

# **3D/2D Registration of Medical Images**

**Dejan Tomažević**

## Colophon

This book was typeset by the author using Microsoft Word.

The cover was designed by the author with the assistance of Domen Pupič. It displays cubes forming a 3D optical illusion and the corresponding 2D “projections”. The 3D optical illusion is itself a 2D image, which requires the reader to mentally reconstruct its 3D appearance. In spite of the additional 2D projections which mislead the observer to believe in its exsistence, the 3D cube structure does not exist in reality. The cover was created to illustrate the non-trivial and ill-defined problem of registering 3D images and corresponding 2D projection images of real world 3D structures.

CIP - Kataložni zapis o publikaciji  
Narodna in univerzitetna knjižnica, Ljubljana

616-073:004.93

TOMAŽEVIĆ, Dejan  
[Three]  
3D/2D registration of medical images /  
Dejan Tomažević. -  
Ljubljana : Fakulteta za elektrotehniko, 2008

ISBN 978-961-243-088-7

238884608

Copyright © 2008 Založba FE in FRI. 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 publisher.

Published by “Založba FE in FRI”, Faculty of Electrical Engineering, University of Ljubljana, Slovenia, 2008 (Ed., Peter Šega, MSc)

Printed by Birografika BORI d.o.o.

# **3D/2D Registration of Medical Images**

**3D/2D Registratie van Medische Beelden**

(met een samenvatting in het Nederlands)

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR AAN DE UNIVERSITEIT UTRECHT  
OP GEZAG VAN DE RECTOR MAGNIFICUS PROF.DR. J.C. STOOF,  
INGEVOLGE HET BESLUIT VAN HET COLLEGE VOOR PROMOTIES  
IN HET OPENBAAR TE VERDEDIGEN OP  
DINSDAG 1 JULI 2008 DES MIDDAGS TE 2.30 UUR

DOOR

**Dejan Tomažević**

geboren op 18 januari 1976 te Novo mesto, Slovenië

Promotoren:      **Prof. dr. ir. M. A. Viergever**  
**Prof. dr. F. Pernuš**

# Contents

<b>1. Introduction .....</b>	<b>1</b>
1.1 3D/2D image registration.....	3
1.2 Outline of the thesis .....	12
<b>2. “Gold standard” data for evaluation and comparison of 3D/2D registration methods.....</b>	<b>17</b>
2.1 Introduction .....	18
2.2 Phantom creation .....	18
2.3 Image acquisition .....	19
2.4 Finding centers of fiducial markers .....	20
2.5 Retrospective calibration of X-ray setup.....	21
2.6 Reconstruction of 3D markers .....	22
2.7 “Gold standard” registration and validation .....	23
2.8 Publicly available “gold standard” data .....	25
2.9 Discussion and conclusion.....	26
<b>3. 3D/2D registration of CT and MR to X-ray images .....</b>	<b>29</b>
3.1 Introduction .....	30
3.2 Theory.....	33
3.2.1 X-Ray image generation .....	33
3.2.2 Image intensity and attenuation coefficient gradients .....	33
3.3 Method.....	35
3.4 Evaluation.....	37
3.4.1 “Gold standard” .....	37
3.4.2 Implementation details.....	38
3.4.3 Experiments.....	39
3.4.4 Results .....	39
3.5 Discussion and conclusion.....	42

3.6 Appendix A .....	46
3.7 Appendix B .....	47
3.8 Appendix C .....	47
<b>4. 3D/2D registration of CT and MR images to X-ray images: The impact of number and projection of X-ray images.....</b>	<b>49</b>
4.1 Introduction.....	50
4.2 3D/2D registration method.....	52
4.3 Images and “gold standard” registrations .....	53
4.4 Experiments .....	55
4.5 Results and discussion .....	57
4.6 Conclusion .....	66
<b>5. Standardized evaluation methodology for 2D-3D registration .</b>	<b>67</b>
5.1 Introduction.....	68
5.2 Evaluation methodology.....	69
5.2.1 Generation of gold standard data.....	69
5.2.2 Method for assessment of 2D-3D registration accuracy...	72
5.2.3 Capture range and failure criteria.....	74
5.2.4 Starting positions .....	74
5.2.5 Data preparation .....	75
5.3 Experiment 1: Comparison of two methods .....	76
5.3.1 Image data .....	78
5.3.2 Experiments and evaluation .....	78
5.4 Experiment 2: Evaluation of the gold standard .....	80
5.5 Results and discussion .....	80
5.5.1 Experiment 1: Comparison of two methods .....	80
5.5.2 Experiment 2: Evaluation of the gold standard .....	89
5.6 Conclusions.....	90
<b>6. 3D/2D registration by integrating 2D information in 3D .....</b>	<b>93</b>
6.1 Introduction.....	94
6.2 Registration of preoperative and reconstructed image .....	97
6.2.1 Asymmetric multi-feature mutual information .....	98

