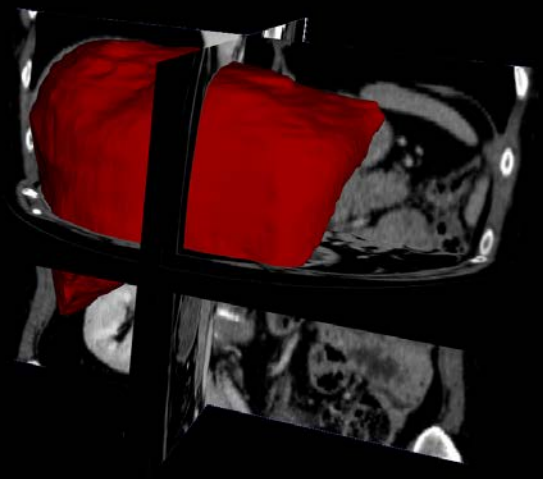


# Advanced techniques in medical image segmentation of the liver

Fernando López Mir

**Tesis doctoral**
Diciembre, 2015

Departamento de Ingeniería electrónica

**Directora**
Dra. Valery Naranjo Ornedo

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

# Advanced techniques in medical image segmentation of the liver

*Autor:* Fernando López Mir

*Directora:* Dra. Valery Naranjo Ornedo

UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Departamento de Ingeniería Electrónica

Programa de Doctorado en Tecnologías para la Salud y el Bienestar

Diciembre, 2015

El trabajo derivado de esta tesis doctoral ha sido financiado parcialmente por el IMPIVA (ref. IMIDTF/2009/83), por el MITYC (ref. TSI-020100- 2009-189) y por el CDTI (ref. IDI-20101153) bajo los proyectos titulados *"Realidad aumentada en cirugía abdominal"*, *"Sistema de realidad aumentada para navegación en cirugía laparoscópica"* y *"Terapias colaborativas asistidas para tratamientos oncológicos"*, respectivamente.

Además, el trabajo de Fernando López Mir para desarrollar esta investigación ha sido financiado por LabHuman, grupo que pertenece al Instituto de Investigación en Bioingeniería y Tecnología orientada al Ser Humano de la Universidad Politecnica de Valencia.

# Agradecimientos

Son muchos los que han hecho posible esta tesis doctoral y espero no olvidar a nadie por nombrar. Quiero ser escueto en mis agradecimientos lo que no quiere decir que no se merezcan unas lineas amigos, familiares, compañeros, etc.

En primer lugar a mis padres, porque soy lo que soy gracias a ellos y si he conseguido esto es porque han sabido inculcarme la filosofía del trabajo, la responsabilidad, el sacrificio pero sobre todo por su apoyo incondicional. También a hermanos (Manolín, Sergio y Maria) y resto de familiares en especial a mi abuela y a mi tío que ya no están, pero que no puedo más que recordar con cariño.

Después de ellos no puedo olvidar a la otra persona responsable que esta tesis salga a la luz. Lo mejor que me llevo de estos 7 años de ella es poder decir que somos amigos, por encima de todo. Su cariño, su paciencia, su guía, tantas cosas que agradecer ... He aprendido tanto, que no tendría páginas suficientes para demostrárselo y agradecérselo. Gracias Valery.

Finalmente a mis compañeros y amigos de LabHuman: Sandra, Fran, Miguel, Rober, Pablo, Adri, Eliseo, Juanjo, Javi, Fran, Félix, Conchi, Jorge, Adrian, etc. a mi tutor de acogida en mi estancia doctoral en *Fontainebleau* (Jesús) y a Andrés junto al resto de *"malakas"* y a mis amigos: Vicente, Víctor, Jesús, Miguel, Carlos, Clemente, Juanto, Iñaqui, David, Jorge, Manolo, Felipe, Miguel Ángel, etc. Y no por último menos importante, a Kike.

*"Si no puedes explicar algo de forma sencilla, es que ni tú mismo lo has entendido lo suficiente"*

A. Einstein

# Resumen

La segmentación de imágenes es, junto al registro multimodal y monomodal, la operación con mayor aplicabilidad en tratamiento digital de imagen médica. Son multitud las operaciones y filtros, así como las aplicaciones y casuística, que derivan de una segmentación de un tejido orgánico. El caso de segmentación del hígado en imágenes radiológicas es, después del cerebro, la que mayor número de publicaciones científicas podemos encontrar. Esto es debido por un lado a la necesidad de seguir innovando en los algoritmos ya existentes y por otro a la gran aplicabilidad que tiene en muchas situaciones relacionadas con el diagnóstico, tratamiento y seguimiento de patologías hepáticas pero también para la planificación clínica de las mismas.

En el caso de imágenes de resonancia magnética, sólo en los últimos años han aparecido soluciones que consiguen buenos resultados en cuanto a precisión y robustez en la segmentación del hígado. Sin embargo dichos algoritmos, por lo general son poco utilizables en el ambiente clínico. En el caso de imágenes de tomografía computarizada encontramos mucha más variedad de metodologías y soluciones propuestas pero es difícil encontrar un equilibrio entre precisión y uso práctico clínico.

Es por ello que para mejorar el estado del arte en ambos casos (imágenes de resonancia magnética y tomografía computarizada) en esta tesis se ha planteado una metodología común a la hora de diseñar y desarrollar sendos algoritmos de segmentación del hígado en las citadas modalidades de imágenes anatómicas. Primero, y tras el estudio del estado del arte, se ha diseñado un algoritmo adecuado a los parámetros de precisión y robustez necesarios para la aplicabilidad deseada. En dichos algoritmos se ha hecho uso de diferentes técnicas y métodos ya publicados y en muchos casos se han mejorado y diseñado algunos nuevos. Además se ha tenido en cuenta la iniciación e interacción de dichos algoritmos para que sea mínima y el coste de computación para que se adapte a las necesidades clínicas.

El segundo paso ha sido la validación de ambos algoritmos. En el caso de imágenes de tomografía computarizada existen bases de datos públicas con imágenes segmentadas manualmente por expertos y que la comunidad científica suele utilizar como nexo común a la hora de validar y posteriormente comparar sus algoritmos. La validación se hace mediante la obtención de determinados coeficientes de similitud entre la imagen segmentada manualmente por los expertos y las que nos proporciona el algoritmo. Esta forma de validar la precisión del algoritmo ha sido la seguida en esta tesis, con la salvedad que en el caso de imágenes de resonancia magnética no existen bases de datos de acceso público. Por ello, y para este caso, lo que se ha hecho es la creación previa de una base de datos propia donde diferentes expertos radiólogos han segmentado manualmente diferentes estudios de pacientes con el fin de que puedan servir como referencia y se pueda seguir la misma metodología que en el caso anterior. Dicha base de datos ha hecho posible que la validación se haga en 17 estudios (con más de 1.500 imágenes), lo que convierte la validación de este método de segmentación del hígado en imágenes de resonancia magnética en una de las más extensas publicadas hasta la fecha. Por el contrario, la comparación de los métodos propuestos con los presentados por otros autores no es directa como sí pasa en el caso del algoritmo de segmentación de hígado en tomografía computarizada.

La validación y posterior comparación han dejado patente una precisión superior al 90 % reflejado en el coeficiente de *Jaccard* y *Dice*, muy en consonancia con valores publicados por la inmensa mayoría de autores que se han podido comparar. Sin embargo, y en general, los algoritmos planteados en esta tesis han obtenido unos criterios de uso mucho mayores, ya que en general presentan menores costes de computación, una interacción clínica casi nula y una iniciación nula en el caso del algoritmo de resonancia magnética y casi nula en el caso de algoritmos de tomografía computarizada. Es por ello, que el diseño y validación de ambos algoritmos suponen una mejora en el estado del arte de la segmentación del hígado en imágenes de alta resolución anatómica.

En esta tesis, también se ha abordado un tercer punto que hace uso

de los resultados obtenidos en la segmentación del hígado en imágenes de resonancia magnética. Para ello, y haciendo uso de algoritmos de realidad aumentada, se ha diseñado y validado un estudio real inocuo y no invasivo para el clínico y para el paciente donde se ha demostrado que la utilización de esta tecnología reporta mayores beneficios en cuanto a mayor precisión y menor variabilidad frente a su no uso en un caso concreto de cirugía con laparoscopia. En particular, se ha implementado un sistema de ayuda para la colocación de los trocares en este tipo de cirugía. A partir de la segmentación previa del hígado, y otros órganos abdominales de interés, se ha obtenido un modelo 3D del paciente. Una vez en el quirófano, gracias a una cámara, se ha adquirido la imagen del abdomen del paciente en tiempo real detectando su posición mediante una marca centrada en su ombligo. De este modo el modelo 3D de los órganos obtenido anteriormente se ha registrado y fusionado con la imagen, permitiendo al clínico tener una herramienta más para ayudar a localizar los puntos donde hacer las incisiones por donde introducirá los trocares.

# Abstract

Image segmentation is, along with multimodal and monomodal registration, the operation with the greatest applicability in medical image processing. There are many operations and filters, as much as applications and cases, where the segmentation of an organic tissue is the first step. The case of liver segmentation in radiological images is, after the brain, that on which the highest number of scientific publications can be found. This is due, on the one hand, to the need to continue innovating in existing algorithms and, on the other hand, to the applicability in many situations related to diagnosis, treatment and monitoring of liver diseases but also for clinical planning.

In the case of magnetic resonance imaging (MRI), only in recent years some solutions have achieved good results in terms of accuracy and robustness in the segmentation of the liver. However these algorithms are generally not user-friendly. In the case of computed tomography (CT) scans more methodologies and solutions have been developed but it is difficult to find a good trade-off between accuracy and practical clinical use.

To improve the state-of-the-art in both cases (MRI and CT), a common methodology to design and develop two liver segmentation algorithms in those imaging modalities has been proposed in this thesis. First, after studying the state-of-the-art, a suitable algorithm with the accuracy and robustness needed for the desired applicability in each case (MRI and CT) has been designed. These algorithms have been designed using different techniques and methods previously published although in many cases they have been improved and new ones have been designed. The initiation and interaction of these algorithms and their computation cost have also taken into account.

The second step has been the validation of both algorithms. In the case of CT images, there exist public databases with images segmented manually by experts that the scientific community uses as a

common link for the validation and comparison of their algorithms. The validation is done by obtaining certain coefficients of similarity between the manual and the automatic segmentation. This way of validating the accuracy of the algorithm has been followed in this thesis, except in the case of magnetic resonance imaging because, at present, there are no databases publicly available. In this case, there aren't public or accessible images. Accordingly, a private database has been created where several expert radiologists have manually segmented different studies of patients that have been used as a reference. This database is composed by 17 studies (with more than 1,500 images), so the validation of this method in MRI is one of the more extensive currently published. By contrast, the comparison of the proposed methods with other state-of-the-art algorithms is not direct as in the case of the algorithm for liver segmentation in CT.

In the validation stage, an accuracy above 90 % in the *Jaccard* and *Dice* coefficients has been achieved. The vast majority of the compared authors achieves similar values. However, in general, the algorithms proposed in this thesis are more user-friendly for clinical environments because the computational cost is lower, the clinical interaction is non-existent and it is not necessary an initiation in the case of the magnetic resonance algorithm and a small initiation (it is only necessary to introduce a manual seed) for the computed tomography algorithm. For this reason, the design and validation of both algorithms are an improvement in the state-of-the-art of liver segmentation in high resolution anatomical images.

In this thesis, a third hypothesis that makes use of the results of liver segmentation in MRI combined to augmented reality algorithms has also been developed. Specifically, a real and innocuous study, non-invasive for clinician and patient has been designed and validated through it has been shown that the use of this technology creates benefits in terms of greater accuracy and less variability versus the non-use in a particular case of laparoscopic surgery. In particular, a support system for trocar placement in this type of surgery was implemented. A 3D model of the patient's abdomen is obtained from the previous segmentation of the

liver, and other abdominal organs of interest. In the operating room, thanks to a camera, a real image of the patient's abdomen is obtained and its position is detected with a mark centered in their navel. Then, the 3D model of the organs is registered and fused with this real image, to provide the clinician with an aid tool for the location of the incision points through the trocars will be introduced.

# Resum

La segmentació d'imatges és, al costat del registre multimodal i monomodal, l'operació amb major aplicabilitat en tractament digital d'imatge mèdica. Són multitud les operacions i filtres, així com les aplicacions i casuística, que comencen en la segmentació d'un teixit orgànic. El cas de segmentació del fetge en imatges radiològiques és, després del cervell, la que major nombre de publicacions científiques podem trobar. Això és degut per una banda a la necessitat de seguir innovant en els algoritmes ja existents i per un altre a la gran aplicabilitat que té en moltes situacions relacionades amb el diagnòstic, tractament i seguiment de patologies hepàtiques però també per a la planificació clínica de les mateixes.

En el cas d'imatges de ressonància magnètica, només en els últims anys han aparegut solucions que aconsegueixen bons resultats quant a precisió i robustesa en la segmentació del fetge. No obstant això aquests algoritmes, en general són poc utilitzables en l'ambient clínic. En el cas d'imatges de tomografia computeritzada trobem molta més varietat de metodologies i solucions proposades però és difícil trobar un equilibri entre precisió i ús pràctic clínic.

És per això que per millorar l'estat de l'art en els dos casos (imatges de ressonància magnètica i tomografia computeritzada) en aquesta tesi s'ha plantejat una metodologia comuna a l'hora de dissenyar i desenvolupar dos algoritmes de segmentació del fetge en les esmentades modalitats d'imatges anatòmiques. Primer, i després de l'estudi de l'estat de l'art, s'ha dissenyat un algoritme adequat als paràmetres de precisió i robustesa necessaris per a l'aplicabilitat desitjada. En aquests algoritmes s'ha fet ús de diferents tècniques i mètodes ja publicats i en molts casos s'han millorat i dissenyat nous. A més s'ha tingut en compte la iniciació i interacció d'aquests algoritmes perquè siga mínima i el cost de computació perquè s'adapte a les necessitats clíniques.

El segon pas ha estat la validació de tots dos algoritmes. En el cas d'imatges de tomografia computeritzada hi ha bases de dades

públiques amb imatges segmentades manualment per experts i que la comunitat científica sol utilitzar com a nexe comú a l'hora de validar i posteriorment comparar els seus algoritmes. La validació es fa mitjançant l'obtenció de determinats coeficients de similitud entre la imatge segmentada manualment pels experts i les que ens proporciona l'algoritme. Aquesta forma de validar la precisió de l'algoritme ha estat la seguida en aquesta tesi, amb l'excepció que en el cas d'imatges de ressonància magnètica no hi ha bases de dades d'accés públic. Per això, i per a aquest cas, el que s'ha fet és la creació prèvia d'una base de dades pròpia on diferents experts radiòlegs han segmentat manualment diferents estudis de pacients amb la finalitat que puguen servir com a referència i es puga seguir la mateixa metodologia que en el cas anterior. Aquesta base de dades ha fet possible que la validació es faja en 17 estudis (amb més de 1.500 imatges), cosa que converteix la validació d'aquest mètode de segmentació del fetge en imatges de ressonància magnètica en una de les més extenses publicades fins a la data . Per contra, la comparació dels mètodes proposats amb els presentats per altres autors no és directa com si passa en el cas de l'algorisme de segmentació de fetge en tomografia computeritzada.

La validació i posterior comparació han deixat patent una precisió superior al 90 % reflectit en el coeficient de textit Jaccard i textit Dice, molt d'acord amb valors publicats per la immensa majoria d'autors en que s'ha pogut comparar. No obstant això, i en general, els algoritmes plantejats en aquesta tesi han obtingut uns criteris d'ús molt més grans, ja que en general presenten menors costos de computació, una interacció clínica quasi nul·la i una iniciació nul·la en el cas de l'algoritme de ressonància magnètica i quasi nul·la en el cas d'algoritmes de tomografia computeritzada. És per això, que el disseny i validació de tots dos algoritmes suposen una millora en l'estat de l'art de la segmentació del fetge en imatges d'alta resolució anatòmica.

En aquesta tesi, també s'ha abordat un tercer punt que fa ús dels resultats obtinguts en la segmentació del fetge en imatges de ressonància magnètica. Per a això, i fent ús d'algoritmes de realitat augmentada, s'ha dissenyat i validat un estudi real innocu i no invasiu per al clínic i

per al pacient on s'ha demostrat que la utilització d'aquesta tecnologia reporta més beneficis pel que fa a major precisió i menor variabilitat enfront del seu no ús en un cas concret de cirurgia amb laparoscòpia. En particular, s'ha implementat un sistema d'ajuda per a la col·locació dels trocars en aquest tipus de cirurgia. A partir de la segmentació prèvia del fetge, i altres òrgans abdominals d'interès, s'ha obtingut un model 3D del pacient. Una vegada al quiròfan, gràcies a una càmera, s'ha adquirit la imatge del abdomen del pacient en temps real i s'ha detectat la seua posició per mig de una marca centrada en el seu melic. D'aquesta manera el model 3D dels òrgans obtingut anteriorment s'ha registrat i fusionat amb la imatge, permetent al clínic tenir una ferramenta més que li ajuda a localitzar els punts on fer les incisions per on introduirà els trocars.

# Índice general

# Capítulo 1

# Introducción

## Índice

## 1.1. Motivación

Tanto la resonancia magnética (RM) como la tomografía computarizada (TC) han constituido grandes avances tecnológicos muy relevantes en la medicina (y en la hepatología en particular) para el diagnóstico de múltiples enfermedades incluso en etapas iniciales. La observación en los estudios anatómicos de un excelente contraste en los tejidos blandos, en el caso de la RM, y las velocidades de adquisición más rápidas, la mayor resolución espacial y la mejor calidad de las imágenes, en el caso de la TC, han hecho de estas metodologías una aplicación de incalculable valor en diferentes fases (detección, seguimiento, intervención, etc.) de numerosas patologías hepáticas (Sorensen, 2006). Gracias a estos tipos de estudios es posible hoy en día obtener un amplio rango de información que abarca desde aspectos como el tamaño, localización, extensión e incluso en algunos casos el grado de heterogeneidad de órganos y posibles anomalías que hubiera en ellos (Hoffman, 2001; Boxerman et al., 1998; Riederer, 1996). Sobre este particular en los últimos años se han realizado numerosas mejoras técnicas creando un ambiente muy prometedor que ha animado al desarrollo de nuevas herramientas, productos de contraste y nuevos métodos para la obtención de imágenes con mayor calidad e información más detallada. Además, y de manera paralela, también se están mejorando la parte del post-procesado de imagen aumentando la calidad de las mismas lo que redunda en una herramienta de gran ayuda para realizar un mejor diagnóstico médico.

La gran variedad de patologías que se pueden evaluar mediante estas metodologías y su alta resolución anatómica han dado lugar a su empleo como pruebas rutinarias en multitud de patologías como indicador de respuesta al tratamiento (Hamstra et al., 2007; Garner et al., 2009; Traboulsee and Li, 2008; Hunt et al., 2009; Huang et al., 2008). Esta aproximación, aunque presenta ciertas limitaciones, continúa siendo en la actualidad la base para las clasificaciones de diagnóstico y terapéutica (Therasse et al., 2000). Por ello, un gran número de pruebas realizadas en las unidades de RM y un gran número de estudios TC realizados a los pacientes, se encuentran destinados a tal propósito. Si se tiene en

cuenta que en primer lugar, la RM y la TC permiten una delineación tridimensional y en segundo lugar, que la utilización de diversos tipos de secuencias permite una caracterización del tejido sometido a estudio junto con el empleo de agentes de contraste, es lógico comprender su amplia demanda y su alta utilización en el campo de la hepatología. Sin embargo tanto la RM como la TC aún presentan ciertas limitaciones, muchas de las cuales vienen determinadas por el tipo de análisis de la información que se obtiene. Esto es debido a que por lo general los estudios son analizados de manera cualitativa y por lo tanto muchas veces poco objetiva. A este hecho debe añadirse el factor de que por lo general no siempre un mismo observador efectúa el estudio de un mismo paciente por lo que a la falta de objetividad del proceso debe unirse el criterio del propio observador que en muchos casos puede divergir. Todo esto plantea la necesidad de una estandarización y validación de los métodos para facilitar sus diferentes usos en las diferentes instituciones dentro del ambiente clínico. Numerosos esfuerzos se están realizando para mejorar y estandarizar aspectos como la obtención de variables cuantitativas así como para mejorar la visualización de las imágenes que permitirán un mayor desarrollo, accesibilidad y aplicabilidad de esta metodología.

## 1.2. Objetivos

El primer objetivo de esta tesis ha sido diseñar, desarrollar y validar dos algoritmos de segmentación del hígado en imágenes anatómicas de RM y de TC respectivamente. En particular dichos algoritmos tienen las siguientes características:

- Automatismo: los algoritmos deben de tener un grado de automatismo alto y la interacción del usuario debe de ser mínima con el fin de no elevar su tiempo de dedicación clínica.
- Fácil iniciación: los parámetros de iniciación (en el caso de que fueran necesarios) de los algoritmos deben ser fácilmente interpretables por los clínicos.

- Coste computacional: el coste de computación de los algoritmos debe ser bajo y pensado para que se puedan ejecutar en un ordenador de sobremesa y de uso cotidiano al que pueda tener acceso la mayoría de profesionales sanitarios.

- Adaptabilidad: los algoritmos deberán poder integrarse bajo alguna plataforma (aplicación, software, sistema, etc.) para su uso clínico.

- Precisión: se validarán ambos algoritmos con la finalidad de que tengan la suficiente precisión y robustez para los fines que se diseñan.

De manera complementaria, otro objetivo de esta tesis ha sido el diseño, desarrollo y validación de un experimento que haga uso de la segmentación del hígado antes propuesta en un caso real de tratamiento de una patología hepática donde sea necesaria dicha segmentación.

Es por ello que la tesis se plantea en tres fases. En la fase uno se procederá, teniendo en cuenta el estado de la investigación, al diseño, desarrollo y validación de un algoritmo de segmentación de imágenes de resonancia magnética. En la segunda fase se procederá, siguiendo el mismo esquema, a diseñar, desarrollar y validar un algoritmo de segmentación de imágenes de tomografía computarizada. En la tercera, y última fase, se procederá al diseño de un experimento clínico que demuestre la utilidad de esta segmentación. Este experimento se llevará a cabo en una situación real pero controlable por el clínico y totalmente trasparente e inocua para él y para el paciente.

Como hemos dicho, en esta tesis se presentan nuevos algoritmos y se combinan con otros ya existentes basados en morfología matemática, en la mayoría de casos, con el fin de extraer el hígado en imágenes de RM y TC.

El método presentado en el capítulo 2 se basa en una variante de la transformada *watershed*[1] estocástica. Para ello, gracias a este nuevo método se segmenta de forma automática el hígado en una primera imagen de la secuencia 3D. A partir de la información de dicha segmentación se extrae el hígado de las imágenes adyacentes de forma secuencial a partir del resultado de la segmentación de la imagen que le precede.

Por otra parte, las imágenes 3D que se obtienen por tomografía computarizada presentan una serie de ventajas que hacen que la segmentación en este tipo de imágenes sea mucho más abordable que en el caso de otro tipo de imágenes anatómicas como las de resonancia magnética. Su mejor resolución y más rápida adquisición hacen que problemas como el ruido o los artefactos sean menores. Si analizamos la literatura existen muchos algoritmos publicados que en general obtienen buenos resultados pero adolecen de algunas características que complican su uso en el ámbito clínico: necesitan de entrenamiento previo, son difíciles de inicializar, necesitan excesiva interacción o tienen un alto coste de computación. El capítulo 3 de esta tesis hace uso de un algoritmo basado en un crecimiento de regiones para posteriormente hacer uso de morfología 3D en una fase de post-procesado, y así obtener una segmentación del hígado lo suficientemente precisa con una fácil interacción por parte del clínico y un bajo coste computacional.

La realidad aumentada es una técnica gráfica que nos permite fusionar elementos virtuales con elementos reales en tiempo real. Actualmente, dicha técnica está altamente extendida en campos temáticos como la educación, el entretenimiento o el marketing, pero cada vez más su uso en la medicina tiene más relevancia pues permite de una forma no invasiva aumentar la información en beneficio del clínico. En el capítulo 4 se presenta un sistema que se utiliza en un acto previo

[1] *Watershed* es un término proveniente de la topografía y representa la línea divisoria entre dos cuencas hidrográficas contiguas. En procesado de imagen, los watershed son las líneas que separan áreas de influencia de mínimos locales adyacentes de la propia imagen.

a la realización de una cirugía con laparoscopia en pacientes. En dicho estudio, además de analizar y valorar diferentes arquitecturas hardware, se presenta un estudio controlado sobre dos grupos de pacientes; en uno se usa el sistema de realidad aumentada y en el otro no. Con este experimento se ha demostrado una mejora en la precisión del cirujano al hacer las incisiones necesarias en este tipo de operaciones cuando se usa el sistema de realidad aumentada y por consiguiente la utilidad que tiene la segmentación en aplicaciones reales de tratamiento hepático.

## 1.3. Principales contribuciones

### CONTRIBUCIONES A LA SEGMENTACIÓN DEL HÍGADO EN RESONANCIA MAGNÉTICA

Si analizamos la literatura hay pocos métodos de segmentación del hígado en imágenes de resonancia magnética en comparación a otras modalidades anatómicas. En esta tesis se presenta un nuevo método para este fin que se basa en el uso de la transformada *watershed* y de particiones estocásticas. La principal contribución del capítulo 2 de esta tesis ha sido la creación de una nueva función de gradiente que para el caso que nos ocupa presenta mejores prestaciones que el gradiente tradicional como valor de entrada para el algoritmo de la transformada *watershed*. El clásico problema que ofrece la transformada *watershed* relacionado con la sobre-segmentación ocasionada por los mínimos regionales se ha reducido con el uso de marcadores. Estos marcadores imponen nuevos mínimos en la imagen controlando la cantidad de regiones a segmentar. La definición de estos marcadores en un órgano como el hígado no es fácil ya que tiene estructuras internas con alto contraste como los vasos. La asignación manual de estos marcadores es ineficiente, por ello se ha diseñado una nueva metodología con el fin de obtener una función densidad de probabilidad que ofrece mejor contraste que el gradiente tradicional. Esta variante hace que de forma automática y de manera sucesiva y aleatoria se generen una serie de marcadores que conforman una función densidad de probabilidad de

contornos que tiene en cuenta el contraste entre regiones adyacentes gracias a un modelo basado en regiones. Se ha demostrado que esta nueva función densidad de probabilidad presenta menos mínimos y mayor contraste que el gradiente tradicional lo que ofrece mejores prestaciones en el caso de la segmentación del hígado. Además en este primer capítulo se han usado una serie de filtros en las fases de pre-procesado y de post-procesado que hacen más robusto el método y finalmente, se ha diseñado un clasificador que permite la segmentación del hígado de una manera automática y suficientemente precisa para su aplicación clínica.

### CONTRIBUCIONES A LA SEGMENTACIÓN DEL HÍGADO EN TOMOGRAFÍA COMPUTARIZADA

La segmentación del hígado tiene muchas aplicaciones como primer paso por ejemplo en cirugía guiada de hígado o en planificación y seguimiento de enfermedades hepáticas. Hoy en día existen métodos publicados que segmentan de una manera razonablemente precisa el hígado en imágenes de tomografía computarizada. El problema es que la mayoría de dichos métodos presentan carencias para su uso clínico debido al elevado coste computacional, a la excesiva dependencia e interacción por parte del clínico o al alto grado de entrenamiento previo que necesitan. El objetivo y principal aportación realizada en el capítulo 3 ha sido presentar un método basado en un crecimiento de regiones 3D que aporta suficiente precisión pero también su utilización es factible en un entorno clínico. Para asegurarse la independencia en la elección de la semilla necesaria para que el algoritmo se inicie, se ha realizado un estudio que demuestra que los resultados son prácticamente invariantes a la semilla seleccionada. Además se ha diseñado un nuevo filtro adaptativo a los parámetros intrínsecos de las imágenes hepáticas y que se ha utilizado previamente a la segmentación. Dicho filtro ha demostrado mejores resultados que otros publicados en la literatura. Finalmente, la metodología presentada aporta un post-procesado basado en filtros morfológicos 3D que demuestran un mejor comportamiento que sus análogos 2D. La finalidad de estos

filtros ha sido la eliminación de estructuras adyacentes al hígado y que con un nivel de gris similar han sido clasificadas erróneamente y la re-clasificación de algunas zonas que en un principio el algoritmo caracteriza como tejido no hepático. Con todas estas aportaciones se mantiene el objetivo de tener una precisión y robustez razonables, pero manteniendo un grado alto de automatismo, un bajo coste computacional y una casi nula interacción por parte del usuario respecto a los métodos publicados en el estado del arte. Por ello, se ha demostrado así que el grado de usabilidad es alto.

