

# Improved Image Guidance in TACE Procedures

Improved Image Guidance in TACE Procedures

Pierre Ambrosini

Pierre Ambrosini



# Improved Image Guidance in TACE Procedures

Verbeterde beeldgeleiding in TACE interventies

## Thesis

to obtain the degree of Doctor from the  
Erasmus University Rotterdam  
by command of the  
rector magnificus

Prof.dr. R.C.M.E. Engels

and in accordance with the decision of the Doctorate Board.

The public defense shall be held on  
the 19<sup>th</sup> of February 2019 at 13.30 hours

by

**Pierre Ambrosini**

born in Oullins, France

**Erasmus University Rotterdam**





## **Doctoral Committee**

Promotor:	<b>Prof.dr. W.J. Niessen</b>
Other members:	<b>Prof.dr.ir. A.F.W. van der Steen</b> <b>Prof. H. Kobeiter</b> <b>Prof. D. Stoyanov</b>
Copromotors:	<b>Dr.ir. T. van Walsum</b> <b>Dr. A. Moelker</b>

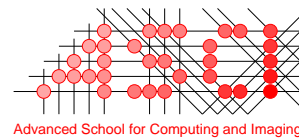


# Improved Image Guidance in TACE Procedures

Pierre Ambrosini



## Colophon



The research described in this thesis was carried out at the Erasmus MC – University Medical Center Rotterdam (the Netherlands), under the auspices of the Advanced School for Computing and Imaging (ASCI): dissertation series number 403.

# PHILIPS

This research has been funded by Philips Healthcare, Image Guided Therapy Systems Innovation (Best, the Netherlands)

It has also been supported by ITEA project 13031, Benefit.

The printing of this thesis was financially supported by the Department of Radiology and Nuclear Medicine of Erasmus MC University Medical Center Rotterdam, the Erasmus University Rotterdam (the Netherlands), and the ASCI graduate school.

Copyright © 2018 by Pierre Ambrosini. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the author.

ISBN: 978-94-6323-480-1

Printed by Gildeprint.



---

# Contents

---

<b>Colophon</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Hepatocellular carcinoma and treatment . . . . .	1
1.2 Minimally invasive interventions . . . . .	2
1.3 Transcatheter arterial chemoembolization . . . . .	5
1.4 Purpose and content of this thesis . . . . .	8
<b>2 Continuous Roadmapping in Liver TACE Procedures Using 2D-3D Catheter-based Registration</b>	<b>11</b>
2.1 Introduction . . . . .	12
2.2 Method . . . . .	13
2.3 Experiments . . . . .	20
2.4 Results . . . . .	25
2.5 Discussion and conclusion . . . . .	26
<b>3 A Hidden Markov Model for 3D Catheter Tip Tracking with 2D X-ray Catheterization Sequence and 3D Rotational Angiography</b>	<b>33</b>
3.1 Introduction . . . . .	34
3.2 Method . . . . .	36
3.3 Experiments . . . . .	43
3.4 Results . . . . .	47
3.5 Discussion and conclusion . . . . .	51
<b>4 Fully Automatic and Real-Time Catheter Segmentation in X-Ray Fluoroscopy</b>	<b>57</b>
4.1 Introduction . . . . .	58
4.2 Method . . . . .	59
4.3 Experiments and results . . . . .	61
4.4 Discussion and conclusion . . . . .	62
<b>5 Fast Prospective Detection of Contrast Inflow in X-ray Angiograms with Convolutional Neural Network and Recurrent Neural Network</b>	<b>65</b>
5.1 Introduction . . . . .	66
5.2 Methods . . . . .	67
5.3 Experiments . . . . .	69



5.4 Results and discussion . . . . .	70
<b>6 Summary and General Discussion</b>	<b>73</b>
6.1 Summary . . . . .	73
6.2 Challenges towards use in clinical practice and future studies . . . . .	75
<b>Bibliography</b>	<b>77</b>
<b>Samenvatting</b>	<b>85</b>
<b>PhD Portfolio</b>	<b>91</b>
<b>Publications</b>	<b>93</b>
<b>Curriculum Vitae</b>	<b>95</b>



# Introduction

---

## 1.1 Hepatocellular carcinoma and treatment

According to the 2015 estimates of the World Health Organization, cancer is responsible of 15.5% of total deaths, while liver cancer is the second most common cause of death among all cancers [100]. In 2012, liver cancer was the fifth most common cancer in men and the ninth in women, and half of the cases and deaths were estimated to occur in China [23]. Hepatocellular carcinoma (HCC) represents about 80% of liver cancer cases [90]. The principal cause of HCC is chronic liver disease which arises in developing countries from mostly Hepatitis B Virus or Hepatitis C Virus and, in developed countries, from cirrhosis due to non-alcoholic fatty liver disease and alcohol abuse [24, 84]. HCC unfortunately has a poor prognosis. Different staging systems have been proposed to provide a clinical classification of HCC [21] and different treatments are possible after the diagnosis of the disease stage, tumor size and spreading of HCC. The Barcelona Clinic Liver Cancer (BCLC) treatment scheme is endorsed by the European Association for the Study of the Liver and European Organisation for Research and Treatment of Cancer (EASL-EORTC) and the American Association for the Study of Liver Diseases (AASLD). Based on the results of several clinical studies, this methodology enables the physician to choose the right treatment for the patient (Fig. 1.1). Curative treatment is possible at early stages of the disease. Resection, liver transplantation, radio frequency ablation (RFA) and percutaneous ethanol injection have resulted in a 5-year survival in more than 50% of the cases [24]. When the disease is not curable, at intermediate stage, untreated patients present a median survival of 16 months. Transcatheter arterial chemoembolization (TACE) extends the survival of the patients to a median of up to 19-20 months [21]. It has also recently been shown to be as effective as resection and RFA procedure with single-nodule HCC of 3cm or smaller in terms of 5-year overall survival [44, 99]. In 2008, the effectiveness of treatment with chemotherapeutic agent sorafenib has been studied and a median survival of 10.7 months for advanced stage HCC has been demonstrated [52].

In the context of this thesis, we will focus on improvements in image guidance for the minimally invasive TACE procedure. Before describing the procedure in detail, minimally invasive interventions, medical images and image guidance are introduced.

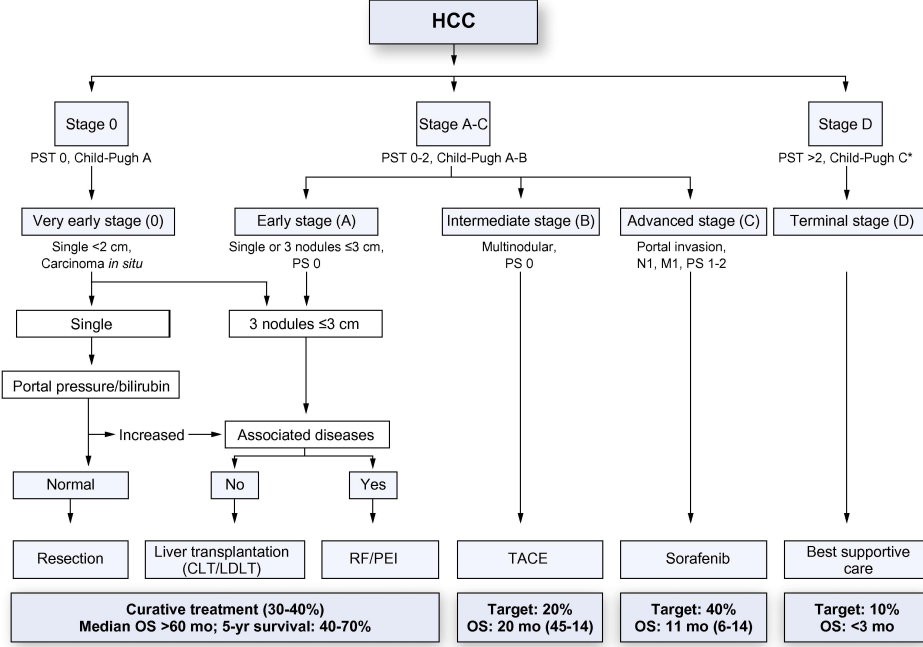


Figure 1.1: BCLC staging system and treatment strategy (Figure from [21]).

## 1.2 Minimally invasive interventions

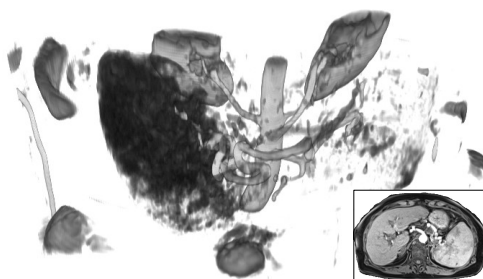
Minimally invasive interventions aim to limit harm during the procedure to surrounding anatomical structures of the region of interest. They are an alternative of open-surgery for many applications. Access to organs and region of interest is performed via small incisions, using thin instruments such as needles, catheters and guidewires. Endoscopy, percutaneous procedures with catheterization, laparoscopy and arthroscopy are examples of minimally invasive interventions. They are usually preferred over conventional surgery because of the benefits for the patient and the society: less complications and infections [28], shorter recovery time and ultimately lower healthcare costs [20]. To plan the interventions and guide instruments through the patient's body with limited or no direct eyesight, different medical image modalities can be used.

### 1.2.1 Medical images

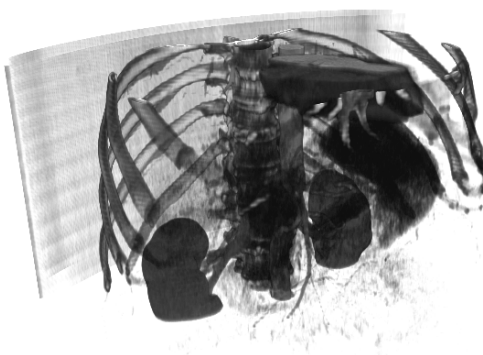
Medical imaging modalities are visualizing anatomy and physiology from microscopic scale to cellular scale and up to full body images. They are used by physicians for disease diagnosis, prognosis, therapy planning, guidance and research. In this manuscript, we focus only on the radiological imaging modalities that are commonly used for the TACE procedure:



- **Magnetic Resonance Imaging (MRI)** provides a non-invasive 3D image acquisition technique that uses strong magnetic fields. Soft tissue contrast is superior compared to the other imaging modalities. One disadvantage is that patients with metal parts inside the body cannot or not reliably be examined. Also, some patients may not be able to lie and stay immobile in a narrow confined space with loud noise for a long time due to claustrophobia. An MR Angiography (MRA) is an MRI where the blood vessels are visible using contrast agents usually injected into a vein in the arm or the hand [53] (Fig. 1.2). It is also possible to enhance the vessel signal without contrast agent [32].
- **Computed Tomography (CT)** is a 3D image reconstructed from multiple X-ray projection images. This imaging modality uses ionizing radiation. CT images provide good contrast for bony structures. A CT Angiography (CTA) or contrast enhanced CT is a CT with contrast agent injected intravenously for increasing the visibility of the vasculature (Fig. 1.3). Cone Beam Computed Tomography (CBCT) and 3D Rotational Angiography (3DRA) (Fig. 1.6 c) use the same principle as in CT/CTA, but with a C-arm so that it can be used in interventions (Fig. 1.4). If contrast agent is used to enhance blood vessels, it can be locally administered for the purpose of the intervention. For example in a liver catheterization procedure, contrast is injected directly in the liver hepatic artery via a catheter to enhance only liver vasculature. This usually results in a high visibility of the blood vessels of interest. Dual Phase CBCT (DP-CBCT) is an acquisition method using two C-arm rotation passes combined with contrast agent. The forward pass highlights the vasculature and the backward pass shows the accumulation of the contrast agent in the tumor.
- **2D X-ray imaging:** 2D X-ray such as fluoroscopic images can be acquired with a C-arm in intervention room with low radiation exposure. Bony structures and radio-opaque structures inside the body such as specific catheters or guidewires are visible in fluoroscopy. A large field-of-view is possible to acquire for example an image of the complete abdomen. Owing to its projective nature, fluoroscopy does not provide depth information (Fig. 1.5). An angiography is a series of 2D X-ray images acquired with contrast agent injection in order to enhance the blood vessels (Fig. 1.5). The patient may hold his or her breath during the acquisition to avoid motion artefacts. A digital subtraction angiography (DSA) is an angiographic imaging method in which both a contrast and non-contrast image are acquired. Background removal is achieved by subtracting the non-contrast image in order to have ideally only the contrast enhanced vessels visible (Fig. 1.5).
- **Ultrasound images:** They are non-invasive and acquired in 2D or 3D via an ultrasound probe that can transmit and receive acoustic waves. The probe is manipulated by the physician during the procedure (Fig. 1.6 a). The manipulation is not straightforward and necessitates training. The acquisition is fast for real-time use but the small field-of-view, the speckles and the difficulty to see behind air and bony structures make ultrasound images quite challenging to use.



**Figure 1.2:** Abdominal MRA with the aorta visible. One 2D slice of the MRA is shown at the bottom-right.

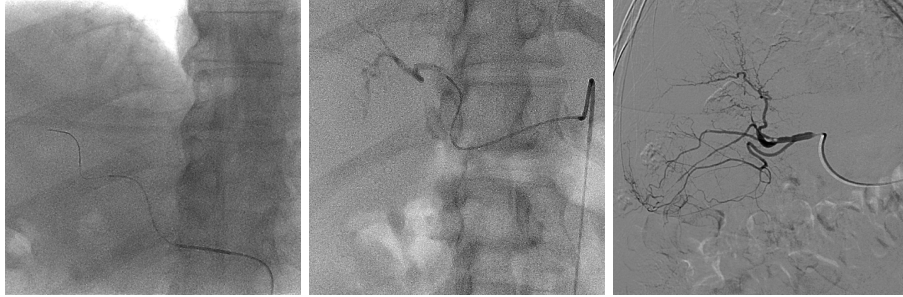


**Figure 1.3:** CTA image with a threshold on more radio-opaque tissue.



**Figure 1.4:** Angiographic C-arm system: Xper Allura, Philips Healthcare, Best, The Netherlands.





**Figure 1.5:** Fluoroscopy with a catheter, micro-catheter and guidewire (left), angiography (middle) and DSA (right).

### 1.2.2 Image guidance in intervention

Minimally invasive interventions are typically performed under image guidance, as direct visualization of the area to be treated is not possible. In most cases, physicians move their instruments inside the patient's body enabled by intra-operative medical imaging modality. X-ray and ultrasound modalities are suitable for image guided procedures as both modalities can be used in a real-time setting. Image guided interventions are common in several clinical areas such as cardiac, abdominal, neuro or orthopedic interventions [18].

Image guidance can be improved with more sophisticated methods combining multiple images and modalities, and other measuring devices. For example, roadmapping by visualizing 3D pre-operative images with 2D intra-operative images enables a better visualization of complex structures like the vasculature. In neuro-navigation where there is no cardiac or respiratory motion, the visualization with combination of 2D and 3D image is particularly appropriate [80]. Optical and electro-magnetics position trackers can facilitate navigation and automate image fusion providing more accurate localization of the instruments or the regions of interest outside and inside the body [26]. As an example, if an ultrasound probe is optically tracked, the ultrasound image can be aligned continuously in the image space of other modalities like 3D pre-operative images [51]. All these approaches to some extent turn traditional image guidance into navigation. These approaches rely on automatic computing for the tracking of the region of interest, fusion of the different images and devices, visualization and robotization. Most of the automation is a combination of fusion/registration [59, 67, 87, 95] and segmentation [72] methods which have been and are still extensively studied in the medical imaging field.

## 1.3 Transcatheter arterial chemoembolization

TACE is a minimally invasive procedure performed to increase survival of patients with an HCC at intermediate BCLC stage. It has been introduced after the popularization of vascular catheterization procedures [31, 102]. In the liver, primary hepatic tumors are mainly fed via the hepatic arteries [14]. The goal of the TACE proce-

ture is to stop this feeding by performing an embolization and also try to kill the tumors with chemotherapy. Embolization consists of blocking the blood supply to the tumors with embolic particles [99]. In chemoembolization, the embolic particles are coated with cytotoxic agents with the aim of killing the tumors cells [76]. Nowadays, in some cases, TACE is replaced by transarterial radioembolization which is a similar procedure except that embolic particles are radio-active rather than being loaded with a chemotherapeutic agent [82]. To perform the embolization, a catheter and then a micro-catheter are introduced as close as possible to the tumor via the vasculature (Fig. 1.6). The embolic particles are then released into the tumor feeding vessels via the micro-catheter. There are two ways to reach the hepatic artery with the catheter: via the femoral or the radial artery (Fig. 1.6 f). Radial catheterization showed lower risk of bleeding complications and also a shorter recovery time for the patient [41, 96]. In both cases, a puncture is performed with a needle in order to insert a catheter sheath. Guidewires, catheters and micro-catheters will be then introduced through the sheath to reach the hepatic artery and finally the tumor blood supply.

Here, we describe a typical TACE procedure using image guidance. To plan and choose the right treatment for HCC, a pre-operative CTA or MRA is acquired some weeks before the intervention (Fig. 1.2, 1.3). Tumors and vessels are enhanced with a contrast agent. This 3D image is used by the physicians to plan the embolization procedure. It is also used to localize the entrance of the hepatic artery with regard to the vertebrae position which will be useful during catheterization.

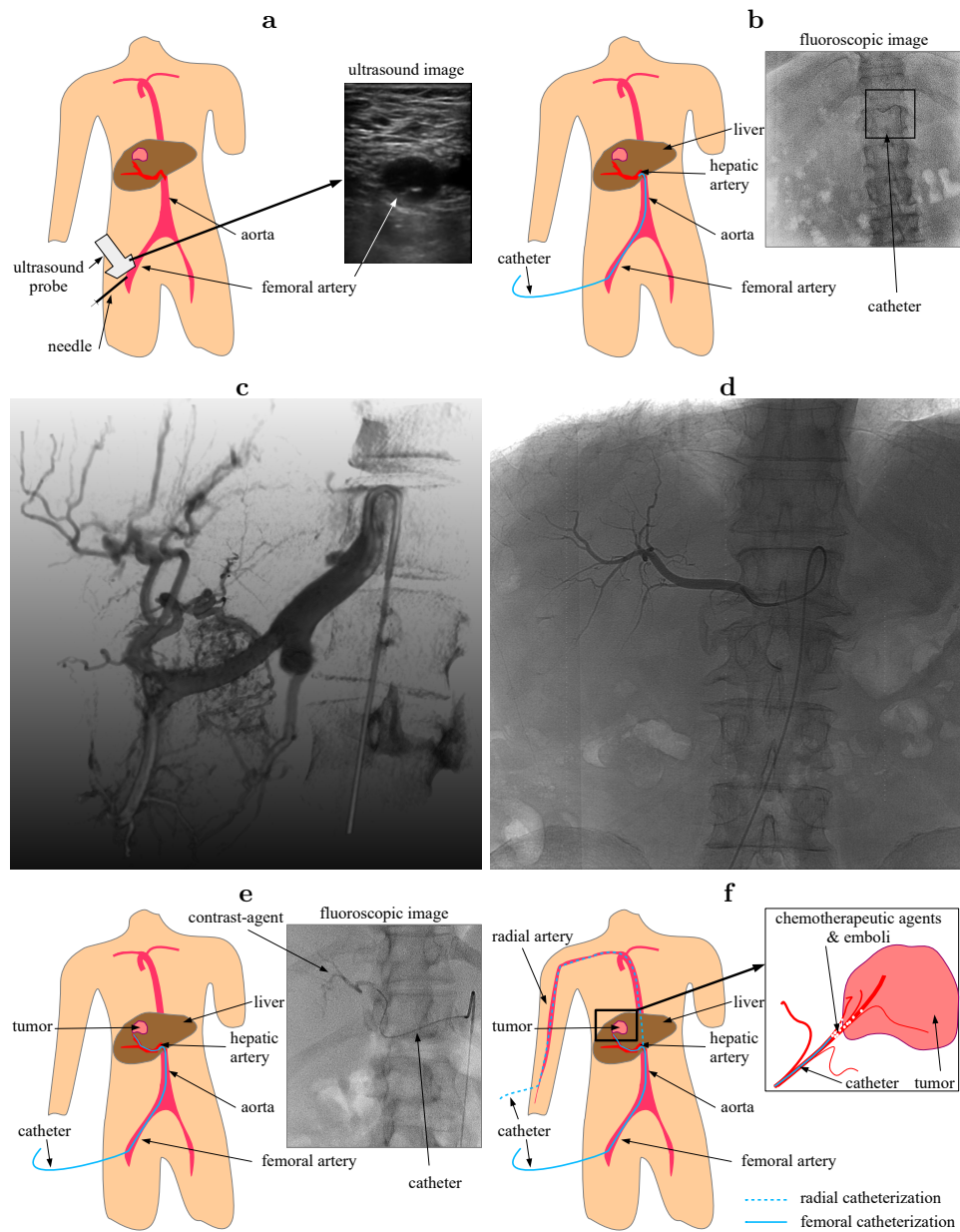
At the beginning of the intervention, a puncture is performed in the femoral or radial artery to insert a sheath, guided with an ultrasound probe to visualize the cross section of the artery and the needle (Fig. 1.6 a).

The procedure is performed in an angiography room where X-ray images can be acquired with an angiographic C-arm system (Fig. 1.4). Such a system can acquire 2D and 3D reconstructed X-ray images either with or without contrast agent to enhance blood vessels in the images.

The procedure is divided in five main steps:

- The catheter is guided using a thin and flexible guidewire into the aorta, then the coeliac trunk and finally the hepatic artery. The coeliac trunk originates from the aorta and divides into the splenic artery to the left side and the common hepatic artery to the right side. This step is performed using 2D fluoroscopic image guidance (Fig. 1.6 b). Sometimes the physician injects contrast agent to visualize the aorta at the end of the catheter in order to localize the coeliac trunk.
- Once the catheter is in the hepatic artery, a micro-catheter is inserted. The physician performs a 2D DSA and/or a 3DRA with contrast agent injected directly into the hepatic artery via the catheter. This enables good visibility of the vasculature in the liver (Fig. 1.6 c and d). 3DRA images allow to resolve overlapping vessels or foreshortening problems that can occur when interpreting 2D DSA images [50, 63]. The 2D or 3D enhanced vessel tree is used as a roadmap to have a precise idea of where to embolize and so where to bring the micro-catheter.





**Figure 1.6:** Main steps during TACE procedure with femoral catheterization. A puncture is performed in the femoral artery (a). The catheter is inserted into the hepatic artery guided with 2D fluoroscopic images (b). A 3DRA (c) or a 2D angiography (d) is acquired to visualize the vessel tree. Then, the micro-catheter is moved to the tumor vessel feeder (e). Finally, emboli with chemotherapeutic agent are injected (f). The alternative radial catheterization is depicted with the blue dotted line (f).

- With the roadmap in mind, the physician guides the micro-catheter to the embolization point with a guidewire through the vasculature (Fig. 1.5). In order to help the guidance of the micro-catheter, the physician may use contrast agent to enhance the blood vessels (Fig. 1.6 e).
- When the micro-catheter is close to the tumor, embolic particles are injected (Fig. 1.6 f). Contrast agent injection is used to check the embolization whether forward flow is still present.
- Finally the micro-catheter is pulled off and an angiography is generally acquired to check if all the feeding vessels are blocked as planned.

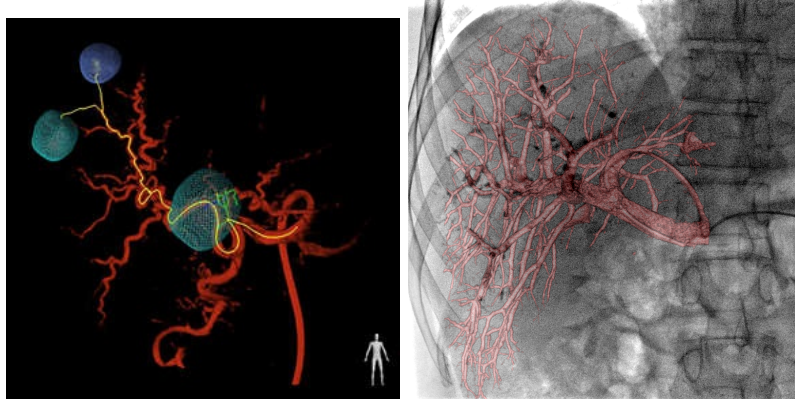
The first challenge for the physicians during the TACE procedure lies in the correct selection of the vessels feeding the tumor. The selected vessels should be as close as possible to the tumor to avoid compromising blood supply to healthy tissue. The common way to select the vessels is to perform 2D DSA but due to the lack of depth information, complex vasculatures appearing with multiple vessel overlappings are difficult to understand. The recent use of 3DRA/CBCT provides a better understanding and visibility of the vasculature because of the third dimension [10, 50, 63]. It has been also shown that small tumors have been detected only with CBCT [66] and DP-CBCT [103]. Finally, studies demonstrated that the use of tumor-feeder detection software using CBCT image helped vessel detection in small HCC embolization [65] and could reduce the number of total image acquisitions and the overall procedure time while keeping a comparable treatment efficacy compared to a procedure without software assistance [40] (Fig. 1.7). A navigated intervention would benefit from integrated feeder information.

The second challenge is the navigation of the catheters and the guidewires inside the patient's body. As the navigation is guided with the help of 2D fluoroscopic images, the longer the navigation is the more the patient and the physician are exposed to radiation. This task is dependent on the vasculature complexity and the skills of the physician. Some angiographic C-arm systems (such as Xper Allura, Philips Healthcare, Best, The Netherlands) provide the possibility to continuously overlay the vasculature onto the 2D fluoroscopic images where the vasculature has been extracted from pre-operative 3D images or intra-operative 3DRA or CBCT (Fig. 1.7). The 2D vasculature overlay is taking into account the C-arm position but is not updated with regards to any deformation caused by patient, respiratory and catheter motion. Such motion can be large in the liver, hampering the static roadmap to be aligned correctly and, as a consequence, assistance to the physician is limited.

## 1.4 Purpose and content of this thesis

Purpose of the work in this thesis is to improve the image guidance in TACE procedures. More specifically, we intend to develop and evaluate technology that permits dynamic roadmapping based on a 3D model of the liver vasculature. In the context of liver vasculature catheterization, studies have been done to register 2D DSA images acquired during the intervention with 3D pre-operative CTA or MRA in order





**Figure 1.7:** Tumor-feeder detection software EmboGuide from Philips (left) and projected roadmap with Innova Vision from GE Healthcare (right).

to project a 3D static vessel roadmap [49, 60] on single-plane [30, 43] or bi-plane [42] X-ray images. This 3D pre-operative aligned vessel roadmap or intra-operative CBCT may be continuously aligned by correcting for the respiratory motion [5]. By also tracking the position of the guidewire and catheter tip, the instruments can then be visualized with respect to the 3D vasculature, providing improved visual feedback to the interventional radiologist [13]. Also, preliminary studies have been performed to provide a roadmap during the guidance of the catheter to reach the hepatic artery. They propose manual or automatized registration of the 2D intra-operative fluoroscopic images with the 3D aorta extracted from pre-operative CTA image [11, 93].

Inspired by these approaches, this thesis will present methods that may bring TACE from a purely standard image guided intervention to a navigated intervention with dynamic 2D and 3D roadmapping and possibly integrating feeder information into the navigation. The following chapters include methods that could give physicians a continuously aligned 3D vasculature roadmap during the catheterization with a tracking of the catheter inside the 3D blood vessel tree.

In the second chapter of this thesis, a feasibility study demonstrates that 3D vessels extracted from intra-operative liver 3DRA can be aligned and projected onto 2D X-ray fluoroscopic images using only the 2D catheter shape and position. A rigid 3D/2D registration method between 3D vessels and 2D catheter is proposed followed by an evaluation on clinical data.

The third chapter describes a fast probabilistic method that allows tracking the catheter tip inside the 3D vasculature and obtaining an overlay of the 3D vessels onto the 2D X-ray fluoroscopic images. A hidden Markov model is used to track over time the catheter tip enabling a probability map of the tip position in the 3D vessels.

In the fourth chapter a new catheter segmentation method in 2D fluoroscopic images is proposed. This method using convolutional neural networks (CNN) is fast and does not require any interaction by the physician. Combined with the methods proposed in the previous chapters, we obtain a continuous overlay of the 3D vessels with a real-time and fully-automatic computation.

The fifth chapter describes two approaches to automatically detect contrast inflow in 2D coronary X-ray angiographic sequences. The first approach uses CNN to detect frames with contrast and the second one proposes a vessel enhancement with layer separation followed by a level of contrast feature fed in a recurrent neural network (RNN). Automated contrast detection in X-ray frame is useful in order to know which methods to apply on the current frame: vessel or catheter extraction and registration.

Finally, in the last chapter we summarize all the methods in the thesis and discuss their benefits, drawbacks and future improvements for practical use.

---

# Continuous Roadmapping in Liver TACE Procedures Using 2D-3D Catheter-based Registration.

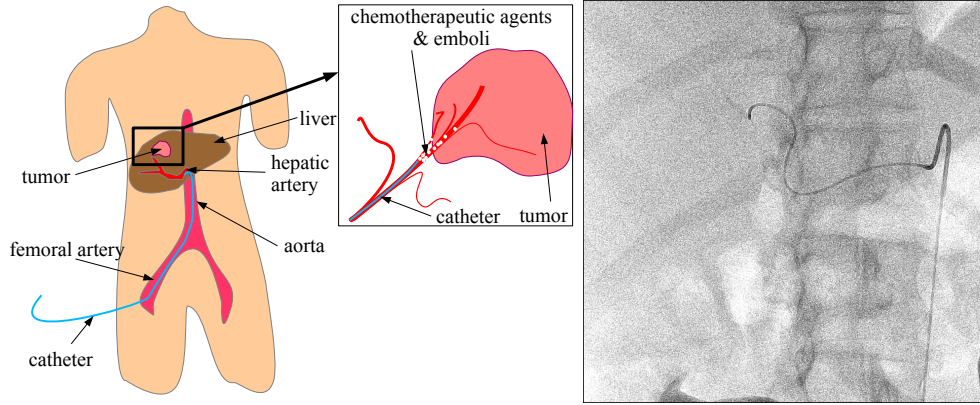
---

**Abstract** — *Purpose* Fusion of pre/peri-operative images and intra-operative images may add relevant information during image guided procedures. In abdominal procedures, respiratory motion changes the position of organs, and thus accurate image guidance requires a continuous update of the spatial alignment of the (pre/peri-operative) information with the organ position during the intervention.

*Methods* In this paper, we propose a method to register in real-time peri-operative 3D Rotational Angiography images (3DRA) to intra-operative single plane 2D fluoroscopic images for improved guidance in TACE interventions. The method uses the shape of 3D vessels extracted from the 3DRA and the 2D catheter shape extracted from fluoroscopy. First, the appropriate 3D vessel is selected from the complete vascular tree using a shape similarity metric. Subsequently, the catheter is registered to this vessel, and the 3DRA is visualized based on the registration results. The method is evaluated on simulated data and clinical data.

*Results* The first selected vessel, ranked with the shape similarity metric, is used more than 39% in the final registration and the second more than 21%. The median of the closest corresponding points distance between 2D angiography vessels and projected 3D vessels, is 4.7-5.4 mm when using the brute force optimizer and 5.2-6.6 mm when using the Powell optimizer.

