



Universidad
Zaragoza

Trabajo de Fin de Máster

SLAM visual para AR contextual en laparoscopia

Autor

Richard Elvira López-Echazarreta

Director

José María Martínez Montiel

Máster en Ingeniería Biomédica

Escuela de Ingeniería y Arquitectura

29 de Junio 2018



DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD

(Este documento debe acompañar al Trabajo Fin de Grado (TFG)/Trabajo Fin de Máster (TFM) cuando sea depositado para su evaluación).

D./D^a. Richard Elvira López-Echazarreta

con nº de DNI 16618876L en aplicación de lo dispuesto en el art.

14 (Derechos de autor) del Acuerdo de 11 de septiembre de 2014, del Consejo de Gobierno, por el que se aprueba el Reglamento de los TFG y TFM de la Universidad de Zaragoza,

Declaro que el presente Trabajo de Fin de (Grado/Máster)
Máster en Ingeniería Biomédica, (Título del Trabajo)
SLAM visual para AR contextual en laparoscopia

es de mi autoría y es original, no habiéndose utilizado fuente sin ser citada debidamente.

Zaragoza, 25 de Junio de 2018

Fdo: _____

Índice general

Resumen	4
1. Introducción	5
1.1. Estudio del caso médico	5
1.2. Técnicas de localización para AR	6
1.3. Objetivos	7
2. Métodos y materiales	9
2.1. Sistema de SLAM aplicado a laparoscopia	9
2.2. Imagen PET-TAC	10
3. Anclaje 3D en el sistema SLAM	12
3.1. Sistema de anclas	12
3.1.1. Cálculo del plano del ancla	13
3.1.2. Representación en el mapa del ancla	16
3.1.3. Refinamiento de la posición del ancla con todos los keyframes del mapa .	17
3.2. Inserción de anotaciones de AR	19
4. Alineamiento de la notación AR en el sistema SLAM	21
5. Creación de la anotación AR	23
6. Adquisición de datos de laparoscopia	25
6.1. Adquisición de la secuencia	25
6.2. Calibración del laparoscopio	25
7. Experimentos	27
8. Conclusiones	29
9. Anexo	30
9.1. Confidencialidad de información del paciente	30
9.2. Formato de fichero OpenGL	30
9.3. Control de versiones	31
Bibliografía	32

Resumen

La realidad aumentada en cirugías mínimamente invasiva es una funcionalidad largamente buscada durante los últimos años. La principal barrera que se encuentra es la estimación precisa del endoscopio respecto de la cavidad del paciente.

Abordamos la localización espacial de las anotaciones de realidad aumentada, para lo cual se ha desarrollado un método de localización de la anotación que no requiere de elementos externos como marcadores, fijaciones o localización óptica.

Se han realizado una validación experimental mediante la colaboración con el servicio de urología del Hospital Clínico Lozano Blesa, focalizándose en un paciente real. El caso es una intervención en la cavidad abdominal para la extracción de una tumor en la zona de la próstata de la que se procesa la imagen médica PET-TAC para generar la anotación de AR y la secuencia tomada por el endoscopio en la que se va a integrar la anotación.

Las anotaciones que se agregan a la vista son creadas como modelos 3D a partir de la información de la imagen médica PET-TAC del paciente, en donde se realiza una segmentación del tumor en 3D. También se definen unas anclas anatómicas que definen la ubicación en la secuencia del endoscopio.

El sistema de posicionamiento de la anotación procesa el vídeo del endoscopio en tiempo real y mediante un algoritmo de SLAM se construye un mapa 3D de la cavidad abdominal, en el cual se ubica el endoscopio para conocer donde se encuentra respecto al mapa. En el mapa se definen unas anclas en referencias anatómicas designadas también en la anotación de AR, para realizar el alineamiento entre ellas y obtener la posición de la anotación en el mapa de la cavidad respecto a las anclas fijadas. Las anclas se pueden fijar en cualquier espacio del mapa de la cavidad, se utilizan estructuras anatómicas ya que son más sencillas de identificar a simple vista. La anotación de AR contiene información sobre la escala, por lo que una vez agregada se actualiza el mapa de la cavidad para establecer la verdadera escala de la zona de trabajo.

El sistema de localización de anotaciones AR a través del vídeo de la cirugía laparoscopia se ha realizado sin modificar el procedimiento quirúrgico y sin añadir ningún elemento adicional. Únicamente se necesita un computador estándar que procesa el flujo del vídeo del endoscopio monocular normalmente empleado en la intervención.

Capítulo 1

Introducción

Un gran número de cirugías se realizan usando técnicas de cirugía mínimamente invasiva (MIS), en este tipo de procedimiento se busca minimizar el impacto que tiene una operación tradicional en el paciente. El trabajo va a tratar un procedimiento de extirpación de una adenopatía en la cavidad abdominal. En una intervención de MIS en la cavidad abdominal el equipo quirúrgico realiza unas pequeñas perforaciones en la cavidad, por donde insertan la trocares por donde se introducen las herramientas para operar. Antes de comenzar la cirugía se insufla en el abdomen gas de dióxido de carbono (CO_2) para crear un espacio de trabajo donde realizar la intervención (figura 1.1a).

En todas las intervenciones de MIS se necesita un endoscopio con el que el cirujano puede ver la zona a operar, esta herramienta lleva su propia fuente de luz. La imagen del laparoscopio se envía digitalmente a un monitor.

El hecho de tener una imagen digitalizada del interior del paciente permite la implementación de nuevas técnicas. La anotación de realidad aumentada (de aquí en adelante AR) permite agregar información digital y que no existe en el mundo real a la imagen digitalizada del paciente.

La anotación proporciona ayuda visual para facilitar el desarrollo de la intervención, sin embargo depende de la anotación que se quiera insertar, en intervenciones quirúrgicas se quiere resaltar las estructuras anatómicas de interés, pero si la anotación no se encuentra acoplada con la estructura que esta representando entonces no resulta de utilidad.

Los casos por lo que una anotación de AR puede ayudar durante en una operación son diversos:

- Ubicación de estructuras anatómicas en zona homogénea.
- Localización de estructuras ocluidas.
- Facilitar la para localizar la posición del endoscopio respecto a la zona de trabajo.
- Liberar al cirujano de la tarea de reconocimiento de la estructura.
- Facilitar la toma de decisiones al tener una información precisa de la zona de trabajo.

1.1. Estudio del caso médico

El caso que se va a estudiar es un caso de linfadenectomía de rescate se realiza en aquellos pacientes con cáncer de próstata en cuyo seguimiento se detecta una recidiva bioquímica (mediante PSA) y en los que por métodos de imagen (TAC, RMN, PET-TAC) se diagnostica de enfermedad adenopática tumoral.

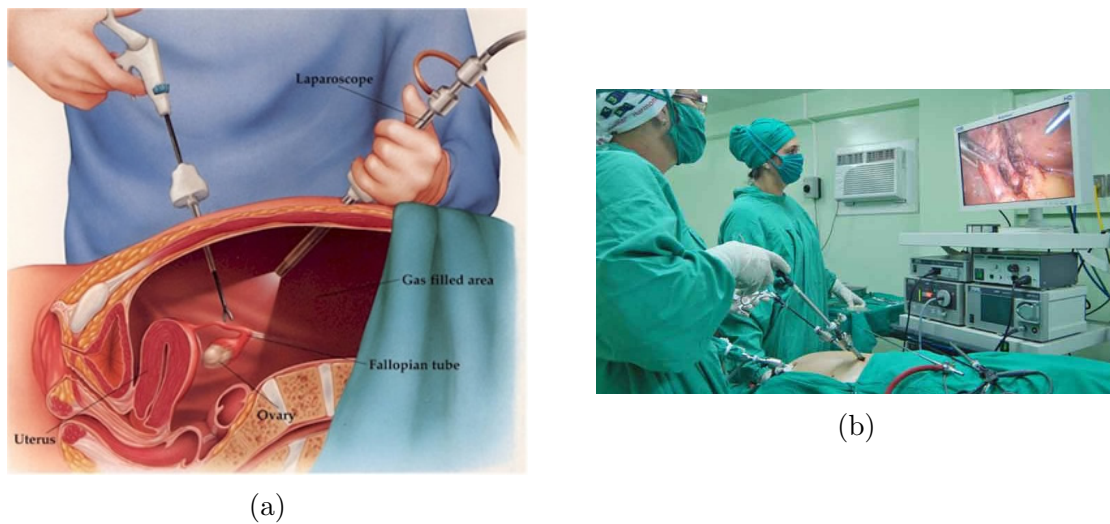


Figura 1.1: Funcionamiento interno de una laparoscopia (a). Intervención mediante laparoscopia (b).

El cirujano tiene que interpretar mentalmente la imagen tomográfica para poder inferir la localización del tumor sobre las imágenes del laparoscopio. El objetivo del proyecto es mostrar el tumor como una anotación de realidad aumentada sobre la imagen del laparoscopio.

El procedimiento quirúrgico sería:

1. Segmentación manual del tumor en sobre el TAC-PET. Será un modelo sólido estimado a partir de unos 10 cortes tomográficos equidistantes.
2. También del TAC-PET se extrae la localización 3D de 3 referencias anatómicas, que llamaremos anclas, que deben sean fácilmente identificables sobre las imágenes de laparoscopia. Estas referencias estará asociadas a partes que no sufran deformaciones no rígidas.
3. Durante la cirugía, después de liberar la zona pélvica, se deben identificar las tres referencias anatómicas correspondientes a las anclas pulsando sobre la imagen del laparoscopio.
4. Después de definir las 3 anclas, la localización del tumor se mostrará en tiempo real sobre la imagen del laparoscopio sirviendo de guía al cirujano para su localización.

La principal barrera para la visualización de la AR es la estimación precisa de la posición del endoscopio respecto del cuerpo del paciente. Para estimar esta posición relativa se empleará el sistema un sistema de SLAM que se ha demostrado capaz de procesar secuencias de laparoscopia y de gastroscopia [1] en vivo con cerdos.

1.2. Técnicas de localización para AR

El posicionamiento de la anotación de AR en el interior del paciente es una tarea que requiere de una gran precisión en la estimación de la posición relativa del endoscopio y de la anotación respecto de la cavidad observada. Se ha estudiado las técnicas empleadas[2] para utilizar AR en cirugía laparoscópica. Todas los métodos que proyectan anotaciones de AR a la visión del endoscopio requieren conocer la posición de la cámara respecto a la escena o a unos marcadores desde los que se define la posición de la anotación.