6.3 Experiments.....	100
6.3.1 Experimental data sets.....	100
6.3.2 Image preprocessing and implementation details.....	103
6.3.3 Registration experiment .....	104
6.4 Results.....	106
6.5 Discussion .....	111
6.6 Conclusion .....	115
6.7 Appendix.....	115
<b>7. Evaluation of similarity measures for reconstruction-based registration in image-guided radiotherapy and surgery .....</b>	<b>117</b>
7.1 Introduction .....	118
7.2 Materials and methods .....	121
7.2.1 Experimental data.....	121
7.2.2 Cone-beam CT .....	122
7.2.3 Similarity measure evaluation protocol.....	122
7.2.4 Similarity measures .....	124
7.3 Experiments and results .....	125
7.3.1 The impact of pre-treatment imaging modalities .....	126
7.3.2 The impact of the number of 2D images used for reconstruction.....	128
7.4 Discussion .....	131
7.5 Conclusion .....	135
<b>8. Robust gradient-based 3D/2D registration of CT and MR to X-ray images.....</b>	<b>137</b>
8.1 Introduction .....	138
8.2 Method.....	141
8.2.1 Calculation of gradient fields .....	141
8.2.2 Gradient correspondence maximization .....	143
8.2.3 Gradient reconstruction-based registration .....	144
8.3 Registration evaluation.....	146
8.3.1 Experimental data.....	147
8.3.2 Evaluation methodology .....	148

8.3.3 Implementation details .....	149
8.4 Discussion.....	155
8.5 Conclusion .....	158
<b>Bibliography .....</b>	<b>159</b>
<b>Samenvatting .....</b>	<b>177</b>
<b>Publications.....</b>	<b>181</b>
<b>Acknowledgements .....</b>	<b>185</b>
<b>About the author .....</b>	<b>187</b>

The beginning is half of the whole.

*Plato*

## Chapter 1

# Introduction

The history of medical imaging started with the discovery of X-ray imaging by the German physicists Wilhelm Conrad Röntgen at the end of the year 1895. From its early beginnings, X-ray imaging became an important part of medical diagnosis and treatment, while it took more than 60 years for other medical imaging techniques to emerge. The development of the sonar during World War II led to the first Ultrasound (US) imaging in the early 1960's, while digitalization and increased computing power enabled the development of Computed Tomography (CT) and Positron Emission Tomography (PET) in the 1970's and Magnetic Resonance Imaging (MRI) in the 1980's.

It is an established fact that different medical imaging techniques can improve diagnosis and treatment planning, as well as help the physician to control, conduct and evaluate a treatment. Image-supported medical interventions can reduce risks, improve quality, and decrease invasiveness of current interventions. Each medical imaging technique uses the advantages of a specific physical property or modality to acquire real world information about the patient and create a corresponding image. On the basis of the information provided, medical imaging techniques can be divided into anatomical and functional techniques. Anatomical imaging provides information on a patient's anatomy, while functional imaging provides information on its functionality. According to the number of image dimensions, imaging techniques are either two-dimensional (2D) or three-dimensional (3D). Three-dimensional imaging usually tries to present the information acquired in a 3D real world while preserving 3D world geometrical relations. Three-dimensional images can provide an excellent view of the details of an anatomy of interest. Widely used 3D imaging modalities like CT or MR imaging contain high-resolution information on the imaged anatomy. 3D imaging techniques, such as Positron Emission Tomography (PET) and Functional Magnetic Resonance Imaging (fMRI) complement anatomical knowledge with metabolic, i.e. functional,

information. Two-dimensional imaging of a 3D world is performed either by: projection imaging, where 3D real world information is projected to a 2D imaging plane, e.g. X-ray imaging; or by cross-section imaging, where a 2D cross-section image of a 3D real world is acquired, e.g. ultra sound (US) or single slice CT or MR. Furthermore, imaging techniques can be classified as diagnostic (pre-operative) and treatment (intraoperative) techniques. Although most of the imaging techniques can be used for both, diagnosis and treatment, different criteria such as limitations and requirements of specific medical procedures, image quality, invasiveness, speed and costs of imaging define their main field of use. Due to high costs and relatively low speed, 3D imaging usually assists in establishing a diagnosis, in treatment planning of surgical or radiological procedures, and in post-treatment evaluation. Projection 2D X-ray imaging techniques, such as fluoroscopy and digital radiology or cross-section 2D ultrasound (US) are usually cheaper and real-time. These techniques are used for both, diagnosis and planning in the pre-operative stage of medical interventions and for control of surgical or radiological interventions in the intraoperative phase. Although helpful, X-ray and US images are 2D, and as such lack spatial information contained in 3D images. Besides, there are a number of anatomical features which can be well visualized by, for example, CT or MR, but are hard to be seen by X-ray imaging or US.

Information that the anatomical and functional imaging modalities provide is usually essential for detecting abnormal conditions and relating these to diseases. However, the real strength of medical imaging is in acquiring and using the information from multiple images. For instance, multiple images of the same modality enable detection of pathological changes in anatomy and function in the course of time, while combining images of complementary modalities provides additional knowledge about normal and pathologic states of a patient's anatomy and function. Furthermore, images taken during surgical and radiological treatments can help to control the procedure, while a comparison of intraoperative images with preoperative ones can assist the physician to follow the preoperative plan and observe anatomical changes caused by the procedure. Moreover, medical science can benefit from comparison studies of images of different subjects, and gain new knowledge and understanding of the human body and its functionality.

