

Cuantificación de movimientos faciales de individuos sanos usando Viola-Jones y Otsu

Javier Caldera-Miguel, Francisco J. Renero-Carrillo

Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE), Puebla, México

{jcaldera, paco}@inaoep.mx

Resumen. La simetría del rostro humano juega un papel importante en las actividades de la vida diaria del individuo. Por otro lado, la asimetría del rostro puede ser sinónimo de enfermedades en el individuo, algunas se manifiestan en reposo y otras en movimiento. En este trabajo cuantificamos cuatro de los movimientos propios del rostro humano. Los movimientos cuantificados son el levantamiento de cejas, cerrar ojos, la sonrisa y el beso. Identificamos los puntos de seguimiento (PS), con Viola-Jones y Otsu, dentro de los tercios o quintos del rostro, dependiente del movimiento. Extraemos el rostro en reposo, que usamos como referencia, se parametriza para segmentar el rostro donde se desarrolla el movimiento facial. Localizamos los máximos desplazamientos (MD), en función de los tercios o quintos, de los PS para cada movimiento. La cuantificación la determinamos por los MD promedio de los PS para la ejecución de 10 repeticiones de un movimiento facial. Presentamos resultados de un grupo de 12 individuos sanos. Los desplazamientos máximos promedios son 30% para cejas dentro del primer tercio, 20% para sonrisa en el segundo y cuarto quinto dentro del tercer tercio, 30% para el beso entre el segundo, tercero y cuarto quinto del tercer tercio; y los 12 individuos logran cerrar los ojos y se desarrolla en el segundo y cuarto quinto dentro del segundo tercio.

Palabras clave. Asimetría facial, movimientos faciales, tercios y quintos del rostro.

Quantification of Facial Movements of Healthy Individuals Using Viola-Jones and Otsu

Abstract. The symmetry of the human face plays an important role in the activities of the individual's daily life. On the other hand, the asymmetry of the face can be synonymous with diseases in the individual, some

manifest at rest and others in movement. In this work we quantify four of the movements of the human face. The quantified movements are lifting eyebrows, closing eyes, smiling and kissing. We identify the tracking points (PS), with Viola-Jones and Otsu, within the thirds or fifths of the face, dependent on the movement. We extract the face at rest, which we use as a reference, it is parametrized to segment the face where the facial movement develops. We locate the maximum displacements (MD), depending on the thirds or fifths, of the PS for each movement. The quantification is determined by the average MD of the PS for the execution of 10 repetitions of a facial movement. We present results from a group of 12 healthy individuals. The maximum average displacements are 30% for eyebrows within the first third, 20% for smile in the second and fourth fifth in the third third, 30% for the kiss between the second, third and fourth fifth of the third third; and the 12 individuals manage to close their eyes and it develops in the second and fourth fifth within the second third.

Keywords. Facial asymmetry, facial movements, thirds and fifths of the face.

1. Introducción

El ser humano es asimétrico, podemos tener un pie más grande que el otro o miembros con longitudes diferentes, en la mayoría de los casos imperceptibles al sentido de la visión [1-4]. Pero la asimetría en el rostro humano juega un papel fundamental en las actividades de la vida diaria del individuo, en estado de reposo al comparar el lado derecho con el izquierdo podemos observar cejas u ojos con dimensiones diferentes [enfermedades]. Si observamos con detalle,

veríamos comisuras de ojos o boca con alturas diferentes entre comisuras equivalentes.

Algunas de estas asimetrías faciales son de origen natural, genética, por gesticulación o patologías [5-7].

Por otro lado, la literatura es vasta en el estudio de la antropometría facial, que va desde el uso de cintas métricas, vernier y el uso de tecnologías digitales para simetría facial en reposo, o para asimetrías faciales por patologías [8, 9]. Sin embargo, se da por obvio la cuantificación de lo normal, es decir, cual es el grado de simetría o asimetría del rostro en individuos sanos y mucho menos estudios para los movimientos propios del rostro, sin gesticular, sin pose y sin patología [10].