### CONTRIBUCIONES A LA CIRUGÍA CON LAPAROSCOPIA DE LA SEGMENTACIÓN DEL HÍGADO Y LA REALIDAD AUMENTADA

Otras aportaciones que ha realizado esta tesis son las relacionadas con el capítulo 4. En dicho capítulo se ha diseñado y desarrollado un sistema hardware y software que hace uso de realidad aumentada y de la segmentación del hígado en imágenes de resonancia magnética. Son varias las aportaciones relevantes de dicho estudio. En primer lugar, se ha extraído la precisión de los algoritmos de realidad aumentada con una metodología basada en *phantoms*[2]. Posteriormente, se ha pasado a definir el sistema más adaptado a las necesidades clínicas desde un punto de vista de ergonomía y de la usabilidad eligiendo aquel que mejores prestaciones ofrecía al cirujano y al paciente. En base a la experiencia, se ha definido una nueva métrica que explica los diferentes factores que influyen a la hora de tener que hacer una incisión en una operación de laparoscopía para posteriormente colocar por dicha incisión el trocar correspondiente. Con esta nueva métrica como referencia, se han realizado dos experimentos no invasivos que presentan la ventaja de que han sido probados en situaciones reales, cosa poco común en la literatura. En uno de estos experimentos se usó el sistema de realidad aumentada de ayuda a la colocación de

[2]En el campo de la imagen médica, un *phantom* es un objeto especialmente diseñado que se escanea con el fin de evaluar, analizar o ajustar el rendimiento de un dispositivo de imagen o un algoritmo específico.

trocares y en el otro no. Ambos experimentos fueron en operaciones de colicestomía por el bajo grado de complejidad antes, durante y después de la cirugía. Finalmente, varios clínicos han participado de manera aleatoria en los diferentes experimentos demostrando la mejora en precisión y variabilidad en el caso de usar un sistema de ayuda a la colocación de trocares basado en realidad aumentada frente a su no uso.

## 1.4. Contexto

Esta tesis está enmarcada en diferentes proyectos de investigación: "Realidad aumentada en cirugía abdominal"(IMIDTF/2009/83), "Terapias asistidas colaborativas para el tratamiento oncológico mediante el uso de tecnologías TIC - ONCOTIC"(IDI-20101153) y "Sistema de Navegación mediante realidad virtual en cirugía laparoscópica - NARALAP"(TSI-020100-2009-189). Estos proyectos han sido financiados por el hospital clínica Benidorm a través del Instituto de la Mediana y Pequeña Industria Valenciana (IMPIVA) y del Centro para el Desarrollo Tecnológico Industrial (CDTI) y por el Ministerio de Industria, Turismo y Comercio a través del sub-programa AVANZA I+D en la acción estratégica de telecomunicaciones y sociedad de la información, respectivamente.

El primero de ellos tenía como objetivo desarrollar y mejorar algoritmos de realidad aumentada adaptados a los requisitos propios de un entorno quirúrgico en cuanto a precisión y fiabilidad. Para ello, otro de los objetivos marcados fue el desarrollo de algoritmos semiautomáticos de segmentación de órganos abdominales y la validación del prototipo final en un *phantom*.

El proyecto Naralap tenía como objetivo desarrollar un sistema de navegación con laparoscopia fundamentado en dos pilares: la realidad virtual/aumentada y un modelo biomecánico del hígado. En el caso del uso de la realidad aumentada, se definió su uso en un caso concreto para mejorar la colocación de trocares en cirugía con laparoscopia. La cirugía

laparoscópica tiene una serie de desventajas para el cirujano. Éstas, se traducen en la importancia del posicionamiento preciso del punto de entrada de los trocares, que normalmente se determina mediante palpación externa, la falta de visión directa y la sustitución de ésta por visión 2D en un monitor, falta de percepción táctil o necesidad de destreza y coordinación ojo-mano. Con el uso de la realidad aumentada se pretende mejorar algunas de estas desventajas. El objetivo del proyecto era avanzar en el conocimiento que permitiera aproximar esta tecnología al problema de la navegación con laparoscopia.

El proyecto ONCOTIC se compone de 6 sub-proyectos que aplican diferentes tecnologías gráficas para mejorar el diagnóstico, tratamiento o seguimiento de diferentes patologías oncológicas. Uno de estos sub-proyectos, HepaPlan, tiene como objetivo la creación de un software para el análisis prospectivo y retrospectivo de lesiones evidenciadas en los estudios de RM y/o TC anatómicos. Para ello, es necesario la segmentación y registro de imágenes de tomografía axial computarizada para extraer el hígado en primer lugar y estructuras internas como el árbol hepático (arterial, portal, suprahepático, biliar) y las lesiones si las hubiera.

Estos tres proyectos citados tienen en común que parten de la necesidad de la segmentación del hígado en imágenes anatómicas de alta resolución (resonancia magnética o tomografía computarizada) bien para ayudar en la planificación de cirugías, como parte de un sistema de navegación o como apoyo a clínicos en fases de diagnóstico o seguimiento de determinadas patologías oncológicas.

## 1.5. Estructura

Esta tesis se ha divido en 5 capítulos, de los cuales los tres capítulos centrales corresponden a tres publicaciones en revistas con factor de impacto. En el primer capítulo se ha introducido la motivación y los objetivos genéricos de la tesis, las principales contribuciones así como el contexto y la estructura de la misma.

El capítulo 2 corresponde al artículo con título: *"Liver segmentation in MRI: A fully automatic method based on stochastic partitions"* publicado en la revista *Computer Methods and Programs in Biomedicine* en el volumen 114, número 1 del año 2014, páginas 11-28 por los autores Fernando López Mir, Valery Naranjo Ornedo, Jesús Angulo López, Mariano Luis Alcañiz Raya y Luis Luna. El articulo trata el desarrollo y posterior validación de un método de segmentación de imágenes de resonancia magnética.

El capítulo 3 coincide con el artículo con título: *"A method for liver segmentation on computed tomography images in venous phase suitable for real environments"* publicado en la revista *Journal of Medical Imaging and Health Informatics* del volumen 5 del año 2015, páginas 1208-1216 por los autores Fernando López Mir, Valery Naranjo Ornedo, Pablo González Pérez, Eugenia Pareja Ibars, Sandra Morales Martínez y Jaime Solaz Mínguez. El artículo pretende abordar el desarrollo y validación de un método de segmentación para imágenes de tomografía computarizada que no sólo sea lo suficientemente preciso y robusto sino que también cumpla otros requisitos que lo hagan más útil desde un punto de vista clínico como por ejemplo haciendo fácil su interacción, con un bajo coste de computación, con una fácil iniciación, etc.

El capítulo 4 corresponde al artículo: *"Design and Validation of an Augmented Reality System for Laparoscopic Surgery in a Real Environment"* publicado en la revista *BioMed Research International* en el volumen 2013, páginas 1-12, por los autores Fernando López Mir, Valery Naranjo Ornedo, Juan José Fuertes Cebrián, Eliseo Villanueva Morte, José Bueno Lledó y Eugenia Pareja Ibars. El artículo trata sobre el diseño de un sistema de realidad aumentada partiendo del modelo 3D de estructuras abdominales de cada paciente extraído a partir de imágenes de resonancia magnética.

La estructura de los capítulos 2-4 ha sido común. En primer lugar se ha realizado un resumen, seguido de una introducción que se compone de un estado de la investigación y la problemática a tratar. Después se

ha explicado la metodología a seguir desde un punto de vista teórico. En el tercer punto, se presentan los resultados obtenidos para realizar una comparativa y discusión de los mismos con otras publicaciones previas. Finalmente se presentan unas conclusiones.

En el capítulo 5 se exponen unas conclusiones finales de la tesis y líneas futuras de investigación para finalizar con los méritos de publicaciones en revistas, congresos nacionales e internacionales y premios que han derivado de esta tesis.

# Capítulo 2

# Liver segmentation in MRI: A Fully automatic Method Based on Stochastic Partitions

## Índice



[3]The content of this chapter is the author version of the following published paper: López-Mir F, Naranjo V, Angulo J, Alcañiz ML, Luna L. Liver segmentation in MRI: A fully automatic method based on stochastic partitions. Comput Methods Programs Biomed. 2014;114:11-28

## Abstract

*There are few fully automated methods for liver segmentation in magnetic resonance images (MRI) despite the benefits of this type of acquisition in comparison to other radiology techniques such as computed tomography (CT). Motivated by medical requirements, liver segmentation in MRI has been carried out. For this purpose, we present a new method for liver segmentation based on the watershed transform and stochastic partitions. The classical watershed over-segmentation is reduced using a marker-controlled algorithm. To improve accuracy of selected contours, the gradient of the original image is successfully enhanced by applying a new variant of stochastic watershed. Moreover, a final classifier is performed in order to obtain the final liver mask. Optimal parameters of the method are tuned using a training dataset and then they are applied to the rest of studies (17 datasets). The obtained results (a Jaccard coefficient of* $0{,}91 \pm 0{,}02$*) in comparison to other methods demonstrate that the new variant of stochastic watershed is a robust tool for automatic segmentation of the liver in MRI.*

## 2.1. Introduction

Fully automatic liver segmentation in medical images is currently an unsolved problem (Heimann et al., 2009). An accurate liver segmentation has a direct application in the planning, monitoring, and treatment of different types of pathologies such as cirrhosis or hepatocellular carcinoma diseases. In these cases, hepatic tissue anomalies are treated using qualitative comparison, which is related to physician experience; however, quantitative measures are not widely used. Liver segmentation is the first step to calculate objective measurements and liver/lesion ratios for decisions regarding treatment and planning for the patient. The segmentation of internal organs is also essential for image-guided surgery and virtual reality scenarios for medical training (Simpson et al., 2012; Shevchenko et al., 2010; Lim et al., 2006; Ruska et al., 2013; Lloréns et al., 2012). In addition, the liver segmentation can help in hepatic steatosis quantification because the results of this segmentation can be correlated to measure fat fractions (d'Assignies et al., 2011).

In most applications mentioned above, due to the high accuracy required, a segmentation of the liver is carried out in images with high spatial resolution, i.e., Computed Tomography (CT) or Magnetic Resonance Images (MRI) (Yang et al., 2014; Gloger et al., 2010). Currently, some efforts are focused on the segmentation of the liver in other types of images (such as PET or ultrasound images) that are less damaging for the patient than the CT images and that are cheaper for hospitals than MRI. However, the low spatial resolution of these images is a disadvantage and, in some cases, even with manual corrections, the segmentation is not accurate enough for image-guided surgery or liver volumetry applications that radiologists or surgeons require (Milko et al., 2008; Li et al., 2012). In the literature, there are more segmentation methods that are applied and validated for CT than for MRI. MRI generally has more artefact effects and a lower gradient response and is more costly for hospitals; however, since it is a non-ionizing radiation, it is less damaging for the patient in comparison

with CT. The authors of several studies support the benefits (or the additional information) of MRI considering it to be a primary diagnostic imaging modality for liver lesion detection or for measuring hepatic steatosis (d'Assignies et al., 2011; Dancygier et al., 2010; Hussain and Semelka, 2005; Balci et al., 2008; Hassan et al., 2011; Phongkitkarun et al., 2009). For example, the segmentation of the liver in MRI is important in automating liver perfusion analysis, which provides important information about the blood supply to the liver (Chen et al., 2009a). In any case, hepatic MR certainly is an alternative to CT images for the diagnosis of liver disease offering benefits that make this image technique interesting for clinical purposes. For this reason it is necessary to advance in the development of methods for liver segmentation in MRI in a way similar to the advances in CT methods.

The liver segmentation methods found in the state-of-the-art in MRI are based mainly on level-set methods (Chen et al., 2009a; Yuan et al., 2010; Cheng et al., 2008; Gloger et al., 2011; Dueñas et al., 2008; Goceri et al., 2012), where the drawbacks of these algorithms (difficult training, high computational cost, or high user iteration) are noticeable. Specially, in (Yuan et al., 2010), a level-set method (a fast marching algorithm) and fuzzy theory are applied in the liver segmentation task, but the computational cost of this algorithm needs to be improved (as the authors themselves recognize) and, additionally, non-uniform intensity problems are not solved. In (Cheng et al., 2008; Gloger et al., 2011), level-sets and probabilistic maps are used, and a training process is required with a high user iteration for manual segmentation. In (Dueñas et al., 2008), another level-set method called active contour is applied in T1 MR images of the liver, and the radiologist's knowledge is required to define the region of interest. Finally, in (Goceri et al., 2012), active contours are also applied in T2 MR images, but the results, though promising, are not accurate in some cases. Other methods based on graylevel properties (region growing, thresholding, k-means, etc) produce poor results in images of this type because the great intra-study differences make difficult the generalization of these algorithms. In (Gloger et al., 2010), authors use a region growing method combined

with threshold techniques and prior knowledge that requires a training step with manual segmentation.

In the current paper, the performance and the validation of a new liver segmentation method that is based on the watershed transform and that is applied to MRI is presented. The goal is to obtain a fully automatic method that requires less of the clinician's time, has enough accuracy and robustness for medical environments, and has a reasonable computational cost. The watershed transform is a segmentation tool that is based on graylevel and contour properties of the image. This tool extends each regional minimum of the image as far as its topography allows. An over-segmentation problem usually appears due to the large number of regional minima in the image. There have been improvements in the original watershed transform in order to reduce its drawbacks. These include using marker-controlled watershed paradigm (Meyer and Beucher, 1990) as well as hierarchical watershed paradigms such as the waterfall algorithm (Beucher, 1994). It must be specify that here the standard framework of watershed transform based on the flooding algorithm is adopted despite to there are other alternative frameworks based on a continuous formulation using topographic distance (Meyer, 1994b; Najman and Schmitt, 1994); the topological watershed based on discrete geometry tools (Couprie and Bertrand, 1997; Bertrand, 2005); graph-based watershed using minimum spanning-tree algorithms (Meyer, 1994a; Cousty et al., 2009); the power watershed algorithm (Couprie et al., 2011); the viscous watershed (Vachier and Meyer, 2005); etc.

With the marker-controlled algorithm, a set of markers imposes the new minima, and the number and position of output regions can be controlled and the over-segmentation problem is reduced. The definition of these markers is not an easy task for the segmentation of the liver, which is a large organ that has an enhanced vessel tree that produces high internal gradients. The manual definition of these markers is inefficient and is not a practical option in clinical environments due to the potential benefits of the algorithm (such as low user interaction)

decrease (Lapeer et al., 2002). To deal with problems of this kind, the use of a new variation of the stochastic transform proposed by (Angulo and Jeulin, 2007) is carried out in this paper. This variation is necessary because when the original stochastic transform is applied in MR images of the liver, it enhances internal edges with respect to the external edges of the liver, which is not useful for our purpose. The purpose of the new variant of the stochastic watershed proposed in this work is to obtain a more significant probability density function of contours by taking into account the contrast between adjacent regions thanks to a region-based model. Besides presenting this new version of stochastic watershed, another contribution of this work is the combination of pre-processing, marker extraction, and post-processing filters. This makes possible the liver segmentation of 3D studies in a fully automated and accurate way and with a low computational cost. These features convert the method into a usable tool for clinical purposes.

There have been other liver segmentation algorithms where the watershed transform has been used due to its easy user initialization/interaction, its reasonable computational cost, its intuitiveness, and the accurate results achieved in other casuistic (type of images, organs, etc). In (Masoumi et al., 2012), the watershed transform is combined with neural networks to train and tune watershed parameters for liver segmentation purposes. This training algorithm requires manual segmentation, and since the method is only used for 2D image segmentation, the particularities of a 3D volume are not taken into account. In (Kim et al., 2007), after applying a pre-processing step, the watershed transform is applied but regions with similar intensities may be incorrectly merged due to problems of intensity that are produced by lesions or by illumination that is not uniform.

In our application, the final goal of the segmentation of the liver is to add the 3D model of the liver into a 3D model with other abdominal organs of a patient previously segmented with own algorithms (López-Mir et al., 2011; López-Mir et al., 2013). This virtual 3D model will be registered and merged thanks to augmented reality algorithms by using

an image of the patient that will be taken with an external video camera. The system will be applied in order to place trocars in the patient's body in which the minimum accuracy required is approximately 2 centimetres (Fuertes et al., 2011).

The rest of the paper is divided into three sections. The first section describes a technical explanation of the watershed transform, the contribution of this work, and the final method developed for the liver segmentation. The second section explains the datasets used, the training procedure, the optimal parameters, and the results of this new method. The last section describes a discussion of our conclusions and future work.

## 2.2. Image processing methods

Let $f$ be a grayscale image defined as the mapping

$$f(\mathbf{x}) : \mathrm{E} \rightarrow \mathcal{T}, \tag{2.1}$$

where $\mathbf{x} \in \mathrm{E}$ is the pixel position in the support space of pixels E, e.g., for 2D images $\mathrm{E} \subset \mathbb{Z}^2$. In the case of valued discrete images, $\mathcal{T} = \{t_{min}, t_{min}+1, ...t_{max}\}$ represents the pixel graylevel. Typically, in 8-bit images, $t_{min} = 0$ and $t_{max} = 255$. Furthermore, let $B(\mathbf{x})$ be a subset of $\mathbb{Z}^2$ called flat structuring element (SE) that is centred at point $\mathbf{x}$ with a particular size and shape.

### 2.2.1. Watershed transform using markers

The watershed transform is a technique based on mathematical morphology for image segmentation (Beucher and Meyer, 1993). From a morphological point of view, the function $f(\mathbf{x})$ to be segmented, belonging to $\mathbb{Z}^2 \times \mathbb{Z}$, can be seen as a topographic surface: the lighter the gray value of $f$ at point $\mathbf{x}$, the higher the altitude of the corresponding point $\{\mathbf{x}, f(\mathbf{x})\}$ on the surface. Following this viewpoint, each regional minimum of $f(\mathbf{x})$ represents a region called catchment basin. The

purpose of the watershed transform is to flood these basins in order to increase the water level. When two neighbouring basins come into contact with each other, a dam is built at these contact points to prevent the merging between two different basins. At the end, the union of all complete dams constitute the watershed lines, which separates the lakes (catchment basins associated to each minimum). From a practical viewpoint, watershed transform is computed using fast algorithms based on hierarchical queues (Beucher and Meyer, 1993; Vincent, 1991). A review of watershed algorithms can be found in (Roerdink and Meijster, 2000). Typically for image segmentation, the gradient of the image is considered as the input image for the watershed transform:

$$f \mapsto \rho(f) \mapsto WS(\rho(f)),$$

where the image $WS(\rho(f))$ corresponds to the set of closed contours of the $N$ segmented regions and $\rho(f)$ is the morphological gradient (Soille, 2003). The number of regions $N$ equals the number of minima of the gradient image.

The result of the watershed can be also considered by a dual representation as an image partition of the support space E into $N$ connected classes $\mathcal{C}_n \subset \mathrm{E}$, denoted $\Pi(WS(\rho(f)))$, such that

$$\begin{aligned} &\Pi(WS(\rho(f))) = \{\mathcal{C}_n\}_{1 \leq n \leq N}\,; \\ &\textit{with } \bigcup_{1 \leq n \leq N} \mathcal{C}_n = \mathrm{E} \text{ and } \mathcal{C}_p \cap \mathcal{C}_q = \emptyset, \forall p \neq q. \end{aligned} \tag{2.2}$$

We notice that it is needed that each arc (watershed line) which separates two regions (catchment basin) belongs to one of them. It is well known that without any pre-processing, watershed transform usually leads to an over-segmentation of the image gradient because of the presence of a great number of regional minima due to the noise and the different structures present in images of this type. The classical approach for solving this problem is known as the marker-controlled watershed (Beucher and Meyer, 1993) which involves a reconstruction

by erosion ($R^{\varepsilon}$) (Soille, 2003) of the gradient image $\rho(f)$ by a marker image $m(\mathbf{x})$:

$$WS(\rho(f))_{m(\mathbf{x})} = WS(R^{\varepsilon}_{(\rho(f)\wedge m(\mathbf{x}))}(m(\mathbf{x}))), \tag{2.3}$$

where $m(\mathbf{x})$ is a set of prior markers:

$$m(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} \textit{ belongs to a marker}, \\ 255, & \textit{otherwise}, \end{cases} \tag{2.4}$$

such that the operation $R^{\varepsilon}_{(\rho\wedge m(\mathbf{x}))}(m(\mathbf{x}))$ imposes the new regional minima of the $f$ gradient, $\rho(f)$, to the pixels set to zero in $m(\mathbf{x})$. Thus, the number of obtained regions equals now the number of connected components of marker image $m(\mathbf{x})$.

### 2.2.2. Stochastic watershed transform

An appropriate marker definition is essential to obtain a good segmentation result with the marker-controlled watershed. In the stochastic watershed a different strategy is followed due to a stochastic procedure when markers are randomly generated (Angulo and Jeulin, 2007). This arbitrary choice is compensated by the $M$ marked-controlled watershed realizations of the algorithm in which non-significant fluctuations are filtered out by the stochastic procedure.

More precisely, stochastic watershed is based on applying the marker-controlled watershed $M$ times to the gradient of the input image, but the markers change randomly in each iteration. Let $\{m_j(\mathbf{x})\}_{j=1}^{M}$ be $M$ sets of $N$ random markers, and let $WS_j = WS(\rho(f))_{m_j(\mathbf{x})}$ be the j-th output image of the marker-controlled watershed imposed by $m_j(\mathbf{x})$. Using this approach, a probability density function of the contours of the image can be obtained with the Parzen method (Duda and Hart, 1973) as follows:

$$\text{pdf}(\mathbf{x}) = \frac{1}{M}\sum_{j=0}^{M} WS_j(\mathbf{x}) \ * \ G(\mathbf{x}; s), \tag{2.5}$$

where $G(\mathbf{x}; s)$ represents typically a Gaussian kernel of variance $s^2$ (in our case, $s = 3$) and mean $\bar{m}$ ($\bar{m} = 0$), that is defined as:

$$G(\mathbf{x}; s) = \frac{1}{2\pi s^2} e^{-(\frac{\|\mathbf{x}\|^2}{2s^2})}. \tag{2.6}$$

Besides the pairs of parameters (M, N), the other variable that can affect the final result of the stochastic watershed algorithm is the procedure for generating the random markers. In the case of random uniform markers, as introduced in (Meyer and Stawiaski, 2010), the probability density function of contours can be explicitly computed without using M simulation of markers and corresponding watershed transforms.

### Procedure for generating random markers

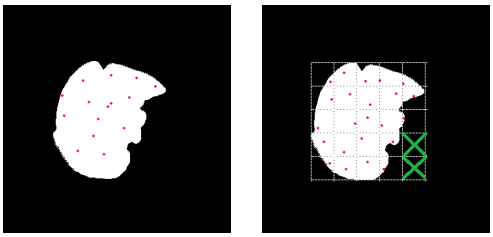
In the basic formulation of stochastic watershed, a uniform distribution in the whole image is used in the marker definition process. Let us consider the problem of liver segmentation tackled in this paper.

Differences in liver regions between adjacent slices are limited and this is the rationale behind our propagative approach of 2D segmentation. If a good segmentation in one slice is obtained, the resulting mask can be dilated and the perimeter of this dilation can be considered as external marker for the adjacent image (on axial axis) with the certain that all the liver regions will be inside of this external marker. It will help to eliminate adjacent structures close to the liver and with similar graylevel. The random markers (or internal markers) will be generated with a uniform distribution under the constraint that they will be inside of the region defined by the previous mask (Fig. 2.1a). Thus it is guaranteed that the most of the markers will belong to liver regions and the stochastic watershed will not enhance edges of other structures, because no marker will be inside of them.

In the method proposed in this paper, another strategy for generating random markers is also considered in the last marker-controlled watershed applied which will produce the final mask of the

liver. The bounding box of the previous mask is divided into a uniform grid, and a random marker is calculated in each region of this grid (Fig. 2.1b). If the mask is not active (pixels set to zero) in a region of the grid, no marker is generated. This generation procedure will be called as stratified random marker generation. The under-segmentation of neighbourhood regions with similar graylevel is reduced with this variation because markers are better distributed but internal boundaries can be enhanced with this second strategy (because the arbitrary choice is biased). For this reason a uniform-distributed random markers is used first in the generation of the pdf($\mathbf{x}$) as it will be explained in section 2.2.4 and a stratified random markers is used finally in the last iteration of the algorithm when internal gradients are less enhanced.

Both strategies proposed in this paper (uniform and stratified random marker distribution) were applied on the previous mask. The different procedures from the original stochastic watershed (Angulo and Jeulin, 2007) were also considered, but the novel procedures proposed in this work are better adapted to the liver segmentation problem.



a) b)

**Figure 2.1:** *a) Uniform-distributed random markers; b) stratified random markers (red) and grid zones in which no internal marker is calculated (green).*

### Probabilistic partitions from stochastic watershed (PPSW)

A new paradigm of the stochastic watershed is introduced here to produce a more reliable density of contours. This new stochastic wa-

tershed version is needed because using classical stochastic watershed, the liver contour is relatively well defined (Fig. 2.2c), but the internal boundaries are also enhanced on account of hepatic tree or illumination differences. This is justified since the probability density function (Eq. 2.5) is extracted directly from the gradient image, which not only contains useful information about the liver edges. This is the reason why the stochastic watershed is not enough selective for the liver segmentation purpose. In order to remove high internal gradients and to keep the external boundaries of the liver enhanced enough, we initially consider the approach suggested in (Angulo and Jeulin, 2007) to compute a gradient-like function defined by a trade-off (a convex combination) between the initial gradient and the probability density of contours, i.e.,

$$\bar{h}_i(\mathbf{x}) = \alpha\rho(f)_i(\mathbf{x}) + (1-\alpha)\,\mathrm{pdf}_i(\mathbf{x}) \tag{2.7}$$

where typically $\alpha = 0{,}5$. However, as can be observed in Fig. 2.2d, the internal boundaries are still enhanced (with high internal boundaries belonging to adjacent regions with poor contrast). Therefore, a new model based on regional graylevel properties for reducing internal gradients, and keeping the liver edges enhanced, is introduced in the stochastic procedure. The idea is based on the evidence that internal gradients of the original image are not as enhanced as the stochastic watershed produces, so if this regional information of internal gradients can be added to the stochastic watershed procedure, the probability density function will enhance liver edges for segmentation purpose.

According to our approach, $\mathrm{pdf}_i(\mathbf{x})$ gives for each pixel $\mathbf{x}$ the probability density estimated from the gradient image which only measures the local energy of the contours. We initially need to calculate from the $\mathrm{pdf}_i(\mathbf{x})$ a watershed transform to obtain an image partition denoted $\Pi_i$, i.e.,

$$\mathrm{pdf}_i(\mathbf{x}) \mapsto WS(\mathrm{pdf}_i)(\mathbf{x}) \mapsto \Pi_i = \Pi(WS(\mathrm{pdf}_i))$$

Each $n$-connected class $\mathcal{C}_{i,n}$ of the partition $\Pi_i$ is then valued with the mean intensity of pixels belonging to this class from the initial image

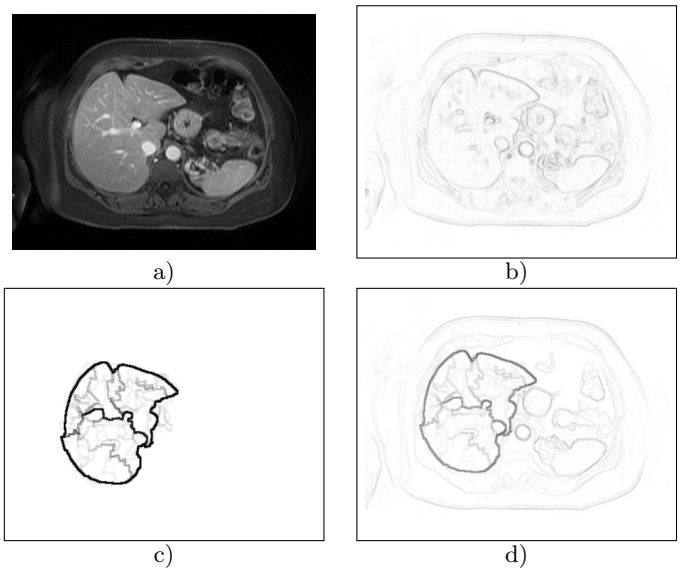



**Figure 2.2:** *a) Original image; b) original image gradient; c) probability density function of image contours* ($\mathrm{pdf}_i(\mathbf{x})$) *with M=20 and N=15; d) probabilistic gradient as proposed in (Angulo and Jeulin, 2007),* $\overline{h}_i(\mathbf{x})$.

$f_i$ in order to construct a "mosaic image", denoted $p_i(\mathbf{x})$, and formally defined as

$$p_i(\mathbf{x}) = \{\mu_{\mathcal{C}_{i,n}}(f_i) : \ \mathbf{x} \in \mathcal{C}_{i,n}\} \tag{2.8}$$

where $\mu_{\mathcal{C}_{i,n}}(f_i) = 1/|\mathcal{C}_{i,n}| \sum_{\mathbf{y}\in\mathcal{C}_{i,n}} f_i(\mathbf{y})$ is the average of pixel values of slice $f_i$ in the connected class $\mathcal{C}_{i,n}$. The morphological gradient of this mean-based simplified model of the image can be interpreted as the regional edgeness energy associated to the probability density function of contours $\mathrm{pdf}_i(\mathbf{x})$, i.e.,

$$\mathcal{E}_i^{regional}(\mathbf{x}) = \rho(p_i(\mathbf{x}))$$

After normalizing this energy to the interval $[0, 1]$ both terms can be multiplied, which leads to a regional-adjusted probability density of

contours

$$\widetilde{\mathrm{pdf}}_i(\mathbf{x}) = \mathrm{pdf}_i(\mathbf{x}) \cdot \mathcal{E}_i^{regional}(\mathbf{x}) \tag{2.9}$$

Nevertheless, we have observed that the results can still be improved by introducing the regional edgeness energy term in each of the $M$ realizations of the stochastic watershed. That is, by integrating the gradient of the mean-based simplified model of each realization instead of directly the watershed lines. Formally, we have

$$\mathrm{pdf}_i^{regional}(\mathbf{x}) = \frac{1}{M} \sum_{j=0}^{M} \rho(p_{i,j}(\mathbf{x})) \ * \ G(\mathbf{x}; s),$$

where $p_{i,j}(\mathbf{x})$ is the mosaic image of realization $j$ obtained as:

$$p_{i,j}(\mathbf{x}) = \{\mu_{\mathcal{C}_{i,j,n}}(f_i) : \ \mathbf{x} \in \mathcal{C}_{i,j,n}\}$$

from the connected classes $\mathcal{C}_{i,j,n}$ of the partition $\Pi_{i,j}$ obtained from:

$$\Pi_{i,j} = \Pi_i \left(WS_j(h_i)\right)$$

Finally, the probabilistic edgeness function for slice $i$ is defined as the image:

$$e_i(\mathbf{x}) = 0{,}5\rho(f)(\mathbf{x}) + 0{,}5\,\mathrm{pdf}_i^{regional}(\mathbf{x}). \tag{2.10}$$

We note that, using the initial probability density function of contours $\mathrm{pdf}_i(\mathbf{x})$, each boundary of the output image $WS_j$ has the same weight during stochastic integration, Fig. 2.3a. With the first proposed modification $\widetilde{\mathrm{pdf}}_i(\mathbf{x})$, the boundaries associated to a low contrasted regions (internal zones of the liver) are less enhanced since their weights are associated to the contrast. However, the effect is still more notable in the case of the final probabilistic edgeness function $e_i(\mathbf{x})$, see Fig. 2.3c and Fig. 2.3d.

