

## VISIÓN ARTIFICIAL PARA DETECCIÓN Y UBICACIÓN ESPACIAL DE INSTRUMENTAL LAPAROSCÓPICO

**César A. Lizarraga<sup>a</sup>, Luis R. Ortega<sup>a</sup>, Bricelis J. Urbina<sup>a</sup> y Omaira Rodríguez<sup>a</sup>**

*<sup>a</sup>Centro de Computación Gráfica, Universidad Central de Venezuela, Paseo Los Ilustres  
Urb. Valle Abajo. Apartado Postal 20513. Caracas 1020-A, Venezuela, bricelis@ciens.ucv.ve,  
<http://ccg.ciens.ucv.ve/>*

**Palabras claves:** Visión Artificial, Reconocimiento de Objeto, Ubicación Espacial.

**Resumen.** El Centro de Computación Gráfica (CCG) en conjunto con el Instituto de Cirugía Experimental (ICE) de la Universidad Central de Venezuela (UCV) y el Instituto Nacional de Bioingeniería (INABIO) están desarrollando un simulador virtual para entrenamiento en cirugía laparoscópica. El componente de hardware del simulador está conformado por un simulador mecánico y un subsistema de rastreo. El subsistema de rastreo está basado en un dispositivo comercial llamado ISOTRAK® II, el cual devuelve al computador la posición y rotación de una pinza quirúrgica cuando esta se encuentra en la cavidad del simulador mecánico. Además del alto costo de este equipo, presenta limitaciones en cuanto al número de objetos rastreados simultáneamente y a la forma de interactuar con estos instrumentos. En este trabajo se presenta un sistema basado en Visión Artificial capaz de reconocer y rastrear el instrumental laparoscópico utilizado en el simulador virtual, presentando una manera alterna de interacción económica, portable e innovadora. El sistema propuesto utiliza un conjunto de cámaras digitales para capturar la imagen del instrumental dentro del simulador mecánico. Estas imágenes pasan por un procesamiento que permite obtener las características relevantes que son la entrada al módulo del sistema que permite obtener la posición y orientación de estos instrumentos en el espacio acotado por el simulador mecánico.

## 1 INTRODUCCIÓN

LAPAROS es un proyecto que se desarrolla multidisciplinariamente entre el Centro de Computación Gráfica (CCG), el Instituto de Cirugía Experimental (ICE) de la Universidad Central de Venezuela y el Instituto Nacional de Bioingeniería (INABIO). Este proyecto presenta un simulador virtual para entrenamiento en cirugía laparoscópica que cuenta con un conjunto de ejercicios virtuales que se deben llevar a cabo para el aprendizaje del uso correcto del instrumental laparoscópico. Estos ejercicios permiten desarrollar y fortalecer las habilidades requeridas por esta cirugía a través de la práctica y el entrenamiento (Urbina, B. et al, 2005).

El simulador está conformado por componentes de hardware y software. El componente de hardware, lo conforma principalmente un simulador mecánico y un subsistema de rastreo. El simulador mecánico fue desarrollado por el INABIO y consta de una estructura llamada SIMULAP V-1 ilustrada en la figura 1, el cual asemeja la forma de un abdomen humano y en donde se insertan diversos instrumentos quirúrgicos que se utilizan durante una intervención laparoscópica. El subsistema de rastreo está conformado por el dispositivo comercial ISOTRAK® II (Polhemus, 2001) como se muestra en la figura 2, el cual devuelve al computador la posición y rotación del instrumento cuando este se encuentra en la cavidad del simulador mecánico.

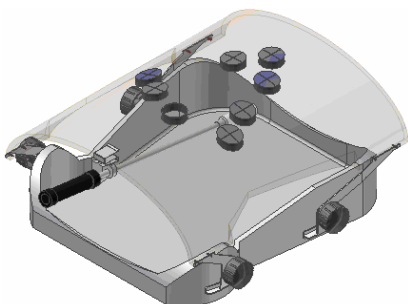


Figura 1: SIMULAP V-1



Figura 2: ISOTRAK® II

Para el simulador la técnica de interacción es primordial, por lo cual el entrenamiento debe ser lo más real posible. Hasta ahora, el ISOTRAK® II es la única manera de interactuar con el simulador, sin embargo su alto costo, las limitantes para simular ciertas acciones sobre el instrumento (por ejemplo, abrir y cerrar de una pinza), y la necesidad de rastrear el movimiento de más de un instrumento a la vez obliga a buscar nuevas alternativas para el subsistema de rastreo del simulador virtual (Climent, J. y Marés, P. 2004).