Los algoritmos de visión artificial permiten identificar y/o reconocer personas, objetos o características de eventos [11, 12]. En este trabajo implementamos el algoritmo de Viola- Jones para identificar el rostro por cada cuadro en un video, así como los ojos, la nariz y la boca [13]. Implementamos filtros de procesamiento puntual y de vecindad [12, 14], para parametrizar el rostro en reposo e identificar los puntos de seguimiento, tales como son los máximos desplazamientos de las cejas y las comisuras de la boca en los respectivos movimientos.

Para determinar los valores de los desplazamientos máximos de los puntos de seguimiento de cada movimiento, dividimos el rostro proporcionalmente en tercios y quintos. Horizontalmente lo dividimos en tercios y verticalmente en quintos. Cada tercio y quinto del rostro es dividido en 10 partes, así la evaluación de los MD se realiza por segmentos (tercios, quintos, y/o tercios con quintos) de cada cuadro del video para cada movimiento. Para obtener el promedio de los máximos desplazamientos, se adquiere video de 12 voluntarios ejecutando 10 repeticiones de un mismo movimiento facial y para los cuatro movimientos.

Los desplazamientos máximos promedios son 30% para cejas dentro del primer tercio, 20% para sonrisa en el segundo y cuarto quinto dentro del tercer tercio, 30% para el beso entre el segundo, tercero y cuarto quinto del tercer tercio; y los 12 individuos logran cerrar los ojos y se desarrolla en el segundo y cuarto quinto dentro del segundo tercio.

2. Metodología

2.1. División del rostro humano por tercios y quintos

Los profesionales de la estética facial, así como los cirujanos plásticos, estiman las medidas del rostro, en reposo, mediante la división por tercios y quintos. La figura 1 esquematiza esta técnica, que consiste en dividir proporcionalmente, el rostro humano en tres partes horizontales y cinco partes verticales. El primer tercio inicia en *trichion* hasta la *glabella*, el segundo tercio inicia donde termina el primer tercio hasta la *columnela*, el tercer tercio inicia donde termina el segundo tercio hasta la *mitad de la sínfisis*.

Los quintos se determinan por la distancia *inter-comisuras* (IC) de los ojos, tal como se muestra en la figura 1, entonces estas distancia determinan el segundo y cuarto quinto, mientras que el tercer quinto es determinado por la *comisuras internas* de cada ojo. Los quintos primero y quinto, quedan determinados por las *comisuras externas* de cada ojo y los bordes externos de los *lóbulos auriculares* correspondientes.

2.2. Identificación de los puntos de seguimiento

El movimiento de levantamiento de cejas se ejerce en el primer tercio, los puntos de seguimientos se encuentran en el segundo y cuarto quinto, como se muestran en la figura 2. El movimiento de la sonrisa y el beso se llevan a cabo en el tercer tercio, los puntos de seguimiento son las comisuras externas de la boca y éstas se encuentran en el 2 y 4 quinto, como se muestra en la figura 2. El movimiento de cerrar ojos se ejecuta en el segundo tercio, en este movimiento los puntos de seguimiento son las pupilas de cada ojo (observe figura 2).

2.3. Cuantificación de movimientos faciales

Grabamos un video para cada uno de los movimientos faciales con una cámara web, donde el voluntario y la cámara están alineados [15] Adquirimos 5 segundos de video en reposo, con el tiempo suficiente para ejecutar 10 repeticiones de

cada movimiento (levantar cejas, cerrar ojos, sonreír y el beso).

La figura 3a muestra uno de los cuadros de los 5 segundos de reposo. Con el algoritmo de Viola-Jones registramos los ojos, la nariz y la boca, con esta información determinamos lo alto y lo ancho del rostro, de los ojos, la nariz y la boca, como se muestra en la figura 3b. Con estos valores determinamos los tercios y los quintos del rostro (ver figura 3c).

