

# Optimización de un método de detección y seguimiento de instrumental quirúrgico en video laparoscópico

M.E. García Regueras<sup>1</sup>, I. Oropesa García<sup>1,2</sup>, P. Sánchez-González<sup>1,2</sup>, A.M. Cano González<sup>1</sup>, F.M. Sánchez Margallo<sup>3</sup>, E.J. Gómez Aguilera<sup>1,2</sup>,

<sup>1</sup>Grupo de Bioingeniería y Telemedicina (GBT), Universidad Politécnica de Madrid, Madrid, España, {megarcia, psanchez, ioropesa, acano, egomez, }@gbt.tfo.upm.es

<sup>2</sup>Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Madrid, España

<sup>3</sup>Centro de Cirugía de Mínima Invasión Jesús Usón (CCMIJU), Cáceres, España, msanchez@ccmijesususon.com

## Resumen

*La localización del instrumental quirúrgico es de gran interés en Cirugía Mínimamente Invasiva (CMI). Determinar su posición con precisión es fundamental para el desarrollo de futuras aplicaciones tanto en la formación y evaluación de cirujanos como en el desarrollo de nuevos sistemas de navegación que asistan al cirujano durante la propia intervención quirúrgica. En la literatura existen técnicas de estimación basadas en el análisis de imágenes, siendo la principal limitación común la gran variabilidad de las imágenes de este tipo de cirugía, que dificultan el diseño de un método genérico capaz de obtener buenas precisiones para todas ellas. En este trabajo de investigación se propone una optimización de un método de detección y seguimiento automático de instrumental quirúrgico basado en el análisis de video endoscópico, incorporando filtros homomórficos, un nuevo algoritmo de segmentación y un seguimiento mediante inventariado. Los resultados obtenidos muestran que el método propuesto consigue una detección precisa y robusta para imágenes con diferentes características sin necesidad de ninguna adaptación previa.*

## 1. Introducción y Antecedentes

La cirugía laparoscópica es un tipo de Cirugía de Mínima Invasión (CMI) en la que el instrumental quirúrgico y la cámara son introducidos a través de pequeñas incisiones realizadas en el abdomen del paciente. Se trata de un tipo de cirugía local y cerrada.

La importancia que ha cobrado la CMI en las últimas décadas se ve reflejada en el continuo aumento de procedimientos que se abordan mediante estas técnicas quirúrgicas. Entre sus principales ventajas destacan el menor trauma tisular y el menor riesgo de complicaciones, que conllevan una reducción del dolor y del tiempo de recuperación lo que se traduce en un ahorro de costes hospitalarios y reducción de listas de espera [1].

En contra, si bien la CMI tiene importantes beneficios para el paciente, presenta por otra parte, ciertas desventajas para el cirujano. A diferencia de las operaciones tradicionales, en este tipo de técnicas el cirujano ve mermada su capacidad de actuación y de sensación [2]: la percepción sensorial se ve notablemente disminuida al desaparecer la interacción directa con los órganos y la información visual pasa a ser monoscópica,

perdiendo así la sensación de profundidad y ocasionando problemas de coordinación ojo-mano[3].

Dada tales limitaciones cobran importancia los procesos de formación y evaluación de destreza de cirujanos noveles. Además, la cirugía guiada por imagen (CGI) también gana interés, pues la inclusión en la escena de información visual mediante realidad aumentada mejora el desarrollo de la propia intervención.

En ambos campos es de gran importancia la extracción de los movimientos realizados por el instrumental. Conocer las trayectorias que el cirujano ha descrito permite realizar una evaluación de sus habilidades en cuanto al análisis y precisión de movimientos. En CGI, la localización del instrumental con respecto al entorno operatorio, permite guiar al cirujano dotándole de información adicional.

En la literatura existen numerosos trabajos que se centran en la estimación de la localización espacial basada exclusivamente en el análisis de las imágenes laparoscópicas. Persiguen la identificación y seguimiento del instrumental mediante un análisis de la forma, color y geometría del mismo. Tonet et al.[5], segmenta el instrumental valiéndose de un marcador artificial, y extrae sus bordes mediante su aproximación a rectas con la ayuda de la transformada de Hough. Doignon et al. [5] emplean un algoritmo de identificación del instrumental basado en un análisis de las componentes HSV de la imagen y en una clasificación de las diferentes regiones que se identifican en la imagen. (3) Feng et al.[6] proponen un método de adquisición de la posición 3D del instrumental a partir de la secuencia de imágenes monoscópicas extraídas por el endoscopio, y las coordenadas 3D del extremo del instrumental determinadas por un sensor electromagnético.

