hacer más eficiente la autenticación, así como al avance de la tecnología móvil, hoy en día es posible fusionar ambos

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

elementos de tal manera que puedan brindar una medida de seguridad que sirva para proteger su contenido.

En este trabajo se propone la implementación de una etapa de preprocesamiento dentro del proceso de autenticación facial. Como se puede observar en la Figura 8, esta metodología consta de cinco pasos:
1) Imágenes del rostro.
2) Detección del rostro.
3) Extracción de características.
4) Coincidencia de características (matching).
5) Decisión.

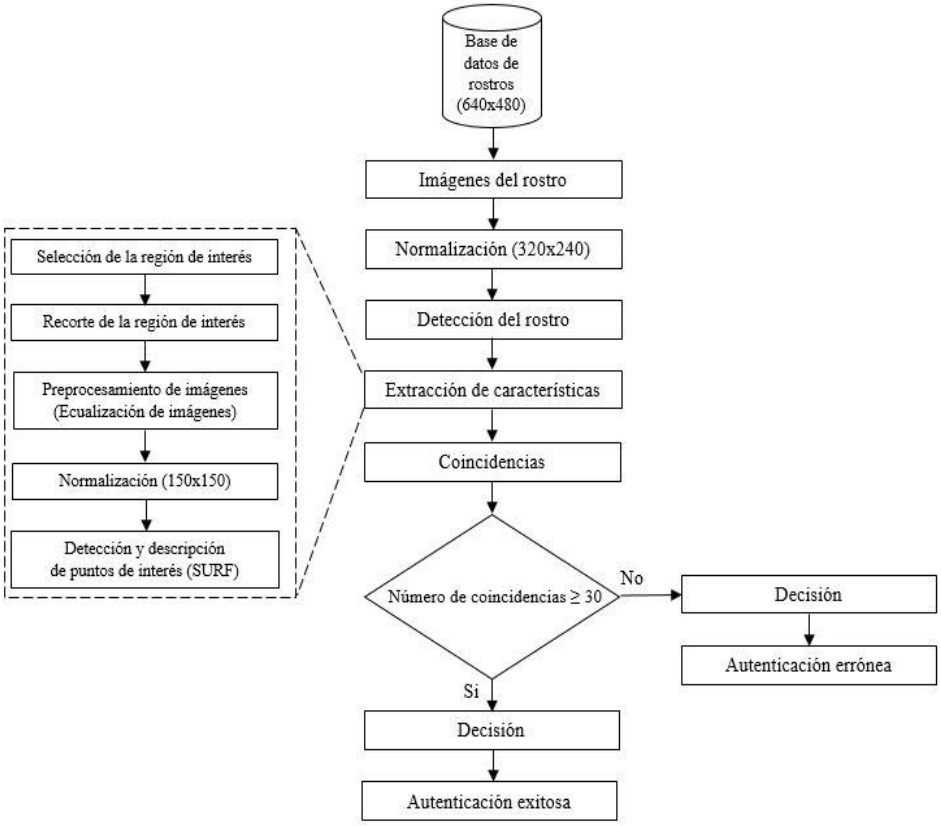

Figura 8 Metodología propuesta

1) Imágenes del rostro. Se deben de adquirir al menos 10 imágenes del rostro de una persona con el propósito de tener la información mínima requerida para el proceso de autenticación facial. Las imágenes se deben tomar con perfil frontal a la cámara del dispositivo móvil, bajo diferentes condiciones de iluminación, expresiones faciales mínimas y poca variación en la posición de la cabeza.
2) Detección del rostro. Una vez seleccionadas las dos imágenes, cada una de ellas se normaliza a 320x240 pixeles y se le aplica el algoritmo Haar como método de detección de rostro, este es considerado como el paso previo al procesamiento de cada una de las imágenes. Información más detallada sobre este método puede ser consultada en [15].
3) Extracción de características. En este paso se implementa una fase de preprocesamiento a las imágenes de entrada que fueron seleccionadas previamente (la imagen del dispositivo y la de la base de datos). Por lo tanto, el objetivo de la extracción de características es obtener sólo las imágenes del rostro (sin el fondo de la imagen, ya que esto facilitará su tiempo de procesamiento) y posteriormente

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

implementar el algoritmo SURF para la detección y descripción de características en las imágenes del rostro. De esta manera se busca la obtención de un mejor resultado de todo el proceso de autenticación facial. Dentro de la etapa de extracción de características, se definen cinco sub-etapas: selección de la región de interés, recorte de la región de interés, preprocesamiento de imágenes, normalización de imágenes, finalmente la detección y descripción de puntos de interés. Una vez que ya se ha recortado la región de interés, la sub-etapa posterior consiste en aplicar a las imágenes de entrada una técnica denominada ecualización de histograma (HE), este proceso en particular se considera como la principal aportación de este trabajo ya que la HE es uno de los métodos más usados para realzar efectivamente el contraste de una imagen [23].