2.3.1. Determinación de los máximos desplazamientos de los puntos de seguimiento

Cuantificación del desplazamiento de levantar cejas

La cuantificación del movimiento de levantamiento de ceja, se lleva a cabo en el primer tercio por lo que en los cuadros de cada movimiento aplicamos el filtro de escala de grises, obtenemos su histograma de frecuencias de intensidad, aplicamos el algoritmo de Otsu para obtener el umbral óptimo y convertimos la imagen a blanco y negro. Esta secuencia de filtros muestra la cavidad de los ojos con tonalidades en negro, calculamos el centro de los ojos a partir de las coordenadas del rostro en reposo, hacemos un barrido vertical hacia el *trichion* hasta que encontrar un pixel y sus ocho vecinos de color blanco, ésta representa el punto de máximo desplazamiento de las cejas. Esta zona la empatamos con las divisiones del primer tercio y obtenemos el máximo desplazamientos promedios para las repeticiones. En la figura 4 mostramos el primer tercio, así como las 10 subdivisiones.

Determinación de ojo cerrado

El cerrar y abrir ojos es un movimiento propio del rostro humano, desde el punto de vista computacional solo determinamos la ejecución de cerrar ojos. Del video de cerrar ojos a cada uno de los 10 cuadros, aplicamos el algoritmo de Otsu y el filtro de binarización. De las coordenadas de los ojos obtenidas del rostro de referencia, dividimos la cavidad de los ojos en una rejilla de 3 x 3 tal como se muestra en la figura 5. Hacemos un barrido por cada subregión para acumular el número de pixeles negros, si las regiones ($f2,c1$),

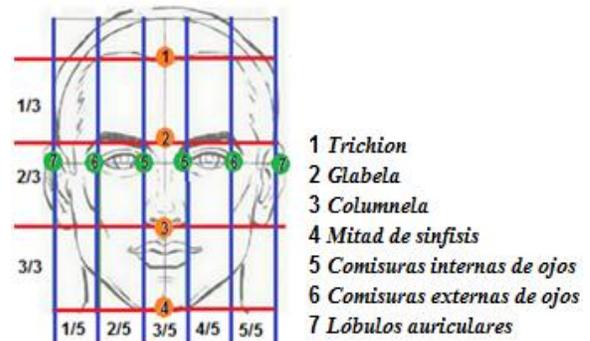


Fig. 1. División del rostro humano por tercios y quintos. Se muestran los puntos anatómicos de referencia

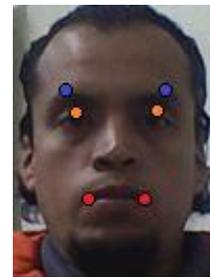


Fig. 2. Puntos de seguimiento de los movimientos, en azul para el levantamiento de cejas, naranja para cerrar ojos y para la sonrisa y el beso de color rojo

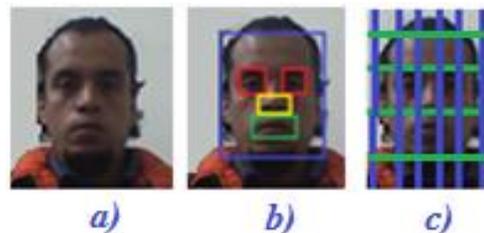


Fig. 3. a) Cuadro del rostro en reposo, b) Identificación de los ojos, la boca y la nariz por Viola-Jones y c) división de tercios y quintos de rostro en estado de reposo

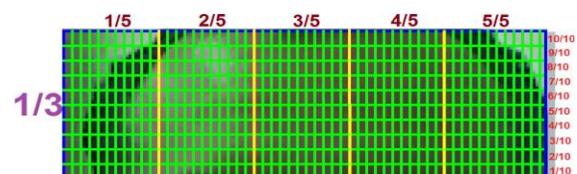


Fig. 4. Primer tercio del rostro humano para cuantificar los desplazamientos máximos de cejas