En un trabajo previo se presentaron el método de localización 3D basado en el video laparoscópico y su precisión [7]. En él se señalaba entre sus futuros trabajos, la mejora del preprocesado de la imagen como clave a la hora de aumentar la fiabilidad y la velocidad del algoritmo. El objetivo del presente trabajo de investigación es, por tanto, mejorar la precisión en la detección y aumentar la robustez del método, generalizándolo para cualquier secuencia quirúrgica. Para ello se modifica el algoritmo de detección incorporando

nuevas técnicas de realce de imágenes y se introduce el seguimiento del instrumental frame a frame.

## 2. Metodología

En el procesado de imagen requerido para la detección y seguimiento de los bordes del instrumental, se diferencian las siguientes etapas: (1) la etapa de **preprocesado** cuyo objetivo es conseguir la imagen óptima sobre la cual aplicar el algoritmo de detección de bordes; (2) la fase de **identificación de bordes**, que extrae la información de la imagen estática; y (3) la parte de **seguimiento** realizado *frame a frame*.

### 2.1. Preprocesado de la imagen

Las diferentes técnicas de realce que se emplean en esta etapa y que permiten una mejora de la imagen de trabajo son:

1. Disminución de la influencia de la iluminación no homogénea.
2. Conversión a un espacio de color adecuado (HSI).
3. Modificación de las componentes de color acentuando aquellas de interés.

La elección del espacio HSI como el más apropiado se debe a: (1) el interés en trabajar en un espacio donde la componente de intensidad (I) está desacoplada de la información cromática; y (2) el hecho de que las componentes de tono (H) y saturación (S) están íntimamente relacionadas con la forma en que los humanos percibimos el color y pueden ser empleados para la identificación y discriminación de objetos.

#### *Disminución de la iluminación no homogénea*

Se propone una técnica que se basa en el modelo de iluminación en el que el valor de intensidad de cada píxel viene dado por:

$$f(x, y) = i(x, y)r(x, y)$$

Donde  $i(x, y)$  es la iluminación de la escena en cada píxel y  $r(x, y)$  es la reflectancia. Típicamente la iluminación será una señal centrada en banda baja mientras que la reflectancia será un proceso de mayor variación espacial, y, por ello, de mayor contenido de alta frecuencia [9].

A partir de esta consideración se pueden separar las dos componentes mediante un filtrado homomórfico. Es decir, se aplica una operación logarítmica que convierte los logaritmos en sumas:

$$\ln(f(x, y)) = \ln(i(x, y)) + \ln(r(x, y))$$

de forma que un simple filtrado lineal permite separar de manera sencilla estas componentes, y, a partir de este momento, llevar a cabo un procesado paralelo por cada uno de los canales resultantes. Se podrá así corregir el problema de la iluminación no homogénea sin distorsionar la información de la imagen contenida en la reflectancia.

#### *Modificación de las componentes de color*

La saturación es una característica importante para la detección de regiones grises en imágenes. Una baja saturación indica píxeles poco coloreados y una alta, se corresponde con píxeles de color muy puros [8].

En esta etapa se emplea el espacio HSV, que al igual que el HSI también pertenece a la familia intuitiva de modelos de color. Se diferencia de este por los valores de saturación y de la componente acromática, que en este caso recibe el nombre de valor.

El cálculo de la saturación para este modelo se deduce a partir de la fórmula descrita en [9].

$$\text{Si } \max(R + G + B) \neq 0 \rightarrow S' = 1 - 3 \frac{\min(R, G, B)}{\max(R + G + B)}$$

$$\text{Si } \max(R + G + B) = 0 \rightarrow S' = 0$$

Esta nueva definición de la saturación,  $S'$ , consigue una mejor discriminación entre los píxeles cromáticos y no cromáticos. En el modelo HSV los colores pierden valor de saturación cuando se oscurecen. Un inconveniente que presenta es ser más sensible a los cambios de brillo, en especial en los píxeles cromáticos.

Debido a los problemas que conlleva la iluminación no homogénea, es interesante trabajar en un espacio donde la iluminación esté desacoplada del color. Por otro parte, la definición de saturación del modelo HSV es un mejor discriminador de píxeles cromáticos y no cromáticos. Así, se define un nuevo modelo HSI'. Este consiste en el modelo HSI donde ahora la componente saturación,  $S'$ , se calcula según la definición del modelo HSV.

