

Análisis de Simetrías y Proporciones Faciales mediante Visión Estereoscópica

Analysis of Symmetries and Facial Proportions through Stereoscopic Vision

Análise de simetrias e proporções faciais por meio de Visão estereoscópica

José Orlando Maldonado Bautista¹, Luis Armando Portilla Granados²

Facultad de Ingenierías y Arquitectura, Departamento de Ingenierías Eléctrica, Electrónica, Sistemas y Telecomunicaciones,

Universidad de Pamplona, Pamplona, Colombia.

¹ormaldonado@unipamplona.edu.co, ²laporti@unipamplona.edu.co

Recibido / Received: 11/08/2015 – Aceptado / Accepted: 01/12/2015

Resumen

En el siguiente documento se explora el uso de técnicas de visión estereoscópica para el análisis de simetrías y proporciones faciales. Inicialmente, se describen algunos puntos del rostro, medidas y proporciones entre ellos, los cuales son de interés en ciertas áreas de la salud que usan la antropometría facial, como es el caso de la fonoaudiología. Posteriormente, se hace una exposición de los fundamentos y técnicas de la visión estéreo utilizados. Mediante el uso de dichas técnicas, se diseña un prototipo de software que permite seleccionar puntos de interés sobre la imagen, las cuales son tomadas por un par de cámaras sintonizadas en estéreo, para realizar los cálculos mediante los algoritmos apropiados. Los algoritmos nos permiten determinar las distancias entre los puntos y las proporciones requeridas. Se hacen pruebas del prototipo tanto en objetos de geometría simple, como en la geometría facial. Los resultados son analizados y se hace una discusión de las posibilidades de la herramienta desarrollada.

Palabras clave: Visión estéreo, antropometría facial, visión por computador.

Abstract

In the following document the use of stereoscopic vision techniques to analyze symmetry and facial proportions is explored. Initially some points of the face, measurements and proportions between them, which are of interest in certain health areas using facial anthropometry, such as phonoaudiology. Subsequently, a presentation of the principles and techniques of stereo vision used is done. By using these techniques a prototype of software that let you select points of interest in the image is designed, which are captured by a pair of stereo cameras tuned to perform the calculations through appropriate algorithms. The algorithms allow us to determine the distances between points and the required proportions. Prototype

testing is performed in both objects of simple geometry and facial geometry. The results are analyzed and a discussion of the possibilities of the tool is developed.

Keywords: Stereoscopic vision, facial anthropometry, computer vision.

Resumo

O seguinte documento descreve o uso de técnicas de visão estereoscópica para análise de simetrias e proporções faciais. Inicialmente, alguns pontos da cara, medidas e proporções entre eles, que são de interesse em certas áreas da saúde com uso de antropometria facial, como é o caso da fonoaudiologia. Posteriormente, uma apresentação dos princípios e técnicas de visão estéreo usado. Ao utilizar estas técnicas, é concebido um protótipo de software que permite que selecione pontos de interesse sobre a imagem, que são tomadas por um par de câmeras sintonizados em estéreo para executar cálculos usando algoritmos apropriados. Os algoritmos nos permite determinar as distâncias entre os pontos e as proporções necessárias. testes do protótipo são feitos nos objetos de geometria simples, como na geometria facial. Os resultados são analisados e uma discussão sobre as possibilidades da ferramenta desenvolvida.

Palavras-chave: visão estéreo, antropometria facial, visão por computador.

I. INTRODUCCIÓN

La antropometría es la ciencia que estudia las medidas de tamaño, peso y proporciones del cuerpo humano. La utilización de la antropometría para describir las proporciones de la cara, data del siglo XVIII. Históricamente, dicha ciencia ha sido usada como un instrumento en la clínica que pretende añadir datos cuantitativos para evaluar el rostro humano. En 1964, Burian et al. [1] empezaron a estandarizar medidas en pacientes con deformidades craneofaciales adquiridas y congénitas, razón por la cual, lograron ser muy reconocidos en el mundo científico.

Gracias al interés de reconocidos autores, la antropometría volvió a tomar fuerza desde hace cuatro décadas aproximadamente. En áreas específicas de la salud, como la fonoaudiología, el registro de medidas concretas de la morfología orofacial para conformar bases de datos con parámetros normalizados, es una práctica que permite la posterior utilización en la determinación de diagnósticos fonoaudiológicos, planteamiento terapéutico y realización de pronósticos clínicos. Aún son utilizados para la toma de las medidas, instrumentos ancestrales tradicionales como el parquímetro, que dependen en gran parte del entrenamiento y habilidades del especialista.