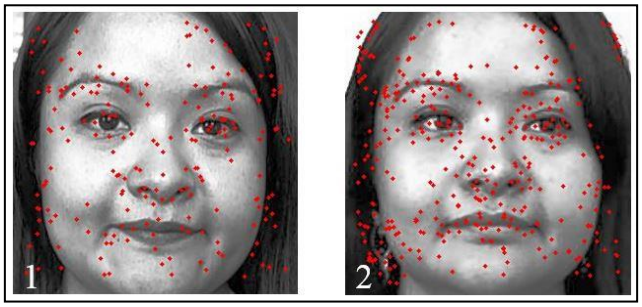

Figura 9 Descriptores de las imágenes de entrada 1 y 2.

Además, la HE modifica el valor de los pixeles de tan manera que la intensidad del histograma de la imagen resultante llegar a ser uniforme, además esta imagen hace uso de todos valores de brillantez posible por lo tanto resulta una imagen realzada en su contaste [24].
4) Coincidencia de características (matching). Una vez extraídas las características (descriptores) de las imágenes de entrada, se continúa con el proceso de coincidencia entre imágenes (en la literatura también se le conoce como match) donde de acuerdo a los descriptores localizados en las imágenes de entrada, se comparan los vectores de la imagen 1 con los vectores de la imagen 2 para determinar cuántos de ellos son similares en ambas imágenes.

Para que un descriptor de una imagen sea considerado como similar en otra imagen, es necesario encontrar la distancia Euclidiana menor entre ellos. Información más detallada sobre este proceso se puede encontrar en (Deza y Deza, 2009). Después, todas las coincidencias que fueron encontradas en ambas imágenes son contadas para determinar su número extracto y de esta manera poder continuar con el paso final de la metodología. Este proceso puede ser apreciado en la Figura 4-10 donde cada descriptor de la imagen 1 que coincide con el descriptor de la imagen 2 es unido por una línea de color verde para representar dicha coincidencia.

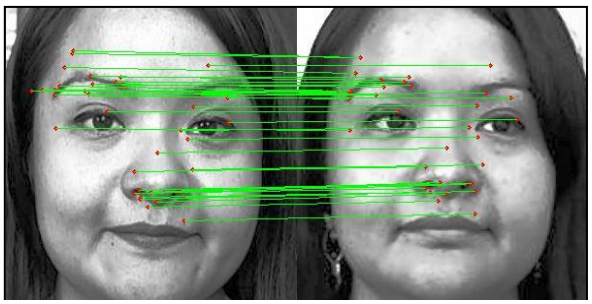
Figura 10 Coincidencias entre las imágenes de entrada.

5) Decisión.

Dado el número total de coincidencias entre las imágenes, el siguiente paso es la decisión donde se ha propuesto un umbral heurístico con valor de 30, esto quiere decir que si el número de coincidencias

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

es mayor o igual a 30, el proceso de autenticación facial se considerará como exitoso (Match). Por el contrario, si es menor a dicho umbral el proceso se considerará como erróneo (Not match).

Con el fin de evaluar esta metodología, los voluntarios, diferentes dispositivos móviles (Smartphones y Tablets), considerando algunas de las imágenes de bases de datos públicas (The Extended Cohn-Kanade Dataset (CK+), Caltech Faces y FERET) fueron sometidos a diferentes pruebas. Los resultados obtenidos se presentan en la siguiente sección.