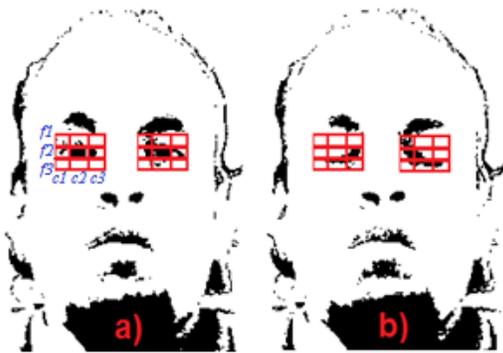


Fig. 5. Esquema para determinar la ejecución de ojos a) si las regiones $(f2,c1)$, $(f2,c2)$ y $(f2,c3)$ son diferentes de cero entonces los ojos están abiertos. b) en caso contrario están cerrados

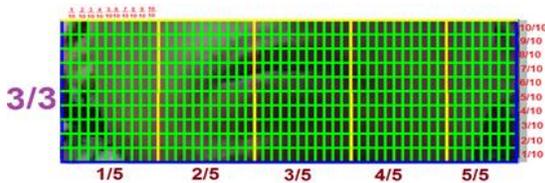


Fig. 6. Tercer tercio del rostro humano para cuantificar el desplazamiento de las comisuras de la boca. Se muestran los puntos de seguimiento para la boca del rostro en reposo

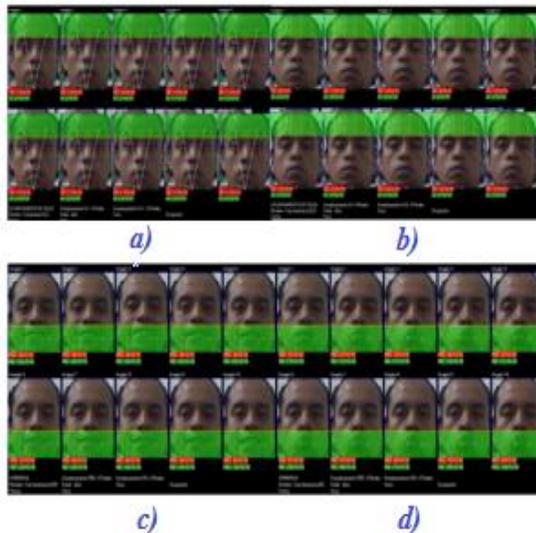


Fig. 7. Resultados del procesado digital de los 4 movimientos de un individuo, a) levantamiento de cejas, b) cerrar ojos, c) sonrisa, d) beso

$(f2,c2)$ y $(f2,c3)$ son diferentes de cero entonces los ojos están abiertos y en caso contrario si al menos una región del centro de la matriz es igual a cero entonces los ojos están cerrados, como se muestra en la figura 6.

Cuantificación del desplazamiento de los movimientos de la sonrisa y el beso

Del video de los movimientos de sonrisa y el beso, para cada cuadro aplicamos la conversión a escala de grises, al histograma de distribución de frecuencias le aplicamos el algoritmo de Otsu y de binarización. Con esto determinamos el centro de la región de la boca, a partir de las coordenadas del rostro en reposo hacemos un barrido de la posición del centro hacia las comisuras de la boca hasta encontrar un pixel y sus 8 vecinos de color blanco.

Este valor es el desplazamiento máximo de las comisuras para los movimientos de sonrisa y beso respectivamente. Empatamos este valor con las subdivisiones del tercer tercio obteniendo así el máximo desplazamientos (ver figura 6).

2.4. Máximos desplazamientos

Cuantificamos los 4 movimientos faciales a un grupo de 12 individuos, 6 hombres y 6 mujeres, de edades de 20 a 40 años de edad, con actividades que corresponden a la docencia o estudiantes de nivel superior. En la imagen 7 se muestran los resultados del procesamiento digital de los 4 movimientos para una persona.

En la tabla 1 muestra los resultados promedios de desplazamiento de los movimientos de levantar cejas, cerrar ojos, sonrisa y el beso para el grupo de 12 individuos. M_i y H_i representan a las 6 mujeres y 6 hombres voluntarios, respectivamente. Cada renglón representa el promedio de las 10 repeticiones de cada uno de los movimientos. En el caso de las comisuras de la boca, el renglón por debajo del promedio de las comisuras representa el quinto en donde se ejecuta el movimiento (sonrisa o beso). En el caso de los ojos 0 y 1 representan abierto y cerrado, respectivamente.