### 2.2. Algoritmo de identificación de bordes

El algoritmo de detección de bordes se diseña teniendo en cuenta las principales características del instrumental en la imagen que mejor permitan su identificación: color oscuro y bordes rectos del instrumental.

De este modo los principios del algoritmo son los siguientes:

1. Detección de bordes en la imagen mediante la aplicación del operador de Sobel.
2. Segmentación del área de interés: eliminación de los bordes que no pertenecen al instrumental.
3. Identificación de los bordes del instrumental.

#### *Segmentación del área de interés*

El algoritmo desarrollado en este trabajo de investigación para segmentar la zona de interés consta de los siguientes pasos:

1. **Selección de píxeles acromáticos.** Se parte de la imagen en el espacio HSI'. Se aplica una segmentación automática, siguiendo el método Otsu [10]. La imagen quedará dividida en dos: por una parte las zonas cromáticas (parte roja de la cavidad) y por otra las acromáticas, es decir, el instrumental, zonas de sombra y zonas de la cavidad con poco color, que presentan tonos grisáceos o blanquecinos.

- Discriminación entre el instrumental y zonas con poco color.** Se segmenta la componente de Intensidad (I). Los píxeles que en el paso anterior fueron clasificados como acromáticos se ponen a 0, y se hace una nueva segmentación según el método Otsu. La imagen resultante se divide del siguiente modo: por un lado estarán aquellas zonas blancas de gran intensidad o con brillo, y por otra parte las zonas oscuras (vástago del instrumental y zonas de sombra).
- Composición de las dos máscaras anteriores.** Se aplica el operador AND a las máscaras resultantes los 2 pasos anteriores. Como resultado se tiene una máscara en la que los píxeles obtenidos son en su mayoría correspondientes al vástago del instrumental.
- Eliminación de líneas y píxeles aislados** mediante una erosión y una dilatación con el mismo elemento.

#### Identificación de los bordes del instrumental:

Combinando la información dada por el operador de detección de bordes y la obtenida de la parte de segmentación, se obtiene la máscara final sobre la cual se buscarán los dos bordes del instrumental quirúrgico. Esto se hace según el método de búsqueda de rectas mediante la transformada de Radon descrito en [7]. A continuación se eliminarán de la máscara estos píxeles, así como los correspondientes a las rectas inmediatamente superiores e inferiores para evitar el efecto de doble borde. Finalmente se aplicará de nuevo la transformada de Radon para la búsqueda del segundo borde.

#### 2.3. Algoritmo de seguimiento de instrumental

El algoritmo de seguimiento propuesto en este trabajo consiste en un método de enventanado en el espacio transformado. La búsqueda de rectas no se hace en el total de la imagen, sino que existe una zona de interés en la que se espera encontrar el instrumental quirúrgico.

La motivación es doble:

- Reducir el coste computacional y por lo tanto el tiempo de procesado.
- Disminuir la probabilidad de detección errónea.

Existe un mecanismo de detección de errores encargado de verificar si la detección y seguimiento son correctos. En caso contrario se reinicia el algoritmo volviendo a la detección simple y el cálculo de una nueva ventana.

#### Cálculo de la ventana o zona de interés

La ventana se aplica en el plano transformado de Radon. Una vez que se ha realizado una detección, se calcula cuales son los valores de  $(\rho, \theta)$  para las dos rectas. La ventana estimada será común para ambas rectas y vendrá dada por las siguientes premisas:

- La variación de la pendiente de las rectas entre frames consecutivos no será superior a 3 grados.

$$[\min(\theta_1, \theta_2) - 3, \max(\theta_1, \theta_2) + 3]$$

- El desplazamiento de una recta entre frames, será menor de 20 píxeles.

$$[\min(\rho_1, \rho_2) - 20, \max(\rho_1, \rho_2) + 20]$$

Si el algoritmo de detección no encuentra rectas en la ventana definida, se aumenta el tamaño de la misma. Así los límites de la región de búsqueda se incrementan en 2 unidades, tanto en  $\rho$  como en  $\theta$ , hasta encontrar los bordes del instrumental.

### 3. Resultados y Discusión

Para la validación de los métodos propuestos en este trabajo de investigación se han utilizado secuencias de imágenes de operaciones por laparoscopia realizadas sobre modelos porcinos. Se han analizado un total de 992 frames pertenecientes a 8 secuencias. Estas han sido elegidas de modo que presenten características diferentes entre si: (1) tomándolas con dos endoscopios diferentes, con lo que el color, la iluminación y el tamaño de las imágenes varían; (2) variando los movimientos descritos por el instrumental cambiando la trayectoria, la velocidad o la distancia a la cámara.

