

CAIS, Congreso Argentino de Informática y Salud

Escaneo corporal 3D de bajo costo para monitoreo y seguimiento remotos de sobrepeso

Navarro, P.^{1,2,3}, Cintas, C.^{1,2,3}, Ramallo, V.¹, de Azevedo, S.¹, Morales, A. L.^{1,2,3}, Trujillo, A.^{1,2}, Pazos, B. A.^{1,2,3}, González-José, R., Delrieux, C.²

¹ Instituto Patagónico de Ciencias Sociales y Humanas (IPCSH). Centro Nacional Patagónico (CENPAT), Puerto Madryn, Chubut.

² Laboratorio de Ciencias de las Imágenes, Departamento de Ingeniería Eléctrica y Computadoras. Universidad Nacional del Sur, Bahía Blanca.

³ Universidad Nacional de la Patagonia San Juan Bosco, Facultad de Ingeniería, Trelew, Chubut.

Abstract. Según la Organización Mundial de la Salud (OMS), el sobrepeso y la obesidad infantil constituyen dos de los principales problemas de salud pública del siglo XXI. De 42 millones de niños y niñas con sobrepeso en el mundo, el 80% vive en países en desarrollo, especialmente Latinoamérica. En Argentina, el 27,4% de los hombres y el 31% de las mujeres sufren de obesidad, es por esto que es reconocida como una enfermedad de interés nacional (Ley 26.396/08) e integrada por el Ministerio de Salud dentro del Programa Médico Obligatorio. Para su tratamiento, es crucial un diagnóstico preciso y un monitoreo y seguimiento regular. Hay creciente consenso en que el índice de masa corporal (BMI), que se basa únicamente en estatura y peso, es insuficiente para tal propósito, requiriéndose en cambio técnicas más elaboradas para poder evaluar la forma y tamaño corporales. Existen diversos sistemas aplicables al escaneo corporal, pero debido a su costo no son de uso masivo o accesibles. En cambio, las mediciones que se realizan usualmente carecen de la suficiente exactitud y precisión antropométricas como para poder brindar información confiable para dicho seguimiento. En este trabajo se presenta una aplicación para dispositivos móviles que, junto con un servicio en la nube, permite realizar un escaneo 3D, sobre el cual se pueden obtener variables antropométricas de calidad y que permitan cuantificar adecuadamente la forma corporal. Mediante un procedimiento simple y no invasivo se analiza la forma y tamaño corporal, permitiendo diagnosticar obesidad o condiciones similares, accediendo a su vez a la historia clínica del paciente para conocer datos específicos, su evolución temporal o valores de referencia de la OMS.

PALABRAS CLAVE: Procesamiento de imágenes, *Structure from Motion* (SFM), Morfometría Geométrica, Reconstrucción 3D, antropometría.

1 Introducción

En un gran número de situaciones y circunstancias, tanto dentro como fuera de la práctica clínica, se requiere contar con mediciones antropométricas precisas (formas, tamaños, perímetros y volúmenes de diferentes partes del cuerpo) de una manera rápida, confiable, y económica. Un apropiado relevamiento de la forma corporal es clave para el diagnóstico, tratamiento y seguimiento clínico de diversas condiciones muy frecuentes como la obesidad y el sobrepeso, y otras no tan frecuentes pero de gran importancia como la escoliosis. Asimismo, un acceso sencillo a la forma corporal y su cuantificación es de interés en la deportología y el *bodybuilding* donde se requiere un monitoreo frecuente y exacto del resultado de diversas rutinas de entrenamiento en el esquema corporal. Es importante tener en consideración que, en la actualidad, las medidas que se utilizan para diagnosticar obesidad varían significativamente, dependiendo del protocolo empleado.