The topic of this thesis is registration of 3D medical images to corresponding projective 2D images, referred to as 3D/2D registration. The goal of 3D/2D registration is to find spatial relations between a 3D image and one or more 2D projective images. There are numerous possible applications of 3D/2D registration in image-aided diagnosis and treatment. In most of the applications, 3D/2D registration provides the location and orientation of the structures in a preoperative 3D CT or MR image with respect to intraoperative 2D X-ray images or video. If the projection of the intraoperative 2D imaging system is known with respect to a reference patient coordinate system, 3D/2D registration can be used for registration of preoperative images to a patient. Proper visualization of surgical instruments and registered preoperative and

intraoperative images can provide the surgeon with information about the position of his instruments relative to the planned trajectory, nearby vulnerable structures, or the ultimate target (Galloway, 2001; Hawkes, 1998). In addition to image guided surgery, 3D/2D registration can assist in interpretation of 3D information from one or more 2D projection images. For example, the independent registration of a 3D image of patient anatomy and a 3D model of an implant to a single post-treatment 2D X-ray image can help in establishing the 3D positions and rotations between the anatomy and the implant.

## 1.1 3D/2D image registration

Image registration is concerned with finding a geometrical transformation that brings one image into the best possible spatial correspondence with another image or physical space. Maurer *et al.* (Maurer and Fitzpatrick, 1993) defined medical image registration as *the determination of a one-to-one mapping between the coordinates of one space and those in another such that points in the two spaces that correspond to the same anatomical point are mapped to each other*. In practice, a registration algorithm tries to find among all allowed spatial transformations between two or more spaces that transformation which fulfils or maximizes the image alignment criterion. The criterion is usually a cost function or similarity measure describing the spatial alignment of features representing the registration basis. The aim of registration is then to find a spatial transformation that optimizes the given criterion. However, in the particular case of 3D/2D registration the condition of one-to-one mapping of Maurer *et al.* (Maurer and Fitzpatrick, 1993) is not fulfilled, since projection of a 3D space to a 2D image is many-to-one mapping between the 3D and 2D space. As such, 3D/2D registration is in general an ill-posed problem. The problem can be avoided by limiting the degree of geometrical transformation, most often to rigid transformation, and by increasing the number of non-correlated 2D projections used in the registration.

Maintz *et al.* (Maintz and Viergever, 1998) proposed classification of registration algorithms according to their dimensionality, nature of the registration basis, nature and domain of geometrical transformation, nature of user interaction, type of optimization procedure, type of modalities involved, and subject and object of registration. Following this classification, the 3D/2D registration problem and the relevant work is categorized and discussed below.

### Dimensionality

According to dimensionality of the image space in which the registration is performed, registrations can be divided into *mono-dimensional* or *multi-dimensional*. Obviously, registration of a 3D image space to a 2D image space is *multi-dimensional* registration. According to Maintz *et al.* (Maintz and

Viergever, 1998) the term 2D-3D registration is referred to direct alignment of 3D spatial data to projective data or alignment of a single tomographic slice to 3D spatial data. In this thesis, we use the term 3D/2D registration for registration of 3D spatial data to 2D projective data, where projection is defined by the pin-hole camera model. Projection and the corresponding many-to-one mapping between the 3D and 2D image space makes 3D/2D registration more difficult than *mono-dimensional* registration. Moreover, the term 3D/2D registration is also used when registering a 3D image to 3D data reconstructed from a few 2D projection images, since such course of 3D reconstruction does not provide enough information to establish a one-to-one mapping between images. On the other hand, the registration of a single tomographic slice to 3D spatial data fulfills the condition of a one-to-one mapping. We therefore referred this type of registration as 2D/3D registration. Such a registration is just an extreme case of 3D/3D registration where one of the images is reduced to a single slice. However, a single 2D slice to 3D image data registration is not trivial but is concerned with different problems than registration of 3D image data to 2D projection data.

## Nature of registration basis

According to the nature of the registration basic, the 3D/2D registration methods can be classified as *extrinsic*, *intrinsic* and *non-image* based.

*Extrinsic* methods rely on artificial objects like, a stereotactic-frame or external point landmarks rigidly attached to patient's anatomy. These usually spherical shaped objects are designed to be well visible in 3D and 2D imaging modalities, and therefore their position can be accurately detected within image space. The registration is based on finding the transformation that minimizes the distance between the positions of corresponding objects detected in both, the 3D and the 2D images (Gall *et al.*, 1993; Tang *et al.*, 2000). To obtain the distance between 2D and 3D positions, these positions should be brought into a common image space. This can be done by projecting the 3D landmark positions into 2D or by reconstructing the 3D positions of landmarks from their positions in 2D images. The latter approach is feasible when more 2D projection images are available. A simple closed form solution for rigid registration is available when registering corresponding 3D positions (Arun *et al.*, 1987; Umeyama, 1991). Registration of images based on *extrinsic* features is easy, fast and does not require complex optimization algorithms. The main drawback of extrinsic methods is that they cannot be applied retrospectively as the markers must be attached in the pre-acquisition phase. Besides, the markers are often screwed or implanted and as such they are inconvenient and invasive to the patient. Less invasive markers can be used, like skin markers (Soete *et al.*, 2002), dental adapters or head mask (Boda-Heggemann *et al.*, 2006; Jin *et al.*, 2006), but these are usually less accurate. *Extrinsic* methods are mostly used for rigid registration, where they enjoy a reputation of means of accurate registration and as such usually serve for defining a reference or "gold standard" registration to which the accuracy of