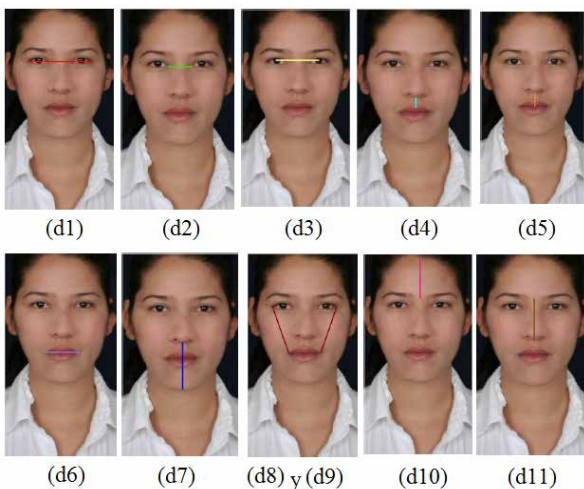
Sin embargo, es cada vez más frecuente el uso exitoso de instrumentos apoyados en la tecnología, para facilitar las labores de los especialistas. Las técnicas de reconocimiento de rostro, apoyadas en la visión por computador, han demostrado gran efectividad en problemas como la identificación de individuos, y las técnicas desarrolladas son susceptibles de aplicar a la inspección detallada del rostro para la toma de medidas; sin embargo, problemas como la iluminación, la oclusión, la escala y el ángulo de observación aún son asuntos por resolver. El problema de la escala puede ser abordado apoyado en la capacidad de la visión estereoscópica para determinar distancias emulando la visión humana. El desarrollo de una herramienta para el análisis automático de la antropometría facial puede contribuir de manera positiva en la labor médica en general.

II. SIMETRÍA Y PROPORCIONES FACIALES

Los puntos de interés tenidos en cuenta para el análisis de proporciones y simetría facial son los sugeridos, según lo consultado en [2]. En el caso de la simetría, se traza una recta vertical sobre la imagen del rostro, de tal forma que pase sobre el punto medio entre los centros de las pupilas. Se calcula la distancia de los pares correspondientes al lado izquierdo y derecho del rostro y se calcula

la razón media como medida de simetría. El rostro ideal debería tener medida uno. Para el cálculo de proporciones, el procedimiento a seguir es calcular las distancias entre los puntos de interés sugeridas por [2], y normalizar las mismas dividiendo por la distancia entre las pupilas. Los resultados obtenidos de forma automática por el sistema son comparados con los resultados obtenidos manualmente. La Fig. 1 muestra las distancias calculadas a partir de los puntos de interés localizados, las cuales se describen a continuación:

- d1: Distancia entre las comisuras externas de los ojos.
- d2: Distancia entre las comisuras internas de los ojos.
- d3: Distancia entre las Pupilas.
- d4: Distancia entre el punto subnasal y el labio superior.
- d5: Distancia del punto subnasal al punto ubicado en la línea central entre los labios.
- d6: Ancho de la Boca.
- d7: Distancia entre el subnasal a la barbilla.
- d8 y d9: Distancia entre los ojos y la boca.
- d10: Distancia del punto ubicado en la parte superior de la frente y el ubicado sobre el entrecejo.
- d11: Distancia del punto sobre el entrecejo y el subnasal.



Fuente: Los Autores, 2014

Fig. 1. Distancias tenidas en cuenta para la estimación de proporciones sobre el complejo orofacial.

III. VISIÓN ESTEREOSCÓPICA

A. Generalidades

La visión estereoscópica constituye un método para la obtención de la forma de los objetos en la escena. Esto se consigue calculando la distancia de dichos objetos a un sistema de referencia. De esta manera, se constituye como un método para la obtención de la tercera dimensión.

Entre las técnicas desarrolladas para la obtención de la distancia, se pueden mencionar [3]:

- Los sensores de ultrasonidos, que utilizan la energía acústica.
- Mediante técnicas de triangulación, primero se emite una fuente de luz sobre un objeto, que al ser reflejado es percibido por un detector. Por triangulación entre el emisor y el receptor, se obtiene la distancia del objeto.
- Mediante luz estructurada, proyectando un haz de luz (ej. láser), se capta la distorsión de la luz mediante una cámara de visión, para así obtener el alcance.
- Mediante telemetría de tiempo de vuelo, se determina la distancia a un objeto midiendo el tiempo que tarde en retornar una señal láser o ultrasonido.
- Mediante visión estereoscópica artificial, se emplean dos o más cámaras para capturar la misma escena y, por triangulación, determinar la distancia.

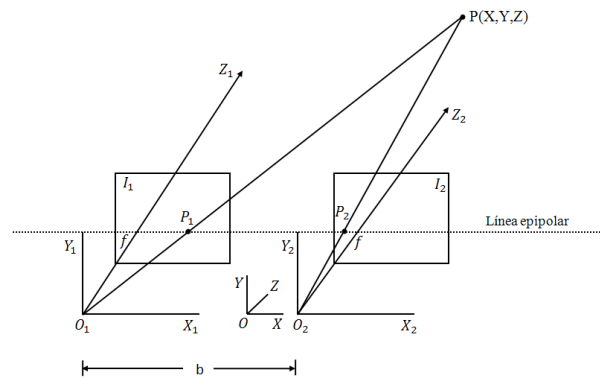
La visión estereoscópica artificial toma como referencia el modelo estereoscópico biológico, en el cual, el desplazamiento relativo de los ojos permite obtener la profundidad de los objetos o tercera dimensión, mediante un proceso de triangulación a partir de dos imágenes generadas del mismo objeto en la escena 3D. En visión estereoscópica artificial, se utilizan dos o más cámaras, separadas entre sí por una cierta distancia. El proceso consiste en captar una imagen de la misma escena por cada cámara, y por triangulación, como en el caso del modelo biológico, se puede encontrar la distancia a la que se encuentra determinado objeto. La obtención de las dos imágenes ligeramente desplazadas se puede ob-

tener, o bien alineando dos o más cámaras de forma que se encuentren levemente desplazadas en el espacio, o desplazando una única cámara a una cierta distancia y captando las imágenes en las diferentes posiciones de desplazamiento.