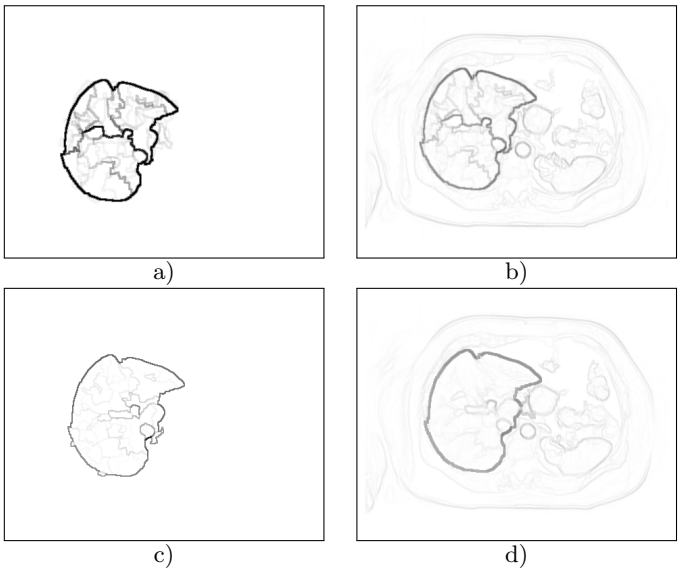



**Figure 2.3:** *a) Probability density function of image contours* ($\mathrm{pdf}_i(\mathbf{x})$) *with N=20; b) probabilistic gradient as proposed in (Angulo and Jeulin, 2007),* $\overline{h}_i(\mathbf{x})$*; c) regional edgeness energy term,* $\mathrm{pdf}_i^{regional}(\mathbf{x})$*; d) final probabilistic edgeness introduced in this paper,* $e_i(\mathbf{x})$.

### 2.2.3. Segmentation algorithm for initial slice

The first processed slice must undergo a special processing. It is important to carry out a precise segmentation of this first slice because its result will be expanded to adjacent slice in order to calculate parameters as explained in section 2.2.2. Manual segmentation is an option for this first slice, but an automatic process based on the marker-controlled watershed has been developed in our case, which is summarized in Fig. 2.4.

In this initial 2D slice of the 3D image, the liver should appear as large as possible and with a homogeneous graylevel. This slice could be selected manually although in this paper an automatic process to this

selection based on the previous assumption on liver features is proposed. First, the graylevel histogram of all the voxels of the whole study is calculated. If the pixels near to zero are excluded, the maximum of the histogram is associated to the liver regions because the liver is the largest organ in the abdominal cavity. Finally, the slice with more pixels whose graylevel is equal to this maximum is selected as the initial slice of the study. With this procedure, in all the processed datasets the initial slice selected was appropriate and fulfilled initial restrictions (the liver appears as large as possible and with homogeneous graylevel).

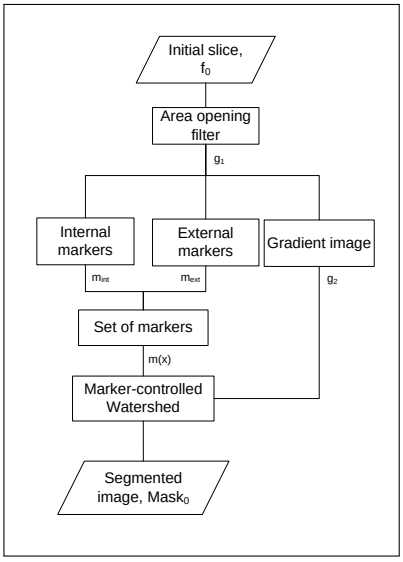



**Figure 2.4:** *Diagram block of the segmentation procedure of the first slice in the study.*

Let $f_0$ be the initial 2D slice to be segment and the goal is obtained a set of markers and a reasonable enhanced gradient for applying the marker-controlled watershed transform. The purpose of the first step is to obtain a well-defined gradient of the liver. If the gradient is calculated on the original slice, i.e., $g_0 = \rho(f_0)$, our algorithm does not produce the required accuracy in many cases (Fig. 2.5b). For this reason, an area opening filter is applied to the original slice, resulting in

$g_1 = \gamma_\lambda(f_0)$ (Soille, 2003). This filter reduces the noise and local bright structures (vessels) while the boundaries of the liver are preserved (Fig. 2.5c).parameter $\lambda$ of this filter corresponds to the area measure. Local bright structures (in our case vessels) with lower area than $\lambda$ will be integrated into local background structures (in our case, liver tissue). The gradient is calculated on the filtered image, $g_2 = \rho(g_1)$, producing an image where the contour of the liver appears more enhanced and less noisy than in $g_0$ (Fig. 2.5d).

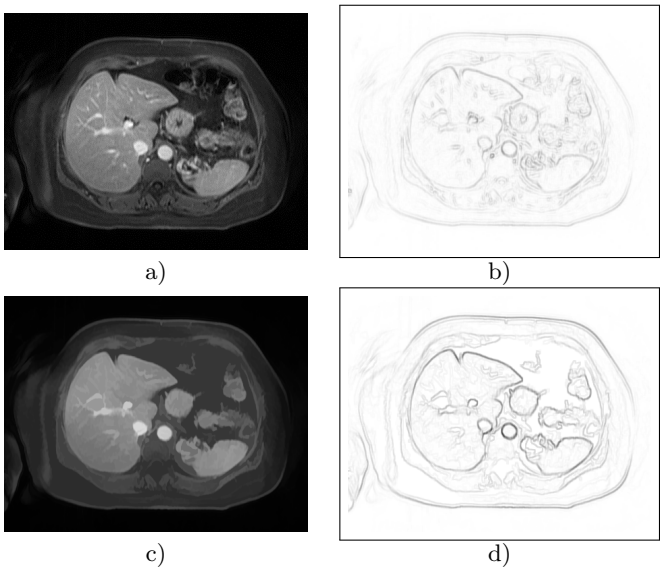


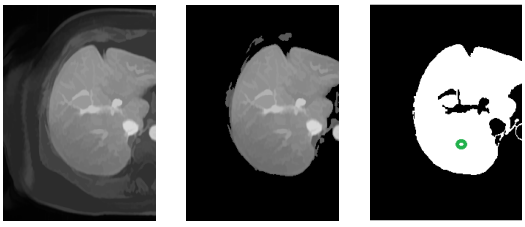

**Figure 2.5:** *a) Original slice, $f_0$; b) gradient image of original slice, $g_0$; c) area opening filtered image, $g_1$; d) gradient of area opening filter output, $g_2$.*

In the next step, before applying the marker-controlled watershed transform, the set of markers is calculated using the output image of the area opening filter as the input of the process. The marker definition not only requires internal markers (that mark the object of interest) but also external markers that are used to constrain the growth of the regions. In this first slice, an external marker image $m_{ext}$ is obtained as

the boundary of the patient's abdomen as previously was proposed in (Campadelli et al., 2009).

Internal marker image ($m_{int}$) is obtained from the output image of the area opening filter, $g_1$ applying some anatomical restrictions. First, it is sure that the liver will appear on the right part of the image and it will be inside of external marker image, so a sub-image can be calculated (Fig. 2.6a), $g_2$. It is well known that the liver is a structure with higher intensities than other structures in MR images (T1 or T2 weighting) as skin, air or bones. If the accumulated histogram in the complement of this sub-image is calculated, a threshold can be obtained when the 50 % of the pixels are accumulated, to obtain the graylevels belonging to the liver. The pixels with a graylevel smaller than this threshold are removed, obtaining image $g_3$ (Fig. 2.6b).

In a second step, in order to remove internal structures (vessels, gallbladder, lesions, etc.) the 10 % of lightest and darkest pixels of $g_3$ are deleted and the image is binarized (Fig. 2.6c). The geodesic distance is calculated and the pixel with the maximum graylevel value is selected as input marker (Fig. 2.6c, in green).



a) b) c)

**Figure 2.6:** *a) Sub-image inside the external marker, $m_{ext}$ (red); b) sub-image after applying the threshold of the accumulated histogram; c) binarized sub-image after removing the 10 % of highest and darkest pixels (the circle in green is the final internal marker selected as the maximun of the geodesic distance).*

For watershed segmentation, the final marker image $m(\mathbf{x})$ is given

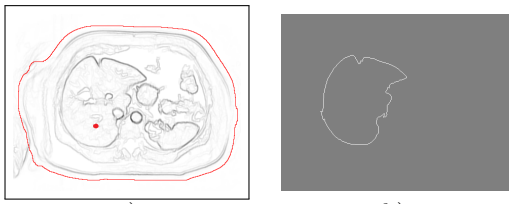
by the union of the internal and the external markers:

$$m(\mathbf{x}) = m_{ext}(\mathbf{x}) \cup m_{int}(\mathbf{x}). \tag{2.11}$$

Finally, with this marker image $m(\mathbf{x})$, the marker-controlled watershed transform is applied to image gradient $g_2$ (Fig. 2.7a). Therefore, the output image of the segmentation process of $f_0$ is obtained as:

$$Mask_0(\mathbf{x}) = WS(g_2(\mathbf{x}))_{m(\mathbf{x})} \tag{2.12}$$

The mask of the liver at this slice is just $Mask_0$ (Fig. 2.7b).



a) b)

**Figure 2.7:** *a) Gradient $g_2$ and set of markers (red) $m(\mathbf{x})$ for initial slice; b) result of marker-controlled watershed transform, $Mask_0$.*

### 2.2.4. Main segmentation algorithm

After the segmentation of the first slice, the different steps presented in Fig. 2.8 are followed to segment the liver into the rest of the slices of the 3D volume. The current slice and the mask of the previous slice are required as input information. The mask of the initial slice is calculated as was just explained.

The goal of the different steps of the process is the same as for the first slice: to obtain an appropriate set of markers and a well-defined image to be flooded by the watershed. The algorithm takes advantage of intrinsic parameters of the MR images: the homogeneity in graylevel between neighbouring pixels and the relative co-location of the liver between adjacent slices.

The disadvantages of using directly the marker-controlled watershed transform and the stochastic watershed transform were discussed in Section 2.1. The algorithm explained for the initial slice not always works correctly in all the images. For that reason, a new paradigm of stochastic watershed introduced in section 2.2.2 is used.

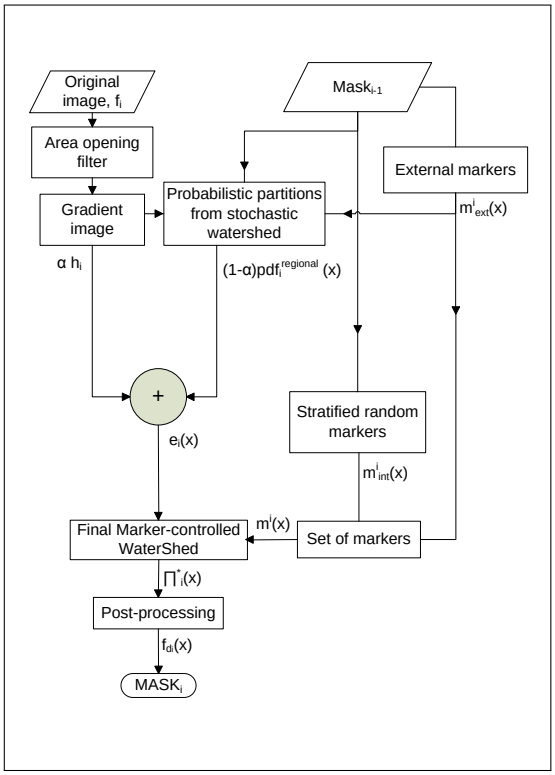



**Figure 2.8:** *Block diagram of main algorithm.*

Let $f_i$ be the i-th slice of the whole study, $i \in \{-R_n, ..., -1, 1, ...R_k\}$ and let $Q = R_n + R_k + 1$ be the number of the slices of the study. Moreover, let $Mask_{i-1}$ be the liver mask of the previously segmented image on the axial axis. First, an area opening filter is applied, then the gradient of the output image of this slice is calculated, in order to

obtain a liver contour image that is more enhanced and less noisy, i.e.,

$$h_i = \rho\left(\gamma_\lambda(f_i)\right).$$

The new version of the stochastic watershed explained in section 2.2.2 is used to reduce internal gradients. The external marker image $m^i_{ext}$ of each slice corresponds to the external perimeter of a dilation of $Mask_{i-1}$ as it was explained in the section 2.2.2, i.e.,

$$m^i_{ext} = \delta_{B_1 \oplus B_1}(Mask_{i-1}) - \delta_{B_1}(Mask_{i-1}).$$

With this external marker image $m^i_{ext}$ and M uniform-distributed realizations of N random markers $m^i_{int_j}$, it is obtained $m^i_j = m^i_{int_j} \cup m^i_{ext}$. These markers are used to calculate the $\text{pdf}_i^{regional}(\mathbf{x})$ of the image contours of slice $i$ and consequently the probabilistic edgeness function, $e_i(\mathbf{x})$ defined in Eq. 2.10.

In the next stage, a final marker-controlled watershed transform is applied to the probabilistic edgeness function. The external marker image is computed as in the previous step: $m^i_{ext} = \delta_{B_1 \oplus B_1}(Mask_{i-1}) - \delta_{B_1}(Mask_{i-1})$. The internal marker image $m^i_{int}$ is calculated with a stratified random markers procedure ($7 \times 7$ grid) using the previous result mask, $Mask_{i-1}$ as it was explained in section 2.2.2. In each region of the grid, an internal random marker is obtained in order to have at least one marker per liver region, which is essential when the liver is broken into several parts (Fig. 2.1b). A final set of markers is obtained by the union of both images: $m_i(\mathbf{x}) = m^i_{ext}(\mathbf{x}) \cup m^i_{int}(\mathbf{x})$. Then, a final marker-controlled watershed is applied to obtain a final partition of slice $i$:

$$\Pi^*_i(\mathbf{x}) = WS(e_i(\mathbf{x}))_{m_i(\mathbf{x})} \tag{2.13}$$

At this point, the problem of high internal gradients and the under-segmentation produced if the liver is divided in various regions are solved. However the problem of illumination is only partially solved with the area opening filter (reducing light objects). Hence, a post-processing

algorithm is applied to finally correct potential minor segmentation errors.

### Post-processing Classifier

Let us consider as initial image the mean-valued mosaic image of the partition associated to $\Pi_i^*(\mathbf{x})$ as it was defined in Eq. 2.8,

$$\Pi_i^*(\mathbf{x}) \mapsto p_i^*(\mathbf{x}).$$

In this mosaic image, the regions belonging to the liver will typically have similar intensity values (under the assumption that the graylevel intensities of the liver have a relatively low variance). However, due to illumination problems, regions belonging to the liver with different graylevel might appear (and the graylevel variance of the liver regions might increase). Therefore, a threshold in the $p_i^*(\mathbf{x})$ image is not enough to obtain the final mask of the liver. For this reason, a more sophisticated algorithm is needed to minimize the graylevel variance of the regions belonging to the liver while it maximizes the graylevel variance between these regions and external regions that are not liver. This algorithm is detailed in Algorithm 2.1:

**Algorithm 2.1:** Intra-region variance correction.

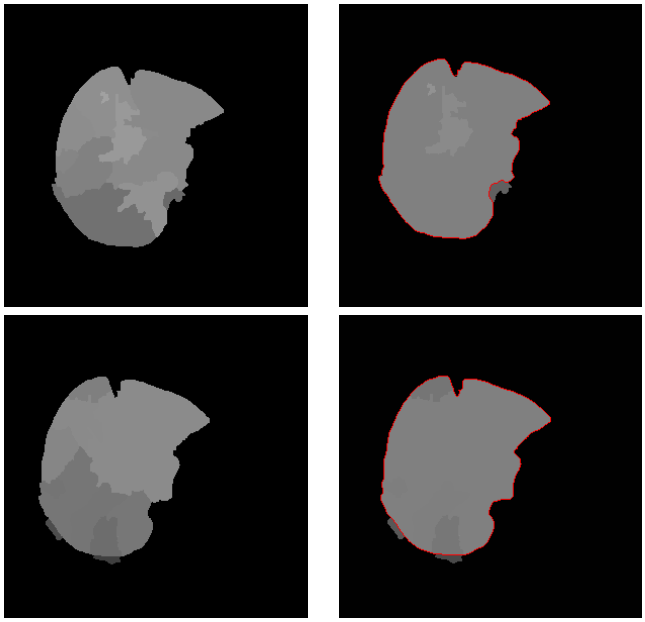
$$f_{a_i} = R^{\delta}_{p_i^*}\left(\delta_{B_2}\left(p_i^*\right)\right)$$
$$f_{b_i} = R^{\varepsilon}_{f_{a_i}}\left(\varepsilon_{B_3}(f_{a_i})\right)$$
$$f_{c_i} = p_i^* - f_{b_i}$$

This algorithm has 3 steps: The reconstruction by dilation will reduce light local zones and the reconstruction by erosion will darken local zones with dark neighbour zones (typically these zones are in the external boundary of the liver). The result of this algorithm is shown in Fig. 2.9b. An intensity-based classifier is then applied to the image

$f_{c_i}(\mathbf{x})$ as follows

$$f_{d_i}(\mathbf{x}) = \begin{cases} 1, & \text{if } f_{c_i}(\mathbf{x}) \in [t_1, t_2], \\ 1, & \text{if } f_{c_i}(\mathbf{x}) \in [t_L, t_1], [t_2, t_U] \\ & \& \; OV > 50\,\%, \\ 0, & \textit{otherwise}. \end{cases} \tag{2.14}$$

Thresholds $[t_L, t_U, t_1, t_2]$ are obtained empirically in the training dataset and $OV > 50\,\%$ means that one of the regions between $[t_L, t_1]$ or $[t_2, t_U]$ will be added to the mask of the liver if the overlap with the mask of the adjacent slice is more than $50\,\%$.



a) b)

**Figure 2.9:** *a)* $p_i^*(\mathbf{x})$ *and* $p_{i+1}^*(\mathbf{x})$ *images; b)* $f_{c_i}$ *and* $f_{c_{i+1}}$ *images (in red,* $f_{d_i}$ *and* $f_{d_{i+1}}$ *contours).*

Finally, the mask that is selected as liver is filtered with the Fourier coefficient technique (Zahn and Roskies, 1972) as previously proposed

in (Gloger et al., 2010) to smooth the contour, obtaining the final mask of the liver for slice $i$, $Mask_i(\mathbf{x})$.

According to our datasets, a criterion for stopping our algorithm is selected. The algorithm ends when the overlap with the adjacent liver mask is inferior to 50 % or when the mask area is inferior to 3.5 $cm^2$. If the inter-slice distance decreases, the overlap percentage could be more restrictive.

## 2.3. Results

### 2.3.1. Dataset information and training procedure

The segmentation method presented in this paper has been evaluated in 17 contrast-enhanced MR studies from *Hospital Clínica Benidorm*. Table 2.1 summarizes the technical parameters of all datasets and pathological information. The x/y/z parameter is the voxel spacing; the w/h/slices are the number of pixels of the series. The last five series (ID 12-16) were acquired with a Philips Achieva machine (3.0 Tesla), the other datasets with a GE Signa HDX machine (1.5 Tesla). The series with ID: 0-11 were a dynamic sequence acquired when the artery or vein received the contrast agent and the series with ID: 12-16 were a late sequence; that is, they were used to observe the elimination of contrast in lesions.

Dataset 0 was used for tuning the algorithm. This dataset has a combination of the most common problems that are presents in the other studies. Fig. 2.10 shows some images of this initial dataset where these problems are observable: the vessel enhancement (Fig. 2.10c), neighbourhood structures with similar graylevel to the liver (Fig. 2.10a,b), the inter-slice non-uniform illumination (Fig. 2.10c), and liver visualization into several parts (Fig. 2.10d).

All the parameters of our method were adjusted with this initial dataset. Empirically, these parameters were changed and different

**Table 2.1:** *Technical and pathological information of the different datasets.*

| ID | x/y-z (mm) | (w/h) x slices | Pathological information |
|---|---|---|---|
| 0(train.) | 0.93-3 | 512x76 | Healthy |
| 1-7(test) | [0.78-1.9, 0.88-3] | 512x59-120 | Healthy |
| 8(test) | 0.86-2 | 512x104 | Metastasis, seg. I |
| 9(test) | 0.82-2 | 512x104 | Hepatocellular carcinoma, left lobe |
| 10(test) | 0.85-1.9 | 512x107 | multiple metastasis |
| 11(test) | 0.88-2 | 512x88 | Metastasis, right lobe |
| 12(test) | 0.71-2.5 | 528x80 | Hepatocel. carcinoma, 2 lesions seg. VIII |
| 13(test) | 0.64-2.5 | 640x80 | Metastasis, seg. V, IVa, VIII |
| 14(test) | 0.71-2.5 | 528x80 | Focal nodular hyperplasia, seg. VII, VIII |
| 15(test) | 0.71-2.5 | 528x80 | Liver hemangioma |
| 16(test) | 0.71-2.5 | 528x80 | Metastasis, seg. V |

coefficients were calculated. For this purpose, a manual segmentation of this training dataset was carried out by radiologist experts. The final parameters are stored when the coefficients are good enough, and then the rest of datasets are segmented with these parameters. The coefficients and the manual segmentation protocol are the same used in the test dataset.

### 2.3.2. Test Dataset and final parameters

The performance of our segmentation method was evaluated comparing the automatic liver mask produced by our algorithm with a manual segmentation carried out by four experts (each expert segmented four different series). Moreover, two series of the dataset were selected to be segmented by all them in order to calculate the Inter-expert Cross Correlation coefficient (ICC), obtaining an $ICC = 0{,}622$. This result proves the goodness of this manual segmentation because

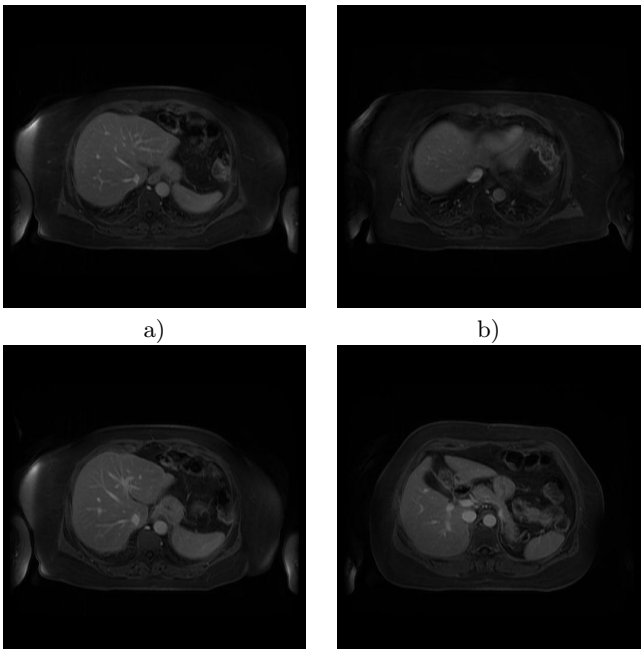



**Figure 2.10:** *Different problems of the training dataset.*

a value greater than 0.61 is considered to be sufficient in order to state that personal differences between the manual segmentation of the experts are not relevant (Kramer and Feinstien, 1981).

The coefficients used for the algorithm validation were the Jaccard index and the Hausdorff distance. These coefficients are the most significant for volume comparison and they give complementary information (Lago et al., 2012; Martínez-Martínez et al., 2013b). Additionally, the Dice Coefficient (DC) is provided for literature comparison (although it provides similar information to Jaccard index). If $X$ is the reference mask and $Y$ is the resulting mask of our segmentation method, these coefficients are defined as:

$$JC = \frac{|X \cap Y|}{|X \cup Y|}$$
$$DC = \frac{2*|X \cap Y|}{|X|+|Y|} \tag{2.15}$$

The Hausdorff distance is defined as Eq. 2.16.

$$d_H = \text{máx}\left\{ \sup_{x \in X} \inf_{y \in Y} d_{Euclid}(x,y), \sup_{y \in Y} \inf_{x \in X} d_{Euclid}(x,y) \right\}, \tag{2.16}$$

where $d_{Euclid}$ is the Euclidean distance and (x,y) are two points of both contours (X,Y): the manual and the automatic segmentation. A $JC = 1$, (or a $DC = 1$) or a $d_H = 0$ means that the segmentation is perfect.

The selected parameters (Table 2.2) were calculated empirically as the optimal parameters for a training dataset and applied to the other studies. The value of $\lambda$ is associated with the minimum area of the liver to be detected and with the internal vessels that the area opening filter integrates in the background. If $\lambda$ increases, vessel contrast is reduced but minimum area to be detected increases. The structuring element used to obtain the external markers is related to the inter-slice distance (a greater distance produces more discontinuities between adjacent slices, so a greater dilation is required to select the liver safely).

The pair of stochastic watershed parameters (M, N) were chosen based on the experience of (Angulo and Jeulin, 2007) and heuristic tests in our training dataset (Fig. 2.12 and 2.11). With few realizations ($M = 20$), the pdf($\mathbf{x}$) (and the pdf$^{regional}$($\mathbf{x}$)) converges to a stable image (Fig. 2.11a). The parameter $M$ can be increased ($M = 100$) but the improvement of the results is not significant in this case and computational cost increases considerably (Fig. 2.11c). The parameter $N$ has a relation to $M$. A high value of $N$ produces many interior regions (Fig. 2.12c), which is very useful for segmenting a large number of different regions. However, if the goal is the segmentation of the

liver tissue (a large organ), the optimal parameter $N$ should be a low value. After different experiments, $N = 15$ is applied to all the datasets to ensure the selection of all the liver regions (even if there are not connected (Fig. 2.12a). If parameters $N$ or $M$ decrease, problems related to gradients definitions could appear as it is appreciated in Fig. 2.11b and 2.12b.

The pair of parameters presented in table 2.2 are very robust against the type of image (only restrictions of size or resolution would be important) for this reason it can be considered as constant values.

**Table 2.2:** *Parameters and optimal values.*

| Filter | Parameters |
|---|---|
| Area opening filter | $\lambda = 0{,}72\ cm^2$ |
| Dilation for external markers | $B_1$: Circular SE, size = 6.8 $cm^2$ |
| N | 15 |
| M | 20 |

In the case of the structuring elements $B_2$ and $B_3$ were selected for reducing bright and dark objects in the final step of our algorithm. Classifier thresholds were selected with the normalized histogram analysis of a set of images, $f_{c_i}$, belonging to different series (some of them are shown in Fig. 2.15c). This strategy to learn the optimal parameter is justified by the fact that graylevel intensities have a tendency to change between datasets. The global histogram of this set of images is depicted in Fig. 2.13. Areas between $t_1$ and $t_2$ fit with liver zones. The areas outside these thresholds cannot always be considered as liver but it is sure that the zones between $[t_1,t_L]$ and $[t_2,t_U]$ are in the external part of the liver tissue. For this reason, an overlap criterion is used to decide if each of these regions belongs to the liver or not. If it has a 50 % overlap with the previous segmented mask it is considered as liver, otherwise it is labelled as background. Table 2.3 summarizes the final thresholds of our algorithm applied in all the datasets processed.

Using the values of tables 2.2 and 2.3, the segmentation was carried

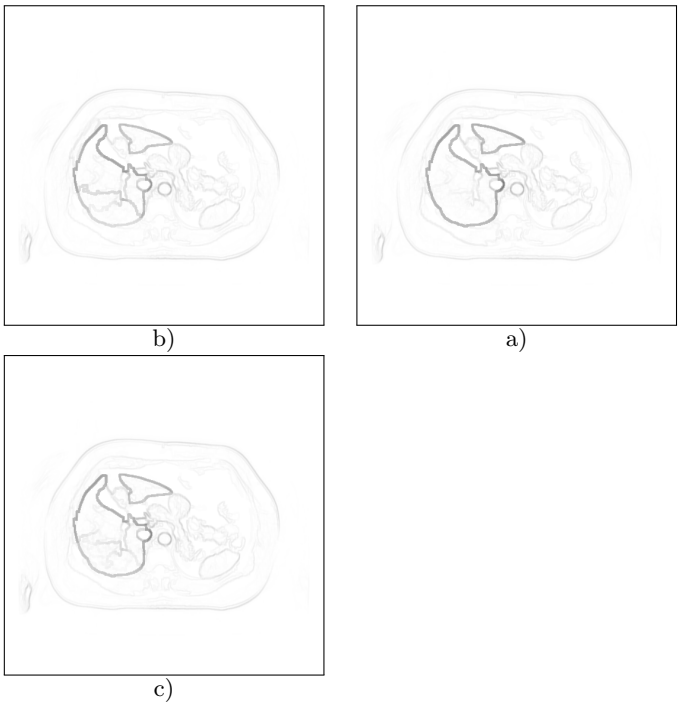



**Figure 2.11:** pdf(**x**) *using different values of parameter M (N = 15): a) M=5; b) M=20; c) M=100.*

out in the 17 datasets (training and test datasets) by the method proposed in this paper.

