



Avances hacia el desarrollo de una interfaz natural e intuitiva para la exploración de datos volumétricos de origen médico.

Vladimir C. Ocegueda-Hernández¹, E. Gerardo Mendizabal-Ruiz¹

¹Departamento de Ciencias Computacionales, Universidad de Guadalajara, Guadalajara, México

Resumen— Las técnicas modernas de exploración médica permiten obtener secuencias de imágenes 2D que corresponden a diferentes secciones de una región corporal. Apilando estas imágenes es posible crear una estructura 3D cuyos elementos de volumen (*voxels*) poseen valores de intensidad luminosa que en su conjunto se denominan *datos volumétricos*. En general, las herramientas disponibles para visualizar esta información son complejas y poco intuitivas. En este artículo, se presentan los avances hacia el desarrollo de una interfaz de usuario basada en algoritmos de visión computacional que permiten suponer que los datos volumétricos existen en un espacio físico definido. Mediante la manipulación de elementos planares dentro de este espacio es posible generar intersecciones con el volumen médico y calcular la imagen correspondiente. El resultado final se traduce en una experiencia de navegación mucho más natural e intuitiva.

Palabras clave— Datos Volumétricos, Exploración Médica, Imagen Médica, Interacción Hombre-Máquina, Interfaz de Usuario

I. INTRODUCCIÓN

Tradicionalmente, el diagnóstico radiológico se ha basado en la inspección de imágenes 2D, aunque existen casos como el análisis de fracturas complejas donde las imágenes 3D se tornan invaluable [1]. Secuencias de cortes o secciones 2D son obtenidos mediante alguna técnica de exploración médica como la tomografía computarizada (TC) o las imágenes por resonancia magnética (IRM). Estas imágenes poseen metadatos que permiten identificar su posición dentro de una secuencia. Por tal razón, es posible apilarlas para generar una representación 3D de la región de interés (Fig. 1). Se denomina *datos volumétricos* a la malla escalar 3D que se forma con los valores de intensidad de los elementos de volumen (*voxels*) contenidos en esta representación.

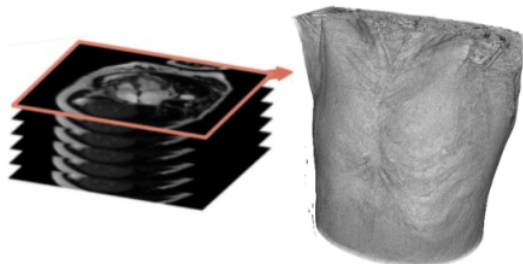


Fig. 1. Cortes o secciones 2D apilados para generar la representación 3D de una región de interés.

Como se plantea en [2], la visualización de imágenes médicas como un modo de explorar los datos de los pacientes es una tarea que los especialistas deberían realizar con rapidez y precisión sin invertir esfuerzos cognitivos considerables. En la práctica, para interactuar con los datos volumétricos se recurre a proyecciones 2D que se visualizan sobre superficies planas (*ej.* pantallas o monitores) y se manipulan con dispositivos mediadores como ratones y teclados de computadora. Las proyecciones pueden procesarse para extraer cortes con diferentes orientaciones, aplicar pseudo-color, medir estructuras presentes en la imagen, etc. [3]. Sin embargo, en general, estos beneficios son obtenidos al precio de incrementar la complejidad de la interfaz de usuario.

En este contexto, la experiencia de exploración podría considerarse más *natural* si la interacción con los datos volumétricos se realizara en un espacio tridimensional y más *intuitiva* si el esfuerzo cognitivo del usuario final se redujera al mínimo. Para lograrlo, se propone un modelo basado en algoritmos de visión computacional que permiten mapear los datos volumétricos a un espacio físico delimitado por cuatro marcadores esféricos. El usuario puede entonces suponer que los datos ocupan un lugar en el mundo real. Luego, utilizando elementos tan simples como una hoja de papel rígido para seccionar este espacio, es posible visualizar en tiempo real la imagen del corte sobre una pantalla de computadora.