En cualquier caso, la geometría del sistema se puede diseñar de manera que las cámaras tengan sus ejes ópticos paralelos o convergentes. En el caso del sistema humano, este tiene la capacidad de ajustar sus ejes de forma tal que para objetos lejanos, dichos ejes sean aproximadamente paralelos. Para ver algo cercano, los ejes convergen hasta llegar a intersectarse. Se pueden encontrar en la literatura diferentes propuestas que buscan emular la capacidad del sistema visual humano, mediante la movilidad de los ejes de las cámaras, permitiendo enfocar las zonas de interés en cada momento. En este trabajo, se estudia el caso en el que los ejes son fijos. En las secciones siguientes se revisará la geometría de un sistema estereoscópico, tanto para ejes convergentes y no convergentes, se revisan los principales enfoques para el proceso de calibración de cámaras y para la obtención de la tercera dimensión mediante triangulación.

A. Geometría de un Sistema Estereoscópico

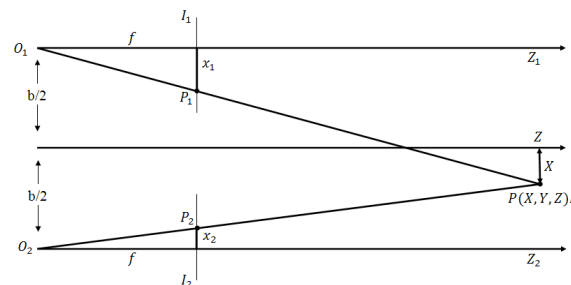
La geometría de un sistema estereoscópico está estrechamente relacionada con la forma de obtener la profundidad de los objetos detectados. En el caso convencional, el sistema está conformado por dos cámaras, con sus ejes ópticos en paralelo y separados una cierta distancia horizontal, la cual se denomina línea base b . Los ejes ópticos de las cámaras son perpendiculares a la línea base, y las líneas epipolares o líneas de exploración son paralelas a esta. Sobre las líneas epipolares, deben coincidir los puntos en las imágenes derecha e izquierda que corresponden a un punto en la escena. Dado que el desplazamiento de los ejes ópticos de las cámaras en este sistema es horizontal, las imágenes de un punto de la escena, captado por cada cámara, difieren solo en la coordenada horizontal de cada imagen. La Fig. 2 muestra la geometría del sistema descrito.



Fuente: Los Autores, 2014

Fig. 2. Geometría de un sistema de visión estereoscópica con ejes ópticos paralelos.

En la figura I_1 e I_2 representan los planos de imagen de cada cámara; y O_1 y O_2 , sus respectivos centros de proyección. El origen del sistema de coordenadas del mundo (sistema de referencia) está en O , la longitud focal efectiva de cada cámara es f . Los ejes de coordenadas del mundo (X , Y y Z) se encuentran entre los ejes de las cámaras. Dado un punto en el espacio, con coordenadas $P(X; Y; Z)$, este produce las proyecciones $P_1(x_1; y_1; z_1)$ y $P_2(x_2; y_2; z_2)$ en el plano de cada imagen. Los rayos PO_1 y PO_2 definen en el plano epipolar, cuya intersección con el plano de la imagen definen la línea epipolar. Para cada par de puntos $P_1(x_1; y_1; z_1)$ y $P_2(x_2; y_2; z_2)$ emparejados, se obtiene un valor de disparidad d , donde $d = x_1 - x_2$.



Fuente: Los Autores, 2014.

Fig. 3. Perspectiva superior de la geometría de un sistema de visión estereoscópica con ejes ópticos paralelos.

De la Fig. 3, mediante semejanza de triángulos, es fácil deducir las siguientes expresiones:

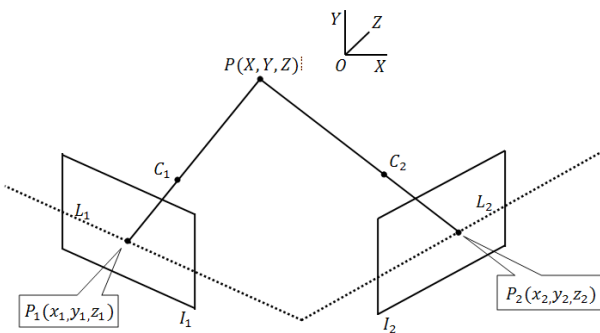
$$O_1 : \frac{\frac{b}{2} + X}{Z} = \frac{x_1}{f} \Rightarrow x_1 = \frac{f}{Z} \left(X + \frac{b}{2} \right), \quad (1)$$

$$O_1 : \frac{\frac{b}{2} + X}{Z} = \frac{x_2}{f} \Rightarrow x_2 = \frac{f}{Z} \left(X + \frac{b}{2} \right), \quad (2)$$

De donde podemos deducir, a su vez:

$$d : x_1 - x_2 = \frac{fb}{Z} \Rightarrow Z = \frac{fb}{d} \quad (3)$$

De esta última expresión, podemos deducir que para un valor de Z fijo, a mayor distancia b , mayor sería el valor en b , lo cual se puede interpretar como mejor precisión en el enfoque. En la Fig. 4 podemos ver el caso en que los ejes ópticos de las cámaras convergen. En forma similar al modelo anterior, un punto P del espacio es proyectado sobre los puntos P_1 y P_2 en el plano de cada imagen. Son estos puntos aquellos en los que encuentran su proyección todos los puntos del espacio de las líneas PP_1 y PP_2 . C_1 y C_2 , son los centros ópticos de cada cámara, de tal forma que P, P_2, C_1 y C_2 se encuentran en el mismo plano (plano epipolar), cuya intersección con I_1 e I_2 determinan la línea epipolar para cada imagen, denotada por L_1 y L_2 . Para determinar las coordenadas (X, Y, Z) de un punto P , a partir de las coordenadas x_p, y_i en las proyecciones de cada imagen, se debe tener en cuenta:



Fuente: Los Autores, 2014.