Los marcadores fiduciales son anclas ajenas al cuerpo que se sitúan en el paciente, las cuales se emplean para establecer un mapa sobre el que proyectar la anotación, el endoscopio se debe encontrar localizado en el espacio en el que se definen estos marcadores, para proyectar el punto

de vista adecuado a la anotación de AR. La posición de los marcadores se puede procesar a través de localización óptica[3] con cámaras fijas en la sala de la intervención, este método de seguimiento presenta inconvenientes, ya que se necesita adaptar todo el espacio de trabajo para visualizar los marcadores, además impide que el personal médico se mueva como necesite, ya que puede generar oclusiones en la visión de las cámaras causando que la anotación no se posicione correctamente.

Los sistemas de seguimiento óptico no tienen problemas con las escenas deformables, ya que se tiene la posición de las marcas en todo momento, siempre que la cámara la vea, de manera que la anotación se desplaza con el movimiento de las anclas.

Las técnicas que no requieren marcadores o sistemas ópticos de seguimiento se basan en la reconstrucción de la escena que proyecta el endoscopio para obtener su ubicación, de manera que se pueden establecer marcas para posicionar la anotación de AR. Este tipo de técnicas requieren de estructuras rígidas a las que anclas la anotación o marcadores fiduciales en el interior del cuerpo. En los primeras aplicaciones se utilizaron los huesos, que son las estructuras más rígidas del cuerpo, para sujetar la anotación de AR y localizarla a partir de ellos. En el caso de usar los marcadores fiduciales se deben colocar en el interior del cuerpo, lo que puede complicar la intervención. Este tipo de localización es más complejo de aplicar a escenas con grandes deformaciones.

1.3. Objetivos

Se va a desarrollar una técnica de poner anotaciones de AR en una intervención de cirugía en la cavidad abdominal, la cual contiene estructuras anatómicas con poca deformación.

Solo se va a emplear el vídeo vivo para definir la posición de la anotación de AR. Para ello se van a emplear herramientas de procesamiento de imagen con las que construir un mapa de la zona de trabajo en la que se van a definir unos puntos de referencia anatómicos que ayudaran a posicionar la anotación de AR. Se ha tomado esta decisión ya que es la menos intrusiva dentro de la metodología empleada en el quirófano, el personal sanitario no necesita emplear herramientas nuevas, por lo que su implementación es más sencilla.

Junto con la herramienta para poner la anotación de AR en el vídeo del endoscopio, se va a crear un visor PET-TAC que permita definir la anotación de AR con respecto a las referencias que se van a emplear durante la laparoscopia para establecer su posición.

Para alcanzar los objetivos se ha dividido el trabajo en bloques más sencillos, de forma que es más sencillo resolver cada problema por separado para luego integrarlo todo:

- Creación de puntos de referencia o anclas en el vídeo en vivo del endoscopio.
- Ajuste de la posición 3D de las anclas ya definidas durante la intervención.
- Integración de técnicas de informática gráfica para representar modelos 3D en la vista de la cámara.
- Alineamiento de las anclas para definir la posición del modelo.
- Creación de la anotación de AR a partir de la imagen PET-TAC.

Cada apartado se va a resolver y validar de manera independiente, ya que facilita la resolución de problemas al tener problemas separados. En todo momento estos puntos se resuelven pensando en la futura integración de todos ellos para evitar problemas de compatibilidad.

Para alcanzar los objetivos se va a utilizar un algoritmo de SLAM (Simultaneous Location And Mapping), que emplea técnicas de visión por computador para construir mapas de entorno 3D a partir de una secuencia de vídeo. Se va a explicar con más detalle en la sección 2.1.

La anotación de AR se va a definir como un objeto 3D (sección 5) que se va a proyectar en la visión endoscopio mediante la librería gráfica de código libre OpenGL[4], que permite definir modelos y entornos 3D en los que realizar renderizados de vista desde una posición definida.

Por último pero no menos importante se quiere facilitar la localización del tumor para ayudar a mejorar la precisión y rapidez con la que se realiza la intervención.

Capítulo 2

Métodos y materiales

2.1. Sistema de SLAM aplicado a laparoscopia

El algoritmo de SLAM del que se parte es ORB-SLAM2[5], el cual construye mapas de entorno 3D en tiempo real a partir de una secuencia de vídeo monocular.

La imagen captada por el endoscopio es procesada en tiempo real para extraer sus características (figura 2.1a) y emparejarlas con las características de imágenes previas. Al conseguir emparejar una gran cantidad de puntos se triangula la posición y orientación en 3D de la trayectoria del endoscopio y, al mismo tiempo, la posición en 3D de los puntos emparejados. La posición de estos elementos se representa en un sistema de referencia global, el cual se define en la posición de la primera imagen en la que se inicializa el mapa de entorno 3D (figura 2.1b), a partir de aquí el mapa de la cavidad abdominal. La inicialización del sistema requiere que haya suficiente paralaje entre los puntos para poder triangular tanto puntos como cámaras.

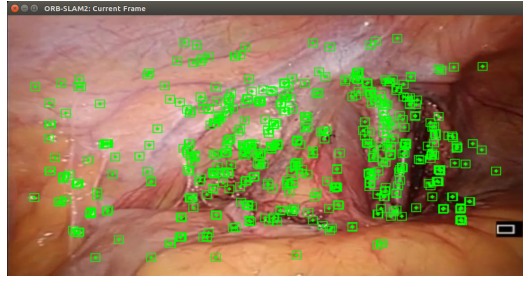
Cuando el mapa de la cavidad se ha iniciado, la posición del endoscopio se puede estimar de forma precisa siempre y cuando en la imagen que esté detectando se observen zonas ya incluidas en el mapa, la imagen que está tomando actualmente el endoscopio se llama *frame*. Los frames del endoscopio que pueden agregar información a la escena se promueven a *keyframes*, se insertan en el mapa junto con todos los puntos que visualiza, de manera que se pueden utilizar para estimar la posición del endoscopio con ellos. De esta manera el mapa de la cavidad crece a la vez que se realiza la exploración. De cada keyframe conocemos su posición y la imagen que lo describe.

La navegación se realiza por zonas conocidas en el mapa, en dichas zonas se puede ubicar el frame del endoscopio mediante el emparejamiento de las características del mapa. A medida que crece el mapa el endoscopio tiene más zonas conocidas donde estimar su posición.

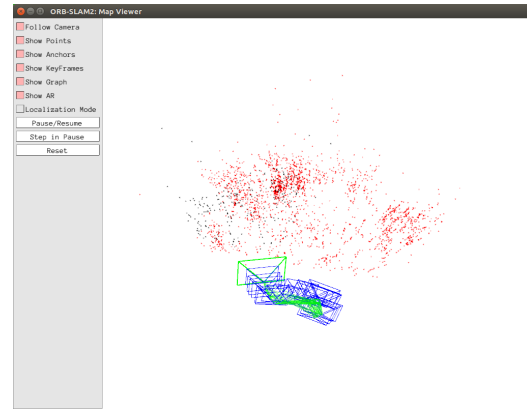
El mapa de la cavidad es persistente aunque se pierda la ubicación del endoscopio. En ese momento el algoritmo trata de buscar coincidencias entre las características del frame actual y las que están en el mapa de la cavidad, el endoscopio se vuelve a ubicar en el mapa cuando encuentre coincidencias. Después de una relocalización exitosa, el algoritmo continúa con las tareas de localización y expansión del mapa.

Durante la operación se utiliza un endoscopio monocular, es decir, que solo tiene un sensor de visión. Los sistemas monoculares son los más extendidos por su reducido coste, sin embargo tienen unas algunas limitaciones:

- La imagen que se obtiene no proporciona información sobre la escala real de la escena que se visualiza.
- La inicialización de un mapa con sensor monocular necesita que las características iniciales se observen en varios frames con suficiente paralaje para triangular la posición de los puntos y las cámaras.



(a) Extracción de características en la imagen.



(b) Mapa de entono en 3D.

Figura 2.1: Información obtenida con el algoritmo ORB-SLAM2.

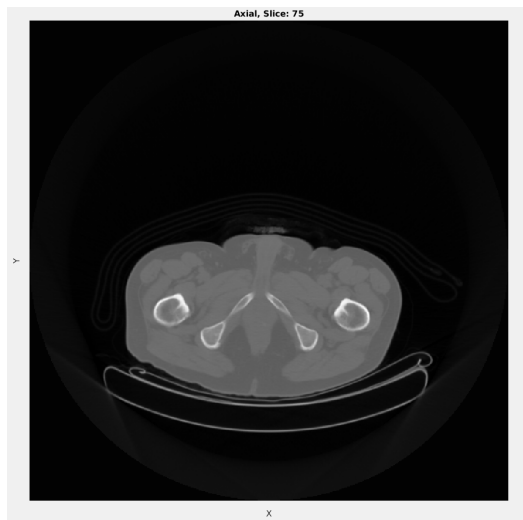
Se asumen conocidos y fijos los parámetros de calibración y distorsión de la cámara.

2.2. Imagen PET-TAC

La imagen tomográfica del PET-TAC se obtiene antes de la intervención, se encuentra en el formato DICOM[6] (Digital Imaging and Communication in Medicine). Este tipo de ficheros almacena la información de obtenida de la imagen médica de forma normalizada.

En este caso la imagen médica que se obtiene es una imagen 3D la cual se representa como un conjunto de imágenes 2D apiladas, en este tipo de representación una coordenada dentro de la imagen se identifica con un voxel, que es la estructura 3D de los píxel, estos voxel se representan como octaedros pero no son necesariamente del mismo tamaño en sus 3 ejes. Los ejes de la imagen se identifican con los planos empleados en medicina coronal, sagittal y axial (figura 2.2). Estas imágenes permiten al personal sanitario analizar el interior del paciente para emitir un diagnostico. Las imágenes permiten a los especialistas identificar en 3D las diferentes estructuras anómalas, como los tumores, y también las estructuras anatómicas.

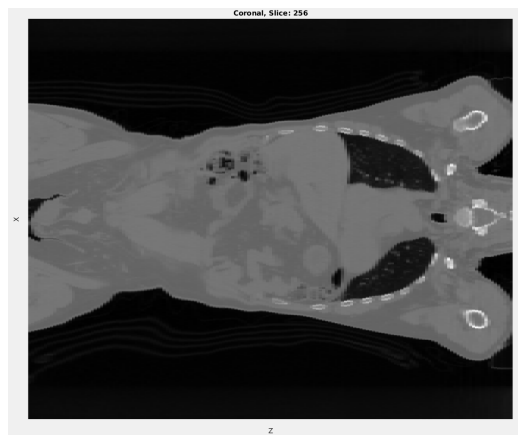
El formato DICOM también guarda el tamaño en milímetros de cada voxel en los 3 ejes, por lo que nos da información sobre la escala.