*intrinsic* 3D/2D registration methods is usually compared (Khamene *et al.*, 2006; Kim *et al.*, 2001; Kim *et al.*, 2007; McLaughlin *et al.*, 2005; Penney *et al.*, 2001; Penney *et al.*, 1998; Rohlfing *et al.*, 2005a; Rohlfing *et al.*, 2005b; Russakoff *et al.*, 2005a; Russakoff *et al.*, 2005b; Tomažević *et al.*, 2004a; van de Kraats *et al.*, 2005a; Weese *et al.*, 1997; You *et al.*, 2001). *Extrinsic* markers can also be used to register non-rigid anatomical structures, for instance the prostate during the course of external beam radiotherapy (Aubry *et al.*, 2004; Litzenberg *et al.*, 2002).

Contrary to *extrinsic* registration, *intrinsic* registration relies on information of patient's anatomy and can be based on a set of geometrical features segmented from images (*segmentation-based*) or on the whole image contents (*whole-image-content-based*).

*Segmentation-based* algorithms require segmentation of corresponding sets of relevant geometrical features. These features may be corresponding points (landmarks) accurately located on the anatomy visible in both 3D and 2D images (Karger *et al.*, 2001; Salter *et al.*, 2001). Corresponding anatomical landmarks are usually selected by user. The registration is obtained by finding the minimal distance between corresponding landmarks. Unfortunately, such 3D/2D registration is generally not reliable since accurate location of anatomical landmarks is difficult especially in 2D projection images. However, this type of registration can be used to establish a rough initial registration which is refined by some more accurate automatic registration method (Penney *et al.*, 1998). The problem of landmark correspondence can be avoided by using 3D surfaces of anatomical structures extracted from pre-operative 3D data and the outlines of these structures on one or more intraoperative 2D projection images. Such an approach was proposed in (Feldmar *et al.*, 1997a; Groher *et al.*, 2006; Guéziec *et al.*, 1998; Kita *et al.*, 1998; Lavallée and Szeliski, 1995). Registration of a 3D surface and a 2D contour is performed by minimizing the distance between the surface model and the lines connecting contour points with the X-ray source. Once the outline is obtained by segmentation, the reduced amount of data makes the registration fast. Unfortunately, the intraoperative segmentation is hard to achieve automatically and errors in segmentation may propagate to errors in registration. Furthermore, a *segmentation-based* algorithms are sensitive to outliers, i.e. geometrical features extracted from one image which have no counterparts in the other image.

The *whole-image-content-based* approach to 3D/2D registration relies solely on information contained in voxels and pixels of 3D images and 2D projection images, respectively. To establish the correspondence between images, 3D pre- and 2D intra-interventional images have to be transformed into a space of the same dimensionality. Nowadays, the most popular *whole-image-content-based* 3D/2D registration method is based on simulated X-ray projection images, produced from the CT image, (Goitein *et al.*, 1983), called digitally reconstructed radiographs (DRRs). The registration of the CT volume relative to the X-ray image is estimated by optimizing the similarity measure

calculated between a DRR and an X-ray image (Birkfellner *et al.*, 2003; Dey and Napel, 2006; Gong *et al.*, 2006; Hipwell *et al.*, 2003; Hristov and Fallone, 1996; Jans *et al.*, 2006; Jaramaz and Eckman, 2006; Jonić *et al.*, 2001; Khamene *et al.*, 2006; Kim *et al.*, 2001; Kim *et al.*, 2007; LaRose *et al.*, 2000a; Lemieux *et al.*, 1994; Lujan *et al.*, 1998; Mahfouz *et al.*, 2003; Munbodh *et al.*, 2007; Murphy, 1997; Penney *et al.*, 2007; Penney *et al.*, 1998; Rohlfing *et al.*, 2005a; Rohlfing *et al.*, 2005b; Russakoff *et al.*, 2005b; Schweikard *et al.*, 2005; Sirois *et al.*, 1999; Weese *et al.*, 1997; You *et al.*, 2001; Zhang *et al.*, 2006). As these algorithms require little or no segmentation their performance does not suffer from segmentation errors. Moreover, since these methods utilize all the information in the images, it is reasonable to expect that they are more accurate and reliable than *segmentation-based* methods. This was proven by the comparative study of McLaughlin *et al.* (McLaughlin *et al.*, 2005), in which the authors showed that DRR-based 3D/2D registration was significantly more accurate and reliable than the *segmentation-based* 3D/2D registration, when registering phase-contrast MRA (PC-MRA) and 2D digital subtraction angiographs. Unfortunately, registration methods employing DRRs are time-consuming due to computational complexity of repeatable DRR generation during registration. This problem may be reduced by calculating DRRs containing only the structures of interest (Murphy, 1997; Sirois *et al.*, 1999; Weese *et al.*, 1997) or by implementing a faster DRR generation method (Birkfellner *et al.*, 2005; Birkfellner *et al.*, 2003; Hong *et al.*, 2006; LaRose *et al.*, 2000a; Li *et al.*, 2006; Rohlfing *et al.*, 2005b; Russakoff *et al.*, 2005b; Weese *et al.*, 1999). Different similarity measures can be used to measure the correspondence between DRRs and X-ray images. The most frequently used and studied measures for DRR-based 3D/2D registrations are normalized cross correlation, mean square difference, mutual information, entropy of difference image, as measures of global intensity correspondence, and gradient correlation, gradient difference and pattern intensity, as measures of correspondence of local intensity changes. Other measures, such as normalized mutual information (Rohlfing *et al.*, 2005a), chi-square and conditional variance, were also applied (Clippe *et al.*, 2003). From studies comparing different similarity measures (Kim *et al.*, 2007; Penney *et al.*, 1998) it can be concluded that similarity measures that measure global intensity correspondence are less suitable for matching DRRs with X-ray images. The measures of correspondences of local intensity changes produce substantially more accurate and reliable results. Unfortunately, DRR-based 3D/2D registration is not suitable for registration of MR and X-ray images because there is practically no correspondence between these imaging modalities, except in cases when contrast agents are used. Another drawback of DRR-based methods is that by projecting a high quality 3D CT image into 2D, valuable 3D information may be lost.