La visión artificial o comprensión de imágenes describe la deducción automática de la estructura y propiedades de un mundo tridimensional, tanto estático como dinámico, a partir de una o varias imágenes bidimensionales de ese mundo (Nalwa, 1993). En nuestro caso específico el sistema se encargará de detectar y obtener la ubicación del instrumental laparoscópico en un espacio tridimensional. En este trabajo se presenta una técnica de Visión Artificial como una alternativa para reemplazar el sistema de rastreo usado actualmente por LAPAROS.

## 2 DESCRIPCION DEL SISTEMA

En esta sección se describe la configuración de hardware del sistema y cada una de sus etapas a nivel de software dentro del proceso de Visión Artificial. El componente de hardware consta de dos cámaras digitales que deben estar colocadas de forma cruzada tal que su rango de visión abarque completamente el objeto de interés (figura 3). Adicionalmente el instrumental laparoscópico debe tener una configuración de colores predeterminedada (Holden, E-J. y Owens, R., 2001), como por ejemplo la figura 4. Para obtener mejores resultados, los colores ideales deben tener poca ocurrencia en el lugar de visualización, ya que el ambiente dentro del SIMULAP V-1 es uniforme, colores puros como rojo, azul o verde son ideales.

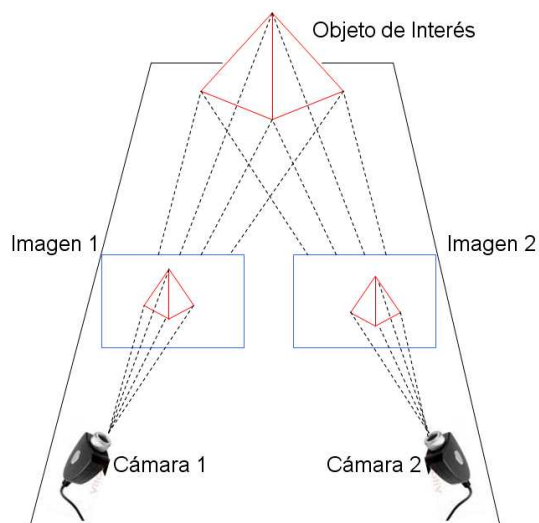


Figura 3: Estructura física del sistema



Figura 4: Configuración de colores

La figura 5 se muestra el diagrama del sistema, señalando cada una de sus etapas, como lo son: Calibración, Adquisición de la imagen, Preprocesamiento, Segmentación, Extracción de Características y Procesamiento de Alto Nivel. Ambos dispositivos pasan por una etapa de calibración para preparar los mismos y poder obtener los colores característicos del objeto. El siguiente paso es adquirir las imágenes de cada uno de los dispositivos, para realizar un preprocesamiento que elimine ruido, contraste y realce bordes. Luego se procede a segmentar las imágenes para detectar la existencia del instrumental laparoscópico en base a la codificación de colores suministrada previamente. En la etapa de extracción de características se identifican las posibilidades y se obtienen sus respectivos atributos en cada una de las imágenes. Por último dichas características serán la entrada para el procesamiento de alto nivel el cual con un conjunto de heurísticas permite diferenciar el objeto de interés de falsos positivos para luego determinar la posición del instrumental en el espacio tridimensional y desplegar los resultados en tiempo real en un ambiente simulado.