(a) Vista axial.



(b) Vista sagittal.



(c) Vista coronal.

Figura 2.2: Ejemplo de las 3 vistas de un mismo paciente según los planos axial, sagittal y coronal.

Capítulo 3

Anclaje 3D en el sistema SLAM

El algoritmo de SLAM nos permite construir el mapa de la cavidad abdominal como se ha explicado en la sección 2.1, cuando se haya construido el mapa se va a integrar información de AR contextualizada.

La información de AR se tiene que integrar en el mapa, para ello es necesario estimar la posición y orientación que debe ocupar en él. Sin embargo, la posición de todo lo que contiene el mapa es relativa respecto a la posición del keyframe con el que se ha inicializado el mapa, por lo que se debe implementar un sistema que permita colocar la anotación de AR en la posición anatómica correcta dentro del mapa y para que se quede contextualizada a pesar de que el endoscopio se mueva por la cavidad. Es muy importante lograr una buena precisión en la posición y orientación de la anotación AR.

El método para localizar la posición de la anotación no requiera de ninguna herramienta extra, tales como trackers posicionales o marcadores fiduciales. Para ello se define la posición de la anotación de AR a partir de tres puntos de referencia, a partir de los cuales se puede establecer el sistema de referencia local desde el que se ha definido la anotación de AR, por lo que es necesario que dichos puntos de referencia (a partir de ahora llamados *anclas*) se puedan ubicar, empleando únicamente imágenes proporcionadas por el endoscopio, en el mapa de la cavidad abdominal y en la imagen PET-TAC.

3.1. Sistema de anclas

Una anotación de AR contextual debe tener una posición y orientación fija dentro del mapa de la cavidad abdominal, el cual se calcula mediante las anclas. Las anclas se definen como referencias anatómicas reconocibles por el personal médico, tanto en la secuencia de vídeo del endoscopio como en la imagen PET-TAC. En el caso de estudio, se ha decidido utilizar por recomendación de los especialistas la epífisis de la pelvis y las dos venas ilíacas en el punto que atraviesan el músculo para dirigirse a las piernas.

El algoritmo de SLAM realiza una reconstrucción de la escena mediante emparejamiento de características ORB[7] en cada frame. Estas características tienen la posición 3D en el mapa de la cavidad, por lo que se podrían usar como anclas con buena precisión. Sin embargo, este tipo de puntos los crea el sistema de SLAM siguiendo un algoritmo aleatorizado, por lo que no se tiene control sobre ellos. Esto conlleva que en la práctica nunca haya un punto del mapa en la posición de la referencia anatómica elegida como ancla. Definir un ancla en una posición que no le corresponde hace que baje la precisión con la que se logra la representación de la anotación de AR, ya que su posición depende de las tres anclas que se van a definir.

Para lograr definir el ancla en cualquier posición dentro de la imagen se define una nueva información en el mapa, las anclas, que se representan como pequeñas superficies localmente planas en el mapa, el plano se define con la punto 3D que ocupa y el vector normal que define

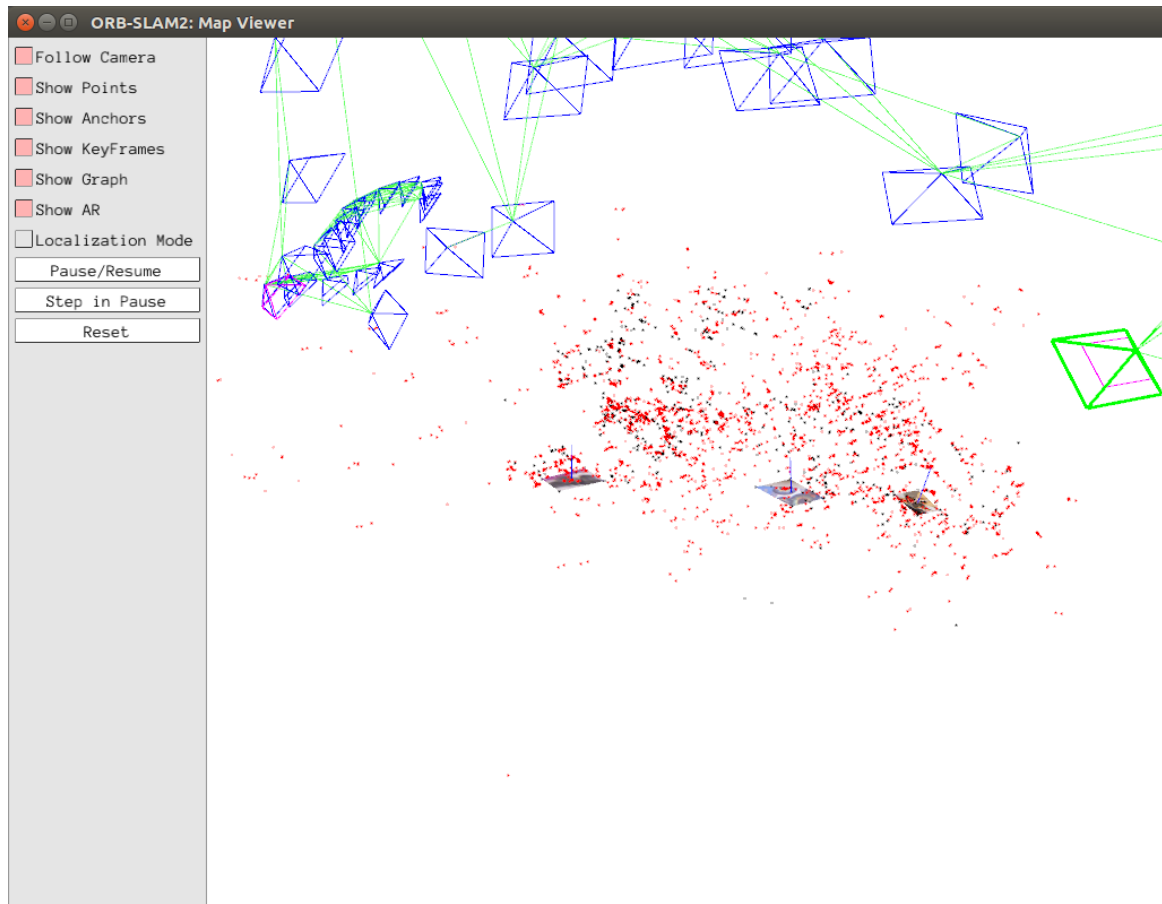


Figura 3.1: Visor 3D con el mapa de entorno con la representación gráfica de las anclas como una imagen y el vector normal que define su plano.

el plano. El punto 3D identifica la posición de la referencia a partir de la cual se calcula donde se debe proyectar del modelo. El vector normal define el plano que contiene el punto 3D, y se emplea para estimar de forma precisa la posición que ocupa el ancla dentro del mapa de la cavidad abdominal.

Cuando el endoscopio este ubicado dentro del mapa de la cavidad, se puede definir un ancla en un punto de la imagen que no este próximo a los bordes. El frame actual donde se va a crear el ancla solo contiene información 2D, por lo que se realiza un procesamiento en varios pasos para determinar si se puede crear el ancla en dicha posición. En caso de que se pueda, se calcula su posición 3D y el vector normal. El ancla se define como un plano ya que se calcula su información 3D a partir de comparaciones de un pequeño parche (de aquí en adelante patch), alrededor de la zona de la imagen donde se crea, con los puntos de vista de otros keyframes. Una vez creada el ancla, el hecho de tratarlo como un plano facilita la búsqueda de correspondencias en otras imágenes mediante correlación normalizada.

3.1.1. Cálculo del plano del ancla

Un ancla se representa como un plano, el cual se compone de un punto 3D contenido en el plano y su vector normal perpendicular al plano. Inicialmente se tiene la información 2D de la imagen donde se ha clicado, sin embargo no es suficiente para establecer su posición 3D. Por ello se realizan varios pasos:

- Estimación una posición 3D inicial con un keyframe próximo mediante una búsqueda por correlación normalizada.

- Se define un vector normal del plano.
- Ajuste de la posición 3D del punto con un keyframe con mayor paralaje.
- Ajuste del vector normal del plano con un keyframe lejano.

La estimación inicial de la posición 3D es una semilla inicial que posteriormente va a ser refinada. Se conoce la posición 2D en el frame, que define un vector con distancia desconocida donde se encuentra el punto en 3D. Se buscan los puntos triangulados en el mapa que estén próximos a la posición 2D del ancla en el frame. De este conjunto de puntos del mapa se guarda el valor máximo y mínimo la coordenada 3D que da información sobre la profundidad del punto respecto al endoscopio. Con la estimación de profundidad se acota el vector donde se encuentra la proyección del punto 2D del frame. Se busca un keyframe pasado que se encuentre cerca de la posición del frame actual, con el objetivo de que la apariencia en la imagen del patch sea similar y no se necesite aplicar ninguna transformación en el patch para el cálculo de correspondencias. Sobre el frame cercano se proyecta el vector acotado donde se estima que se encuentra la posición del ancla (figura 3.2b), se establece una ventana de búsqueda alrededor de esta línea y en la región acotada se busca el patch con correlación normalizada. El punto de la ventana de búsqueda con mayor valor de correlación identifica la coordenada 2D del keyframe de mayor coincidencia con el patch del frame. Con la posición 2D en el frame y keyframe de los cuales se conoce su posición dentro del mapa de la cavidad (figura 3.2), se triangula la posición 3D del punto que identifica la posición inicial del ancla.

El ancla tiene una posición 3D inicial que permite calcular una normal inicial, que va a ser el vector normalizado que va del punto 3D del ancla a la posición del centro óptico del frame. Se define de esta forma ya que, al crear un ancla se encara el endoscopio a la región del ancla para obtener un mejor punto vista. En la figura 3.2d se puede ver la estimación en el mapa de la cavidad abdominal, la cual es suficiente para terminar de ajustar el ancla, ya que permite compensar aproximadamente la ciclorsión y el cambio de escala.