The values of the validation coefficients are presented in Table 2.4. All the results were calculated on an Intel core i5 @ 2.80 GHz, with a RAM of 2 GHz and Windows 7 (32 bits). The computational cost was about 7 seconds per slice so 7-12 minutes are needed for the segmentation of one dataset.

Fig. 2.14 shows the segmentation of some initial slices applying the algorithm explained in section 2.2.3 to different datasets. Several

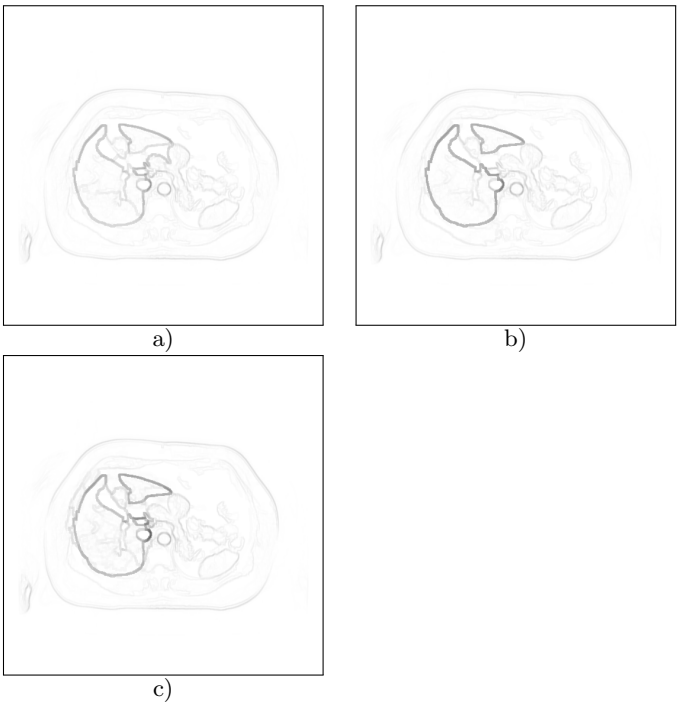



**Figure 2.12:** pdf(**x**) *using different values of parameter N (M = 20): a) N=5; b) N=15; c) N=100.*

intermediate images of healthy datasets for the process explained in section 2.2.4 are shown in Fig. 2.15: the output of the area opening filter ($\gamma_\lambda(f_i)$), the probabilistic edgeness function($e_i(\mathbf{x})$), the illumination-corrected image ($f_{c_i}$), the initial image ($f_i$) and the final mask of the liver ($Mask_i$) (in red), and the 3D liver model obtained for some studies. Fig. 2.16 shows the resulting segmentation of some images applying the algorithm explained in section 2.2.4 to different unhealthy datasets.

**Table 2.3:** *Parameters and optimal values of the post-processing step.*

| | |
|---|---|
| Reconstruction by dilation | $B_2$: Circular SE, size = 10.7 $cm^2$ |
| Reconstruction by erosion | $B_3$: Circular SE, size = 24 $cm^2$ |
| Classifier Thresholds | |
| $(t_1 - t_2)$ | [100-160] Liver |
| $(t_L - t_1,\ t_2 - t_U)$ | [70-100,160-190] Overlap criteria |
| $(0 - t_L,\ t_U - 255)$ | [0-70,190-255] Not Liver |

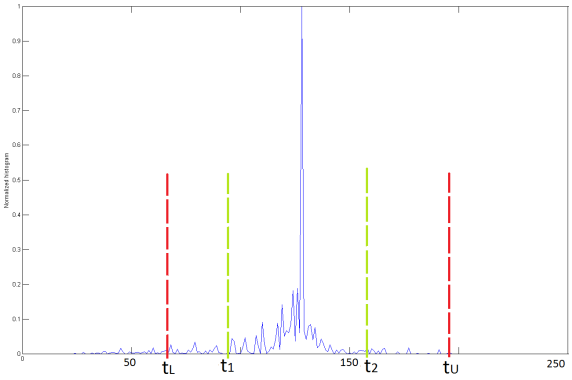



**Figure 2.13:** *Normalized histogram of the set of images for obtaining optimal thresholds.*

## 2.4. Discussion and Conclusions

This paper presents a new fully automatic method for liver segmentation in MR images for clinical use. First, a new input image of the watershed transform (instead of the classical gradient) was calculated with a new variant of the original stochastic watershed in order to obtain a probabilistic edgeness function more robust and less noisy than the original gradient. A final marker-controlled watershed was calculated with a set of markers to obtain a final partition (or "mosaic image") of the liver. Finally, based on the liver graylevel properties and the inter-slice co-location of this organ, a classifier was implemented to obtain the final mask.

**Table 2.4:** *Coefficients results.*

| Dataset | JC | DC | $d_H$ (mm.) |
|---|---|---|---|
| Training(ID:0) | 0.9 | 0.95 | 37.19 |
| Healthy(ID:1-8) | 0,9 ± 0,02 | 0,95 ± 0,01 | 31,61 ± 4,9 |
| Unhealthy(ID:9-16) | 0,91 ± 0,03 | 0,96 ± 0,02 | 35,56 ± 4,9 |
| Total(ID:0-16) | 0,91 ± 0,02 | 0,95 ± 0,01 | 33,58 ± 6,1 |

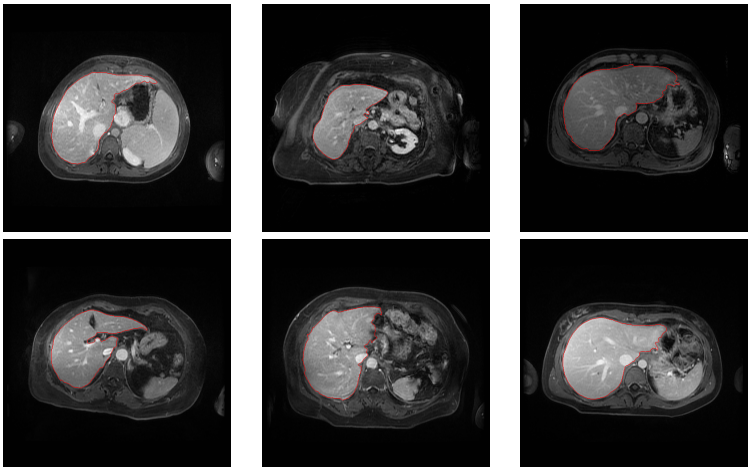


**Figure 2.14:** *First original slice and segmented mask (in red) for different studies.*

The datasets used for the validation were from healthy and unhealthy patients with different properties (spacing, scale, machine, or pathologies). With these considerations, a JC of 0,91 ± 0,02 (or a DC of 0,95 ± 0,01) and the Hausdorff distance of 33,58(±6,1) mm. demonstrate the accuracy and the robustness (low standard deviation of the coefficient) of the method proposed in this paper.

If a high overlap coefficients and a high Hausdorff distance are obtained, this means that in general a good segmentation is carried

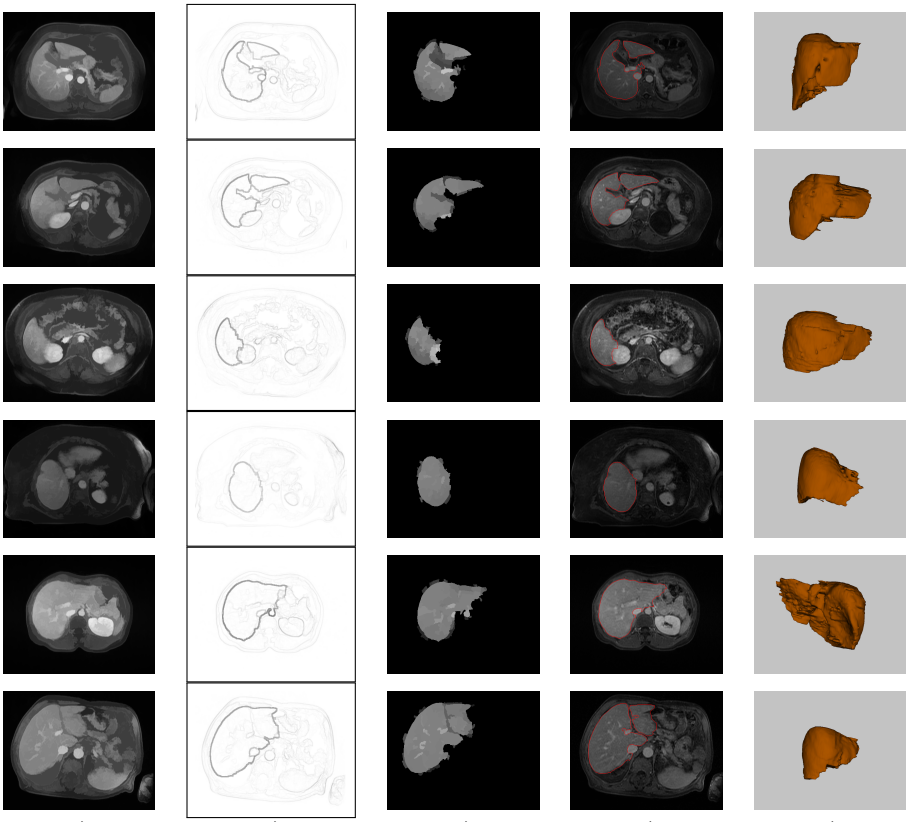



**Figure 2.15:** *a) Area opening filter output, $\gamma_\lambda(f_i)$; b) the probabilistic edgeness function($e_i(\mathbf{x})$); c) corrected image by illumination, $f_{c_i}$; d) original image and final mask ($f_i$, $Mask_i$); e) 3D results of some datasets.*

out in all the images and some particular problem in some images can produce this high Hausdorff distance. The causes of obtaining a high value of overlap (good segmentation) while a high distance value may be these two: the misclassification is produced in cranial or caudal images and the error propagation is stopped by the algorithm when it stops, Fig. 2.17, or if the misclassification is produced in central images,

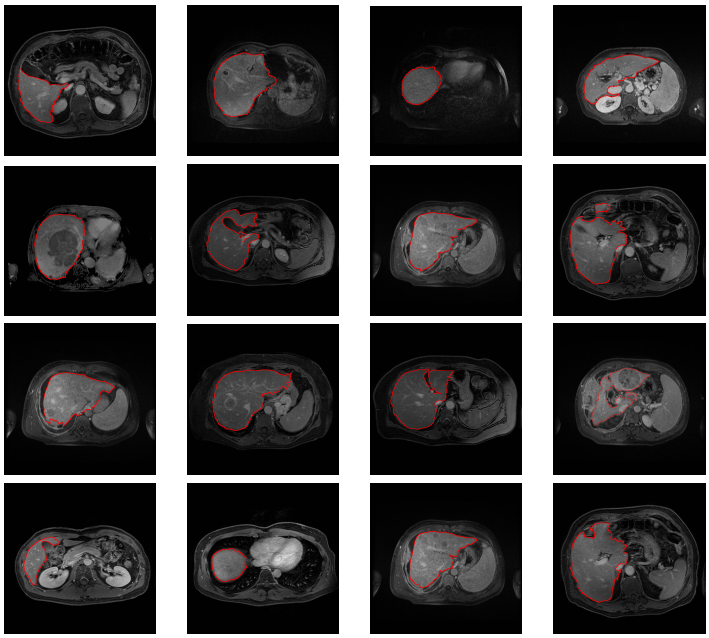
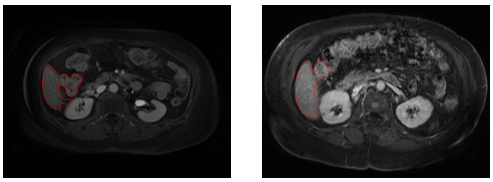


**Figure 2.16:** *Original image and final mask from unhealthy datasets.*

errors are not spread to adjacent slices. To discover the correct cause of this discrepancy a test was carried out and 10 % of caudal and cranial segmented slices were eliminated and the coefficients were calculated. The results in this new scenario for Hausdorff distance and for Jaccard coefficient improved by 50 % and by 25 % respectively. So the conclusion is the high value of Hausdorff distance had the origin in caudal/cranial images and the worst segmentation is produced in cranial and caudal images. One option to solve these problems is to use additional seed points that can be added manually by the user and that can help to improve the results in these images.

Five of the datasets were acquired in a 3 Teslas MR machine and the others in a 1.5 Teslas MR machine. Table 2.1 shows that the



(a) (b)

**Figure 2.17:** *Misclassification: (a) caudal image; (b) cranial image.*

group of unhealthy patients (where the five 3T studies are included) presents more accurate results despite of being pathological cases and it is probably due to the improvements of MRI hardware. Therefore, it is expected that MRI hardware improvements will improve segmentation results.

Our results were compared with the results of others authors in order to establish our conclusions and future lines of work.

In (Heimann et al., 2009), several methods of liver segmentation are presented and evaluated. These methods have been validated with 10 CT datasets whose spatial resolution is higher than the MRI datasets used in this work. So, the results presented in this paper cannot be compared in a direct manner with the results of these authors but initial conclusions can be extracted from this comparison. In the automatic methods, seven authors have a JC in the range (0.91-0.87.5), only two of them have a JC higher than 0.92, and one author obtains a JC smaller than 0.8. Other coefficients have a similar variability. The main drawbacks of the method that obtained the best results in (Heimann et al., 2009) is that it needed more than 100 liver shapes for the training process and besides that a semi-automatic (and manual) iteration was required in the training step. In general, a high computational cost was required in these methods [15-45 minutes per dataset]. In the iterative methods, JC is better than in the automatic ones but a high computational cost and a hard training or initialization is required.

About MRI methods (table 2.5), in the fully automatic algorithms

only healthy patients were used for the method validation (or fat livers in two cases), i.e., patients with tumors were not used (Gloger et al., 2011; Gloger et al., 2010; Masoumi et al., 2012; Goceri et al., 2012). Besides that, two of them used few images for validation (Masoumi et al., 2012; Goceri et al., 2012) and other two used only a type of coefficient for the validation procedure (Gloger et al., 2011; Gloger et al., 2010) so a direct comparison with our results is difficult.

**Table 2.5:** *Results of other authors and main features of their methods. By default runtime is given in minutes per dataset and No. datasets is the number of 3D datasets used in the validation of each method (2D means that only several 2D images were used).*

| Ref. | JC/DC | $dH$ (mm.) | Runtime | No. datasets | User-interaction |
|---|---|---|---|---|---|
| **ours** | $JC = 0{,}91(0{,}03)$ $DC = 0{,}95(0{,}02)$ | 33.58(6.1) | 7-11 $7\ s/image$ | 8 healthy, 9 unhealthy | fully autom. |
| (Cheng 2008) | - | - | - | 3 2D-images | manual segm. train. proc. |
| (Chen 2009a) | - | - | $8\ s/image$ | 12 healthy | init. by seeds manual work |
| (Gloger 2011) | $DC = 0{,}94(0{,}02)$ $DC = 0{,}89(0{,}06)$ | - | - | 20 healthy, 20 fat liver | fully autom. |
| (Gloger 2010) | - | 13,51(7,59) 20,35(8,66) | 11,22(2,78) 15,37(4,96) | 10 healthy, 20 fat liver | fully autom. train. requir. |
| (Goceri 2012) | $DC = 0{,}92(0{,}02)$ | 12,74(3,7) | $14\ s/image$ | 6 2D-healthy | fully autom. |
| (Masuomi 2012) | $JC = 0{,}94$ | - | - | 115 2D-healthy | automatic |

The runtime of these methods was higher than our method in all the cases. In general, overlap measurements (JC and DC) were better in our method. Regarding to the other three MRI methods presented in table 2.5, manual initialization was required (Chen et al., 2009a; Dueñas et al., 2008) or few images and qualitative measurements were provided for validating the algorithms (Cheng et al., 2008). Only one author in

table 2.5 provided overlap (JC or DC) and distance coefficients but few images were used for the validation of this method (Goceri et al., 2012). Summarizing, to accurately validate a segmentation algorithm both coefficients ($d_H$ and overlap, i.e, JC or DC) should be calculated (Lago et al., 2012), so although some authors presented a better Hausdorff distance than our method, we cannot conclude that the accuracy is better. If the high level of automation of our method is taken into account, we can state that our results are promising. A Jaccard coefficient of $0{,}91 \pm 0{,}02$ (or DC of $,95 \pm ,01$) is better than the results obtained by the majority of authors in the studies mentioned above.

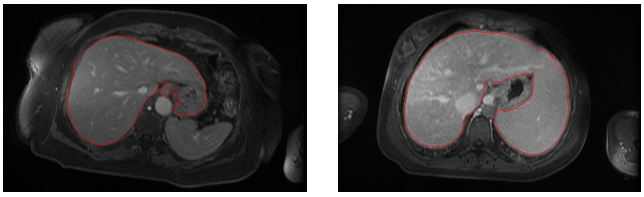


**Figure 2.18:** *Results in some datasets where a 3D variation of our method is applied.*

The algorithm presented in this paper could be applied directly in the 3D data instead of following a 2D strategy (slice by slice) but some problems have been detected in that case. On the one hand, using the 2D strategy, the external marker definition is computed as the dilation of the previous segmented mask and the main problem only consists on obtaining an accurate segmentation in the first image. Fig. 2.14 shows as our algorithm segment the first image with enough accuracy, so the markers chosen with this 2D strategy contribute to the good performance of our algorithm. On the other hand, if the same algorithm is applied to the 3D data, the 3D marker definition for the background is more complicated. Fig. 2.18 shows two cases where adjacent structures are misclassified as liver due to the poor gradient between adjacent organs and where the use of patient's skin is used as external marker. However, if the segmentation is carried out slice by slice, the external marker is extracted from previous segmented slices

where these organs are separated from the liver. When the algorithm segments these problematic slices, the external marker position (a dilation of the previous mask) achieves that the stochastic procedure enhances this poor gradient and improves the performance of the 3D approach.

Our segmentation algorithm fulfils the accuracy objective for trocar placement in laparoscopic surgery (the accuracy required for minimally invasive surgery in the insertion of trocars (Feuerstein, 2007)).

# Capítulo 3

# A Method for Liver Segmentation on Computed Tomography Images in Venous Phase Suitable for Real Environments

## Índice



[4]The content of this chapter is the author version of the following published paper: López-Mir F, González P, Naranjo V, Pareja E, Morales S, and Solaz-Mínguez J. A Method for Liver Segmentation on Computed Tomography Images in Venous Phase Suitable for Real Environments. J Med Imaging Health Inform. 2015;5:1-9

# Abstract

*Liver segmentation is the first step in image-guided liver surgery or hepatic planners. Nowadays, different methods are being published for the segmentation of the liver but, in general, most of them are not suitable for clinical practice due to several inconveniences as high computational cost, excessive user dependence or low accuracy. The purpose of this paper is to present the performance and validation of a liver segmentation method in computed tomography images where automation, easy user interaction, and low computational cost (besides the required accuracy for clinical purposes) have been taken into account for selecting the final algorithm. Firstly, an adaptive filter based on intrinsic parameters of the liver is applied to reduce noise but preserving external liver gradients. In a second step, from a seed or a group of them, voxels with similar intensities are included in an initial 3D mask. Finally, thanks to the combination of morphological operators in different orientations, several non-liver structures (cava vein, ribs, stomach or heart) are removed and the final 3D liver mask is obtained. Thirty public datasets have been used to estimate the accuracy of the proposed algorithm, twenty for training the method and ten for testing it. An average Jaccard index of 0,91 (±0,03), a Hausdorff distance of 26,68 (±10,42) mm, and a runtime of 0.25 seconds per slice, state a promising efficiency and efficacy in the test datasets. Additionally, the low variance demonstrates the robustness of the method. To our knowledge, liver segmentation methods in the state of the art are achieving high accuracy at the expense of requiring an exhaustive training stage or so much clinician interaction time in different steps of the process. In this paper, a method based on intensity properties is carried out with a high grade of automatism, an easy user interaction and a low computational cost. The results obtained for different patients state a low variance and a good accuracy in most images, thus the robustness of the method is demonstrated.*

## 3.1. Introduction

Hepatocellular carcinoma is the responsible of most liver cancers, the sixth most common cancer in the world, the main reason of death in cirrhotic people and it is also the third cause of death by neoplasm. Nowadays, hepatic tissue anomalies are treated with a qualitative comparison of computed tomography (CT) or magnetic resonance images (MRI) thanks to physician experience; however, quantitative measures are not extended. Liver segmentation is the first step to help surgeons or radiologists to delimit lesions in this organ, and thus liver/lesions ratios and measures could be calculated. An accurate liver segmentation has a direct application in planning, monitoring, and treatment of different types of pathologies such as cirrhosis or hepatocellular carcinoma diseases (Li et al., 2013), but liver segmentation in CT images is currently a difficult problem (Heimann et al., 2009; Mharib et al., 2012; Punia and Singh, 2013).

Liver segmentation algorithms are applied to MRI or CT images. In the literature, these methods are more commonly used in CT (Li et al., 2013; Heimann et al., 2009; Mharib et al., 2012; Punia and Singh, 2013; Ji et al., 2013; Park et al., 2003; Wang et al., 2013; Yang et al., 2014; Casciaro et al., 2012; Peng et al., 2014; Ciecholewski, 2014; Campadelli et al., 2009; T. Heimann, 2007; Kainmüller et al., 2007; Ling et al., 2008; Ruskó et al., 2009; Chen et al., 2009b; Soler et al., 2001) than in MRI (Punia and Singh, 2013; Gloger et al., 2010) for different reasons: CT has a better gradient response than MRI, it has less artefact effects because the movement is less (a CT study requires an acquisition time of 2-3 minutes whereas a MR analysis requires 6-8 minutes) and, consequently, CT involves less cost than MR. Currently, some efforts are being focused on the liver segmentation in other image modalities as positron emission tomography (PET) or ultrasounds but the poor spatial resolution is a crucial disadvantage that neither manual corrections can solve in some cases for clinician purposes (Milko et al., 2008; Hsu et al., 2008; Li et al., 2012).

Liver segmentation methods can be classified according to several criteria: user interaction, input image type, algorithm properties, etc. Regarding the algorithm properties, the methods can be divided in greylevel-based and contour-based methods. Contour-based methods generally achieve better results in liver segmentation but a more complex interaction, initialization, and/or training process is required in order to obtain the desired results. These characteristics can be inappropriate in a clinical environment. Probabilistic atlases, level-sets, deformable models, statistical shape models are algorithms used in liver segmentation which are based on contour properties (Li et al., 2013; Heimann et al., 2009; Mharib et al., 2012; Punia and Singh, 2013; Ji et al., 2013; Park et al., 2003; Wang et al., 2013; Yang et al., 2014; Casciaro et al., 2012; Peng et al., 2014; Ciecholewski, 2014; Campadelli et al., 2009; T. Heimann, 2007; Kainmüller et al., 2007). Probabilistic atlases require a manual segmentation of several livers in order to form the atlases (Ji et al., 2013; Park et al., 2003). This part requires a hard work, since the robustness of the method is related to the quality of manual liver segmentations that forms the atlas. Then, a registration algorithm is performed to find out the correspondence between the liver atlas and this structure in the new image. The computational cost of this process can be higher than 30 minutes in some cases (Ji et al., 2013). Level set algorithms fit an initial curve to the boundary of the organ of interest, in our case the liver. These algorithms also need training (Wang et al., 2013) or an initial iteration to form the initial curve, for example with more than 30 seeds (Yang et al., 2014) or drawing this initial curve (Peng et al., 2014). Some efforts are focused on providing this initial curve in an automatic way but the accuracy of the algorithm decreases (Ciecholewski, 2014) or computational cost increases more than 10 minutes for a typical dataset of 100-120 slices (Casciaro et al., 2012). In the case of statistical or deformable models, a hard training is also required in a similar way to the probabilistic atlases (T. Heimann, 2007; Kainmüller et al., 2007). Additionally, the computational cost increases considerably in these methods and therefore the clinical usability decreases. For those reasons (computational cost, training, and hard user interaction) these algorithms can be accurate but their clinical

application is limited.

Greylevel-based methods have a lower computational cost but these algorithms are less robust to noise and gradient changes. Thresholding algorithms, classifiers as k-means, or region growing methods belong to this group of algorithms (Ruskó et al., 2009; Chen et al., 2009b; Soler et al., 2001; Gloger et al., 2010). Region growing methods need some seed points in order to use some greylevel criteria for pixel addition to the liver mask (Ruskó et al., 2009; Chen et al., 2009b; Gloger et al., 2010). Thresholding algorithms use histogram properties to classify pixels, thus these methods are optimized for a determined type of images with similar histograms properties (Soler et al., 2001). These methods can obtain promising results and thanks to the reasonable computational cost and the low training required, they could be optimal for clinical environments (Ruskó et al., 2009; Gloger et al., 2010). Efficient pre- and/or post-processing steps are needed in order to reduce noise and adjacent organs connections in this kind of algorithms to carry out an accurate segmentation but these processes could increase the computational cost and the user interaction and similar problems to contour-based methods may appear (Gloger et al., 2010).

User iteration increases the method accuracy but it reduces their usability due to the consuming of clinician's time and the final algorithm can obtain user-dependent results (Heimann et al., 2009; Wang et al., 2013; Yang et al., 2014). So, it is important to balance the user interaction and the automation of the method (the behaviour of greylevel methods are better in these issues) and the robustness and accuracy of the final segmentation (the behaviour of contour-based methods are better in these cases).

The primary contribution of this paper is the design and validation of some pre- and post-processing steps in order to perform a 3D region growing algorithm for the segmentation of the liver in CT images. Region growing algorithm has the advantages of a low computational cost and no training requirements. These conditions are combined, in

the proposed work, with an easy user interaction, and a considerable robustness and accuracy, provided thanks to the pre- and post-processing filters designed. The selection of this algorithm is based initially on the accurate results of some authors (Ruskó et al., 2009; Gloger et al., 2010). However, the main problems of these algorithms are the pre- or post-processing steps used, that increase the computational cost and therefore, clinical utility is reduced. Besides the selection of an optimum criterion that allows an accurate 3D growing, the success of our proposal lies in the design of these pre and post processing stages which turn a 3D region growing algorithm into a suitable solution for clinical use, i.e., a reasonable balance between accuracy, robustness, computational cost, and user interaction.

The liver segmentation method proposed in this paper is included within the framework of the HepaPlan project (Lopez, 2013). The aim of HepaPlan is the development of a planning system for hepatic surgery, Figure 3.1, that is currently under clinical validation.

After the liver tissue extraction, internal structures (hepatic veins and arteries, or biliary ducts) and lesions will be segmented. These segmented structures are obtained from different studies of the same patient, and, therefore, a registration step is required to have this information in a common 3D space. The correct segmentation and registration of these hepatic structures is needed in order to develop several tools for surgeons and radiologists: to compute different measures (such as tumour volume or area) or ratios (liver/tumour volume); to calculate distance between lesions and other structures; to obtain tumour position in Couinaud classification; and to carry out virtual resections for hepatic transplants.

The rest of the paper is divided into three sections. Section 3.2 reminds some tools used in our algorithm and explains the developed algorithm. Section 3.3 presents the final results and Sections 3.4 and 3.5 state a discussion of our conclusions and outlines the future work.

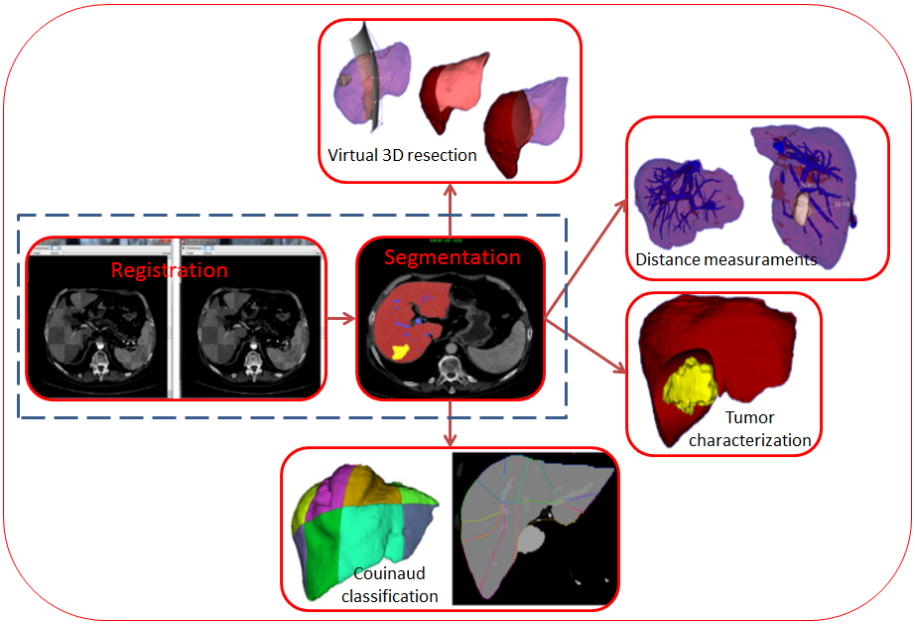



**Figure 3.1:** *HepaPlan project framework.*

## 3.2. Method

### 3.2.1. Theoretical background

Region growing methods are based on neighbourhood greylevel intensities to expand an initial seed (or group of seeds) to obtain a final mask of the region of interest. With this initial seed (or seeds), a pixel is added to the mask (in our case the liver mask) if it fulfils two conditions: it is connected to the pixels that have already been labelled as liver and it satisfies some similitude criterion (different similitude criteria define different region growing algorithms). In our work, this criterion is defined as:

$$f(\mathbf{x}) \in m_0 \pm k * \sigma_0, \tag{3.1}$$

where $f(\mathbf{x})$ is the intensity of a new candidate pixel with coordinates ($\mathbf{x} = (x, y)$), $m_0$ and $\sigma_0$ are the average and the standard deviation of the pixel values of an initial liver mask and $k$ is a tolerance parameter. This formulation is extensive to 3D volumes, only the connexion criteria between voxels changes. If the algorithm begins with a seed point (and not with an initial mask), $m_0$ and $\sigma_0$ is calculated using a neighbourhood around this seed point.