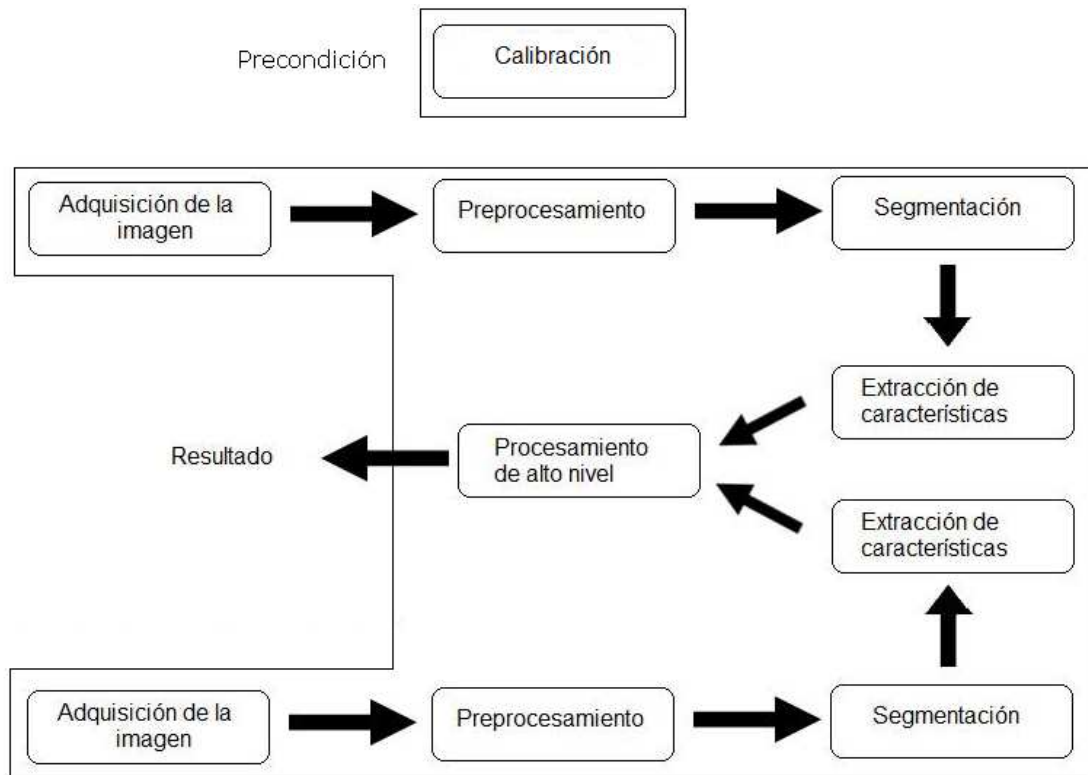


Figura 5: Diagrama global del sistema

### 2.1 Calibración:

En esta etapa del sistema se realiza la selección y configuración de los dispositivos de captura así como también la selección de los colores preestablecidos en el objeto de interés. Debido a la variedad de dispositivos de captura existentes y sus diferentes modos de uso, este módulo tiene como objetivo hacer transparente la gestión del dispositivo de captura a utilizar.

El primer paso del proceso es la selección de los dispositivos a utilizar, gracias al sistema operativo es posible solicitar información de los mismos, desplegarlos y facilitar su elección.

Luego es necesario obtener la matriz de proyección asociada a cada uno de los dispositivos seleccionados, esto se logra utilizando imágenes (capturadas previamente) que contienen un patrón especial que permite el cálculo de dicha matriz, en este caso específico el patrón utilizado es un tablero estándar de ajedrez.

Por último es necesario obtener el color característico del objeto de interés, para lograr esto se captura una imagen que contiene al objeto y se selecciona el color que representa al mismo (figura 6).