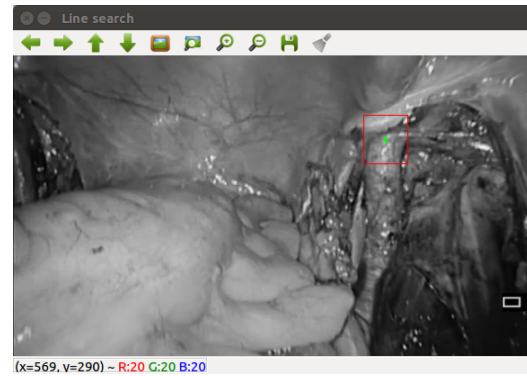
El ajuste definitivo de la posición 3D del ancla se realiza con un keyframe que tenga un mayor paralaje, para estimar la profundidad con mayor precisión. Se busca un keyframe que tenga un punto de vista que se encuentre desplazado entre 5 y 15 grados respecto al punto de vista del frame (figura 3.3b), y que la distancia del keyframe al punto 3D no supere en un 20 % la distancia entre el punto 3D y frame. Estas restricciones hacen que el patch del ancla no tenga un cambio de perspectiva o escala lo suficientemente grande para que se necesite aplicar una transformación al patch, y al mismo tiempo es suficiente para estimar la posición 3D con mayor precisión. Se predice la posición 2D del punto en la imagen del keyframe (figura 3.3a) y se establece una ventana de búsqueda alrededor, en la que se realiza una búsqueda por correlación normalizada con el patch. Se guarda la coordenada 2D con el valor de correlación más alto y se triangula la posición del punto en 3D con las coordenadas 2D en la imagen de los dos keyframes y el frame y sus posiciones en el mapa de la cavidad.

Por último se ajusta la normal, se selecciona un keyframe alejado que pueda proyectar en su imagen la posición de la posición 3D del ancla, para ello se comprueba si el punto esta dentro de la imagen y si la normal de l ancla permite verlo desde el punto de vista de la cámara. se comprueba el ángulo entre el punto de vista y la normal (figura 3.5). Se considera que un keyframe puede ver el plano si esta dentro del área de confianza (60 grados respecto a la normal).

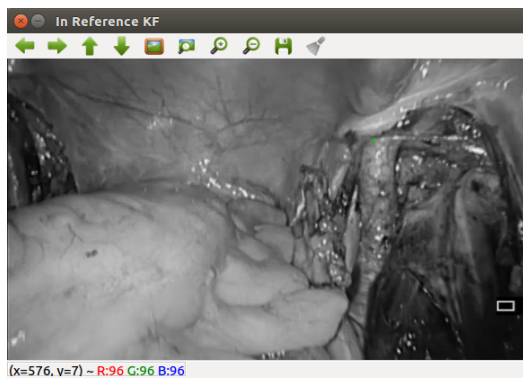
Con el keyframe lejano ya seleccionado, se definen un conjunto de vectores normalizados alrededor del vector normal del ancla, con cada uno de ellos se realiza una búsqueda por correlación en el keyframe. Sin embargo, la diferencia de escala y perspectiva entre el keyframe y el frame donde se ha definido el patch es grande, así que es necesario aplicar una transformación homográfica, la cual define una transformación perspectivas a un plano. La matriz de homografía



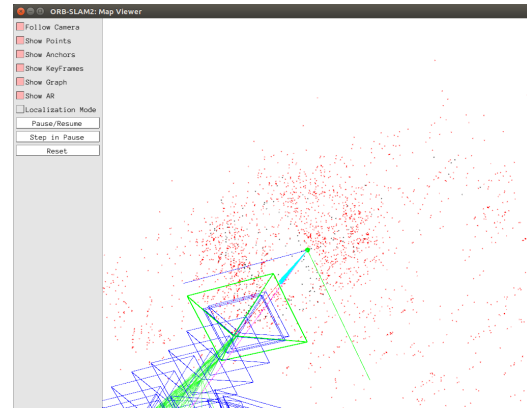
(a) Definición.



(b) Definición.

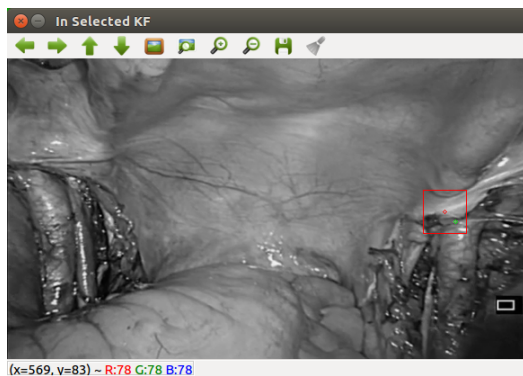


(c) Correlación.

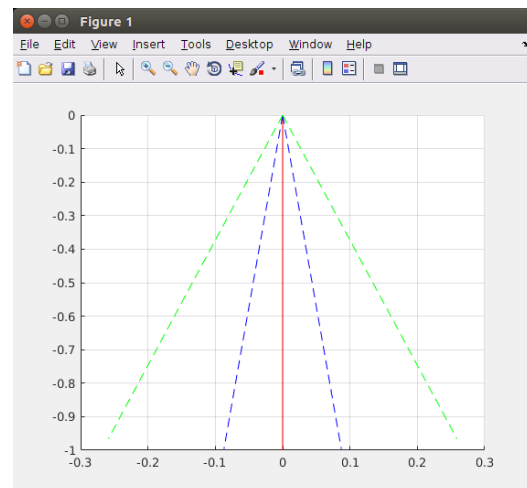


(d) Estimación inicial de posición.

Figura 3.2: Imagen donde se define el ancla(a), el círculo verde el punto definido. Imagen donde se predice el emparejamiento (d), la línea verde es donde se estima, se traza una ventana de búsqueda alrededor. Imagen donde se busca el patch (c), el círculo verde es la punto donde se ha emparejado. Estimación inicial (d) del punto (punto verde) y normal del ancla (vector azul claro del punto) en 3D.



(a) Imagen con ventana de búsqueda (cuadro rojo) y posición con mayor valor en el emparejamiento (punto verde).



(b) Región de punto de vista de entre 5 y 15 grados (área entre líneas discontinuas verde y azul) respecto al punto de vista del frame (línea roja).

Figura 3.3: Imagen a escala similar al patch (a). Paralaje entre la región de la imagen (b)

se construye a partir de información que se conoce. Las ecuaciones (3.1) y (3.2) definen como se construyen. El frame donde se crea el ancla, y por tanto el patch, se define como f mientras que el keyframe donde se realiza la búsqueda por correlación normalizada es la kf . K define los parámetros intrínsecos de cada cámara, R la rotación en la orientación del primer subíndice respecto al segundo, t es la traslación de la posición del primer subíndice respecto al segundo, n es el vector normal del plano del patch desde la referencia del subíndice y D la distancia al plano (3.3).

$$\mathbf{H}_{kf,f} = \mathbf{K}_{kf} * (\mathbf{R}_{kf,f} - \mathbf{t}_{kf,f} * \mathbf{n}_f^T * D) * \mathbf{K}_f^{-1} \quad (3.1)$$

$$\mathbf{H}_{f,kf} = \mathbf{H}_{kf,f}^{-1} \quad (3.2)$$

Para realizar la búsqueda por correlación del patch se debe aplicar la homografía a la imagen del frame para transformar el plano del frame y obtener el plano que se predice que se visualiza desde el keyframe, pero las imágenes pueden tener distorsión, la cual depende su efecto de las coordenadas de la imagen, por ello es necesario eliminar la distorsión de la imagen del frame, luego se aplica la transformación homográfica y por último se aplica la distorsión correspondiente al keyframe, con la imagen resultado se puede realizar la búsqueda por correlación en la imagen del keyframe correspondiente. Realizar esta operación es costosa para solo necesitar comparar un pequeño patch, por ello se opera solo con el tamaño del patch, al cual se le aplican las tres operaciones de una sola vez, debido a que se trabaja con las coordenadas de la imagen y no con la propia imagen. Para cada coordenada del patch de destino se calcula las coordenadas de donde se obtiene la información de color de la imagen del frame (figura 3.6), de esta manera la operación es más eficiente al realizar una sola transformación conjunta en vez de tres transformaciones independientes. Una vez sintetizado el patch que se predice en el keyframe, se realiza una búsqueda por correlación normalizada en la ventana de búsqueda del keyframe y se guarda el valor máximo que se obtiene con la correlación y la normal con la que se ha calculado la transformación de homografía.

Se esta trabajando con un conjunto de vectores normalizados, por lo que el paso anterior se repite con cada normal candidata y se compara el valor máximo de correlación que se obtiene con cada una, para guardar el vector normalizado que devuelve el valor de normalización más alto. El vector con mayor valor de normalización se convierte en la nueva normal del plano.

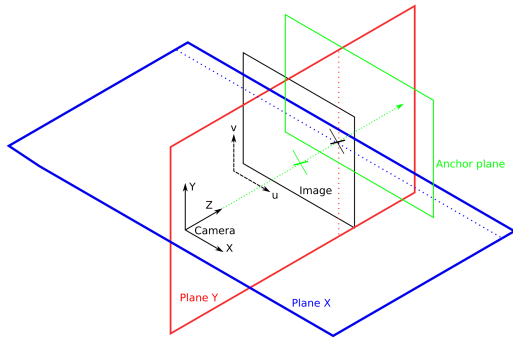
3.1.2. Representación en el mapa del ancla

El ancla se va a representar en el visor del mapa de la cavidad abdominal por una miniatura de tamaño fijo de 64x64 píxeles, que se define en la imagen donde se crea el ancla. Para estos puntos se va a calcular su posición con la normal del plano que los contiene. Estos puntos no se utilizan para ajustar la anotación de AR, solo se utilizan para representar el ancla en el visor de forma precisa.

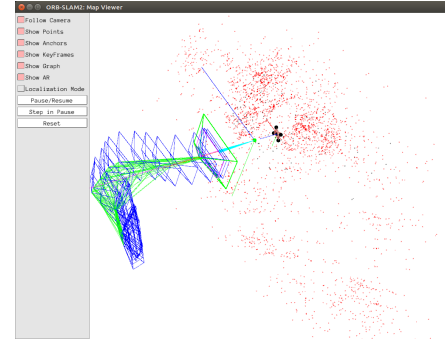
El patch del ancla se ha definido como un plano, por lo que los puntos que identifican las esquinas del patch están dentro del propio plano. Por ello se utiliza la ecuación del plano para calcular su posición 3D en el mapa, la ecuación 3.3 permite encontrar calcular un punto (x,y,z) que pertenece al plano. Para ello se necesita conocer los valores que definen el plano (A,B,C,D) , los cuales son conocidos, tres componentes definen la normal del plano (A,B,C) y el cuarto la distancia del plano respecto a la normal (D) .

$$A * x + B * y + C * z + D = 0 \quad (3.3)$$

El cálculo de la posición de las esquinas se realiza mediante la ecuación de corte de 3 planos en un espacio 3D (figura 3.4a). Dos planos se conocen a partir de la proyección del punto en la



(a) Definición de punto.