*Conclusions* We present a catheter-based registration method to continuously fuse a 3DRA roadmap arterial tree onto 2D fluoroscopic images with an efficient shape similarity.



**Figure 2.1:** TACE intervention overview (left) and fluoroscopy example (right).

## 2.1 Introduction

Transcatheter Arterial Chemo Embolization (TACE) is a minimally invasive procedure to treat liver cancer (mostly hepatocellular carcinoma). In this procedure, a catheter is navigated towards a tumor via the femoral and hepatic artery, after which chemotherapeutic agents are injected. Currently, the interventionalist guides the catheter using single plane 2D X-ray (fluoroscopy), visualizing only the catheter (Fig. 2.1). Frequently, contrast is injected to visualize the arteries. Computed tomography angiography (CTA) or 3D Rotational Angiography (3DRA) are used pre/peri-operatively to visualize the tumor and feeding arteries. The navigation of the catheter using only 2D fluoroscopy is hampered by the inability to continuously visualize the arterial tree.

Purpose of our work is to integrate information of the vasculature from pre/peri-operative 3D images by fusing it with the intra-operative 2D X-ray images. Such an approach enables a continuous up-to-date roadmap and thus may improve the guidance during the procedure and consequently has the potential to reduce intervention time, radiation dose and contrast agent use.

2D-3D registration for improving image guidance has been studied in cardiac, cranial, abdominal and orthopedic procedures. An overview of 2D-3D registration methods is presented by Markelj et al. [60] and Liao et al. [49]. Following [60], 2D-3D registration methods can be classified as extrinsic, intrinsic and calibration-based. Extrinsic methods use markers to register and update the registration [69]. Usually objects visible on X-ray (e.g. small beads) are inserted close to the region-of-interest before 3D image acquisition. Intrinsic methods rely on anatomical structures such as bones or the vasculature and are generally intensity-, gradient-, or feature-based or a combination of them [60]. In abdominal interventions, the vasculature and catheters are mostly the only structures visible on 2D X-ray images that can be used for registration. In cardiac [8, 9, 61, 77, 81], cranial [35, 64, 91, 92] and abdominal [30, 42, 47] interventions, vessel-based registration have been used between pre- or peri-operative 3D/4D CTA (Computed Tomography Angiography), MRA (Mag-

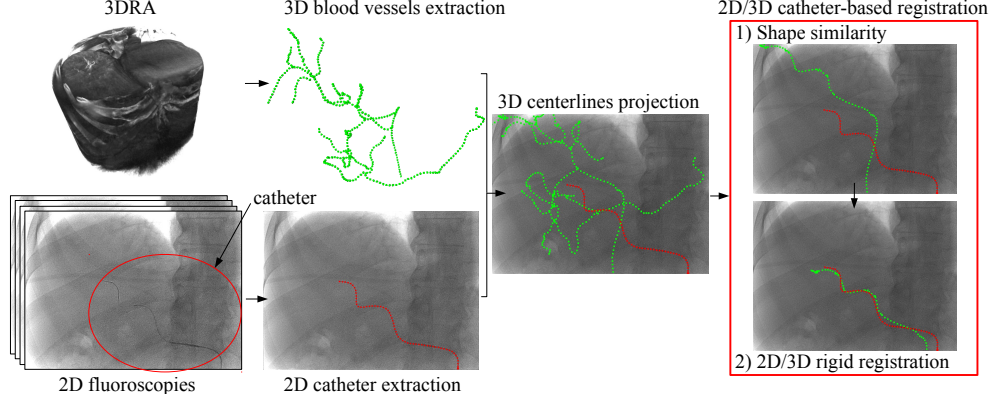


netic Resonance Angiography) or CBCT (Cone Beam Computed Tomography) and 2D DSA (Digital Subtraction Angiography) or 2D fluoroscopies. The X-ray acquisition can either be single plane or bi-plane. Rigid as well as non-rigid registration approaches for aligning the vessels from 3D/4D pre- or peri-operative images with those from DSA or fluoroscopy have been described. These approaches update the 3D vessels position with regard to the C-arm but do not enable a continuous roadmap of the 3D vessels because continuous contrast agent injection during the intervention would be harmful to the patient. Calibration-based methods can be used when the 3D peri-operative image and the 2D images are acquired with the same device. For example, if the 3D position of the C-arm is known accurately, it allows alignment of intra-operative 2D X-ray images with peri-operative 3D images. Atasoy et al. [5] and Ruijters et al. [80] use C-arm information to update the registration between peri-operative 3DRA (or CBCT) and 2D X-ray. This approach has been demonstrated to work accurately in cranial procedures with no head movement. Utilization in abdominal interventions, however, is hampered by the respiratory motion, which invalidates the initial alignment. Atasoy et al. [5] proposed a semi-automatic method to follow one moving region of interest selected by a physician during the intervention (a part of a catheter) and to update the registration with this information. The transformation model contains in-plane translation to correct for shifts caused by respiratory motion. In cardiac interventions, Ma et al. [58] used manual calibration-based methods to achieve an initial alignment and then used features such as diaphragm/heart border, tracheal bifurcation or the catheter to correct for breathing motion. Another method was proposed by Luan et al. [55] for oral cancer treatment. They track the catheter tip with an electromagnetic sensor, reconstruct the catheter path and then register it with a pre-operative image. Although tracking the 3D catheter tip tracking is valuable, breathing motion may hamper the reconstruction of the path in e.g. the abdomen. Unlike most of the other methods, our previous method [1] performs a 2D/3D catheter-based registration using a 3DRA and the complete catheter visible in the 2D X-ray images. It does not require 2D angiographic images nor user interaction for the initial alignment. However, computation times were not interactive, hampering interventional use.

The major contribution of our current work is to propose a method for generating an automatic continuous roadmap during abdominal catheterization using 2D/3D registration with single plane 2D X-ray images and peri-operative 3DRA. This paper is an extension of our previous work: the metric for alignment has been improved, the registration is faster, and the evaluation has been performed on a larger set of data, containing synthetic images, clinical images and additional evaluation metrics.

## 2.2 Method

The method is based on the registration of a 3D vessel tree with a 2D catheter shape. Therefore, in a pre-processing step, the arterial tree is extracted from the 3DRA image and the catheter shape position is determined from the single plane fluoroscopic images. The extraction of the vessel tree itself is relatively straightforward for high-contrast 3DRA images. The segmentation of the catheter in the fluoroscopic images,



**Figure 2.2:** Global overview: vessels/catheter extraction and 2D/3D registration.

albeit more challenging, has been subject of other studies [34, 68, 70, 71, 86, 101, 104]. These steps are not addressed in this paper.

Given the 3D vascular model and the 2D catheter centerline, the method consists of two steps (Fig. 2.2). First a shape similarity metric is used to find the vessel centerlines from the 3DRA that are most similar to the 2D catheter shape. Subsequently, a constrained 2D-3D registration is applied to find the corresponding rigid transformation between the 2D catheter and the 2D projections of the best ranked vessel centerlines. In the following, we first define our coordinate systems and transformations, then we describe each registration steps.

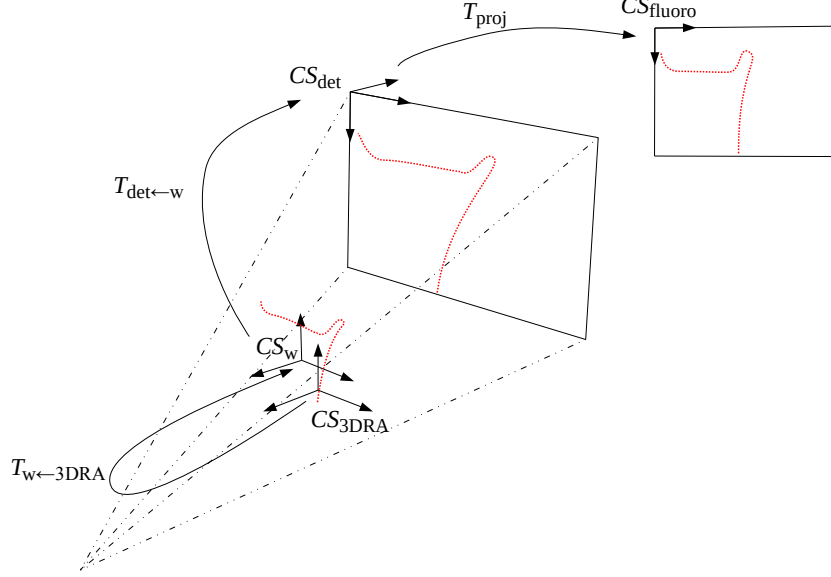
### 2.2.1 Definitions

We define the following coordinate systems (CS) for our setup in the intervention room (Fig. 2.3):

- $CS_w$ , denotes the world 3D CS, with the origin at the iso-center of the C-arm, and oriented along the C-arm in its default position
- $CS_{det}$ , the detector 3D CS (X-ray image plane)
- $CS_{fluoro}$ , is the 2D CS of the fluoroscopic image
- $CS_{3DRA}$ , 3D CS of the 3DRA

Accordingly, the following coordinate transformations are defined:

- $T_{det \leftarrow w}$ , transformation from  $CS_w$  to  $CS_{det}$
- $T_{proj}$ , cone-beam projection from  $CS_{det}$  to  $CS_{fluoro}$
- $T_{w \leftarrow 3DRA}$ , transformation that aligns the 3DRA space with the patient, from  $CS_{3DRA}$  to  $CS_w$



**Figure 2.3:** Coordinate systems and transformations of the C-arm space.

$T_{\text{det} \leftarrow w}$  and  $T_{\text{proj}}$  are known for each X-ray image because the geometry and orientation of the C-arm are known.  $T_{w \leftarrow 3\text{DRA}}$  is unknown and is the result of our registration.

With the projection function  $F_{\text{proj}}$  (in homogeneous coordinates):

$$F_{\text{proj}}(p_{3D}, T) = T_{\text{proj}} \cdot T_{\text{det} \leftarrow w} \cdot T \cdot p_{3D} \quad , \quad (2.1)$$

we have a 3D point in the 3DRA space,  $p_{\text{CS}_{3\text{DRA}}}$ , which can be projected on the fluoroscopic image space  $\text{CS}_{\text{fluoro}}$  using the following equation:

$$p_{\text{CS}_{\text{fluoro}}} = F_{\text{proj}}(p_{\text{CS}_{3\text{DRA}}}, T_{w \leftarrow 3\text{DRA}}) \quad .$$

The catheter centerline extracted from a 2D fluoroscopic image is defined as an ordered set of  $n_C$  points:

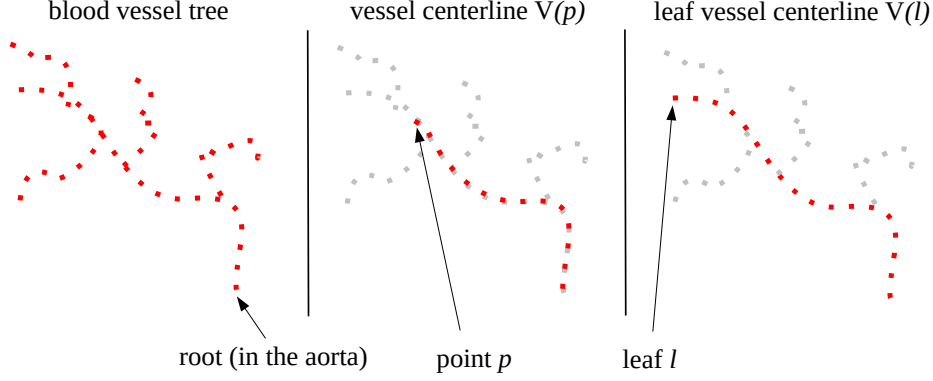
$$C_{2D} = \{c_1, c_2, \dots, c_i, \dots, c_{n_C}\} \quad ,$$

where  $c_i \in \mathbb{R}^2$  are 2D points at the center of the catheter in  $\text{CS}_{\text{fluoro}}$  and  $c_1$  denotes the tip of the catheter.

The blood vessel tree centerline extracted from the 3DRA is represented as a directed tree:

$$G_{3D} = (\mathcal{P}, \mathcal{E}) \quad ,$$

where  $\mathcal{P}$  is the set of 3D points on the centerlines of the vessels in  $\text{CS}_{3\text{DRA}}$ ,  $\mathcal{E}$  the set of directed edges between points. The root of  $G_{3D}$  is in the aorta and the branches of the tree are in the liver.



**Figure 2.4:** Terminology: blood vessel tree, vessel centerline and leaf vessel centerline.

We define a vessel centerline  $V(p)$  as an ordered set of points in  $G_{3D}$ , from any point  $p \in \mathcal{P}$  along the directed edges to the root (Fig. 2.4):

$$V(p) = \{p, p_1, p_2, \dots, p_i, \dots, p_{n_P}\} \quad ,$$

where  $p_i \in \mathbb{R}^3$  in  $CS_{3DRA}$  and  $p_{n_P}$  is the root of  $G_{3D}$ .

Similarly, we define the 2D projection of the 3D vessel centerline  $V(p)$ :

$$V_{proj,T}(p) = \{F_{proj}(p, T), F_{proj}(p_1, T), \dots, F_{proj}(p_i, T), \dots, F_{proj}(p_{n_P}, T)\} \quad .$$

Additionally, we define  $V_{proj,T}(p, u)$  with  $u \in [0, V_l]$ , a linearly interpolated version of the projected centerline, with  $V_l$  the length of the projected vessel centerline  $V(p)$ .

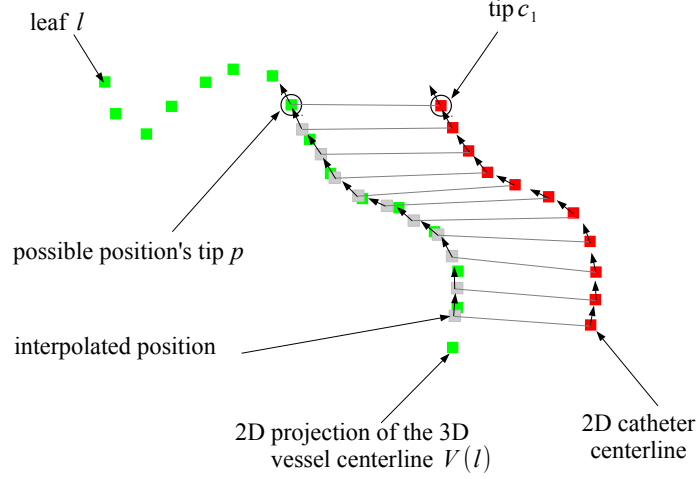
### 2.2.2 Shape-based vessel centerline selection

The registration is performed on a vessel centerline running from a leaf to the root (Fig. 2.4). Before performing the registration, the vessels that are the most likely to contain the catheter are selected. This selection is based on the shape similarity metric between the catheter and the projected 3D vessel. The metric quantifies the alignment of the tangent vectors of the catheter and the projected vessel. Therefore it is not sensitive to the distance between centerlines. The underlying assumption is that the orientation of the vasculature changes little between the 3DRA and the fluoroscopy acquisition, which is valid for our application. The shape similarity for a vessel from a point  $p \in \mathcal{P}$  is defined as:

$$S(p) = \int_0^{C_l} \vec{C}_{2D}(u) \cdot \vec{V}_{proj,I_4}(p, u) du \quad , \quad (2.2)$$

where  $C_l$  is the length of the 2D catheter,  $I_4$  the 4x4 identity matrix and  $\vec{C}_{2D}(u)$  (resp.  $\vec{V}_{proj,I_4}(p, u)$ ) the tangent of  $C_{2D}$  (resp.  $V_{proj,I_4}(p)$ ) at the position  $u$ .  $C_{2D}(0)$  is the





**Figure 2.5:** Discretized sum of dot products between tangents of catheter and the vessel centerline  $V(p)$  with  $p \in V(l)$ .

tip of the catheter and  $V_{proj, I_4}(p, 0)$  is the possible location of the tip in the tree  $G_{3D}$ .  $S(p) \in [0, C_l]$  with  $C_l$  denoting the maximum similarity. As the catheter centerline is represented as a set of points, the integral over  $S$  is approximated by summing the dot products over all catheter positions, thereby interpolating the corresponding vessel positions (Fig. 2.5).

In order to select the leaf vessel centerline for which the registration needs to be performed, for each leaf  $l$ , the maximum similarity over all points in  $V(l)$  is determined:

$$S_{max}(l) = \max_{p \in V(l)} S(p) \quad . \quad (2.3)$$

Based on the values of  $S_{max}$ , we selected the  $k$  leafs with largest  $S_{max}$  for the registration. When several leafs share the same common part with the catheter, only one is kept.

### 2.2.3 Rigid 2D/3D registration with forward projection

To register the 2D catheter with the vessel centerline, we need to find the rigid transform  $T_{w \leftarrow 3DRA}$  that yields the best match with the 2D catheter in  $CS_{fluoro}$ . We decompose the transformation as follows:

$$T_{w \leftarrow 3DRA} = T_{w \leftarrow det} \cdot T_{trans} \cdot T_{det \leftarrow w} \cdot T_{rot} \quad ,$$

where  $T_{rot}$  is a rotation matrix with three unknowns (Euler angles,  $\alpha, \beta$  and  $\gamma$ ),  $T_{trans}$  a translation matrix with three unknowns ( $x, y, z$ ) with the translations aligned in  $CS_{det}$ . A translation along the projection axis in  $CS_{det}$  will only have a very minor effect in the projection. We therefore exclude  $z$  from the registration parameters, leaving us with a five degrees of freedom transformation.

Our registration metric is the sum of distances between points on the catheter and closest points on the projected leaf vessel centerline. The catheter tip  $c_1$  thereto is matched to the closest point of the projected vessel  $V(l^{sel})$ , where  $l^{sel}$  is a leaf selected thanks to the shape similarity. The distance between the catheter tip and the vessel centerline  $V(l^{sel})$ , given a rigid transformation  $T$ , is given by:

$$D_1(l^{sel}, T) = \min_{p \in V(l^{sel})} \|c_1 - F_{proj}(p, T)\| \quad . \quad (2.4)$$

Each next point of the catheter is similarly matched with a point of the projected vessel. To ensure continuity of the vessel (and simultaneously reducing computation time), the search range is limited to only a few points proximal to the point closest to the previous catheter point. Thus, let  $p_{prev} \in V(l^{sel})$  be the point matched with  $c_{i-1}$ , then the distance to the subsequent catheter point  $c_i$  is defined as:

$$D(c_i, l^{sel}, T, p_{prev}) = \min_{p \in [p_{prev}, \dots, p_{prev+h}]} \|c_i - F_{proj}(p, T)\| \quad , \quad (2.5)$$

where  $[p_{prev}, \dots, p_{prev+h}]$  are the  $h + 1$  consecutive points in  $V(l^{sel})$ , starting at  $p_{prev}$ , and  $h$  is determined such that all points in that range are within a distance  $d_{max}$  of  $p_{prev}$ .

Given these definitions, the final registration metric  $M$  of our registration is a weighted sum of these distances (Fig. 2.6):

$$M(C_{2D}, l^{sel}, T) = D_1(l^{sel}, T) + \sum_{c_i \in [c_2, c_{n_C}]} W(\|c_i, c_1\|) \cdot D(c_i, l^{sel}, T, p_{prev}) \quad , \quad (2.6)$$

where  $W(x)$  is a weight function  $\in [0, 1]$  and  $\|c_i, c_1\|$  is the length of the catheter between  $c_1$  and  $c_i$ . As the registration accuracy close to the tip is more important than at the proximal part of the catheter, we use a weight to decrease the distance values that are further from the tip. We use a Gaussian with an offset:

$$W(x) = \lambda + (1 - \lambda) \cdot e^{-\frac{x^2}{2\sigma^2}} \quad , \quad (2.7)$$

where  $\sigma$  is a parameter to control how fast the weight decrease (Fig. 2.7).

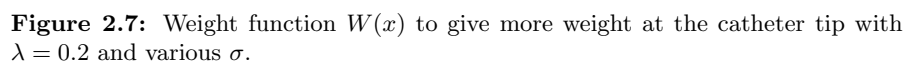
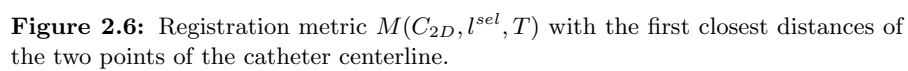
This metric  $M$  has two advantages: first, it is fast because we only look for the closest point in a specific neighbourhood; second, by only matching points that are locally connected the continuity of the vessel centerline is respected.

Lastly, the final transformation is the one with the smallest cumulative distance:

$$T_{w \leftarrow 3DRA} = \operatorname{argmin} TM(C_{2D}, l^{sel}, T) \quad , \quad (2.8)$$

where  $T$  represents the 5 degrees of freedom rigid transformation matrix.

Every selected leaf vessel centerline  $V(l^{sel})$  is registered and the pair  $(V(l^{best}), T_{w \leftarrow 3DRA})$  with the optimal similarity  $M$  is kept.



## 2.3 Experiments

We performed two series of experiments. In the first one, we used clinical data from TACE interventions. As we do not have a ground truth available in this data, we rather evaluate the registration on the alignment of the vessels distal to the catheter tip using angiography and we investigate the effect of varying parameter settings. In the second experiment, registration has been performed on the same clinical data, but with the catheter position simulated. The simulation allows us to have a ground truth for the catheter position. The error in the localization of the registered catheter tip position was used for evaluation.

### 2.3.1 Data

We retrospectively acquired anonymized data of 19 TACE interventions (Table 2.1). The 16 first sets were acquired in the Erasmus MC, University Medical Center, Rotterdam, the Netherlands, between 2012 and 2014 in two different intervention rooms with angiographic C-arm systems (Xper Allura, Philips Healthcare, Best, the Netherlands). The last 3 sets were acquired in the Hôpitaux Universitaires Henri Mondor, Créteil, Paris, France and the Ospedale di Circolo e Fondazione Macchi, Varese, Italy. For each intervention, we have a set of images consisting of one 3DRA image where the catheter was inside the hepatic artery, a set of fluoroscopic sequences with contrast agent and a set of Digital Subtraction Angiographies (DSA) (Fig. 2.8). In these sequences, both the catheter and a part of the vasculature distal to the catheter tip is visible (by using the contrast agent). For each sequence, we selected the image with most of the vasculature visible and we manually annotated both the 2D catheter centerline and the 2D vasculature centerlines. The 3D arterial tree from 3DRA is segmented with a semi-automatic method based on thresholding and skeletonization [83].

We divided our data in two different groups depending on the 3DRA acquisition: complete and incomplete acquisition. Incomplete acquisition occurs when the patient's liver is not aligned with the C-arm rotation iso-center. In that case, the aorta and the hepatic artery are not visible in the 3DRA which hampers the registration.

### 2.3.2 Implementation

The method described was implemented in C++ and run on a computer with a 3.4Ghz Intel Core i7. We set  $k$ , the number of selected leaf vessel centerline to register, to 5.

In order to minimize our metric, we evaluated two different optimizers: a brute force and the Powell optimizer [73]. The brute force is exhaustive and is more likely to find the global minimum, whereas Powell is faster but because of its local search is more likely to converge to local minimum.

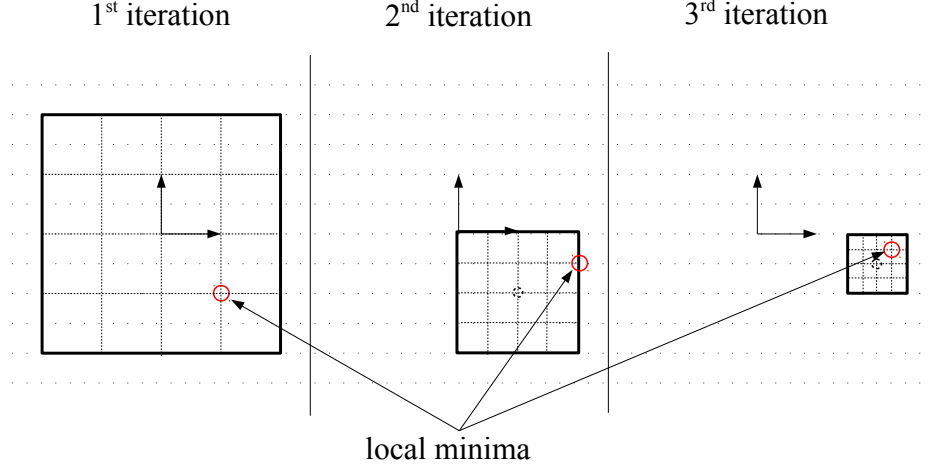
Our brute force optimizer has  $n = 7$  iterations and for each iteration  $i$ , the search space is centered at the minimum found in iteration  $i - 1$  with an interval size  $s_i = c * s_{i-1}$  whereby coefficient  $c$  was set to 0.5 (Fig. 2.9). Each dimension in the search space interval is subdivided in  $d = 7$  steps. We set the initial intervals  $s_0$  of



**Table 2.1:** Number of clinical data.

Patients	3DRA	Number of angios	Number of DSAs
01	complete	4	2
02	complete	2	3
03	complete	0	2
04	complete	1	4
05	complete	3	1
06	complete	2	2
07	complete	5	1
08	complete	2	3
09	incomplete	4	3
10	incomplete	2	3
11	incomplete	3	2
12	complete	2	2
13	complete	1	4
14	complete	0	2
15	complete	3	2
16	complete	3	5
17	complete	1	2
18	complete	1	0
19	complete	0	2
Total	-	39	45

**Figure 2.8:** Fluoroscopy with contrast agent (left) and DSA (right).



**Figure 2.9:** Brute force optimizer in a 2D space with  $n = 3$  iterations, reduction coefficient  $c = 0.5$ , number of steps  $d = 5$  and initial interval size  $x, y = \pm 2$ .

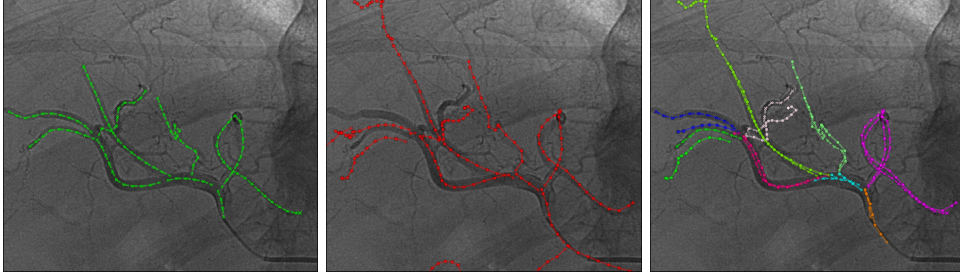
our brute force search to  $\pm 50$  mm for  $x$  and  $y$  and  $\pm 7^\circ$  for  $\alpha$ ,  $\beta$  and  $\gamma$ . These intervals are sufficiently large to capture breathing motion.

For the Powell optimizer, we use a two-stage approach. We first optimize the in-plane translation and subsequently use that translation to initialize the full 5 degrees of freedom registration.

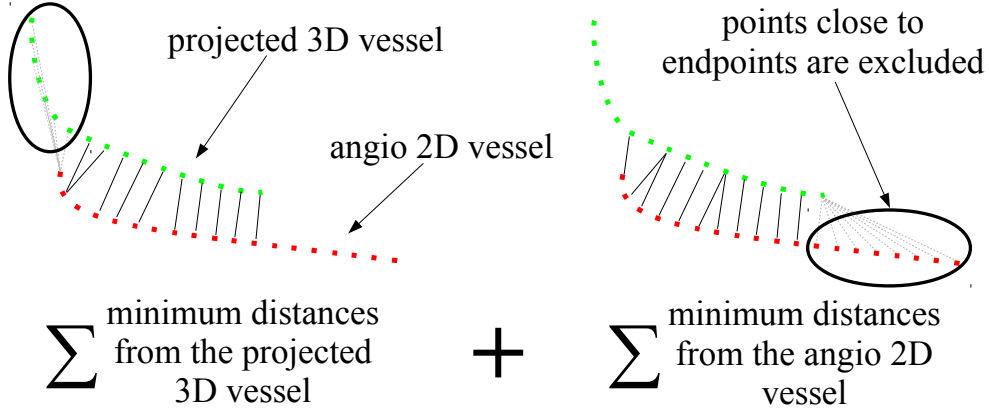
### 2.3.3 Clinical data and parameters optimization

In the first experiment, we investigated optimal parameter settings for the method, and evaluated how well the resulting registration aligns the vasculature distal to the catheter tip. To determine optimal parameter settings, and evaluate the effect of changing parameters, we applied the method with a large set of different settings ( $\lambda$ ,  $\sigma$  and  $d_{max}$ ) in a leave-one-out cross validation scheme (determine the optimal parameter values over the set containing all patients except the one on which the evaluation is done). We tested the following settings;  $\lambda$ : 0, 0.1, 0.2, 0.3;  $\sigma$ : 20, 40, 60, 80, 100;  $d_{max}$ : 10, 20, 30, 40, 50 mm. Because the 2D catheter centerline and 3D blood vessel tree are discretized we also investigated the effect of using different samplings: 1.5, 3 and 6 mm between each point. Finally, computation time is recorded.

As we do not have a ground truth for the registration, we used the vasculature visible on the angiographies as a reference. Thus, for validating, we compared how well the projected 3DRA matches the arteries visible in the projection images. To this end, we projected the 3DRA vasculature on the 2D image with the registered  $T_{w \leftarrow 3DRA}$  and we computed the closest corresponding points distance between the projected 3D vasculature and the 2D vasculature. The most relevant region for the roadmapping is the area close to the catheter tip, we therefore only evaluated in a circular region (3 cm radius) around the catheter tip. Before computing the distance, we manually labelled the vessels such that distances are computed between correspond-



**Figure 2.10:** 2D vasculature from contrast agent (left), 2D projection of 3DRA vasculature after manual registration (center) and manual paired vessels (right): same labels have same color vessels.



**Figure 2.11:** Closest corresponding points distance between paired vessels.

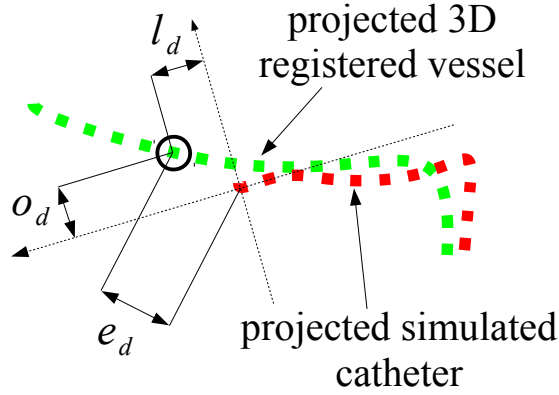
ing vessels (Fig. 2.10). To prevent bias in this assignment, the manual annotation was done without registration, thus only using the initial projection of the 3DRA. To aid in the annotation, the observer could manually register 3DRA and 2D vasculature by changing translations and rotations of the 3DRA. Vessels that can not be manually adequately linked were not used in the evaluation. Using the labelled corresponding vessels, we computed the closest corresponding points distance for each pair of vessels (excluding distances to endpoints). The distance was computed both from the 2D angio vessel and from its registered projected 3D vessel pair (Fig. 2.11). Our evaluation metric for one image is the average of distances over all pairs of vessels.