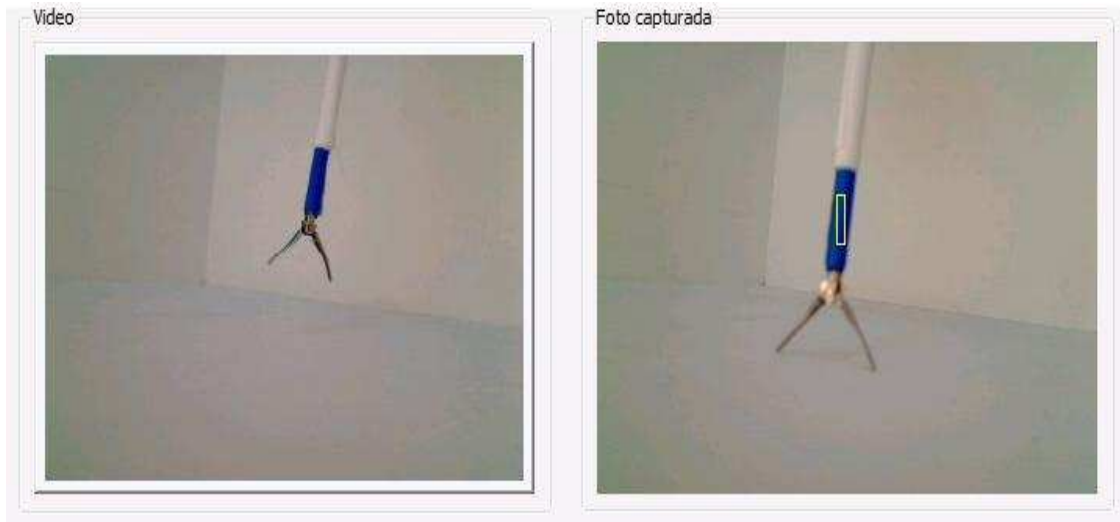


Figura 6: Configuración y selección de los colores a detectar.

## 2.2 Adquisición de la imagen:

En esta etapa del sistema se asume que los dispositivos ya fueron configurados y calibrados, en este módulo se verifica el estado y la disponibilidad de los mismos. Luego se procede a adquirir y representar la imagen capturada en forma digital. La cantidad de imágenes que se pueden procesar por iteración dependen del dispositivo de captura.

## 2.3 Preprocesamiento

Debido a la necesidad de corregir los errores introducidos por los dispositivos de adquisición de imágenes y el ambiente es necesario realizar un tratamiento previo. Algunos de estos errores pueden ser ruido, iluminación inadecuada, poca definición, entre otros. Por lo tanto es necesario corregir estos defectos y filtrar usando técnicas conocidas la información de la imagen, para facilitar el tratamiento en las etapas posteriores figura 7 y 8. En el sistema actual se utilizó un filtrado gaussiano con núcleo de tamaño 3x3.



Figura 7: Imagen capturada



Figura 8: Imagen suavizada

## 2.4 Segmentación

En esta etapa se procede a aplicar segmentación por umbral en base al color (definido en la etapa de calibración) obteniendo como resultado los posibles elementos a detectar, clasificados y ordenados. En esta parte del proceso se determina la existencia del instrumental laparoscópico en base a la presencia o ausencia de las marcas de colores. Para lograr un tamaño de muestra representativo y uniforme se aplica una erosión del fondo (con un elemento estructurante cuadrado de tamaño 6x6) al resultado anterior, esto con el fin de eliminar pequeñas estructuras logrando estructuras más grandes y manejables (figura 9).

En esta etapa se procede a aplicar una apertura (Opening) sobre la imagen obtenida en la etapa anterior, esto con el fin de distanciar y definir los colores de interés con respecto al resto de la imagen, logrando de esta manera una descomposición en objetos más simples y permite eliminar ciertos falsos positivos figura 9. El siguiente paso es realizar una segmentación por umbral en base a los colores (definidos en la etapa de calibración) obteniendo como resultado los elementos a detectar clasificados, ordenados y sin ambigüedades figura 10. En esta parte del proceso se determina la existencia del instrumental laparoscópico en base a la presencia o ausencia de las marcas de colores.