Mathematical morphology is a non-linear processing technique in spatial structures that is based on maxima and minima operators (Serra, 1983). Erosion and dilation are the two basic operators:

$$Dilation : \delta_B(f)(\mathbf{x}) = \max_{\mathbf{b} \in B(\mathbf{x})} f(\mathbf{x} - \mathbf{b}) \tag{3.2}$$

$$Erosion : \varepsilon_B(f)(\mathbf{x}) = \min_{\mathbf{b} \in B(\mathbf{x})} f(\mathbf{x} + \mathbf{b}). \tag{3.3}$$

where $B(\mathbf{x})$ is the structuring element centred at point ($\mathbf{x}$) with a particular size and shape and $f(\mathbf{x}) : \mathrm{E} \to \mathcal{T}$ is a greyscale image where $\mathbf{x} \in \mathrm{E}$ is the pixel position and $T$ represents the pixel greylevel.

Other filter based on mathematical morphology that is used in this paper is the reconstruction by dilation and the partial reconstruction by dilation (Soille, 2003). The reconstruction by dilation of an image $g(\mathbf{x})$ from a marker image $f(\mathbf{x})$ is defined as a geodesic dilation of the image $f$ with respect to other image $g$, called reference, up to idempotence:

$$R_g^{\delta}(f) = \delta_g^{(i)}(f), \tag{3.4}$$

where $\delta_g^{(i)}(f)$ is such that $\delta_g^{(i)}(f) = \delta_g^{(i+1)}(f)$, and where $\delta_g^{(i)}(f) =$

$\delta_g^{(1)}(\delta_g^{(i-1)}(f)))$ being $\delta_g^{(1)}(f) = \delta_B(f)$ the geodesic dilation with the unitary structuring element (a disk of radius equal to one pixel).

The partial reconstruction by dilation instead of applying the geodesic dilation $(\delta_g^{(i)}(f))$ up to the idempotence, it is carried out $N$ times, i.e. $i \in N, i = 1, 2, ..., N$.

The close-hole operator can be defined from the reconstruction by dilation. In greylevel images, any set of connected voxels surrounded by connected components of higher values than the surrounding is considered a hole. This operator fills all the holes in an image $f(\mathbf{x})$ that do not touch the marker image $f_\partial$:

$$\chi^{ch}(f) = [R^{\delta}_{f^c}(f_\partial)]^c, \tag{3.5}$$

where $f^c$ is the complement image of $f$, and $f_\partial$ is the image border defined in 2D as:

$$f_\partial(\mathbf{x}) = \begin{cases} 255, & \text{if } x = 0 \mid\mid y = 0 \mid\mid x = width \mid\mid y = height \\ 0, & otherwise. \end{cases} \tag{3.6}$$

### 3.2.2. Algorithm

The diagram block of our algorithm is presented in Figure 3.2. The first purpose of the algorithm is to obtain a CT volume with a liver tissue as uniform as possible but preserving the gradient between adjacent organs.

Several smoothing filters were evaluated: a curvature flow filter (CF) (Ibáñez et al., 2003), a morphological centre filter (MC) (Soille, 2003), and the adaptive filter (AF) proposed in this paper that is based on first-order statistics of the image, such as the average of the greylevel

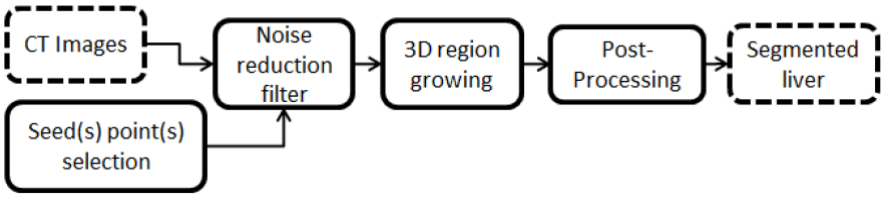



**Figure 3.2:** *Segmentation algorithm flowchart.*

values ($m$) and their standard deviation ($\sigma$). This filter changes the pixel intensity as:

$$f(\mathbf{x}) = \begin{cases} m_5, & \text{if } m_5 \in m \pm \sigma \\ f(\mathbf{x}), & otherwise, \end{cases} \tag{3.7}$$

where $m_5$ is the average of the circular pixel neighbourhood of radius 5 and $m$ and $\sigma$ are an initial estimation of the average greylevel of the liver values and their standard deviation, respectively.

These parameters $(m, \sigma)$ are calculated using a 2D region growing with a seed point selected by the user in one slice where the liver appears as large as possible and with homogeneous greylevel in the liver tissue (Fig. 3.3a). For this region growing 2D, the $k$ parameter has been extracted experimentally ($k$ = 1,2) and the $m_0$ and $\sigma_0$ parameters that control the region growing (Eq. 3.1) are calculated in a circular neighbourhood of radius 25, centred in this selected seed. Each pixel in the final mask of this 2D region growing is considered for calculating the $m$ and $\sigma$ parameters needed by the adaptive filter. With these parameters, the adaptive filter is applied to the whole 3D study according to Eq. 3.7. Several experiments demonstrate that the adaptive filter obtains a more uniform grey-level in the liver reducing the noise and preserving better its boundaries than other approaches. Quantitative comparisons will be presented in Results section.

In the next step, a 3D region growing is applied to the smoothed CT volume (Figs 3.3b-c), Eq. 3.1. The whole mask obtained as result of the previous 2D region growing is used as seed for the 3D region growing algorithm, so $m_0$ and $\sigma_0$ parameters are set to $m$ and $\sigma$ respectively (the adaptive filter parameters). The tolerance ($k = 1{,}85$) is extracted empirically using several datasets for evaluating the optimal tolerance parameter.

At this point, the 3D liver mask states two problems. On one hand, an over-segmentation is produced by organs connected to the liver and with similar greylevel. It is the case of the ribs, cava/portal vein, heart and stomach. These connections appear frequently in all patients, but not always all structures are over-segmented as it depends on the type of study, the contrast agent and the anatomical patient's features. On the other hand, the method obtains the under-segmentation of some lesions with a greylevel lower than liver tissue.

Organ over-segmentation is solved with several post-processing filters focused on two main steps: an erosion to break connections between organs followed by a partial reconstruction to recover and refine the liver boundaries, as it is depicted in Figure 3.4.

The problem of the under-segmentation is solved with a close-hole operator. Only lesions in external zones of the liver tissue, the close-hole filter in axial view doesn't have effect because the lesion touches the liver border. Taking into account to the datasets used for the validation of the method, this situation is not common but to minimize this drawback the filter is applied in the three main views of the 3D study (axial, coronal and sagittal).

### 3.2.3. Post-Processing

Figure 3.4 shows the flowchart of the post-processing stage and Figure 3.5 depicts the results of the different steps. Firstly, 2D erosion is applied to the axial masks with a structuring element, $B_1(\mathbf{x})$ (Fig.

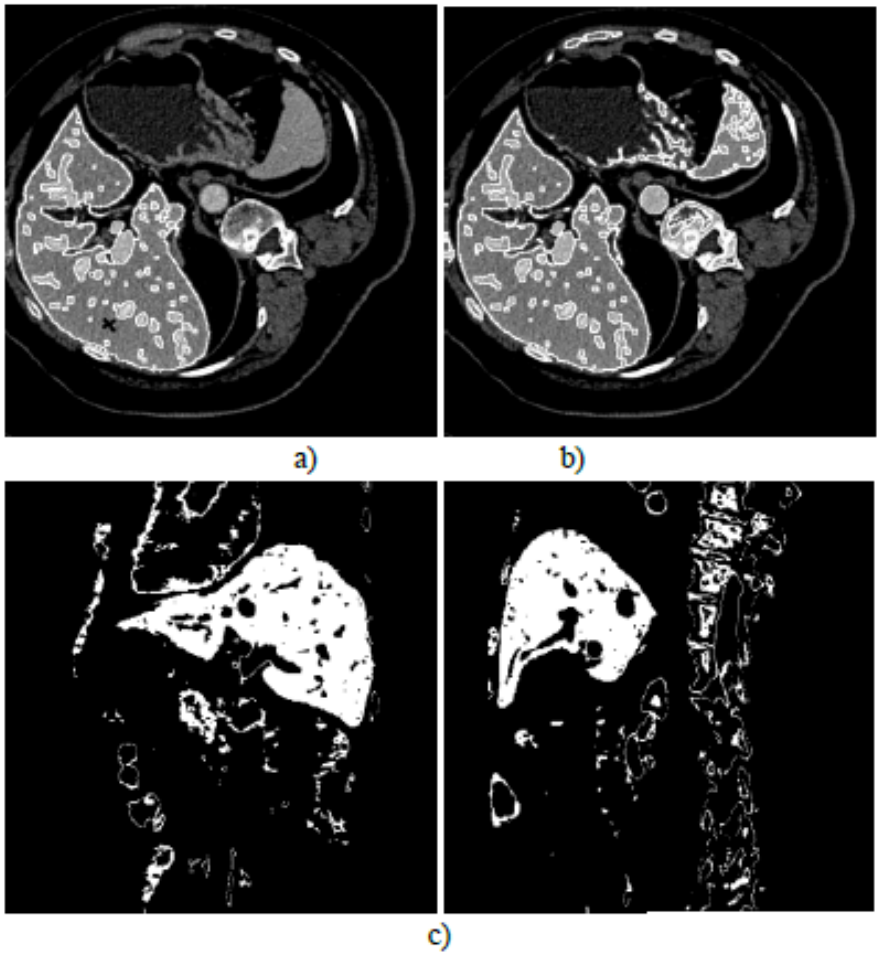



**Figure 3.3:** *a) 2D region growing output (original slice and contour superimposed and seed in black cross position); b) one slice of the 3D region growing (original slice and contour superimposed); c) output 3D region growing masks in sagittal/coronal views.*

3.5a). In the eroded masks, $M_E$, the holes are closed using the close-

hole operator in the three main directions of the eroded masks (axial, sagittal and coronal), the largest 3D object (the liver) is selected being 2D dilated in axial direction, with the same structuring element than the 2D erosion to try to recover the original size of the liver, obtaining the mask $M_D$ (Fig. 3.5b). Erosion and dilation are not lineal operators, so it is not possible to recover completely the original size with this 2D dilation. For this reason, a partial reconstruction is performed under $M_D$ using $M_C H$ masks (where the contours are better defined) as marker images, to refine the resulting contour of the 2D dilation (Fig. 3.5c).

Finally, the mask contour is smoothed using Fourier descriptor filtering technique (Zahn and Roskies, 1972), as previously proposed in (Gloger et al., 2010), obtaining the final liver mask (Fig. 3.5d).

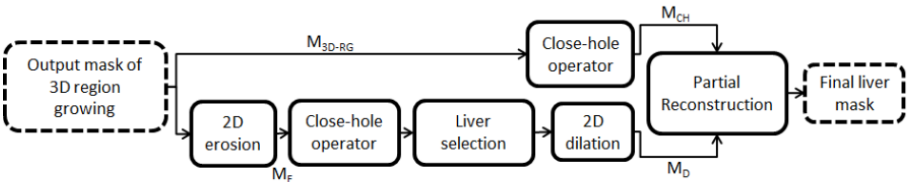



**Figure 3.4:** *Post-processing flowchart.*

## 3.3. Results

The goodness of our method has been evaluated with 30 public studies (sliver07 dataset), the same way that most authors follow to validate their algorithms (Heimann et al., 2009; Ji et al., 2013; Yang et al., 2014; Peng et al., 2014; T. Heimann, 2007; Kainmüller et al., 2007; Ruskó et al., 2009). A collection of 20 studies were provided with manual reference in order to tune the algorithm. Other 10 dataset were provided (without manual reference) in order to submit the obtained masks and then, an independent observer evaluates the liver segmentation method with five coefficients and a score system (Heimann et al., 2009). This approach ensures the validity and independence of the final results.

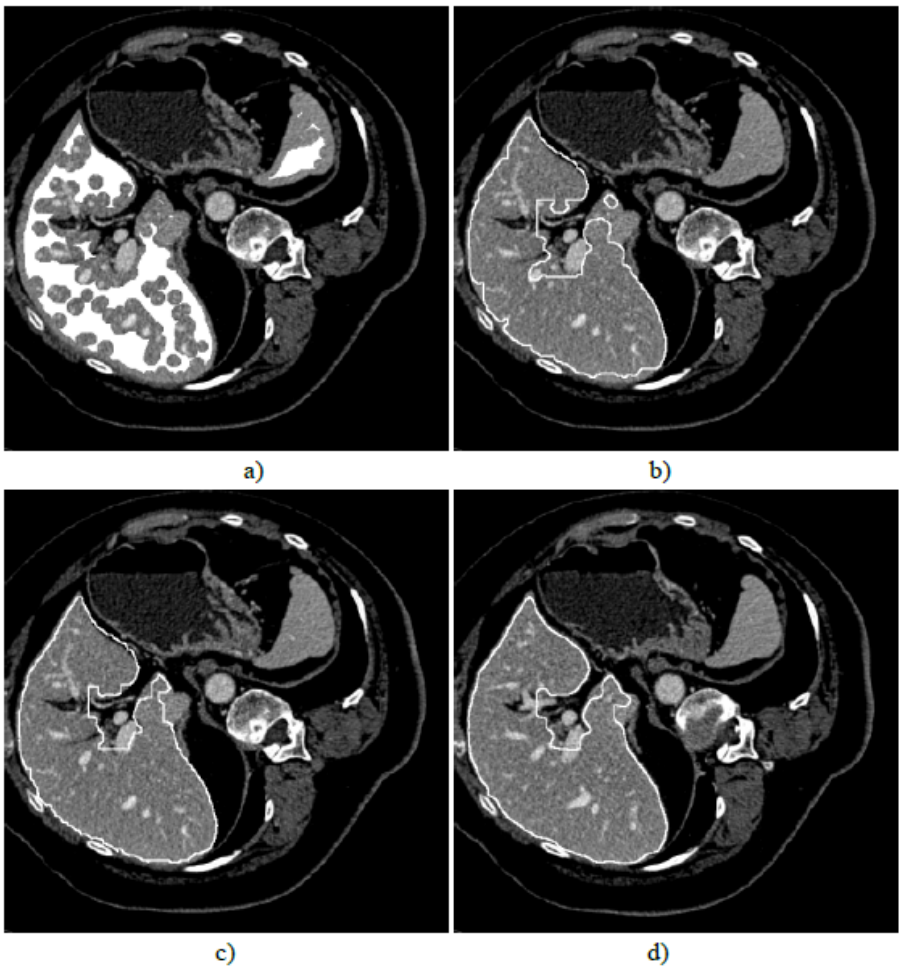



**Figure 3.5:** *Liver masks superimposed to original image resulting to algorithm steps: a) 2D eroded mask; b) dilated contour; c) contour after partial reconstruction; d) final contour.*

The size of each slice was $512 \times 512$ pixels. The number of slices of each volume varied from 64 to 348, depending on the dataset. Pixel spacing

was between 0.55 and 0.88 mm in x/y direction and slice distance was between 1 and 3 mm (Heimann et al., 2009).

In the sliver07 dataset, the five coefficients used to obtain the goodness of the method have been: Volumetric Overlap Error (VOE) or Jaccard Coefficient (JC), Relative Volume Difference (RVD), Average Symmetric Surface Distance (ASSD), Root Mean Square Symmetric Surface Distance (RMSD) and Maximum Symmetric Surface Distance or Hausdorff Distance (HD) (Heimann et al., 2009).

The parameters to be tuned in our method are the structuring element of the 2D erosion (and 2D dilation), $B_1(\mathbf{x})$, and the number of iterations in the partial reconstruction, N, both used in the post-processing stage. After different tests with the 20 training sliver07 datasets, the best results were achieved for a circular $B_1(\mathbf{x})$ of radius 5 and $N = 3$.

Table 3.1 shows the average ($m_{test}$) and the standard deviation ($\sigma_{test}$) of the five coefficients of the 10 test studies evaluated by an independent observer and the average ($m_{train}$) and the standard deviation ($\sigma_{train}$) of the five coefficients of the 20 training studies whose manual reference is open access (Heimann et al., 2009; Miccai:Organicers, 2013). Additionally, the Dice Coefficient (DC) has been also computed for comparison purposes.

All the results were calculated on an Intel core i5 @ 2.80 GHz, with a RAM of 2 GHz and Windows 7 (32 bits). The runtime average of the 30 datasets (20 training and 10 test) is 0.25 seconds per slice. A typical study, 120-140 slices, is processed in 30-35 seconds.

Additionally, Figures 3.6 and 3.7 show respectively the results of different steps of the algorithm and 3D liver results of a private dataset. These CT studies are provided by the "Hospital Universitario y Politécnico La Fe de Valencia"but a manual segmentation is not available and therefore, quantitative coefficients can't be computed. These images are acquired in Toshiba Aquilion and Philips Brilliance

**Table 3.1:** *Final results of the 20 training and 10 test dataset.*

| | VOE/JC | RVD | ASSD | RSMD | $d_H$ | DC |
|---|---|---|---|---|---|---|
| | %/(0-1) | % | mm | mm | mm | % |
| $m_{test}$ | 8.78/0.91 | -4.83 | 1.49 | 3.2 | 26.68 | 0.95 |
| $\sigma_{test}$ | (2.78)/(0.03) | (2.21) | (0.72) | (1.84) | (10.42) | (0.02) |
| $m_{train}$ | 8.48/0.92 | -3.68 | 1.62 | 3.65 | 31.83 | 0.96 |
| $\sigma_{trsain}$ | (1.67)/(0.02) | (2.83) | (0.62) | (1.57) | (10.7) | (0.1) |

CT machines and image resolution and spacing have not significant differences with the sliver07 dataset. This dataset contains healthy and unhealthy patients and the sequence used for the segmentation of the liver is the portal phase because in this sequence the liver appears better contrasted than in others. The qualitative results are very similar to the public studies shown in Table 3.1 and are depicted in Figures 3.3 and 3.5. The parameters used for these private CT studies are the same that in the sliver07 dataset, which demonstrates the good performance of the method in different situations and that is not needed new parameter settings to work properly with different input conditions.

### 3.3.1. Influence of the noise reduction filter

The influence of the noise reduction (or smoothing) filter was calculated by applying different smoothing filters and obtaining the five coefficients in the 20 training datasets. Table 3.2 shows the average of the most representative coefficients for the different filters applied in the pre-processing step.

### 3.3.2. Influence of user-iteration in the seed selection

To evaluate the dependence of the algorithm accuracy with the seed selection, a final test was carried out. The selected seed by one expert in one dataset was stored (called from now the expert seed). Ten random

**Table 3.2:** *Results with different noise reduction filters.*

| Smooth. Filter | VOE (%) | $d_H$ (mm.) |
|---|---|---|
| Morphological center | 11.37 | 32.23 |
| Curvature flow filter | 10.24 | 32.85 |
| Adaptive filter(our proposal) | 8.47 | 31.83 |

seeds were generated. The restrictions to select the random seeds were to have a similar greylevel (between ± standard deviation in a circular neighbourhood of radius 25 pixels) to the seed selected by the expert and to be in a position relative close (a sphere of radius 5 cm was selected) to the expert seed. This procedure tried to emulate the seed selection by a new user who will probably select the seed in a slice where the liver appears as large as possible and with homogenous level intensity (in a region with similar greylevel and not inside a vessel or in a caudal or cranial slice). The algorithm was run ten times using one of the ten random seeds each time and the average and the standard deviation of the most representative coefficients can be observed in Table 3.3. The first row shows the results using the expert seed (previously included in the Table 3.2). The dataset selected for this experiment was the one that got the closest coefficients to the coefficient average of the twenty training datasets.

**Table 3.3:** *Results obtained with the expert seed and the average and standard deviation of the results obtained with the random seeds.*

| Seeds | VOE (%) | $d_H$ (mm.) |
|---|---|---|
| Radiol. seed | 8.47 | 24.41 |
| $m_{10}(\sigma_{10})$ | 8.44 (0.77) | 21.97 (1.33) |

## 3.4. Discussion

Table 3.4 provides the most promising results and features of methods presented in the state of the art for comparison purposes. It has been demonstrated that the Jaccard Coefficient and the Hausdorff Distance are the most significant coefficients for volume comparison (Lago et al., 2012), so these two coefficients (and the dice coefficient for comparative purposes) and the computational cost are provided. The authors (Ji et al., 2013; Yang et al., 2014; Peng et al., 2014; T. Heimann, 2007; Kainmüller et al., 2007) use the same sliver07 dataset (10 test studies) than in this work, so the results are directly comparable. The last 5 authors of this Table 3.4 (Wang et al., 2013; Casciaro et al., 2012; Ciecholewski, 2014; Campadelli et al., 2009; Ruskó et al., 2009) use not publicly available datasets and the comparison is not direct. It is observable that methods with better accuracy than ours required a high user interaction (Yang et al., 2014; Peng et al., 2014) or a high computational cost (Ji et al., 2013; Wang et al., 2013; Casciaro et al., 2012; Campadelli et al., 2009; T. Heimann, 2007; Kainmüller et al., 2007). In (Campadelli et al., 2009; Ruskó et al., 2009) no user interaction or training is required but accuracy is lower than our method (Ruskó et al., 2009) or computational cost is higher (Campadelli et al., 2009).These methods (Campadelli et al., 2009; Ruskó et al., 2009) use the region growing method in some part of the whole process of the liver segmentation, so it is demonstrated the benefits that this algorithm can produce in our purpose.

In (Heimann et al., 2009), several authors published their methods and results in the same ten test datasets (whose manual segmentation is not published). These methods are divided in automatic and interactive methods for comparison purpose. To compare our algorithm with interactive methods it is important to specify that these methods require a hard initialization in comparison to our method. In most cases an initial contour has to be provided for initialization purpose or several seeds located in specific zones of the liver. Additionally, these methods require a computational cost between 7 and 60 minutes per dataset.

These features make these methods accuracy higher than ours (a VOE less than 7 % in most cases) but they reduce considerably the clinical usability of these methods in comparison to ours.

If the ten automatic methods published in (Heimann et al., 2009) are analysed, the accuracy is lower than the one obtained by interactive methods and only two of them (T. Heimann, 2007; Kainmüller et al., 2007) have a VOE lower than our method (other coefficients have similar tendency). However, these two methods need more than 115 and 35 training datasets, respectively, and they require 7 and 15 minutes, respectively, for the segmentation of one patient, in any case higher than our method.

Finally, other region growing algorithms have been analysed and compared with our approach. In (Heimann et al., 2009), an interactive method and an automatic method are based on this algorithm. In both cases, they produce lower accuracy and higher computational cost than our method. In (Chen et al., 2009b), a 2D region growing method is applied to liver segmentation and a runtime of 3.53 seconds per slice is obtained, but quantitative results are not provided. Finally, in (Ruskó et al., 2009) a sophisticated 3D region growing algorithm is applied on CT images and the results in the 20 training datasets that we use in our work are provided. The same five coefficients are computed; the VOE is 8.2 % (in our training datasets is similar, 8.48 %) but the standard deviation (±4,1) is higher than our method (±1,67) so the variability and the robustness of our method is higher.

For these reasons, we can conclude that our approach fulfills the requirements needed for clinical practice. It has a high accuracy (in the same order than the best automatic methods with which it has been compared) and a lower variability, which makes it a robust solution. Besides that, the proposed method has a high usability in comparison with others, measured in terms of computational cost and user interaction.

**Table 3.4:** *Accuracy and computational cost of different liver segmentation methods.*

| Method | VOE (%) | HD (mm) | DC (0-1) | Comput. Cost | Algorithm features |
|---|---|---|---|---|---|
| OUR METHOD | 8.78 | 26.68 | 0.95 | 25-30 s | 1 seed |
| (Ji et al., 2013) | 8.3 | 29.3 | | 35 min | - |
| (Yang et al., 2014) | 5.7 | 20.02 | | 2 min | ≈ 30 seeds in specific positions |
| (Peng et al., 2014) | 5.5 | 18.6 | - | - | High user interaction: initial contour in heart, portal vein and liver |
| (T. Heimann, 2007) | 7.7 | 30.1 | - | 7 min | Training required |
| (Kainmüller et al., 2007) | 6.6 | 23.05 | - | 15 min | Training required with manual segmentation |
| (Ruskó et al., 2009) | 11.4 | 28.12 | | 25 s | |
| (Wang et al., 2013) | | 21.49 /7.68 | | 18/10 min | Training required with manual segmentation |
| (Casciaro et al., 2012) | | | 0.96/ 0.95 | ≈ 18 min | |
| (Ciecholewski, 2014) | | | 0.81 | | Only for healthy patients |
| (Campadelli et al., 2009) | | | 0.95 | 50 s | Image resolution 256x256 |

## 3.5. Conclusions and future work

A liver segmentation method in CT images is presented in this paper. A 3D region growing algorithm has been evaluated with promising results. The selection of this method is based on easy user interaction, automation (only a seed point introduced by the user is needed), and low computational cost. These are typical requirements in a medical environment. A pre-processing filter to reduce noise is applied due to intrinsic CT images properties. An adaptive filter based on statistical parameters is proposed in this paper and it is compared with other

filtering approaches obtaining the best results (Table 3.2). After, a post-processing step is carried out to refine the final results. An averaged VOE coefficient of 8,78 ± 2,78 % (or Jaccard index of 0,91 ± 0,03) and a Hausdorff Distance of 26,68 ± 10,42 mm. in the test datasets sliver07 demonstrate that results are promising above all if the high level of automation of the method and the low computational cost is taken into account. The computational cost is about 0.25 seconds per slice, in a typical study of 120 slices, the software needs 30 seconds for obtaining the 3D final mask. The user interaction is only related to a seed selection, the other parameters were tuned with the training datasets. For this reason, the variability of this seed selection has been measured (Table 3.3) and the low coefficients variability state user independence of the algorithm.

Regarding future work lines, new approaches of our method will be focused on applying some variation of the watershed transform, like the marker-controlled paradigm (Meyer and Beucher, 1990) or the stochastic watershed (Angulo and Jeulin, 2007) in order to improve the accuracy in some areas of the actual region growing method presented in this work. Watershed algorithm does not increase considerably the computational cost and it has easy user interaction. The problem of watershed transform is the initial separation between foreground and background which is equivalent to compute regional minima in the gradient of the image and therefore, an over-segmentation may appear in the original transform. To define these minima and reduce the over-segmentation problem, the final mask obtained with the approach presented in this paper could be used to obtain the needed internal markers in order to apply a 3D marker controlled algorithm for improving the accuracy of our method and for segmenting other areas that are currently not accurately segmented.