### 2.3.4 Clinical data with a simulated catheter

When evaluating the registration on patient data, no accurate ground truth registration is available. We therefore also evaluated our method on simulated data with a known ground truth. To stay as close as possible to the reality, our simulation was based on the clinical data (Table 2.1) where we used the 3D extracted vessel tree in the

**Table 2.2:** Randomizations for the simulation experiments.

	<b>Slight</b>	<b>Moderate</b>	<b>Large</b>
Translation $x$ (in mm)	$[-30, 30]$	$[-30, 30]$	$[-30, 30]$
Translation $y$ (in mm)	$[-20, 20]$	$[-40, -20] \cup [20, 40]$	$[-50, -40] \cup [40, 50]$
Translation $z$ (in mm)	$[-30, 30]$	$[-30, 30]$	$[-30, 30]$
Rotation $\alpha, \beta, \gamma$ (in $^\circ$ )	$[-6, 6]$	$[-6, 6]$	$[-6, 6]$
Catheter smoothing $\sigma_{simu}$ (in mm)	$[1, 5]$	$[5, 10]$	$[10, 15]$

**Figure 2.12:** Evaluation of the registered tip position with  $e_d$  the Euclidean distance,  $l_d$  the longitudinal distance and  $o_d$  the orthogonal distance.

3DRA registered space (using  $T_{w \leftarrow 3DRA}$ ) as well as the fluoroscopic sequences with their projection information. We choose the position of the 3D simulated catheter tip in the 3D vessel tree such that it matches the catheter tip in the fluoroscopic image after a registration of the 3D vessel tree to the fluoroscopic image. Next, we extract a 3D simulated catheter centerline following the 3D vessel path from the tip to the root and project it on the fluoroscopic image using the angles from the fluoroscopic sequences acquired. Those ground truth projections are used to quantify the accuracy of the registration results. To simulate a smooth catheter that may be stretching the vessel, and that may be partially outside the vasculature, we smooth the 2D projection centerline with a Gaussian kernel (with a standard deviation  $\sigma_{simu}$ ).

We applied random transformations  $T_{w \leftarrow 3DRA}$  to the 3DRA volume, divided over three sets depending on the magnitude of the transformations and the Gaussian smoothing parameter (Table 2.2).

In this experiment we used the algorithm parameters obtained in the previous leave-one-out cross-validation. We quantified the Euclidean distance  $e_d$  between the known projected 3D tip and the registered projected 3D tip. We also computed the longitudinal  $l_d$  and orthogonal  $o_d$  distances from the point of view of the known tip (Fig. 2.12).



## 2.4 Results

### 2.4.1 Clinical data and parameters optimization

The results of the leave-one-out cross-validation with the two optimizers Powell and brute force, 3 mm sampling and all images are presented in Table 2.3. The optimal parameter settings are consistent over the leave-one-out experiments. Based on these results, unless noted otherwise, we used 3 mm sampling and with Powell:  $\lambda = 0.2$ ,  $\sigma = 20$ ,  $d_{max} = 40$  mm and with brute force:  $\lambda = 0.1$ ,  $\sigma = 80$ ,  $d_{max} = 20$  mm.

The average paired vessels distance results are summarized in Figure 2.13. Results are grouped depending on the 3DRA acquisition. Compared to our previous method [1], using the same sampling, the new method has a median that is smaller. Also the brute force optimizer performs better than Powell and is more robust. With the complete 3DRA set, medians are around 5 mm for brute force and 6 mm for Powell. Varying the sampling density between 1.5 mm and 6 mm does not clearly affect the accuracy.

In Table 2.4, we present the distribution of the best registered leaf vessel centerline  $V(l^{best})$  among the  $k$  ranked and selected leaf vessel centerlines. This Table shows that the best registration result is generally obtained with the leaf vessel centerlines that ranked best using the shape-based metric  $S_{max}$ . This demonstrates that the metric can effectively be used to reduce the number of potential vessels to register. The low percentage for the fifth ranked vessel also suggests that the choice of  $k = 5$  is a good compromise between registration speed and the robustness of the method. The first ranked leaf vessel centerline is also the one giving the best registration for 39-49% of the images.

Figure 2.14 shows the paired vessels distance after the registration as function of the distance from the catheter tip. The median becomes less accurate after 3 cm between 1 and 5 mm.

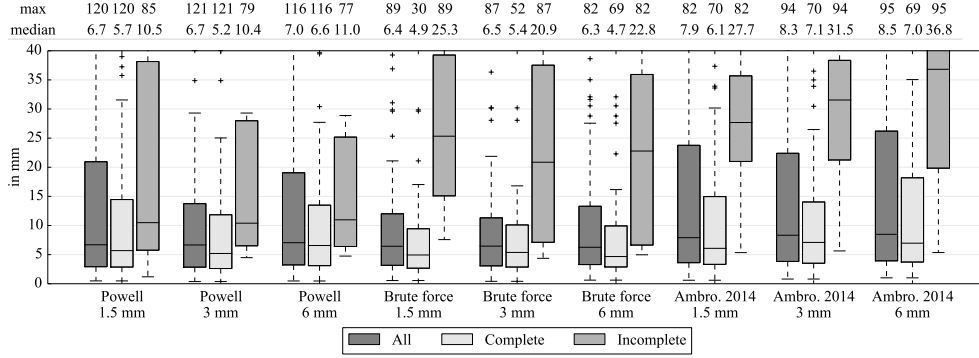
The sampling and the local distance  $d_{max}$  in the metric  $M$  are the parameters that affect the computation time. We show in Figure 2.15 the relation between accuracy and computation time. Brute force is slower than Powell optimization. Our previous method [1] did the registrations with the samplings: 1.5, 3 and 6 mm in 95, 25 and 10 s, which is at least twice as slow as the brute force of our current approach. After 20 mm,  $d_{max}$  does not seem to change the accuracy with brute force. Powell is less stable with both the sampling and  $d_{max}$ .

Figure 2.16 shows examples of correct and incorrect registrations. We note that when there is a small part of the catheter visible on the image, the optimizers are more likely to yield misregistrations because of the lack of information. A correct tip position and distance metric  $M$  do not imply a perfect match of the vasculature due to the deformation of the liver and the catheter.

We visually inspected all registrations from the complete 3DRA set with optimal settings and labelled them as correct, visually close and incorrect for both the registered tip and the registered vessels distal to the tip (Table 2.5). Visually close implies that the registration is sufficient to know where the catheter is in the 3D vasculature while incorrect is of no use for the intervention. For each incorrect case, we also report the likely cause of failure (Table 2.6). The two main reasons of wrong

**Table 2.3:** Optimal settings for Powell (left) and brute force (right) with 3 mm sampling after a leave-one-out cross-validation.

Patients	$\lambda$	$\sigma$	$d_{max}$	Patients	$\lambda$	$\sigma$	$d_{max}$
01,02,03,04,05,06, 08,09,10,11,12,13, 14,15,16,17,19	0.2	20	40	02,03,04,05,06,07, 08,09,11,12,13, 14,15,16,17,18	0.1	80	20
07	0.1	40	40	01	0.2	80	10
18	0	40	20	10	0.1	80	30
				19	0.1	40	20

**Figure 2.13:** Average distance between paired vessels (in 3 cm radius from the tip) for each image after registration (in mm). Comparisons between Powell, brute force optimizer and our previous method Ambrosini et al. 2014 [1].

registrations are a too small part of the catheter visible in the 2D X-ray image and large deformation of both the catheter and the vessels distal to it.

## 2.4.2 Clinical data with a simulated catheter

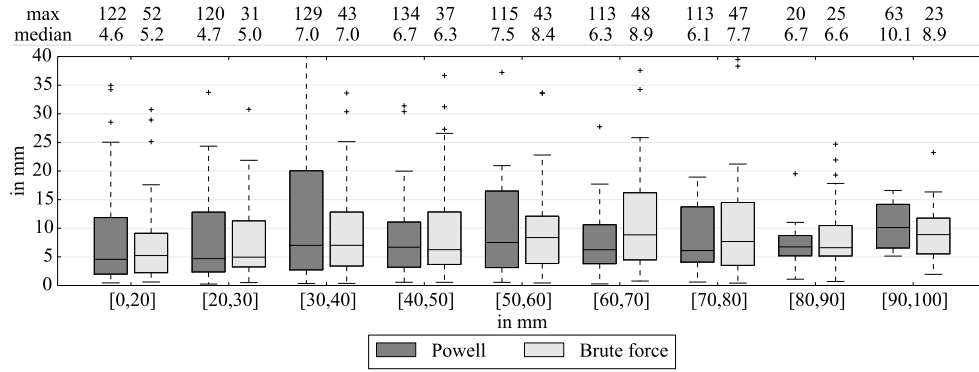
Figure 2.17 shows the distance between the real tip in the simulated catheter (without smoothing) and the tip after registration. Figure 2.18 shows the results with catheter smoothing. Without catheter smoothing, for the brute force optimizer, the median of the Euclidean distance  $e_d$  is below 1 mm whereas for Powell the distance is below 3 mm. With catheter smoothing, the registered tip is less accurate and less robust with both Powell and brute force optimization. The longitudinal and orthogonal distance are similar with slight, moderate or large transformation. Overall, the longitudinal distance is slightly more robust than the orthogonal.

## 2.5 Discussion and conclusion

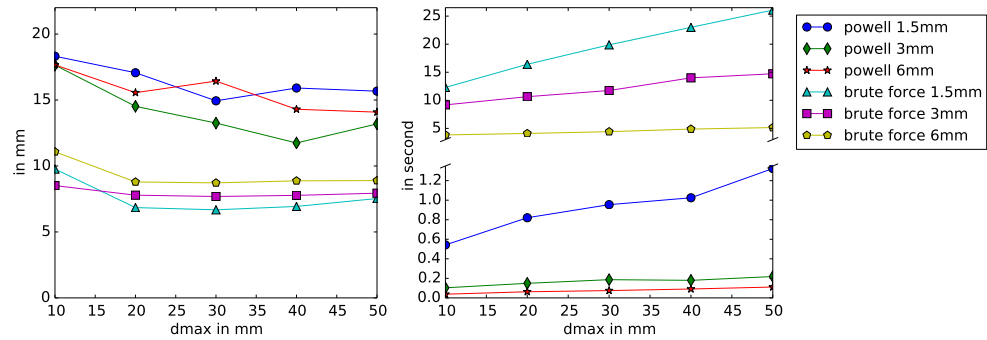
We proposed and evaluated a method that enables a continuous roadmap during abdominal catheterization. The method registers a 3D vessel model obtained from

**Table 2.4:** Distribution of the best registered leaf vessel centerline  $V(t^{\text{best}})$  among the  $k = 5$  ranked and ordered selected leaf vessel centerlines; with the optimal settings and the complete 3DRA set.

sampling (in mm)		1 <sup>st</sup> leaf	2 <sup>nd</sup> leaf	3 <sup>rd</sup> leaf	4 <sup>th</sup> leaf	5 <sup>th</sup> leaf
Powell	1.5	40%	29%	10%	8%	13%
Powell	3	49%	24%	12%	9%	6%
Powell	6	45%	24%	10%	16%	5%
Brute force	1.5	46%	21%	9%	16%	8%
Brute force	3	39%	24%	8%	19%	10%
Brute force	6	39%	22%	12%	17%	10%



**Figure 2.14:** Average distance between paired vessels for all images (in mm) with optimal settings, 3 mm sampling and the complete 3DRA set. Paired vessels are grouped following their distance from the catheter tip (from 0 to 100 mm).



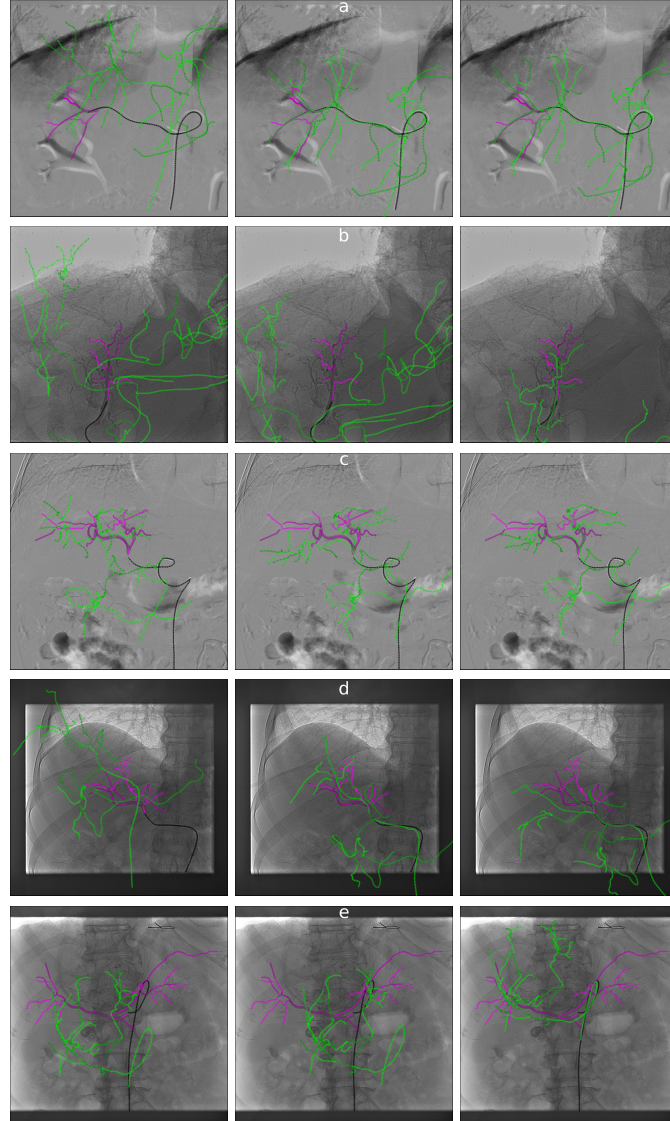
**Figure 2.15:** Average distance between paired vessels for all images in mm (left) and average time in second (right) according to the neighbourhood distance  $d_{max}$ , with optimal settings and the complete 3DRA set.

**Table 2.5:** Visual registration results with optimal settings, 3 mm sampling and the complete 3DRA set.

	Powell	Brute force
<b>Selected registered vessel</b>		
Correct	84%	94%
Incorrect	16%	6%
<b>Match angio/registered vessels</b>		
Visually correct	38%	35%
Visually close	30%	52%
Incorrect	32%	13 %
<b>Registered tip</b>		
Visually correct	59%	73%
Visually close	20%	22%
Incorrect	21%	5%

**Table 2.6:** Registration error details among incorrect match with optimal settings, 3 mm sampling and the complete 3DRA set.

	Powell	Brute force
Small catheter part visible	40%	38%
Large vessels and catheter deformation	25%	25%
Catheter shape not sufficiently distinctive	5%	12%
Rotate too much to fit the best the catheter		25%
Powell stops in a local minimum	20%	
Catheter only in the aorta (missing informations)	5%	
Large part of the aorta is not visible in the 3DRA	5%	



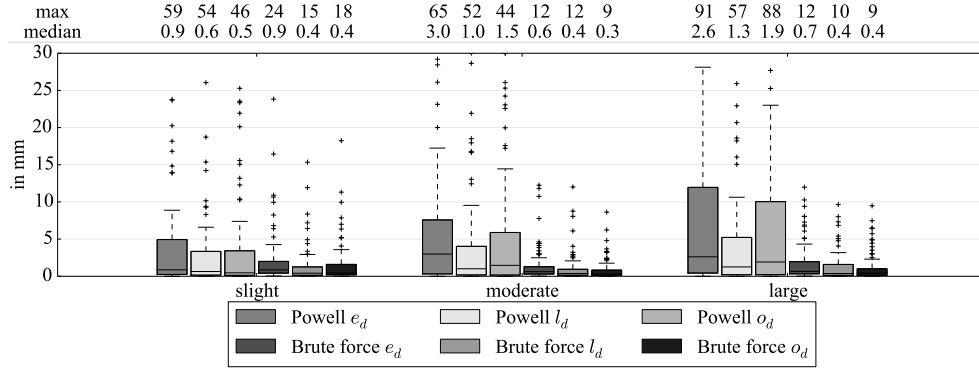
**Figure 2.16:** Projection of the 3DRA blood vessel (in green) with the catheter (in black) and the contrast agent (in purple). Initial position (left). Registered position with Powell (middle). Registered position with brute force (right).

**a** The registration is correct. Here the catheter is long enough to give information.  
**b** The catheter part is too short. Powell registered with a good distance metric but the result is wrong. Brute force is correct.

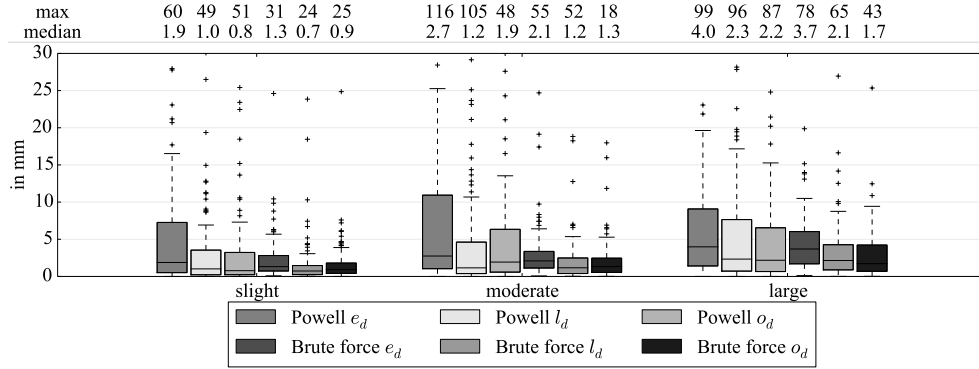
**c** The catheter tip position is correct for both optimizers. The vessels and catheter deformation prevent to have a perfect match.

**d** Here the distance metric and the tip is correct with both optimizers but brute force rotates too much.

**e** As a long part of the aorta is missing in the 3DRA, Powell stops in a local minimum while brute force is more exhaustive and reach the global minimum.



**Figure 2.17:** Euclidean distance  $e_d$ , longitudinal distance  $l_d$  and orthogonal distance  $o_d$  between the real tip and the registered one (in mm) with no catheter smoothing, 3 different simulations (Table 2.2), optimal settings and 3 mm sampling.



**Figure 2.18:** Euclidean distance  $e_d$ , longitudinal distance  $l_d$  and orthogonal distance  $o_d$  between the real tip and the registered one (in mm) with catheter smoothing, 3 different simulations (Table 2.2), optimal settings and 3 mm sampling.



3DRA imaging using a catheter that is extracted from single-plane 2D X-ray images with or without contrast agent. The method first selects the vessels using shape similarity and then rigidly registers the selected vessels to the catheter.

With the complete 3DRA set and optimal settings, the median of average paired vessels distances of the roadmap distal to the catheter tip and within a radius of 3 cm from the tip is 5.4 mm for the brute force optimizer and 5.2 mm for the Powell optimizer. The first selected vessel during shape similarity is used more than 39% in the final registration and the second more than 21%.

We investigated two optimizers for the registration approach: Brute force and Powell. In our setup, with less than 200 ms computation time on average, the registration is real-time with the Powell optimizer and a 3 or 6 mm sample interval. Though the brute force optimizer is slower, it could be improved with parallelization and a dedicated implementation. The brute force optimizer tends to be more accurate and robust than the Powell optimizer. Powell is more sensitive to the initial position of the registration (end up in local minima) as well as the length and distinct shape of the catheter.

The simulation experiments with catheter deformation demonstrate that the registration is robust for both optimizers with slight deformation. They also show that larger deformation leads to less accurate registration. In the simulated data, the longitudinal distance from the tip shows how well the tip is registered along the catheter direction. This distance is more significative than the orthogonal and is slightly more robust.

An important source of error in our experiments was the lack of vessel information in the 3DRA (especially missing the aorta and the hepatic artery), partly caused by the retrospective nature of our study. Optimization of the 3DRA acquisition protocol could remedy this. Another source of registration errors is the lack of information in the 2D X-ray because only very short part of the catheter is visible. This shows the limitation of the method working with no prior knowledge other than the current image. This may be addressed, during the intervention, by slightly increasing the field-of-view, moving the patient table, or by adding more a priori knowledge into the registration such as previous image registration transformations. If we take into account the previous registered transformations and the table motion (which in principle could be obtained from the C-arm system, but is not available in our acquired fluoroscopic images), the registration should have a better initialization and thus use a smaller search space and both Powell and brute force optimizers will perform more robustly while reducing computation time.

Most related 2D/3D registration methods register angiography with CTA or 3DRA. As the complete vasculature is visible on both 2D and 3D images and non-rigid registration is performed, they reach submillimeter accuracies. Our method, dealing only with the catheter visible on the 2D image, has lower accuracy. However, we are interested in improving guidance and the fusion provides a continuous roadmap of sufficient accuracy to the clinician to reliably estimate the catheter tip position in the 3D vasculature. As far as we know, in abdominal studies, the presented method can be compared only with the method proposed by Atasoy et al. [5]. They evaluate their method with the overlap of the 3DRA vessels onto the catheter. In our case, the overlap is our distance metric so a comparison will be biased towards our method.

During abdominal catheterization, knowledge of the position of the tip in the 3D vasculature is of crucial importance. Table 2.5 shows a small percentage of incorrect registered tip positions. This implies that if we use the presented fusion method as a roadmap, combining any of the optimizers, the resulting fused visualization is sufficient to guide the interventionists in localizing the tip and identifying the subsequent bifurcations, also in case of slight misalignment.

A robust automatic 2D catheter segmentation is required after initialization to integrate our method into the interventional workflow. The accuracy of the segmentation will influence the registration method. For example, Heibel et al. [34] obtain a median error of real-time automatic catheter tracking less than 1.5 pixels for abdominal fluoroscopies. Those results are sufficiently accurate for our registration.

Registration studies often lack ground truth for clinical data. In our case, this also prevented us to evaluate the accuracy of the registration method directly. However both the simulation experiments and the validation with angiographic images demonstrate the good performance of the method.

In our current setup, each registration is independent from previous registrations. During continuous roadmapping, only slight motion should occur between two registrations. In the future, we intend to use previous registration results to further improve the robustness (especially when the visible catheter part is too small to do an accurate independent registration), and to limit the computation time by reducing the space search. A source of registration errors was due to large vessels and catheter deformation. A non-rigid registration to match the catheter deformation could also improve the accuracy close to the catheter tip.

To conclude, we presented a catheter-based registration method to fuse continuously 3DRA roadmap arterial tree onto 2D fluoroscopic images. We evaluated our work with clinical and simulated data demonstrating an efficient shape similarity and a median accuracy, evaluated on close by vessels, of 4.7-6.6 mm and below 4 mm on simulated experiments.

**Acknowledgements** This research is funded by Philips Healthcare, Best, The Netherlands. We thank the Hôpitaux Universitaires Henri Mondor, Créteil, Paris, France and the Ospedale di Circolo e Fondazione Macchi, Varese, Italy for providing image datasets.

# A Hidden Markov Model for 3D Catheter Tip Tracking with 2D X-ray Catheterization Sequence and 3D Rotational Angiography.

---

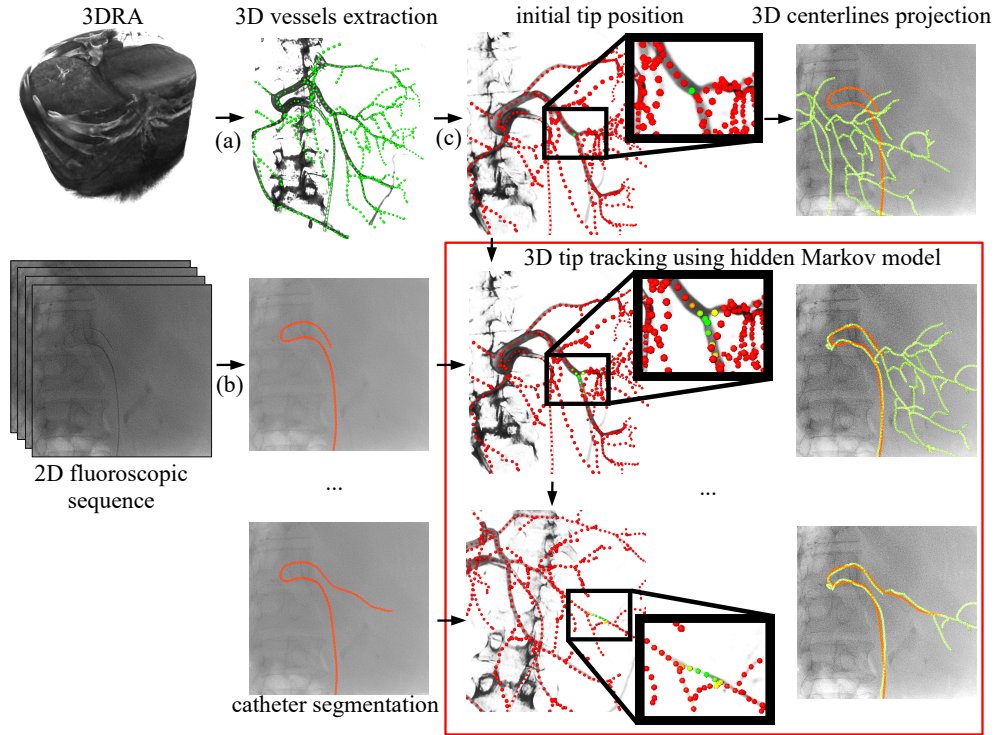
**Abstract** — In minimal invasive image guided catheterization procedures, physicians require information of the catheter position with respect to the patient’s vasculature. However, in fluoroscopic images, visualization of the vasculature requires toxic contrast agent. Static vasculature roadmapping, which can reduce the usage of iodine contrast, is hampered by the breathing motion in abdominal catheterization. In this paper, we propose a method to track the catheter tip inside the patient’s 3D vessel tree using intra-operative single-plane 2D X-ray image sequences and a peri-operative 3D rotational angiography (3DRA). The method is based on a hidden Markov model (HMM) where states of the model are the possible positions of the catheter tip inside the 3D vessel tree. The transitions from state to state model the probabilities for the catheter tip to move from one position to another. The HMM is updated following the observation scores, based on the registration between the 2D catheter centerline extracted from the 2D X-ray image, and the 2D projection of 3D vessel tree centerline extracted from the 3DRA. The method is extensively evaluated on simulated and clinical datasets acquired during liver abdominal catheterization. The evaluations show a median 3D tip tracking error of 2.3 mm with optimal settings in simulated data. The registered vessels close to the tip have a median distance error of 4.7 mm with angiographic data and optimal settings. Such accuracy is sufficient to help the physicians with an up-to-date roadmapping. The method tracks in real-time the catheter tip and enables roadmapping during catheterization procedures.

### 3.1 Introduction

Nowadays, minimally invasive procedures are common because of the associated benefits for the patients, such as shorter recovery times. For example, catheterization procedures are executed to non-invasively reach locations via the vasculature. During catheterization, image guidance is commonly performed using intra-operative 2D X-ray fluoroscopy. On this imaging modality, the catheter is visible but the vasculature is not. Physicians need to know where the catheter (particularly the tip) is to navigate to a specific target. To this end, 2D intra-operative images are conventionally enhanced using contrast agent to visualize the vasculature, which permits physicians to localize the catheter inside the vasculature. However, contrast agent cannot be used continuously due to its toxicity. Also, the projected vasculature can sometimes be difficult to interpret. To have a continuous roadmap, physicians use 2D overlays of Digital Subtraction Angiography (DSA) onto the X-ray images. 3D projection overlays, e.g. from pre-operative CTA or MR images, have also been used [80]. Unfortunately, such 2D and 3D roadmaps are generally static and the registration between 3D vasculature extracted from CTA/MRA and 2D images is not straightforward. In e.g. abdominal catheterization, respiration induced motion and the catheter stiffness lead to motion and deformation of the vasculature, invalidating the roadmap.

The purpose of our work is to continuously localize the catheter tip inside the 3D model of the patients vasculature during the catheterization procedures. The paper focuses particularly on liver catheterization interventions such as Transcatheter Arterial Chemoembolization (TACE). We propose a method to track the catheter tip inside a patient-specific contrast-enhanced 3D abdominal vasculature, obtained from peri-operative 3D Rotational Angiography (3DRA), using single-plane 2D X-ray images with no contrast agent. The proposed tip tracking method enables 2D as well as 3D roadmapping that can easily be integrated in the intervention and permits continuous and contrast-free image guidance (Fig. 3.1). Such image guidance would potentially reduce toxic contrast agent use and may decrease procedure time as well. In liver catheterization, peri-interventional 3DRA (which is a form of Cone Beam Computed Tomography, CBCT) is acquired at the beginning of the procedure when the catheter is in the common hepatic artery. The contrast agent is injected directly into the liver vessels to offer a better visibility of the vascular morphology. 3DRA provides a better understanding of the vasculature compared to DSA because of its 3D nature. It also helps to position the C-arm to obtain an optimal view on the vessels, taking foreshortening, overlap and bifurcations into account. 3DRA is more and more acquired during TACE procedure because it increases the confidence of the physician and it helps for the planning and the guidance [10].

Most current approaches in image guidance for catheterization procedures focus on registering a 3D pre-operative angiographic image, such as CTA or MRA, with intra-operative 2D images, such as single-plane/bi-plane X-ray images or DSA. Thorough reviews of 3D/2D registration methods have been presented by Markelj et al. [60] and Liao et al. [49]. Various methods have been proposed for cardiac [8, 9, 61, 77, 81], cranial [35, 64, 91, 92] and abdominal [30, 42, 47] procedures. These methods use the image intensity, gradient or features such as bones and more generally vessels, to spatially align the 3D image to the 2D image. In cardiac and abdominal procedures, due



**Figure 3.1:** Tip tracking workflow:

- (a) 3D vessel tree is extracted from 3DRA.
- (b) 2D catheter is extracted for each frame of the X-ray sequence.
- (c) The HMM is initialized with the initial tip position in the 3D vessel tree. The HMM is then updated depending on the 3D/2D registration metric. The tip position and the registered transformation are obtained which can be used as a 3D/2D roadmap to guide the physician. The tracking using HMM (red frame) is the scope of this paper. The other steps are approached in the discussion.

to respiratory motion, the registration has to be updated continuously. Methods that utilize vessel information for alignment can not be used continuously, as the vessels are not continuously visible because of the contrast agent toxicity. Ambrosini et al. [2], Atasoy et al. [5] and Ma et al. [58] correct for breathing motion using other features in the 2D X-ray image such as diaphragm/heart border, tracheal bifurcation or the catheter. The use of such features gives reasonably accurate results but the robustness depends on the complete visibility of the features. In abdominal X-ray catheterization images, only the catheter is visible and the catheter-based registration fails if the visible catheter part is too short [2]. Using the catheter-based registration combined with a hidden Markov model (HMM) [75], our feasibility study [3] showed that 3D catheter tip tracking over the time was possible and could potentially overcome the problem of short feature visibility such as short part of the catheter visible on 2D X-ray images. This study was limited but showed encouraging results for 3D/2D roadmapping in abdominal catheterization procedures.