Para el movimiento de levantamiento de cejas determinamos que los máximos desplazamientos de los puntos de seguimiento se ejecutan, en promedio, en el 3/10 del primer tercio.

Tabla 1. Promedios de los movimientos de levantar cejas, cerrar ojos, sonrisa y beso de un grupo de 12 individuos (6 mujeres y 6 hombres). En el movimiento de cerrar ojos 1 corresponde a logrado y 0 a no logrado

Mujeres						
Levantar Cejas						
	M1	M2	M3	M4	M5	M6
Ceja Der.	2.9/10	3.0/10	3.2/10	3.3/10	3.5/10	3.2/10
Ceja Izq.	3.0/10	3.2/10	3.0/10	3.1/10	3.4/10	3.2/10
Cerrar ojos						
Ojo Der.	1	1	1	1	1	1
Ojo Izq.	1	1	1	1	1	1
Sonrisa						
Comisura Der.	4.1/10 2/5	3.9/10 2/5	4.3/10 2/5	4.5/10 2/5	4.0/10 2/5	4.2/10 2/5
Comisura Izq.	5.2/10 4/5	5.5/10 4/5	5.5/10 4/5	5.6/10 4/5	5.4/10 4/5	5.7/10 4/5
Beso						
Comisura Der.	9.7/10 2/5	9.8/10 2/5	9.7/10 2/5	9.9/10 2/5	10.0/10 2/5	9.8/10 2/5
Comisura Izq.	1.2/10 4/5	1.1/10 4/5	1.3/10 4/5	1.2/10 4/5	1.4/10 4/5	1.5/10 4/5
Hombres						
Levantar Cejas						
	H1	H 2	H3	H 4	H 5	H6
Ceja Der.	2.8/10	2.9/10	3.2/10	3.1/10	3.4/10	3.2/10
Ceja Izq.	3.0/10	3.2/10	3.0/10	3.1/10	3.4/10	3.2/10
Cerrar Ojos						
Ojo Der.	1	1	1	1	1	1
Ojo Izq.	1	1	1	1	1	1
Sonrisa						
Comisura Der.	4.5/10 2/5	4.0/10 2/5	4.3/10 2/5	4.4/10 2/5	4.4/10 2/5	4.3/10 2/5
Comisura Izq.	5.5/10 4/5	5.4/10 4/5	5.5/10 4/5	5.6/10 4/5	5.5/10 4/5	5.7/10 4/5
Beso						
Comisura Der.	9.6/10 2/5	9.6/10 2/5	9.8/10 2/5	9.7/10 2/5	9.7/10 2/5	9.9/10 2/5
Comisura Izq.	1.0/10 4/5	1.4/10 4/5	1.4/10 4/5	1.3/10 4/5	1.1/10 4/5	1.4/10 4/5

Para el movimiento de la sonrisa se ubican, en promedio, en el 2/10 del 2/5 y 8/10 del 4/5, del tercer tercio.

En el movimiento del beso se encuentran en el 9/10 del 2/5 y en el 2/10 del 4/5, del tercer tercio. El movimiento de cerrar ojos solo se determina si los ojos están cerrados o abiertos.

2.5. Comparación con dos metodologías para la identificación de expresiones del rostro en base a movimientos faciales

Para identificar y clasificar las expresiones faciales primero se identifican los puntos anatómicos del rostro y posteriormente se analizan los movimientos de estos, de acuerdo a cada movimiento o combinación de movimientos. Existen varias metodologías para dar seguimiento a los puntos anatómicos, pero dos son las más relevantes y utilizadas para la identificación de emociones.