Fig. 4. Geometría de un sistema de visión estereoscópica con ejes ópticos convergentes.

La calibración de la cámara: se refiere al establecimiento del modelo de la cámara, que

permite la inferencia de las coordenadas de un punto 2D, a partir de un punto 3D en la escena. Este modelo no es más que la transformación que proyecta un punto 3D del mundo real es su correspondiente punto 2D de la imagen.

La transformación perspectiva inversa: consiste en la recuperación de las coordenadas de un punto en la escena 3D, a partir de un punto 2D en la imagen. Con una sola imagen, lo que podemos obtener es la línea de visión que pasa por el punto P y su proyección 2D en la imagen, que en la figura será el punto P_1 . Si se cuenta con una segunda cámara desde donde P sea visible, esta proporciona otra imagen donde se proyecta P , en el punto P_2 , de tal forma que obtenemos una segunda línea de visión. El punto P es el punto en el cual se intersectan las dos rectas. Debido a los posibles errores de digitalización y calibración, dichas rectas pueden no cruzarse, por lo cual el punto P se calcula como aquel que minimiza (en términos de los mínimos cuadrados) la distancia entre las dos rectas.

A. Calibración de las Cámaras

Para poder determinar las coordenadas de un punto en el espacio, dadas las coordenadas de sus proyecciones 2D en cada una de las cámaras, es necesario realizar el proceso de calibración de cada cámara. Mediante la calibración se determina la correspondencia entre los puntos de la imagen y el espacio físico. Esto se realiza mediante la estimación de los parámetros del modelo de cada cámara, los cuales modelan su funcionamiento desde el punto de vista geométrico y la relación existente entre el sistema de referencia del mundo real y el del plano de la imagen. El modelo de la cámara viene dado por las ecuaciones que relacionan los puntos del espacio con su proyección en el plano de la imagen. El modelo estima parámetros propios de la cámara como distancia focal o distorsiones, así como parámetros propios del sistema estereó como la posición y orientación espacial. Mediante la calibración se calculan dichos parámetros y, al ser reemplazados en las ecuaciones del modelo, se obtiene la relación existente en una cámara en particular. En forma simplificada, se puede considerar que el modelo de la cámara y la transformación perspectiva inversa

son lineales. Es decir, no se tienen en cuenta las posibles distorsiones. Otras propuestas como la formulada por Tsai [4], tienen en cuenta la distorsión radial que se produce al formar la imagen, proporcionando un método eficiente y de alta precisión.

1) *Transformación Lineal Directa:* Para realizar la calibración, se debe resolver un sistema de ecuaciones lineales. Si tenemos el par (U, V) representando al par de coordenadas reales de un punto en la imagen, y sean (u, v, t) , las coordenadas homogéneas correspondientes. Estas se relacionan por la siguiente expresión:

$$U = u/t \quad (3)$$

$$V = v/t \quad (4)$$

Definimos el modelo de la cámara a la transformación perspectiva directa C , que está relacionada con un punto 3D del espacio mediante la siguiente expresión:

$$\begin{pmatrix} u \\ v \\ t \end{pmatrix} = C \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (5)$$

$$\begin{aligned} c_{11}x + c_{12}y + c_{13}z + c_{14} - Uc_{31}x - Uc_{32}y - Uc_{33}z - Uc_{34} &= 0 \\ c_{21}x + c_{22}y + c_{23}z + c_{24} - Vc_{31}x - Vc_{32}y - Vc_{33}z - Vc_{34} &= 0 \end{aligned} \quad (9)$$

Para un punto en la imagen del que se conozcan su dos coordenadas (U, V) y las coordenadas correspondientes en el espacio 3D $(x; y; z)$, se tienen dos ecuaciones con 12 incógnitas. Podemos establecer el valor de $c_{34} = 1$, dado que, por la formulación homogénea, no importa la escala, con lo cual quedamos con 11 incógnitas.