To enable registration of CT as well as MR images to X-ray images and to keep the 3D pre-interventional image intact, Tomažević *et al.* (Tomažević *et al.*, 2006) proposed that correspondence between 3D and 2D images is measured in the 3D image space. The approach requires coarse reconstruction of a 3D image from a few 2D projection images and implementation of multi-feature

mutual information similarity measure of image intensities and intensity gradients.

Due to specific requirements of 3D/2D registration, such as limited time available to perform the registration, registration methods that cannot be directly classified as *segmentation-based* or *whole-image-content-based* have been proposed. These methods coarsely segment either the 2D projection images (Gilhuijs *et al.*, 1996a; Livyatan *et al.*, 2003) or the 3D image (Jomier *et al.*, 2006; Tomažević *et al.*, 2003), but not both as the *segmentation-based* methods. Segmentation serves only for data reduction. Similarly to *whole-image-content-based* methods, registration is established by measuring the similarity between segmented and unsegmented image.

The 3D/2D registration between medical images can also be *non-image* based. Registration of MR and X-ray images, proposed by (Rhode *et al.*, 2003; Rhode *et al.*, 2005), was based on carefully pre-calibrated imaging devices and establishment of the position of the operation table with respect to the imaging device during image acquisition. The position of the operation table was established by means of an optical tracking device and infrared emitting diodes attached to the table and MR and X-ray imaging system. Since the position of patient anatomy relative to the operation table had to be fixed during both image acquisitions, this type of registration required proper patient immobilization and short times between pre- and intra-intervention image acquisitions.

## Nature of transformation

According to the nature of spatial transformation and its degrees of freedom, the methods can be classified as *rigid* and *non-rigid*. *Non-rigid* registration can further be divided into *affine*, *projection* and high-order *elastic* registration. The majority of published 3D/2D registrations employ a *rigid* transformation model, composed of translations and rotations. *Rigid* registration is generally applied when it is assumed that target anatomy fulfills the criterion of rigidity and no spatial distortions are induced in the image acquisition process. In case of *rigid* 3D/2D registration, the number of transformation parameters is six. The limited number of degrees of freedom makes the 3D/2D registration problem better defined, in spite of the fact that the condition for one-to-one mapping is not fulfilled. In contrast to the large number of *rigid* 3D/2D registrations, only a few *non-rigid* applications have been published. *Non-rigid* registration is required when the imaged anatomy may be non-rigidly deformed for example when internal organs deform due to breathing. *Non-rigid* 3D/2D registration is a challenging problem since the condition for one-to-one mapping between images is not fulfilled. *Non-rigid* 3D/2D registration was applied by (Behrenbruch *et al.*, 2000), to registered contrast-enhanced MR breast images with X-ray mammographs by a high-order thin-plate spline transformation. In the following work (Benameur *et al.*, 2005a; Benameur *et al.*, 2005b; Benameur *et al.*, 2003; Fleute and Lavallée, 1999; Sadowsky *et al.*,

2006; Zheng *et al.*, 2007; Zheng *et al.*, 2006) *non-rigid* 3D/2D registration was applied to obtain a 3D shape reconstruction of the anatomy of interest from a small number of X-ray projection images. In these studies the applied transformations were limited to the transformations allowed by statistical shape models. In the work of (Schweikard *et al.*, 2005) *non-rigid* 3D/2D registration was performed between *elastically* deformed 3D CT images and 2D portal images for respiratory tracking during radiosurgery.