Por su parte, los escaneos 3D se usan en medicina por su capacidad de medir tamaño, forma, área y volumen de un cuerpo, pero sus aplicaciones en epidemiología o diagnóstico y monitoreo se encuentran limitadas por el costo de adquisición y de operación del equipamiento, las dificultades de traslado de los instrumentos, y el ambiente controlado que se requiere para su correcto funcionamiento. En este trabajo, se presenta una implementación de un sistema de escaneo móvil, basado en un *smartphone* o tablet, que permita la toma de datos con máxima calidad y de manera rápida, amigable y económica. Sobre la superficie escaneada detectarán automáticamente *landmarks* anatómicos, derivando medidas directas, superficies y volúmenes de interés para el diagnóstico de diferentes condiciones, permitiendo además el manejo remoto de datos. El mecanismo de



Fig. 1. Reconstrucción 3D a partir de una secuencia de fotogramas.

adquisición, al ser muy sencillo, puede ser realizado por el paciente, y al subirse los datos a la nube, pueden ser analizados por el especialista a distancia.

2 Procedimiento propuesto

Se desarrolló un prototipo de la aplicación, basado en un pipeline de procesamiento que se describe más abajo. Este prototipo cuenta con el proceso de generación de una *mesh 3D* comenzando por la captura del video, y finalizando con la visualización de la reconstrucción en tres dimensiones del paciente. Dicho pipeline esta basado en siete (7) pasos bien diferenciados, los cuales involucran tecnologías y técnicas particulares de cada etapa.

1. **Video:** para la obtención de la filmación contamos con una página web responsiva que le permite al paciente grabar un video o seleccionar del dispositivo uno ya grabado, además de permitirle ingresar sus datos personales como DNI, nombre, edad, etc. Actualmente el protocolo para la captura involucra filmar en 360 grados a la persona (con ropa interior o deportiva lo más ajustada al cuerpo que se pueda) de pie, con las piernas separadas a una distancia fija de 40 cm, los brazos completamente extendidos a la altura de la cintura escapular y las manos en pronación (ver Fig. 1).
2. **Preprocesamiento:** en esta etapa se divide el video en un conjunto de *frames* (aproximadamente cien), a los cuales se le aplica un filtro gaussiano para eliminar el ruido de las imágenes producido por el movimiento de la cámara, la iluminación, etc. Para la implementación de dicho filtro se utiliza una matriz de convolución (en este caso la matriz que representa una función Gaussiana determinada), la cual es aplicada a cada *pixel* de la imagen junto con sus vecinos mas cercanos [1]. Este procesamiento es importante ya que produce contornos más

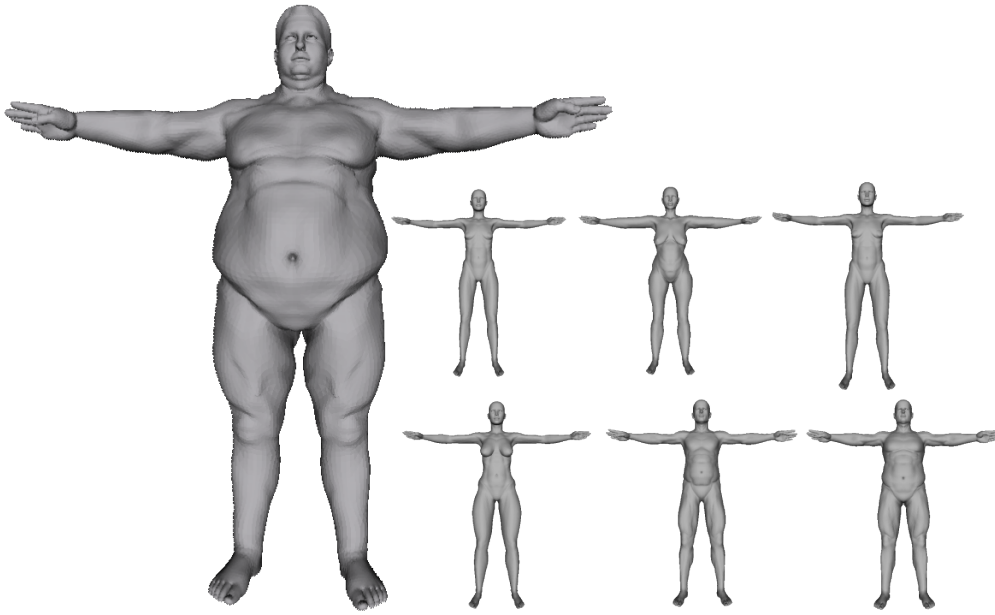


Fig. 2. Modelos sintéticos.

pronunciados, lo cual permite diferenciar al objeto a reconstruir del fondo y el ruido que puede presentar la escena.