La transformación C resulta ser la matriz de dimensiones 3×4 , cuyos componentes plasman la correspondencia entre:

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{pmatrix} \quad (6)$$

La expresión anterior se puede reescribir como un sistema de ecuaciones, con lo cual se obtiene:

$$\begin{aligned} u &= c_{11}x + c_{12}y + c_{13}z + c_{14} \\ v &= c_{21}x + c_{22}y + c_{23}z + c_{24} \\ t &= c_{31}x + c_{32}y + c_{33}z + c_{34} \end{aligned} \quad (7)$$

De las ecuaciones anteriores, se pueden derivar las siguientes expresiones:

$$\begin{aligned} u - Ut &= 0 \\ v - Vt &= 0 \end{aligned} \quad (8)$$

Sustituyendo y reescribiendo las expresiones anteriores, se obtiene el siguiente sistema de ecuaciones:

Se requieren al menos 6 puntos para realizar la calibración y obtener 12 ecuaciones. Completando y reformulando el sistema de ecuaciones en forma matricial cuyas incógnitas son los coeficientes c_{ij} , se tiene:

$$\begin{pmatrix} x^1 & y^1 & z^1 & 1 & 0 & 0 & 0 & -U^1x^1 & -U^1y^1 & -U^1z^1 \\ 0 & 0 & 0 & x^1 & y^1 & z^1 & 1 & -V^1x^1 & -V^1y^1 & -V^1z^1 \\ x^2 & y^2 & z^2 & 1 & 0 & 0 & 0 & -U^2x^2 & -U^2y^2 & -U^2z^2 \\ 0 & 0 & 0 & x^2 & y^2 & z^2 & 1 & -V^2x^2 & -V^2y^2 & -V^2z^2 - V \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x^6 & y^6 & z^6 & 1 & 0 & 0 & 0 & -U^6x^6 & -U^6y^6 & -U^6z^6 \\ 0 & 0 & 0 & x^6 & y^6 & z^6 & 1 & -V^6x^6 & -V^6y^6 & -V^6z^6 \end{pmatrix} \begin{pmatrix} c_{11} \\ c_{12} \\ c_{13} \\ c_{14} \\ c_{21} \\ c_{22} \\ c_{23} \\ c_{24} \\ c_{31} \\ c_{32} \\ c_{33} \end{pmatrix} = \begin{pmatrix} U^1 \\ V^1 \\ U^2 \\ V^2 \\ \cdot \\ \cdot \\ \cdot \\ U^6 \\ V^6 \end{pmatrix} \quad (10)$$

Ahora, si notamos la matriz de coeficientes como X , la matriz de incógnitas como C y a la matriz de términos independientes como Y ; el sistema puede escribirse en la forma:

$$X \cdot C = Y. \quad (11)$$

Para obtener la solución que minimice el error cuadrático, calculamos la matriz pseudo-inversa. Así, podemos hallar C en la forma dada por la siguiente expresión:

$$C = (X^T X)^{-1} X^T Y = X^+ Y \quad (12)$$

Donde X^+ es la matriz pseudoinversa de X .

Por tratarse de un sistema sobre determinado, se requieren mínimo 6 puntos para poder resolverlo, sin embargo, si se toma un conjunto de puntos mayor se conseguirá un mejor resultado en la calibración. Para el proceso de calibración, es habitual utilizar una plantilla en la cual se distinga un gran número de puntos, cuyas posiciones respecto a un sistema de referencia en el espacio 3D son conocidas. Las posiciones de la proyección sobre la imagen 2D a calibrar, junto con las coordenadas en el espacio real, son utilizadas para resolver el sistema de ecuaciones planteado. Los puntos sobre la plantilla deben ser no coplanares y se sugiere tomar diferentes posiciones de la plantilla sobre el campo de visión de la cámara. En Tsai [4], se propone un método de calibración

no lineal que permite mejorar la precisión obtenida con la transformación lineal, mediante el cálculo de la longitud focal, la distorsión radial de la lente, y los demás parámetros intrínsecos (propios de la cámara).

2) *Transformación Perspectiva Inversa:*

Mediante esta transformación, es posible determinar la información 3D a partir de la información obtenida del plano de la imagen 2D. En forma general, se pueden considerar el siguiente método:

Método lineal. Una vez determinado el modelo C de la cámara, es posible determinar las coordenadas en el espacio 3D de un punto dadas sus coordenadas 2D en la imagen. Dado un punto P_p , con coordenadas (U, V) , se puede obtener la línea de visión sobre la que se proyectan todos los puntos sobre el rayo PP_p , siendo P el punto en el espacio 3D.

De las ecuaciones en la sección anterior, podemos obtener:

$$c_{11}x + c_{12}y + c_{13}z + c_{14} = u = Ut \quad (13)$$

$$c_{21}x + c_{22}y + c_{23}z + c_{24} = v = Vt \quad (14)$$

$$c_{31}x + c_{32}y + c_{33}z + c_{34} = t \quad (15)$$

Sustituyendo t en las dos primeras, y reagrupando, podemos obtener:

$$(c_{11} - Uc_{31})x + (c_{12} - Uc_{32})y + (c_{13} - Uc_{33})z + (c_{14} - Uc_{34}) = 0 \quad (16)$$

$$(c_{21} - Vc_{31})x + (c_{22} - Vc_{32})y + (c_{23} - Vc_{33})z + (c_{24} - Vc_{34}) = 0 \quad (17)$$

Si denominamos las expresiones entre paréntesis por $a_1; b_1; c_1; d_1; a_2; b_2; c_2; d_2$; el sistema toma la forma:

$$a_1x + b_1y + c_1z + d_1 = 0 \quad (18)$$

$$a_2x + b_2y + c_2z + d_2 = 0 \quad (19)$$

Las anteriores son las ecuaciones de los planos 2D que, al llegar a intersectarse, nos dan la línea de visión buscada. Al fijar el valor de z , se pueden encontrar las coordenadas (x, y) en 3D reales, solucionando las ecuaciones de la siguiente forma:

$$x = \frac{z(b_1c_2 - b_2c_1) + (b_1d_2 - b_2d_1)}{(a_1b_2) - (a_2b_1)} \quad (20)$$

$$y = \frac{z(a_2c_1 - b_1c_2) + (a_2d_1 - a_1d_2)}{(a_1b_2) - (a_2b_1)} \quad (21)$$

En general, el valor z es desconocido, pero se puede determinar incluyendo otra vista estéreo (par estéreo). Para ello, se utiliza una segunda cámara que registra la misma escena, para determinar una segunda línea de visión, cuya intersección con la primera determina el punto z buscado. Es posible que las líneas no lleguen a la intersección, por lo cual se busca aproximar el punto de cruce mediante mínimos cuadrados.