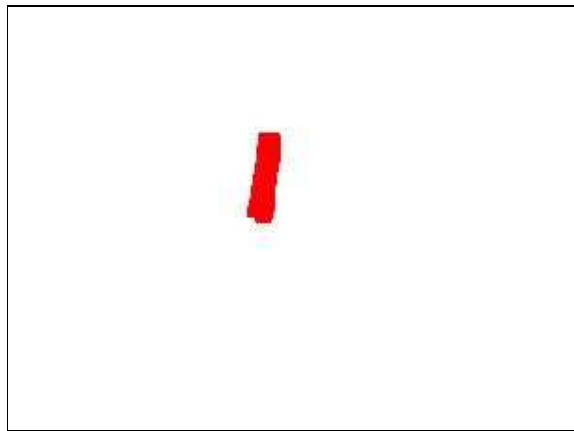


Figura 9: Imagen segmentada

### 2.5 Extracción de características.

Una vez culminada la etapa de segmentación se procesa la información obtenida y se genera como salida un vector de características el cual representa los objetos en la imagen. Cada una de estas posibilidades se basa en delimitar el área segmentada, calcular la posición relativa de estas áreas en la imagen y el tamaño de las mismas figura 10.

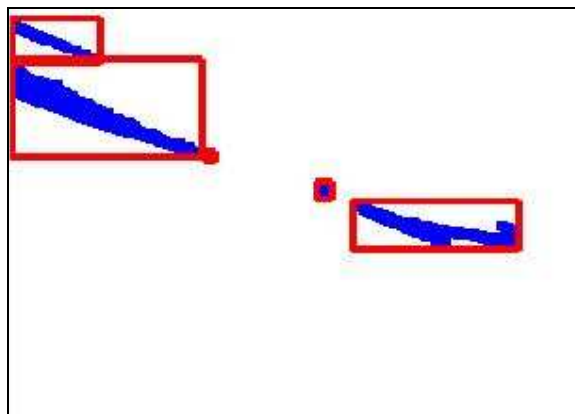


Figura 10: Resultado extracción de características (múltiples objetos)

### 2.6 Procesamiento de alto nivel.

En esta etapa se toman ambos vectores de características a los cuales se le aplican filtros heurísticos para descartar falsos positivos y encontrar una solución única. Luego se procede a estimar la posición en el espacio tridimensional utilizando geometría epipolar y se despliegan los resultados en un ambiente simulado.

La geometría epipolar (Hartley y Sisserman, 2000) es un mecanismo que permite estimar la posición tridimensional de un objeto utilizando dos vistas diferentes que contengan al mismo. Es decir, dado un punto  $x$  de la primera imagen y su correspondiente  $x'$  en la segunda imagen queremos encontrar la posición  $X$ . figura 11.

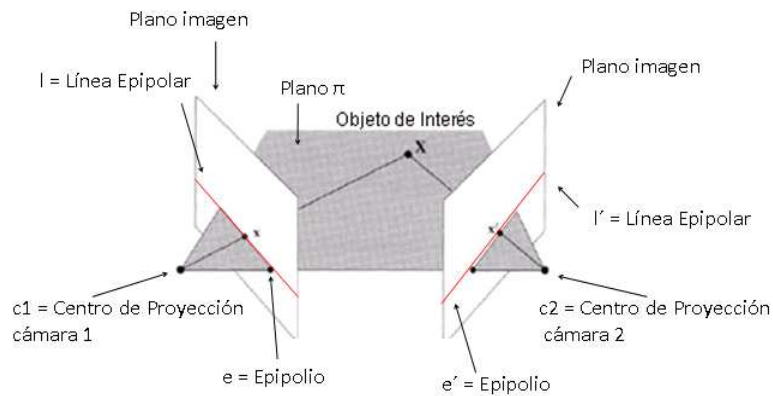


Figura 11: Geometría epipolar

En base a este mecanismo es posible adaptar el problema al caso actual. figura 12.

