



Estudio comparativo de métodos para el análisis del movimiento en 2D: ventajas y desventajas del uso de marcadores.

K. M. Pérez-Alba, A. N. León-Aguilar, R. A. Salido-Ruiz

División de Electrónica y Computación, Coordinación de Ingeniería Biomédica, CUCEI, Guadalajara, México.

Resumen— Se presentan y discuten las ventajas y desventajas de distintos métodos de procesamiento de imágenes aplicados al análisis del movimiento humano. Para el estudio dividimos estos métodos en dos tipos: 1) aquellos que utilizan marcadores pasivos y 2) aquellos que no utilizan ningún tipo de marcador. Se seleccionaron 2 métodos existentes uno basado en el sensor Kinect™ y otro en el software Kinovea además se propusieron e implementaron 2 nuevos métodos: con marcadores y sin marcadores. Un estudio comparativo se llevó a cabo con la finalidad de evaluar la exactitud de cada uno de estos métodos. Los métodos propuestos mostraron ser una opción viable para el análisis del movimiento humano, obteniendo resultados similares a los métodos populares existentes en el mercado. Los resultados apuntan a que Kinovea presenta las mejores estimaciones superando ligeramente al algoritmo propuesto con marcadores. Sólo el método sin marcadores propuesto superó al método basado en el sensor Kinect™ que realiza sus estimaciones sin marcadores. Creemos que las estimaciones de Kinect™ en 2D pudieran ser mejoradas en un futuro.

Palabras clave— Análisis de Video, Captura de movimiento, Kinect, Marcadores, Procesamiento de imágenes.

I. INTRODUCCIÓN

El análisis del movimiento humano tiene como objetivo la obtención y procesamiento de datos, para una mejor comprensión del movimiento humano, su uso en sistemas de rehabilitación [1], e incluso en la captura de movimientos utilizada en animación [2]. Diversos sistemas y programas (software) han sido desarrollados para llevar a cabo análisis de movimiento que van desde electrogoniómetros [3], hasta equipos de captura de movimiento con marcadores y sin marcadores [4]. En este trabajo se proponen métodos que realicen el análisis del movimiento a través del procesamiento de imágenes con y sin marcadores de una manera más eficiente en términos de la exactitud (lo más cercano a mediciones realizadas manualmente sobre las imágenes), así como un estudio comparativo para evaluar el margen de error de éstos algoritmos respecto a algunos de los más populares disponibles en la literatura. Un estudio de la eficiencia del algoritmo en cuanto a recursos o rapidez no se persigue en este trabajo. Lo que se busca de ésta comparación es exhibir que varios de los softwares que están siendo ampliamente utilizados para este tipo de estudios presentan errores que pueden ser reducidos implementando otras técnicas que se basan en el procesamiento de imágenes.

II. METODOLOGÍA

1) *Metodología del estudio*: Se elaboró una base de datos de videos de movimiento humano, conformada por

imágenes de un sujeto realizando cuatro movimientos diferentes repetidos 5 veces. Los movimientos inician en posición anatómica y comprenden las siguientes variantes: a) El sujeto levanta el brazo derecho a 90 grados. b) El sujeto levanta ambos brazos a 90 grados. c) El sujeto eleva la pierna derecha lo más alto posible manteniendo el equilibrio. d) El sujeto extiende sus piernas lateralmente hasta donde sus capacidades le permiten.

2) *Montaje y captura del movimiento*: La captura del movimiento fue realizada con la cámara de un Smartphone y un sensor “Kinect”. Con el objetivo de evitar ruido en las imágenes, la captura fue realizada en una habitación de piso y paredes blancas. El montaje fue centrado a fin de lograr observar todo el cuerpo del sujeto sin tener la necesidad de modificar la posición de los dispositivos durante el proceso de captura, logrando así que ambos videos fueran lo más parecidos entre sí. La grabación con la cámara del Smartphone se controló mediante una conexión WiFi, enlazando MATLAB® y la aplicación IP Webcam. El proceso consistió en obtener las capturas mediante la dirección IP, y leerlas en directo en MATLAB®. De igual forma las grabaciones hechas con el sensor “Kinect” fueron controladas desde una interfaz gráfica que se implementó en MATLAB®. Para fines de procesamiento, se grabó el entorno de cada movimiento. Las imágenes obtenidas contaban con una resolución de 8 bits. Los dos algoritmos propuestos con y sin marcadores fueron implementados en el software MATLAB® utilizando un equipo con Sistema Operativo Windows 7 Ultimate (64 bits) Procesador: Intel Core i3-3227U CPU 1.90 GHz y RAM: 4GB (3.87 utilizable).