B. Triangulación y Obtención de la Tercera Dimensión

Mediante la captura de una segunda escena de la imagen con una cámara cuyo modelo también es conocido, es posible calcular los puntos 3D, mediante sus proyecciones, utilizando técnicas de visión estereoscópica binocular. La imagen capturada con la segunda cámara desde otro punto de vista, que contenga la proyección del punto de interés cuyas coordenadas en el espacio se desean calcular, permite incluir dos ecuaciones más al sistema previo, de tal forma que se tiene:

$$a_1x + b_1y + c_1z + d_1 = 0 \quad (22)$$

$$a_2x + b_2y + c_2z + d_2 = 0 \quad (23)$$

$$p_1x + q_1y + r_1z + s_1 = 0 \quad (24)$$

$$p_2x + q_2y + r_2z + s_2 = 0 \quad (25)$$

Siendo este un sistema de 4 ecuaciones, con 3 incógnitas, y dado que se pueden determinar los coeficientes con ayuda de los parámetros conocidos de cada uno de los modelos de las cámaras, el sistema puede ser solucionado mediante el mínimo error cuadrático, que determina las coordenadas (x, y, z) del punto en el que convergen ambas líneas de proyección.

IV. ANÁLISIS ANTROPOMÉTRICO FACIAL MEDIANTE VISIÓN ESTÉREO

A. Materiales y Métodos

Para la realización de las pruebas de análisis del complejo facial, mediante técnicas de visión estereoscópica, se utilizaron los siguientes implementos de tipo hardware y software:

Hardware

- Computador portátil Samsung, Procesador Core i5 1.6 GHz, Memoria RAM 6 Gb.
- (2) Webcam Genius Eye 13, Sensor CMOS, resolución máxima de 640 x 480 (VGA).

Software

- Sistema Operativo Windows 7 Ultimate.
- Software de desarrollo: MATLAB Versión 7.10.0.499 (R2010a)
- Camera Calibration Toolbox for Matlab (Update July 9. 2010)

La Toolbox para calibración de cámaras, desarrollada por Jean-Yves Bouguet, del Instituto Tecnológico

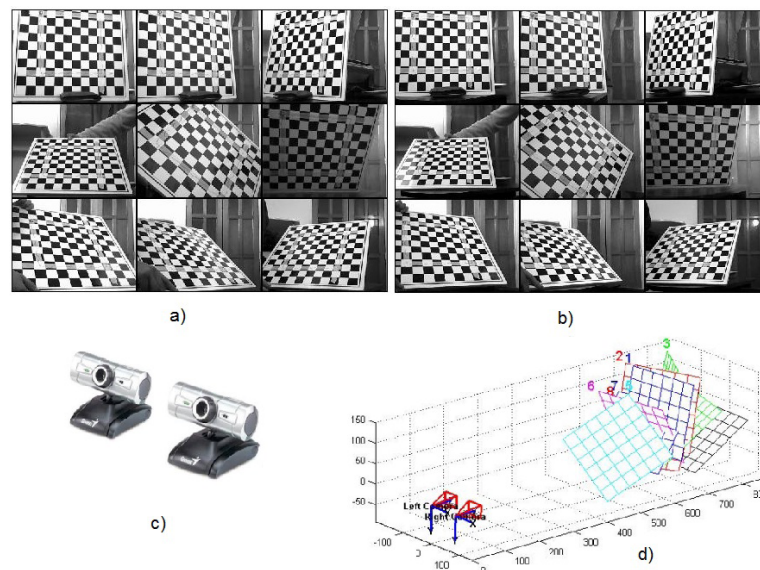
co de California, está conformada por un conjunto de herramientas que permiten establecer los parámetros tanto para el modelo extrínseco como el modelo intrínseco de la cámara, la calibración de pares en estéreo y la determinación de puntos en el espacio mediante triangulación estéreo. Los algoritmos desarrollados están principalmente basados en las propuestas para calibración realizadas en [4]-[6].

A. Pruebas Preliminares

Inicialmente, se realiza la calibración de cámaras; y posteriormente, se realizan pruebas con figuras de geometría simple para verificar la precisión en la calibración.

1) *Calibración de las Cámaras:* las cámaras son dispuestas en paralelo, sobre el mismo eje

horizontal, con una separación de 8 cm. entre los lentes. Para el proceso de calibración, se elaboró un tablero ajedrezado con rectángulos de 30 milímetros en blanco y negro, que será fotografiado por ambas cámaras en diferentes posiciones. Inicialmente, se realiza el proceso de calibración de cada cámara por separado para obtener el modelo de cada una de ellas. Posteriormente, se realiza la calibración estéreo, mediante la cual se optimiza los parámetros de las cámaras respecto a los modelos de referencia de cada una. Los detalles sobre el proceso de optimización son descritos en la página de la Camera Calibration Toolbox. La Fig. 5 muestra la disposición de las cámaras, el mosaico de imágenes tomadas para las calibraciones y el resultado de la calibración.