La primera lleva por nombre Codificación Facial, esta consiste en identificar 34 puntos anatómicos del rostro, 2 por cada ceja, 3 por cada ojo (comisuras internas, externas y centro), 3 puntos en la nariz (centro y extremos), 8 puntos en la boca (2 comisuras, 3 para el labio superior y 3 para el labio inferior) y 13 puntos (tomados desde los lóbulos de las orejas y rodeando a la barbilla) [16].

Cabe destacar que esta identificación se realiza por medio de una plantilla que tiene un tamaño de 96x96 pixeles, es decir cada imagen se escala a esta medida y se implementa el algoritmo de histogramas por gradiente, y dividen la imagen en 3 x 3, se analiza cada región para identificar cada punto y posteriormente se hace uso de una base de datos de entrenamiento en donde tienen más de 15000000 de muestras, hacen una comparativa para identificar los movimientos de las comisuras y posteriormente se determina la expresión facial.

El segundo método lleva por nombre vectores de distancias para medición de intensidad de expresiones faciales, éste, consiste en identificar primero 83 puntos de referencia facial obtenidas de una base de datos (son todos los puntos que rodean a cabeza, ojos, nariz y boca) y posteriormente se identifican 11 puntos anatómicos 1 punto para la ceja izquierda, 4

puntos para el ojo derecho (comisuras externa, interna punto superior e inferior), 2 puntos de las orejas (lóbulos inferiores), 4 puntos para boca (Comisuras Izquierda, derecha, arriba y abajo) [17].

Posteriormente se colocan en un vector los desplazamientos que deben de tener estos puntos de acuerdo a un movimiento o conjunto de movimientos según sea una expresión (en este caso 6). Para reconstruir el rostro normalizan pares de los 83 puntos que describen al rostro, una vez realizada la reconstrucción en 3D y en base a los datos de clasificación se determinan si los movimientos de los 11 puntos pertenecen a una expresión.

Los resultados obtenidos por estos dos métodos indican si los movimientos de los puntos anatómicos del rostro están desplazados de acuerdo a patrones ya determinados y almacenados. Sin embargo no indican cuanto se están desplazando (unidades de medidas) los puntos anatómicos, no hace una cuantificación como tal, los algoritmos implementados utilizan muchos puntos por lo que el procesamiento y la generación de resultado es más tardado, además las búsquedas se hacen por comparación con otros rostros y no con la simetría del mismo rostro, estos algoritmos deben de hacer una conexión con servidores o computadoras para poder hacer la clasificación ya que los algoritmos de identificación y procesamiento digital de imágenes son muy pesados para hacer en dispositivos móviles.

La propuesta que nosotros damos es posible implementarla en dispositivos con poca capacidad ya que los algoritmos implementados no analizan toda la imagen solo zonas.

3. Conclusiones

Con Viola-Jones y Otsu identificamos y determinamos los puntos de seguimiento para la cuantificación de los máximos desplazamientos de movimientos faciales de individuos sanos.

El individuo tiene movimientos simétricos para las cejas, las comisuras de la boca y cerrar ojos.

Los desplazamientos máximos promedios son 30% para cejas dentro del primer tercio, 20% para sonrisa en el segundo y cuarto quinto dentro del tercer tercio, 30% para el beso entre el segundo,

tercero y cuarto quinto del tercer tercio; y los 12 individuos logran cerrar los ojos y se desarrolla en el segundo y cuarto quinto dentro del segundo tercio.