3) *Procesamiento de Imágenes con Marcadores para el Análisis del Movimiento (PICMAM)*:

a) Inserción manual de marcadores utilizando un software de edición de imagen, en cada cuadro de video se trazaron marcadores circulares de 20 píxeles de diámetro sobre 15 puntos de interés. El color de estos píxeles se determinó en función de su posición, siendo los marcadores del miembro inferior rojos, los del miembro superior azules y los de cabeza, clavícula y centro de gravedad de color verde (ver Fig. 1).

b) Cargado y organización de los cuadros del video.

c) Separación en componentes RGB de cada cuadro de video para separar marcadores.

d) Binarizado de las imágenes de las componentes RGB con umbral de entre 220 y 255 niveles de cuantificación [5], para obtención de marcadores del color seleccionado.

e) Acondicionamiento de las imágenes binarias para uniformar y eliminar píxeles aislados.

f) Refinado de estructuras mediante dilatación binaria y elemento estructural disco plano de radio 3 [6].

g) Localización (coordenadas) de marcadores por aplicación iterativa de encogimiento/erosión [7] para reducción de objetos (marcadores) a un solo pixel localizado en su centroide (ver Fig. 2).

h) Identificación de marcadores por etapas:

Etapa 1 (Marcadores de la línea central): se comparan y organizan según su valor en el eje Y (vertical), siendo el valor más pequeño el centro de gravedad, el intermedio la clavícula y el más alto la cabeza.

Etapa 2 (Miembros superior e inferior): se separan los marcadores en miembro izquierdo y miembro derecho a través del análisis de las coordenadas en el eje X (horizontal), organizándolos de menor a mayor, definiendo los primeros tres elementos parte del miembro derecho, y los tres últimos parte del miembro izquierdo.

Etapa 3 (Marcadores de extremidades inferiores): se categorizan de acuerdo a su valor en el eje Y, siendo los de valor menor los tobillos, los de valor intermedio las rodillas y los de valor mayor las articulaciones de cadera.

Etapa 4 (Marcadores en extremidades superiores): se determinó primero el hombro como el punto más cercano a la clavícula y posteriormente según su distancia con los dos elementos restantes se define el codo como el más cercano y la muñeca como el más lejano.

i) Graficado de los puntos articulares identificados y estimación del ángulo entre vectores (ver Fig. 3).

j) Estimación de los ángulos mediante tangente inversa.

4) *Procesamiento de Imágenes sin Marcadores para el Análisis del Movimiento (PISMAM)*:

a) Captura un cuadro de video del entorno sin el sujeto.

b) Carga video del movimiento y se separa en cuadros.

c) Se resta a cada cuadro del video capturado el cuadro del entorno y se convierte a escala de grises.

d) Se binariza la imagen con un umbral de entre 50 y 255 niveles de cuantificación obteniéndose así imágenes con fondo negro y la silueta del sujeto en color blanco.

e) Se rellenan huecos en la imagen empleando técnicas de dilatación utilizando una estructura en forma de disco, esto debido a que la figura era bastante irregular.

f) Se estiman las medidas antropométricas de las siluetas del sujeto. En el cuadro 1 del movimiento (sujeto en posición anatómica) se calcula la altura del sujeto, restando el pixel blanco máximo y mínimo encontrados en el eje y, con un margen de error de aproximadamente 10 pixeles debido al error que produce la incertidumbre de los pies del sujeto. Después se estiman las líneas centrales a partir de los contornos encontrados al buscar cambios de color en los pixeles en el eje x, y se separan en conformidad con su posición en el eje x en, línea central, brazo izquierdo, brazo derecho, pierna izquierda y pierna derecha.

g) A partir de las medidas antropométricas estimadas y la altura del sujeto, se calcula la longitud de cada parte del cuerpo sobre las líneas centrales y se marcan posibles puntos articulares.



Fig. 1. Trazado de marcadores en imágenes sin marcadores físicos para comparación. a) Imagen original, b) Imagen con marcadores.

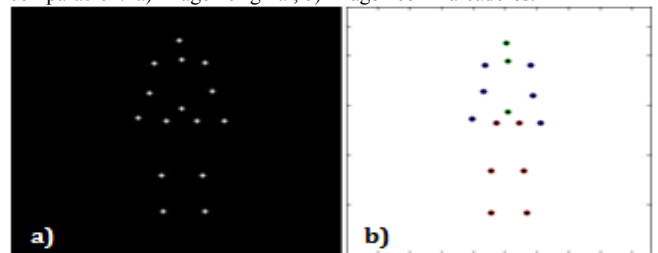


Fig. 2. Gráfica de los pixeles encontrados a) Marcadores de asterisco b) Graficado posterior de puntos articulares.