II. METODOLOGÍA

II.A. Localización de los marcadores del mundo real

La tecnología disponible en los dispositivos Kinect permite capturar imágenes RGB y estimar la posición de los objetos en escena con respecto a su sistema de referencia por medio de un sensor de profundidad [4].

Auxiliados por un equipo Kinect, localizamos los marcadores del mundo real mediante un proceso de dos etapas: segmentación y estimación de centroides. En la primera etapa, captamos la imagen de cuatro marcadores esféricos de color rojo dispuestos en forma cuadrangular. Luego, convertimos la imagen del modelo RGB al modelo HSV para facilitar su segmentación por umbrales de color [5]. Etiquetamos entonces las regiones obtenidas usando el algoritmo de componentes conectados [6]. Por último, depuramos la imagen segmentada conservando únicamente las regiones R , $\forall r \in \{1,2,3,4\}$ más grandes en escena (a saber, las regiones pertenecientes a los marcadores).

En la segunda etapa, empleamos el cálculo de centroides para asignar una posición a los marcadores del mundo real. Por definición, para un conjunto finito de n puntos $p_1, p_2, \dots, p_n \in \mathbb{R}^3$, las coordenadas de su centroide están dadas por (1).

$$x_n = \frac{1}{n} \sum_{p=1}^n x_p \quad y_n = \frac{1}{n} \sum_{p=1}^n y_p \quad z_n = \frac{1}{n} \sum_{p=1}^n z_p \quad (1)$$

Operando sobre los píxeles de cada R_r (los únicos objetos presentes en la imagen binaria), transformamos sus coordenadas al sistema de referencia del Kinect. Así formamos cuatro conjuntos de puntos 3D cuyos centroides C_r se obtuvieron por medio de (1). Con las coordenadas espaciales de C_r definimos las posiciones de los marcadores del mundo real.

II.B. Mapeo de puntos entre diferentes espacios

Utilizando las librerías de ITK, almacenamos en memoria los cortes o secciones 2D en formato DICOM procedentes de un equipo de exploración médica. Apilando estas imágenes, reconstruimos un volumen médico que denominamos *espacio virtual* V (Fig. 2). Las dimensiones de V se acotaron por los vértices $V_v \forall v \in \{1,2,3,4,5,6,7,8\}$ expresados en términos de los índices i, j, k .

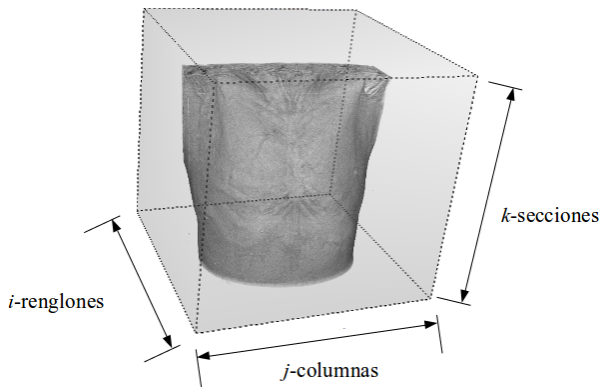


Fig. 2. Espacio virtual definido por un volumen médico.

Agregando la dimensión vertical L al arreglo de marcadores del mundo real, definimos un *espacio físico* F (Fig. 3). Para preservar la similitud entre F y V , calculamos la longitud L (en milímetros) considerando las proporciones entre ambos espacios. En (2), F_x representa la dimensión horizontal de F (en milímetros), y E_j y E_k representan los espaciamentos inter-voxel (en milímetros) sobre los ejes horizontal y vertical de V .

$$L = \frac{F_x \cdot k_{secciones} \cdot E_k}{j_{columnas} \cdot E_j} \quad (2)$$

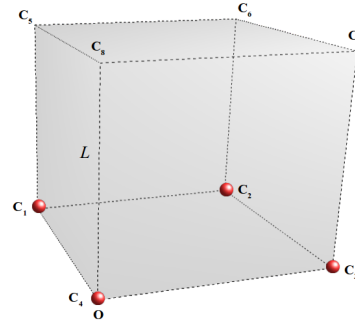


Fig. 3. Espacio físico definido por los marcadores del mundo real.