## Nature of user interaction

While research activities make every effort to develop fully automated 3D/2D registration algorithms, some user interaction is still needed. The level of user interaction required can be *interactive*, *semi-automatic* or *automatic*. Fully *automatic* registration algorithms require that the user provides only the image data. Contrary to *automatic* registration, *interactive* registration is assisted by a user using software tools which provide visual feedback of the current transformation. Most often the level of user interaction is somewhere between these two extremes. For *semi-automatic* registration it is essential that the user provides some initialization to the algorithm, such as segmentation and/or the initial guess of registration, and/or to validate the results of registration. It can not be claimed that any of the published 3D/2D registration algorithms is fully *automatic*. An *interactive* registration performed by tuning translation and rotation parameters was implemented by (De Buck *et al.*, 2005; Gilhuijs *et al.*, 1996b). Unfortunately, accurate *interactive* registration is time consuming and depends on the skills of the human operator. However, a coarse *interactive* 3D/2D registration can be fast and can provide an initial guess of registration for subsequent *semi-automatic* procedures. Most of 3D/2D registration methods are *semi-automatic* and require different levels of user interaction. In landmark-based methods, the user has to provide a set of corresponding anatomical landmarks (Karger *et al.*, 2001; Salter *et al.*, 2001). For 3D/2D registration based on 3D surfaces and corresponding 2D outlines of anatomical structures (Feldmar *et al.*, 1997a; Groher *et al.*, 2006; Guéziec *et al.*, 1998; Kita *et al.*, 1998; Lavallée and Szeliski, 1995) a user may be needed to adjust the parameters of the automatic 3D and 2D segmentation methods and to validate and/or correct segmentation results. Minimal user interaction may be required by *whole-image-content-based* 3D/2D registration algorithms. In these registration algorithms it is essential that the initial guess of the registration is close to true registration. This condition may be fulfilled by accurate positioning of the patient with respect to imaging devices by means of laser markers or by *interactive* coarse 3D/2D registration performed by the user (Hipwell *et al.*, 2003; Khamene *et al.*, 2006; Penney *et al.*, 2007; Penney *et al.*, 1998). Since the 3D/2D registration problem is by its nature an ill-defined problem, 3D/2D registration algorithms may provide registration results that are false. To avoid the risk of guiding a medical procedure with false registration results, the results of 3D/2D registration should be validated by the user. Initial *interactive* user registration and final validation of registration results my be

supported by proper visualization such as superimposing the outlines of a 3D surface (Feldmar *et al.*, 1997a; Guéziec *et al.*, 2000; Hipwell *et al.*, 2003; Jomier *et al.*, 2006) or edges extracted from DRRs (Gilhuijs *et al.*, 1996a; Khamene *et al.*, 2006; Livyatan *et al.*, 2003) onto 2D X-ray images.

## Optimization procedure

By its definition, image registration is concerned with finding a geometrical transformation that brings one image into the best possible spatial correspondence with another image or physical space by optimizing a registration criterion. The parameters that describe a geometrical transformation can be *computed* directly or *searched for*. Direct *computation* of transformation parameters is possible only when the one-to-one mapping between a limited set of points found in both images is known (Gall *et al.*, 1993; Tomažević *et al.*, 2004a). Generally, point pairs are not available. Therefore, registration parameters have to be *searched for* iteratively by minimizing the distance between corresponding feature sets. The Levenberg-Marquardt (Guéziec *et al.*, 1998; Lavallée and Szeliski, 1995; Tang *et al.*, 2000) or Kalman Filter (Feldmar *et al.*, 1997a) optimization methods may be used. In the case of *whole-image-content-based* registration the registration criterion is formulated as a similarity measure defined in multidimensional space of searched parameters. It is desired that the similarity measure is well-behaved, i.e. is monotone and quasi-convex in the vicinity of the true registered position. The commonly implemented search techniques are Powell's method (Birkfellner *et al.*, 2003; Rohlfing *et al.*, 2005b; Tomažević *et al.*, 2006; Tomažević *et al.*, 2003; You *et al.*, 2001), the Downhill Simplex method (Kim *et al.*, 2001; Kim *et al.*, 2007) and gradient decent (Munbodh *et al.*, 2007; Murphy, 1997; Penney *et al.*, 2001; Penney *et al.*, 2007; Penney *et al.*, 1998; Rohlfing *et al.*, 2005a; Russakoff *et al.*, 2005b). Moreover, the Monte-Carlo random sampling in conjunction with Powell's optimization method was investigated in (Dey and Napel, 2006), while (Gong *et al.*, 2006) implemented the Kalman Filter and (Mahfouz *et al.*, 2003) employed simulated annealing optimization. When the formulation of a similarity measure allows an implicit estimation of its second order derivatives the Levenberg-Marquardt least square optimization can be implemented (Jonić *et al.*, 2001; Zhang *et al.*, 2006).

To avoid that registration is trapped in false local optima and to speed up the registration process different *whole-image-content-based* methods often implement a hierarchical coarse-to-fine, multi-scale and multi-resolution search strategy (Birkfellner *et al.*, 2003; Jonić *et al.*, 2001; Khamene *et al.*, 2006; Lemieux *et al.*, 1994; McLaughlin *et al.*, 2005; Penney *et al.*, 2001; Penney *et al.*, 2007; Penney *et al.*, 1998; Rohlfing *et al.*, 2005b; Russakoff *et al.*, 2005a; Russakoff *et al.*, 2005b; Zheng *et al.*, 2006).