Cada punto es centrado según su cercanía al contorno del sujeto (ver Fig. 4). Los datos obtenidos son utilizados como referencia para el cuadro siguiente.

h) Se identifica la cabeza en todas las imágenes, y en base a ello se encuentran la clavícula y los hombros.

i) Se delimita un área de búsqueda alrededor del segmento del miembro superior y se encuentran las articulaciones correspondientes a codos y las muñecas mediante la realización de búsquedas pixel a pixel (barridos) en el área delimitada aplicando el teorema de Pitágoras:

$$b^2 + a^2 = c^2 \quad (1)$$

de forma que, al dibujar un triángulo rectángulo imaginario entre las coordenadas de la extremidad, donde la hipotenusa c es la longitud en pixeles de la extremidad, (i.e., en el brazo c , sería la longitud entre el hombro y el codo, por lo tanto a y b son las longitudes en pixeles de los catetos de este triángulo[8]. El resultado de (1) proporciona valores muy cercanos a la medida de cada extremidad cuando fue medida durante la posición anatómica, de tal forma que podemos delimitar áreas que no superen esa longitud para hacer los barridos correspondientes a cada articulación.

j) De manera similar se procedió a encontrar las articulaciones del miembro inferior (tobillo, rodilla y cadera) utilizando (1) y con ello se encontró el centro de gravedad, sobre la cadera en un punto medio.

k) Los puntos encontrados son centrados y tomados como referencia para el siguiente cuadro.



Fig. 3. Puntos articulares y visualización del esqueleto, a) Puntos sobre I. RGB; b) Esqueleto sobre I. RGB; c) Esqueleto sobre fondo blanco.

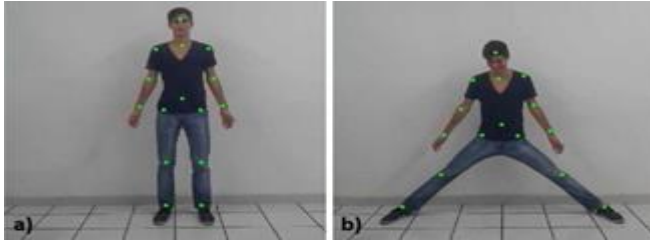


Fig. 4. Puntos articulares encontrados, a) En posición anatómica; b) Durante movimiento.

Una vez estimados todos los puntos se grafica sobre diferentes opciones, la imagen binaria, la imagen en RGB o un fondo negro, (ver Fig. 5) tanto con puntos como con líneas (ver Fig. 6).

1) Por último se eligieron los puntos con más movimiento y con ayuda de la función trigonométrica de la tangente inversa y haciendo uso de las coordenadas de los puntos encontrados, se calculó el ángulo de movimiento de los hombros, los codos y la cadera.

5) *Procesamiento con Kinect*: La grabación se inicia definiendo las cámaras del dispositivo como entradas de video. Se selecciona la modalidad de seguimiento, que lleva a cabo el trazado del esqueleto y se selecciona la posición del sujeto en estado de pie. Posteriormente se establece la tasa de grabación de 60 cuadros a 10 fps. Se crea un objeto de video a través del cual se obtienen los cuadros del video (ver Fig. 7). Se mantiene la imagen capturada por la cámara RGB como el fondo del video. Los videos se almacenan en formato mp4 y los datos de las coordenadas de las articulaciones se organizan en arreglos de celdas que contienen los valores en los 3 ejes coordenados (x, y, z).

6) *Procesamiento con Kinovea*: Se carga cada uno de los videos y se selecciona manualmente la ubicación de los puntos articulares, y posteriormente se activa la opción de seguimiento de trayectoria. Se avanza el video cuadro a cuadro se edita la trayectoria de cada punto para mayor precisión (ver Fig. 8).

7) *Cálculo del error entre mediciones*: Como lo que se pretende es exhibir el error de la estimación de los algoritmos respecto a una medición, solo se tomaron medidas manuales de los ángulos de apertura formados por hombros, codos, y cadera en los diferentes videos imprimiendo en papel cada cuadro y usando como herramientas un transportador y una escuadra. De igual forma, los mismos ángulos fueron estimados por cada



Fig. 5. Graficado de puntos articulares en posición anatómica en distintos fondos, a) I. binaria; b) Imagen RGB; c) Fondo negro.