Las dimensiones de F se acotaron por los vértices $C_r \forall r \in \{1,2,3,4,5,6,7,8\}$ expresados con respecto al sistema de referencia del Kinect.

Formando las matrices $P_{4 \times 8}$ y $Q_{4 \times 8}$ con los puntos C_r y V_r respectivamente (en coordenadas homogéneas), inferimos una transformación afín T capaz de minimizar el error de ajuste espacial \mathcal{E} entre F y V (3).

$$\mathcal{E} = \|T(P) - Q\|^2 \quad (3)$$

Dado que $T(P) = MP$, el problema se redujo a estimar la matriz de transformación $M_{4 \times 4}$. La solución especial M_0 que minimiza \mathcal{E} se obtuvo calculando la matriz pseudoinversa de P , denotada como P^+ [7]. La descomposición en valores singulares DVS proporciona un método para computar P^+ [8]. Esta operación debe realizarse una sola vez por cada cambio en la posición de los marcadores del mundo real. Aplicando (4), encontramos M_0 , que es la matriz de transformación utilizada para mapear puntos de F a V .

$$M_0 = QP^+ \quad (4)$$

II.C. Cálculo de la intersección plano-volumen

Segmentamos un elemento planar de color azul S de acuerdo al mismo principio descrito en II.A. A continuación, mapeamos la posición de sus píxeles a F y luego a V empleando una secuencia de transformaciones geométricas. En otras palabras, con la intersección de F por S , generamos un corte sobre V (Fig. 4).

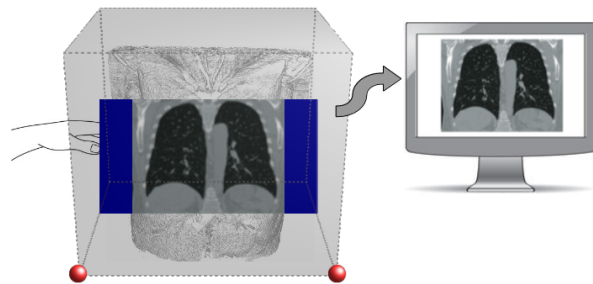


Fig. 4. Seccionamiento de espacios por medio del elemento planar.

III. RESULTADOS

Finalmente, usando interpolación trilineal, calculamos la intersección plano-volumen y desplegamos la imagen correspondiente sobre la pantalla de una computadora.

Para ejecutar el proceso anterior, requerimos rastrear el movimiento de S . Esto se logró empleando un enfoque basado en la detección en tiempo real de sus esquinas y centroide (Fig. 5).

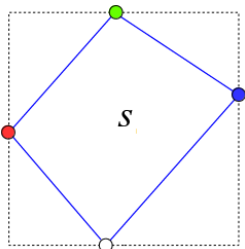


Fig. 5. Detección de esquinas del elemento planar. La línea punteada indica su mínima caja delimitadora.

Sobre la imagen binaria de S (el único objeto en escena) estimamos su centroide C_s y su mínima caja delimitadora D_s . Como se observa en la Fig. 5, con $S \cap D_s$ formamos un conjunto de puntos integrado por las esquinas de S . Utilizando C_s como elemento de referencia, ubicamos la posición de estas esquinas y, por lo tanto, podemos rastrear el movimiento de S . Cuando la cardinalidad $|S \cap D_s| > 4$, sabemos que S y D_s son idénticos y sus esquinas coinciden.

II.D. Mejoramiento de la calidad de la visualización

Para mejorar la visualización de las secciones o cortes calculados, fue necesario rectificar y redimensionar las imágenes. En general, tratamos aquí con el problema de mapear puntos de un rectángulo a un cuadrilátero irregular convexo. En 2D, esta transformación proyectiva H queda definida si se conocen las coordenadas de las cuatro esquinas en ambas figuras (Fig. 6).