This paper has two main contributions. First, we present a novel and robust 3D catheter tip tracking method, based on an HMM, and that extends our previous preliminary work [3]. We propose a new transition probability between the states in the HMM. Furthermore, a new cost function and different initialization is introduced for the 3D/2D registration. Second, we perform extensive quantitative and qualitative evaluations with both simulated and clinical data, using clinically relevant measures, i.e. tip distance error and vessel distance error close to the tip. In summary, this method presents the following advancements compared to the state of the art:

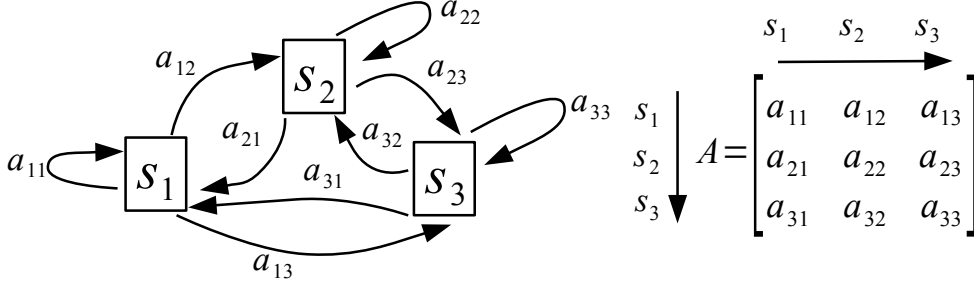
- The tip tracking as well as the roadmapping can be performed continuously using fluoroscopic imaging without the need for contrast agent and with no distinctive features except the catheter,
- 3D roadmapping is provided next to 2D roadmapping. This would add more information for the physician with difficult cases,
- The nature of the tracking method provides temporal consistency of the tracking results, and as such is robust to e.g. field-of-view changes.

## 3.2 Method

The purpose of our method is to track the catheter tip inside the 3D vessel tree extracted from the peri-operative 3DRA using information from 2D fluoroscopic imaging. A probability distribution of the 3D catheter tip position is computed using an HMM. The probability distribution update is based on the previous distribution and the 3D/2D registration results of the 3D vessel tree and the 2D catheter centerline extracted from the 2D intra-operative X-ray image. In this process, a 3D/2D registration is performed for the most probable catheter tip locations.

In the following, we first describe the HMM that was used and then show how it applies to our application with the catheter tip tracking.





**Figure 3.2:** HMM with 3 states and its matrix  $A$  of state transition probabilities.

### 3.2.1 Hidden Markov model

A hidden Markov model is a system with a set of  $N$  states  $S = \{s_1, \dots, s_N\}$  (Fig. 3.2). The HMM state changes at each discrete time point  $t$  according to the probabilities associated with the state transitions and the current measurement/observation  $Z_t$ . The transition probabilities between states are defined in a matrix  $A$  (dimension  $N \times N$ ) where each element  $a_{ij}$  of  $A$  is the probability that the HMM moves from state  $i$  to state  $j$  ( $a_{ij} \geq 0$  and  $\sum_j a_{ij} = 1$ ). To employ an HMM, we define an observation likelihood (observation score)  $O_t(i) = P(Z_t|s_i)$ . It represents the likelihood that the state  $s_i$  at time  $t$  produced the observation  $Z_t$ .

Following Rabiner et al. [75], the Viterbi algorithm selects at time  $t$  the optimal path over time (called Viterbi path) through the state space based on the maximum  $\delta_t(i)$  which is the best score along a single path that starts in any state at time 0 and ends in state  $s_i$  at time  $t$ . The Viterbi algorithm takes into account the  $t$  previous system states. Starting from an initial distribution of the probabilities over the states  $\Pi = \{\pi_1, \dots, \pi_N\}$  where  $\sum_{j=1}^N \pi_j = 1$ , the algorithm initializes  $\delta_0(i)$  as follows:

$$\delta_0(i) = \pi_i \quad . \quad (3.1)$$

Next,  $\delta_t(i)$  can be computed using recursion [75]:

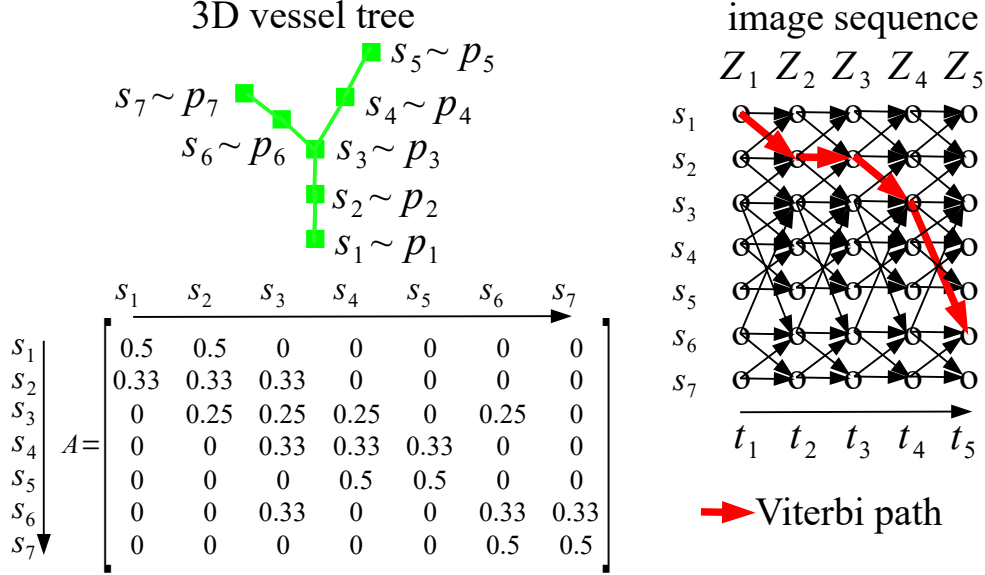
$$\delta_t(i) = \max_j [\delta_{t-1}(j) a_{ji}] O_t(i) \quad . \quad (3.2)$$

The next section explains the relation between HMM and tracking of a catheter tip in 3D using fluoroscopic images.

### 3.2.2 Catheter tip tracking

#### 3.2.2.1 Timepoint

The timeline is associated with the frames of the 2D X-ray image acquisition (around 7 Hz). Each discrete time point  $t$  of the HMM corresponds to a 2D X-ray image  $Z_t$ . The measurement/observation for the HMM is the 2D X-ray image.



**Figure 3.3:** HMM with a vessel tree. In this example, the transitions between possible tip locations exist only with direct neighbours and are equiprobable. The Viterbi path goes to the optimum tip position knowing the images/observations  $Z_1 \dots Z_5$  in the 5-images sequence.

### 3.2.2.2 States

The 3D catheter tip is tracked inside the 3D vessel tree centerline extracted from the 3DRA. The vessel tree is discretized as a set of 3D points  $P = \{p_1, \dots, p_N\}$ . The probability that the catheter tip is at position  $p_i$  is the probability that the HMM is in the state  $s_i$ . The function  $\delta_t(i)$  is then the score that the tip is at the position  $p_i$  in the tree for the discrete time point  $t$  (Fig. 3.3). In our application, the liver arteries move because of respiration. For our model, we define  $p_i^0$  the position of the 3D point  $p_i$  in the world coordinate system at time  $t = 0$  (starting position at 3DRA acquisition), and  $p_i^t$  its position at time  $t$ .

### 3.2.2.3 Matrix A of state transition probabilities

The matrix  $A$  of state transition probabilities describes all the transition probabilities  $a_{ij}$  for the catheter tip to move from one point  $p_i$  to another point  $p_j$  in the vessel tree  $P$  (state  $s_i$  to  $s_j$ ), between two time points (i.e. between two subsequent 2D X-ray images). As the interval of time is relatively short, the probabilities of transition should be high in the vicinity of the point  $p_i$  because the catheter (and so the tip) is not expected to move very far away. Therefore, the transition probabilities are set according to the distance along the vessel path between points  $p_i$  and  $p_j$  of the 3D vessel tree. Two distributions are proposed to model the catheter tip motion: one,  $a'_{ij}$ , restrictive around the point  $p_i$  and one,  $a''_{ij}$ , flexible in the close vicinity of  $p_i$ .

The first one models a catheter moving relatively slow with a Gaussian function [3]:

$$a'_{ij} = e^{-\frac{D(p_i, p_j)^2}{2\sigma_a^2}} \quad (3.3)$$

where  $\sigma_a$  controls how fast and far the catheter tip can move, and where  $D(p_i, p_j)$  models the distance (along the vasculature) between  $p_i$  and  $p_j$ . I.e. let  $\{g_1, \dots, g_n\}$  be the set of points representing the vessel centerline between the points  $p_i = g_1$  and  $p_j = g_n$ , then  $D(p_i, p_j)$  is defined as the sum of the distances between each neighboring pair  $g_k$ , and  $g_{k+1}$ :

$$D(p_i, p_j) = \sum_{k=1}^{n-1} \|g_k, g_{k+1}\| \quad (3.4)$$

The second distribution gives the same probability for moving to any point in the vicinity of the point  $p_i$ , where vicinity is defined by a maximum distance  $\theta$  (in mm). We have  $a''_{ij} = 1$  when  $D(p_i, p_j) \leq \theta$  and  $a''_{ij} = 0$  when  $D(p_i, p_j) > \theta$ . Because the matrix  $A$  defines probabilities,  $\sum_j a_{ij}$  has to be equal to 1, thus we normalize the coefficients  $a'$  (resp.  $a''$ ) to obtain  $a_{ij} = a'_{ij} \cdot (\sum_j a'_{ij})^{-1}$  (resp.  $a_{ij} = a''_{ij} \cdot (\sum_j a''_{ij})^{-1}$ ).

#### 3.2.2.4 Observation scores

During the tracking, for each point  $p_i$ , we need to compute the score  $\delta_t(i)$  (Equation 3.2), showing how likely the observation is following the previous system states and under the condition that the tip is at the position  $p_i$  (the system in the state  $s_i$ ) at time  $t$ . Thus each observation score (likelihood)  $O_t(i)$  needs to provide information on the current state, based on the X-ray image  $Z_t$ . In our method, we evaluate the observation score via 3D/2D registration. In other words, the better the 3D/2D registration is (following our registration metric), the better the observation score will be. Each observation score  $O_t(i)$  is based on the 3D/2D registration between the catheter shape extracted from the 2D X-ray image  $Z_t$  and the unique 3D catheter path centerline  $V_i = \{v_1, \dots, v_{n_{V_i}}\}$ , starting from the tip  $p_i$  and going to the root of the tree. The catheter shape is defined as a set of 2D points in the image coordinate system  $C_t = \{c_1, \dots, c_{n_{C_t}}\}$ . The observation score  $O_t(i)$  is a likelihood between 0 and 1 stating how the 3D/2D registration performed at time  $t$  given that the tip is at the position  $p_i$  in the vessel tree. It is defined as a Gaussian function:

$$O_t(i) = e^{-\frac{M(C_t, V_i)^2}{2\sigma_s^2}} \quad (3.5)$$

where  $M$  is the metric of the 3D/2D registration and  $\sigma_s$  controls the scaling of the registration metric.

The 3D/2D registration between the vessel path  $V_i$  and the catheter  $C_t$  determines the transformation matrix  $\tau$  that minimizes the metric  $M$ .  $M$  quantifies the alignment of the 3D vessel with every point of the 2D catheter:

$$M(C_t, V_i) = \min_{\tau} \left( \sum_{c \in C_t} F_{\text{cost}}(c, V_i, \tau) \right) \quad (3.6)$$

where  $\tau = T_{\text{rot}_z} T_{\text{rot}_y} T_{\text{rot}_x} T_{\text{trans}_x}$  is a 4 degrees of freedom rigid transformation matrix (three rotations and one translation), and  $F_{\text{cost}}$  is a function that represents how close is a 2D point of the catheter to the projected 3D vessels knowing the transformation  $\tau$ . The transformation  $\tau$  only has 4 degrees of freedom because the projection of the 3D tip  $p_i^t$  at time  $t$  has to match the 2D catheter tip  $c_1$  and therefore the transformation  $\tau$  translates only along one axis, the line from  $c_1$  to the origin of the X-ray projection, and rotates around it (Fig. 3.4).

In order to define the cost function  $F_{\text{cost}}$ , we first explain how the 3D vessel tree is projected from the world coordinate system to the image coordinate system. We compute (in the world coordinate system) the intersection of the line projecting  $c_1$  (i.e. the line from X-ray source to  $c_1$  on the X-ray image/detector plane) with a plane that is parallel to the image plane, and that contains the 3D tip  $p_i^0$ . This intersection is represented by the point  $l_i$  which will be the starting point as a 3D catheter tip for the registration search of  $\tau$ . As any point on the line along the X-ray source projects on  $c_1$ , the only translation allowed to find  $p_i^t$  is along this line. Furthermore, the center of rotation of  $\tau$  is the same point  $l_i$ , such that the rotations also guarantee that  $p_i^t$  projects on  $c_i$ . Therefore, the projection transformation is parametrized as follows (Fig. 3.4):

$$F_{\text{proj}}(v, \tau) = T_{\text{proj}} T_{\text{det} \leftarrow w} T_{w \leftarrow l_i} \tau T_{l_i \leftarrow w} T_{p_i^0 \rightarrow l_i} v \quad (3.7)$$

where  $T_{p_i^0 \rightarrow l_i}$  is the translation along the line  $p_i^0$  to  $l_i$  in the world coordinate system and  $T_{l_i \leftarrow w}$  is the transformation from the world coordinate system to the coordinate system centered around  $l_i$ . The transformation  $T_{\text{proj}}$  is the cone-beam projection and  $T_{\text{det} \leftarrow w}$  the transformation matrix from the world to the C-arm detector (X-ray image plane). Both transformations are known because of the C-arm geometry (given in the DICOM file). Because the projection of the 3D tip  $p_i^t$  is the 2D catheter tip  $c_1$ , we have  $F_{\text{proj}}(p_i^t, \tau) = c_1$ .

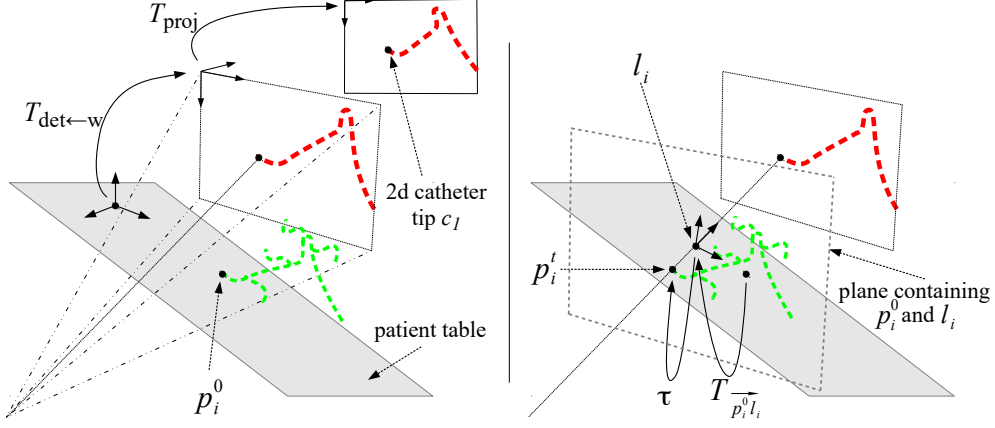
We propose two cost functions to measure how far a 2D catheter point is from a projected 3D vessel. The first one is the minimal distance between each point of the 2D catheter  $C_t$  and any projected point of the 3D centerline  $V_i$ :

$$F_{\text{cost}_1}(c, V_i, \tau) = \min_{v \in V_i} \|c - F_{\text{proj}}(v, \tau)\| \quad (3.8)$$

where  $F_{\text{proj}}(v, \tau)$  is the projection of the 3D point  $v$  onto the 2D images. This cost function ignores the vessel diameter when registering the centerline of the vessel path  $V_i$ . As some vessels may have a large diameter (e.g. in the hepatic artery or the aorta), we defined a second cost function that takes into account the vessel diameter and reduces the cost of the alignment when the catheter is inside a vessel  $V_i$ . Using the known radiuses  $R_i = \{r_1, \dots, r_{n_{V_i}}\}$  associated with every point of the vessel  $V_i$ , the second cost function based on sigmoid is (Fig. 3.5):

$$F_{\text{cost}_2}(c, V_i, \tau) = \frac{F_{\text{cost}_1}(c, V_i, \tau)}{1 + e^{\alpha_1 \cdot (-F_{\text{cost}_1}(c, V_i, \tau) + \alpha_2 \cdot r)}} \quad (3.9)$$

where  $\alpha_1$  and  $\alpha_2$  control the cost inside the vessel and  $r$  is the radius of the point  $v \in V_i$ .  $v$  is the point which has the minimum distance with the catheter point  $c$  in



**Figure 3.4:** Initial position of the 3D vessel tree (in green) with the 2D projected catheter (in red) at time  $t$  (left figure). 3D points are transformed from the 3D world coordinate system to the 3D X-ray detector plane coordinate system ( $T_{\text{det} \leftarrow w}$ ) and then projected to the 2D image plane ( $T_{\text{proj}}$ ).

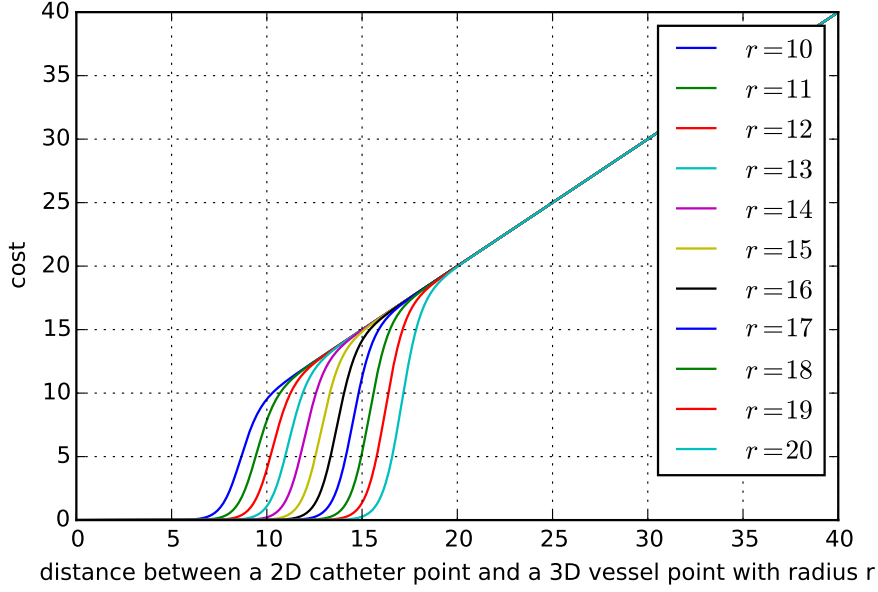
Registered position of the 3D vessel tree (knowing that the catheter tip is at the position  $p_i$ ) after the computation of the optimal transformation  $\tau$  (right figure).

the first cost function  $F_{\text{cost}_1}(c, V_i, \tau)$ . We want  $F_{\text{cost}_2}$  to start to penalize when the minimum distance reach the radius  $r$ . The variables  $\alpha_1 = 2$  and  $\alpha_2 = 0.85$  fit to obtain this behaviour (Fig. 3.5).

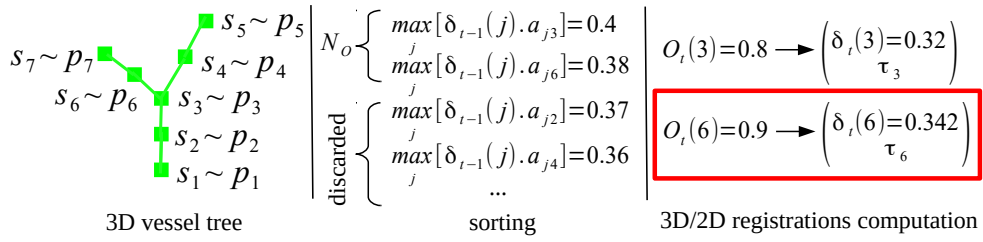
### 3.2.2.5 Viterbi algorithm

For each image, the Viterbi path [75] is computed from the initial state position at the first image to the current image. This results in the most likely 3D tip point  $p_i$  for each image, and also (as a 3D/2D registration has been performed during the observation scores evaluation) the transformation  $\tau$  that aligns the projected 3D vessel tree with the 2D fluoroscopy.

Theoretically, at time  $t$ , the observation score  $O_t(i)$  for every state  $i$  should be evaluated. For large vessel trees, that may contain hundreds of points, this would not permit real-time use. Therefore, to reduce the computational effort, the observation scores of only a relatively small number of  $N_O$  states are evaluated and for all the other states, the observation scores are set to 0. To this end, we first sort all the states  $s_i$  following the score  $\max_j [\delta_{t-1}(j) a_{ji}]$ . The  $N_O$  states with the best scores will be selected to have their observation scores  $O_t(i)$  evaluated. Thus, only the most likely tip positions will be evaluated and the other possible tip positions will be discarded (Fig. 3.6).



**Figure 3.5:**  $F_{\text{cost}_2}$  in function of the distance  $F_{\text{cost}_1}$  and various radius  $r$ .



**Figure 3.6:** Process example with  $N_O = 2$ . At time  $t$ ,  $p_6$  is the most likely position for the catheter tip.  $\tau_6$  is the result of the minimization for the 3D/2D registration when evaluating  $O_t(6)$ .



### 3.3 Experiments

The method presented in the previous section was qualitatively and quantitatively evaluated, considering the accuracy as well as the robustness. The evaluation was carried out on both simulated data and clinical data. In the following, we describe respectively the data acquisition and implementation, the simulation experiments and the experiments using clinical data.

#### 3.3.1 Data acquisition and implementation

We retrospectively acquired data of 28 Transcatheter Arterial Chemoembolization (TACE) procedures. The data, which were anonymized prior to use in this study, comes from three different hospitals (Erasmus MC, University Medical Center, Rotterdam, the Netherlands; the Hôpitaux Universitaires Henri Mondor, Créteil, Paris, France; and the Ospedale di Circolo e Fondazione Macchi, Varese, Italy) using intervention rooms with angiographic C-arm systems (Xper Allura, Philips Healthcare, Best, the Netherlands). In total, we acquired 10 long fluoroscopic sequences (46-76 frames) from 10 procedures and 74 angiographic sequences (4-11 frames) from 19 procedures (one is in common with the long sequences set). A 3DRA image was acquired at the beginning of each intervention when the catheter was in the common hepatic artery. For each fluoroscopic image  $Z_t$  in each sequence (i.e. all the frames), the 2D catheter centerline  $C_t$  was manually segmented by annotating points on the 2D catheter and then fitting a continuous spline to these points. The 3D vessel tree centerline  $P$  in the 3DRA was extracted with a semi-automatic method based on thresholding and skeletonization [83]. Each point  $p \in P$  is associated with the radius of the vessel at that position. The catheter  $C_t$  and the vessel tree  $P$  are discretized with a sampling distance of  $\gamma$  mm between every consecutive point. The effect of the sampling distances  $\gamma = \{0.5, 1.5, 3, 6, 9\}$  mm is evaluated in the experiments.

To minimize the metric  $M$ , we used the Powell optimizer [73] where the rotation  $\text{rot}_x$ ,  $\text{rot}_y$  and  $\text{rot}_z$ , and the translation along the projected line  $\text{trans}_x$  (from the transformation  $\tau$ ) are the optimization parameters. As the 3D vessel tree will translate and rotate around the projected line of the 2D catheter tip position, the initial position of the 3D vessel tree is already close to the optimal position. Thus, we can have a rather limited search space during the registration. The rotations have been constrained between  $\pm 2^\circ$  and the translation between  $\pm 2$  mm. Larger rotations around  $\text{rot}_y$  and  $\text{rot}_z$  can lead to overlapping vessels and less reliable registrations.

When the HMM updates  $\delta_t(i)$  are computed, values can become very small over the time and as a result, computations may be numerically unstable. As Rabiner et al. [75] suggest, updates can also be computed in log scale in order to get numerical results more stable. The equation in log scale becomes:

$$\delta_t(i) = \max_j [\delta_{t-1}(j) + \log(a_{ji})] + \log(O_t(i)) \quad (3.10)$$

The experiments were run on an laptop Intel Core i7 (2 Ghz). The HMM tracking method source code is available.<sup>1</sup>

<sup>1</sup><https://github.com/pambros/HMM-3D-Catheter-Tip-Tracking>

### 3.3.2 Patient vessel tree with a moving simulated catheter

We performed experiments on simulated data; the main purpose of these experiments was to apply and evaluate the tracking algorithm in the context of data with ground truth. With this simulation, we know the exact position of the 3D catheter and the 3D vessel tree, with respect to the C-arm system. We thus can project the simulated catheter and obtain the position of the 2D simulated catheter. Then, similar to the real data scenario, the algorithm tracks the 3D catheter tip position in the 3D vasculature (using the 3D vessel tree centerline, the 2D catheter centerline at each frame and the first 3D tip position to initialize the distribution of the probabilities over the states:  $\Pi$ ). Below, we first describe how the simulation data was generated, and then we detail the experiments running these data.

#### 3.3.2.1 Catheter simulation

To obtain realistic data, a 3D catheter is simulated using the clinical data i.e. 3D vessel tree and fluoroscopic image projection. In the simulation, we included catheter motion inside the vasculature and respiration induced motion. This was done as follows: For each fluoroscopic sequence, the 3D vessel tree (extracted from the 3DRA) was manually registered with a frame where the 2D catheter tip is at the most distal position in the vasculature in order to get the longest catheter visible. The goal is not to obtain an accurate registration but a simulation as close as the real clinical case. Because of the manual registration, the catheter tip position in the 3D vessel tree is obtained. A catheter is generally not at the vessel center, but rather positioned at the boundaries of the vessels. So, from the tip position, an initial catheter shape is determined by extracting a minimum cost path inside the vessel tree binary mask from the catheter tip position to the root of the vessel tree; the resulting path is the simulated catheter  $C_m$ . The catheters  $C_0, C_1, \dots, C_{m-1}$  are constructed from  $C_m$ . The catheter and thus the tip move from the hepatic artery, where the contrast agent has been injected during the 3DRA acquisition, to the tip position of  $C_m$ . The displacement has a fixed speed for each sequence, randomly chosen in a range from 1 to 10 mm/frame.  $C_0$  is the first catheter shape, running from the root until the hepatic artery.  $C_1, \dots, C_{m-1}$  are similar parts of  $C_m$ , always starting at the root and ending at a position in-between the hepatic artery and the tip position of  $C_m$ . These positions are computed such that the catheter tip advances with a fixed speed. A Gaussian smoothing with a random  $\sigma \in [1, 2]$  mm is applied on every constructed catheter to add a slight deformation. Additionally, the catheter motion caused by respiration is simulated by translating the catheter in the cranial-caudal direction. The translation amplitude is defined as:

$$\text{translation}(i) = \lambda \sin\left(\frac{2\pi}{\beta} i \Delta_t\right) \quad (3.11)$$

where  $\lambda = 10$  mm is the peak amplitude,  $\beta = 4$  s the respiration period,  $i$  the number of the current image in the sequence and  $\Delta_t = 0.133$  s the time between two images, obtained from the clinical fluoroscopic image frequency.

### 3.3.2.2 Experiments and parameters optimization

The catheter simulation is based on the 74 angiographic sequences. 6 sequences have been discarded because the catheter was too short (less than 8 cm) to simulate a tip motion. The sequences with catheter simulation are divided in one training set with 20 sequences originating from 7 procedures and one test set with 48 sequences originating from 12 different procedures. First, experiments have been done to choose optimal parameters. We used the simulated catheters to evaluate the tracking accuracy. To this end, we computed the 3D distance between the ground truth catheter tip and the tip after the registration for each frame of each simulated sequence. For each experiment, the two cost functions ( $F_{\text{cost}_1}$  and  $F_{\text{cost}_2}$ ), and the two transition distributions ( $a'$  and  $a''$ ) were evaluated.

The parameters  $\sigma_a$  and  $\theta$  from the transition matrix  $A$ , and  $\sigma_s$  from the observation score were optimized with the training set. The purpose is to evaluate the tip tracking accuracy of the method as a function of  $\sigma_a$ ,  $\theta$  and  $\sigma_s$  and therefore we report the 3D tip distance over all the frames of sequences. The parameter values that gave the shortest median distance were used for all the following experiments. The evaluation has been done with the values:  $\sigma_a = \{6, 9, 12, 15, 18, 20\}$  mm,  $\theta = \{6, 9, 12, 15, 18, 20\}$  mm and  $\sigma_s = \{0.25, 0.5, 1, 1.5, 2, 2.5\}$  mm. For these experiments, the sampling  $\gamma$  of the 2D catheter and the 3D vessel tree centerline is set to 3 mm and the number of observation scores to evaluate  $N_O$  is set to 200.

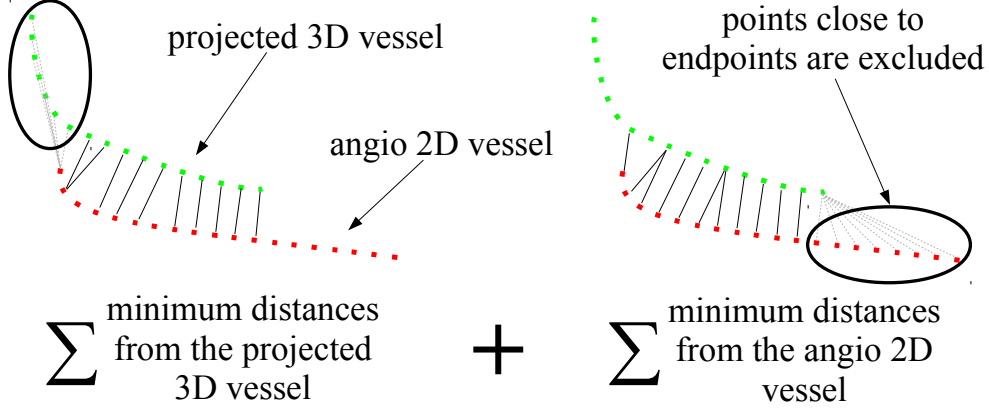
The sampling  $\gamma$  and the number of observation scores to evaluate,  $N_O$ , affect computation time, and may also affect method accuracy. We investigated the effect of using different samplings  $\gamma$  and different  $N_O$  with the optimal values of  $\theta$ ,  $\sigma_s$ , the cost function  $F_{\text{cost}_1}$  and the transition distributions  $a''$  (determined for each sampling  $\gamma$  and  $N_O = 200$ , like the previous experiment). The samplings  $\gamma = \{0.5, 1.5, 3, 6, 9\}$  mm and  $N_O = \{10, 25, 50, 100, 200\}$  have been evaluated. The 3D distance between the real and registered tip is reported to find the optimal accuracy/computation time tradeoff.

With the optimal parameters, the test set is evaluated and the 3D tip distance is reported. Also, the robustness of the tracking was evaluated by determining whether the catheter tip is tracked correctly until the end of the sequence; we consider a tracking failed when the registration results of the last 5 images of the sequence yield a 3D distance between the real catheter tip and the registered one greater than 3 mm. The percentage of tracking failures is reported.

### 3.3.3 Clinical data

The optimal parameters computed with simulated catheters were used in the experiments with clinical data. As clinical data is missing ground truth for tip tracking, we conducted two different experiments:

1. One experiment with quantitative evaluation on sequences with contrast agent: 74 sequences (between 1 and 2 s length) where the catheter is not advanced significantly and the contrast agent is visible. The main motion to be recovered is the motion of the liver caused by respiration. The 2D enhanced vessels close to the tip were compared with the 2D vessels projected from the 3D vessels tree.