(b) Proyección del ancla en el mapa.

Figura 3.4: Corte de 3 planos no paralelos para definir el punto que pertenece a ellos (a). Posición del ancla (b) ajustada (plano con puntos negros en las esquinas) frente a la posición estimada inicial (punto verde) .

imagen de la cámara donde se crea el ancla, cada coordenada (X, Y) proporciona una ecuación del plano, el tercer plano lo define la normal del ancla donde está contenido el punto, de tal forma que el corte de los 3 planos da la posición 3D del punto. Para calcular la posición de forma precisa se utiliza la descomposición de matrices con las 3 ecuaciones del plano definida en una matriz (3.4), este cálculo minimiza el error cuadrático en la posición del punto que es contenido en los 3 planos. El resultado del SVD nos devuelve 3 matrices (3.5), la solución que identifica el punto de corte de los 3 planos se encuentra en la última columna de la matriz \mathbf{V} .

$$\mathbf{M} = \begin{bmatrix} A_x & B_x & C_x & D_x \\ A_y & B_y & C_y & D_y \\ A_a & B_a & C_a & D_a \end{bmatrix} \quad (3.4)$$

$$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = SVD(\mathbf{M}) \quad (3.5)$$

Este método se aplica a las cuatro esquinas del patch en la imagen. Se obtienen las coordenadas 3D que permiten la representación del ancla en la imagen mediante su miniatura (figura 3.4b).

3.1.3. Refinamiento de la posición del ancla con todos los keyframes del mapa

Cuando se tiene la información de posición y normal que definen el ancla, se crea e introduce en el sistema de refinamiento de posición. Para ello se identifican todas los keyframes anteriores en las que se puede proyectar el punto dentro de su imagen y que su punto de vista respecto a la normal del plano no tenga un ángulo mayor de 60 grados (figura 3.5). Las coordenadas 2D del punto en la imagen de cada keyframe se guarda para triangular la posición del punto en 3D.

Se va a triangular la posición 3D donde se encuentra el ancla con todos los keyframes del mapa, para ello se calcula para cada keyframe la posición 2D en la imagen en la que se proyecta. La búsqueda del patch del ancla en la imagen del keyframe se realiza mediante correlación normalizada, como el punto de vista del plano en cada posición es muy diferente se calcula la nueva apariencia mediante la transformación de homografía para cada imagen, utilizando la ecuación 3.1. La transformación siempre parte de la posición del frame que ha creado el ancla, y se transforma el patch desde ese punto de vista al punto de vista del keyframe que se está procesando.

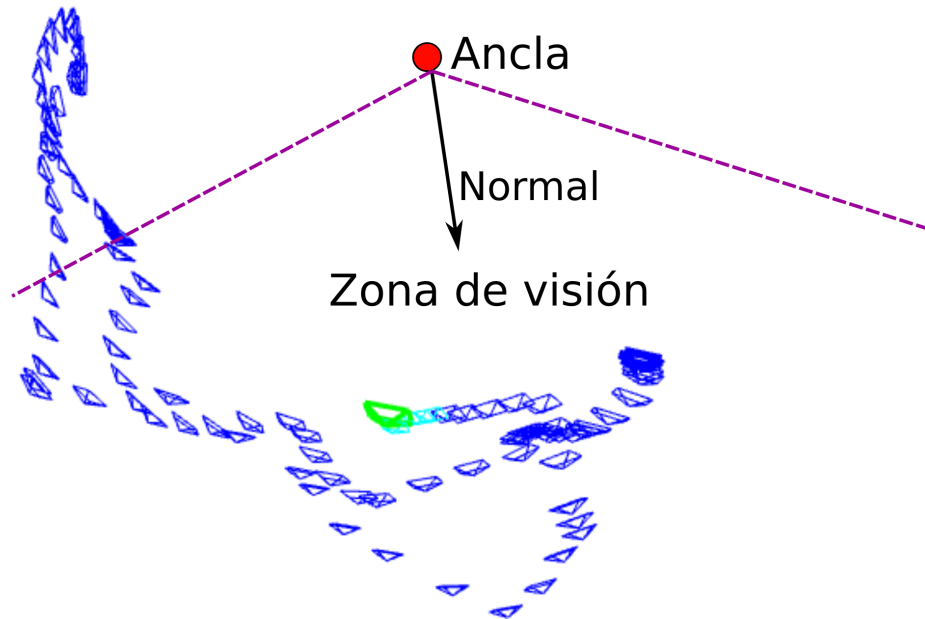


Figura 3.5: Zona de admisión de puntos de vista en función de la normal del plano que define del ancla, solo las vista que quedan dentro del cono definido por la zona de visión son admitidos, siempre que el punto se pueda proyectar en la escena.

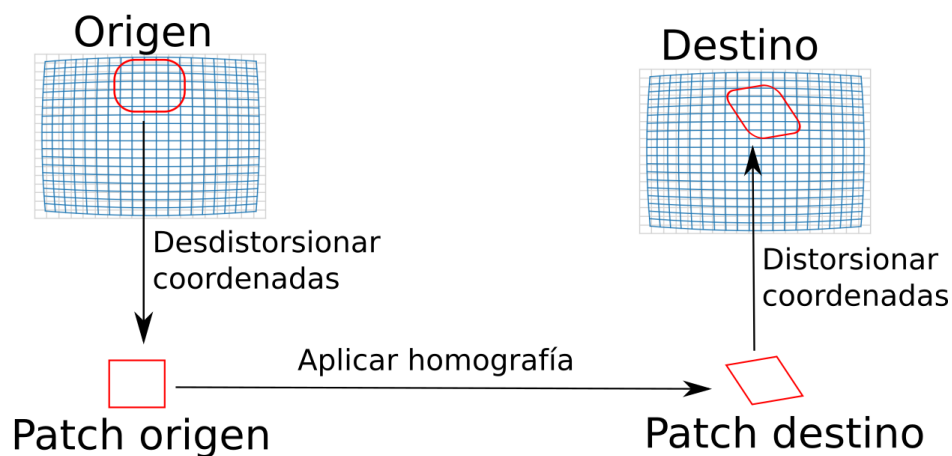


Figura 3.6: Cálculo de la vista del patch en el resto de imágenes a partir de la inicial, se debe corregir la distorsión de la cámara para realizar la búsqueda por correlación.

Para cada posición de keyframe candidato se calcula el punto de mayor similitud entre el patch transformado y la ventana de búsqueda. La transformación del patch requiere corregir la distorsión del patch, después aplicar la transformación homográfica y aplicar la distorsión de la cámara del keyframe candidato, para después realizar la búsqueda por correlación. Tal y como se ha explicado en la sección 3.1.1, aplicar tres transformaciones a la imagen completa y para cada keyframe candidato no es eficiente, por lo que se trabaja con las coordenadas del patch destino. Para ello se realizan las transformaciones sobre las coordenadas del patch destino (figura 3.6), de manera que al procesar el patch del keyframe candidato solo se ha calculado la coordenada de donde debe coger la información de los píxeles en el patch de origen.

La ventana de búsqueda en la que se aplica la correlación normalizada con el patch transformado se define a partir de la posición 3D del ancla. Si el keyframe es candidato es porque ve el punto y se ha calculado la posición 2D en la que se proyecta dentro de su imagen, se establece una ventana de búsqueda de tamaño fijo alrededor de la posición 2D donde se proyecta en la imagen. La búsqueda por correlación solo se realiza en esta ventana, la cual tiene un mayor tamaño que el patch.

Al terminar la búsqueda por correlación se obtienen las coordenadas 2D en la imagen donde se ha obtenido el valor más grande durante el emparejamiento entre el patch transformado y la ventana de búsqueda en el keyframe. Estas coordenadas se definen en un espacio discreto, que puede tener error hasta de 1 píxel. Para realizar un ajuste más preciso se ha decidido reducir el error por debajo del píxel, para ello se emplea el algoritmo de *Enhanced Correlation Coefficient*[8], que a partir de una semilla inicial próxima a la solución óptima realiza el emparejamiento entre el patch y la ventana de búsqueda a nivel de sub-píxel.

Con la posición del ancla calculada en todas los keyframes candidatos se realiza una triangulación de la posición del ancla con todas las vistas, para ello se asume que las posiciones de las cámaras y las coordenadas 2D en cada imagen son fijos para que solo se ajuste la posición 3D del ancla minimizando el error cuadrático.

Cuando se ha terminado este proceso el ancla no se deja fija hasta la finalización del programa, al insertar nuevos keyframes al mapa se comprueba si ven el ancla para agregarlos al refinamiento de la posición

3.2. Inserción de anotaciones de AR

Las anotaciones de AR son gráficos 3D que deben proyectar información en la vista de la cámara de forma contextualizada, la información no es visible en el mundo real por lo que se enriquece la visión con información adicional y útil. La anotación puede ser de diferentes tipos por lo que se ha trabajado en implementar una solución que pueda utilizar en distintos problemas de AR.

Para generar información gráfica 3D que se utiliza para estas anotaciones se emplea un estándar ampliamente usado en la informática gráfica OpenGL. Esta librería permite la representación del mundo 3D en un ordenador, de forma que se pueda visualizar en la pantalla y navegar por él. Tiene una representación de estructuras 3D basada en objetos mediante triángulos, lo que permite representar cualquier forma que se represente con una malla compuesta por triángulos conectados.

Un objeto 3D en OpenGL se define como un conjunto de vértices y caras, los vértices ocupan una posición en el espacio 3D, mientras que las caras identifican los triángulos uniendo los vértices, de esta forma se forma una malla de triángulos, los cuales se almacenan en un fichero mediante un etiquetado estandarizado. El fichero permite modelar cualquier objeto que se quiera representar en el entorno de visualización. En el capítulo 5 se explica en mayor detalle como es generada la anotación.

El visualizador de OpenGL renderiza una imagen en 3 dimensiones para obtener su representación en un plano a partir de un punto de vista que se puede mover por la escena. El punto de vista identifica el plano sobre el que se proyecta la escena, para definir que se proyecta se proyecta un rayo a través de cada píxel, de forma que recógela información de color en la zona donde impacta. El sistema de SLAM crea un mapa de cavidad con la posición del endoscopio en cada frame, esta información de posición se usa para proyectar los objetos 3D de la anotación de AR en el frame actual, de forma que la anotación se queda contextualizada con respecto a la posición del frame.