Fig. 6. Visualización de esqueleto. A través del graficado de segmentos entre los puntos articulares, a) I. binaria; b) Imagen RGB; c) Fondo negro.

método, tomando como referencia los puntos articulares encontrados por cada uno. La diferencia entre las mediciones y las estimaciones de los métodos fueron promediadas y así conseguimos encontrar el promedio de error de cada método en el punto indicado.

III. RESULTADOS

Se observa que el procesamiento con Kinovea fue el que tuvo una mayor exactitud ya que presenta estimaciones con un promedio de error menor a todos los demás algoritmos, seguido del *PICMAM*, el cual únicamente mantiene una ligera diferencia con el primero. Después de ellos, tal y como se puede observar a en la Tabla 1, se encuentra el método sin marcadores, que de igual forma tiene un error promedio muy cercano al anterior, por último se puede observar el procesamiento con Kinect, que tiene el mayor error promedio. Si se ve detenidamente podemos destacar que para encontrar la cadera, el *PICMAM* resultó mucho más exacto que los otros métodos, así como en la estimación de los codos Kinovea tiene un menor promedio de error pero no muy lejano al *PISMAM*. En el caso de los hombros de igual forma, Kinovea obtiene más exactitud, pero se mantiene cerca el método con marcadores. Al comparar los métodos que usan marcadores, se puede observar que Kinovea tiene un menor error promedio en general, aunque también debe ser tomada en cuenta la practicidad de cada algoritmo, pues al usar Kinovea, el proceso debe ser vigilado cuadro a cuadro para irse corrigiendo cada punto conforme avanza el video, mientras que al usar el *PICMAM* el proceso se realiza de forma automática, sin embargo, los cuadros deben ser marcados antes de procesarse, de igual forma, al usar el *PICMAM*, obtenemos un esqueleto completo del sujeto, mientras que al usar Kinovea, únicamente se obtienen las trayectorias de los puntos señalados. Por otro lado, comparando los métodos que no utilizan marcadores, el promedio de error es mayor en el procesamiento con Kinect, siendo más exacto el *PISMAM*,

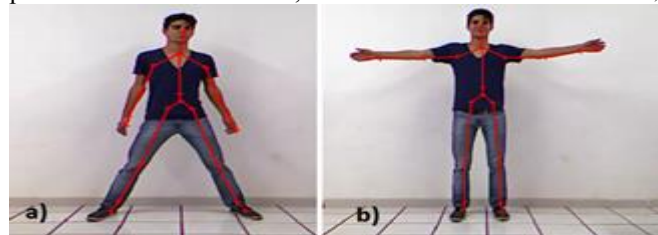


Fig. 7. Captura de video obtenido con Kinect. Muestra el esqueleto creado por el algoritmo de Kinect sobrepuesto al cuerpo del sujeto, a) Movimiento de ambas piernas; b) Movimiento de ambos brazos.

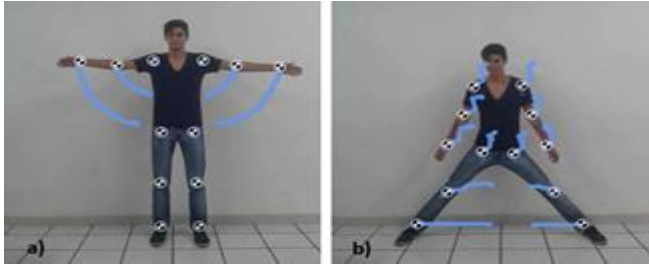


Fig. 8. Captura de video procesado con Kinovea. Muestra marcadores de Kinovea y las trayectorias registradas en un tono azul, a) Movimiento de ambos brazos; b) Movimiento de ambas piernas.

también se puede destacar que este último contiene más opciones de visualización pero es un procesamiento que se realiza en videos previamente grabados, mientras que Kinect realiza el procesamiento en tiempo real. A pesar de que Kinovea presenta mejores resultados, estos son solo ligeramente mejores al *PICMAM* (ver Fig. 9 y Fig. 10).