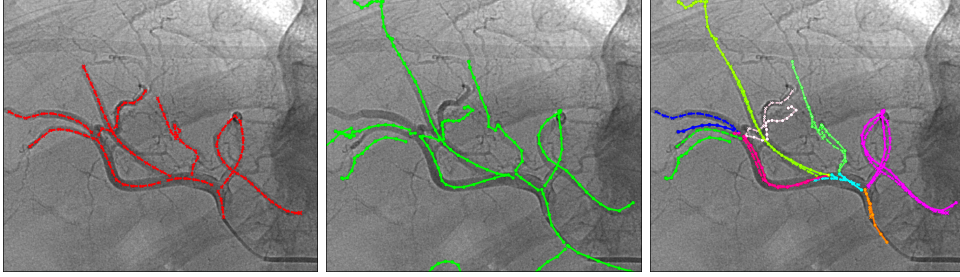
**Figure 3.7:** Closest corresponding points distance between paired vessels used for the sequences with contrast agent evaluation.

2. One experiment with qualitative evaluation on motion sequences: 10 sequences where the catheter is advanced (between 7 and 10 s length). The tracking was visually checked.

For each sequence, in the first frame, the 3D tip is manually annotated in the 3D vessel tree to initialize the method. Although annotating the exact position is difficult, the annotation is sufficient for the initialization of the HMM. If the point  $p_i$  is the manual 3D tip position annotation, then  $\pi_i$  is set to 1 and all the other position probabilities in the initial distribution  $\Pi$  are set to 0.

### 3.3.3.1 Experiments with sequences with contrast agent

Angiographies or DSA were acquired for the sequences with contrast agent. The last X-ray images of the sequences show the vasculature in the vicinity of the tip of the catheter. These vessels can be compared with the 2D projection of the 3D registered vessel tree. We used the same evaluation as described in our previous method [2]. The evaluation metric is the average distance of the closest points between the 2D vessels and the 2D projection of their corresponding 3D vessels (Fig. 3.7). The correspondence between the projections of the 3D vessels and the 2D vessels in the angiographies is unknown. Therefore, we manually paired the vessels visible in the X-ray angiography to the vessels projected from the 3D tree (Fig. 3.8). The distances between the paired vessels close to the tip are then computed. The most relevant region for the tip tracking is the area close to the catheter tip; we therefore only evaluated in a circular region (3 cm radius) around the catheter tip. The 2D vessels from the angiographies are manually segmented and the 3D vessel tree is projected onto the image using the C-arm geometry. In order to help the visual pairing, the 2D projection of the 3D vessel tree was manually aligned, this alignment was not used in the tracking experiments.



**Figure 3.8:** Annotation example for the sequences with contrast agent evaluation: 2D segmented vasculature from contrast agent (left), 2D projection of 3DRA vasculature after manual registration (center) and manual paired vessels used for the evaluation (right). The paired vessels (between the 2D vasculature and the 2D projection of 3DRA vasculature) are presented in the same color.

### 3.3.3.2 Experiments with motion sequences

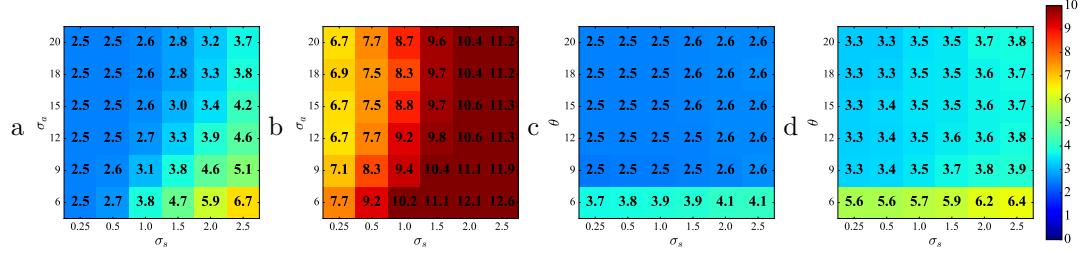
The method is applied on every sequence and the tracking results are evaluated visually. For each second of each sequence, we report one of the three following observations: “Correct tracking”, “Incorrect tracking but clinically relevant”, “Incorrect tracking and no clinical relevance”. Clinical relevancy means that the the 2D projected roadmapping of the vasculature still gives useful visual information.

During the procedure, the physician regularly moves the table and the C-arm detector, and changes magnification or the field of view. Most of the time, when the catheter is closer to the tumor, the field of view is centered at the tip of the catheter. In those cases, only the distal part of the catheter is visible and the registration is challenging. To evaluate the tracking method also in those cases, for each fluoroscopic sequence, part of the proximal segmented catheter centerline was removed. At one third (resp. two third) of the sequence, 33% (resp. 66%) of the segmented catheter is removed. The visual evaluation of the tracking method on these sequences with field of view simulation is reported.

## 3.4 Results

### 3.4.1 Patient vessel tree with a moving simulated catheter

Figure 3.9 shows the tracking results of the simulated training set with the two cost functions ( $F_{\text{cost}_1}$  and  $F_{\text{cost}_2}$ ), the two transition distributions ( $a'$  and  $a''$ ), and for various  $\sigma_a$ ,  $\theta$  and  $\sigma_s$ . The median distance between the 3D tip and the 3D registered tip is below 2.6 mm for more than half of the experiments (67%) with the cost function  $F_{\text{cost}_1}$ . Results with the transition distribution  $a''$  have a median distance smaller than those with the distribution  $a'$  except for a few experiments (4%). For the following experiments, we choose the optimal settings with the experiments giving the shortest median distance. Table 3.1 summarizes the chosen parameters. The experiment could have been done with  $\sigma_s$  smaller than 0.25 mm to check if the median distance would



**Figure 3.9:** Median distance between the 3D tip and the 3D registered tip (in mm) for each frame of the simulated sequence training set with different  $\sigma_a$ ,  $\theta$  and  $\sigma_s$  (in mm). Experiments with the cost function  $F_{\text{cost}_1}$  and the transition distribution  $a'$  (a), with  $F_{\text{cost}_2}$  and  $a'$  (b), with  $F_{\text{cost}_1}$  and  $a''$  (c), with  $F_{\text{cost}_2}$  and  $a''$  (d).

**Table 3.1:** Optimal settings resulting from the simulated experiments. The bold line is the optimal one.

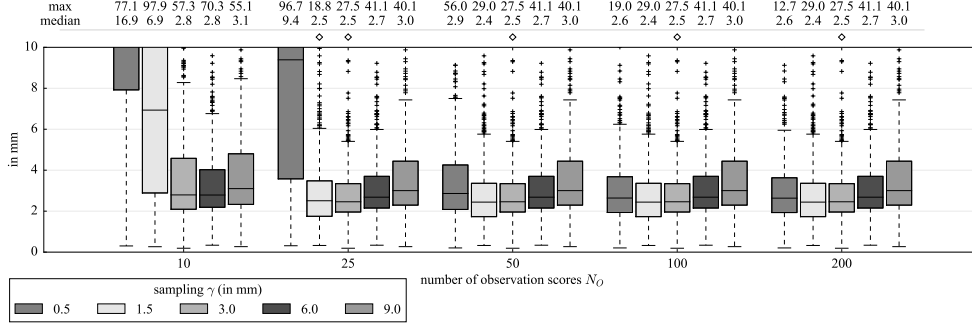
Transition distribution	$\sigma_a$	$\theta$	Registration cost function	$\sigma_s$	$N_O$	Sampling $\gamma$
$a'$	9 mm	N/A	$F_{\text{cost}_1}$	0.25 mm	25	3 mm
$a'$	12 mm	N/A	$F_{\text{cost}_2}$	0.25 mm	25	3 mm
<b><math>a''</math></b>	<b>N/A</b>	<b>9 mm</b>	<b><math>F_{\text{cost}_1}</math></b>	<b>0.25 mm</b>	<b>25</b>	<b>3 mm</b>
$a''$	N/A	18 mm	$F_{\text{cost}_2}$	0.25 mm	25	3 mm

decrease. However, the pixel size of the fluoroscopic sequences is on average 0.23 mm and thus a smaller  $\sigma_s$  would be beyond the pixel resolution, and the accuracy of the catheter centerline extraction. The cost function  $F_{\text{cost}_1}$  and the transition distribution  $a''$  were chosen for the subsequent experiments. Though the median distance is similar for both transition distribution  $a'$  and  $a''$ ,  $a''$  is less sensitive to the parameter settings  $\sigma_s$  and  $\theta$ . Additionally, the implementation of  $F_{\text{cost}_1}$  and  $a''$  is also simpler and the computation time is lower.

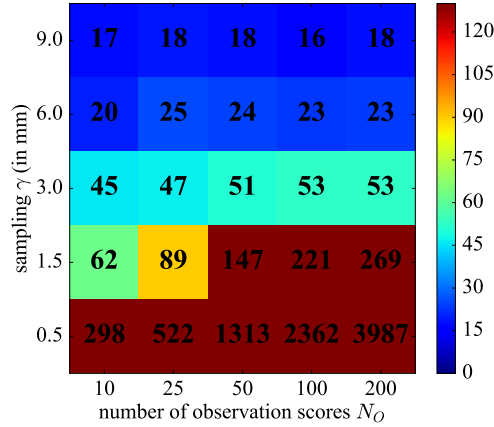
The optimal settings have also been computed for the samplings  $\gamma = \{0.5, 1.5, 6, 9\}$  mm with  $N_O = 200$ , cost function  $F_{\text{cost}_1}$  and transition distribution  $a''$ . For all the samplings  $\gamma$ , we obtain the optimal settings  $\theta = 9$  mm and  $\sigma_s = 0.25$  mm. With these settings, the impact of the centerlines sampling  $\gamma$  and the number of observation scores  $N_O$  is shown in Figure 3.10. The experiments with the sampling  $\gamma = 1.5$  and 3 mm obtain the smallest median distances (less than 2.5 mm) and standard deviations, with a number of observation scores  $N_O$  greater than 50, and greater than 25 for the 3 mm sampling. The average registration time is less than 53 ms for all the experiments with a sampling greater than 3 mm (Fig. 3.11). For all the following experiments, the 3 mm sampling and  $N_O = 25$  are chosen as a tradeoff between accuracy, robustness and computation time (see Table 3.1).

Figure 3.12 shows the distance between registered tip and real tip during the tracking with the test set and optimal settings. With the settings that result from the previous experiments and the definition described in the ‘‘Experiments and pa-

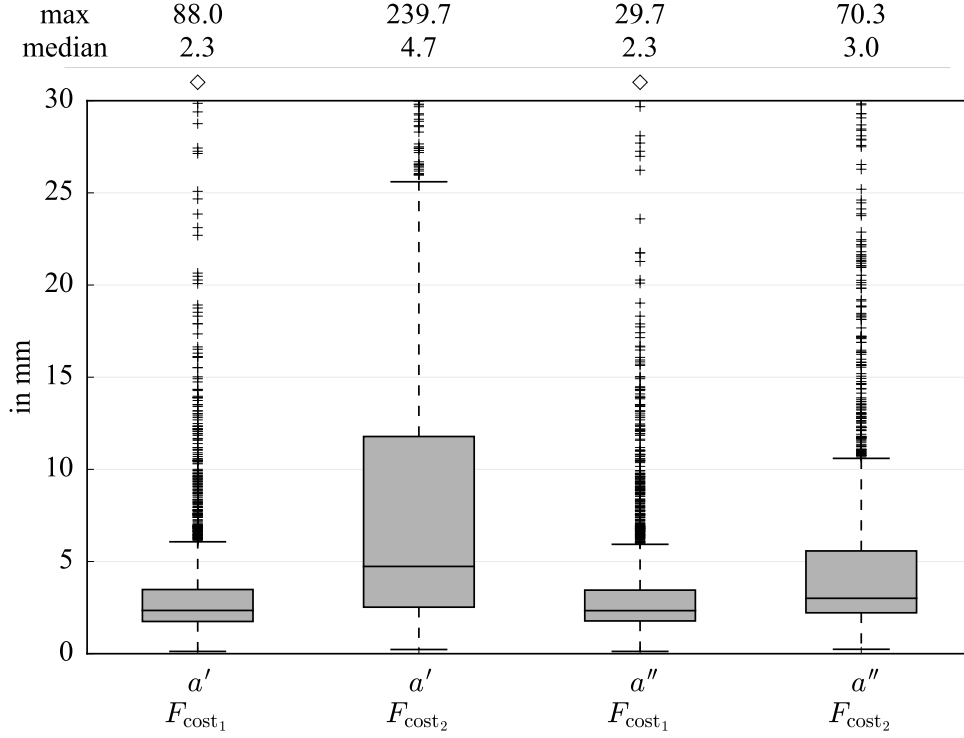




**Figure 3.10:** Distance between the 3D tip and the 3D registered tip (in mm) for each frame of the simulated sequence training set with different samplings  $\gamma$  (in mm), number of observation scores  $N_O$  and optimal settings ( $\theta = 9$  mm,  $\sigma_s = 0.25$  mm,  $F_{\text{cost}_1}$  and  $a''$ ).  $\diamond$  indicates that there is no statistically significant difference (p-value  $> 0.05$  with a Wilcoxon signed-rank test) between this experiment and the experiment with optimal settings (with  $\gamma = 3$  mm sampling and  $N_O = 25$ ). The boxes of the boxplot report first and third quartiles, and median values. The whiskers report  $\pm 1.5$  times the inter-quartile range around the box.



**Figure 3.11:** Average registration time (in ms) for each frame of the simulated training set with different samplings  $\gamma$  (in mm) and number of observation scores  $N_O$ .



**Figure 3.12:** Distance between the 3D tip and the 3D registered tip (in mm) for each frame of the simulated sequence test set with the optimal settings (see Table 3.1).  $\diamond$  indicates that there is no statistically significant difference (p-value  $> 0.05$  with a Wilcoxon signed-rank test) between this experiment and the experiment with optimal settings (cost function  $F_{\text{cost}_1}$  and transition distribution  $a''$ ). The boxes of the boxplot report first and third quartiles, and median values. The whiskers report  $\pm 1.5$  times the inter-quartile range around the box.

rameters optimization” section, the tip tracking failed to reach the last tip position in less than 16.7% of the sequence test set.

### 3.4.2 Clinical data

Figure 3.13 shows the average distance between paired vessels (within 3 cm radius from the tip) for all the sequences with contrast agent after the tracking and registration. The optimal settings are used and the two cost functions and two transition distributions are evaluated. We also compare with our previous method [2]. The transition distribution  $a''$  with the cost function  $F_{\text{cost}_2}$  gives the lowest median average distance (4.6 mm) but there is no statistically significant difference (p-value  $> 0.05$  with a Kruskal-Wallis test) between the different transition distributions and cost functions. Our previous none-tracking method yielded a larger median average distance and the distribution is wider as well. The average computation time of the

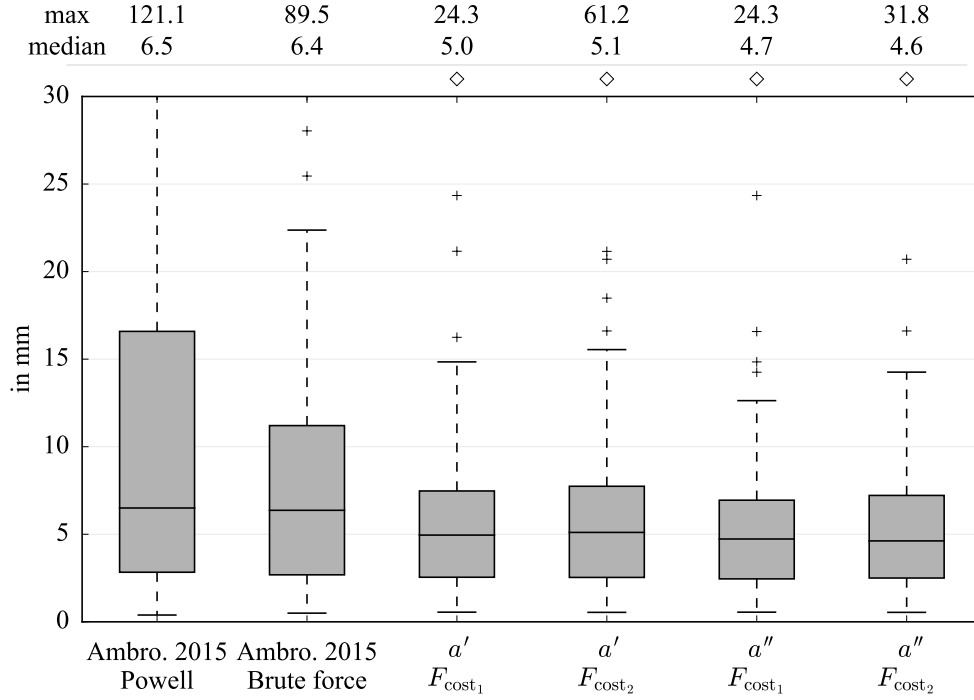
previous method is 236 ms with Powell optimizer and 5.2 s with brute force. The four combined tracking methods have an average computation time of 64 ms.

The tracking method with optimal settings (see Table 3.1) has been applied on the original 10 motion sequences and on the 10 motion sequences with field of view simulation (part of the catheter removed) (Fig. 3.14). Tracking results based on visual observations are depicted in Figure 3.15. Two video clips showing the tracking results are available as supplementary material.<sup>2</sup> The 3D vasculature (in green) is projected on the fluoroscopy and only the vessel branches after and in the vicinity of the catheter tip are displayed. The tip tracking is on average visually correct on eight motion sequences (Fig. 3.15). “Incorrect tracking but clinically relevant” means that the tip tracking is incorrect but the 2D projected roadmapping of the vasculature still gives useful visual information. For example, in the end of the sequences 9 and 10, the tracked tip is shifted or enters the wrong vessel but the resulted registration can still be used as a roadmap. In sequence 9, the tracked tip enters in a loop-shaped vessel parallel to the correct path. In sequence 10, the 3DRA image has to be rotated significantly (at least  $10^\circ$  around the cranial-caudal axis) from its initial position to fit with the catheter position. The limited search space during the 3D/2D registration prevents the optimizer to reach the correct fit and therefore the tracking tip is not very smooth. Sequences 3 and 6 have the same problem but also have a catheter shape which is considerably deformed compared to the vessel shape in the 3DRA image. The conclusions for the tracking with the simulated field of view are the same, except for sequence 8 at eight seconds, where the remaining simulated catheter part is short. In this case, the lack of discriminative shape information in the remaining visible part of the catheter prevents to have a correct 3D/2D registration. On the other hand, in sequence 3 (compared to the tracking with unaltered motion sequences), after four seconds, the simulated field of view is helping the tracking because the beginning of the catheter is not visible and this part is actually difficult to align due to its large deformation compared to the vessel extracted from the 3DRA. Here we note that the HMM method mainly kept two different tracking paths more than 3 cm apart. The tracking in the first four seconds is in the wrong path and as soon as the score is getting higher in the correct path, the tracking switches immediately to this one.

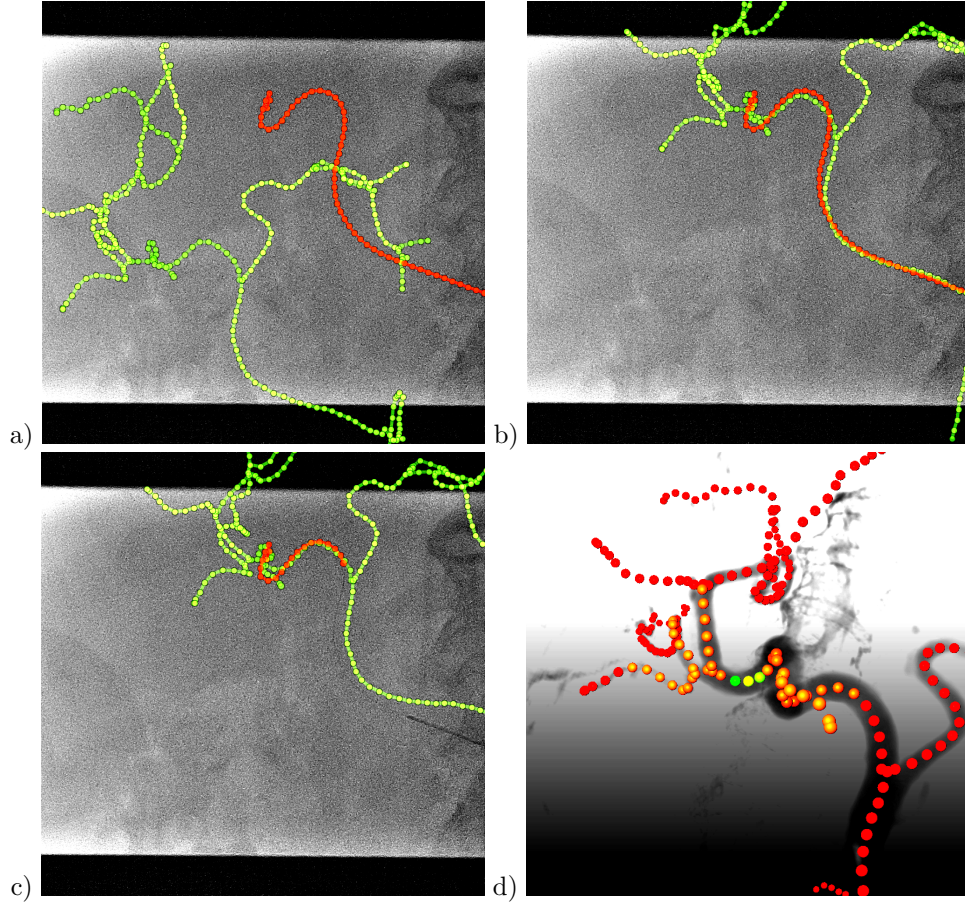
### 3.5 Discussion and conclusion

We proposed a method to track the tip of catheter in the 3D vasculature (extracted from 3DRA) that uses an HMM to model the catheter tip probabilities, and 2D catheter sequences (manually segmented from 2D catheterization single-plane fluoroscopic images). In the HMM, every 3D position in the 3D blood vessel is associated with the probability that the catheter tip is at that position; additionally we model the transitions for the catheter to move from that 3D position to a different one.

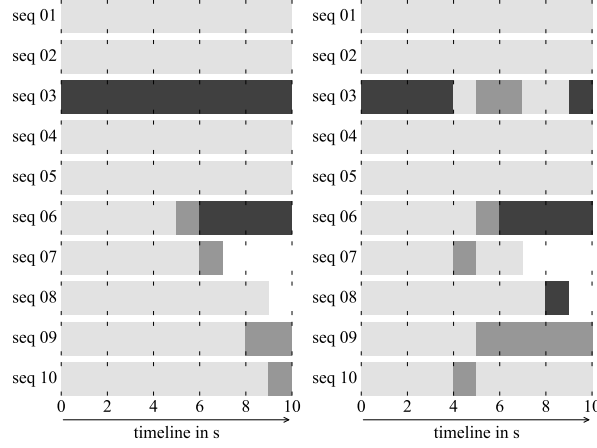
<sup>2</sup>Two video clips are available as supplementary downloadable material on <https://sites.google.com/view/ambrosini>



**Figure 3.13:** Average distance between paired vessels (within 3 cm radius from the tip) for each sequence with contrast agent after registration (in mm). The optimal settings were used (see Table 3.1).  $\diamond$  indicates that there is no statistically significant difference (p-value  $> 0.05$  with a Kruskal-Wallis test) between this experiment and the experiment with optimal settings (cost function  $F_{\text{cost}_1}$  and transition distribution  $a''$ ). Kruskal-Wallis test has been used because we do not have paired samples (number of paired vessels can differ following the tip position). The boxes of the boxplot report first and third quartiles, and median values. The whiskers report  $\pm 1.5$  times the inter-quartile range around the box.



**Figure 3.14:** Tip tracking with the catheter (in red) and the projection of the vessel tree (in green). Before the tracking with 3DRA at its initial position (a), breathing and table motion are the causes for the misalignment. After the tracking (b). After the tracking with the simulated field-of-view (shorter catheter visible) (c). Tracking in the 3D view (d), from a different angle than the projection, with the colored score  $\delta_t$  for every point (scale between red (score = 0) and green (best score)).



**Figure 3.15:** Visually-based observations of the tip tracking over time for the 10 motion sequences: tracking of the unaltered sequences (left) and tracking of the sequences with field of view simulation (right). For each second of the trackings, we attribute one of the three following observations: “Correct tracking” (light gray), “Incorrect tracking but clinically relevant” (dark gray) and “Incorrect tracking and no clinical relevance” (black).

With the optimal settings, the simulated catheter results have a median 3D distance of 2.3 mm between the real-tip and the registered tip position. The results of the experiments with simulated catheters show that a transition matrix with constant probabilities ( $a''$ ) is less sensitive to parameter changes (we obtain a smaller median distance with most of the parameters) than when using Gaussian distributed transition probabilities ( $a'$ ). This may be explained by the fact that in case of constant transition probabilities, the probabilities derived from the registration cost function are more discriminative and thus will lead to select the best registration instead of choosing a less well registered one because that one has a higher prior probability due to the non-constant transition matrix. In the simulated catheter experiments, the error is larger with the cost functions  $F_{\text{cost}_2}$  (taking into account the radius) than with  $F_{\text{cost}_1}$ . The cost function  $F_{\text{cost}_2}$  apparently does not constrain the catheter/vessels alignment sufficiently, which may lead to an inaccurate tip position. The catheter and vessel deformation simulation may also be insufficiently realistic to demonstrate added-value of taking into account the vessel diameter during the registration.

The number of observation scores  $N_O$  is directly related to the discretization  $\gamma$  of the 3D vessel tree and the 2D catheter. Intuitively, if the sampling  $\gamma$  is denser, the number of observation scores  $N_O$  must be increased to maintain similar results. This assumption is verified with the experiments. The accuracy and the robustness of the tip tracking are optimal when the vessels and the catheter are sampled with  $\gamma = 3$  mm (resp. 0.5 or 1.5 mm), and the number of observation scores to evaluate  $N_O$  is more than 25 (resp. 50 or 200). In our case, a sampling of  $\gamma = 3$  mm and  $N_O = 25$  provide a reasonable tradeoff for real-time registration.

The experiments with clinical data with a catheter that is not advanced, and

where contrast agent is introduced, yield a median average distance between paired vessels of 4.7 mm using optimal settings. Here, both cost functions  $F_{\text{cost}_1}$  and  $F_{\text{cost}_2}$  have similar results. The deformations of vessels and catheter in clinical data are much larger than our simulated data, therefore the registration cost function  $F_{\text{cost}_2}$  which takes the radius into account is probably more effective than with simulated catheters.

The curvature of the vasculature and catheter is relevant in the tracking. A unique catheter shape (with high curvature) originating from a unique shape of the vasculature will facilitate the registration. In our previous method [2] we propose a method based solely on the catheter shape. The tracking was generally successful in cases with a unique catheter shape but our experiments also demonstrated that the registration can fail when the shape is not sufficiently discriminant. The tip tracking in our current method enables a more robust registration when the catheter has a simple shape or is rather short.

The tracking and the associated 2D or 3D roadmap can be used for guidance to navigate the catheter through vasculature, without additional contrast agent injections. According to our clinical partners, a median vessel distance error of 4.7 mm is sufficient for them to localize the instrument in the vessel tree, and helps them guiding the catheter to the correct location. In practice, for the physicians, consistency in tip tracking is more important for guidance than the actual vessel position around the tip. Also, the experiments with clinical data with the catheter being advanced, demonstrated that the tracking is robust even when the tracking is lost for some time or when only a short part of the catheter is visible. Although the lack of a reference standard for these cases prevents us from quantifying the results, the tracking and resulting roadmap provide visual information which may help physicians to maneuver the catheter more efficiently, and possibly to use less contrast agent. Future clinical studies need to be conducted to validate this assumption.

Most of the methods in the literature register a 3D pre-operative image with 2D X-ray contrast enhanced images. Some of them obtain submillimeter accuracy. We cannot directly compare our results with those methods because we use 2D fluoroscopic images with no contrast in which only the catheter is visible, making it a much more challenging task. In order to compensate the breathing motion, Atasoy et al. [5] updated the initial 3D vasculature position using also only a small part of the catheter (extracted in a manually selected region of the X-ray image). They evaluated on one clinical sequence by computing the overlap of the catheter with the vasculature, resulting in 70% overlap. Our tracking method with the motion sequences obtains on average 80.3% overlap. However, this evaluation metric does not give information about the catheter tip and the vessels around the tip.

To apply this method in the angiography suite, the 2D segmentation of the catheter has to be done automatically during each 2D X-ray frame. This task is challenging but it has been already addressed in several studies, such as Heibel et al. [34]; they demonstrated a catheter tracking accuracy of less than 1.5 pixels on liver catheterization fluoroscopic sequences. Such accuracy should be sufficient for application in clinical practice. Other methods propose to track the catheter in fluoroscopic images such as [86, 97, 98]. The 2D catheter detection as well as the 2D tip detection are important because the registration is solely based on the catheter

centerline. We did not evaluate the error resulting from a missing tip part. However, given our approach, we expect that a missing tip part will result in an additional tip position error (in 3D) of approximately the length of the missing part. On the other hand, the roadmap should still be correct and could still help the physician to guide the catheter. In addition, the catheter tip position in the 3D vessel tree needs to be initialized once, after the 3DRA acquisition, by manually indicating the tip. This task is simple and fast, and could also be automated because the catheter can be easily identified in 3DRA.

Our method could be extended in several directions. A non-uniform sampling of the catheter and the vessel tree centerline could be used to guide the optimizer during the registration to align features such as high-curvature regions. The matrix  $A$  of state transition probabilities could similarly be adapted, e.g. at region of high curvature and also branching points. In the liver, as motion of the vessels is mainly due to respiration, the motion transformation should be small between each frame and a periodic respiratory motion is expected. It could be useful to include a motion prior in the registration transformation model.

To conclude, we presented a 3D catheter tip tracking method using an HMM with 2D fluoroscopic sequence and 3DRA. Experiments on simulated data demonstrated a median tip tracking distance error of up to 2.3 mm. On clinical data, the results demonstrate a robust tracking in cases where the catheter is advanced. In case of no advancement of the catheter, the registration yields a median distance error of less than 4.7 mm on vessels close to the tip. These accuracies indicate that the method could become a promising tool for improving image guidance in liver catheterization procedures.

**Acknowledgements** This research was funded by Philips Healthcare, Best, The Netherlands. We thank the Hôpitaux Universitaires Henri Mondor, Créteil, Paris, France and the Ospedale di Circolo e Fondazione Macchi, Varese, Italy for providing image datasets.



# Fully Automatic and Real-Time Catheter Segmentation in X-Ray Fluoroscopy

---

**Abstract** — Augmenting X-ray imaging with 3D roadmap to improve guidance is a common strategy. Such approaches benefit from automated analysis of the X-ray images, such as the automatic detection and tracking of instruments. In this paper, we propose a real-time method to segment the catheter and guidewire in 2D X-ray fluoroscopic sequences. The method is based on deep convolutional neural networks. The network takes as input the current image and the three previous ones, and segments the catheter and guidewire in the current image. Subsequently, a centerline model of the catheter is constructed from the segmented image. A small set of annotated data combined with data augmentation is used to train the network. We trained the method on images from 182 X-ray sequences from 23 different interventions. On a testing set with images of 55 X-ray sequences from 5 other interventions, a median centerline distance error of 0.2 mm and a median tip distance error of 0.9 mm was obtained. The segmentation of the instruments in 2D X-ray sequences is performed in a real-time fully-automatic manner.