Capítulo 4

Alineamiento de la notación AR en el sistema SLAM

Se ha definido las anclas tanto en el modelo 3D de la anotación de AR como en el mapa de la cavidad abdominal construido a partir de la laparoscopia, de manera que se puede calcular la posición de la anotación de AR respecto al mapa de la cavidad. Para realizar el alineamiento se calculan los ejes de referencia a partir de las anclas, los cuales se definen con el algoritmo de Horn [9]. Los ejes permiten que se realice la transformación de similitud ($Sim(3)$) que alinea las anclas del modelo 3D con las del mapa, y con ello la anotación de AR.

El algoritmo de Horn se basa en calcular un sistema de referencia local dentro de un sistema de referencia global a partir de tres puntos contenidos en la referencia global, de manera que se obtiene la transformación de similitud entre los dos sistemas, con el que transformar los objetos contenidos en un sistema al otro.

Se establecen los ejes de referencia del sistema local a partir de los tres puntos, el eje \hat{x} se establece a partir de las anclas 1 y 2:

$$\mathbf{x} = a_2 - a_1 \quad (4.1)$$

$$\hat{x} = \mathbf{x} / \|\mathbf{x}\| \quad (4.2)$$

Después se calcula el eje \hat{y} a partir del anterior y las anclas 1 y 3:

$$\mathbf{y} = (a_3 - a_1) - [(a_3 - a_1) \cdot \hat{x}] \hat{x} \quad (4.3)$$

$$\hat{y} = \mathbf{y} / \|\mathbf{y}\| \quad (4.4)$$

Con los ejes \hat{x} e \hat{y} calculados se puede sacar el \hat{z} como producto entre ellos

$$\hat{z} = \hat{x} \times \hat{y} \quad (4.5)$$

De esta forma se obtienen los tres ejes de referencia del sistema local dentro del global, con ellos se puede construir la rotación entre modelos, no obstante falta obtener la traslación respecto al sistema de referencia global. Para calcularlo se define el ancla a_1 como el punto de origen del sistema local, la traslación se realiza a partir del punto de origen del sistema global (o_g):

$$\hat{t} = a_1 - o_g \quad (4.6)$$

El origen del sistema de referencia global toma la posición (0, 0, 0) cuando se observa desde el propio sistema. Con los ejes y la traslación se construye la matriz de transformación entre sistemas de referencias:

$$\mathbf{M}_{g,l} = \begin{bmatrix} \hat{x} & \hat{y} & \hat{z} & \hat{t} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.7)$$

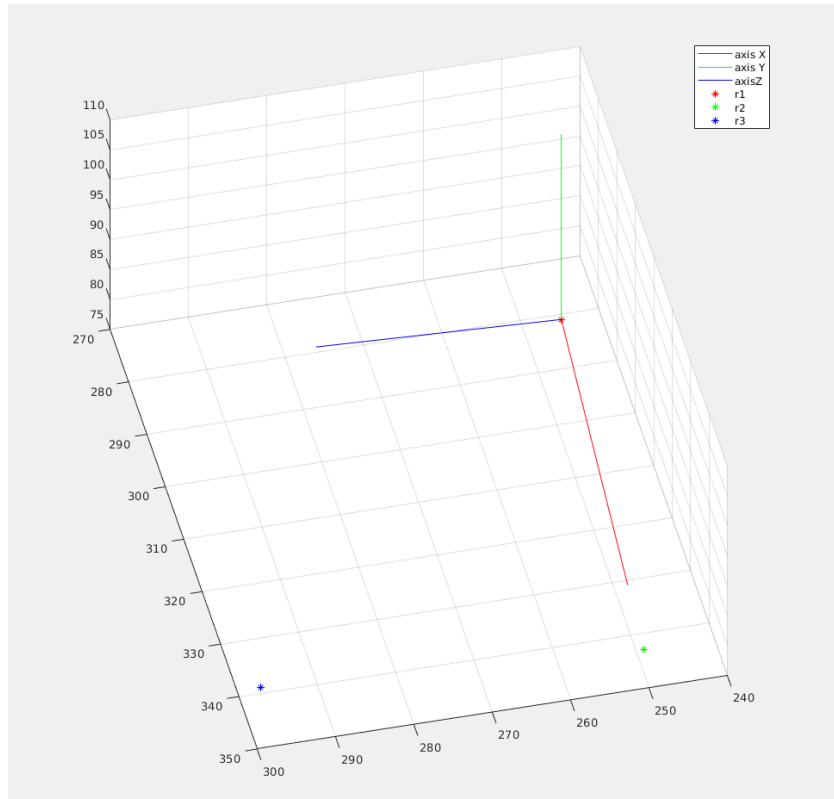


Figura 4.1: Sistema de referencia local calculado a partir de los tres puntos de referencia.

$$\mathbf{M}_{l,g} = \mathbf{M}_{g,l} - \mathbf{1} \quad (4.8)$$

La matriz de transformación permite transformar las coordenadas de la anotación 3D entre diferentes sistemas de referencia. Siempre se va a utilizar tres puntos para calcularlo, el único requisito es que dichos puntos sean los mismos cuando se definen como anclas en el TAC y en el mapa de la cavidad abdominal.

El mapa de la cavidad obtenido a partir de la secuencia de laparoscopia no tiene escala, sin embargo la anotación se ha generado a partir de una imagen PET-TAC que contiene información de escala. Una vez alineada la anotación de AR al mapa de la cavidad, se aplica una transformación al mapa para representarlo con su verdadera escala. Se emplea la distancia conocida entre las anclas en la imagen tomográfica del PET-TAC y la distancia de las mismas anclas en el mapa de la cavidad, se calcula el factor de escala que se debe aplicar para dejar el mapa con sus verdadera escala. El factor de escala se obtiene como el promedio de las 3 factores de escala calculados. El factor de escala se aplica a todo el mapa, de esta forma se consigue un mapa con la escala real construido con un sensor monocular.

El alineamiento se recalcula después de cada optimización de las anclas en el mapa.

Capítulo 5

Creación de la anotación AR

La anotación de AR se crea a partir de la imagen PET-TAC que se realiza al paciente antes de la operación para emitir el diagnóstico, el especialista debe generar la anotación utilizando la aplicación desarrollada en la cual puede segmentar la zona de interés. Esta tarea es complicada de realizar por una persona no experta por la dificultad que tiene distinguir las estructuras anatómicas en la imagen PET-TAC.

Se ha creado una herramienta interactiva para trabajar con imágenes PET-TAC, que permite realizar la generación de la anotación de AR de manera interactiva (figura 5.1). El visor permite navegar por las diferentes capas de cada vista de forma independiente. Las vistas se muestran de forma simultanea en la herramienta, estas tres vistas corresponden a las vistas normalizadas en el ámbito médico (*Axial, Coronal y Sagittal*).

La anotación de AR va a identificar el tumor que se quiere extirpar, debido a la forma que tiene se ha decidido modelarlo como un elipsoide, ya que es la forma básica a la que mejor se ajusta. El elipsoide se define en el visor estableciendo su centro, a partir del cual se forman 3 ejes (figura 5.1b), cada uno paralelo a los ejes de la vista. Cada eje identifica uno de los 3 radios del elipsoide. La anotación se define en el visor como un conjunto de voxel de tamaño conocido.

En el visor el elipsoide se define como un conjunto de voxel, para para crear la anotación AR se debe transformar en un objeto OpenGL, el cual define la superficie del objeto como un conjunto de triángulos.

La posición del elipsoide se realiza usando anclas, las cuales se establecen en estructuras anatómicas visibles en la imagen PET-TAC y el vídeo vivo del endoscopio. Las estructuras anatómicas que se emplean son las dos venas iliacas cuando atraviesan el musculo hacia las piernas y la epífisis de la pelvis. Los puntos se definen en el mismo visor, por lo que tienen una posición conocida. Los tres puntos van a identificar el eje de referencia sobre el que se va a definir el modelo 3D del tumor, de forma que al emparejar los puntos en el vídeo vivo del endoscopio se calcula la posición y orientación que ocupa en el mapa.

Cuando el modelo 3D de OpenGL ha sido generado se visualiza su aspecto respecto a las anclas que definen su eje de referencia (figura 5.2), el modelo se encuentra ubicado respecto a este eje de referencia. Posteriormente, el modelo se guarda en un fichero de formato OpenGL para ser representado en el visualizador del algoritmo de SLAM.

La imagen PET-TAC contiene información sobre el tamaño de cada voxel, lo que proporciona información sobre la escala del tumor y la distancia entre las anclas. Esta información también se guarda en el fichero de OpenGL.

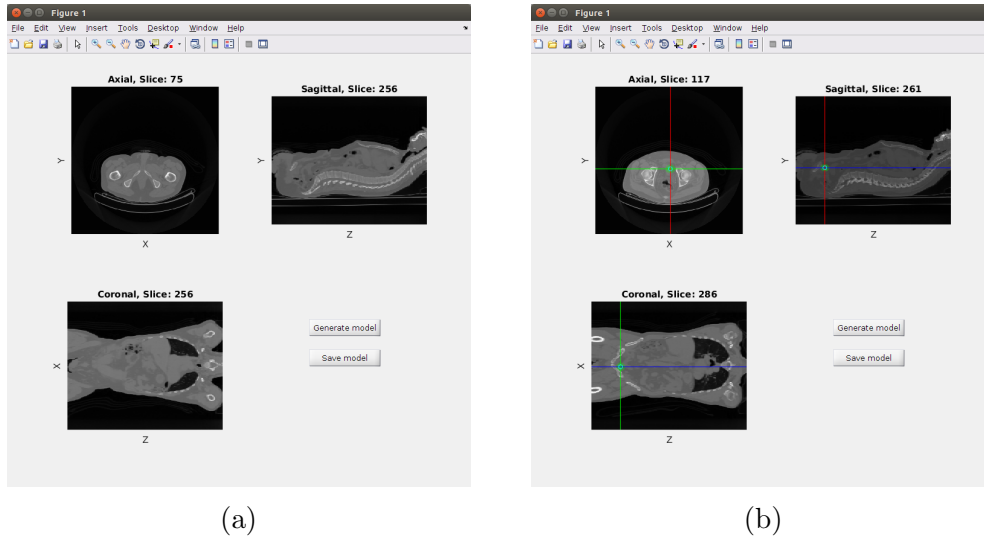


Figura 5.1: Visor con las tres vistas (a). Visor con el modelo definido en las tres vistas (b), las líneas de color identifican cada eje de referencia que cortan en el centro del modelo.