4. Pruebas y Resultados

Empleando el diagrama de la metodología propuesta en la Figura 8 se realizaron diferentes pruebas en dispositivos móviles, donde en conjunto con Eclipse se instaló la aplicación de autenticación facial en cada uno de ellos, así todo el procesamiento se realiza de forma independiente en cada dispositivo. Las versiones de Android utilizadas fueron distintas para cada dispositivo móvil, variando desde la versión 4.0, hasta la 5.1, dependiendo de la versión que soportaba el dispositivo móvil de acuerdo a sus características internas. A continuación se presentan los resultados de dicha metodología usando las bases de datos públicas The Extended Cohn-Kanade Dataset (CK+), Caltech Faces y FERET en las Tablets Samsung Galaxy Tab 4 y Note 10.1.

Imágenes del rostro: Se seleccionaron las mismas 40 imágenes de cada base de datos que habían sido empleadas en pruebas anteriores, las cuales fueron usadas en escala de grises y normalizadas a 320x240 pixeles.

Detección del rostro: La Figura 11 presenta algunos de los rostros que fueron detectados correcta (95%) e incorrectamente (5%). En ambos dispositivos se obtuvieron los mismos resultados de este proceso.

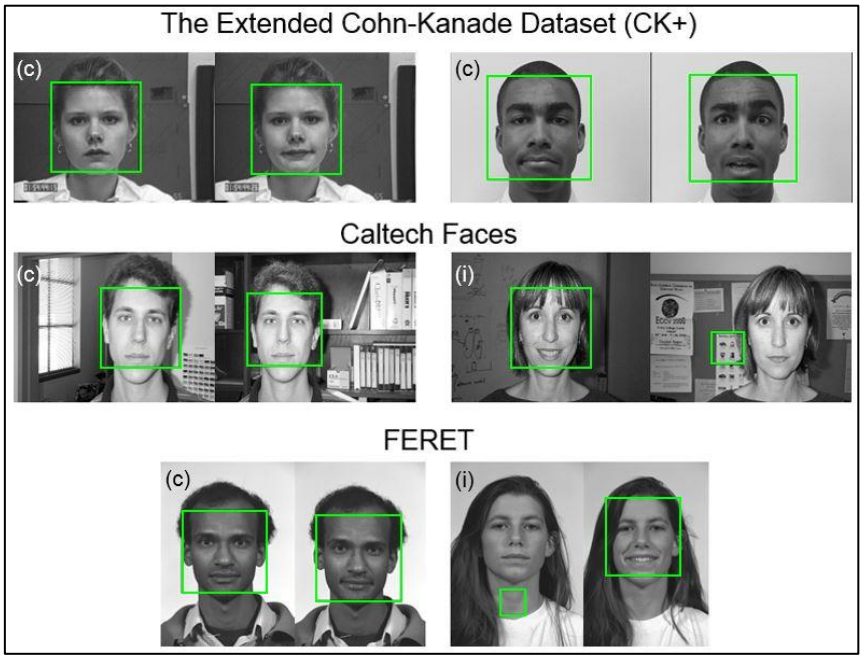

Figura 11 Detección del rostro correcta (C) e incorrecta (I).

Extracción de características: La Figura 12 muestra un ejemplo del proceso de extracción de características en las imágenes de entrada, la cual se considera desde el recorte de la región de interés hasta la detección y descripción de puntos de interés tanto para una detección de rostro correcta como incorrecta.