## 4.1 Introduction

Minimally invasive procedures are generally preferred over open surgery interventions, as these localized and accurate interventions lead to less trauma and shorter recovery times than conventional procedures. Minimally invasive procedures require real-time imaging to visualize the relevant anatomy and the instruments. Particularly, in catheterization procedures, a catheter<sup>1</sup> is inserted into the body via the vasculature and fluoroscopic imaging is used to continuously visualize the catheter. The vasculature is only visible in X-ray images when contrast agent is injected, and contrast agent is used sparingly because of its toxic nature. Therefore, recent approaches for virtual roadmapping that permit the visualization of a 3D vessel tree from pre-operative images have been presented [4, 80]. Such methods benefit from automated extraction of the instruments from fluoroscopic images. The purpose of this work was therefore to develop and evaluate a method that segments fully automatically the catheter in 2D single-plane X-ray fluoroscopic sequences in real-time.

Automatic catheter segmentation is not straightforward, as the catheter is a thin, moving structure with low contrast in noisy images. Segmentation methods for electrophysiology (EP) electrodes and EP catheter in 2D X-ray images have been reported [12, 101]. EP electrodes are clearly visible, and their location is often used to obtain a full segmentation of EP catheters. Segmentation of catheters without features such as electrodes has been studied less frequently. Most methods enhance the instruments with Hessian-based filters, which are followed by a spline fitting approach, starting from the catheter shape of the previous frame [6, 15, 16, 34, 86, 98]. These methods have two drawbacks: the first frame of the fluoroscopic sequence has to be manually annotated and the curvature and length of the catheter should not change much between frames. [15, 34] propose semi-automatic methods to segment the first frame. Recently, a fully automatic method using directional noise reduction and path extraction, with segments and similarity from the previous frame cost function, has been proposed [97]. The method was evaluated on the last frame of 7 sequences from one canine study on which it performs well; it is, however, not a real-time method. We summarize in Table 4.1 the methods proposed in the literature in order to show results they obtained. Note that the results, the accuracy metrics and computation times cannot be directly compared but they give an idea of the performances.

Our method utilizes deep convolutional neural networks (CNNs) for the segmentation. CNNs have been demonstrated to be very effective in image classification and image segmentation [54], also in case of medical images with a limited set of annotations [62, 79]. Ronnerberger et al. [79] introduced an end-to-end biomedical imaging segmentation network called U-net: a model with a fully convolutional part (downsampling part) and a deconvolutional part (upsampling part) which outputs after appropriate thresholding a binary segmented image. Extensive data augmentation enables the neural network to generalize well, even in case of small training sets. Several improvements w.r.t. the network and the training process have been introduced more recently. The downsampling (resp. upsampling) part has been shown to

<sup>1</sup>For clarity, in the remainder of the paper, the word “catheter” also refers to the micro-catheter and guidewire instruments. Although they have quite different appearances, they are handled altogether as one instrument in the proposed method.

**Table 4.1:** Summary of the methods in the literature and the method of this paper.

References	Fully Auto.	Time	Accuracy	Tip Accuracy
2003 Baert et al.[6]	No	5 s	mean 0.9 px <sup>a</sup>	mean < 2 mm <sup>a</sup>
2007 Slabaugh et al.[86]	No	175 ms	-	-
2009 Wang et al.[98]	No	500 ms	mean 2 px (0.4 mm)	mean 5.4 px
2012 Heibel et al.[34]	No	60 ms	mean 0.8-3.9 px	-
2016 Chang et al.[15]	No	-	-	-
2016 Chen et al.[16]	No	-	mean 2.1 px (0.5 mm)	-
2016 Wagner et al.[97]	Yes	> 1 min	mean 0.5 mm	-
This work	Yes	125 ms	median 0.2 mm	median 0.9 mm

<sup>a</sup>The failed segmentations are not included in the evaluation

be more effective with strided convolution (resp. transposed convolution) than with max pooling [62, 88]. Strided convolution enables to learn how the features should be downsampled/upsampled. Moreover, batch normalization [39] and residual learning [33] have been proposed to improve training convergence.

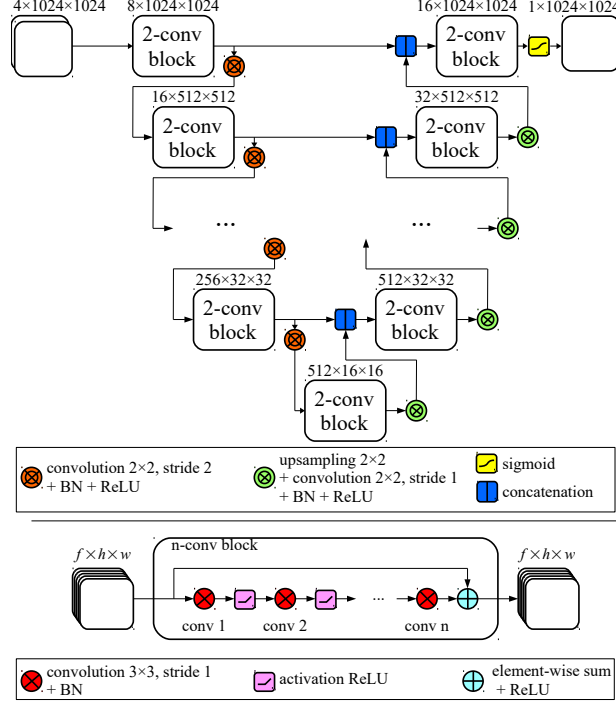
2D X-ray fluoroscopic images are very noisy and are used in liver catheterization procedures to guide catheter inside the liver vessel tree. The catheter used does not have specific features, such as electrodes, nor specific shapes that may facilitate their segmentation. The framerate is around 7Hz, so, during the catheter manoeuvre, its shape and length may change considerably between two consecutive frames. Tracking the catheter over time is therefore quite challenging. In this paper, we propose a fully automatic segmentation method based on the U-net model combined with recent strategies to improve the training of the network, such as batch normalization, residual learning and a data augmentation scheme to increase the size of the training dataset. The catheter centerline is then extracted using skeletonization and linking of the extracted branches. Our work differs from previous approaches in that we introduce a fully automatic approach that can be run in real-time.

## 4.2 Method

The catheter is segmented using the CNN and then the centerline is extracted from the result of the CNN using skeletonization and subsequent linking of branches.

### 4.2.1 Data

A 2D X-ray image sequence is a set of  $s$  2D images  $S = \{I_1, I_2, \dots, I_s\}$ . Each image  $I_i$  is associated with a binary image  $B_i$  where the catheter pixels have a value of 1 and the background pixels 0. We also associate the output from the neural network, the image prediction  $B_i^p$ , where each pixel is between 0 and 1, and a pixel closer to 1 is considered as a catheter pixel. Our neural network model is trained to predict a mask  $B_i^p$ , given an image  $I_i$  as input.



**Figure 4.1:** Neural network model (top):  $I_i, I_{i-1}, \dots, I_{i-3}$  images are fed into the model and segmented image  $B_i^p$  is predicted. The model is composed of n-conv block at each layer. An n-conv block is  $n$  consecutive convolutions of input features (with  $f$  filters, height  $h$  and width  $w$ ) with a residual connection to the output [33, 62] (bottom).

#### 4.2.2 CNN Model

The model is an adapted version of the well-known U-net model [79]. The input of the model is the current image  $I_i$  and previous frames  $I_{i-1}, \dots, I_{i-3}$ . The output is the image prediction  $B_i^p$ . For the network topology, see Figure 4.1. To improve convergence speed during training, we add batch normalization (BN) after every convolution [39]. In order to also learn how to downsample/upsample the features we use strided convolutions [62, 88]. To prevent overfitting we add dropout at the end of the two last blocks in the downsampling part [89].

#### 4.2.3 Training

The loss function is based on the Dice overlap metric [62] between the ground truth mask  $B_i$  and the output image  $B_i^p$  of the model, defined as:

$$L_{\text{Dice}}(B_i, B_i^p) = - \frac{2 \sum_k B_{ik} B_{ik}^p}{\sum_k (B_{ik}) + \sum_k (B_{ik}^p)} \quad (4.1)$$

where  $B_{ik}$  and  $B_i^p$  are respectively the pixels of the mask  $B_i$  and the output image  $B_i^p$ .

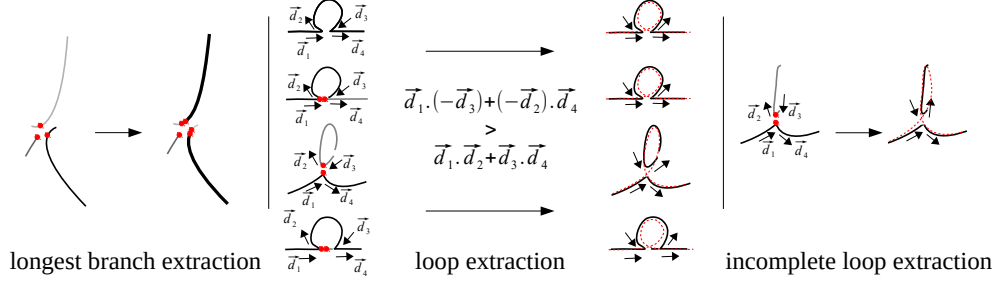
In order to have more data to train and to generalize well, large data augmentation is used during the training. Data augmentation is done on the fly. For every image in the set of training images, there is a 50% probability to augment the image. If the image is augmented, we apply all of the following transformations to both the X-ray image inputs and the corresponding binary image output: 50% probability on a horizontal and vertical flip, a random rotation (around the image center) in a range of  $\pm 9$  degrees, a random scale with a factor in a range of 0.9 to 1.1, a random horizontal and vertical translation in a range of  $\pm 16\%$  of the image size, a random intensity shift with a factor in a range of  $\pm 0.07$  (in normalized image between 0 and 1) and a Gaussian noise with  $\sigma = 0.03$ .

#### 4.2.4 Centerline extraction

The output of the neural network is first thresholded with a threshold  $\alpha$  (between 0 and 1) and then skeletonized [106]. Next, the branches (ordered sets of pixels) are determined based on connectivity. Connection points are created between close branches. If the closest points between two branches are within a distance  $D_{\max}$  pixels, we consider this to be a possible connection, and there can be only one connection between two particular branches. Then, to link the branches, three steps are done (Fig. 4.2). First, for each connection, we divide and merge branches in order to have the longest branches. Second, loops are detected and merged following the direction at the crossing point. We have a loop in a branch when two points, within a distance  $D_{\max}$  pixels, have their distance along the branch of at least  $B_{\min}$  pixels. Before the third step, the first and second steps are repeated a second time with a distance  $D'_{\max}$  superior to  $D_{\max}$ . Finally, in the last step, the remaining connected branches larger than  $B_{\min}$  pixels are considered as incomplete loops or straight loops due to foreshortening. We process them similarly as the second step by closing the two endpoints of their branch. When all the potential links have been processed, we keep the longest connected set of branches and choose amongst the two endpoints the farthest from the image border as the tip of the catheter centerline. Finally, the centerline is smoothed by fitting a spline.

### 4.3 Experiments and results

2D single plane X-ray fluoroscopic sequences have been acquired during 28 liver catheterization procedures in three different hospitals (Erasmus MC, Rotterdam, the Netherlands; Hôpitaux Universitaires Henri Mondor, Créteil, Paris, France; and Ospedale di Circolo e Fondazione Macchi, Varese, Italy) with angiographic C-arm systems (Xper Allura, Philips Healthcare, Best, the Netherlands). Every image is normalized on the range  $[0, 1]$ , mapping the intensities between the 2<sup>nd</sup> and the 98<sup>th</sup> percentile of the image histogram. In 182 sequences from 23 procedures, we manually segmented the catheter in four consecutive images by annotating points and fitting a spline. From the catheter spline we constructed the binary segmented image using



**Figure 4.2:** The three steps for the branches linking. Branches are depicted with different colors and connection points are in red. The red dash curve is the centerline after spline smoothing. The first step links the longest branch parts, the second step links and corrects the loops and the last step links the remaining branches longer than  $B_{\min}$ , considered as an incomplete loop or a straight loop due to foreshortening.

a dilation operator with a  $5 \times 5$  pixel kernel. These sequences were used as training data for the parameter optimization. In 55 sequences from the other 5 procedures, we also segmented four frames per sequence. These 55 sequences will be used as testing data after the model optimization and training.

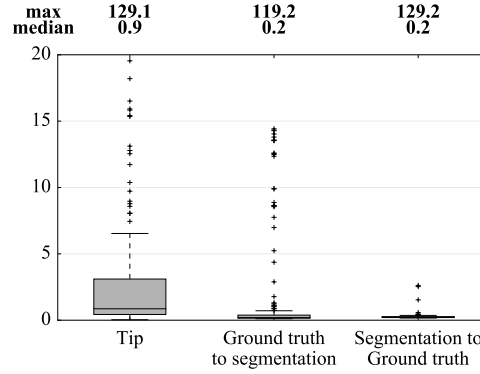
The loss function  $L_{\text{Dice}}$  of the model is optimized using stochastic gradient descent with a learning rate of 0.01, a decay of  $5 \cdot 10^{-4}$  and a momentum of 0.99. Following the training, we set the threshold  $\alpha$  to 0.01, the maximum distance to connect two branches during the first pass (resp. second pass)  $D_{\max}$  to 5 pixels (resp.  $D'_{\max}$  to 20 pixels) and the minimum loop length  $B_{\min}$  to 30 pixels. Using an Nvidia GTX 1080, the average time to segment one image was 125 ms which is suitable for real-time processing. Our method is publicly available <sup>2</sup>.

We evaluate using the tip distance error (i.e. the distance between the annotated catheter tip and the tip of the segmented catheter), and the average distance between the manually segmented catheter and the automatically segmented catheter. Figure 4.3 shows the tip and catheter distances results. We compute the precision of the tip between consecutive frames. The median, average, minimum and maximum of the standard deviation per sequence of the tip distance error are respectively 0.7 mm, 4.9 mm, 0.1 mm and 55.7 mm. Five examples of segmentation are shown in Figure 4.4. In frame **c**, the segmentation is going too far and follows part of the vertebrae. Frame **d** misses the proximal part of the catheter and frame **e** is the only sequence with significant false positives. It is less noisy because it has been acquired with higher radiation dose. The neural network was not trained for such sequence.

## 4.4 Discussion and conclusion

We proposed a fully automatic method to segment catheter on 2D X-ray fluoroscopic images using CNNs. The segmentation on testing data gives a median tip distance error of 0.9 mm and a median centerline distance error of 0.2 mm where 85% of the

<sup>2</sup>Available at <https://github.com/pambros/CNN-2D-X-Ray-Catheter-Detection>



**Figure 4.3:** Tip distance error, ground truth to segmented centerline distance error and segmented centerline to ground truth distance error (in mm) on the test set: 4 frames per sequence on 55 sequences.

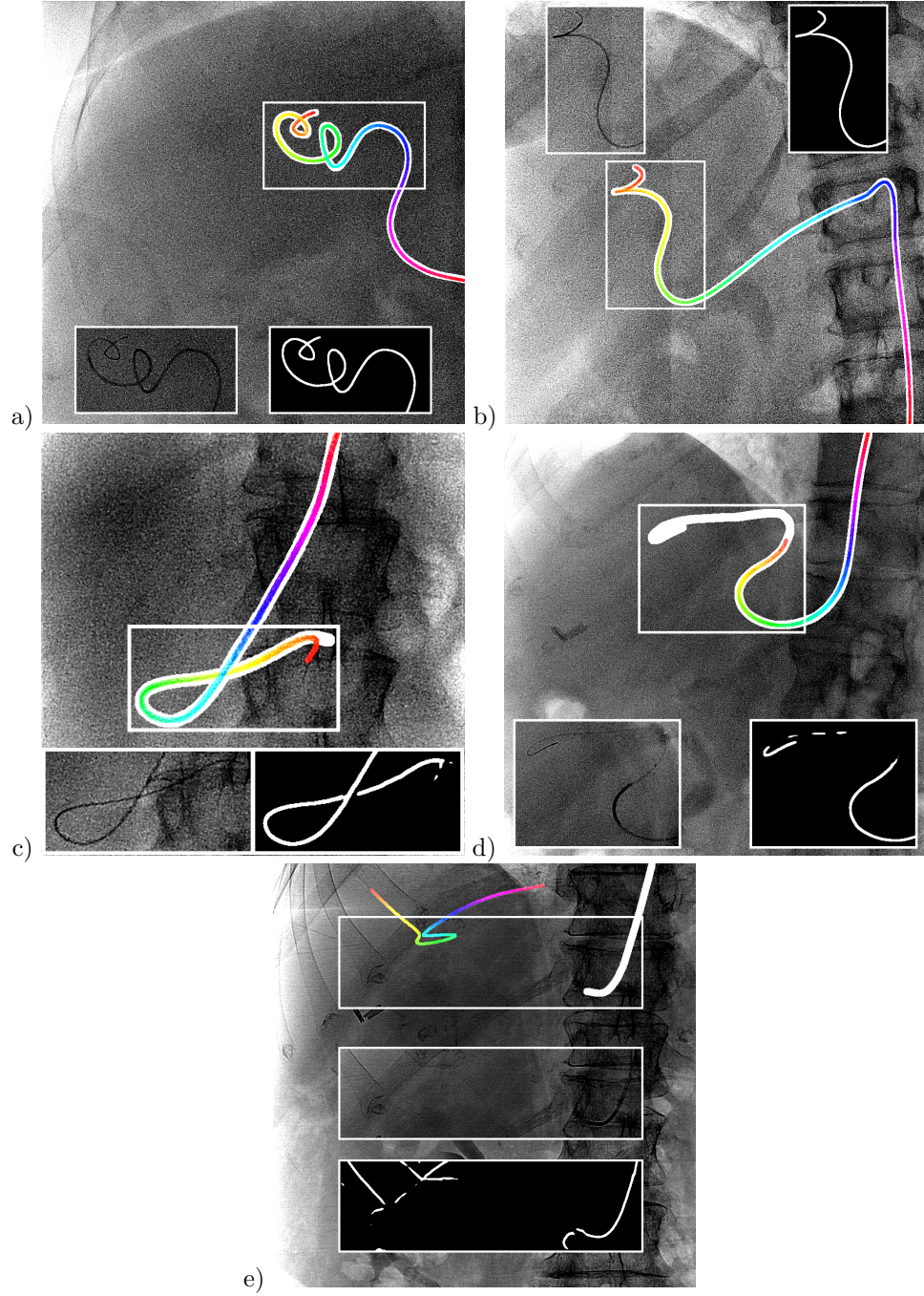
frames have less than 1 mm of centerline distance error. We note that the distance errors are in mm at the X-ray detector scale. The real distance errors at the patient scale are smaller.

Very few images have false positives after the CNN segmentation. Therefore, we can use simple criteria to extract the catheter centerline from the CNN segmentation. The results show that it works well and can handle self-intersections. The main problem in the extracted catheters are sometimes large gaps in the segmentation due to false negatives. As a consequence, occasionally the proximal part of the catheter is missing. With a larger training set, the model is expected to generalize better.

Previous studies show a higher success rate, probably because they manually initialize the tracking process. We, in contrast, do not employ a tracking approach. Whereas it is clear that a stronger incorporation of the time dimension (beyond using consecutive frames in the segmentation) may provide a more robust result, our current results demonstrate that even without tracking good results can be obtained. A major advantage of not utilizing tracking is that the method is not hampered by previously incorrectly segmented frames, and thus automatically can recover from previous failures.

The catheter and guidewire have different thickness and appear quite differently on fluoroscopic images. Whereas we trained one network that segments both, it could be interesting to use two different models and retrospectively combine their results to obtain a more accurate segmentation. The current model is using the previous images to segment the catheter but it could also be useful to use the previous segmentation in the model. Both strategies are future work.

To conclude, we developed and evaluated a CNN-based-approach to fully automatically segment catheters in live fluoroscopic images. With execution times within 125 ms, this method is promising for use in real-time catheter detection.



**Figure 4.4:** 5 segmented frames of the test set. The segmented centerline (dilated for visual purpose) appears from red (tip) to rose. The ground truth (thicker) appears in white. The windows show the original image and the output of the neural network.



# Fast Prospective Detection of Contrast Inflow in X-ray Angiograms with Convolutional Neural Network and Recurrent Neural Network

---

**Abstract** — Automatic detection of contrast inflow in X-ray angiographic sequences can facilitate image guidance in computer-assisted cardiac interventions. In this paper, we propose two different approaches for prospective contrast inflow detection. The methods were developed and evaluated to detect contrast frames from X-ray sequences. The first approach trains a convolutional neural network (CNN) to distinguish whether a frame has contrast agent or not. The second method extracts contrast features from images with enhanced vessel structures; the contrast frames are then detected based on changes in the feature curve using long short-term memory (LSTM), a recurrent neural network architecture. Our experiments show that both approaches achieve good performance on detection of the beginning contrast frame from X-ray sequences and are more robust than a state-of-the-art method. As the proposed methods work in prospective settings and run fast, they have the potential of being used in clinical practice.

## 5.1 Introduction

During percutaneous coronary interventions (PCI), X-ray angiography (XA) is commonly used by clinicians to identify the sites of plaque and navigate devices through the arteries of patients with advanced coronary artery disease. As X-ray imaging has poor soft tissue contrast, coronary arteries are normally visualized by injecting radio-opaque contrast agent in the vessels.

Approaches for improving image guidance in such procedures have been reported, for example fusion of coronary models from CTA [81]. Such methods can only be applied if vessels are visible in the XA, thus automated application of such methods requires detection of presence of contrast agent. Similarly, automated detection of catheter and guidewires, which can also be used for virtual roadmapping [7], is generally only possible in non-contrast enhanced frames. Therefore, an automatic way to detect contrast inflow online is relevant for further automating advanced image guidance methods for coronary interventions, reducing interactions of clinicians with computers during procedures.

Existing works for detection of contrast inflow in X-ray images fall into two categories: enhancement-based and learning-based. Enhancement-based methods [19, 46, 48, 105, 107] enhance contrasted structures, followed by a step to extract features that indicate the change of contrast throughout the sequence. The contrast-enhanced frames are then detected via analysis of the feature. Learning-based approaches [17, 37] train a classifier to detect contrast or non-contrast frames based on handcrafted image features. Among these works, [17, 48, 105] need an entire sequence to detect contrast inflow, and thus only work retrospectively. [46] does not rely on a complete sequence, but retrospectively runs on a sliding segment of a few new X-ray frames, thus there is a trade-off between the possible delay of the contrast inflow detection and the overall processing efficiency. In addition, this method was designed specifically for TAVI procedures on aorta: their contrast detection method involves aligning a predefined aorta shape model to X-ray images and a step of TEE probe detection, which is not relevant for coronary interventions. [107] uses a heuristic approach to detect the first contrast-enhanced frame from X-ray sequences of left atrium (LA) used for electrophysiology (EP) ablation procedures. [37] developed a learning-based framework on X-ray images of LA for EP procedures. The method used a SVM classifier with the heuristic features introduced in [19] and [107]. Out of these methods, [19] is the only one that may be directly used for coronary interventions and work in prospective settings.

The purpose of our work is to develop and evaluate solutions for prospective detection of contrast inflow in XA images that can fit into the clinical work-flow of coronary interventions. Specifically, we aim at prospectively detecting if a frame has contrast agent. To this end, two different approaches were developed. Due to the exceptional performances that convolutional neural networks (CNN) have in image classifications [45], and medical applications, such as tissue segmentation and surgical tools detection [29], we propose a learning-based method using CNN to classify each frame of an XA sequence into two classes: with or without contrast. Additionally, we propose a hybrid of enhancement- and learning-based. It computes a temporal contrast feature from vessel-enhanced sequences based on which contrasted frames

are detected with long short-term memory (LSTM) [36], a recurrent neural network (RNN) architecture. To the best of our knowledge, this is the first work that applies deep learning for contrast inflow detection in X-ray images. To validate the detection, the position of the beginning contrast frame (BCF) [17, 37] in a sequence (where contrast starts being visible) was used in the experiment.

## 5.2 Methods

### 5.2.1 The CNN-based method

Let  $S = \{I_1, I_2, \dots, I_n\}$  denote a sequence of  $n$  frames in which  $I_c$  is the beginning contrast frame. All frames  $I_1, \dots, I_{c-1}$  are associated with the label “without contrast”. The other frames  $I_c, \dots, I_n$  have the label “with contrast”.

In order to classify the fluoroscopic frames, we used a CNN to learn the difference between the contrast frame and non-contrast frame (Fig. 5.1, top). The input of the CNN has 5 images: the current frame  $I_i$  to be classified, its 3 previous frames  $I_{i-1}, I_{i-2}, I_{i-3}$ , and the first frame  $I_1$  (normally non-contrasted). There are 7 intermediate layers directly after the input layer, each of which has a  $n$ -conv block with  $n$  consecutive convolutions (Fig. 5.1, bottom). The last  $n$ -conv block is connected with two fully-connected layers. The final output is a softmax layer with two nodes: “with contrast” and “without contrast”. The model was trained with binary cross-entropy as the loss function. In order for a faster convergence, batch normalization was used after every convolution, residual connection at every layer and the strided convolutions instead of pooling layers.

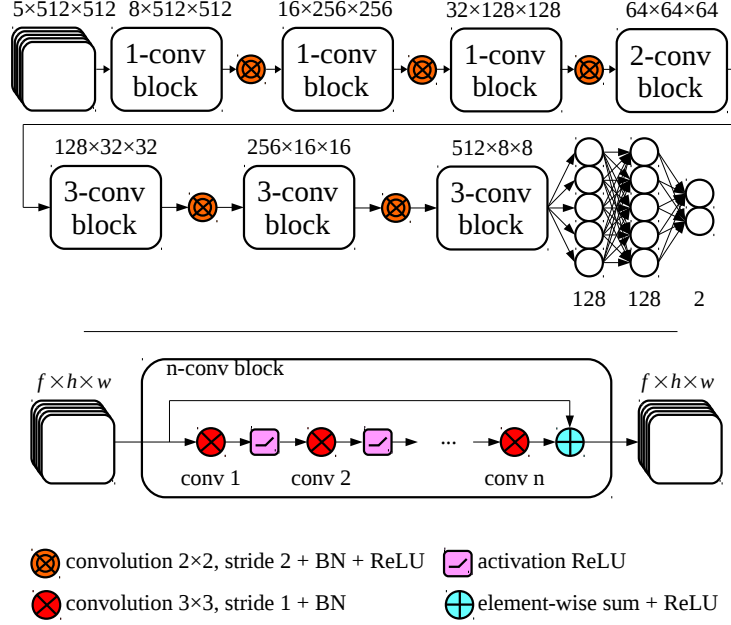
To detect the BCF of an XA sequence online using the trained model, frames of the sequence were classified one by one in a chronological order. The first frame labeled as “with contrast” in the sequence is considered as BCF.

### 5.2.2 The RNN-based method

The RNN-based method consists of two major steps: vessel enhancement and contrast frame detection. An overview of this method is illustrated in Fig. 5.2.

**Vessel enhancement** The vessel enhancement step is crucial for accurate approximation of contrast changes in XA sequences. This step removes most non-vessel background structures using a previously developed online layer separation technique [57] followed by multi-scale Frangi-vesselness filtering [25].

The online layer separation method prospectively separated an XA sequence into three layers: a breathing layer, a quasi-static background layer, and a vessel layer in which vessels have better visibility. First, the breathing layer was separated via morphological closing. After this layer was removed from the original image, online robust PCA (OR-PCA) [22] was applied to separate the low-rank quasi-static layer and sparse vessel layer through alternatively projecting the new data sample (frame) to the underlying low-rank subspace basis and updating the basis using the new estimation of the layers. After layer separation, the structures that may cause artefacts in the next step, such as diaphragm, spine, were removed from the vessel layer.



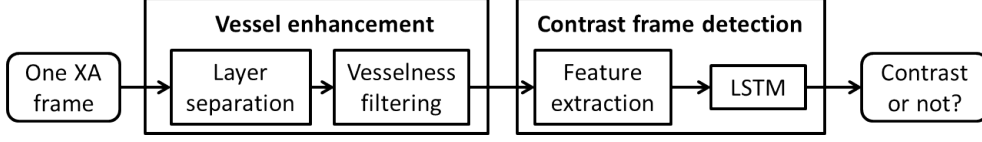
**Figure 5.1:** The neural networks (top) connects the 5 input images (the first, the current and its 3 previous X-ray frames) to the 2 output nodes (“with contrast” and “without contrast”). The model consists of several  $n$ -conv blocks. They are a succession of CNNs with a skip connection between the input and the output of the block (bottom).  $f \times h \times w$  is the dimension of the data (feature number times image height times image width).

Following the layer separation, a multi-scale vesseness filter [25] was applied on the separated vessel layer to further enhance the tubular structures. In the end, after the vessel enhancement step, for each incoming frame, a new image was created where vessel structures are enhanced.

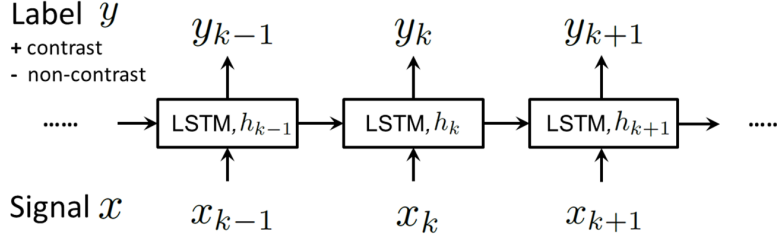
**Contrast frame detection** Once the image with enhanced vessel structures is obtained, the feature that indicates the level of contrast agent was extracted from the image. In this work, we used the average pixel intensity of the complete vessel-enhanced image as the contrast feature. This results in a 1D signal for a complete sequence.

The last step is to detect contrast frames from the previously obtained 1D contrast signal. In order to fully use the temporal relation between frames, each signal point is classified as “contrast” or “non-contrast” with a recurrent neural network. The long short-term memory (LSTM) network [36] was used due to its good performance on modeling long-term temporal relations in time-series data.

Let  $x_k$  denote the feature for the  $k$ th frame  $I_k$ . The single-direction LSTM takes  $x_k$  as the input. A hidden state  $h_k$  in the LSTM network is recurrently updated through nonlinear interactions between the input signal  $x_k$ , the LSTM units and its



**Figure 5.2:** The overview of the RNN-based method.



**Figure 5.3:** Each signal point is classified from a contrast frame or a non-contrast frame with a LSTM network.

state of the last time point  $h_{k-1}$ . The output label  $y_k$  of  $x_k$  is the outcome of a nonlinear function of  $h_k$ . This process is illustrated in Fig. 5.3.

### 5.3 Experiments

We retrospectively obtained anonymized data that was acquired during clinical routine with a Siemens AXIOM-Artis biplane system. The data were 120 XA sequences from 26 patients who underwent a PCI procedure. The frame rate of all sequences is 15 frames per second. The length of sequence varies from 24 to 244 frames. The size of images in our dataset are  $512 \times 512$ ,  $600 \times 600$ ,  $776 \times 776$  and  $1024 \times 1024$ . In all sequences, contrast inflow can be observed. In our experiments, 40 sequences from 20 patients were used as training data, the 80 sequences from the other six patients were used for validation.