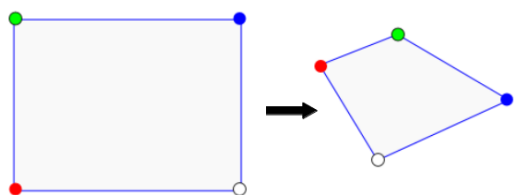


Fig. 6. Transformación proyectiva en 2D.

Con las esquinas de una nueva imagen I y las esquinas de S , de forma similar a II.B., estimamos H , donde $H(I)=AI$. De nuevo, utilizando DVS, encontramos la matriz de transformación $A_{3 \times 3}$. Proyectando los píxeles de I sobre S , estimamos sus posiciones en coordenadas del Kinect. Aplicando T , llevamos estos puntos hacia V y, finalmente, calculamos los valores de intensidad usando interpolación trilineal.

Los métodos computacionales descritos en este trabajo fueron implementados con OpenNI2, ITK y VTK. En la Fig. 7 se muestra una secuencia de cortes desplegados por S al intersectar a F (y por lo tanto, a V). Como puede observarse, un objeto rígido planar despliega cortes en el espacio definido por los marcadores del mundo real. Consideramos que este proceso de navegación de los datos volumétricos resulta más natural e intuitivo para el usuario final. En este caso, las imágenes de los cortes se superponen sobre la superficie de S .

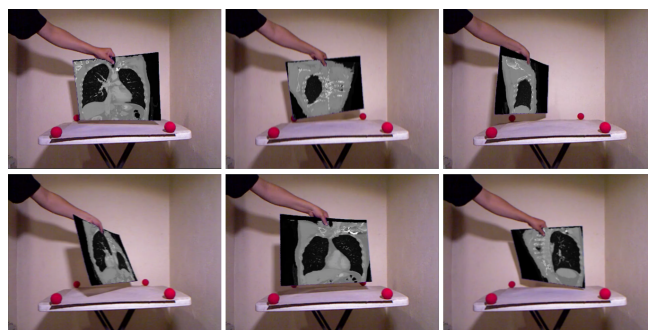


Fig. 7. Imágenes de algunas secciones desplegadas por el elemento planar al intersectar al volumen médico.

Hasta este punto, sin embargo, el tamaño reducido y la la distorsión de perspectiva de las imágenes, limitan la utilidad de la herramienta. En la Fig. 8 se observan algunos resultados obtenidos al implementar las mejoras descritas en II.D. Las secciones rectificadas se ajustaron a un tamaño de 640x480.

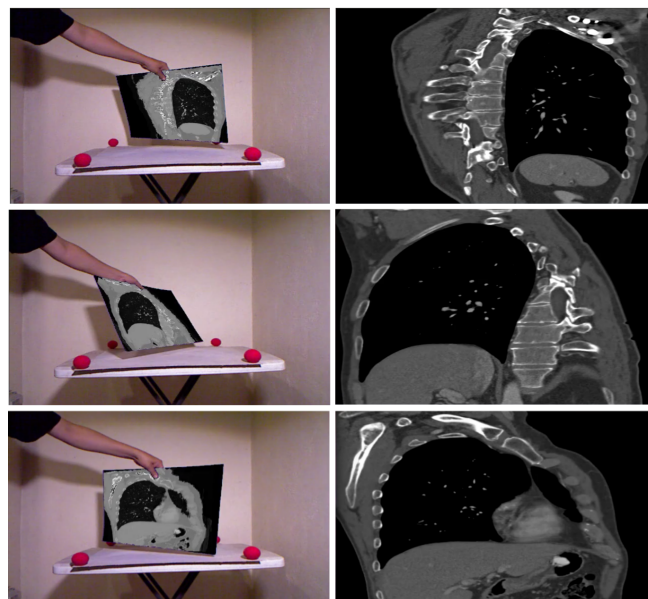


Fig. 8. Secciones rectificadas y redimensionadas.