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

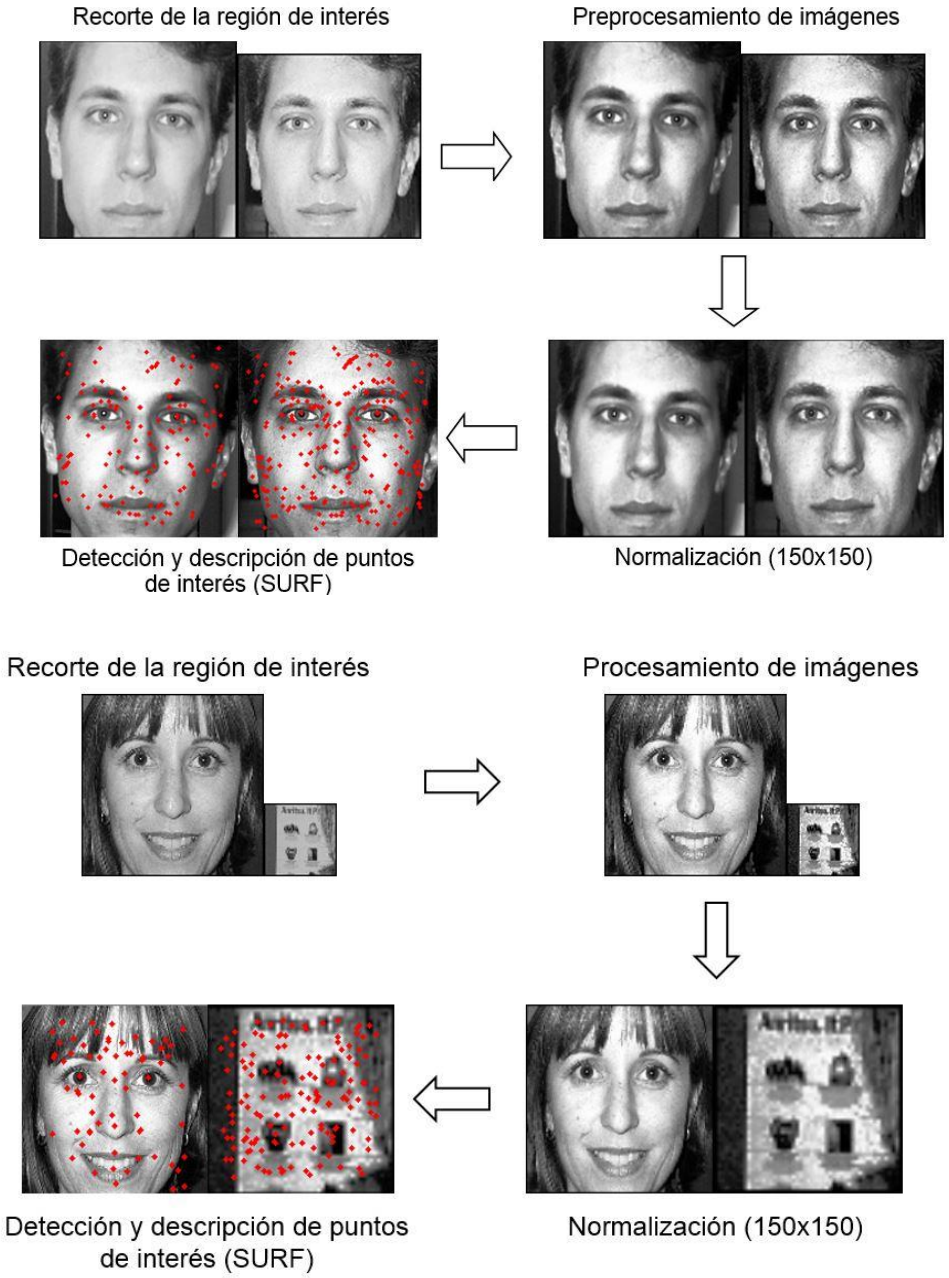

Figura 12 Extracción de características en dispositivos móviles (parte 1).

La Tabla 1 despliega los resultados obtenidos en este proceso además del número de características en promedio por cada base de datos. Cabe resaltar que se adquirieron los mismos resultados en ambos dispositivos.

Coincidencias: La Tabla 2 presenta los resultados obtenidos en este proceso además del número de coincidencias en promedio por cada base de datos. Estos resultados fueron los mismos en ambos dispositivos. Por otro lado, la Figura 13 despliega algunos de los resultados tanto exitosos como erróneos del proceso de autenticación facial en las imágenes de las bases de datos The Extended Cohn-Kanade Dataset (CK+), Caltech Faces y FERET, respectivamente.

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

	CK+		Caltech Faces		FERET	
Prueba	FImg1	FImg2	FImg1	FImg2	FImg1	FImg2
1	103	107	125	122	110	114
2	121	104	129	129	142	134
3	091	107	93	113	144	152
4	110	142	125	136	113	130
5	111	105	137	126	121	127
6	127	129	112	137	101	094
7	106	120	111	107	113	106
8	125	124	130	120	107	092
9	115	107	102	088	099	092
10	086	080	119	115	112	130
11	120	140	098	110	104	106
12	103	098	128	134	139	130
13	093	094	130	120	138	134
14	102	097	115	123	125	111
15	119	116	132	154	141	142
16	134	142	100	105	099	114
17	111	106	129	132	125	130
18	119	117	133	112	125	117
19	124	115	111	099	116	096
20	117	137	128	114	108	106
Promedio	111	114	119	119	119	117

Tabla 1 Número de características en dispositivos móviles (parte 1).