Fuente: Los Autores, 2014.

Fig. 5. Proceso de calibración de cámaras. a) Vista cámara izquierda tablero ajedrezado b) Vista cámara derecha tablero ajedrezado c) Cámaras web en paralelo d) Resultado del ajuste de parámetros extrínsecos de los pares de cámaras.

2) *Verificación sobre Objetos de Geometría Simple:* para determinar la exactitud conseguida con el proceso de calibración, se han capturado imágenes de objetos con una geometría simple, sobre las cuales se puede realizar una sencilla medición manual entre distancias de puntos específicos como aristas, bordes, esquinas, etc. Luego, se capturan las coordenadas de los puntos de cada imagen en cada

cámara, correspondientes a las proyecciones de cada punto del espacio. Con dichas coordenadas, se determina, mediante triangulación de las cámaras en estéreo, la distancia relativa entre los puntos respecto a los marcos de referencia de cada cámara. Con los datos obtenidos manualmente, y los calculados mediante el par estéreo, se calcula el error, como la diferencia de las medidas (dadas en milímetros). La

escena tomada desde ambas cámaras contiene tres objetos con forma prismática. Del primero, se toman cuatro distancias entre pares de puntos sobre el objeto, tratando de obtener diferentes direcciones y profundidad. De los restantes, se toman 3 distancias

entre pares de puntos sobre los objetos, tratando de que cada punto esté localizado en diferentes zonas del espacio 3D, esto es, que varíen en sus tres coordenadas. Los resultados son mostrados en la Fig. 6.

		Dist.Real	Dist.Cam1	Dist. Cam2	Error. Cam1	Error. Cam2	Err. Medio	Error. Norm
Obj1	Par1	52,50	2,46	2,46	0,04	0,04	0,04	0,08
	Par2	29,50	27,38	27,38	2,12	2,12	2,12	7,19
	Par3	59,50	60,38	60,38	0,88	0,88	0,88	1,49
	Par4	65,00	65,03	65,03	65,03	65,03	0,03	0,05
Obj2	Par1	80,00	80,41	80,41	0,41	0,41	0,41	0,52
	Par2	118,00	120,98	120,98	2,98	2,98	2,98	2,52
	Par3	30,50	31,01	31,01	0,51	0,51	0,51	1,67
Obj3	Par1	55,00	56,07	56,07	1,07	1,07	1,07	1,95
	Par2	108,00	10,22	10,22	2,22	2,22	2,22	2,06
	Par3	55,00	53,38	53,38	1,62	1,62	1,62	2,95
							1,19	2,05

Fuente: Los Autores, 2014.

Fig. 6. Resultados del cálculo de distancia entre pares de puntos de la imagen, con sus respectivas coordenadas 3D en el espacio.

La Fig. 7 muestra las imágenes con los objetos sobre los que se calcularon las distancias. El error medio porcentual encontrado entre las distancias calculadas por el sistema, y las tomadas en forma manual, es de 2.05%. La forma en que se deberá interpretar es que, el error medio para una distancia de 100 mm es de 2 mm. El error suele variar de acuerdo a las coordenadas de los pares de puntos, ya que si en la imagen se toman dos puntos cuya perspectiva en

profundidad es alta, el error tiende a ser máximo. En el cálculo del error, sin embargo, no se ha tenido en cuenta el error mismo en la toma de los datos, tanto en el momento de realizar la medición manual, ya que siendo en milímetros, suele haber mayor sensibilidad a variaciones en la forma de medir, como en la selección de los píxeles de cada imagen, lo cual también viene afectado por factores diversos como por ejemplo, la resolución de la misma cámara.



Fuente: Los Autores, 2014.

Fig. 7. Fotografías tomadas por el par de cámaras de los objetos sobre los cuales se realizaron medición manual entre distancias y cálculo de las mismas mediante triangulación.

B. Análisis Antropométrico Mediante Visión Estéreo Binocular

El error obtenido para el cálculo de distancia en puntos sobre objetos de geometría simple, deja ver

que el sistema está aceptablemente calibrado, con lo cual se realizaron experimentos para determinar la eficiencia del sistema en el momento de tomar medidas sobre el complejo facial. Para ello, se toma la medida manual de las distancias descritas en la

sección 2 y, posteriormente, se calcularon dichas medidas sobre las imágenes obtenidas con el par de cámaras, mediante triangulación de los puntos

adecuadamente seleccionados en cada imagen. Los resultados se muestran en la Fig. 8.