3. **Structure from Motion(SFM):** Esta técnica permite, a partir de una colección de imágenes, recrear una escena en tres dimensiones, es decir, a través de imágenes 2D y calculando la posición aproximada de la cámara en el espacio, obtener las coordenadas X, Y, Z de los puntos más destacados de la escena. En primer lugar el método necesita, de cada imagen, puntos característicos (*features*) invariantes en escala, rotación e iluminación. Los puntos así etiquetados en diferentes cuadros de la toma son puestos en correspondencia, utilizándose algoritmos específicos para resolver ambigüedades en el caso en que los features de los puntos generen confusión. Contando con grandes cantidades de puntos correctamente identificados y puestos en correspondencia, es posible determinar por fotogrametría su posición absoluta en 3D (y por paralaje, la posición relativa de la cámara en cada uno de los cuadros de la toma), lo cual da como resultado una nube de puntos 3D de la persona que estaba siendo filmada [2,3].
4. **Limpeza de la nube de puntos:** Como la reconstrucción generada a partir de SFM contiene la escena completa, es necesario filtrar el objeto principal que se quiere reconstruir, en este caso a la persona, por lo cual en esta etapa se elimina el fondo de la escena, junto con los puntos atípicos de la reconstrucción. Para realizarlo se calcula el promedio y desviación estándar de los vecinos más cercanos, y se eliminan los puntos que quedan fuera del promedio más/menos el desvío multiplicado por un valor definido por la cantidad de vecinos que se utilicen [4].
5. **Alineación de modelo:** La reconstrucción 3D calculada a partir de imágenes tomadas de un dispositivo móvil es de menor precisión que una generada a partir de un escáner 3D dedicado, por lo que se debe lidiar con una reconstrucción incompleta. Es por esto que se utilizan modelos existentes del objeto que se quiere escanear con el objetivo de completar esta nube. Como la forma (peso, altura, ancho de la cadera, cintura, etc) de las personas escaneadas pueden variar, se debe encontrar un modelo que se ajuste con el menor esfuerzo posible, es decir, se debe alinear con el que menos transformaciones se le deba aplicar para llegar a un modelo perfectamente alineado. Actualmente se utilizan dos modelos sintéticos para esta etapa, los cuales corresponden al hombre y mujer promedio. Además, se realizó una colecta de 150 personas para obtener datos y entrenar el software en el landmarking 3D automático, y ponerlo a prueba en diferentes ambientes (color de ropa, estatura, tipos de iluminación, etc.)¹. En un futuro se planea incorporar más de 600 modelos sintéticos con diversa forma y tamaño (Fig. 2) y duplicar la cantidad de datos con humanos reales.
6. **Extracción de esqueleto:** Para generar un mesh 3D a partir de la nube de puntos completa, es necesario contar con las normales a cada punto de la nube. Las técnicas existentes para

¹ Realizada en el hospital de la ciudad de Puerto Madryn con el correspondiente aval del comité de ética.
47JAIIO - CAIS - ISSN: 2451-7607 - Página 12