Prueba	CK+	Caltech Faces	FERET
1	49	31	16
2	62	16	63
3	40	37	51
4	44	37	44
5	45	06	15
6	32	21	33
7	59	08	34
8	31	39	29
9	61	07	31
10	49	05	02
11	78	33	63
12	49	04	57
13	65	31	89
14	55	51	57
15	62	37	97
16	62	31	33
17	65	39	04
18	71	06	54
19	47	53	54
20	51	41	35
Promedio	53	26	43

Tabla 2 Número de coincidencias en dispositivos móviles (parte 1).

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

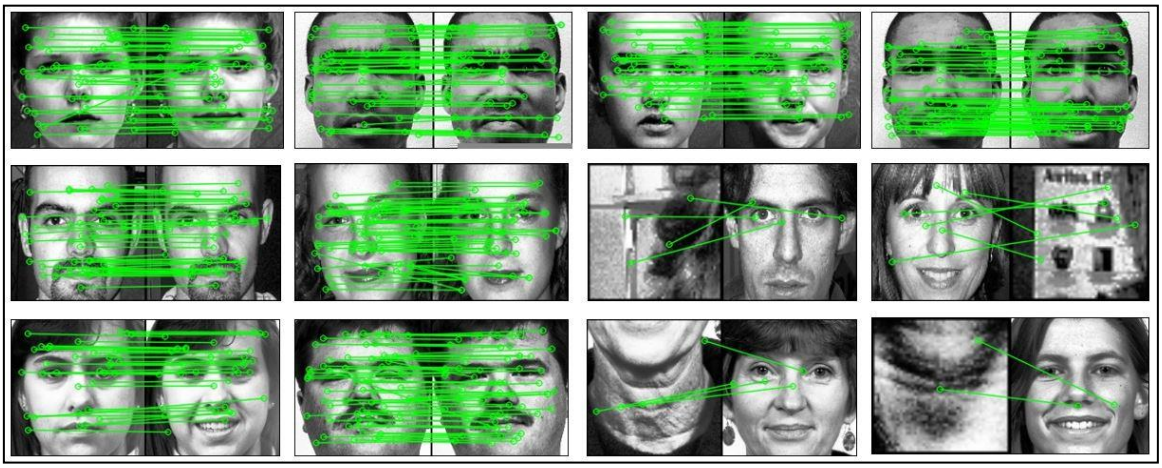

Figura 13 Autenticación facial exitosa y errónea en dispositivos móviles.

Decisión: Los resultados se presentan en la Tabla 3 donde se muestran los verdaderos positivos (TP), falsos negativos (FN) y el porcentaje de autenticidad de las imágenes evaluadas, con el cual se determina si esta fue exitosa o errónea. Estos resultados fueron los mismos para ambos dispositivos.

Base de datos	TP	FN	Autenticidad
CK+	100%	0%	100%
Caltech Faces	60%	40%	60%
FERET	75%	25%	75%
Promedio	78%	22%	78%

Tabla 3 TP y FN en imágenes evaluadas en dispositivos móviles (parte 1).

En las Figuras 14 y 15 se puede observar un ejemplo de la implementación de la metodología propuesta, la cual fue ejecutada en las Tablets Samsung Galaxy Tab 4 y Galaxy Note 10.1, respectivamente.

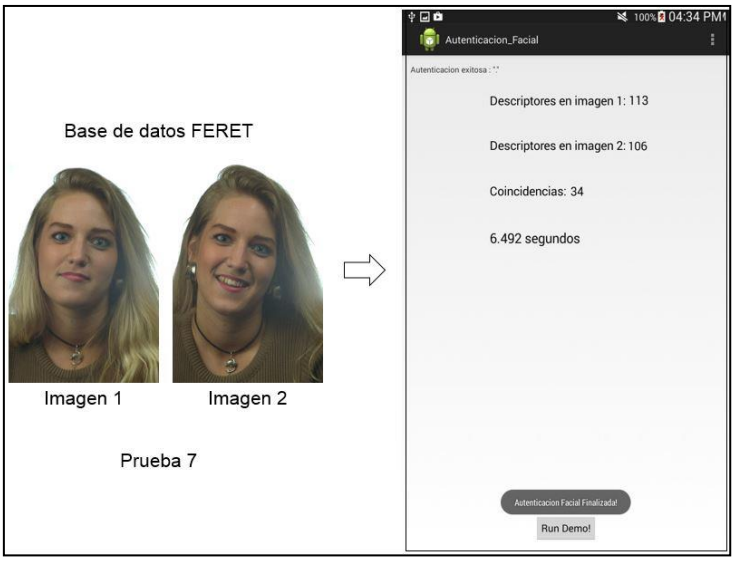

Figura 14 Autenticación facial en Tablet Samsung Galaxy Tab 4.