For the CNN-based method, all images were resized to  $512 \times 512$  before training. The parameters of the CNN model were optimized using stochastic gradient descent with a learning rate 0.0001, a decay of 0.0005 and a momentum of 0.99. The model was trained with a batch size of 15 during 33,000 iterations. For each sequence, the six frames before and after the BCF were chosen to ensure an even number of contrast and non-contrast training images. The BCF was discarded to assist the CNN to learn more differences between contrast and non-contrast frames. As the dataset used to train the model is small, data augmentation was applied during the training to virtually create more data: translation ( $\pm 100$  pixels), rotation ( $\pm 5$  degrees), scaling ( $\pm$  factor 0.1), intensity shift ( $\pm 0.2$ ), Gaussian noise ( $\sigma_g = 0.01$ ) on the normalized image between 0 and 1, and vertical flip were used to transform images.

For the RNN-based method, we manually tuned the parameters based on visual check and quantitative evaluation on the training data; the same parameters were used

for testing. The images were first down-scaled 2 or 4 times to  $256 \times 256$  or  $300 \times 300$  or  $388 \times 388$  depending on the original image size for speeding up the image processing. The parameters for layer separation were set following the approach with the sliding window option in [57] using the closed-form solution of OR-PCA. To improve the convergence of OR-PCA, we used a mini-batch of 5 frames (before contrast agent was injected) to get an initial estimate of the low-rank subspace basis. This was done using the layer separation method in [56] with fast principal component pursuit [78]. The scale of Frangi vesselness filter was set ranging from 0.6 mm to 2.8 mm according to the size of coronary arteries. The  $\beta$  and  $c$  parameter of the vesselness filter were 0.5 and 15. The dimension of LSTM units was set to 7 with a dropout probability being 0.2. The nonlinear activation function of the hidden layer is sigmoid function. The LSTM network was trained using RMSprop optimizer with a learning rate being 0.005 during 100 epochs. At last, the BCF was detected as the first frame in a sequence being classified as contrasted by LSTM.

In the experiments, we also compared our methods with the state-of-the-art approach of Condurache et al.[19]. For setting the parameters of the method, the first 3 feature values from non-contrast frames were modeled as a Gaussian  $N_0(\mu_0, \sigma_0^2)$ . The threshold  $T$  for choosing contrast frames was set to  $\mu_0 + 3\sigma_0$ .

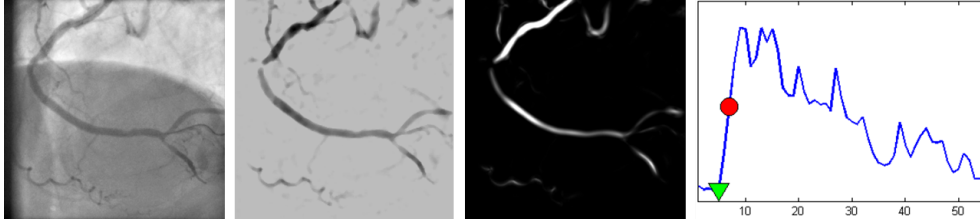
The evaluation metric we used is the absolute difference between the frame index of the ground truth BCF and the frame predicted by different methods.

The image processing steps in the RNN-based method and the method of Condurache et al. were implemented in MATLAB with a single CPU core (Intel Core i7-4800MQ 2.70 GHz). LSTM and CNN were implemented in Keras with TensorFlow as backend. LSTM was running on the CPU due to its small dimension. CNN was trained and tested on an Nvidia GeForce GTX 1080 GPU.

## 5.4 Results and discussion

Fig. 5.4 shows an example to illustrate steps in the RNN-based method. The statistics of the absolute errors made by the three methods are shown in Table 5.1. The results of the mean and median errors show that the two proposed approaches have smaller errors than the state-of-the-art method, especially, the RNN-based method is able to achieve a median absolute error of 2 frames. The median of non-absolute errors (prediction minus ground truth) indicates the prediction bias of each method. The method of Condurache et al. makes late predictions, while the others have a minor bias. The table also lists the number of sequences with a small prediction error (3 frames, being about 0.2 seconds) and a large error ( $>10$  frames). The method of Condurache et al. has mis-detection on 7 sequences (the first entry in the last two columns in Table 5.1), which was also reported in [37]. While the two proposed methods both have 55 sequences with a small error ( $\leq 3$  frames) out of 80, the CNN-based approach has the smallest numbers of sequences with a large error ( $> 10$  frames) among the three methods.

The median error of the RNN-based method is similar to the results reported in [17]. While they achieved a mean error of less than one frame, their detection step requires the knowledge of complete sequences, hence it will not work in a prospective



**Figure 5.4:** An example to illustrate the RNN-based method. From left to right are the original XA frame (left), the vessel layer after layer separation (middle left), the vesselness image (middle right), the contrast signal for the whole sequence (right). The color markers in the signal show the prediction of BCF with LSTM (red) and the ground truth (green). Note that the artefact of diaphragm does not appear in the vesselness image thanks to layer separation.

Methods	Mean (std)	Median (*)	#(Error $\leq 3$ )	#(Error $> 10$ )
Condurache et al.[19]	6.2 (7.1)	5 (4)	29 / 73	10 / 73
CNN-based	3.9 (4.9)	2.5 (1)	<b>55</b> / 80	<b>5</b> / 80
RNN-based	<b>3.6</b> (4.6)	<b>2</b> (-0.5)	<b>55</b> / 80	7 / 80

**Table 5.1:** The statistics of the absolute error for the 3 methods. The two columns in the middle show the mean, standard deviation, median of the absolute errors and the median of non-absolute errors (\*) in frames. The last two columns show the number of sequences on which the method made an absolute error no larger than 3 frames or larger than 10 frames.

scenario. The learning-based method in [37] can be used for prospective detection, but some of the proposed features were heuristically designed for X-ray images of LA for EP procedure, which have different image features from the XA of coronary interventions. Compared to these methods, our approaches were designed for prospective settings and the CNN-based method is a general framework that could potentially be applied in different clinical procedures.

The RNN-based learning with a handcrafted feature has slightly lower mean and median error than the CNN-based method, although the latter has a more complex and deeper architecture. This might contradict to what is commonly known about the performance of deep learning. The possible reasons may be two-fold. First, the size of training data was small, even with data augmentation and a reduced CNN model, some over-fitting was observed. Second, the CNN treats frames independently rather than modeling their temporal relations. Although CNNs perform excellent in many classification tasks, detecting BCF requires a classifier that has good accuracy for data on the border between two classes.

In terms of computation efficiency (test time), the method of Condurache et al. needed 111 ms to 443 ms to process a frame. While the CNN-based method ran very fast and used on average only 14 ms to process one frame. The RNN-based method ran on average 64 ms/frame on images of the original size  $512 \times 512$  or  $1024 \times 1024$ , and 140 ms/frame on images of the original size  $776 \times 776$ . As the test time of the RNN-

based method was based on a MATLAB implementation with a single CPU core, it has large potential to run in real-time (<66 ms) with an optimized implementation running on a modern GPU.

In conclusion, we have developed two novel approaches for prospective detection of contrast inflow in XA sequences, a CNN-based and a RNN-based approach. The proposed methods perform well in BCF detection tasks in XA sequences, and outperform a previous state-of-the-art method. Both methods work in prospective settings and run fast, therefore have the potential to be integrated in advanced image guidance systems for PCI.

**Acknowledgement** This work was supported by Technology Foundation STW, IMAGIC project under the iMIT program (grant number 12703).



## Summary and General Discussion

---

### 6.1 Summary

In this thesis we proposed methods to improve image guidance during TACE procedures. The goal was to provide better guidance to the physician which may lead to shorter procedure times, improved selection of feeder vessels, less contrast use and less radiation exposure for both the patient and physician. We proposed tools and investigated the possibility to track and continuously align the catheter and guidewire inside the liver vasculature. Such tracking offers improved image guidance to the physician with dynamic 2D overlays and 3D roadmaps.

In the second chapter, in the context of liver catheterization, we presented a method to register intra-operative single-plane 2D fluoroscopic images with peri-operative 3DRA images. Catheter and guidewire centerline were manually segmented from the 2D X-ray image and the 3D blood vessel tree centerline was extracted from 3DRA. The method is based on the alignment of the catheter with the vessel tree and is divided in two parts. The first part uses a shape similarity metric to sort the projected 3D vessel centerlines of the vessel tree according to their similarity with the catheter shape. In the second part, the most similar vessel centerlines are registered with the catheter centerline and the best registration with regards to its metric is kept. The registration metric is based on the closest corresponding points distance between the catheter and the projected vessel paths. To align the vessel paths with the catheter, the 3D vessel tree is translated and rotated in order to minimize the registration metric using two different optimizers: brute force and Powell. The evaluation on clinical data with 2D angiographic images showed a median of the closest corresponding points distance between 2D angiography vessels and projected 3D vessels of 5.4 mm when using the brute force optimizer and 5.2 mm when using the Powell optimizer. Experiments on simulated data and visualization of results on real data showed a small percentage of incorrectly registered tip positions. Despite these few cases of vessel misalignment, the registered overlay was generally sufficient for the physician to localize the tip and identify the subsequent bifurcations. The registration was failing when large deformations occurred and when there was a lack of information such as imaging data in which only a small part of the catheter is visible in the fluoroscopic images and in cases in which the aorta and hepatic artery are not visible in the 3DRA.

To avoid misregistrations in these cases, the previous registrations and catheter positions need to be taken into account as prior knowledge during the registration.

In the third chapter, we used this idea and propose to continuously track the tip of the catheter inside the 3D vessel tree and keep a probability map of the tip position over time. To this end, we utilize a hidden Markov model (HMM) where states of the model are the possible positions of the catheter tip inside the 3D vessel tree. The transitions from state to state model the probabilities for the catheter tip to move from one position to another. The HMM is updated following the observation scores, based on the registration between the 2D catheter centerline extracted from the 2D X-ray image, and the 2D projection of 3D vessel tree centerline extracted from the 3DRA. We evaluated the method with the same protocol as in the previous chapter and obtained a median of the closest corresponding points distance between 2D angiography vessels and projected 3D vessels of 4.7mm. The distribution of distance errors is significantly smaller compared to the previous method. Experiments on complete X-ray sequences showed that the method is robust even when only a small part of the catheter is visible in the image.

The methods in chapter two and three relied on manual extraction of catheter and guidewire centerline in 2D X-ray images. In the fourth chapter, we proposed a method to segment the catheter and guidewire in a fully-automatic and real-time manner. The X-ray image to segment and the three previous ones of the sequence are fed into a convolutional neural network (CNN) trained with manually segmented catheter images. From the segmented catheter mask image output, the catheter centerline is extracted using skeletonization with subsequently pixel connectivity and branch linking. Evaluation on liver catheterization data demonstrated a median tip distance error of 0.9 mm and a median centerline distance error of 0.2 mm. The catheter segmentation was robust but may fail when parts of the guidewire hardly have any contrast. In such cases, the tip part may be missing in some of the extracted centerline. As the registration uses the complete shape of the centerline, a missing segment does not hamper the dynamic roadmapping, but it may result in inaccurate 3D tracking of the tip.

Automatic contrast-agent detection in X-ray sequences enables to choose the appropriate segmentation and registration method to use without requiring input from the physician. If vessels are present in the fluoroscopic images, vessel-based registration may be used [30, 43] instead of our catheter-based registration described in the previous chapters. The last chapter describes two approaches to automatically detect contrast agent injection in X-ray sequences. The methods are based on supervised learning and neural networks. The first method uses a CNN to classify frames as contrasted or not contrasted. The input of the neural network contains the current frame, the three previous frames and the first frame of the sequence which does not have contrast. The second approach is composed of two steps. The first step separates background, breathing and vessel layer using online robust PCA on the current and previous frames of the X-ray sequence. The vessel layer is subsequently enhanced with a multi-scale vesselness filter. In the second step, the vessel layer is converted in a 1D signal corresponding to the average intensity of the image. This signal is recurrently fed into a recurrent neural network (RNN) to learn and detect contrast frames. The method was applied to detect the first frame with contrast on X-ray sequences derived from percutaneous coronary interventions. In a prospective scenario, the two methods showed similar results with a mean absolute error of 3.6-3.9 frames. These numbers

are significantly better than the state of the art for prospective cases. The RNN method showed slightly better results but the CNN method is much faster. Although the experiment was not performed on abdominal data, coronary data is similar and the proposed methods are general enough to be translated to other interventions such as TACE.

## 6.2 Challenges towards use in clinical practice and future studies

In chapter two to five, methods have been presented and evaluated for navigation guidance. A next step would be to integrate these methods in a solution and evaluate the robustness and usefulness of the tracking and roadmapping during complete TACE procedures. This necessary step is beyond the scope of this thesis. However, we have suggestions for potential future studies and how to solve practical issues. A challenge that needs to be addressed is the fact that combining methods leads to combining errors of each method. If a method is not robust, it will impact the following step of the complete method so we have to insure that each step is sufficiently robust.

The clinical imaging data and instruments used in TACE procedures are very heterogeneous between hospitals and physicians performing the intervention. For example, our catheter and guidewire extraction method has been developed to detect instruments fully radio-opaque. In some hospitals, during TACE procedure, the guidewires that are used have only the tip radio-opaque. With these instruments, the guidewire extraction is failing and only the visible catheter part is extracted. A practical solution would be to use only fully radio-opaque catheter and guidewire during intervention. Another example of heterogeneity between procedures is related with the 3DRA acquisition. Registration is based on the assumption that the visible part of the catheter on the X-ray image is inside the vessel tree of the 3DRA. However, sometimes the aorta and hepatic artery part are not in the 3DRA field of view, as these are not always of clinical interest for a common TACE procedure. Medical imaging and image guidance in particular may benefit from standardization and protocols in interventions.

The two issues could be solved with changes to the intervention protocol. However other approaches dealing with these challenges could also be studied. 3DRA images could be augmented to add the missing artery and aorta using registration with 3D pre-operative CTA. This combination may add information on the tumor position as well. Also, it would be beneficial to improve the catheter and guidewire extraction method. Due to the noise of low radiation dose fluoroscopic images, parts of the guidewire are sometime not visible even with radio-opaque material. Humans have no problem to fill the gaps and segment the instrument because they are able to keep track of the object and also easily predict the next position and shape of the structure. The flexible structure of the catheter is intrinsically used as a prior knowledge. Such knowledge may be incorporated in the segmentation method in order to increase its robustness. Most of the catheter segmentation papers use a spline model [6, 15, 16, 34, 38, 86, 98]. This may not work properly as a 2D projection of a 3D thin instrument can have almost any arbitrary shapes. Integration of a 3D

shape model, based on the extracted vasculature, may be an option in this case. A different approach with merging of segments has been studied [94, 97] but it cannot deal with a long none-visible part of the catheter in the image. Our fourth chapter enables full-automatic and real-time computation using CNNs. However, unlike the previously cited methods, no prior knowledge is used in this method. Combining CNNs, catheter knowledge and temporal consistency would be one way to improve the segmentation. Traditional model optimization problems always need fine-tuning for most of the parameters. These may be learnt by a neural network. So, a 3D spline model or a model that fits the projective nature of the 3D catheter shapes could potentially be directly integrated into the neural network. Prior knowledge of the catheter position inside the 3D vessel tree may reduce the optimization space. As an example, a spline model could be incorporated in the neural network using a spline matrix representation [15, 74]. Other paradigms not using images at all could be investigated such as shape sensing [27, 85].

In order to improve the alignment metric between the 2D catheter and the 3D vasculature for the registration and tracking method, a non-rigid approach could be investigated. However, a more pragmatic idea could be explored. The C-arm detector position does not always provide the optimal views of the patient anatomy to easily perform the procedure. Following the complexity of the vasculature, the traditional coronal and sagittal planes may not be the best views to remove ambiguity between branching, vessels overlapping and foreshortening problems, and to optimally project vessels for catheter navigation. These problems make the guidance difficult for the physician but also for the registration method and the catheter extraction. Manually choosing a better view with the C-arm is not easy due to the complexity of the 3D vasculature. With the 3D vessel tree extracted from the 3DRA, an optimal C-arm position could be automatically proposed using various criteria. In the vessel tree, we should make sure that the vessel path between the hepatic artery and the tumors has a distinct shape compared to its neighborhood paths in the imaging projection space. It should also have the smallest amount of foreshortening. These different metrics could provide a final score for every different primary and secondary angle of the C-arm. Following the advancement of the catheter in the vasculature, the software could propose another C-arm position. Such solution could help the physician and the registration at the same time.

To conclude, we presented in this thesis, novel methods to improve image guidance during TACE interventions. We demonstrated and evaluated catheter tracking and roadmapping using non-contrasted 2D X-ray images and a 3D vessel tree extracted from peri-operative images. These developed methods could help physicians to be more accurate, efficient and confident, which could in turn lead to faster interventions with less contrast agent use and less radiation exposure for both patients and the physicians. Ultimately, this would lead to better patient outcome. To be integrated in clinical routine, physicians should follow specific protocols in order to ensure that the methods can be used in an optimal way.

---

# Bibliography

---

- [1] Ambrosini, P., Ruijters, D., Moelker, A., Niessen, W. J., and van Walsum, T. (2014). 2d/3d catheter-based registration for image guidance in tace of liver tumors. In *Information Processing in Computer-Assisted Interventions*, pages 246–255. Springer.
- [2] Ambrosini, P., Ruijters, D., Niessen, W. J., Moelker, A., and van Walsum, T. (2015a). Continuous roadmapping in liver tace procedures using 2d–3d catheter-based registration. *Int J Comput Assist Radiol Surg*, 10(9):1357–1370.
- [3] Ambrosini, P., Smal, I., Ruijters, D., Niessen, W. J., Moelker, A., and van Walsum, T. (2015b). 3d catheter tip tracking in 2d x-ray image sequences using a hidden markov model and 3d rotational angiography. In *Augment Environ Comput Assist Interv*, pages 38–49. Springer.
- [4] Ambrosini, P., Smal, I., Ruijters, D., Niessen, W. J., Moelker, A., and van Walsum, T. (2017). A hidden markov model for 3d catheter tip tracking with 2d x-ray catheterization sequence and 3d rotational angiography. *IEEE Trans. Med. Imag.*, 36(3):757–768.
- [5] Atasoy, S., Groher, M., Zikic, D., Glocker, B., Waggershauer, T., Pfister, M., and Navab, N. (2008). Real-time respiratory motion tracking: roadmap correction for hepatic artery catheterizations. In *Medical imaging*, pages 691815–691815. International Society for Optics and Photonics.
- [6] Baert, S. A., Viergever, M. A., and Niessen, W. J. (2003). Guide-wire tracking during endovascular interventions. *IEEE Trans. Med. Imag.*, 22(8):965–972.
- [7] Baka, N., Lelieveldt, B., Schultz, C., Niessen, W., and van Walsum, T. (2015). Respiratory motion estimation in x-ray angiography for improved guidance during coronary interventions. *Physics in Medicine & Biology*, 60(9):3617.
- [8] Baka, N., Metz, C., Schultz, C., Neefjes, L., van Geuns, R. J., Lelieveldt, B. P., Niessen, W. J., van Walsum, T., and de Bruijne, M. (2013). Statistical coronary motion models for 2d+ t/3d registration of x-ray coronary angiography and cta. *Medical image analysis*, 17(6):698–709.
- [9] Baka, N., Metz, C., Schultz, C., van Geuns, R., Niessen, W., and van Walsum, T. (2014). Oriented Gaussian mixture models for non-rigid 2d/3d coronary artery registration. *IEEE Transactions on Medical Imaging*, 33(5):1023–1034.
- [10] Bapst, B., Lagadec, M., Breguet, R., Vilgrain, V., and Ronot, M. (2015). Cone Beam Computed Tomography (CBCT) in the field of interventional oncology of the liver. *Cardiovasc. Intervent. Radiol.*, 39(1):8–20.
- [11] Bargellini, I., Turini, F., Bozzi, E., Lauretti, D., Cicorelli, A., Lunardi, A., Cioni, R., and Bartolozzi, C. (2013). Image fusion of preprocedural cta with real-time fluoroscopy to guide proper hepatic artery catheterization during transarterial chemoembolization of hepatocellular carcinoma: a feasibility study. *Cardiovascular and interventional radiology*, 36(2):526–530.
- [12] Baur, C., Albarqouni, S., Demirci, S., Navab, N., and Fallavollita, P. (2016). Cathnets: Detection and single-view depth prediction of catheter electrodes. In *International Conference on Medical Imaging and Virtual Reality*, pages 38–49. Springer.

- [13] Bender, F., Groher, M., Khamene, A., Wein, W., Heibel, T. H., and Navab, N. (2008). 3d dynamic roadmapping for abdominal catheterizations. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 668–675. Springer.
- [14] Breedis, C. and Young, G. (1954). The blood supply of neoplasms in the liver. *The American journal of pathology*, 30(5):969.
- [15] Chang, P.-L., Rolls, A., De Praetere, H., Vander Poorten, E., Riga, C. V., Bicknell, C. D., and Stoyanov, D. (2016). Robust catheter and guidewire tracking using b-spline tube model and pixel-wise posteriors. *IEEE Robot. Autom. Lett.*, 1(1):303–308.
- [16] Chen, B.-J., Wu, Z., Sun, S., Zhang, D., and Chen, T. (2016). Guidewire tracking using a novel sequential segment optimization method in interventional x-ray videos. In *Proc IEEE Int Symp Biomed Imaging*, pages 103–106.
- [17] Chen, T., Funka-Lea, G., and Comaniciu, D. (2011). Robust and fast contrast inflow detection for 2d x-ray fluoroscopy. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 243–250. Springer.
- [18] Cleary, K. and Peters, T. M. (2010). Image-guided interventions: technology review and clinical applications. *Annual review of biomedical engineering*, 12:119–142.
- [19] Condurache, A., Aach, T., Eck, K., and Bredno, J. (2004). Fast detection and processing of arbitrary contrast agent injections in coronary angiography and fluoroscopy. In *Bildverarbeitung fuer die Medizin 2004*, pages 5–9. Springer.
- [20] Epstein, A. J., Groeneveld, P. W., Harhay, M. O., Yang, F., and Polsky, D. (2013). Impact of minimally invasive surgery on medical spending and employee absenteeism. *JAMA surgery*, 148(7):641–647.
- [21] European Association For The Study Of The Liver and others (2012). Easl–eortc clinical practice guidelines: management of hepatocellular carcinoma. *Journal of hepatology*, 56(4):908–943.
- [22] Feng, J., Xu, H., and Yan, S. (2013). Online robust pca via stochastic optimization. In *Advances in Neural Information Processing Systems*, pages 404–412.
- [23] Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D. M., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in globocan 2012. *International journal of cancer*, 136(5).
- [24] Forner, A., Llovet, J. M., and Bruix, J. (2014). Hepatocellular carcinoma. *The Lancet*, 379(9822):1245–1255.
- [25] Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 130–137. Springer.
- [26] Franz, A. M., Haidegger, T., Birkfellner, W., Cleary, K., Peters, T. M., and Maier-Hein, L. (2014). Electromagnetic tracking in medicine: a review of technology, validation, and applications. *IEEE transactions on medical imaging*, 33(8):1702–1725.
- [27] Fuerst, B., Sutton, E. E., Ghotbi, R., Cowan, N. J., and Navab, N. (2016). Bioelectric navigation: a new paradigm for intravascular device guidance. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 474–481. Springer.
- [28] Gandaglia, G., Ghani, K. R., Sood, A., Meyers, J. R., Sammon, J. D., Schmid, M., Varda, B., Briganti, A., Montorsi, F., Sun, M., et al. (2014). Effect of minimally invasive surgery on the risk for surgical site infections: results from the national surgical quality improvement program (nsqip) database. *JAMA surgery*, 149(10):1039–1044.

- [29] Greenspan, H., van Ginneken, B., and Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159.
- [30] Groher, M., Zikic, D., and Navab, N. (2009). Deformable 2d-3d registration of vascular structures in a one view scenario. *Medical Imaging, IEEE Transactions on*, 28(6):847–860.
- [31] Guan, Y.-S., He, Q., and Wang, M.-Q. (2012). Transcatheter arterial chemoembolization: history for more than 30 years. *ISRN gastroenterology*, 2012.
- [32] Hartung, M. P., Grist, T. M., and François, C. J. (2011). Magnetic resonance angiography: current status and future directions. *Journal of Cardiovascular Magnetic Resonance*, 13(1):19.
- [33] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*, pages 770–778.
- [34] Heibel, H., Glocker, B., Groher, M., Pfister, M., and Navab, N. (2013). Interventional tool tracking using discrete optimization. *Medical Imaging, IEEE Transactions on*, 32(3):544–555.
- [35] Hipwell, J. H., Penney, G. P., McLaughlin, R. A., Rhode, K., Summers, P., Cox, T. C., Byrne, J. V., Noble, J. A., and Hawkes, D. J. (2003). Intensity-based 2-d-3-d registration of cerebral angiograms. *Medical Imaging, IEEE Transactions on*, 22(11):1417–1426.
- [36] Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8):1735–1780.
- [37] Hoffmann, M., Müller, S., Kurzidim, K., Strobel, N., and Hornegger, J. (2015). Robust identification of contrasted frames in fluoroscopic images. In *Bildverarbeitung für die Medizin 2015*, pages 23–28. Springer.
- [38] Honnorat, N., Vaillant, R., and Paragios, N. (2011). Graph-based geometric-iconic guide-wire tracking. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 9–16. Springer.
- [39] Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *Proc ICML*, volume 37, pages 448–456.
- [40] Iwazawa, J., Ohue, S., Hashimoto, N., and Mitani, T. (2013). Comparison of the number of image acquisitions and procedural time required for transarterial chemoembolization of hepatocellular carcinoma with and without tumor-feeder detection software. *Radiology research and practice*, 2013.
- [41] Jolly, S. S., Amlani, S., Hamon, M., Yusuf, S., and Mehta, S. R. (2009). Radial versus femoral access for coronary angiography or intervention and the impact on major bleeding and ischemic events: a systematic review and meta-analysis of randomized trials. *American heart journal*, 157(1):132–140.
- [42] Jomier, J., Bullitt, E., Van Horn, M., Pathak, C., and Aylward, S. R. (2006). 3d/2d model-to-image registration applied to tips surgery. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2006*, pages 662–669. Springer.
- [43] Kim, J., Lee, J., Chung, J. W., and Shin, Y.-G. (2016). Locally adaptive 2d-3d registration using vascular structure model for liver catheterization. *Computers in biology and medicine*, 70:119–130.
- [44] Kim, J. H., Shim, J. H., Lee, H. C., Sung, K.-B., Ko, H.-K., Ko, G.-y., Gwon, D. I., Kim, J. W., Lim, Y.-S., and Park, S. H. (2017). New intermediate-stage subclassification for patients with hepatocellular carcinoma treated with transarterial chemoembolization. *Liver International*.
- [45] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553):436.

- [46] Liao, R., Miao, S., and Zheng, Y. (2013a). Automatic and efficient contrast-based 2-d/3-d fusion for trans-catheter aortic valve implantation (tavi). *Computerized Medical Imaging and Graphics*, 37(2):150–161.
- [47] Liao, R., Tan, Y., Sundar, H., Pfister, M., and Kamen, A. (2010). An efficient graph-based deformable 2d/3d registration algorithm with applications for abdominal aortic aneurysm interventions. In *Medical Imaging and Augmented Reality*, pages 561–570. Springer.
- [48] Liao, R., You, W., Liu, Y., Yan, M., John, M., and Shea, S. (2013b). Integrated spatiotemporal analysis for automatic contrast agent inflow detection on angiography and fluoroscopy during transcatheter aortic valve implantation. *Medical physics*, 40(4).
- [49] Liao, R., Zhang, L., Sun, Y., Miao, S., and Ched'Hotel, C. (2013c). A review of recent advances in registration techniques applied to minimally invasive therapy. *IEEE transactions on multimedia*, 15(5):983–1000.
- [50] Liapi, E., Hong, K., Georgiades, C. S., and Geschwind, J.-F. H. (2005). Three-dimensional rotational angiography: introduction of an adjunctive tool for successful transarterial chemoembolization. *Journal of vascular and interventional radiology*, 16(9):1241–1245.
- [51] Lindseth, F., Langø, T., Selbekk, T., Hansen, R., Reinertsen, I., Askeland, C., Solheim, O., Unsgård, G., Mårvik, R., and Hernes, T. A. N. (2013). Ultrasound-based guidance and therapy. In *Advancements and breakthroughs in ultrasound imaging*. InTech.
- [52] Llovet, J. M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J.-F., de Oliveira, A. C., Santoro, A., Raoul, J.-L., Forner, A., et al. (2008). Sorafenib in advanced hepatocellular carcinoma. *New England journal of medicine*, 359(4):378–390.
- [53] Lohrke, J., Frenzel, T., Endrikat, J., Alves, F. C., Grist, T. M., Law, M., Lee, J. M., Leiner, T., Li, K.-C., Nikolaou, K., et al. (2016). 25 years of contrast-enhanced mri: developments, current challenges and future perspectives. *Advances in therapy*, 33(1):1–28.
- [54] Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proc IEEE Comput Soc Conf CVPR*, pages 3431–3440.
- [55] Luan, K., Ohya, T., Liao, H., Kobayashi, E., and Sakuma, I. (2013). Vessel bifurcation localization based on intraoperative three-dimensional ultrasound and catheter path for image-guided catheter intervention of oral cancers. *Computerized Medical Imaging and Graphics*, 37(2):113–122.
- [56] Ma, H., Dibildox, G., Banerjee, J., Niessen, W., Schultz, C., Regar, E., and van Walsum, T. (2015). Layer separation for vessel enhancement in interventional x-ray angiograms using morphological filtering and robust pca. In *Workshop on Augmented Environments for Computer-Assisted Interventions*, pages 104–113. Springer.
- [57] Ma, H., Hoogendoorn, A., Regar, E., Niessen, W. J., and van Walsum, T. (2017). Automatic online layer separation for vessel enhancement in x-ray angiograms for percutaneous coronary interventions. *Medical image analysis*, 39:145–161.
- [58] Ma, Y., King, A. P., Gogin, N., Gijsbers, G., Rinaldi, C., Gill, J., Razavi, R., and Rhode, K. S. (2012). Clinical evaluation of respiratory motion compensation for anatomical roadmap guided cardiac electrophysiology procedures. *Biomedical Engineering, IEEE Transactions on*, 59(1):122–131.
- [59] Maintz, J. A. and Viergever, M. A. (1998). A survey of medical image registration. *Medical image analysis*, 2(1):1–36.
- [60] Markelj, P., Tomaževič, D., Likar, B., and Pernuš, F. (2012). A review of 3d/2d registration methods for image-guided interventions. *Medical image analysis*, 16(3):642–661.