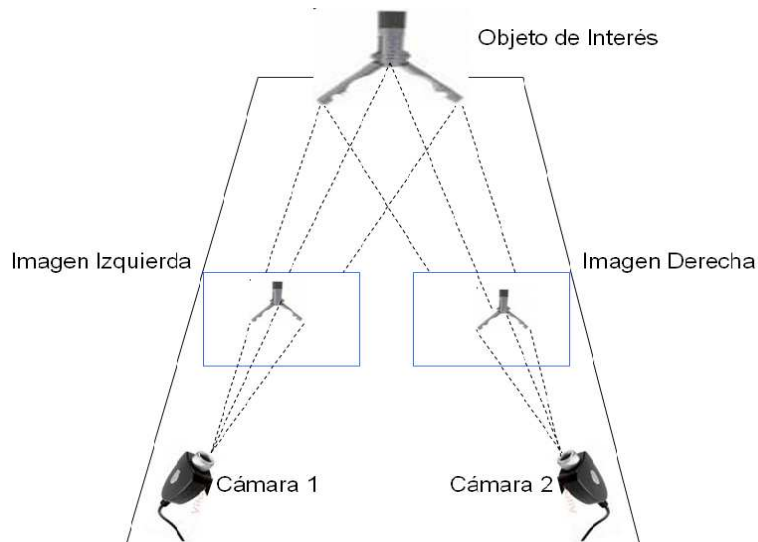


Figura 12: Caso específico

Cada uno de los dispositivos tiene asociada una matriz de proyección. Denotemos por  $P_1$  la matriz de proyección de la primera cámara y  $P_2$  la matriz de la segunda cámara.



$$P_1 = \begin{bmatrix} f_1 x & 0 & 0 & 0 \\ 0 & f_1 y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}; P_2 = \begin{bmatrix} f_2 x & 0 & 0 & 0 \\ 0 & f_2 y & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$f_1 x$  = distancia focal en x para la primera cámara.

$f_1 y$  = distancia focal en y para la primera cámara.

$f_2 x$  = distancia focal en x para la segunda cámara.

$f_2 y$  = distancia focal en y para la segunda cámara.

La relación entre un punto X en el espacio tridimensional y su proyección P se describe por las ecuaciones usuales de proyección perspectiva. Así, un punto x en una imagen de la primera vista se puede expresar como  $x = P_1 \cdot X$  y homológicamente  $x = P_2 \cdot X$  para la segunda vista

$$X = \begin{bmatrix} a \\ b \\ c \\ 1 \end{bmatrix}; x = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}; x' = \begin{bmatrix} x_1' \\ y_1' \\ 1 \end{bmatrix}$$

X, x, x' expresados en coordenadas homogéneas.

La idea es determinar el punto X del objeto de interés, expresado en coordenadas homogéneas.

Realizando el producto cruz en la ecuación  $x \times P_1 \cdot X = 0$  el sistema de ecuaciones resultante es:

$$\begin{aligned} f_1 y \cdot b - y_1 \cdot c &= 0 \\ f_1 x \cdot a - x_1 \cdot c &= 0 \\ f_1 x \cdot a \cdot y_1 - f_1 y \cdot b \cdot x_1 &= 0 \end{aligned}$$

Aplicando el mismo procedimiento para  $x' \times P_2 \cdot X = 0$  el sistema de ecuaciones resultante es:

$$\begin{aligned} f_2 y \cdot b - y_1' \cdot c &= 0 \\ f_2 x \cdot a - x_1' \cdot c &= 0 \\ f_2 x \cdot a \cdot y_1' - f_2 y \cdot b \cdot x_1' &= 0 \end{aligned}$$

Las matrices asociadas a cada uno de estos sistemas tienen rango 2. Luego, el sistema que relaciona las dos vistas con el objeto de interés lo podemos expresar como:

$$A \cdot X = 0 \iff \begin{bmatrix} 0 & f_1 y & -y_1 & 0 \\ f_1 x & 0 & -x_1 & 0 \\ 0 & f_2 y & -y_1' & 0 \\ f_2 x & 0 & -x_1' & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

La matriz  $A$  es singular, para resolver el sistema  $A.X = 0$  se utilizó la descomposición SVD (Golub y Van Loan, 1996). La solución del sistema obtiene una coordenada en tres dimensiones que representa la ubicación del objeto de interés expresado en coordenadas homogéneas.

Dicha coordenada es transformada a una escala fija para su posterior despliegue dentro de una escena cúbica, es posible visualizar este escenario desde diferentes perspectivas.