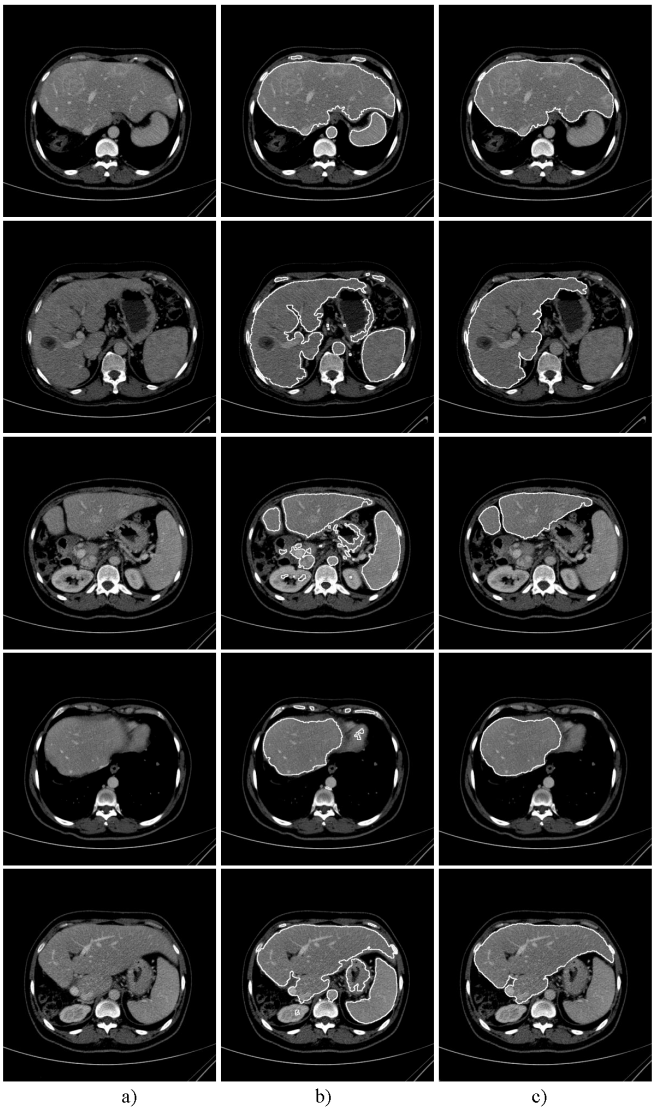



**Figure 3.6:** *Liver masks overlayed on the original image resulting to algorithm steps: a) original image; b) after 2D erosion; c) final contour.*

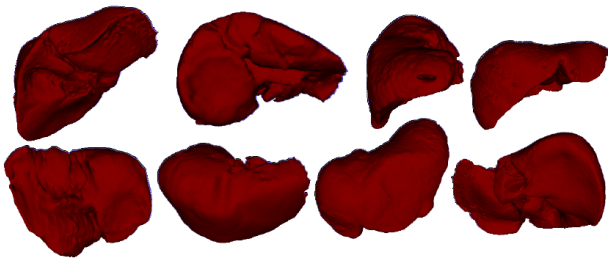


**Figure 3.7:** *3D Liver models of the private dataset.*

# Capítulo 4

# Design and Validation of an Augmented Reality System for Laparoscopic Surgery in a Real Environment

## Índice



[5]The content of this chapter is the author version of the following published paper: López-Mir F, Naranjo V, Fuertes JJ, Alcañiz ML, Pareja E, Bueno J. Design and Validation of an Augmented Reality System for Laparoscopic Surgery in a Real Environment. Biomed Res Int. 2013:1–12

## Abstract

*Purpose. This work presents the protocol carried out in the development and validation of an augmented reality system which was installed in an operating theatre to help surgeons with trocar placement during laparoscopic surgery. The purpose of this validation is to demonstrate the improvements that this system can provide to the field of medicine, particularly surgery. Method. Two experiments that were noninvasive for both the patient and the surgeon were designed. In one of these experiments the augmented reality system was used, the other one was the control experiment, and the system was not used. The type of operation selected for all cases was a cholecystectomy due to the low degree of complexity and complications before, during, and after the surgery. The technique used in the placement of trocars was the French technique, but the results can be extrapolated to any other technique and operation. Results and Conclusion. Four clinicians and ninety-six measurements obtained of twenty-four patients (randomly assigned in each experiment) were involved in these experiments. The final results show an improvement in accuracy and variability of 33 % and 63 %, respectively, in comparison to traditional methods, demonstrating that the use of an augmented reality system offers advantages for trocar placement in laparoscopic surgery.*

## 4.1. Introduction

Laparoscopic surgery has proven to be an alternative to traditional open surgery since smaller incisions are made in the abdomen of the patient (Gordon et al., 1993). The laparoscopic camera and the different endoscopic instruments are introduced through trocars, the hollow cylindrical instruments that are placed into these incisions. Thanks to these smaller incisions, this surgery offers many advantages to the patient, such as less chance of infection, less subsequent operations to remake the abdominal muscle, and so forth. Consequently, the recovery time for the patient is faster both physically and psychologically, which means a lower postoperative cost for the hospital (Courtney K. Rowe, 2012).

The main drawbacks of laparoscopic surgery in contrast to open surgery are the lack of direct vision, the need for hand-eye coordination, and the lack of tactile feedback to the surgeon. Another problem is related to the trocar placement since it may result in more invasive surgery. Currently, the incisions are made by palpation based on the experience and skill of the surgeon. The improper placement of trocars in an operation, such as in the lymph node dissection in the hepatoduodenal ligament, can complicate the operation. In these cases, a relocation of the trocar might be necessary (and more incisions than strictly necessary will be made) thereby limiting the advantages of laparoscopic surgery mentioned above (Feuerstein, 2007).

Augmented reality (AR) is a 3D computer vision technique that is characterized by the real-time fusion of virtual elements on a real space (Azuma et al., 1997). Currently, augmented reality offers enormous potential in many fields such as education, simulation, architecture, advertising, navigation devices, medicine, and rehabilitation (Azuma et al., 1997). Surgery is the branch of medicine where augmented reality has more potential for application because it can provide surgeons with preoperative information (magnetic resonance imaging or MRI, radiography, 3D reconstructions, etc.) in the same place and at the same

time that they are operating. Thus, some of the drawbacks previously cited are alleviated (Samset et al., 2008; Shuhaiber, 2004).

In (Kersten-Oertel et al., 2012), a taxonomy of augmented reality systems in image-guided surgery is proposed. The work compares and analyzes several systems which use augmented reality technology in surgery applications. The analysis is based on the type of input data, the visualization format, and the way in which data is displayed in the operating theatre. The objectives of this comparison are to establish the syntax for defining a system of these characteristics and to show the principal components of an AR system for image-guided applications. In our case, following the analysis suggested in the work (Kersten-Oertel et al., 2012), three components have been chosen.

- Specific data of the patient: our system uses MRI of the patient and generates a 3D model of internal structures.
- A visualization format based on color coding for anatomical structures: transparency has been used in the models to give more realistic depth in the virtual model.
- A full HD monitor for displaying data.

The main limitation of AR systems is the registration technique employed. In our case, the registration and fusion are done between a 3D volume from the segmented magnetic resonance images of the patient and the real-time image that is recorded by a webcam placed over the patient in the operating theater, specifically above of the patient's abdomen (Figure 4.1).

Some authors have developed techniques to improve and automate preoperative placement of trocars. Based on 3D information extracted from computed tomography (CT) images or MRI, the surgeons must remember this information once they are in the operating theatre. In (Chiu et al., 2000), an optimal access system with virtual endoscopic

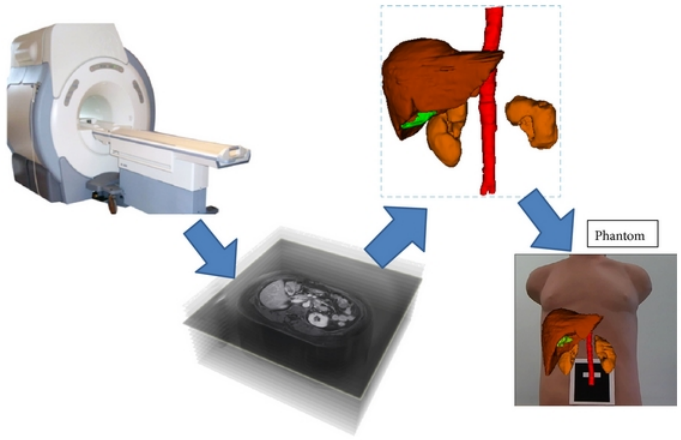



**Figure 4.1:** *Magnetic resonance machine. Magnetic resonance images and 3D model of abdominal organs. Registration and fusion with real-time video in a phantom.*

views is proposed, making the simulation with a phantom. In (Cannon et al., 2003; Adhami and Coste-Manière, 2003), the problem is addressed in image-guided surgery, and trocar placement is optimized from a robotic point of view. The validation is performed on animals. In (Scheuering et al., 2003), the system requires the use of fiducials that have to be in the same position as when the CT was acquired. In addition, the position and orientation of the patient have to be the same in the operating theater.

Other authors deal with the problem in the operating theater similarly to the method presented in this work, but they are focused on the navigation during the intervention using the image that the endoscopic camera provides. In (Bichlmeier et al., 2009), a registration with fiducials is carried out to monitor the camera. These fiducials must be placed over the patient in the same positions in the operating theater and when the CT is acquired. In (Feuerstein et al., 2008), 3D

information is merged into the laparoscopic video. In (Volonte et al., 2010), the validation was done in animals, and the registration and fusion processes were done manually thanks to the surgeon's anatomical knowledge. In (Ferrari et al., 2009), a head-mounted display is used, and the validation was carried out in a commercial phantom.

The experiments carried out in this work were performed in laparoscopic cholecystectomies. This type of intervention is a common solution for diseases such as symptomatic gallstones (McSherry, 1989). It is a common operation with low probability of pre- and postoperative complications. The placement of trocars is usually performed using either French or American techniques (Kum et al., 1996). The choice of one technique or another does not determine the outcome of the experiments, which can be extrapolated from each other. For our experiments, the surgeons chose the French technique because they are accustomed to using it. Both techniques are based on placing four abdominal trocars. Three of them are placed in the same positions in both cases; the fourth trocar differs in the American and the French technique from the area below the sternum in the American and in the opposite side of the liver in the French. In both cases, the surgeon draws four marks with a biocompatible pen taking into account external anatomical references. These marks serve as the initial references of where to make the final incisions. The first trocar (which is called the Veress needle or Hasson cannula and is different from others) is always located at the same position that has been marked with the pen (one centimeter above the navel after making an incision of about 10 mm with the scalpel). By means of this trocar, the pneumoperitoneum technique is performed, and the abdominal cavity is deformed (Martínez-Martínez et al., 2010). Subsequently, the endoscopic camera is inserted through this same trocar to visualize the abdominal cavity (keeping the insufflator hose connected to maintain pneumoperitoneum throughout the entire surgery). The other three incisions are made with the internal vision of the camera and palpation, correcting the position of the marks made with the pen. For two of the incisions, the primary surgeon inserts the surgical instruments (scalpel,

forceps, scissors, etc.), and the other incision is used by the secondary surgeon according to the principal surgeon instructions.

The goodness of our system using augmented reality in the operating theatre was determined by measuring the precision offered by the system compared to not using it. Four distances relating to the four incisions made in the patient ($\parallel d_i \parallel, i = [1,2,3,4]$)were obtained. In this work, the position of the incision ($P_i^* =< x_i^*, y_i^*, z_i^* >$) is equal to the position of the pen market ($P_i =< x_i, y_i, z_i >$) plus an offset ($d_i =< d_x i, d_y i, d_z i >$) (Eq. 4.1). The measured distance $\parallel d_i \parallel = \sqrt{d_{x_i}^2 + d_{y_i}^2 + d_{z_i}^2}, i = 1,2,3,4$ is due to different factor such as displacements and deformities of the pneumoperitoneum technique ($d_{neum}$), the distinctive features of the internal anatomy of the patient ($d_{patient}$)that the surgeon does not notice at the time of making the marks with the pen, and the experience and skill of the surgeon ($d_{surgeon}$):

$$P_i^* = P_i + d_i, \tag{4.1}$$

$$d_i = d_{neum} + d_{patient} + d_{surgeon} + d_{other}. \tag{4.2}$$

The distance $d_{other}$ is related to any error not taking into account by the other variables, for example, operating theatre characteristics (light, position of the patient in the stretcher). When an augmented reality system is used, $d_i$ may be decomposed into the elements of (Eq. 4.3) (similar to Eq. (4.2), but decomposing $d_{other}$ into three new corrections). The augmented reality system (introduces an error due to the applied registration method ($d_{AR}$) which is related to the precision offered by the AR system (Section 4.2.2.1). There is also another error related to the accuracy in the segmentation procedure ($d_{segment}$) of the 3D model which will be projected onto the patient. In our experiments, this segmentation was previously done by an expert and then reviewed by a second expert to check it. This assumption leads us to conclude that the segmentation error can be considered zero or equal to the pixel resolution. The distance $d_{other'}$ is similar to $d_{other}$ (and other errors as the manual alignment of the marker are also included in this distance).

In any case, the hypothesis of this work is that the errors introduced by the augmented reality system will be compensated by the global improvement in $d_i$:

$$d_i = d_{neum} + d_{patient} + d_{surgeon} + d_{AR} + d_{segment} + d'_{other}. \qquad (4.3)$$

The purpose of this work is to measure the $\| d_i \|$ distance for the four trocars which $P_i^*$ (incisions where the trocar is inserted) and $P_i$ (marks made with the pen) are known. These distances are measured in some patients when the system is used and on other patients when the system is not used. The goal is to verify if the use of augmented reality system minimizes $\| d_i \|$.

Several authors have attempted to measure the error caused by an augmented reality system in an operating theater. Most validate it on phantoms and in the maxillofacial and neurosurgery fields. The high resolution of these images and the rigidity of these structures indicate that this error can be explained mainly by (but not limited to) a registration error associated with augmented reality algorithms (Schönfelder et al., 2008; Simitopoulos and Kosaka, 2001; Neil Glossop et al., 2004; Mischkowski et al., 2006; King et al., 2007; Kawamata et al., 2002; Vogt et al., 2006; Fuertes et al., 2011). This error is measured qualitatively (Mischkowski et al., 2006) or quantitatively as being on the order of several millimeters. Other authors validate their algorithms using abdominal operations. In (Nicolau et al., 2011), an AR system is applied in a liver phantom limiting the measured error. The AR system presented in (Scheuering et al., 2003) is validated for liver surgery on pigs where registration with fiducials is used to measure its accuracy.

The rest of this paper is organized as follows. Section 4.2 is divided into two parts, the first part explains the AR system, and the second part describes the protocol of the experiments that were carried out. Section 4.3 presents the results, and Section 4.4 presents conclusions and discussions. The primary contribution of this paper is the design,

performance, and validation of an augmented reality system, the ergonomic study of the visualization devices and the protocol definition for its validation on real patients in an operating theatre.

## 4.2. Methodology

### 4.2.1. Augmented reality system

**4.2.1.1. Virtual 3D Model**. When MR images are acquired, the patient must lie on a stretcher with his/her back straight and centered on both sides to calculate the position and the orientation relative to an initial coordinate system. A virtual model of the patient's organs is extracted from these images using techniques of digital image processing, especially our own image segmentation algorithms (López-Mir et al., 2011) and others developed in (Ibáñez et al., 2003). With this model, the clinician selects the patient's navel in the MRI images to establish the origin of 3D space at that point in order to perform the registration with the real-time image (Eq. 4.4). The new coordinate system is:

$$x' = x + \alpha y' = y + \beta z' = z + \gamma \tag{4.4}$$

where $\alpha$, $\beta$, and $\gamma$ are the coordinates of the center of the patient's navel with respect to the initial coordinate system (x, y, z), (Figure 4.2).

**4.2.1.2. Camera Calibration and Real-Time Image**. The real-time images are recorded with a camera that shows the area of interest throughout the entire surgery. Initially, the intrinsic parameters of the camera are obtained to calibrate it. To do this, it is necessary to have different captures of planar checkerboard patterns (see Figure 4.3), which should be different for each calibration image. Zhang's method is used for the calibration step, taking the correspondence between 2D image points and 3D scene points over a number of images (Zhang, 2000).

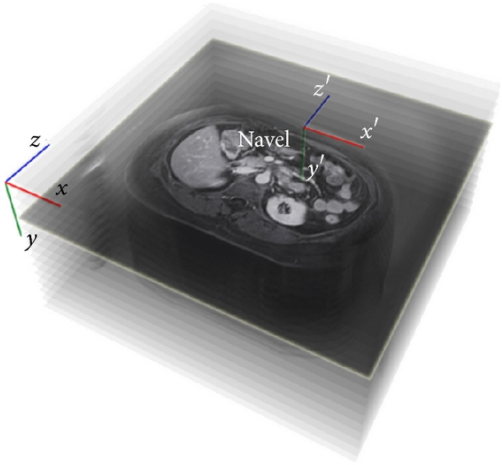



**Figure 4.2:** *Change in the coordinate system.*

The $3 \times 3$ intrinsic matrix K and the vector $\gamma$ of the camera with the distortion parameters have the following form:

**Figure 4.3:** *Checkerboard that is used to calibrate the camera. A Logitech QuickCam Pro 9000 webcam.*

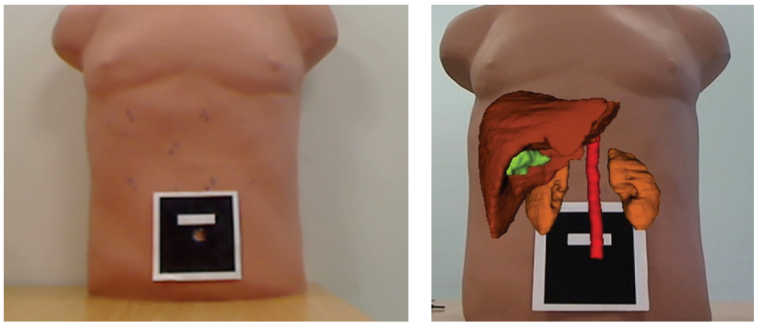
$$K = \begin{bmatrix} f & s & u \\ 0 & af & v \\ 0 & 0 & 1 \end{bmatrix}, \qquad \gamma = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \tag{4.5}$$

where $f$ is the focal length, $(u, v)$ is the optical center of the camera, $a$ is the aspect ratio, $s$ is the camera skew between the x- and y-axes, $\alpha_1$ and $\alpha_2$ are the radial distortion parameters, and $\beta_1$ and $\beta_2$ are the tangential distortion parameters. The values of these parameters for camera calibration were (all in mm)

$$K = \begin{bmatrix} 518{,}55 & 0 & 349{,}727 \\ 0 & 1{,}33 * 517{,}48 & 279{,}897 \\ 0 & 0 & 1 \end{bmatrix}, \ \gamma = \begin{bmatrix} -0{,}323276 \\ 0{,}112309 \\ -0{,}000341309 \\ -0{,}00175445 \end{bmatrix}. \tag{4.6}$$

Then, a hexadecimal mark is placed on the navel and centered and oriented as shown in Figure 4.4. It is advisable to keep the camera parallel to the patient's trunk in order to improve the accuracy of the system, but it is not mandatory (as explained in Section 4.2.1.3) because the system takes into account the inclination between the patient and the camera position. The next steps are the hexadecimal mark detection and the registration and fusion of the real image with the virtual model of the patient.

**4.2.1.3. Registration, Fusion, and Hexadecimal Mark Detection**. A binary hexadecimal code marker of 8,45 × 8,45 centimeters is used in this step. First, the RGB captured image is converted to a binary image, and the edge of the marker is detected thanks to an adaptive threshold algorithm based on the technique of Pintaric (Pintaric, 2003). Basically, *"this technique evaluates the mean pixel luminance over a thresholding region of interest, which is defined as a bounding rectangle around the marker axis-aligned corner vertices in the screen-space."*



(a) (b)

**Figure 4.4:** *(a) Hexadecimal binary mark and (b) fusion of virtual model in a phantom.*

Afterwards, the relative marker position and orientation with respect to the camera (view point) can also be estimated from a planar structure when the internal parameters are known, in order to apply them to the virtual model. First, a 3D/2D homography matrix must be calculated to later obtain the projective matrix, as detailed in (Martín-Gutiérrez et al., 2010).

A 3D/2D correspondence $(m, M)$ includes a 3D point M and a 2D pixel point m, which are represented as $(X, Y, Z, 1)$ and $(x, y, 1)^T$, respectively. $(m, M)$ is related by the $3 \times 3$ projective matrix $P_i$ as (Hartley and Zisserman, 2003) shows:

$$m = \lambda_i P_i M, \qquad P_i = K \left[ R_i \mid t_i \right], \tag{4.7}$$

where $R_i$ is a $3 \times 3$ rotation matrix, $t_i$ is the translation vector of the camera, and $\alpha_i$ is the homogeneous scale factor that is dependent on $P_i M$. Specifically, considering the $z = 0$ plane, the expression of the homography that maps a point onto this plane, and its corresponding 2D point $m$ under the perspective can be recovered by writing

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = m = \lambda_i P_i M = \lambda_i K(R_1 R_2 R_3 t) \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \lambda_i K(R_1 R_2 t) \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \tag{4.8}$$

where $R_1$, $R_2$, and $R_3$ are the columns of the matrix R. Thus, $(m, M)$ is related by a $3 \times 3$ matrix $H_w^i$ that is called homography matrix:

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda_i H_\omega^i \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad H_\omega^i = K(R_1 R_2 t_i). \tag{4.9}$$

Conversely, once $H_w^i$ and K are known, the patient's pose can be recovered from (Eq. 4.7) and (Eq. 4.9), because $R$ is a unit orthogonal matrix, as is explained in (Simon et al., 2000) ("the last column $R_3$ is given by the cross-product $R_1 \times R_2$"):

$$K^{-1} = H_\omega^i = (R_1 R_2 t_i), \quad P_i = K(R_1 R_2 R_3 t_i) \tag{4.10}$$

Generally, the patient's pose can be refined by nonlinear minimization, since the anterior processes are sensitive to noise, and, therefore, a lack of precision and the "jitter"phenomenon are produced.

In this case, the sum of the reprojection errors is minimized, which is the squared distance between the projection of the 3D points and their measured 2D coordinates. We can therefore write that

$$\left[R_i \mid t_i\right] = \underset{\left[R_i \mid t_i\right]}{\arg\min} \sum_j \left\| PM_j - m_j \right\|. \tag{4.11}$$

This equation will be solved using the Levenberg-Marquardt (LM)

algorithm proposed by (Marquardt, 1963), providing a solution for the problem "Nonlinear Least Squares Minimization."

In this way, the 3D virtual model and the patient's image can be registered and fused. Just then, it is important for the patient to maintain his/her position to avoid possible registration errors.

### 4.2.2. Experimentation

**4.2.2.1. Error Introduced by the AR System**. Before the system was validated by the surgeon in the hospital, to test how the AR module works and to determine its accuracy ($d_{AR}$),the following experiment was performed. Initially, 512 × 512 CT images with a spacing-resolution of $0{,}488 \times 0{,}488 \times 0{,}625$ mm per pixel were extracted from a jar by means of a GE LightSpeed VCT-5124069 machine. The model used was a 500 mL DURAN GLS 80 jar with a diameter of 101 mm. The 3D virtual model was obtained by applying a region growing algorithm taking the pixels between thresholds 150 and 2200 Hounsfield Units (HU).

The camera was placed at a 90 degree angle relative to the real jar. Then, the middle point of the jar was selected in the CT images as the new origin, and the marker was centered on the jar. The registration and fusion were performed at that moment, taking an image of the real jar and the virtual jar to validate the system's accuracy. A full graphic example of the experiment is shown in Figure 4.5.

Different positions of the camera and measures were taken. Finally, it was proved that if the camera was placed at a 90 degree angle relative to the real jar, the system was introduced an error of 3 pixels (the minima of all cases). The real width of the jar and the image width provided by the camera were known, so a direct correspondence was made, and a measurement of $d_{AR} = 2{,}91$ mm, was obtained (Fuertes et al., 2011). As mentioned in the introduction, the augmented reality systems introduce an error in the virtual pose calculation with respect

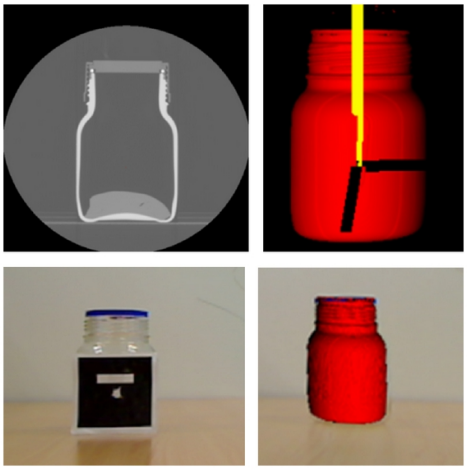


**Figure 4.5:** *CT Jar images, 3D Jar model, marker, and registration and fusion of 3D virtual jar and real jar.*

to the real space. This error is irrelevant in most of the domains where an augmented reality system is used; however, it is of great importance in medicine. The main causes of this error (but not the only ones) are related to the following: the camera (the internal configuration and different lighting conditions that produce different behaviors), the accuracy of registration algorithms, and the accuracy of segmentation methods. Since it is difficult to give the results for each of these errors separately, this measure is usually given as a whole, and in our case it is defined as $d_{AR}$.

**4.2.2.2. Real Patient Experiments**. We carried out two experiments out on real patients. All the documents required for the adoption of the experiments were presented to the research ethics committee of the hospital. This documentation included the following.

- A certificate of commitment related to the ethical principles

of clinical trials: it includes the fundamental human rights and ethical principles related to biomedical research on humans of the Helsinki and Tokyo declaration.

- A certificate of commitment from the researchers that take part in these experiments: in this certificate, the researchers agree to follow the rules and the protocol approved by the research ethics committee of the hospital.
- Informed consent, which is delivered to the patient: this document explains the purpose of study, the procedure, confidentiality, the cost, and the right to leave the study at any time without their final treatment being affected.
- A manual of the developed research: this document specifies the sample selection, the protocol used for the randomization of the sample into the two experiments, the protocol of the whole experiment, and the collection and analysis of the data.
- A validation and data collection protocol: this document has the templates and protocols necessary for the data collection of these studies.
- A request to the hospital committee for approval of the protocol: this document summarizes all the information explained above and is mandatory in order to apply for approval of the experiments.

Initially, the experiments were to be performed through a segmentation of gadolinium contrast MR images. The use of this agent improves the image contrast and facilitates the segmentation of different organs to extract the patient's 3D model. Even though it is safe, there is always the possibility of small allergic reactions in the patient. For this reason and since this contrast agent is not commonly used for this type of pathology, the committee rejected its use in the MRI acquisition. This change caused more difficulties in the segmentation procedure of

abdominal organs, but it did not affect the results or conclusions of the experiments. After making this change, the clinical research committee approved the study, and the experiments were carried out.

In the first experiment, the augmented reality system is not used. The selected sample consists of 12 patients chosen randomly (eight women and four men).The following protocol was used:

- Before the operation (the first time the surgeon visits the patient), the informed consent approved by the research ethics committee of the hospital and common information related to the MRI exam are given to the patient.
- The day of the surgery, the patient goes to the pre-surgery room and then passes to the operating theater.
- The surgeon performs the usual protocol until the operation ends. This protocol can be summarized as follows.
  - First, with a biocompatible pen, the surgeon marks the points where he/she will make the four incisions through which trocars will be inserted (Figure 4.6, left).
  - Second, the surgeon performs the four incisions based on his/her skill, experience, and traditional palpation techniques as explained in Section 4.1 (Figure 4.6, right).
  - When the four trocars are placed, the surgeon begins the operation according to the specific protocol for this type of surgery.
  - Once the gallbladder has been extracted and the four incisions are sutured, the surgeon measures the four values or distances $d_i$ (Figure 4.6, right). These four distances measure the difference between the initial pen marks and the real incisions or, in other words, the correction that has to be made for the technique of pneumoperitoneum, the anatomical differences of patients, and the skill of the surgeon.

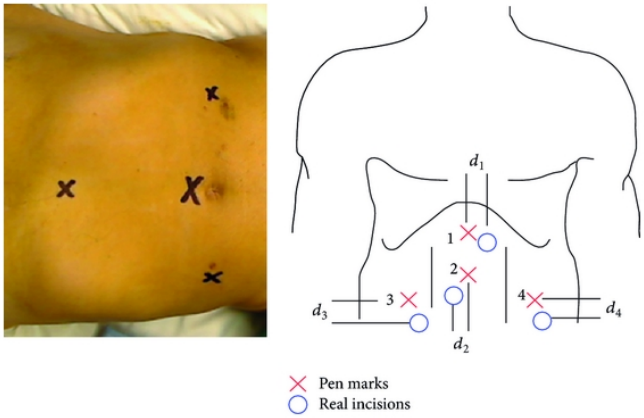



**Figure 4.6:** *Marks in a simulation patient and marks made with the biocompatible pen (red) and real incisions (blue).*

- Finally, the four incisions are bandaged.

- The surgery ends, and the patient leaves the operating theatre and goes to the post-operation room where he/she wakes up and continues with the recovery protocol.

In the second experiment, the augmented reality system was used. The system hardware, as shown in Figure 4.7 (left), is composed of a display device and a camera. The goal of the camera is to capture the image in real-time in order to register and merge this sequence with the 3D virtual model of the patient. The display device is responsible for showing the fusion of the video and the virtual object. In this experiment, different display devices were evaluated. Table 4.1 summarizes the advantages and the disadvantages that different displays offer; this information was obtained by a usability study taking in account operating theatre restrictions.

We chose a 23 inch full HD monitor as the display device based on

**Table 4.1:** *Advantages and disadvantages of different visualization devices.*

| Device | Advantages | Disadvantages |
|---|---|---|
| Projector | Direct vision of 3D model onto the patient, free movement of the surgeon in the scene, and easy hand-eye training. | Low resolution, ambient light in the operating theater that reduces projection sharpness, heat and air ventilation that are dangerous in sterile environments, shadow effect, and set-up just above the scene that entails an associated risk. |
| AR glasses | High grade of immersion and good mobility of the surgeon in the scene but more reduced in comparison to the projector. | Possibility of dizziness, decrease of reality perception and loss of depth, additional material in the surgeon's field of view, and feeling of stress until surgeon gets used to it. |
| Monitor | Harmless for patient and surgeon, free movement in the scene (the screen can be placed farther from the doctor), easy to sterilize, and high resolution. | Hand-eye training since the surgeon must look up to see the monitor. |

the criteria of minimal interaction with the patient, minimal discomfort to the surgeon, and low cost. A dual core i3 computer with graphics card "Nvidia GT 240"was used. The screen and the camera were mounted on a stand as shown in Figure 4.7 (left). The stretcher with the patient was positioned between the stand and the surgeon. The actual image of the abdomen of the patient was captured by the camera which was positioned perpendicularly to the patient as shown in Figure 4.7.

The software (Figure 4.7, right) loads preoperative imaging and 3D models. A mark is placed in the navel for registration and fusion, and the final result is shown in one window of the software (Fuertes et al., 2011).

The sample selected for this experiment also consisted of 12 patients chosen randomly (seven men and five women). The protocol used was similar to the one used in Experiment 1.