- [61] Metz, C., Schaap, M., Klein, S., Rijnbeek, P., Neeffes, L., Mollet, N., Schultz, C., Serruys, P., Niessen, W., and van Walsum, T. (2012). Alignment of 4d coronary cta with monoplane x-ray angiography. In *Augmented Environments for Computer-Assisted Interventions*, pages 106–116. Springer.
- [62] Milletari, F., Navab, N., and Ahmadi, S.-A. (2016). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *Proc Int Conf 3D Vis*, pages 565–571. IEEE.
- [63] Minami, Y., Yagyu, Y., Murakami, T., and Kudo, M. (2014). Tracking navigation imaging of transcatheter arterial chemoembolization for hepatocellular carcinoma using three-dimensional cone-beam ct angiography. *Liver Cancer*, 3(1):53–61.
- [64] Mitrović, U., Spiclin, Z., Likar, B., and Pernuš, F. (2013). 3d-2d registration of cerebral angiograms: a method and evaluation on clinical images. *Medical Imaging, IEEE Transactions on*, 32(8):1550–1563.
- [65] Miyayama, S., Yamashiro, M., Ikuno, M., Okumura, K., and Yoshida, M. (2014). Ultraslective transcatheter arterial chemoembolization for small hepatocellular carcinoma guided by automated tumor-feeders detection software: technical success and short-term tumor response. *Abdominal imaging*, 39(3):645–656.
- [66] Miyayama, S., Yamashiro, M., Okuda, M., Yoshie, Y., Sugimori, N., Igarashi, S., Nakashima, Y., and Matsui, O. (2009). Usefulness of cone-beam computed tomography during ultraslective transcatheter arterial chemoembolization for small hepatocellular carcinomas that cannot be demonstrated on angiography. *Cardiovascular and interventional radiology*, 32(2):255–264.
- [67] Oliveira, F. P. and Tavares, J. M. R. (2014). Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93.
- [68] Pauly, O., Heibel, H., and Navab, N. (2010). A machine learning approach for deformable guide-wire tracking in fluoroscopic sequences. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 343–350. Springer.
- [69] Pawiro, S. A., Markelj, P., Pernuš, F., Gendrin, C., Figl, M., Weber, C., Kainberger, F., Nöbauer-Huhmann, I., Bergmeister, H., Stock, M., Georg, D., Bergmann, H., and Birkfellner, W. (2011). Validation for 2d/3d registration i: A new gold standard data set. *Medical Physics*, 38(3):1481–1490.
- [70] Petkovic, T. and Loncaric, S. (2010a). Guidewire tracking with projected thickness estimation. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 1253–1256. IEEE.
- [71] Petkovic, T. and Loncaric, S. (2010b). Using x-ray imaging model to improve guidewire detection. In *Signal Processing (ICSP), 2010 IEEE 10th International Conference on*, pages 805–808. IEEE.
- [72] Pham, D. L., Xu, C., and Prince, J. L. (2000). Current methods in medical image segmentation. *Annual review of biomedical engineering*, 2(1):315–337.
- [73] Powell, M. J. (1964). An efficient method for finding the minimum of a function of several variables without calculating derivatives. *The computer journal*, 7(2):155–162.
- [74] Qin, K. (2000). General matrix representations for b-splines. *The Visual Computer*, 16(3-4):177–186.
- [75] Rabiner, L. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286.
- [76] Rammohan, A., Sathyanesan, J., Ramaswami, S., Lakshmanan, A., Senthil-Kumar, P., Srinivasan, U. P., Ramasamy, R., and Ravichandran, P. (2012). Embolization of liver tumors: Past, present and future. *World journal of radiology*, 4(9):405.

- [77] Rivest-Henault, D., Sundar, H., and Cheriet, M. (2012). Nonrigid 2d/3d registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. *Medical Imaging, IEEE Transactions on*, 31(8):1557–1572.
- [78] Rodriguez, P. and Wohlberg, B. (2013). Fast principal component pursuit via alternating minimization. In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pages 69–73. IEEE.
- [79] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, pages 234–241.
- [80] Ruijters, D., Homan, R., Mielekamp, P., Van de Haar, P., and Babic, D. (2011). Validation of 3d multimodality roadmapping in interventional neuroradiology. *Physics in medicine and biology*, 56(16):5335.
- [81] Ruijters, D., ter Haar Romeny, B. M., and Suetens, P. (2009). Vesselness-based 2d-3d registration of the coronary arteries. *International journal of computer assisted radiology and surgery*, 4(4):391–397.
- [82] Sangro, B., Iñarrairaegui, M., and Bilbao, J. I. (2012). Radioembolization for hepatocellular carcinoma. *Journal of hepatology*, 56(2):464–473.
- [83] Selle, D., Preim, B., Schenk, A., and Peitgen, H.-O. (2002). Analysis of vasculature for liver surgical planning. *Medical Imaging, IEEE Transactions on*, 21(11):1344–1357.
- [84] Sherman, M. (2010). Hepatocellular carcinoma: epidemiology, surveillance, and diagnosis. In *Seminars in liver disease*, volume 30, page 3.
- [85] Shi, C., Luo, X., Qi, P., Li, T., Song, S., Najdovski, Z., Fukuda, T., and Ren, H. (2017). Shape sensing techniques for continuum robots in minimally invasive surgery: A survey. *IEEE Transactions on Biomedical Engineering*, 64(8):1665–1678.
- [86] Slabaugh, G., Kong, K., Unal, G., and Fang, T. (2007). Variational guidewire tracking using phase congruency. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2007)*, pages 612–619. Springer.
- [87] Sotiras, A., Davatzikos, C., and Paragios, N. (2013). Deformable medical image registration: A survey. *IEEE transactions on medical imaging*, 32(7):1153–1190.
- [88] Springenberg, J. T., Dosovitskiy, A., Brox, T., and Riedmiller, M. (2014). Striving for simplicity: The all convolutional net. *arXiv preprint arXiv:1412.6806*.
- [89] Srivastava, N., Hinton, G. E., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res*, 15(1):1929–1958.
- [90] Stewart, B., Wild, C. P., et al. (2014). World cancer report 2014, chapter 5.6. *World Health Organization*.
- [91] Sundar, H., Khamene, A., Xu, C., Sauer, F., and Davatzikos, C. (2006). A novel 2d-3d registration algorithm for aligning fluoro images with 3d pre-op ct/mr images. In *Proc. SPIE*, volume 6141, pages 61412K–61412K–7. International Society for Optics and Photonics.
- [92] Van der Bom, I., Klein, S., Staring, M., Homan, R., Bartels, L., and Pluim, J. (2011). Evaluation of optimization methods for intensity-based 2d-3d registration in x-ray guided interventions. In *SPIE Medical Imaging*, volume 7962, pages 796223–796223–15. International Society for Optics and Photonics.
- [93] van der Cammen, M. (2017). 2d fluoroscopy and 3d computed tomography registration for minimally invasive liver procedures. Master’s thesis, Delft University of Technology, the Netherlands.

- [94] Vandini, A., Glocker, B., Hamady, M., and Yang, G.-Z. (2017). Robust guidewire tracking under large deformations combining segment-like features (seglets). *Medical image analysis*, 38:150–164.
- [95] Viergever, M. A., Maintz, J. A., Klein, S., Murphy, K., Staring, M., and Pluim, J. P. (2016). A survey of medical image registration—under review. *Medical image analysis*, 33:140–144.
- [96] Wagener, J. F. and Rao, S. V. (2015). A comparison of radial and femoral access for cardiac catheterization. *Trends in cardiovascular medicine*, 25(8):707–713.
- [97] Wagner, M. G., Strother, C. M., and Mistretta, C. A. (2016). Guidewire path tracking and segmentation in 2d fluoroscopic time series using device paths from previous frames. In *Proc. SPIE*, pages 97842B–97842B.
- [98] Wang, P., Chen, T., Zhu, Y., Zhang, W., Zhou, S., and Comaniciu, D. (2009). Robust guidewire tracking in fluoroscopy. In *Proc IEEE Computer Vision and Pattern Recognition (CVPR)*, pages 691–698.
- [99] Wáng, Y.-X. J., De Baere, T., Idée, J.-M., and Ballet, S. (2015). Transcatheter embolization therapy in liver cancer: an update of clinical evidences. *Chinese Journal of Cancer Research*, 27(2):96.
- [100] World Health Organization (2016). Global health estimates 2015: Deaths by cause, age, sex, by country and by region, 2000–2015. geneva: Who; 2016. *World Health Organization*.
- [101] Wu, X., Housden, J., Ma, Y., Razavi, B., Rhode, K., and Rueckert, D. (2015). Fast catheter segmentation from echocardiographic sequences based on segmentation from corresponding x-ray fluoroscopy for cardiac catheterization interventions. *IEEE Trans. Med. Imag.*, 34(4):861–876.
- [102] Yamada, R., Sato, M., Kawabata, M., Nakatsuka, H., Nakamura, K., and Takashima, S. (1983). Hepatic artery embolization in 120 patients with unresectable hepatoma. *Radiology*, 148(2):397–401.
- [103] Yao, X., Yan, D., Jiang, X., Li, X., Zeng, H., Liu, D., and Li, H. (2018). Dual-phase cone-beam ct-based navigation imaging significantly enhances tumor detectability and aids superselective transarterial chemoembolization of liver cancer. *Academic radiology*.
- [104] Yatziv, L., Chartouni, M., Datta, S., and Sapiro, G. (2012). Toward multiple catheters detection in fluoroscopic image guided interventions. *Information Technology in Biomedicine, IEEE Transactions on*, 16(4):770–781.
- [105] You, W., Liao, R., Yan, M., and John, M. (2011). Spatio-temporal analysis for automatic contrast injection detection on angiography during trans-catheter aortic valve implantation. In *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*, pages 702–706. IEEE.
- [106] Zhang, T. and Suen, C. Y. (1984). A fast parallel algorithm for thinning digital patterns. *Commun ACM*, 27(3):236–239.
- [107] Zhao, X., Miao, S., Du, L., and Liao, R. (2013). Robust 2-d/3-d registration of ct volumes with contrast-enhanced x-ray sequences in electrophysiology based on a weighted similarity measure and sequential subspace optimization. In *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*, pages 934–938. IEEE.



---

# Samenvatting

---

## Samenvatting

In dit proefschrift hebben we methoden voor verbetering van de beeldgeleiding bij TACE procedures voorgesteld. Het doel was betere beeldgeleiding voor de arts te bewerkstelligen, wat kan leiden tot kortere proceduredtijden, betere selectie van de vaten die de tumor voeden, gebruik van minder contrastmiddel en minder blootstelling aan ioniserende straling voor de arts en de patiënt. We hebben de mogelijkheden onderzocht om de katheter en voerdraad te volgen in de slagaders van de lever. Dit soort technieken verbeteren de beeldgeleiding door gebruik te maken van dynamische 2D en 3D roadmaps.

In het tweede hoofdstuk hebben we een methode gepresenteerd om intra-operative 2D doorlichtingsbeelden (ook wel fluoroscopie genoemd) te registreren met 3D rotatie angiografie (RA) beelden die tijdens de interventie verkregen zijn. De katheter en de voerdraad werden handmatig gesegmenteerd in het 2D angiografie beeld en de 3D vaatboom werd uit de 3DRA beelden gehaald. De methode is gebaseerd op het over elkaar leggen van de katheter en de vaatboom, en wordt gedaan in twee stappen. In de eerste stap worden, op basis van vorm, delen van de (geprojecteerde) 3D vaatboom gesorteerd. In de tweede stap worden de delen van de vaatboom die qua vorm het meest lijken op de katheter geregistreerd met de katheter, en de beste registratie wordt gekozen. De maat voor de registratie is gebaseerd op de afstand tussen dichtstbijzijnde corresponderende punten tussen de geprojecteerde vaatboom en de katheter. Tijdens deze registratie wordt de 3D vaatboom zodanig verschoven en gedraaid, dat de registratiemaat minimaal is. Hiervoor wordt geëxperimenteerd met twee verschillende optimalisatie-methoden: brute force en Powell. De evaluatie met klinische beelddata laat een mediane afstand zien (tussen de dichtstbijzijnde corresponderende punten van de katheter en de geprojecteerde vaatboom) van 5.4 mm voor de brute kracht optimalisatie, en 5.2 mm voor de Powell optimalisatie. Experimenten met gesimuleerde gegevens en de visualisatie van de resultaten met echte beelddata laten zien dat in een klein percentage van de gevallen de tip van de katheter niet juist gepositioneerd is in het vat. Ondanks deze gevallen waar de registratie niet goed is, was de overprojectie van de vaatboom in het algemeen voldoende om de arts te laten weten waar de katheter zich in de vaatboom bevindt. De registratie ging vooral fout in gevallen van grote vervorming van de vaatboom, of als er slechts een klein deel van de katheter in beeld was, en ook als de aorta of leverslagader niet zichtbaar waren in de 3DRA beelden.

Misregistraties als hierboven beschreven zouden voorkomen kunnen worden als eerdere registraties en katheter posities meegenomen worden als voorkennis in de registratie. In het derde hoofdstuk stellen we, op basis van dit idee, voor om de tip van de katheter continu te volgen in de 3D vaatboom, en een waarschijnlijkheidsverdeling bij te houden die aangeeft waar de tip zich bevindt. Hiervoor gebruiken we een hidden Markov model (HMM), waarbij de toestanden de mogelijke posities van de tip van de katheter in de 3D vaatboom zijn. De overgangen tussen verschillende toestanden modelleren dan de beweging van de tip van de katheter. Het HMM wordt telkens aangepast op basis van waarnemingen, die in ons geval bestaan uit de registratie van de 2D katheter vorm (gehaald uit het 2D doorlichtingsbeeld) en de 2D projectie van de 3D vaat boom (gehaald uit de 3DRA beelden). We evalueerden de methode op dezelfde manier als de methode beschreven in hoofdstuk 2, en verkregen een mediane afstand van 4.7 mm. De spreiding van de fouten (afstanden) is significant kleiner vergeleken met de methode uit hoofdstuk 2. Experimenten uitgevoerd met complete doorlichtingsbeeldseries laten zien dat de methode robuust is, zelfs in het geval slechts een klein deel van de katheter zichtbaar is.

De methoden uit hoofdstuk 2 en 3 gebruiken katheter en voerdraad posities die handmatig aangegeven zijn in de 2D doorlichtingsbeelden. In het vierde hoofdstuk beschrijven we een volledig automatische en real-time methode voor de segmentatie van de katheter en de voerdraad. Het te segmenteren doorlichtingsbeeld, en drie eerdere beelden uit dezelfde serie, zijn de invoer voor een neurale netwerk met convoluties (CNN) dat getraind is met handmatig gesegmenteerde voorbeelden. Uit het segmentatieresultaat van de CNN wordt de middenlijn van de katheter gehaald via een skelet-operatie, waarna vervolgens het resultaat wordt nabewerkt. Bij deze nabewerking wordt bekeken welke pixels verbonden zijn, en worden losse delen verbonden met elkaar. Evaluatie van de methode op beelden van lever katherisatie procedures laten een mediane tip afstand zien van 0.9 mm, en een mediane afstand van de middenlijn van 0.2 mm. De segmentatie was robuust, maar ging fout in gevallen wanneer nauwelijks beeldcontrast was tussen de katheter en de omgeving. In zulke gevallen ontbrak de tip soms. Omdat de registratie die hierop volgt de hele katheter vorm gebruikt, heeft het ontbreken van de tip geen nadelige gevolgen voor de projectie van de vaatboom, maar het kan leiden tot onnauwkeurigheid in het volgen van de 3D positie van de tip in de vaatboom.

Automatische detectie van contrastmiddel in doorlichtingsbeelden maakt het mogelijk automatisch te schakelen tussen verschillende segmentatie- en registratiemethoden zonder tussenkomst van de arts. Als bijvoorbeeld de bloedvaten zichtbaar zijn door de aanwezigheid van contrastmiddel, zou een bloedvaten-gebaseerde methode [30, 43] gebruikt kunnen worden in plaats van de katheter-gebaseerde methode uit de vorige hoofdstukken. Het laatste hoofdstuk beschrijft daarom twee methoden om automatisch het injecteren van contrastmiddel te detecteren in doorlichtingsbeeldseries. De methoden zijn gebaseerd op het leren met referentiewaarden, en neurale netwerken. De eerste methode gebruikt een CNN om voor elk doorlichtingsbeeld te bepalen of er contrast aanwezig is of niet. De invoer voor de CNN bevat het huidige beeld, de drie vorige beelden en het eerste beeld (zonder contrastmiddel) van de serie. De tweede methode bestaat uit twee stappen. In de eerste stap wordt het beeld gesplitst in drie verschillende beelden: de achtergrond, structuren die bewegen

met ademhaling, en de bloedvaten. Het vaatbeeld worden vervolgens gefilterd met een vaatfilters van verschillende afmetingen. In de tweede stap wordt dit gefilterde vaatbeeld omgezet in een eendimensionaal signaal wat correspondeert met de gemiddelde intensiteit van het beeld. Dit signaal is de invoer voor een recurrent neuraal netwerk (RNN) wat beelden met contrast detecteert. De methode is gebruikt om het eerste beeld met contrastmiddel te selecteren uit een serie doorlichtingsbeelden van interventies aan de kransslagaders. Bij prospectief gebruik laten de twee methode vergelijkbare resultaten zien, met een gemiddelde absolute fout van 3.6 - 3.9 beelden. Dit is significant beter dan de resultaten van bestaande methoden. De RNN methode is iets beter dan de CNN methode, maar laatstgenoemde is veel sneller. Hoewel de methode niet is geëvalueerd op beelden van de buik, verwachten we dat ze ook voor TACE beelden zullen werken, omdat de methoden generiek zijn, en ook omdat de doorlichtingsbeelden van de kransslagaders vergelijkbaar zijn met die van de buik.

## **Uitdagingen voor toepassing in de klinische praktijk en toekomstige studies**

In hoofdstuk twee tot en met vijf hebben we methoden gepresenteerd en geëvalueerd voor beeldgeleiding. Een volgende stap zou zijn om dit werk te integreren in een compleet systeem, en de robuustheid en bruikbaarheid van het volgen van de instrumenten en de overprojectie van de 3D vaatboom tijdens TACE interventies te onderzoeken. Deze noodzakelijke stap valt buiten dit proefschrift. We hebben wel suggesties voor mogelijke toekomstige studies en ideeën hoe praktische problemen opgelost kunnen worden. Een uitdaging die aangepakt moet worden is het feit dat een combinatie van methoden leidt tot een combinatie van fouten in deze methoden. Als een onderdeel niet robuust werkt, zal dat gevolgen hebben voor de volgende stappen. Dus elke onderdeel moet zo robuust mogelijk werken.

De klinische beelddata en de instrumenten die gebruikt worden bij TACE procedures verschillen tussen ziekenhuizen en artsen die de TACE procedure uitvoeren. Onze methode voor het vinden van de katheter en voerdraad in de doorlichtingsbeelden is ontwikkeld voor instrumenten die over de hele lengte duidelijk zichtbaar zijn in de beelden. In sommige ziekenhuizen worden TACE procedures uitgevoerd met voerdraden waarvan alleen de tip duidelijk zichtbaar is in de beelden. Met zulke instrumenten zal onze detectie minder goed werken. Een pragmatische oplossing is om alleen instrumenten die over de hele lengte zichtbaar zijn te gebruiken. Een ander voorbeeld van verschillen is de opname van het 3DRA beeld. De registratie is gebaseerd op de aanname dat het zichtbare deel van de katheter in het doorlichtingsbeeld in de vaatboom in het 3DRA beeld zit. Bij sommige 3DRA beelden vallen de aorta en de leverslagader buiten het beeld, omdat ze voor een normale TACE procedure niet van belang zijn. Beeldgeleide interventies zijn daarom in het bijzonder gebaat bij standaardisatie en protocollering.

De bovenstaande twee verschillen zouden opgelost kunnen worden met aanpassing van het protocol. Er zijn echter ook andere mogelijkheden die onderzocht kunnen worden. Zo zouden de aorta en de leverslagader in een 3DRA mogelijk aangevuld kunnen worden door een registratie met een 3D CTA (computed tomographic angiography).

Deze combinatie zou ook relevante informatie over de tumor positie kunnen inbrengen. Verder zou het de katheter en voerdraad extractie methoden kunnen verbeteren. Door de ruis in de lage dosis fluoroscopische beelden zijn delen van de voerdraad soms niet zichtbaar, zelfs als ze gemaakt zijn van materiaal wat zichtbaar behoort te zijn in deze Röntgenbeelden. Voor mensen is het geen probleem om de ontbrekende delen in te vullen, en ze kunnen ook eenvoudig de positie en vorm van het instrument voorspellen. De flexibele structuur van de katheter wordt dan gebruikt als voorkennis. Zulke kennis zou in de segmentatie methode geïntegreerd kunnen worden om de methode robuuster te maken. De meeste artikelen over segmentatie van katheters gebruiken een mathematisch model van een gladde kromme om de katheter te modeleren [6, 15, 16, 34, 38, 86, 98]. Dit zou voor ons geval niet goed kunnen werken, omdat de 2D projectie van een dun instrument in 3D bijna elke willekeurige vorm kan aannemen. Het integreren van een 3D vormmodel, gebaseerd op de vaatboom die uit de beeld-data gehaald is, zou een mogelijkheid zijn in dit geval. Een andere aanpak van het combineren van segmenten is elders bestudeerd [94, 97], maar deze methode kan niet overweg met een katheter met een lang niet-zichtbaar deel in het beeld. Ons vierde hoofdstuk maakt volledig-automatische en real-time berekening met CNNs mogelijk. In tegenstelling tot de eerder genoemde methoden wordt geen gebruik gemaakt van voorkennis. De combinatie van CNNs, kennis van de katheter en consistentie over de tijd zou een manier zijn om de segmentatie te verbeteren. Traditionele model optimalisaties hebben altijd een fijne afstelling van de parameters nodig. Deze zouden geleerd kunnen worden door een neurale netwerk. Dus, een 3D gladde kromme model, of een model dat past bij de projectie van de 3D katheter vormen zou mogelijk direct in het neurale netwerk geïntegreerd kunnen worden. Voorkennis van de positie van de katheter in de 3D vaatboom kan de optimalisatie zoekruimte verkleinen. Een gladde kromme model kan bijvoorbeeld in een neurale netwerk geïntegreerd worden door een matrix representatie van de kromme [15, 74]. Aanpakken die geen beelden gebruiken, maar rechtstreeks de vorm bepalen [27, 85], zouden ook onderzocht kunnen worden.

Om te zorgen dat de 2D projectie van de katheter en de 3D vaatboom beter passen tijdens de registratie, zouden ook niet-rigide aanpakken onderzocht kunnen worden. Een meer pragmatische aanpak zou ook onderzocht kunnen worden. De doorlichtingsbeelden die gemaakt worden, worden niet altijd genomen onder een hoek die optimaal is voor de procedure. Afhankelijk van de complexiteit van de vaatboom zouden de traditionele afbeeldingen van de voorkant of zijkant niet altijd de ambiguïteit die ontstaat door splitsingen van vaten en overprojectie kunnen oplossen, en dus niet de optimale beelden geven voor het navigeren van de katheter. Dit maakt de beeldgeleiding lastig, zowel voor de arts als voor onze registratie- en segmentatiemethode. Handmatig een betere projectie-richting bepalen is niet eenvoudig, gegeven de complexiteit van de 3D vaatboom. Met behulp van de 3D vaatboom uit de 3DRA beelden zou een optimale (volgens bepaalde criteria) projectierichting automatisch bepaald kunnen worden. Bijvoorbeeld een projectierichting zodanig dat het geprojecteerde pad van de leverslagader naar de tumor een duidelijk andere vorm heeft dan vaten in de buurt, en een projectierichting die zo veel mogelijk loodrecht op de vaten is. Met deze criteria zou elke projectierichting een score kunnen krijgen. Tijdens de procedure zou op deze manier zelfs een aanpassing van de projectierichting voorgesteld kunnen worden. Zon werkwijze zou zowel de arts als de automatische



registratie kunnen helpen.

Tenslotte, in dit proefschrift hebben we nieuwe methoden voor verbeterde beeldgeleiding in TACE procedures gepresenteerd. We hebben het volgen van de katheter en overprojectie gebruik makend van 2D doorlichtingsbeelden en een 3D vaatboom laten zien, en geëvalueerd. Deze methoden zouden de artsen kunnen helpen om nauwkeuriger, efficiënter en zelfverzekerder te zijn, wat kan leiden tot kortere interventietijden, met gebruik van minder contrast middel en minder blootstelling aan straling voor de patiënt en de arts. Uiteindelijk zou dit moeten leiden tot betere resultaten voor de patiënt. Om deze methoden te integreren in de klinische praktijk zouden artsen specifieke protocollen moeten volgen, zodanig dat de methoden optimaal ingezet kunnen worden.



---

# PhD Portfolio

---

## Research

- PhD period: 2012 - 2017
- Biomedical Imaging Group Rotterdam (BIGR), Department of Radiology & Nuclear Medicine and the Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
- Advanced School for Computing and Imaging (ASCI), TU Delft, Delft, The Netherlands.

## In-depth courses

- A22 Knowledge driven Image Segmentation, ASCI, LUMC, Leiden, The Netherlands, March 12-16, 2012.
- Presentation course, Erasmus MC, Rotterdam, The Netherlands, November-December 2012.
- Philips workshop on iSDK, Philips Healthcare, Best, The Netherlands, February 15, 2013.
- A1 Advanced Pattern Recognition, ASCI, TU Delft, Delft, The Netherlands, April 8-12, 2013.
- 2nd Biomedical Imaging Summer School: Modalities, Methodologies & Clinical Research, Ecole Centrale de Paris and Ecole des Ponts-ParisTech, Institut Henri Poincaré, Paris, France, July 8-12, 2013.
- 7th International Computer Vision Summer School (ICVSS): Computer Vision and Machine Learning, University of Catania, Le Castella, Italia, July 14-20, 2013.
- A8 Frontend Vision and Multiscale Image Analysis, ASCI, TU Eindhoven, Eindhoven, The Netherlands, November 11-29, 2013.
- Anatomy session, Erasmus MC, Rotterdam, The Netherlands, November 25, 2013.

- Roadshow Integrity in research (Wetenschappelijke integriteit), Erasmus MC, Rotterdam, The Netherlands, February 6, 2014.
- A24 A Programmer's Guide for Modern High-Performance Computing, ASCI, TU Delft, Delft, The Netherlands, December 9-12, 2014.
- Biomedical English Writing in Communication, Erasmus MC, Rotterdam, The Netherlands, February-April 2015.

### **International conferences (poster/oral presentation)**

- The 5th International Conference on Information Processing in Computer-Assisted Interventions (IPCAI), Fukuoka, Japan, June 28, 2014 (poster and oral presentation).
- The 10th MICCAI Workshop on Augmented Environments for Computer-Assisted Interventions (AE-CAI), Munich, Germany, October 9, 2015 (oral presentation).
- The 28th Conference of the International Society for Medical Innovation and Technology (SMIT), Delft, The Netherlands, October 6, 2016 (oral presentation).
- The 20th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), Quebec city, Quebec, Canada, September 12, 2017 (poster)

### **Miscellaneous**

- 7th International Computer Vision Summer School (ICVSS): Computer Vision and Machine Learning, Le Castella, Italia, July 15, 2013 (poster).
- The 10th Medical Imaging Symposium for PhD students (MISP), Amsterdam, The Netherlands, March 5, 2015 (oral presentation).
- Attendance at the Medical Imaging Symposium for PhD students (MISP), Leiden and Amsterdam, The Netherlands, 2014 and 2015.
- Attendance seminar on Where your PhD can take you; Career opportunities, Erasmus MC, The Netherlands, June 4, 2015.
- 3x BGR seminar presentations, 4x Medical Informatics Research Lunch presentations, >20x Image Guided Intervention meetings, 2012-2016
- Attendance: BGR seminar, Radiology Research Lunch, Medical Informatics Research Lunch, Image Guided Intervention Meetings, 2012-2016
- Attendance and presentation: literature meetings, 2012-2016

---

# Publications

---

## International Journals

- **P. Ambrosini**, I. Smal, D. Ruijters, W.J. Niessen, A. Moelker and T. van Walsum: A Hidden Markov Model for 3D Catheter Tip Tracking with 2D X-ray Catheterization Sequence and 3D Rotational Angiography. *IEEE Transaction on Medical Imaging*, vol. 36(3), pp. 757-768, 2017.
- **P. Ambrosini**, D. Ruijters, W.J. Niessen, A. Moelker and T. van Walsum: Continuous Roadmapping in Liver TACE Procedures Using 2D-3D Catheter-based Registration. *International Journal of Computer Assisted Radiology and Surgery*, vol. 10, pp. 1357-1370, 2015.

## International Conference Proceedings

- **P. Ambrosini**, D. Ruijters, W.J. Niessen, A. Moelker and T. van Walsum: Fully Automatic and Real-Time Catheter Segmentation in X-Ray Fluoroscopy. The 20th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), *Lecture Notes in Computer Science*, vol. 10434, pp. 577-585, 2017.
- H. Ma, **P. Ambrosini** and T. van Walsum: Fast Prospective Detection of Contrast Inflow in X-ray Angiograms with Convolutional Neural Network and Recurrent Neural Network. The 20th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), *Lecture Notes in Computer Science*, vol. 10434, pp. 453-461, 2017.
- **P. Ambrosini**, I. Smal, D. Ruijters, W.J. Niessen, A. Moelker and T. van Walsum: 3D Catheter Tip Tracking in 2D X-ray Image Sequences Using a Hidden Markov Model and 3D Rotational Angiography. The 10th MICCAI Workshop on Augmented Environments for Computer-Assisted Interventions (AE-CAI), *Lecture Notes in Computer Science*, vol. 9365, pp. 38-49, 2015.
- **P. Ambrosini**, D. Ruijters, A. Moelker, W.J. Niessen and T. van Walsum: 2D/3D Catheter-based Registration for Improved Image Guidance in TACE of Liver Tumors. The 5th International Conference on Information Processing in Computer-Assisted Interventions (IPCAI), *Lecture Notes in Computer Science*, vol. 8498, pp. 246-255, 2014.

### Conference Abstracts

- **P. Ambrosini**, D. Ruijters, W.J. Niessen, A. Moelker and T. van Walsum: Catheter Segmentation in X-ray Fluoroscopy using Convolutional Neural Network. The 31th International Congress on Computer Assisted Radiology and Surgery (CARS), Barcelona, Spain, June 22, 2017.
- **P. Ambrosini**, I. Smal, D. Ruijters, W.J. Niessen, A. Moelker and T. van Walsum: 3D Catheter Tip Tracking in 2D X-ray Image Sequences Using 3D Rotational Angiography. The 28th conference of the international Society for Medical Innovation and Technology (iSMIT), Delft, The Netherlands, October 6, 2016.
- E. Vast, **P. Ambrosini** and T. van Walsum: Spatial calibration of 2d/3d ultrasound with the PLUS framework and electro-magnetic tracking. 5th Dutch Conference on Bio-Medical Engineering, Egmond aan Zee, The Netherlands, January 22, 2015.

### Supervision

- Master thesis supervision of M.A. van der Cammen: 2D Fluoroscopy and 3D Computed Tomography Registration for Minimally Invasive Liver Procedures. TU Delft, 2016-2017.

---

# Curriculum Vitae

---

Pierre Ambrosini received a Master in Computer Science from the University of Lyon (France). He then worked three years in France as a video game developer and moved to the Netherlands to work as a research engineer at the Biomedical Imaging Group Rotterdam in Erasmus MC, University Medical Center Rotterdam (the Netherlands). With his group, he prototyped a workstation that aims at improving image guidance during medical interventions. The goal was to propose calibration methods in order to combine different imaging modalities and instruments such as computed tomography scan, magnetic resonance imaging, ultrasound imaging and optical/electromagnetic trackers. Afterwards, he continued his research in the same group as a PhD candidate on image guidance for Transcatheter Arterial ChemoEmbolization procedure (this thesis). His research mainly revolved around real-time registration, tracking and segmentation with X-ray images. From September 2017, he has been in a postdoctoral position at the Imaging Physics department of Delft University of Technology (TU Delft, the Netherlands) working on automatic detection of tumor growth patterns in prostate histopathology images.