Referencias

- Codari, M., Pucciarelli, V., Stangoni, F., & Zago, M. (2016).** Facial thirdse-based evaluation of facial asymmetry using. *Journal of Cranio-Maxillo-Facial Surgery*, Vol. 45, pp. 76–81. DOI: 10.1016/j.jcms.2016.11.003.
- Bock, M. T. & Bowman, A. W. (2006).** On the measurement and analysis of asymmetry. *Royal Statistical Society*, Vol. 55, pp. 77–91. DOI: 10.1111/j.1467-9876.2005.00532.x.
- Thornhill, R. & Gangestad, S. W. (1999).** Facial attractiveness. *Trends in Cognitive Sciences*, Vol. 3, pp. 452–460. DOI: 10.1016/S1364-6613(99)01403-5.
- Komori, M., Kawamura, S., & Ishihara, S. (2009).** Averageness or symmetry: Which is more important for facial attractiveness?. *Acta Psychologica*, Vol. 131, pp. 136–142. DOI: 10.1016/j.actpsy.2009.03.008.
- Schmid, K., Marx, D., & Samal, A. (2008).** Computation of a Face Attractiveness Index Based on Neoclassical Canons, Symmetry, and Golden Ratios. *Journal Articles*, Vol. 92. DOI: 10.1016/j.patcog.2007.11.022.
- Lee, A. J., Mitchemb, D. G., Wright, M. J., Nicholas, G. M., Kellerbc, M. C., & Zietschad, B. P. (2016).** Facial averageness and genetic quality: testing heritability, genetic. *Evolution and Human Behavior*, Vol. 37, No. 1, pp. 61–66. DOI: 10.1016/j.evolhumbehav.2015.08.003.
- Henderson, A. J., Holzleitner, I. J., Talamas, S. N., & Perrett, D. I. (2015).** Perception of health from facial cues. *Philosophical Transactions B*, Vol. 371, pp. 1–9. DOI: 10.1098/rstb.2015.0380.
- Santosa, K. B., Fattah, A., Gavilán, J., Hadlock, T. A., & Snyder-Warwick, A. K. (2017).** Photographic Standards for Patients with Facial Palsy and Recommendations by Members of the Sir. Charles Bell Society. *JAMA Facial Plastic Surgery*, Vol. 19, No. 4, pp. 275–281. DOI: 10.1001/jamafacial.2016.1883.
- Zago, M., Tarabbia, F., Bassetti, L., Biglioli, F., & Sforza, C. (2017).** Photographic technique for the quantitative assessment of lagophthalmos and eyelid position in patients with facial palsy. *Br. J. Oral Maxillofac Surg*, Vol. 55, No. 5, pp. 548–550, DOI: 10.1016/j.bjoms.2017.01.019.
- Penkea, L., Bates, T. C., Gow, A. J., Pattie, A., Starr, J. M., Jones, B. C., Perrett, D. I., & Deary, J. (2009).** Symmetric faces are a sign of successful cognitive aging. *Evolution and Human Behavior*, Vol. 30, No. 6, pp. 429–437. DOI: 10.1016/j.evolhumbehav.2009.06.001.
- Sadrô, J., Jarudi, I., & Sinhaô, P. (2003).** The role of eyebrows in face recognition. *Perception*, Vol. 32, pp. 285–293. DOI: 10.1068/p5027.
- Gonzalez, R. G. & Woods, R. E. (2008).** *Digital Image Processing*. Upper Saddle River, New Jersey: Pearson Prentice Hall.
- Viola, P. & Jones, M. J. (2001).** Rapid object detection using a boosted cascade of simple features. *Conference on Computer Vision and Pattern Recognition*, pp. 511–518. DOI: 10.1109/CVPR.2001.990517.
- Otsu, N. (1979).** A threshold selection method from gray-level histogram. *IEEE Transactions on System, Man, Cybernetics*, Vol. SMC-9, No. 1, pp. 62–66.
- Romero, R. & Renero, F. (2016).** *Evaluación de herramientas digitales para el desarrollo de la antropometría facial*. Tesis Doctoral. Instituto Nacional de Astrofísica, Óptica y Electrónica.
- McDuff, D. & Kaliouby, R. (2017).** Applications of Automated Facial Coding in Media Measurement. *IEEE transactions on affective computing*, Vol. 8, No. 2, pp. 148–160. DOI: 10.1109/TAFFC.2016.2571284.
- Chiek-Ying, A. C., Ujir, H., & Hipiny, I. (2017).** 3d facial expression intensity measurement analysis. *Proceedings of the 6th International Conference on Computing and Informatics (ICOCI'17)*, pp. 43–48.

Article received on 05/06/2017; accepted on 11/10/2017.
Corresponding author is Javier Caldera-Miguel.