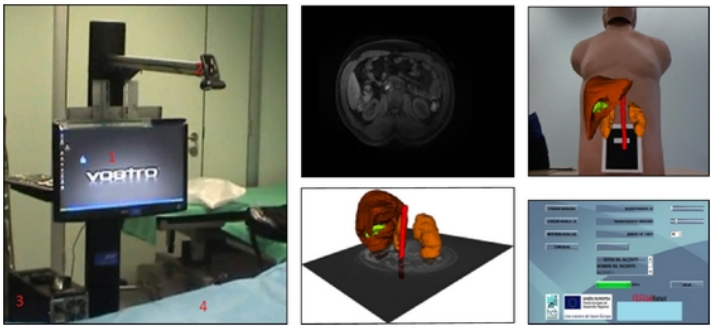


**Figure 4.7:** *Hardware device and software system.*

- Before the operation, the same informed consent as in the first experiment is given to the patient. Then, the MRI is acquired.
- Thanks to different segmentation algorithms, a 3D model of the patient's organs is obtained with the MR images. Specifically, in all cases, the liver and kidneys were segmented; in some cases the gallbladder and aorta were extracted (for surgeon requirements). The tool to perform the segmentation was made ad hoc (López-Mir et al., 2011; Ibáñez et al., 2003).
- On the day of the surgery, all the steps were similar to the first experiment, with only one difference: when the surgeon marks with the pen (Figure 4.8), he/she used the AR system that registers and merges the 3D model with the real-time image (Figure 4.9). The result of this process is shown on the screen that is directly in front of the surgeon.
- Once the 4 marks are drawn, the system is removed, and the surgeon continues the usual protocol until the surgery ends.
- Finally, the same four values or distances ($d_i$) as in the first experiment are measured, and the patient goes to the

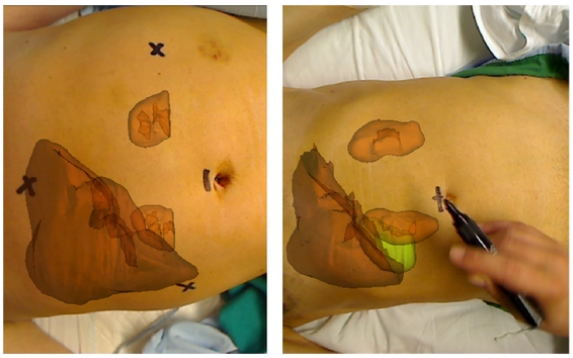


**Figure 4.8:** *Performing the marks in two different simulation patients.*

postoperation room to wake up and continue the recovery protocol.

## 4.3. Results

Ninety-six distances/measures were obtained (four per patient) half of them using the system and the other half without it. The protocol described in both experiments has been followed without major problems. The usual procedure for cholecystectomy surgery was only to be modified when the four distances were measured, after the operation had been completed and before the incisions were bandaged. If any unexpected complication appeared, these distances would not be measured in the patient. During the twenty-four surgeries, no complications occurred, so the measures were taken in all cases. Table 4.2 shows the mean and standard deviation of the four distances measured in Experiment 1 on twelve patients. In this case, the traditional procedures (palpation and the skill of the surgeon) were used in the placement of trocars.

**Figure 4.9:** *Surgeon view and image fusion in a real environment with a simulation patient.*

**Table 4.2:** *Results of Experiment 1 (without the system).*

| | $d_1$ | $d_2$ | $d_3$ | $d_4$ |
|---|---|---|---|---|
| Average(cm) | 1.13 | 0.13 | 1.88 | 1.00 |
| Stan. deviat.(cm) | 1.13 | 0.35 | 1.96 | 1.07 |

Table 4.3 shows the mean and standard deviation of the twelve cases in Experiment 2, that is, when the augmented reality system was used (a new procedure was added to the traditional protocol).

The Mann-Whitney U test was used to validate the null hypothesis of equal medians at the default 5 % significance level relating to the distance measures of both experiments. A p-value higher than 0,05 indicates that there is a non-significance difference, and therefore, the measures can be added. In Experiment 1, the p-values between the four distances were $P_{d1-d3} = 0{,}56$, $P_{d1-d4} = 0{,}93$, $P_{d3-d4} = 0{,}5$, $P_{d1-d2} = 0{,}003$, $P_{d3-d2} = 0{,}048$, and $P_{d4-d2} = 0{,}03$. If these values are analyzed the distance $d_2$ has significance differences with the other three distances. However, the median of the other three measurements has non-significance changes, and it can be assumed that the three distances has the same distribution.

**Table 4.3:** *Results of Experiment 2 (with the system).*

| | $d_1$ | $d_2$ | $d_3$ | $d_4$ |
|---|---|---|---|---|
| Average(cm) | **0.87** | **0.12** | **1.00** | **0.98** |
| Stan. deviat.(cm) | **0.57** | **0.35** | **0.53** | **0.35** |

**Table 4.4:** *Global result of the experiments.*

| | Without system | With system |
|---|---|---|
| Average(cm) | 1.33 | **0.87** |
| Stan. deviat.(cm) | 1.43 | **0.53** |

These conclusions are the same for the Experiment 2 distance measures. Table 4.4 shows the average of the three distances ($d_1$, $d_3$, $d_4$) as a global measure for both experiments. It represents the required correction when an augmented reality system is used and when it is not used.

## 4.4. Conclusions

This paper shows the protocol followed for the validation of an augmented reality system to help surgeons in the placement of trocars on patients in a real environment. First of all, the documentation that is normally required for patient involvement was presented to the hospital. The difference with other experiments, where new drugs or therapies are applicable, lies in the non-impact that this process has on the patient or on the clinical staff because it does not introduce any additional risk to the surgery and no protocols are changed.

The hypothesis of this paper is that an augmented reality system can improve the placement of trocars in laparoscopic surgery. The results confirm this hypothesis since the average accuracy improved and its variability decreased when the AR system was used.

The Experiment 2 of our work was validated with a 3D model extracted from MR images. The reason for this modeling is that MR

images are acquired in the normal protocol of the hospital where we performed the experiments for the laparoscopic cholecystectomies. However, the 3D model can be segmented from other types of images such as CT images. We used the navel as the anatomical structure of reference in the registration procedure of our AR system, but other external reference may be used for other surgeries if it is visible in the CT/MR images.

When Tables 4.1 and 4.2 are analyzed, the first trocar that is introduced in the patient ($d_2$) shows a null improvement. This result is consistent since the surgeon has no additional information (i.e., laparoscopic camera view) to change the position since the first mark is drawn with the pen until the first incision is made as explained in the introduction; it is usually located one centimeter above the navel, so a minimal correction is made with or without the system.

As shown in Table 4.3, the system improved the trocar placement accuracy by 33 %, while variability was reduced by 65 %. The use of an augmented reality system can be helpful in complex situations by providing additional information, where even the use of an internal camera view is not enough for the required accuracy (and more incisions than necessary may be made as mentioned in Section 4.1). Another advantage of using the augmented reality system is its low cost and its applicability. When the surgeon has the internal information provided by the laparoscopic camera which has already been introduced into the patient, they have a finite and limited time to make the rest of the incisions. The longer the decision time, the higher the costs and the greater the risks for the patient. The augmented reality system is useful even in the hours before the operation (when the patient is awake and out of risk) making it possible to plan and reduce the time spent in the placement of the trocar in the operating theater. The augmented reality system also has direct application for automating and optimizing the trocar placement for guided surgery.

When (Eq. 4.2) and (Eq. 4.2) are analyzed, errors introduced by

the augmented reality system ($d_{RA}$, $d_{segment}$, and $d_{other}$) are much lower than the correction offered by the system ($d_i$). The correction (in absolute value) achieved by the system in this work was on the order of $d = 25$ mm, while our AR algorithms introduced an error $d_{AR} = 2{,}91$ mm. If the correction or distance "$d_i$" is the consequence of displacement produced by the technique of pneumoperitoneum and/or the subjectivity of the surgeon and/or the patient anatomical particularities, the system helps to correct the subjectivity and the particularities. The deformity and displacement that the pneumoperitoneum technique ($d_{neum}$) produces could be solved if the 3D model was deformed the same way as the real deformations using a biomechanical or predictive model (Martínez-Martínez et al., 2013a). Since the augmented reality system offers an internal view of the patient's organs, it is hoped that the system can help to accurately determinate the displacement ($d_{patient}$) that is currently corrected with the laparoscopic camera by introducing references that are not visible when the initial marks are made.

It is difficult to make a direct comparison with the literature results because different scenario particularities, the methods, the surgeries, and the type of "*patients*" involved in each validation (phantom, animals, or humans) as it was introduced in Section 4.1. Most of the authors validate their systems using numerical methods and/or using phantoms but few of them are evaluated in clinical settings with real patients, showing that the integration of augmented reality technology into the clinical environment and workflow is not common (Cannon et al., 2003; Adhami and Coste-Manière, 2003; Scheuering et al., 2003; Bichlmeier et al., 2009; Feuerstein et al., 2008; Volonte et al., 2010; Ferrari et al., 2009; Schönfelder et al., 2008; Simitopoulos and Kosaka, 2001; Neil Glossop et al., 2004; Mischkowski et al., 2006; King et al., 2007; Kawamata et al., 2002; Vogt et al., 2006; Fuertes et al., 2011). It is often not feasible to evaluate a system based on surgical outcome or the impact of the system on the patient, but it is possible to evaluate these systems indirectly in phantoms and/or controlled environments (Kersten-Oertel et al., 2012). The contribution of our work is that the

system is validated and evaluated in a real environment with patients, and the benefits of using an AR system are demonstrated in a more realistic manner.

# Capítulo 5

# Conclusiones

En esta tesis se han diseñado, desarrollado y validado dos algoritmos de segmentación del hígado en imágenes anatómicas de alta resolución (resonancia magnética y tomografía computarizada) que cumplen los objetivos de precisión y robustez necesarios para los fines establecidos pero también tienen un alto grado de uso clínico. Esta característica se ha medido con el coste de computación y la interacción e iniciación de los algoritmos en comparación con otros algoritmos publicados. Además, en esta tesis también se ha diseñado y validado un experimento real clínico donde se ha hecho uso de dicha segmentación que, junto a tecnología de realidad aumentada, ha demostrado los beneficios que aporta tanto a los usuarios directos (los clínicos) gracias al ahorro en tiempo, la disminución de subjetividades y la obtención de coeficientes objetivos como a los usuarios indirectos (los pacientes) pues todos los beneficios clínicos redundan en una mejora asistencial, más precisa y personalizada.

En el capítulo 2 se presenta un nuevo método de segmentación automático para el hígado en imágenes de resonancia magnética que se basa en la creación de una nueva imagen para usar como entrada en la transformada *watershed* (en lugar del clásico gradiente). Esta variante hace uso de una función probabilística de contornos de la

imagen más robusta y menos ruidosa que hace que dicha transformada presente buenos resultados de precisión y robustez en comparación con otros algoritmos estudiados en el estado de la investigación. Además el algoritmo presenta un bajo coste de computación y una fácil iniciación e interacción.

Este método ha sido validado en 17 estudios (más de 1.000 imágenes) de resonancia magnética con realce de contraste, en concreto 8 de estos estudios son de pacientes sanos y los otros 9 de pacientes con algún tipo de patología. De los 17 estudios, uno ha sido utilizado para obtener los parámetros óptimos que luego han sido aplicados al resto de estudios. La obtención de estos parámetros óptimos y de la validación del algoritmo se ha llevado a cabo gracias a la previa segmentación manual por expertos de las diferentes imágenes que han hecho posible la comparación directa con los resultados automáticos que nos proporcionaba el algoritmo. La calidad de esta segmentación manual ha sido calculada con el coeficiente de correlación inter-experto.

Los resultados más representativos de la validación de las diferentes bases de datos (coeficiente de Jacard de 0.91, Coeficiente de Dice de 0.95 y distancia de Hausdorff de 33.58 mm.) demuestran la precisión y robustez (reflejada en la baja desviación típica de estos coeficientes) antes comentados. El tiempo de procesado ha sido de aproximadamente de 7 segundos por imagen, es decir 7-10 minutos por estudio. Mientras que los valores de los coeficientes de solape son en general muy precisos, el valor no tan ajustado de la distancia de Hausdorff es debido sobre todo a la menor precisión que tiene el método con las imágenes caudales/craneales. Este coeficiente mejora un 50 % si no se tienen en cuenta el 25 % de las imágenes caudales y craneales.

En la comparación del método propuesto en el capítulo 2 con otros métodos publicados por otro autores, sólo el propuesto en esta tesis usa pacientes patológicos en la validación. Además, la mayoría de métodos usan sólo un tipo de coeficiente en su validación o usan muy pocas imágenes por lo que la comparación directa es difícil pero en cualquier

caso el proceso de validación llevado a cabo en este capítulo es más completo que la mayoría de los publicados. El tiempo de procesado es más alto en general en todos los métodos publicados en comparación al aquí propuesto y los coeficientes de solape son también mejores en el algoritmo presentado en esta tesis. El uso de una estrategia 2D con propagación imagen a imagen es debido a que en cada iteración los resultados son utilizados para obtener los marcadores de la siguiente imagen a segmentar. El uso del algoritmo directamente en 3D es de difícil aplicación precisamente por la definición de los marcadores 3D que se requeriría.

En el capítulo 3 se presenta un método para la segmentación del hígado en imágenes de tomografía computarizada. La utilización de un método basado en un crecimiento de regiones se fundamenta en la fácil interacción e iniciación (sólo se necesita una semilla), su bajo coste computacional y la precisión demostrada por otros autores en situaciones similares (Ruskó et al., 2009), (Gloger et al., 2010). El método se fundamenta en la realización de un pre-procesado basado en un filtro auto-adaptativo y un procesado morfológico 3D que presenta mejores resultados que los mismos filtros en 2D. Los resultados demuestran una buena precisión en comparación con métodos previamente publicados, y además son más utilizables con un coste computacional bajo, una interacción por parte del usuario nula y la introducción de una única semilla por parte del clínico al inicio del proceso. La bondad del método se ha evaluado con 30 estudios siguiendo el protocolo e imágenes que usan la mayoría de autores (Heimann et al., 2009). Además de diferentes coeficientes de solape y distancia, se ha calculado el coste computacional del método siendo de 0.25 segundos por imagen, por lo que un estudio típico se procesa en 30-35 segundos.

Adicionalmente el método se ha validado de una manera cualitativa con imágenes pertenecientes a una base de datos privada. Esto ha sido debido a que no se dispone de una segmentación manual y no se han podido calcular los correspondientes coeficientes de solape o distancia cuantitativos. Los resultados que ofrece el método en estas imágenes

(según expertos) son similares a los resultados objetivos obtenidos con las 30 bases de datos públicas antes comentadas. Los parámetros usados por el algoritmo en este caso han sido los mismos que se utilizaron en el caso de las bases de datos públicas. Ello demuestra la robustez de la metodología planteada ante imágenes de diferentes fuentes y diferentes tipologías.

En el procedimiento de preprocesado planteado en el capítulo 3 se ha demostrado que de entre los diferentes filtros utilizados para la eliminación del ruido, el filtro auto-adaptativo desarrollado en esta tesis es el que mejores prestaciones ofrece. Otra característica que se ha validado en este capítulo es la influencia de la semilla inicial en todo el proceso, concluyendo que dicha influencia es muy baja (prácticamente nula) siempre que la semilla cumpla unas mínimas características fácilmente asumibles por el usuario.

En la comparación con otros métodos publicados, los que ofrecen mejor precisión requieren una alta interacción, un alto coste computacional y/o un extenso entrenamiento. En los casos que tienen un bajo coste computacional y baja interacción, los resultados de precisión son peores que en nuestro caso. Por ello podemos concluir que el método propuesto en el capítulo 3 tiene una precisión o robustez en general del mismo orden que la mayoría de los métodos publicados pero además también tiene un alto grado de usabilidad.

Finalmente, en el capítulo 4 se presenta un experimento clínico que hace uso de los resultados obtenidos con la segmentación de imágenes de resonancia magnética y de tecnología de realidad aumentada. En dicho experimento se demuestra que la utilización de ambas tecnologías mejora la precisión y variabilidad en la realización de las incisiones necesarias en una cirugía con laparoscopia frente a su no uso. Con ello se demuestra en una aplicación real, la verdadera utilidad que puede presentar la previa segmentación de imágenes de alta resolución anatómica del hígado para un proceso, en este caso, de intervención quirúrgica. En este proceso se han obtenido 96 medidas (cuatro por

paciente) repartidas en dos experimentos donde en uno de ellos se ha usado la tecnología antes mencionada y en el otro no. En el experimento donde no se usó la tecnología propuesta, se midió la distancia entre el punto en el que inicialmente el cirujano pensaba hacer la incisión (y que marca con un rotulador biocompatible) y dónde realmente la hizo. En el experimento en que se usó la tecnología la medición fue la misma; sólo varia el momento en que el cirujano iba a marcar con el rotulador las incisiones, momento en el cual se aplicaron algoritmos de realidad aumentada que permitían ver en tiempo real la imagen del abdomen del paciente y sus órganos superpuestos previamente segmentados.

La distancia media (calculada como la media aritmética entre cada una de las marcas iniciales realizadas por el cirujano con el rotulador biocompatible y la incisión real realizada) al usar el sistema fue de 0.87 cm ($\pm$0,53) y cuando no se usó fue de 1.33 cm ($\pm$1,43) lo que quiere decir que la mejora es de un 33 % en precisión y en robustez la mejora es de un 66 %. De las 4 incisiones realizadas al paciente, solo una no presentó ninguna mejora (la primera que se realiza) ya que siempre se hace un centímetro encima del ombligo. Con ello se ha demostrado la hipótesis de partida, que es que un sistema de realidad aumentada puede aportar mejoras al cirujano en el emplazamiento de los trocares en cirugía con laparoscopía. Esta validez es extensible si se usan otro tipo de imágenes como referencia (siempre que la segmentación sea correcta), se aplica a otro tipo de cirugía más compleja o se usa una técnica diferente en el momento de hacer las incisiones. También los resultados y aplicabilidad pueden ser usados en otros ámbitos, por ejemplo en el del entrenamiento quirúrgico o de la simulación y optimización del emplazamiento de trocares en cirugía robótica guiada.

Otra conclusión a extraer de este capítulo 4 es que la mejora introducida por el sistema (del orden de 25 mm.) es muy superior en valor absoluto al error que puede introducir un sistema de realidad aumentada de manera intrínseca (del orden de 2.91 mm.). El sistema planteado ayuda a corregir las particularidades del propio paciente y también parte de la subjetividad del cirujano, si bien no la deformidad

en los órganos que genera la técnica del pneuroperitoneo; por eso el error introducido todavía tiene margen para poder reducirse más. Respecto a la comparativa con otros experimentos realizados por otros autores es muy difícil por las particularidades de cada escenario (real o laboratorio), tipo de paciente (phantom, animales o humanos), protocolo de validación, etc. Por ejemplo, muy pocos usan pacientes o situaciones reales; de ahí que ésta pueda ser considerada una de las más importantes contribuciones de esta tesis.

## 5.1. Trabajo Futuro

Respecto al método de segmentación presentado para las imágenes de resonancia magnética en el capítulo 2, el valor de la distancia hausdorff indica que la precisión del método no es igual en todas las imágenes. Después de hacer un análisis visual de las imágenes segmentadas, parece claro que el método es menos preciso en las imágenes más craneales y más caudales. De hecho, queda demostrado que la eliminación del 10 % de las imágenes craneales y caudales mejora esta medida en un 50 %. Una opción para mejorar este caso sería el uso de una metodología complementaria basada en puntos anatómicos que delimiten otros órganos como el corazón o el estómago y que haría que la propagación no se extendiera a esos órganos. Respecto a la metodología planteada para el caso de imágenes de tomografía computarizada, una vez obtenida una segmentación base, podría usarse una aproximación más simple basada en una variante de la transformada watershed 3D donde los marcadores internos y externos podrían quedar definidos por la máscara inicial que nos ofrece el método aquí presentado.

En cuanto a las aplicaciones directas de los trabajos desarrollados en esta tesis se pueden enumeras varias. Una aplicación que ha hecho uso de los resultados de esta tesis ha sido un planificador de cirugía hepática. Este software parte de los algoritmos de segmentación de masa hepática en imágenes de tomografía computarizada presentados en esta tesis para posteriormente segmentar los vasos hepáticos y las lesiones si las hubiera. En la segmentación de estas estructuras hepáticas se

usan diferentes estudios que realzan el contraste entre tejidos ya que cada una de estas estructuras internas aparece más realzada en una fase o secuencia diferente. Estos estudios son adquiridos de manera continua pero no hay una correspondencia exacta entre ellos, por lo que es necesario realizar un registro previo antes de la fusión de la información en un escenario común.

Con una correcta segmentación del hígado y después de abordar la segmentación de sus estructuras internas y lesiones si las hubiera, previo registro de cada una de estas estructuras, otro paso es el cálculo de medidas objetivas de ayuda al clínico. Éstas siempre están relacionadas con la forma y el tamaño del tumor (volumen, área, diámetro) así como con su posición respecto a otras estructuras (distancias a la vena porta o a las venas supra hepáticas). Con esta información se pueden obtener parámetros objetivos sobre la patología a tratar y hacer comparaciones con otros pacientes o realizar un seguimiento en el tratamiento a seguir por uno de ellos. También es posible hacer cálculos en las volumetrías del hígado que se va a preservar previo a plantear una hepatectomía mayor con el fin de determinar la proporción de hígado que se puede quitar en una resección o cualquier otra operación hepática.

Este planificador capaz de realizar un seguimiento de la morfología tisular, simulaciones de resecciones hepáticas, cálculo de segmentos funcionales del hígado y cálculo de la posición relativa y distancias a vasos sanguíneos del tejido tumoral, es un prototipo CAD/CAM llamado HepaPlan y que se encuentra en fase de validación clínica en la unidad de radiología y en la unidad de Cirugía Hepatobiliopancreática y Trasplante del Hospital Universitario y Politécnico La Fe de Valencia. Se está adaptando el algoritmo de segmentación del hígado en resonancia magnética presentando en esta tesis y se prevé poder ampliar con los correspondientes algoritmos de segmentación de estructuras internas tal y como se ha hecho en el caso de las imágenes de tomografía computarizada así como adaptar los algoritmos de registro necesarios. La finalidad será que el planificador esté adaptado a las dos tipologías de imagen anatómica de alta resolución de uso más común para

el tratamiento, seguimiento o diagnóstico de diferentes patologías hepáticas. Con este logro, también se podrá disponer en un único sistema de información proveniente de diversas fuentes para que el clínico pueda emitir un juicio de una forma más eficiente.

# Méritos

## Journal papers

***López-Mir F**, González P, Naranjo V, Pareja E, Morales S, and Solaz-Mínguez J. A Method for Liver Segmentation on Computed Tomography Images in Venous Phase Suitable for Real Environments. J Med Imaging Health Inform. 2015;5:1-9.*

**López-Mir F**, Naranjo V, Verdú R, Morales S, Brugger S, Pareja E. HepaPlan: a CAD software for planning hepatic surgeries. Int J Comput Assist Radiol Surg. 2015;10:238-239.

**López-Mir F**, Naranjo V, Villanueva E, et al. Comparing algorithms for automated vessel segmentation in computed tomography scans of the lung: the VESSEL12 study. Med Image Anal. 2014;18:1217-1232.

***López-Mir F**, Naranjo V, Angulo J, Alcañiz ML, Luna L. Liver segmentation in MRI: A fully automatic method based on stochastic partitions. Comput Methods Programs Biomed. 2014;114:11-28.*

***López-Mir F**, Naranjo V, Fuertes JJ, Alcañiz ML, Pareja E, Bueno J. Design and Validation of an Augmented Reality System for Laparoscopic Surgery in a Real Environment. Biomed Res*

[6] In italic font, the three main papers of this thesis

*Int. 2013:1–12. (Retitled from the Journal of Biomedicine and Biotechnology in 2013).*

Lloréns R, Naranjo V, Alcañiz ML, **López-Mir F**. Jaw tissues segmentation in dental 3D CT images using fuzzy-connectedness and morphological processing. Comput Methods Programs Biomed. 2012;108:832-843.

Naranjo V, Lloréns R, Alcañiz ML, **López-Mir F**. Metal artifact reduction in dental CT images using polar morphology. Comput Methods Programs Biomed. 2011;102:64-74.

**López-Mir F**, Martínez F, Fuertes JJ, Lago MA, Rupérez Moreno MJ, Naranjo V, Monserrat C. NaRALap: augmented reality system for navigation in laparoscopic surgery. Int J Comput Assist Radiol Surg. 2011;6:98-99.

## Book chapters

Legaz A , Verdú R, Larrey J, **López-Mir F**, Naranjo V, Bernabéu A. Multimodal 3D Registration of Anatomic (MRI) and Functional (fMRI and PET) Intra-Patient Images of the Brain. In: Ferrández Vicente JM, Álvarez-Sánchez JR, de la Paz López F, et al, eds. Chapter 35 of Artificial Computation in Biology and Medicine. Ed. Switzerland: Springer International Publishing, 2015. p 340-347.

## International conferences

Morales S, Naranjo V, **López-Mir F**, Fuentes FJ, Bernabéu A, Luna L. Automatic Computer-Based Brain Tumor Volume Detection. In: International Conference on Recent Advances in Neurorehabilitation (ICRAN 2015), pages 95-96.

Peñaranda FJ, **López-Mir F**, Naranjo V, Angulo J, Kast L, Schnekenburger J. New Spectral Representation and Dissimilarity Mea-

sures Assessment for FTIR-spectra using Unsupervised Classification. In: the International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS 2015), pages 172-177.

**López-Mir F**, Naranjo V, Morales S, Angulo J. Probability density function of object contours using regional regularized stochastic watershed. In: IEEE International Conference on Image Processing (ICIP 2014), pages 4762-4766.

Martínez S, Rupérez MJ, Lago MA, Martínez F, **López-Mir F**, Monserrat C. In vivo characterization of the patient-specific biomechanical behavior of the abdominal aortic aneurysm. In: 12th International Symposium on Computer Methods in Biomechanics and Biomedical Engineering (CMMBE 2014).

Morales S, Naranjo V, Angulo J, **López-Mir F**, Alcañiz ML. Determination of Retinal Network Skeleton through Mathematical Morphology. In: 22nd European Signal Processing Conference (EUSIPCO 2014), pages 1691-1695.

Montoro A, Morales S, Naranjo V, **López-Mir F**, Alcañiz ML. Feature extraction for retinal vascular network classification. In: 2nd IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2014), pages 404-407.

Ureña N, **López-Mir F**, Morales S, Naranjo V, Ruiz JL. Image-Assisted System for the Diagnosis of Bladder Tumor Recurrence. In: 2nd IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2014), pages 141-144.

Naranjo V, Villanueva E, Rhys G, Stone N, **López-Mir F**, Alcañiz M. Stained and Infrared Image Registration as First Step for Cancer Detection. In: 2nd IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI 2014), pages 420-423.

Morales S, Naranjo V, **López-Mir F**, Navea A, Alcañiz ML.

Computer-Assisted System for Hypertensive Risk Determination through Fundus Image Processing. In: XIII Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICOM 2013), volume 41 of IFMBE Proceedings, pages 273-276.

Naranjo V, **López-Mir F**, Marin C, Morales S, Fuertes JJ, Villanueva E. Marker-Controlled Watershed for Volume Countouring in PET Images. In: XIII Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICOM 2013), volume 41 of IFMBE Proceedings, pages 281-284.

**López-Mir F**, Verdú R, Naranjo V, Larrey J, Gonzalez P, Lopez R, Morales J. Efficient Variational Registration Method applied to CT Abdominal Images. Effect of interpolation. In: International Conference on Medical Imaging using Bio-inspired and Soft Computing (MIBISOC 2013), pages 85-90.

Fuertes JJ, Naranjo V, Verdú R, Bernabéu A, **López-Mir F**, Larrey J, Morales J, Sánchez J, Alcañiz ML. Incorporation of a variational registration method into a spectroscopy tool. In: International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2013), pages 473-482.

**López-Mir F**, Gonzalez P, Naranjo V, Pareja E, Alcañiz ML, Solaz J. Liver Segmentation on CT Images. A fast Computational Method Based on 3D Morphology and a Statistical Filter. In: International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2013), pages 483-490.

Verdú R, Larrey J, Morales J, **López Mir F**, Naranjo V, Alcañiz ML, López R. Efficient Image Registration for the Analysis of Different Phases of Contrast-Enhanced Liver CT Data. In: 20th European Signal Processing Conference (EUSIPCO 2012), pages 2596-2599.

**López-Mir F**, Naranjo V, Angulo J, Villanueva E, Alcañiz ML,

López S. Aorta segmentation using the watershed algorithm for an augmented reality system in laparoscopic surgery. In: 18th IEEE International Conference on Image Processing (ICIP 2011), pages 2705-2708.

Fuertes JJ, **López-Mir F**, Naranjo V, Ortega M, Villanueva E, Alcañiz ML. Augmented reality system for keyhole surgery: Performance and Accuracy Validation. In: International Conference on Computer Graphics Theory and Applications (GRAPP 2011), pages 273-279.

Martínez F, Rupérez MJ, Lago MA, **López-Mir F**, Monserrat C, Alcañiz ML. Pneumoperitoneum Technique Simulation in Laparoscopic Surgery on Lamb Liver Samples and 3D Reconstruction. In: 18th Medicine Meets Virtual Reality Conference (MMVR 2011), pages 348-350.

## National conferences

Aparicio A, Verdú R, Larrey J, **López-Mir F**, Naranjo V, Bernabéu A. Registro 3D multimodal de imágenes anatómicas (MRI) y funcionales (fMRI y PET) intra-paciente del cerebro. In: XXXII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2014).

Gallego A, Ruiz JL, **López-Mir F**, Ureña N, Naranjo V, Ordaz DG, Boronat F. Sistema asistido por imagen para cuantificar objetivamente el grado de actividad neo-angiogénica en mucosa de pacientes con cáncer vesical no invasivo. In: LXXIX Congreso Nacional de Urología 2014.