En primer lugar se muestran los resultados del método de segmentación diseñado, encargado de eliminar aquellos bordes que no pertenecen al área de interés.

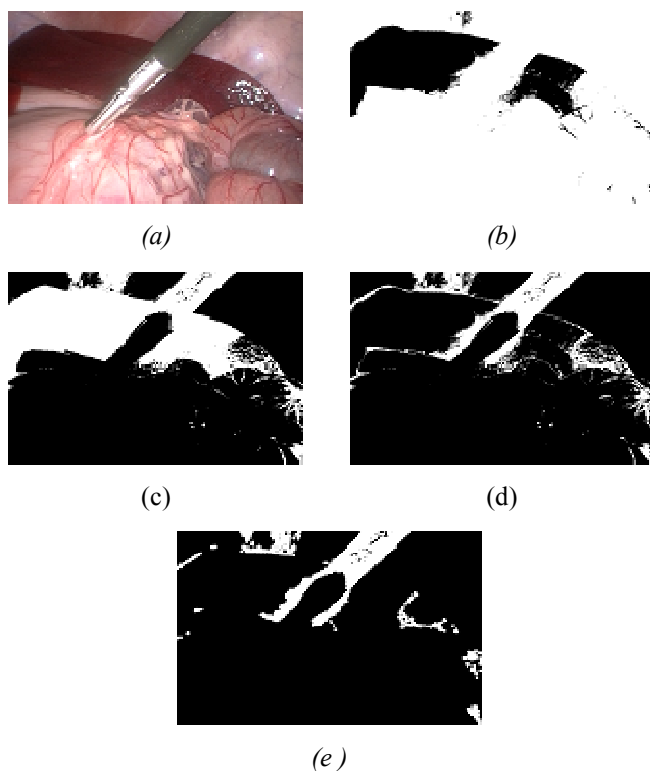


Figura 1: Método de segmentación del área de interés: (a) Imagen original. (b) Primera segmentación. (c) Segunda segmentación. (d) Composición de máscara 1 y máscara 2. (e) Máscara final

Como se observa en la máscara final obtenida, el instrumental se ha extraído correctamente. Sin embargo el método presenta un inconveniente: no es capaz de filtrar fácilmente aquellas regiones de la imagen con mucha sombra, lo cual podrá dar lugar a zonas de error.

A continuación se muestra la mejora conseguida con el algoritmo de seguimiento. En los siguientes frames se

comprueba como el seguimiento (rectas verdes) corrige en el *frame#329*, el borde erróneo que se obtendría solo con detección (rectas azules), al estar limitada ahora la región de búsqueda.

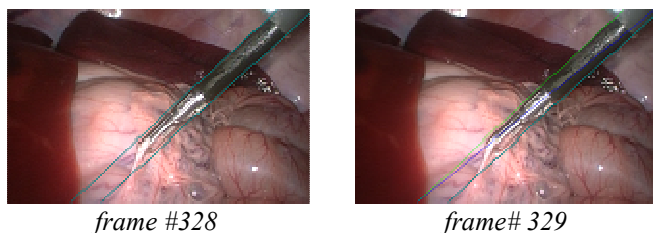


Figura 2: Detección simple (azul) vs. Detección y seguimiento(verde)

Finalmente se presentan los resultados obtenidos con el método de detección y seguimiento diseñado para el total de las secuencias. La siguiente figura contiene un *frame* representativo de cada una de las 8 secuencias analizadas. Se comprueba que las detecciones son correctas y la precisión buena. Además los resultados son similares independientemente de las características de la secuencia analizada y sin necesidad de ningún tratamiento previo dependiente de la secuencia. Se demuestra así la propiedad de “detección automática” del algoritmo.