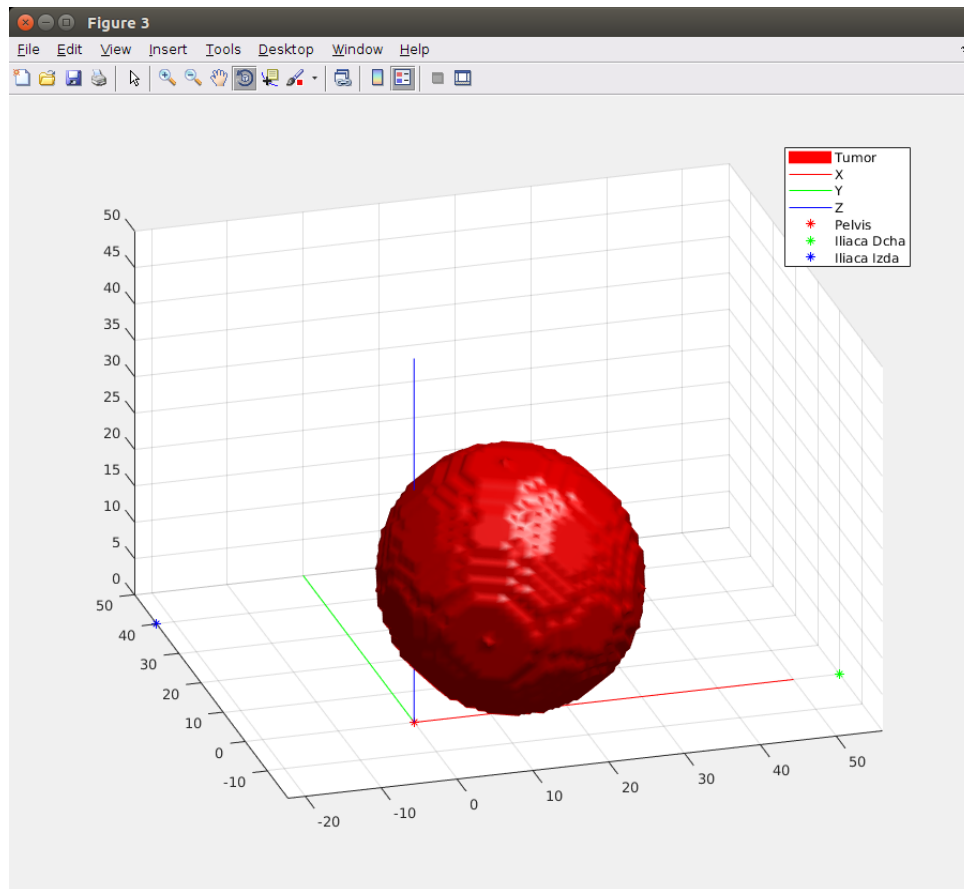


Figura 5.2: Modelo 3D generado a partir de la segmentación, se define a partir del eje de referencia (lineas) creado a partir de las anclas (asteriscos).

Capítulo 6

Adquisición de datos de laparoscopia

Los datos de los que se dispone es la imagen PET-TAC del paciente, la cual se obtiene antes de la operación, y el vídeo de la laparoscopia y otro con una secuencia de calibración.

6.1. Adquisición de la secuencia

El vídeo de la cirugía laparoscopia se realiza con un endoscopio monocular que captura la escena en alta definición (resolución de 1920x1080) con una tasa de refresco de 25Hz. La imagen se envía a un monitor para visualizar el interior del paciente. La secuencia de la cirugía tiene una duración cercana a las 2 horas. Se ha necesitado emplear una capturadora de vídeo y un disco duro para grabar la secuencia completa.

Para obtener una buena secuencia se deben seguir unas recomendaciones. Para construir un mapa de entorno denso a partir de la secuencia se necesita realizar barridos de toda la cavidad a cierta distancia, de esta forma se inicia un mapa con suficiente paralaje y se densifica el mapa para que el endoscopio se pueda localizar por toda la cavidad. Durante el crecimiento del mapa se evitan los movimientos bruscos que pueden hacer que se pierda la localización, cuando el mapa se ha densificado esto no supone un problema ya que tiene muchas zonas en las que relocalizarse. Es necesario visualizar las estructuras anatómicas que definen las anclas en detalle, para tener gran variedad de puntos de vista que ayuden con el ajuste de la posición. Por último, se necesita visualizar las estructuras anatómicas que definen las anclas en una posición cercana y con poco movimiento, para evitar el desenfoque cuando se define el patch del ancla.

Los requisitos para grabar una buena escena no han variado la ejecución de la cirugía, se ha mantenido la metodología que emplea el personal sanitario sin interferir con su trabajo.

Al finalizar la operación se graba una secuencia extra con el laparoscopio para visualizar el patrón de calibración desde diferentes puntos de vista para obtener la calibración. Se ha realizado al final de la operación, debido a que previamente las herramientas se encuentran esterilizadas y no está permitido que el personal ajeno interactúe con ellos.

6.2. Calibración del laparoscopio

Para la calibración del endoscopio se ha empleado un patrón conocido y usado en la librería de software libre OpenCV [10] que se puede ver en la figura 6.1a.

La secuencia de calibración necesita ver el patrón conocido desde diferentes puntos de vista, para calcular la calibración de la cámara de forma precisa se extraigan los parámetros de calibración del sensor.

La calibración de una cámara es necesario para obtener una reconstrucción 3D precisa.

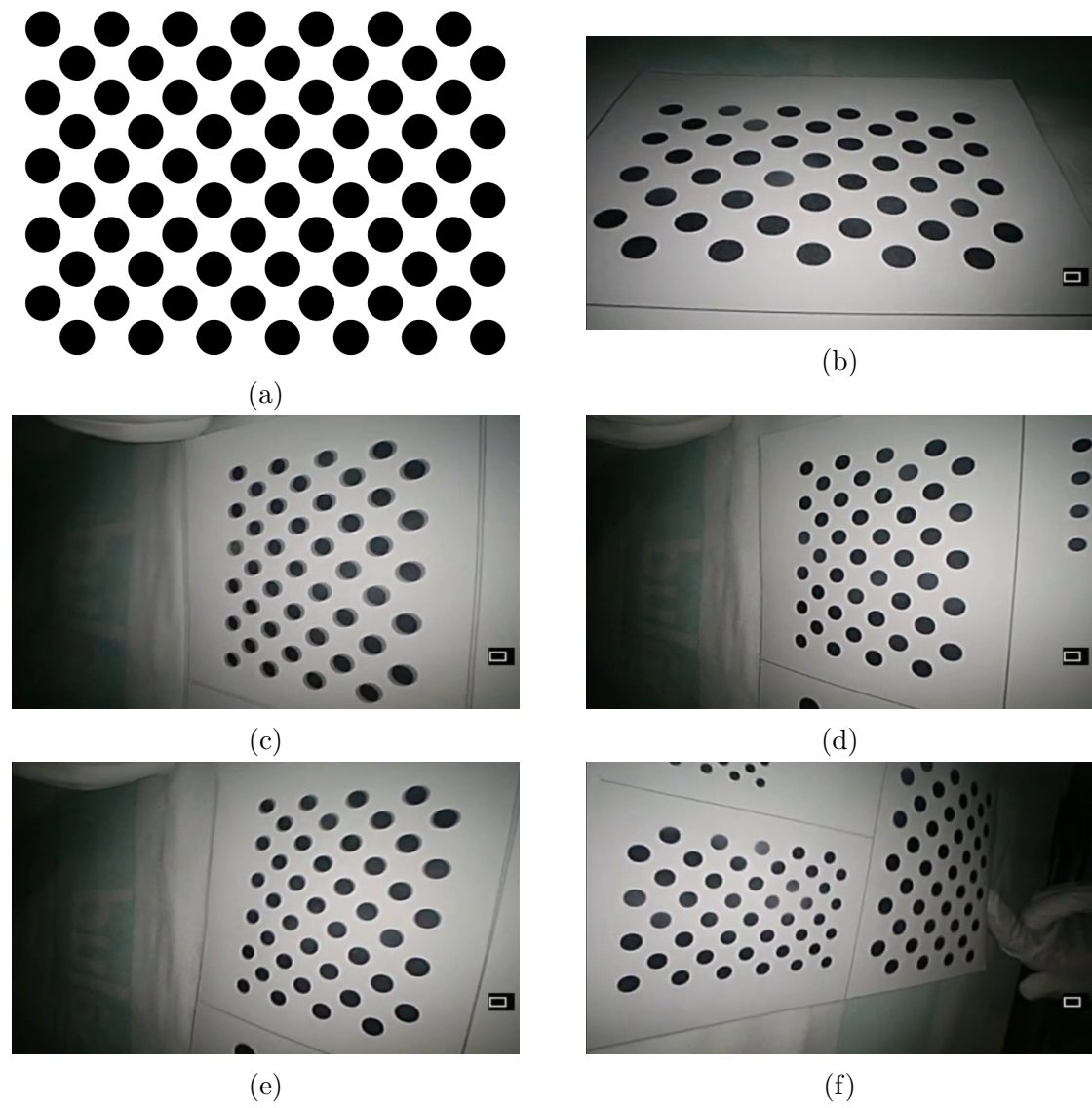


Figura 6.1: Patrón de calibración (a). Diferentes vistas del patrón(b-e).

Capítulo 7

Experimentos

El experimento se ha realizado con los datos obtenidos del paciente. Primero se ha generado la anotación de AR referenciada a partir de las anclas. Ante la dificultad de identificar el tumor para un ojo inexperto se ha colocado una anotación de unos 2cm entre las tres anclas, los 3 radios son diferentes para que no sea perfectamente esférico. Las estructuras anatómicas que identifican las anclas no se han modificado.

La secuencia del endoscopio capturada es muy extensa, por lo que se ha extraído una secuencia en la que se realiza una exploración de la cavidad abdominal con la que iniciar el mapa de la cavidad abdominal. Después se realiza una aproximación a las referencias anatómicas que definen las anclas para tener diferentes puntos de vista con los que ajustar la posición 3D del ancla.

Se procesa el vídeo para construir el mapa de la cavidad. Cuando se ha inicializado se realizan barridos para captar diferentes keyframes y puntos que lo vuelvan más robusto, ya que tiene mas keyframes con los que generar los ajustes. Se definen las anclas de las tres referencias anatómicas, al obtener la posición de las anclas en el mapa se representan con el plano de su miniatura dentro del mapa (figura 7.1b). En el visor de anclas (figura 7.1a) se visualiza la miniatura que identifica el ancla del mapa de la cavidad en la parte superior, mientras que en la parte inferior se encuentran las anclas anatómicas de la anotación de AR. Para insertar el modelo se emparejan las anclas, el color del recuadro de alrededor del ancla permite saber como se encuentran emparejadas. La anotación de AR solo se puede agregar cuando las tres anclas se encuentran emparejadas, el usuario debe insertarlo manualmente, de esta manera se tiene la opción de eliminar un ancla para crear otra.