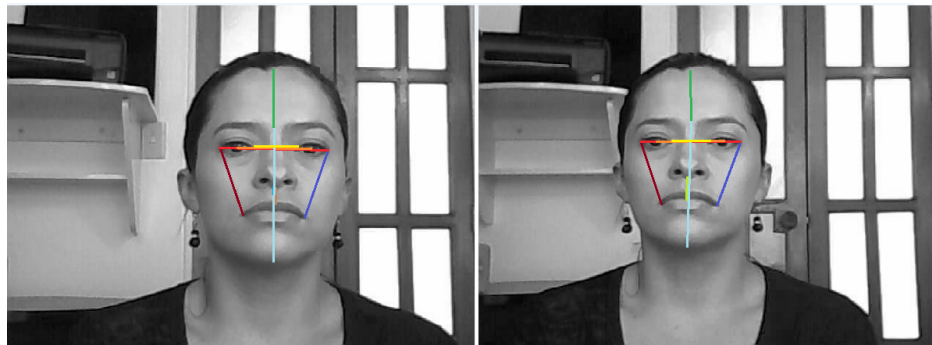
	Manual	Dist. Cam1	Dist. Cam2	Error. Cam1	Error. Cam2	Error medio	Error porcentual
d1	98,0000	96,0763	96,0763	1,9237	1,9237	1,9237	1,96
d2	34,0000	35,5586	35,5586	1,5586	1,5586	1,5586	4,58
d3	64,0000	64,7009	64,7009	0,7009	0,7009	0,7009	1,10
d4	12,0000	12,1227	12,1227	0,1227	0,1227	0,1227	1,02
d5	20,0000	21,9546	21,9546	1,9546	1,9546	1,9546	9,77
d6	55,0000	53,7703	53,7703	1,2297	1,2297	1,2297	2,24
d7	60,0000	63,5213	63,5213	3,5213	3,5213	3,5213	5,87
d8	71,0000	71,5885	71,5885	0,5885	0,5885	0,5885	0,83
d9	69,0000	69,1583	69,1583	0,1583	0,1583	0,1583	0,23
d10	55,0000	61,0812	61,0812	6,0812	6,0812	6,0812	11,06
d11	60,0000	61,0558	61,0558	1,0558	1,0558	1,0558	1,76
						Error porcentual medio	3,67

Fuente: Los Autores, 2014.

Fig. 8. Tabla con los resultados de realizar las medidas manuales de las 11 distancias antropométricas del complejo facial y su aproximación mediante la selección de puntos de las imágenes en el par estereó.

En la Fig. 8, se puede apreciar que el error medio porcentual es del 3,67 %, esto quiere decir que en una distancia de 100 milímetros, el error promedio está en +/-3 milímetros. Como en el caso de la toma de medidas de objetos de geometría simple, no se ha tenido en cuenta el error debido a la falta de precisión en la toma manual de medidas, así como

en la selección de los puntos sobre las imágenes, pero seguramente la suma de estos errores aporta al error medio porcentual. En la Fig. 9, se muestra un ejemplo de la toma del complejo facial, con las medidas tomadas en forma manual, y con la selección de puntos sobre las imágenes de cada cámara.



Fuente: Los Autores, 2014.

Fig. 9. Ejemplo del par de imágenes tomadas desde las cámaras sintonizadas en estereó, y de las medidas que han sido calculadas en forma manual y mediante la selección de puntos sobre cada imagen.

V. CONCLUSIONES

Se ha realizado un estudio de las técnicas de visión artificial y visión estereoscópica que pueden ser aplicadas al desarrollo de herramientas para el análisis de la antropometría facial, tanto en lo re-

lacionado con el análisis de proporciones, como de distancias antropométricas que pueden ser de interés en ciertas áreas de las ciencias de la salud, tal como la fonoaudiología. Para tomar en cuenta las proporciones y medidas de interés, se han tomado

como base documental el texto de Gámez [7]. Se han encontrado gran cantidad de propuestas para la calibración de cámaras y configuración de pares de cámaras en estéreo, con una precisión aceptable. Se ha desarrollado un prototipo basado en la Toolbox de uso público con fines académicos, desarrollada por Bouguet, y se ha probado que el cálculo de distancias sobre el complejo facial, con selección manual de puntos arroja resultados aceptables, con un error promedio bajo. La detección automática de los puntos de interés sobre el complejo facial para las cámaras estéreo requiere el desarrollo de algoritmos robustos a variaciones en iluminación y pose. Esta es un área en constante desarrollo de donde se pueden derivar proyectos futuros. Dentro de estos posibles proyectos futuros, está la revisión y/o desarrollo de algoritmos robustos para la detección de puntos de interés (esto es con poca sensibilidad a variaciones en iluminación y la pose), y la revisión de técnicas de optimización para reducir el error en el proceso de calibración de cámaras en estéreo y determinación de puntos en el espacio por triangulación.

REFERENCIAS

- [1] K. Burian, G. F. Gestring, and M. Haider, "Objective speech audiometry?", *International J. Audiol.*, vol. 5, no. 2, pp. 183-187, 1969.
- [2] D. M. Cattoni, "O uso do paquímetro na morfocidade orofacial: pro cedimentos de acaliacao", *PROFONO*, 2006.
- [3] G. P. Martinsanz, and J. M. de la Cruz, "Visión por computador", *Imágenes digitales y aplicaciones*, 2a. ed., México: Alfaomega, Ra-Ma, 2008.
- [4] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using o -the-shelf TV cameras and lenses", *Radiometry, USA: Jones and Bartlett Publishers, Inc*, 1992, pp. 221-244.
- [5] J. Heikkila, and O. Silven, "A four-step camera calibration procedure with implicit image correction", *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*, 1997, p. 1106.
- [6] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations", *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1999, vol. 1, pp. 666-673.
- [7] C. V. Gámez Jimenez, "Diseño y desarrollo de un sistema de reconocimiento de caras", *Universidad Carlos III de Madrid*, 2009.