La Mecatrónica en México, Enero 2017, Vol. 6, No. 1, páginas 21 –35.
Disponible en línea en www.mecamex.net/revistas/LMEM
ISSN: 2448-7031, Asociación Mexicana de Mecatrónica A.C.

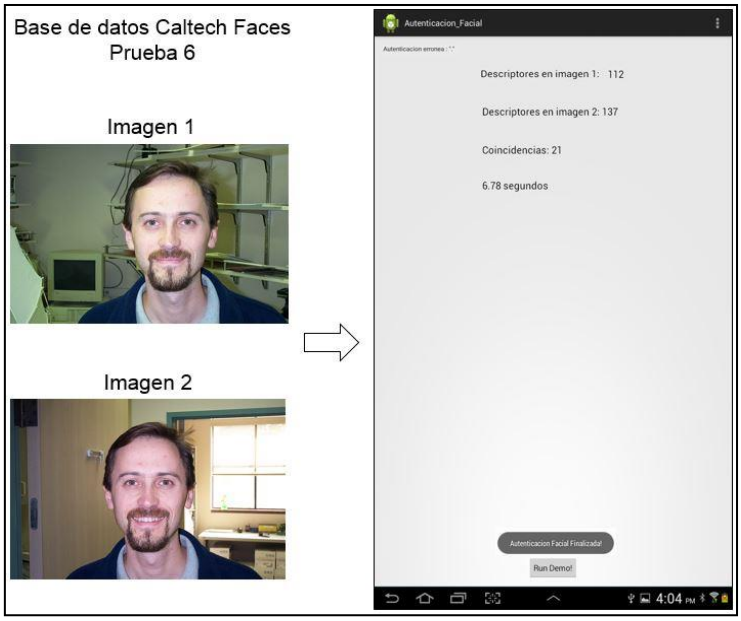

Figura 15 Autenticación facial en Tablet Samsung Galaxy Note 10.1.

5. Conclusiones

En el presente proyecto se presentó una metodología de autenticación facial. Esta metodología incluye el uso del algoritmo SURF y adicional a éste se desarrolló una etapa de preprocesamiento basada en la ecualización de histograma, con el propósito de aumentar el porcentaje de autenticación facial, haciendo con esto que el sistema sea capaz de identificar a la persona que está utilizando este dispositivo y que no le permita a otra persona ingresar a él.
Para validar este trabajo se implementó la metodología en dispositivos móviles que cuentan con el sistema operativo Android y se llevaron a cabo pruebas del desempeño de esta metodología en algunos dispositivos móviles.
Los resultados obtenidos mejoran el porcentaje de autenticación utilizando técnicas similares, e inclusive aumentan ese porcentaje en algoritmos que utilizan exclusivamente el algoritmo SURF. El número de coincidencias reportado hasta antes de este trabajo era de veinte para identificar que se trata de la misma persona, y con el método propuesto se aumentó ese número de coincidencias a 30, lo cual es una mejora significativa con respecto a trabajos similares en el área.

Referencias

[1] Chao, W.-L. 2010. Face Recognition. GICE, National Taiwan University, Taiwan.
[2] Iglesias, G. 2007. Sistema de Autenticación para Dispositivos Móviles basado en Biometría de comportamiento de Tecleo. Tesis de grado de Ingeniería en Sistemas Computacionales. Instituto Tecnológico de Morelia. México, D.F.
[3] Yin, Q., Tang, X., and Sun, J. 2011. An associate-predict model for face recognition. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, p. 497–504
[4] Ruiz, S., S. Lee, S. R. Soekadar, A. Caria, R. Veit, T. Kircher, and R. Sitaram. 2013. Acquired self-control of insula cortex modulates emotion recognition and brain network connectivity in schizophrenia. Human brain mapping, 34(1), p. 200–212.
[5] Valstar, M., J. Girard, T. Almaev, G. McKeown, M. Mehu, L. Yin, and J. Cohn. 2015. Fera 2015-second facial expression recognition and analysis challenge. Proc. IEEE ICFG.