realizar este tipo de procesamiento se basan en aproximar a un plano una cantidad determinada de vecinos a cada punto. El resultado son las dos posibles normales a dicho plano. Para desambiguar este proceso, utilizamos el cálculo de centroides, que difiere dependiendo la parte del cuerpo (tronco y cada una de las extremidades). Para calcular cada centroide se divide la nube por cada parte del cuerpo, mediante un proceso de esqueletonización [5]. El mesh 3D es contraído hasta dar como resultado el esqueleto de la nube de puntos, cabe aclarar que este esqueleto no es anatómicamente correcto, pero sí es útil para el cálculo de los centroides. Con cada parte dividida, el cálculo del centroide es un promedio. Además la generación y extracción del esqueleto genera información en forma de estructuras curvilíneas que permite por un lado tener información para futuros análisis, tanto de morfología como de topología del cuerpo, y por otro, tener un modelo geométrico articulado que permite fácilmente el desarrollo de *avatars* antropométricamente exactos que sean inmediatamente manipulables por las plataformas de videojuegos 3D. En próximas versiones esta etapa será substituida por la utilización de landmarks ubicados automáticamente (véase [6]), ya que, a diferencia de la esqueletización, las divisiones serán anatómicamente válidas. Con las normales calculadas, y utilizando la técnica *Poisson Surface Reconstruction* [7] se obtiene la primera aproximación al mesh 3D. Esta reconstrucción será la base del último paso del pipeline.

7. **Ajuste volumétrico del modelo:** Con el mesh 3D resultante de la etapa anterior y el modelo sintético utilizado para completar la nube, se procede a realizar un nuevo ajuste [8], que resultará en el mesh final. Este ajuste se realiza deformando el modelo sintético, de acuerdo a los valores reales del mesh obtenido por reconstrucción de Poisson *cita* en el punto anterior (que corresponden al individuo filmado) buscando vértices del modelo que tienen un vértice "objetivo" en el mesh obtenido.

Cabe mencionar que, además de todas las consideraciones de seguridad informática necesarias para preservar datos sensibles, el sistema borrea la información visual del rostro y demás partes que puedan afectar la privacidad de la persona durante este proceso.

3 Resultados y Discusión

Actualmente se cuenta con un prototipo funcionado en condiciones controladas, lo cual permite poner a prueba el mismo con el objetivo de incrementar la cantidad de caso relevados, así como mejorar la precisión y exactitud de este *pipeline*. Si bien existen sensores 3D en el mercado (mayormente en Europa y EEUU), tienen varias desventajas comparadas con este desarrollo, una de ellas, es que no generan de manera automática y amigable datos o variables de interés en la práctica médica. En su mayoría son de gran tamaño y costo, y además no están integrados a una nube que contemple el análisis de los datos, su persistencia, acceso remoto, etc. Futuras versiones podrán incluir la evaluación y descripción de malformaciones congénitas y síndromes. El acopio de datos permitirá obtener parámetros estadísticos demográficos de utilidad para el diseño de políticas sanitarias.

References

1. R. Szeliski, "Computer Vision : Algorithms and Applications," *Computer*, vol. 5, p. 832, 2010.
2. F. Dellaert, S. M. Seitz, C. E. Thorpe, and S. Thrun, "Structure from motion without correspondence," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 2, pp. 557–564 vol.2, 2000.
3. N. Micheletti, J. H. Chandler, and S. N. Lane, "Structure from Motion (SfM) Photogrammetry," *Geomorphological Techniques*, vol. 2, pp. 1–12, 2015.
4. R. B. Rusu, Z. C. Marton, N. Blodow, M. Dolha, and M. Beetz, "Towards 3D Point cloud based object maps for household environments," *Robotics and Autonomous Systems*, vol. 56, no. 11, pp. 927–941, 2008.
5. A. Tagliasacchi, I. Alhashim, M. Olson, and H. Zhang, "Mean curvature skeletons," *Computer Graphics Forum*, vol. 31, no. 5, pp. 1735–1744, 2012.
6. Z. B. Azouz, C. Shu, and A. Mantel, "Automatic locating of anthropometric landmarks on 3D human models," in *Proceedings - Third International Symposium on 3D Data Processing, Visualization, and Transmission, 3DPVT 2006*, pp. 750–757, 2007.
7. M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson Surface Reconstruction," *Eurographics Symposium on Geometry Processing*, 2006.
8. T. H. Kwok, K. Y. Yeung, and C. C. L. Wang, "Volumetric template fitting for human body reconstruction from incomplete data," *Journal of Manufacturing Systems*, vol. 33, no. 4, pp. 678–689, 2014.