Cuando se inserta el modelo se realiza la transformación de similaridad ($Sim(3)$) que nos transforma las coordenadas de los vértices de la anotación de AR a las coordenadas del mapa de entorno 3D. La anotación se puede ver tanto en el mapa de la cavidad (figura 7.1c) como en la vista del endoscopio (figura 7.1d).

La posición dentro del mapa 3D se corresponde con la zona donde se ha definido respecto a las anclas, al mismo tiempo mientras la cámara se mueve por la cavidad abdominal se comprueba que la anotación cambia en función del punto de vista.

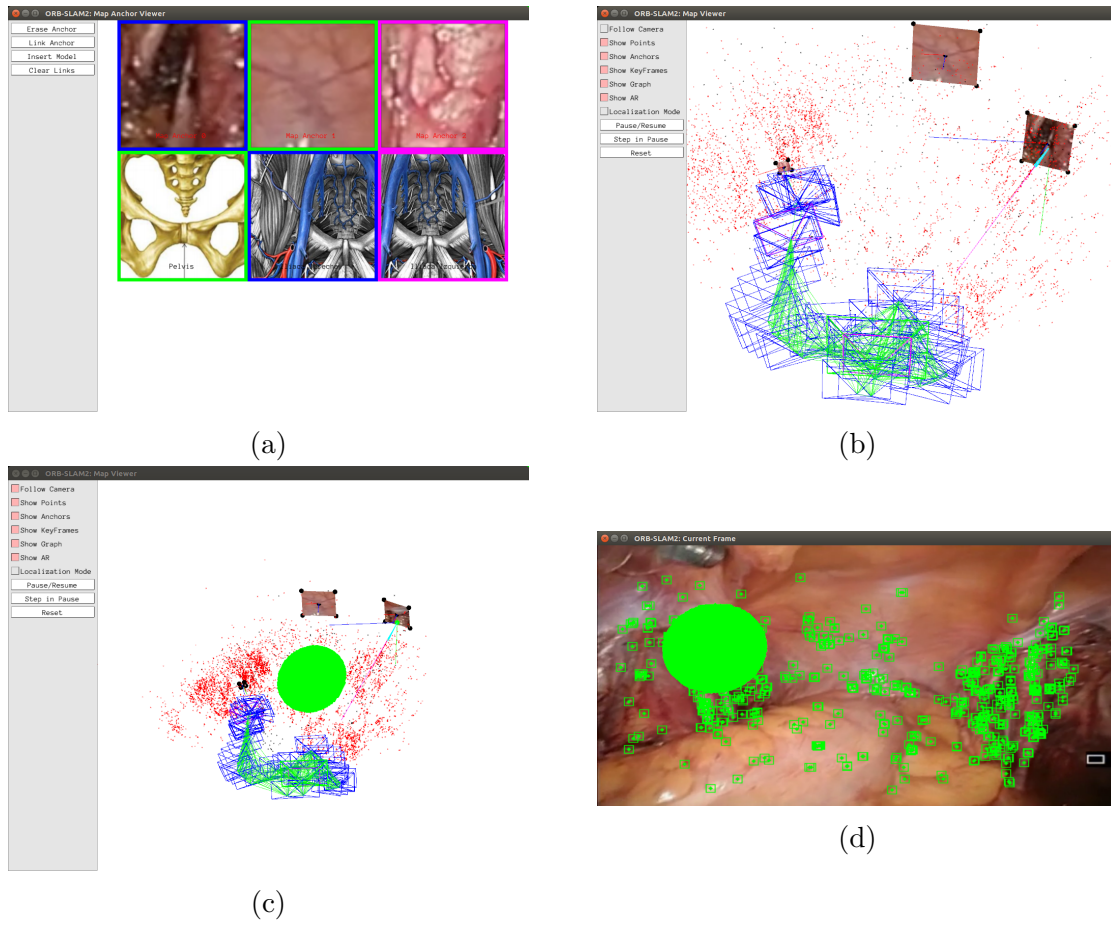


Figura 7.1: Visor de anclas con su emparejamiento(a). Mapa de la cavidad con las anclas definidas (b). Mapa de la cavidad abdominal con las anclas definidas y la anotación de AR (c). Vista del endoscopio con la proyección de la anotación de AR (d).

Capítulo 8

Conclusiones

Se ha completado el desarrollo del sistema de posicionamiento de anotaciones de AR sin necesidad de integrar nuevos protocolos o herramientas adicionales durante el desarrollo de la intervención. Se ha demostrado que en una escena con zonas rígidas se puede establecer un sistema de alineamiento sin modificar el protocolo que se sigue durante la intervención.

A pesar de que el cuerpo del paciente tiene movimientos involuntarios como la respiración, el algoritmo desarrollado se puede emplear en zonas que tengan referencias anatómicas principalmente rígidas.

Se han definido puntos en la escena independientes de las características del mapa, de forma que se puede definir un ancla en cualquier parte del mapa, la única condición necesaria es que se haya explorado la zona previamente y que la región que define el ancla tenga textura para realizar el ajuste de la posición.

La anotación de AR se puede ubicar en la posición que se necesite, la única condición es que debe estar referenciado desde tres puntos para establecer su posición y orientación.

Se han agregado características al algoritmo de SLAM del que se ha partido, al que se le han agregado la posibilidad de importar cualquier tipo de anotación definida como un objeto 3D modelado en OpenGL.

El mapa adquiere la escala real al incluir la anotación de AR con escala, esto crea la posibilidad de agregar mejores herramientas dentro del sistema SLAM, tales como herramientas de medida en la vista del endoscopio, o una distancia a la que se encuentra la anotación respecto del endoscopio en el caso, esto último puede ser útil con las anotaciones que identifican un tumor que están por detrás de otras estructuras.

Cuando el algoritmo de SLAM realiza correcciones en el mapa se aplican también una corrección a las anclas, mediante las búsquedas por correlación normalizada, de manera que la posición de las anclas se ajusta durante todo el proceso, logrando que la posición de la anotación AR sea más precisa.

Capítulo 9

Anexo

9.1. Confidencialidad de información del paciente

Se hace uso secundario de datos, esto es, no se hace ningún procedimiento adicional ni se modifican los procedimientos existente, simplemente se hace una grabación en alta calidad sin compresión de las secuencias de endoscopia.

En el caso de la información tomográfica se emplean las misma pruebas que se emplearon en el diagnóstico.

Los aspectos éticos se reducen a incluir el uso secundario de datos en el consentimiento informado y a garantizar la confidencialidad del paciente en en el uso de los datos durante el proyecto. Al ser imágenes del interior del cuerpo basta con emplear claves para su identificación.

Todos los datos proporcionados por el hospital para su estudio han sido debidamente anonimizados para ocultar la identidad del paciente.

9.2. Formato de fichero OpenGL

Un objeto OpenGL se guarda en un fichero de texto con extensión *obj* el cual contiene una descripción de todos los elementos que caracterizar el modelo con una etiqueta que identifica el tipo de campo y los valores que contiene a continuación.

La lectura del fichero se debe realizar en orden, ya que la referencia entre vértices, vectores normales y textura se aplican con el índice del vector, que empieza en 1 y se asigna su índice en función del orden en el fichero.

Los tipos de cabecera estándar son los siguientes, la primera palabra identifica su tipo y el resto de la línea son los datos:

- $v\ x\ y\ z \Rightarrow$ Identifica el vértice i -ésimo con sus coordenadas.
- $vt\ u\ v \Rightarrow$ Identifica la textura i -ésima con las coordenadas de la imagen de donde se obtiene la textura.
- $vn\ x\ y\ z \Rightarrow$ Identifica el vector normal i -ésimo con su dirección.
- $f\ v1/vt1/vn1\ v2/vt2/vn2\ v3/vt3/vn3 \Rightarrow$ Identifica la cara de un triángulo con tres datos, cada grupo de 3 componentes identifica las características de un vértice (coordenadas/normal/textura). Solo es obligatorio definir los 3 vértices, la normal y la textura se puede ignorar.

Ante la necesidad de incluir la información de las anclas se han agregado más campos siguiendo el mismo estilo:

- $r [i] [x] [y] [z] \Rightarrow$ Identifica el ancla i con sus coordenadas.
- $d [i] [j] [m] \Rightarrow$ Identifica la distancia en valor absoluto entre las anclas i y j en m milímetros.

9.3. Control de versiones

Para el desarrollo del software se ha partido de la versión publica de ORB-SLAM2 y se creado una nueva versión que parte de ella en el repositorio privado de la universidad en GitHub, el cual mantiene el código con control de versión, el cual permite monitorizar los cambios a lo largo del tiempo y revertir cambios si fuera necesario.

Este tipo de herramientas es muy empleada en el desarrollo de software ya que permite la edición de grandes proyectos entre diferentes personas.

El código se va a mantener en el repositorio de la universidad para que sea accesible para todos los miembros del departamento.

Bibliografía

- [1] Nader Mahmoud, Toby Collins, Alexandre Hostettler, Luc Soler, Christophe Doignon, and J.M.M.Montiel. Live tracking and dense reconstruction for hand-held monocular endoscopy. *IEEE Transaction on Medical Imaging*, 2018.
- [2] Sylvain Bernhardt, Stéphane A Nicolau, Luc Soler, and Christophe Doignon. The status of augmented reality in laparoscopic surgery as of 2016. *Medical image analysis*, 37:66–90, 2017.
- [3] Ronald T Azuma. A survey of augmented reality. *Presence: Teleoperators & Virtual Environments*, 6(4):355–385, 1997.
- [4] Khronos Group. Open source graphic library. <https://www.opengl.org/>, 2017.
- [5] Raúl Mur-Artal and Juan D. Tardós. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Transactions on Robotics*, 33(5):1255–1262, 2017.
- [6] Peter Mildnerberger, Marco Eichelberg, and Eric Martin. Introduction to the DICOM standard. *European radiology*, 12(4):920–927, 2002.
- [7] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: An efficient alternative to sift or surf. In *Computer Vision (ICCV), 2011 IEEE international conference on*, pages 2564–2571. IEEE, 2011.
- [8] Georgios D Evangelidis and Emmanouil Z Psarakis. Parametric image alignment using enhanced correlation coefficient maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(10):1858–1865, 2008.
- [9] Berthold KP Horn. Closed-form solution of absolute orientation using unit quaternions. *JOSA A*, 4(4):629–642, 1987.
- [10] Itseez. Open source computer vision library. <https://github.com/itseez/opencv>, 2015.