[6] Mendoza-Martinez, C., J. C. Pedraza-Ortega, and J. M. Ramos-Arreguin. 2014. A Novel Approach for Face Authentication Using Speeded Up Robust Features Algorithm. In Human-Inspired Computing and Its Applications, p. 356–367. Springer International Publishing.
[7] Fedias, M., and D. Saigaa. 2011. A New Fast method of face Authentication based on First order Statistical Feature. In International Journal of Computer Applications (IJCA), 14(8), p. 32-37.
[8] Bairagi, B. K., S. C. Das, A. Chatterjee, and B. Tudu. 2012. Expressions invariant face recognition using SURF and Gabor features. In Third International Conference on Emerging Applications of Information Technology (EAIT), IEEE, p. 170-173.
[9] Ren, J., X. Jiang, and J. Yuan. 2013. A complete and fully automated face verification system on mobile devices. In Pattern Recognition, 46(1), p. 45-56.
[10] Stokkenes, M., K. B. Raja, R. Raghavendra, and C. Busch. 2015. Multi-modal Authentication System for Smartphones Using Face, Iris and Periocular. International Conference on Biometrics, IEEE, p.143-150
[11] Lowe, D. G. 2004. Distinctive Image Features from Scale-Invariant Keypoints. In International journal of computer vision, 60(2), p. 91–110.
[12] Boullosa, Ó. 2011. Estudio comparativo de descriptores visuales para la detección de escenas cuasi-duplicadas. Universidad Autónoma de Madrid, Madrid, España.
[13] Oyallon, E., and J. Rabin. 2013. An analysis and implementation of the SURF method, and its comparison to SIFT. In Image Processing On Line, ISSN 2105–1232.
[14] Derpanis, K. 2007. Integral image-based representations. Department of Computer Science and Engineering, York University, p. 1–6.
[15] Viola, P., and M. Jones. 2002. Robust real-time object detection. In International Journal of Computer Vision, vol. 57, no. 2, p. 137–154.
[16] Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. 2008. Speeded-up robust features (SURF). In Computer vision and image understanding, 110(3), p. 346–359.
[17] Terriberry, T. B., L. M. French, and J. Helmsen. 2008. GPU accelerating speeded-up robust features. In Proceedings of 3DPVT, Vol. 8, p. 355–362.
[18] Svab, J., T. Krajnik, J. Faigl, and L. Preucil. 2009. FPGA based speeded up robust features. In Technologies for Practical Robot Applications. IEEE International Conference on, p. 35–41.
[19] Bouris, D., A. Nikitakis, and I. Papaefstathiou. 2010. Fast and efficient FPGA-based feature detection employing the SURF algorithm. In Field-Programmable Custom Computing Machines, 18th IEEE Annual International Symposium on, p. 3–10.
[20] Murali, Y., and M. MITS. 2012. Image mosaic using speeded up robust feature detection. In International Journal of Advanced Research in Electronics and Communication Engineering, 1(3), p. 40–45.
[21] Thakoor, K. A., S. Marat, P. J. Nasiatka, B. P. McIntosh, F. E. Sahin, A. R. Tanguay, J. D. Weiland, and L. Itti. 2013. Attention biased speeded up robust features (AB-SURF): A neurally-inspired object recognition algorithm for a wearable aid for the visually-impaired. In Multimedia and Expo Workshops, IEEE International Conference on, p. 1–6.
[22] Wang, Y. T., C. T. Chi, and Y. C. Feng. 2014. Robot mapping using local invariant feature detectors. In Engineering Computations: International Journal for Computer-Aided Engineering and Software 31(2), p. 297–316.
[23] Paik, J.K. 2011. Image processing method and system using gain controllable clipped histogram equalization. U.S. Patent No 7, 885, 462, p. 1–15.
[24] Han, J.H., S. Yang, and B.U. Lee. 2011. A novel 3-D color histogram equalization method with uniform 1-D gray scale histogram. IEEE Transactions on Image Processing, 20(2), p. 506–512.
[25] Deza, M. M., and E. Deza. 2009. Encyclopedia of Distances. Springer, Berlin. DIGINFO. 2012.