## Modalities involved

According to the type of modalities involved, 3D/2D registrations can be classified into *modality-to-modality* and *model-to-modality*. 3D/2D registrations are thus always multi-modal since information from different modalities are involved in the registration process. In general, the preinterventional imaging modalities are 3D CT or MR, while the intra-interventional modalities are 2D X-ray or video. For CT, MR and X-ray imaging, contrast agents may be used, while for MR imaging different types of weighting may be implemented. X-rays may be acquired at various energies using different types of X-ray detectors (image intensifier, digital X-ray detector). From the modalities above, CT and X-ray imaging employ the same physical properties. Therefore, these two modalities are more closely related and easier to be registered than, for example, MR to X-rays. Most medical applications perform 3D/2D image registration between CT and X-rays (Birkfellner *et al.*, 2003; Dey and Napel, 2006; Gong *et al.*, 2006; Groher *et al.*, 2006; Guéziec *et al.*, 1998; Hristov and Fallone, 1996; Jans *et al.*, 2006; Jonić *et al.*, 2001; Khamene *et al.*, 2006; Kim *et al.*, 2001; Kim *et al.*, 2007; LaRose *et al.*, 2000a; Lemieux *et al.*, 1994; Livyatan *et al.*, 2003; Lujan *et al.*, 1998; Munbodh *et al.*, 2007; Murphy, 1997; Nakajima *et al.*, 2007; Penney *et al.*, 1998; Rohlfing *et al.*, 2005a; Rohlfing *et al.*, 2005b; Russakoff *et al.*, 2005b; Schweikard *et al.*, 2005; Sirois *et al.*, 1999; Weese *et al.*, 1997; You *et al.*, 2001; Zhang *et al.*, 2006). There are fewer applications of MR to X-ray registration (Behrenbruch *et al.*, 2000; De Buck *et al.*, 2005; Hipwell *et al.*, 2003; Jomier *et al.*, 2006; Kita *et al.*, 1998; McLaughlin *et al.*, 2005; Rhode *et al.*, 2003; Rhode *et al.*, 2005; Tomažević *et al.*, 2006; Tomažević *et al.*, 2003; van de Kraats *et al.*, 2005b), and 3D model to video registration (Deligianni *et al.*, 2006). 3D/2D registration may also be a *model-to-modality* registration. A model can be a 3D statistical model of an anatomy (Benameur *et al.*, 2005a; Benameur *et al.*, 2005b; Benameur *et al.*, 2003; Fleute and Lavallée, 1999; Sadowsky *et al.*, 2006; Zheng *et al.*, 2007; Zheng *et al.*, 2006) or an exact 3D geometrical model of an implant (Jaramaz and Eckman, 2006; Penney *et al.*, 2007; Yamazaki *et al.*, 2004).

## Subject of registration

According to the subject, 3D/2D registrations can be divided into *intrasubject* and *atlas* registrations. *Intrasubject* registration is performed between images of the same patient. This type of 3D/2D registration is most common, since registration between 3D and 2D images is usually used to establish the position of 3D imaged anatomy of the patient in the coordinate system of the treatment room. *Atlas* registration is performed between images of the patient and the corresponding atlas, where atlas is not an image of a normal subject but is constructed from the image database of many subjects. If the atlas provides the expected difference in scale, shape and structure between the subjects in the form of a parametrical model, such a 3D/2D registration can be better defined than by general *elastic* registration. In 3D/2D registration applications *atlas* registration is usually applied to obtain a 3D shape

reconstruction of the anatomy of interest from a limited number of X-ray projection images (Benameur *et al.*, 2005a; Benameur *et al.*, 2005b; Benameur *et al.*, 2003; Fleute and Lavallée, 1999; Sadowsky *et al.*, 2006; Zheng *et al.*, 2007; Zheng *et al.*, 2006).

## Object of registration and applications

Numerous medical interventions, like surgery, radiotherapy, radiosurgery, biopsy, ablation, catheterization, etc., can benefit from proper 3D/2D medical image registration. In these interventions different parts of human anatomy, as the head, thorax, abdomen, pelvis, limbs, spine, vascular structures, etc., can be objects of 3D/2D registration.

The intervention that currently most benefit from 3D/2D image registration is radiotherapy. In radiotherapy, it is crucial that the treated anatomy is accurately aligned with radiation beams. The alignment of a patient has to be performed on a daily basis before each individual treatment in the whole series of treatments. In the current clinical practice two X-ray portal images are acquired to verify the patient position. The 3D/2D registration between diagnostic CT or MR images and acquired X-ray images can help correcting patient position errors. 3D/2D registration can be used for treatment of brain, spine, abdomen and prostate tumors. For radiotherapy of brain tumors the skull is usually used as the object of registration (Gall *et al.*, 1993; Gilhuijs *et al.*, 1996a; Gilhuijs *et al.*, 1996b; LaRose *et al.*, 2000a; Lemieux *et al.*, 1994; Murphy, 1997; Sirois *et al.*, 1999) since it can be expected that the relation between brain and skull is almost rigid. In spine procedures individual vertebrae are registered (Rohlfing *et al.*, 2005a; Russakoff *et al.*, 2005a). For abdomen radiotherapy movements of internal organs and tumor have to be compensated. These movements can be defined by combining 3D/2D registration of the rib cage and an optical tracker (Schweikard *et al.*, 2005). The alignment of prostate can be based on pelvis registration (Jans *et al.*, 2006; Munbodh *et al.*, 2007) or registration of extrinsic markers implanted in the prostate (Aubry *et al.*, 2004; Litzenberg *et al.*, 2002).