IV. DISCUSIÓN y CONCLUSIÓN

Los datos obtenidos a través de los algoritmos desarrollados entregaron muy buenos resultados, el *PICMAM* propuesto permitió trazar de manera muy precisa el esqueleto sobre el video, encontrando la ubicación exacta de cada articulación marcada y el *PISMAM* pudo estimar las ubicaciones que difieren muy poco con las ubicaciones de las articulaciones reales. Se pudo observar que los métodos de estimación con marcadores presentan un error promedio menor a la de los métodos de estimación sin marcadores y por ello poseen una mayor precisión en sus resultados. Esto se debe a que muchos de los errores que pueden presentarse durante la estimación en los algoritmos sin marcadores se evitan al determinar manualmente la ubicación de las distintas articulaciones con el colocado de marcadores. Los tiempos de procesamiento fueron bastante diferentes para cada uno de los algoritmos, se realizaron mediciones y se calculó el promedio del tiempo en segundos para el procesamiento de 60 cuadros de video. Los sistemas con tiempos mayores fueron el *PICMAM propuesto* y Kinovea, ya que se incluyó el tiempo necesario para la colocación y ajustes manuales de marcadores como parte del tiempo total de procesamiento, Kinovea tuvo un promedio de 186.30 segundos y nuestro algoritmo 167.02 segundos. Y por su parte los procesamientos sin marcadores tuvieron tiempos menores, *PISMAM* tuvo 54.28 segundos y Kinect 13.96 segundos. Si bien Kinect tuvo el menor tiempo del conjunto,



Fig. 9. Comparación de resultados en movimiento de ambas piernas, a) Resultado del Kinect; b) Resultado de Kinovea; c) Resultado del algoritmo sin marcadores; d) Resultado del algoritmo con marcadores.

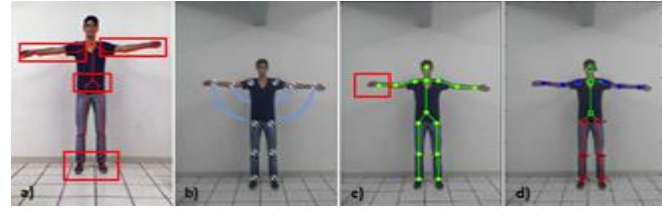


Fig. 10. Comparación de resultados en levantamiento de ambos brazos, a) resultado del Kinect; b) Resultado de Kinovea; c) Resultado del algoritmo sin marcadores; d) Resultado del algoritmo con marcadores.

TABLA 1
ERROR PROMEDIO EN GRADOS ENCONTRADO EN LAS DISTINTAS ARTICULACIONES.

Articulación	Error promedio calculado				
	Método	Con marcadores	Sin marcadores	Con Kinect	Con Kinovea
Hombro derecho		3.5°	2.62°	11.04°	0.75°
Hombro izquierdo		2.5°	4.05°	3.92°	1.25°
Codo derecho		5°	2.33°	6.51°	1°
Codo izquierdo		1.75°	2.92°	9.42°	3.5°
Cadera derecha		0.75°	2.18°	2.98°	1.25°
Cadera izquierda		1°	1.68°	3.8°	2.25°
Promedio		2.41°	2.63°	6.27°	1.66°

también tuvo el mayor error promedio entregando un esqueleto desproporcionado por el desplazamiento de articulaciones a ubicaciones distintas a las reales, mientras que el *PISMAM* destacó por su fidelidad a las ubicaciones reales al realizar la estimación del esqueleto a través de la silueta corporal. Es posible que la deformación presentada en los datos de Kinect se deba a que el dispositivo hace uso de su sensor de profundidad para estimar también la profundidad, es decir, crea un esqueleto tridimensional, lo cual posiblemente sea la causa del incremento del error en mediciones bidimensionales como las de éste estudio. Se planea en un futuro realizar un nuevo estudio comparativo en 3D y proponer un método sin marcadores que sea capaz de llevar a cabo dichas estimaciones.

BIBLIOGRAFÍA

- [1] Muñoz-Cardona, J. E., Henao-Gallo, O. A., & López-Herrera, J. F. (2013). Sistema de Rehabilitación basado en el Uso de Análisis Biomecánico y Videojuegos mediante el Sensor Kinect. *TecnoLógicas*, 43-54.
- [2] Sharma, A., Agarwal, M., Sharma, A., & Dhuria, P. (2013). Motion capture process, techniques and applications. *Int. J. Recent Innov. Trends Comput. Commun*, 1, 251-257.
- [3] Braidot, A., Giménez, D., Kipen, E., & Ramos, F. (2002). Diseño de un electrogoniómetro telemétrico y microcontrolado. *Bioingeniería y Física Médica Cubana*, 3, 13-21.
- [4] Michoud, B., Guillou, E., Briceno, H., & Bouakaz, S. (2007, October). Real-time marker-free motion capture from multiple cameras. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on* (pp. 1-7).
- [5] Gonzalez, R. S., & Wintz, P. (1977). *Digital image processing*.
- [6] Torres, W. J., & Bello, R. J. Procesamiento de imágenes a color utilizando morfología matemática.
- [7] Young, I. T., Gerbrands, J. J., & Van Vliet, L. J. (1998). *Fundamentals of image processing*. Delft: Delft University of Technology.
- [8] Dixon, R. A. (1987). *Mathographics*. Courier Corporation. 92-95.