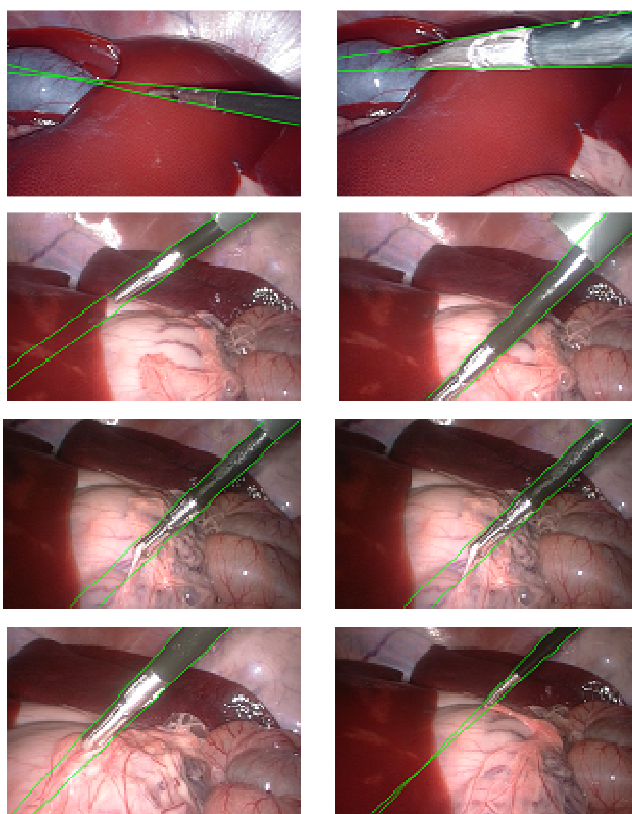


Figura 3: Resultados del método de detección y seguimiento para frames pertenecientes a 8 secuencias diferentes

La principal limitación son los cambios de color que experimenta el instrumental a lo largo de las secuencias. El método está diseñado para imágenes donde el color del vástago del instrumental es oscuro. Por lo tanto en aquellas imágenes en los que por algún motivo este se aclare, el algoritmo puede fallar.

Una extensión interesante a este trabajo sería el desarrollo de un mecanismo de detección de fallos robusto. Si siempre que el algoritmo fallase o se perdiese se reiniciase la ventana de seguimiento, el número de detecciones erróneas a lo largo de una secuencia disminuiría.

## 4. Conclusiones

Se ha optimizado un método automático de detección y seguimiento de instrumental quirúrgico a lo largo de los vídeos laparoscópicos sin introducir marcadores adicionales.

Las principales aportaciones de este trabajo de investigación son: (1) la extracción automática de los bordes del instrumental con buena precisión. Esto se consigue gracias al tratamiento previo de la imagen para la corrección de artefactos y al nuevo método de segmentación planteado. (2) La robustez introducida por el algoritmo de seguimiento por enventanado. Este corrige en su mayoría, las posibles detecciones erróneas obtenidas mediante detección simple.

## Referencias

- [1] Lamata, P, Methodologies for the Analysis, design and evaluation of laparoscopic surgical simulators. *Similar Network of Excellence*. UCL Presses Universitaires de Louvain, 2006.
- [2] Somolinos Sánchez JA, Avances en robótica y visión por computador. Univ de Castilla La Mancha, 2002, pp. 247.
- [3] Rattner DW, Apeltgren KN, W.S Eubanks. The need for training opportunities in advanced laparoscopic surgery. The residents' perspective. *Surgical Endoscopy* vol.15, num 10, 2001, pp. 1066-1070.
- [4] Tonet O, Ramesh TU, Megali G, Dario P, Tracking endoscopic instruments without localizer: image analysis-based approach, *Stud. Health Technol. Inform.*, vol. 119 2006, pp. 544-549
- [5] Doignon C, Nageotte F, Mathelin M, Segmentation and Guidance of Multiple Rigid Objects for Intra-operative Endoscopic Vision, *Lecture Notes in Computer Science*, 2007., vol. 4358, pp. 314-327,
- [6] Feng C, Rozenblit JW, Hamilton AJ, A hybrid in a laparoscopic surgery training system. *Proc. of the 14th Annual IEEE Int. Conf. & workshops on the Eng. of Computer-based systems (ECBS'07)*, 2007.
- [7] Cano AM, Gayá F, Lamata P, del Pozo F, Sánchez-Margallo FM, Gómez EJ, Método de seguimiento 3D del instrumental quirúrgico mediante análisis de vídeo laparoscópico, *Libro de Actas del XXVI Congreso Anual de la Sociedad Española de Ingeniería Biomédica, Valladolid*, 2008, pp. 257-260.
- [8] Ohta Y, Kanade T, Sakai T. Color information for region segmentation. *CVGIP*, 13:222-242, 1980.
- [9] Gonzalez RC, Woods RE, Eddins SL, *Digital Image Processing*, Prentice Hall 2nd Edition, 2001.
- [10] Otsu N, A threshold selection method from gray-level histogram, *IEEE Transactions on System Man Cybernetics*, Vol. SMC-9, No. 1, 1979, pp. 62-66.