**López-Mir F**, Naranjo V, Villanueva E, Luna L, Alcañiz ML, López S. Protocolo de validación de un sistema de realidad aumentada para ayuda a la colocación de trocares en quirófano. In: XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica

(CASEIB 2012).

**López-Mir F**, Verdú R, Naranjo V, Morales J, González P, Larrey J. Registro variacional de imágenes con contraste de TC de hígado: implementación eficiente y efecto de la interpolación. In: XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2012).

**López-Mir F**, Naranjo V, González P, Cortés M, Villanueva E, Verdú R, Brugger S, Larrey J, Alcañiz ML. HepaPlan: Planificador Hepático. In: XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2011), pages 407-410.

Fuertes JJ, **López-Mir F**, Naranjo V, Ortega M, Alcañiz ML. Sistema de Realidad Aumentada para ayuda en la colocación de trocares en cirugía laparoscópica. In: XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica (CASEIB 2010), pages 28-31.

## Research awards

**López-Mir F**, Naranjo V, Morales M, Angulo J. Probability density function of object contours using regional regularized stochastic watershed. Top 10 % Paper Award in the IEEE International Conference on Image Processing (ICIP 2014).

XI edición del premio Radio Castellón en la categoría de Sanidad a: Proyecto Oncotic, 01/10/2014, Radio Castellón - Cadena ser. La presente tesis se enmarca dentro del proyecto HepaPlan, uno de los seis subproyectos de Oncotic.

"Best innovation capsules for Technology 2013ïn MiHealth Innovation Awards al proyecto HepaPlan, 28/06/2013, Mihealth Forum.

# Índice de figuras

# Índice de tablas

# Bibliografía


Adhami, L. and Coste-Manière, È. (2003). Optimal planning for minimally invasive surgical robots. *Robotics and Automation, IEEE Transactions on*, 19(5):854–863. 2 citations in pages 79 and 99.

Angulo, J. and Jeulin, D. (2007). Stochastic watershed segmentation. In *PROC. of the 8th International Symposium on Mathematical Morphology*, pages 265–276. 9 citations in pages 18, 21, 23, 24, 25, 27, 39, 71, and 115.

Azuma, R. T. et al. (1997). A survey of augmented reality. *Presence*, 6(4):355–385. One citation in page 77.

Balci, N. C., Befeler, A. S., Leiva, P., Pilgram, T. K., and Havlioglu, N. (2008). Imaging of liver disease: Comparison between quadruple-phase multidetector computed tomography and magnetic resonance imaging. *Journal of Gastroenterology and Hepatology*, 23(10):1520–1527. One citation in page 16.

Bertrand, G. (2005). On topological watersheds. *Journal of Mathematical Imaging and Vision*, 22:217–230. One citation in page 17.

Beucher, S. (1994). Watershed, hierarchical segmentation and waterfall algorithm. In *PROC. of Mathematical Morphology and its Applications to Image Processing*, pages 69–76. Kluwer. One citation in page 17.

Beucher, S. and Meyer, F. (1993). The morphological approach to segmentation: the watershed transformation. mathematical morphology in image processing. *Optical Engineering*, 34:433–481. 2 citations in pages 19 and 20.

Bichlmeier, C., Heining, S. M., Feuerstein, M., and Navab, N. (2009). The virtual mirror: a new interaction paradigm for augmented reality environments. *Medical Imaging, IEEE Transactions on*, 28(9):1498–1510. 2 citations in pages 79 and 99.

Boxerman, J. L., Mosher, T. J., McVeigh, E. R., Atalar, E., Lima, J., and Bluemke, D. A. (1998). Advanced mr imaging techniques for evaluation of the heart and great vessels. *Radiographics*, 18(3):543–564. One citation in page 2.

Campadelli, P., Casiraghi, E., and Esposito, A. (2009). Liver segmentation from computed tomography scans: A survey and a new algorithm. *Artificial Intelligence in Medicine*, 45:185–196. 5 citations in pages 30, 53, 54, 68, and 70.

Cannon, J. W., Stoll, J., Selha, S. D., Dupont, P. E., Howe, R. D., Torchiana, D. F., et al. (2003). Port placement planning in robot-assisted coronary artery bypass. *Robotics and Automation, IEEE Transactions on*, 19(5):912–917. 2 citations in pages 79 and 99.

Casciaro, S., Franchini, R., Massoptier, L., Casciaro, E., Conversano, F., Malvasi, A., and Lay-Ekuakille, A. (2012). Fully automatic segmentations of liver and hepatic tumors from 3-d computed tomography abdominal images: Comparative evaluation of two automatic methods. *Sensors Journal, IEEE*, 12(3):464–473. 4 citations in pages 53, 54, 68, and 70.

Chen, G., Gu, L., Qian, L., and Xu, J. (2009a). An improved level set for liver segmentation and perfusion analysis in MRIs. *IEEE Transactions on Information Technology in Biomedicine*, 13(1):94–103. 2 citations in pages 16 and 48.

Chen, Y., Wang, Z., Zhao, W., and Yang, X. (2009b). Liver segmentation from ct images based on region growing method. In *Bioinformatics and Biomedical Engineering , 2009. ICBBE 2009. 3rd International Conference on*, pages 1–4. 3 citations in pages 53, 55, and 69.

Cheng, K., Gu, L., and Xu, J. (2008). A novel shape prior based level set method for liver segmentation from MR images. In *International Conference on Information Technology and Applications in Biomedicine*, pages 144 –147. 2 citations in pages 16 and 48.

Chiu, A. M., Dey, D., Drangova, M., Boyd, W., and Peters, T. M. (2000). 3-d image guidance for minimally invasive robotic coronary artery bypass. In *Heart Surgery Forum*, volume 3, pages 224–231. FORUM MULTIMEDIA PUBLISHING. One citation in page 78.

Ciecholewski, M. (2014). Automatic liver segmentation from 2d ct images using an approximate contour model. *Journal of Signal Processing Systems*, 74(2):151–174. 4 citations in pages 53, 54, 68, and 70.

Couprie, C., Grady, L., Najman, L., and Talbot, H. (2011). Power watershed: A unifying graph-based optimization framework. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33:1384 – 1399. One citation in page 17.

Couprie, M. and Bertrand, G. (1997). Topological gray-scale watershed transform. In *PROC. of SPIE Vision Geometry V, volume 3168,*, page 136?146. One citation in page 17.

Courtney K. Rowe, Michael W. Pierce, K. C. T. C. S. H. J. M. A. B. R. H. T. N. (2012). A comparative direct cost analysis of pediatric urologic robot-assisted laparoscopic surgery versus open surgery: Could robot-assisted surgery be less expensive? *Journal of Endourology*, 26(7):871–877. One citation in page 77.

Cousty, J., Bertrand, G., Najman, L., and Couprie, M. (2009). Minimum spanning forests and the drop of water principle.

*IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31:1362–1374. One citation in page 17.

Dancygier, H., Erturk, S. M., Cay, E., and Ros, P. R. (2010). Computed tomography and magnetic resonance imaging. In *Clinical Hepatology*, pages 405–423. Springer Berlin Heidelberg. One citation in page 16.

d'Assignies, G., Kauffmann, C., Boulanger, Y., Bilodeau, M., Vilgrain, V., Soulez, G., and Tang, A. (2011). Simultaneous assessment of liver volume and whole liver fat content: a step towards one-stop shop preoperative MRI protocol. *European Radiology*, 21(2):301–309. 2 citations in pages 15 and 16.

Duda, R. O. and Hart, P. E. (1973). *Pattern Classification and Scene Analysis*. John Willey & Sons, New Yotk. One citation in page 21.

Dueñas, C. P., Gonzalez, P. M., del Carmen Tobar, M., Poncela, J. M., Sanguino, F. J., Asensio, G., and Santos, E. (2008). Automatic method to segment the liver on multi-phase MRI. In *Proceedings of the Computer Assisted Radiology and Surgery 22nd International Congress and Exhibition, CARS 2008*, pages 404– 406, New York, EEUU. Springer Verlag. 2 citations in pages 16 and 48.

Ferrari, V., Megali, G., Troia, E., Pietrabissa, A., and Mosca, F. (2009). A 3-d mixed-reality system for stereoscopic visualization of medical dataset. *Biomedical Engineering, IEEE Transactions on*, 56(11):2627–2633. 2 citations in pages 80 and 99.

Feuerstein, M. (2007). *Augmented Reality in laparoscopic surgery. New concepts for intraoperative multimodal imaging*. PhD thesis, Fakultät für Informatik. Technische Universität München. 2 citations in pages 50 and 77.

Feuerstein, M., Mussack, T., Heining, S. M., and Navab, N. (2008). Intraoperative laparoscope augmentation for port placement and resection planning in minimally invasive liver resection. *Medical*

*Imaging, IEEE Transactions on*, 27(3):355–369. 2 citations in pages 79 and 99.

Fuertes, J. J., López-Mir, F., Naranjo, V., Villanueva, E., , Ortega, M., and Alcañiz, M. (2011). Augmented reality system for keyhole surgery performance and accuracy validation. In *PROC. of International Conference on Computer Graphics Theory and Applications*, pages 273–279. SciTePress – Science and Technology. 5 citations in pages 19, 82, 88, 93, and 99.

Garner, H. W., Kransdorf, M. J., Bancroft, L. W., Peterson, J. J., Berquist, T. H., and Murphey, M. D. (2009). Benign and malignant soft-tissue tumors: Posttreatment mr imaging 1. *Radiographics*, 29(1):119–134. One citation in page 2.

Gloger, O., Kuhn, J., Stanski, A., Volzke, H., and Puls, R. (2010). A fully automatic three-step liver segmentation method on lda-based probability maps for multiple contrast MR images. *Magnetic Resonance Imaging*, 28(6):882 – 897. 9 citations in pages 15, 16, 36, 48, 53, 55, 56, 63, and 103.

Gloger, O., Toennies, K., and Kuehn, J.-P. (2011). Fully automatic liver volumetry using 3D level set segmentation for differentiated liver tissue types in multiple contrast MR datasets. In *Image Analysis*, volume 6688, pages 512–523. Springer Berlin / Heidelberg. 2 citations in pages 16 and 48.

Goceri, E., Unlu, M., Guzelis, C., and Dicle, O. (2012). An automatic level set based liver segmentation from MRI data sets. In *3rd International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 192–197. 3 citations in pages 16, 48, and 49.

Gordon, A., Taylor, P., and Royston, C. (1993). *Practical Laparoscopy*. Blackwell Scientific. One citation in page 77.

Hamstra, D. A., Rehemtulla, A., and Ross, B. D. (2007). Diffusion magnetic resonance imaging: a biomarker for treatment response

in oncology. *Journal of clinical oncology*, 25(26):4104–4109. One citation in page 2.

Hartley, R. and Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press. One citation in page 86.

Hassan, A., Al-Ajami, R., Dashti, K., and Abdoelmoneum, M. (2011). Sixty-four multi-slice computed tomography and magnetic resonance imaging in evaluation of hepatic focal lesions. *The Egyptian Journal of Radiology and Nuclear Medicine*, 42(2):101 – 110. One citation in page 16.

Heimann, T., van Ginneken, B., Styner, M., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., Bello, F., Binnig, G., Bischof, H., Bornik, A., Cashman, P., Chi, Y., Cordova, A., Dawant, B., Fidrich, M., Furst, J., Furukawa, D., Grenacher, L., Hornegger, J., Kainmuller, D., Kitney, R., Kobatake, H., Lamecker, H., Lange, T., Lee, J., Lennon, B., Li, R., Li, S., Meinzer, H.-P., Nemeth, G., Raicu, D., Rau, A.-M., van Rikxoort, E., Rousson, M., Rusko, L., Saddi, K., Schmidt, G., Seghers, D., Shimizu, A., Slagmolen, P., Sorantin, E., Soza, G., Susomboon, R., Waite, J., Wimmer, A., and Wolf, I. (2009). Comparison and evaluation of methods for liver segmentation from CT datasets. *IEEE Transactions on Medical Imaging*, 28(8):1251 –1265. 10 citations in pages 15, 47, 53, 54, 55, 63, 65, 68, 69, and 103.

Hoffman, J. M. (2001). New advances in brain tumor imaging. *Current opinion in oncology*, 13(3):148–153. One citation in page 2.

Hsu, C.-Y., Liu, C.-Y., and Chen, C.-M. (2008). Automatic segmentation of liver {PET} images. *Computerized Medical Imaging and Graphics*, 32(7):601 – 610. One citation in page 53.

Huang, M. Q., Pickup, S., Nelson, D. S., Qiao, H., Xu, H. N., Li, L. Z., Zhou, R., Delikatny, E. J., Poptani, H., and Glickson, J. D. (2008). Monitoring response to chemotherapy of non-hodgkin's lymphoma

xenografts by t2-weighted and diffusion-weighted mri. *NMR in Biomedicine*, 21(10):1021–1029. One citation in page 2.

Hunt, S. J., Yu, W., Weintraub, J., Prince, M. R., and Kothary, N. (2009). Radiologic monitoring of hepatocellular carcinoma tumor viability after transhepatic arterial chemoembolization: estimating the accuracy of contrast-enhanced cross-sectional imaging with histopathologic correlation. *Journal of Vascular and Interventional Radiology*, 20(1):30–38. One citation in page 2.

Hussain, S. M. and Semelka, R. C. (2005). Hepatic imaging: Comparison of modalities. *Radiologic Clinics of North America*, 43(5):929–947. One citation in page 16.

Ibáñez, L., Schroeder, W., Ng, L., and Cates, J. (2003). *The ITK Software Guide*. Kitware. 3 citations in pages 59, 83, and 94.

Ji, H., He, J., Yang, X., Deklerck, R., and Cornelis, J. (2013). Acm-based automatic liver segmentation from 3-d ct images by combining multiple atlases and improved mean-shift techniques. *Biomedical and Health Informatics, IEEE Journal of*, 17(3):690–698. 5 citations in pages 53, 54, 63, 68, and 70.

Kainmüller, D., Lange, T., and Lamecker, H. (2007). Shape constrained automatic segmentation of the liver based on a heuristic intensity model. 6 citations in pages 53, 54, 63, 68, 69, and 70.

Kawamata, T., Iseki, H., Shibasaki, T., and Hori, T. (2002). Endoscopic augmented reality navigation system for endonasal transsphenoidal surgery to treat pituitary tumors: technical note. *Neurosurgery*, 50(6):1393–1397. 2 citations in pages 82 and 99.

Kersten-Oertel, M., Jannin, P., and Collins, D. L. (2012). Dvv: a taxonomy for mixed reality visualization in image guided surgery. *Visualization and Computer Graphics, IEEE Transactions on*, 18(2):332–352. 2 citations in pages 78 and 99.

Kim, P., Lee, Y., Jung, Y., Cho, J., and Kim, M. (2007). Liver extraction in the abdominal CT image by watershed segmentation algorithm. In *PROC. of World Congress on Medical Physics and Biomedical Engineering*, volume 14, pages 2563–2566. Springer Berlin Heidelberg. One citation in page 18.

King, B. W., Reisner, L., Klein, M. D., Auner, G. W., Pandya, A. K., et al. (2007). Registered, sensor-integrated virtual reality for surgical applications. In *Virtual Reality Conference, 2007. VR'07. IEEE*, pages 277–278. IEEE. 2 citations in pages 82 and 99.

Kramer, M. and Feinstien, A. (1981). The biostatistics of concordance. *Clinical biostatistics*, January:111–123. One citation in page 38.

Kum, C.-K., Eypasch, E., Aljaziri, A., and Troidl, H. (1996). Randomized comparison of pulmonary function after the ?french?and ?american?techniques of laparoscopic cholecystectomy. *British journal of surgery*, 83(7):938–941. One citation in page 80.

Lago, M. A., Martínez-Martínez, F., Rupérez, M. J., Monserrat, C., and Raya, M. A. (2012). A study about coefficients to estimate the error in biomechanical models used to virtually simulate the organ behaviors. In Westwood, J. D., Westwood, S. W., Felländer-Tsai, L., Haluck, R. S., Robb, R. A., Senger, S., and Vosburgh, K. G., editors, *MMVR*, volume 173 of *Studies in Health Technology and Informatics*, pages 250–256. IOS Press. 3 citations in pages 38, 49, and 68.

Lapeer, R. J., Tan, A. C., and Aldridge, R. (2002). A combined approach to 3D medical image segmentation using marker-based watersheds and active contours: the active watershed method. In *PROC. Medical Image Understanding and Analysis*, pages 165–168. One citation in page 18.

Li, C., Wang, X., Eberl, S., Fulham, M., Yin, Y., Chen, J., and Feng, D. (2013). A likelihood and local constraint level set model for liver tumor segmentation from ct volumes. *Biomedical Engineering,*

*IEEE Transactions on*, 60(10):2967–2977. 2 citations in pages 53 and 54.

Li, C., Wang, X., Xia, Y., Eberl, S., Yin, Y., and Feng, D. D. (2012). Automated pet-guided liver segmentation from low-contrast {CT} volumes using probabilistic atlas. *Computer Methods and Programs in Biomedicine*, 107(2):164 – 174. 2 citations in pages 15 and 53.

Lim, S.-J., Jeong, Y.-Y., and Ho, Y.-S. (2006). Automatic liver segmentation for volume measurement in CT images. *Journal of Visual Communication and Image Representation*, 17(4):860 – 875. One citation in page 15.

Ling, H., Zhou, S., Zheng, Y., Georgescu, B., Suehling, M., and Comaniciu, D. (2008). Hierarchical, learning-based automatic liver segmentation. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. One citation in page 53.

Lloréns, R., Naranjo, V., López, F., and Alcañiz, M. (2012). Jaw tissues segmentation in dental 3d ct images using fuzzy-connectedness and morphological processing. *Computer Methods and Programs in Biomedicine*, 108(2):832 – 843. One citation in page 15.

Lopez, F. (2013). Hepaplan project. One citation in page 56.

López-Mir, F., Naranjo, V., Angulo, J., Villanueva, E., Alcañiz, M., and López-Celada, S. (2011). Aorta segmentation using the watershed algorithm for an augmented reality system in laparoscopic surgery. In *18th IEEE International Conference on Image Processing*, pages 2649 –2652. 3 citations in pages 18, 83, and 94.

López-Mir, F., Naranjo, V., Fuertes, J. J., Alcañiz, M., Bueno, J., and Pareja, E. (2013). Design and validation of an augmented reality system for laparoscopic surgery in a real environment. *BioMed Research International*, 2013:1–12. One citation in page 18.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2):431–441. One citation in page 88.

Martín-Gutiérrez, J., Saorín, J. L., Contero, M., Alcañiz, M., Pérez-López, D. C., and Ortega, M. (2010). Design and validation of an augmented book for spatial abilities development in engineering students. *Computers & Graphics*, 34(1):77–91. One citation in page 86.

Martínez-Martínez, F., Lago, M., Rupérez, M., and Monserrat, C. (2013a). Analysis of several biomechanical models for the simulation of lamb liver behaviour using similarity coefficients from medical image. *Computer Methods in Biomechanics and Biomedical Engineering*, 16(7):747–757. One citation in page 99.

Martínez-Martínez, F., Rupérez, M., Martín-Guerrero, J., Monserrat, C., Lago, M., Pareja, E., Brugger, S., and López-Andújar, R. (2013b). Estimation of the elastic parameters of human liver biomechanical models by means of medical images and evolutionary computation. *Computer Methods and Programs in Biomedicine*, 111(3):537 – 549. One citation in page 38.

Martínez-Martínez, F., Rupérez, M. J., Lago, M., López-Mir, F., Monserrat, C., and Alcañíz, M. (2010). Pneumoperitoneum technique simulation in laparoscopic surgery on lamb liver samples and 3d reconstruction. *Studies in health technology and informatics*, 163:348–350. One citation in page 80.

Masoumi, H., Behrad, A., Pourmina, M. A., and Roosta, A. (2012). Automatic liver segmentation in MRI images using an iterative watershed algorithm and artificial neural network. *Biomedical Signal Processing and Control*, 7(5):429 – 437. 2 citations in pages 18 and 48.

McSherry, C. K. (1989). Cholecystectomy: the gold standard. *The American journal of surgery*, 158(3):174–178. One citation in page 80.

Meyer, F. (1994a). Minimum spanning forests for morphological segmentation. In *PROC. of Mathematical Morphology and its Applications to Image Processing*, pages 77–84. Kluwer. One citation in page 17.

Meyer, F. (1994b). Topographic distance and watershed lines. *Signal Processing*, 38:113–125. One citation in page 17.

Meyer, F. and Beucher, S. (1990). Morphological segmentation. *Journal of visual communication and image representation*, 1(1):21–46. 2 citations in pages 17 and 71.

Meyer, F. and Stawiaski, J. (2010). A stochastic evaluation of the contour strength. In *Pattern Recognition*, volume 6376 of *Lecture Notes in Computer Science*, pages 513–522. Springer Berlin Heidelberg. One citation in page 22.

Mharib, A., Ramli, A., Mashohor, S., and Mahmood, R. (2012). Survey on liver ct image segmentation methods. *Artificial Intelligence Review*, 37(2):83–95. 2 citations in pages 53 and 54.

Miccai:Organicers (2013). Liver segmentation challenge. One citation in page 65.

Milko, S., Samset, E., and Kadir, T. (2008). Segmentation of the liver in ultrasound: a dynamic texture approach. *International Journal of Computer Assisted Radiology and Surgery*, 3(1-2):143–150. 2 citations in pages 15 and 53.

Mischkowski, R. A., Zinser, M. J., Kübler, A. C., Krug, B., Seifert, U., and Zöller, J. E. (2006). Application of an augmented reality tool for maxillary positioning in orthognathic surgery–a feasibility study. *Journal of Cranio-Maxillofacial Surgery*, 34(8):478–483. 2 citations in pages 82 and 99.

Najman, L. and Schmitt, M. (1994). Watershed of a continuous function. *Signal Processing*, 38:99–112. One citation in page 17.

Neil Glossop, Z. W., Wedlake, C., Moore, J., and Peters, T. (2004). Augmented reality laser projection device for surgery. *Medicine Meets Virtual Reality 12: Building a Better You: the Next Tools for Medical Education, Diagnosis, and Care*, 98:104. 2 citations in pages 82 and 99.

Nicolau, S., Soler, L., Mutter, D., and Marescaux, J. (2011). Augmented reality in laparoscopic surgical oncology. *Surgical oncology*, 20(3):189–201. One citation in page 82.

Park, H., Bland, P., and Meyer, C. (2003). Construction of an abdominal probabilistic atlas and its application in segmentation. *Medical Imaging, IEEE Transactions on*, 22(4):483–492. 2 citations in pages 53 and 54.

Peng, J., Wang, Y., and Kong, D. (2014). Liver segmentation with constrained convex variational model. *Pattern Recognition Letters*, pages 81–88. 5 citations in pages 53, 54, 63, 68, and 70.

Phongkitkarun, S., Srianujata, T., and Jatchavala, J. (2009). Supplement value of magnetic resonance imaging in small hepatic lesion detected on routine computed tomography. *Journal of the Medical Association of Thailand*, 92(5):677–686. One citation in page 16.

Pintaric, T. (2003). An adaptive thresholding algorithm for the augmented reality toolkit. In *Augmented Reality Toolkit Workshop, 2003. IEEE International*, pages 71–. One citation in page 85.

Punia, R. and Singh, S. (2013). Review on machine learning techniques for automatic segmentation of liver images. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(4):666–670. 2 citations in pages 53 and 54.

Riederer, S. J. (1996). Recent technical advances in mr imaging of the abdomen. *Journal of Magnetic Resonance Imaging*, 6(5):822–832. One citation in page 2.

Roerdink, J. and Meijster, A. (2000). The watershed transform: Definitions, algorithms and parallelization strategies. *FUNDINF: Fundamenta Informatica*, 41. One citation in page 20.

Ruskó, L., Bekes, G., and Fidrich, M. (2009). Automatic segmentation of the liver from multi- and single-phase contrast-enhanced {CT} images. *Medical Image Analysis*, 13(6):871 – 882. Includes Special Section on Computational Biomechanics for Medicine. 8 citations in pages 53, 55, 56, 63, 68, 69, 70, and 103.

Ruska, L., Mateka, I., and Kriston, A. (2013). Virtual volume resection using multi-resolution triangular representation of b-spline surfaces. *Computer Methods and Programs in Biomedicine*, 111(2):315 – 329. One citation in page 15.

Samset, E., Schmalstieg, D., Vander Sloten, J., Freudenthal, A., Declerck, J., Casciaro, S., Rideng, Ø., and Gersak, B. (2008). Augmented reality in surgical procedures. In *Electronic Imaging 2008*, pages 68060K–68060K. International Society for Optics and Photonics. One citation in page 78.

Scheuering, M., Schenk, A., Schneider, A., Preim, B., and Greiner, G. (2003). Intraoperative augmented reality for minimally invasive liver interventions. In *Medical Imaging 2003*, pages 407–417. International Society for Optics and Photonics. 3 citations in pages 79, 82, and 99.

Schönfelder, C., Stark, T., Kahrs, L., Raczkowsky, J., Wörn, H., Belherazem, N., Müller, B., and Gutt, C. (2008). Port visualization for laparoscopic surgery-setup and first intraoperative evaluation. *International Journal of Computer Assisted Radiology and Surgery*, 3(1):141–142. 2 citations in pages 82 and 99.

Serra, J. (1983). *Image Analysis and Mathematical Morphology*. Academic Press, Inc., Orlando, FL, USA. One citation in page 58.

Shevchenko, N., Seidl, B., Schwaiger, J., Markert, M., and Lueth, T. (2010). Mimed liver: A planning system for liver surgery. In *Engineering in Medicine and Biology Society*, pages 1882–1885. One citation in page 15.

Shuhaiber, J. H. (2004). Augmented reality in surgery. *Archives of surgery*, 139(2):170–174. One citation in page 78.

Simitopoulos, D. and Kosaka, A. (2001). An augmented reality system for surgical navigation. In *Int. Conf. Augmented, Virtual Environments and 3D Imaging*, pages 152–157. 2 citations in pages 82 and 99.

Simon, G., Fitzgibbon, A., and Zisserman, A. (2000). Markerless tracking using planar structures in the scene. In *Augmented Reality, 2000. (ISAR 2000). Proceedings. IEEE and ACM International Symposium on*, pages 120–128. One citation in page 87.

Simpson, A., Dumpuri, P., Jarnagin, W., and Miga, M. (2012). Model-assisted image-guided liver surgery using sparse intraoperative data. In Payan, Y., editor, *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, volume 11, pages 7–40. Springer Berlin Heidelberg. One citation in page 15.

Soille, P. (2003). *Morphological Image Analysis: Principles and Applications*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2 edition. 5 citations in pages 20, 21, 29, 58, and 59.

Soler, L., Delingette, H., Malandain, G., Montagnat, J., Ayache, N., Koehl, C., Dourthe, O., Malassagne, B., Smith, M., Mutter, D., and Marescaux, J. (2001). Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery. *Computer Aided Surgery*, 6(3):131–142. 2 citations in pages 53 and 55.

Sorensen, A. G. (2006). Magnetic resonance as a cancer imaging biomarker. *Journal of clinical oncology*, 24(20):3274–3281. One citation in page 2.

T. Heimann, H.P. Meinzer, I. W. (2007). A statistical deformable model for the segmentation of liver ct volumes. 6 citations in pages 53, 54, 63, 68, 69, and 70.

Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C., et al. (2000). New guidelines to evaluate the response to treatment in solid tumors. *Journal of the National Cancer Institute*, 92(3):205–216. One citation in page 2.

Traboulsee, A. and Li, D. K. (2008). Conventional mr imaging. *Neuroimaging clinics of North America*, 18(4):651–673. One citation in page 2.

Vachier, C. and Meyer, F. (2005). The viscous watershed transform. *J. Math. Imaging Vis.*, 22:251–267. One citation in page 17.

Vincent, L. Soille, P. (1991). Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE PAMI, 1991*, 13(6):583–598. One citation in page 20.

Vogt, S., Khamene, A., and Sauer, F. (2006). Reality augmentation for medical procedures: System architecture, single camera marker tracking, and system evaluation. *International Journal of Computer Vision*, 70(2):179–190. 2 citations in pages 82 and 99.

Volonte, F., Bucher, P., Pugin, F., Carecchio, A., Sugimoto, M., Ratib, O., and Morel, P. (2010). Mixed reality for laparoscopic distal pancreatic resection. *International Journal of Computer Assisted Radiology and Surgery*, 5(1):122–130. 2 citations in pages 80 and 99.

Wang, G., Zhang, S., Li, F., and Gu, L. (2013). A new segmentation framework based on sparse shape composition in liver surgery planning system. *Medical Physics*, 40(5):–. 5 citations in pages 53, 54, 55, 68, and 70.

Yang, X., Yu, H. C., Choi, Y., Lee, W., Wang, B., Yang, J., Hwang, H., Kim, J. H., Song, J., Cho, B. H., and You, H. (2014). A

hybrid semi-automatic method for liver segmentation based on level-set methods using multiple seed points. *Computer Methods and Programs in Biomedicine*, 113(1):69 – 79. 7 citations in pages 15, 53, 54, 55, 63, 68, and 70.

Yuan, Z., Wang, Y., Yang, J., and Liu, Y. (2010). A novel automatic liver segmentation technique for MR images. In *International Congress on Image and Signal Processing*, volume 3, pages 1282 –1286. One citation in page 16.

Zahn, C. T. and Roskies, R. Z. (1972). Fourier descriptors for plane closed curves. *IEEE Trans. Comput.*, 21(3):269–281. 2 citations in pages 35 and 63.

Zhang, Z. (2000). A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(11):1330–1334. One citation in page 83.