IV. DISCUSIÓN

Para alcanzar los objetivos propuestos en este trabajo aún quedan problemas por resolver. Primero, los algoritmos de visión computacional utilizados no son robustos frente a cambios en las condiciones de iluminación y la oclusión de los marcadores. Segundo, con tomas monoculares no es posible estimar todas las orientaciones del elemento planar S en escena. En tales casos, el modelo falla y no recupera la imagen correcta del corte. Para superar estas limitaciones, es previsible que parte del trabajo futuro se concentre en la búsqueda de algoritmos de visión computacional más robustos, así como en el estudio de métodos que fusionen la información proveniente de otras fuentes (ej. cámaras RGB adicionales y/o sensores auxiliares).

Entre las mejoras consideradas para extender la funcionalidad del modelo se encuentra el uso de perfiles selectivos de segmentación. La idea básica sería poner al alcance del usuario la capacidad de segmentar estructuras específicas durante el procedimiento exploratorio (ej. tejido óseo). El ajuste tendría que ser sencillo para mantener el carácter intuitivo de la herramienta.

Finalmente, aún es necesario medir la eficacia del modelo propuesto bajo condiciones reales de trabajo. En adelante, resultará crucial recabar la opinión de especialistas médicos para establecer métricas de desempeño que valoren los resultados cuantitativamente.

V. CONCLUSIÓN

El diagnóstico radiológico certero y oportuno es crucial para salvaguardar la salud de los pacientes. En este artículo se presentan los avances hacia un modelo de navegación que debería permitir al usuario final explorar datos volumétricos de origen médico de una manera más natural e intuitiva. La herramienta fue concebida para prescindir de interfaces complejas y concentrar el esfuerzo cognitivo del especialista en el análisis de los datos del paciente. Creemos que en etapas futuras será posible construir un instrumento de diagnóstico de verdadera utilidad en clínicas y consultorios.

RECONOCIMIENTOS

Los autores agradecen el apoyo del Consejo Nacional de Ciencia y Tecnología (CONACYT) para la realización de este trabajo.

BIBLIOGRAFÍA

- [1] Bernhard Preim, Alexandra Baer, Douglas Cunningham, Tobias Isenberg, Timo Ropinski. "A survey of perceptually motivated 3d visualization of medical image data." *Computer Graphics Forum*, vol. 35, no. 3, pp 501-525, 2016.

- [2] Kamyar Abhari, John S. H. Baxter, Ali R. Khan, Terry M. Peters, Sandrine De Ribaupierre, Roy Eagleson. "Visual enhancement of MR angiography images to facilitate planning of arteriovenous malformation interventions." *ACM Transactions on Applied Perception*, vol. 12, no. 1, pp 1-15, 2015.
- [3] Daniel Haak, Charles-E. Page, Thomas Martin Deserno. "A Survey of DICOM Viewer Software to Integrate Clinical Research and Medical Imaging." *Journal of Digital Imaging*, vol. 29, no. 2, pp 206-215, 2016.
- [4] Jungong Han, Ling Shao, Dong Xu, Jamie Shotton. "Enhanced computer vision with microsoft kinect sensor: A review". *IEEE transactions on cybernetics*, vol. 43, no. 5, pp 1318-1334, 2013.
- [5] Savita Agrawal, Deepak Kumar Xaxa. "Survey on image segmentation techniques and color models." *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp 3025-3030, 2014.
- [6] Hanan Samet, Markku Tamminen. "Efficient component labeling of images of arbitrary dimension represented by linear bintrees." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp 579-586, 1988.
- [7] João Carlos Alves Barata, Mahir Saleh Hussein. "The Moore–Penrose pseudoinverse: A tutorial review of the theory." *Brazilian Journal of Physics*, vol. 42, no. 1-2, pp 146-165, 2012.
- [8] Gene Golub, William Kahan. "Calculating the singular values and pseudo-inverse of a matrix." *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, vol. 2, no. 2, pp 205-224, 1965.