### 3 PLATAFORMA

El sistema esta conformado por componentes de hardware y software. Para el componente de hardware se utilizaron dos cámaras Web modelo Logitech Quickcam Pro 9000, un instrumental laparoscópico codificado con colores predeterminados para su detección, un simulador mecánico y un computador con al menos un procesador INTEL de 3 Ghz. Se desarrollo un sistema de software que controla y administrar los dispositivos de captura, procesa imágenes, software especializado en proyecciones para ubicación tridimensional en un espacio determinado y una interfaz donde se despliegan gráficamente los resultados obtenidos.

### 4 RESULTADOS

Una técnica sencilla de Visión Artificial permite adquirir la posición de un objeto de un sistema virtual. El sistema se basa en marcas de colores para realizar la detección y en geometría epipolar para estimar la posición. El sistema se desarrolló en C++, utilizando la librería OpenCV (Intel, 2006) para el procesamiento digital de imágenes, la librería QT (Trolltech, 2008) para la creación de interfaces de usuario y OpenGL (Silicon Graphics, 2006) para el despliegue en tres dimensiones.

El uso de la Librería OpenCV fue ideal ya que está optimizado para el procesamiento de imágenes sobre procesadores Intel lo cual acelera el procesamiento y gestión de las mismas obteniendo así un mejor tiempo de respuesta con respecto a la implementación propia o de otras librerías. De esta manera es posible procesar las imágenes y obtener la posición en tiempo real.

El uso de QT permitió el desarrollo rápido, sencillo y usable de interfaces lo cual aceleró y facilitó el proceso de desarrollo. El lenguaje utilizado para la programación fue C++ el cual permitió crear un sistema portable, escalable, reutilizable, interorperable y orientado a objetos.

Es recomendable utilizar colores poco convencionales en el entorno para las marcas del instrumental laparoscópico debido a que facilitan el procesamiento y ayudan a eliminar falsos positivos creando un sistema mas eficaz.

El computador a utilizar debe presentar al menos dos interfaces de conexión con los dispositivos de captura que cuenten con un BUS de datos distinto para no saturar al mismo, también es recomendable para obtener resultados satisfactorios el uso de dos dispositivos de captura iguales.

Actualmente el sistema es capaz de reconocer dos (2) codificaciones de colores diferentes simultáneamente (objetos de interés), esto se debe a que en las prácticas médicas se utilizan dos instrumentos laparoscópicos por usuario. Esta sistema es adaptable para N codificaciones de colores dependiendo del caso.

## 5 AGRADECIMIENTOS

Agradecemos al Consejo de Desarrollo Científico y Humanístico de la UCV por el Trabajo financiado parcialmente. PTP 08-12-5294-2005.

## 6 REFERENCIAS

- Climent, J. y Marés, P. Sistema de Seguimiento en Tiempo Real para Intervenciones Quirúrgicas Asistidas. *Universidad Politécnica de Catalunya*. (2004).
- Golub, H., and Van Loan, F. Matrix Computations, *Johns Hopkins University Press*. (1996).
- Hartley, R. y Zisserman, A. Multiple View Geometry in Computer Vision. *Cambridge University Press*. (2000).
- Holden, E-J. y Owens, R. Visual Sign Language Recognition. *Lecture Notes in Computer Science*. (2001).
- Intel. Consultado el 14 de julio de 2008 de la World Wide Web: <http://www.intel.com/technology/computing/opencv/index.htm>. (2006).
- Nalwa, V. S. A Guided Tour to Computer Vision. Addison – Wesley. (1993).
- Polhemus. Consultado el 14 de julio de 2008 de la World Wide Web: <http://www.inition.co.uk/inition/pdf/isotrak.pdf>. (2001).
- Silicon Graphics. Consultado el 14 de julio de 2008 de la World Wide Web: <http://www.opengl.org>. (2006).
- Trolltech. Consultado el 14 de julio de 2008 de la World Wide Web: <http://trolltech.com/products/qt/>.(2008).
- Urbina, B., Coto, E., Rodríguez, O., Miquilarena, R. y Cerrolaza, M. Laparos: Computer Assisted Laparoscopic Surgery Training. *Universidad Central de Venezuela*. (2005, Septiembre).