A lot of current orthopedic procedures require 3D planning from preoperative CT or MR images, while guidance is usually performed by interventional C-arm X-ray imaging. In such procedures 3D/2D registration can provide the surgeon with information about the position of his instruments relative to the planned trajectory. In spine surgery, reinforcement of failing vertebra by placement of cement and pedicle screw placement are just some possible applications of 3D/2D image registration. Since spine is a non-rigid anatomy, registration is based on single vertebra rigid registration (Livyatan *et al.*, 2003; Penney *et al.*, 2001; Penney *et al.*, 1998; Tomažević *et al.*, 2006; Tomažević *et al.*, 2003; van de Kraats *et al.*, 2005a; Zhang *et al.*, 2006). The 3D/2D registration can also improve the total hip replacement procedure. The 3D/2D registration of femur images can improve positioning of femoral implant (Guéziec *et al.*, 1998), while 3D/2D registration of pelvis and implant can be

applied to analyze postoperative X-ray images for possible errors in cup positioning (Jaramaz and Eckman, 2006; Penney *et al.*, 2007). 3D/2D image registration of lower extremity bones and knee implants can be used to analyze knee kinematics before (You *et al.*, 2001) and after (Mahfouz *et al.*, 2003; Yamazaki *et al.*, 2004) a total knee arthroplasty procedure, respectively. Orthopedic diagnostics may also benefit from a 3D analysis of 2D images supported by 3D/2D registration. Such applications are measurement of 3D curvature of scoliotic spine (Benameur *et al.*, 2005b) or 3D reconstruction and classification of scoliotic rib cage (Benameur *et al.*, 2005a) from biplanar X-ray images.

There are a number of serious illnesses, which can be treated by the use of catheters that are maneuvered in blood vessels. In those procedures a clinician follows the path of the catheter in the patient's body with the help of dynamic intra-operative 2D X-ray imaging. For the purpose of diagnosis and planning, 3D CT or MR images may also be acquired prior to treatment. To provide 2D and 3D images with high contrast vessel structures an opaque contrast material has to be injected into the patient just before imaging. Due to the complex 3D structure of vessel trees, catheter navigation supported by 2D imaging may not be trivial. Navigation may be improved by 3D/2D registration. By incorporating 3D/2D registration the "road map" obtained from 3D images may be projected and visualized on dynamic 2D X-ray images or the position of the catheter detected on 2D X-ray images may be visualized on preoperative 3D images (Baert *et al.*, 2003). The support of 3D/2D registration was proposed for neuro (Hipwell *et al.*, 2003; McLaughlin *et al.*, 2005), cardiac (De Buck *et al.*, 2005; Rhode *et al.*, 2003; Rhode *et al.*, 2005) and liver (Groher *et al.*, 2006; Jomier *et al.*, 2006) interventions.

The 2D X-ray mammography is a widespread and efficient tool for breast pathology assessment, while the emerging 3D techniques like contrast-enhanced magnetic resonance imaging may provide complementary diagnostics information. Behrenbruch *et al.* (Behrenbruch *et al.*, 2000) proposed that diagnostics may benefit from data fusion of 2D X-ray mammography images and contrast enhanced 3D MR images supplied by 3D/2D registration.

## 1.2 Outline of the thesis

The proposed doctoral dissertation tries to find original solutions for the problem of multi-dimensional 3D/2D medical image registration of 3D images with 2D projection X-ray images. The dissertation also addresses the important problems of validation and comparison of different registration methods.

A necessary step, required prior to the widespread clinical use of any novel registration technique, is the evaluation and validation of the method (Jannin *et al.*, 2002). Evaluation and comparison with other methods is an important

problem, but have so far received little attention in the literature. Chapter 2 of the dissertation focuses on the challenging problem of generating reliable “gold standard” data for evaluating the accuracy and reliability of 3D/2D rigid registration methods. Major steps in a construction of “gold standard” data are the following; 1. Definition of the registration basis, for highly accurate registration. 2. Preparation of a patient for *in-vivo* experiment or physical phantom that mimics the chosen anatomy for *in-vitro* experiment. 3. Calibration of image acquisition system if necessary. 4. Image acquisition. 5. Image registration. 6. Evaluation of registration accuracy. Chapter 2 describes a construction of a lumbar spine phantom and the acquisition of CT, MR and digital X-ray images of the phantom. The calibration of 2D projection and “gold standard” registration were obtained by means of fiducial markers attached to the phantom, while the “gold standard” registration accuracy was evaluated by analyzing the residual error of marker registration (Fitzpatrick *et al.*, 1998).

In Chapter 3, we propose a novel method for registering 3D CT or MR images to 2D X-ray images with the final goal to estimate the position and orientation of patient’s anatomy during surgery or therapy. The originality of the approach is in using normals to bone surfaces, preoperatively segmented from 3D MR or CT data, and gradients of intraoperative X-ray images at locations defined by the X-ray source and 3D surface points. The method uses a novel criterion function to measure the match between normals to the surface and the corresponding intensity gradients of X-ray images. Such matching avoids intraoperative segmentation of X-ray images and thus eliminates misregistration induced by intraoperative segmentation errors, while the segmentation of 3D images reduces the amount of data, which makes the registration faster. For validation of the proposed 3D/2D registration method we introduce a new validation methodology. The proposed evaluation protocol requires determination of a wide range of initial registration positions, translations and rotations, whose distribution is random and uniform in a given interval from the “gold standard” position. The *accuracy* is defined as root mean square (RMS) target registration error (TRE) of all successful registrations. The criterion for successful registration is the maximum TRE below the clinically tolerable error. The *reliability* of the method is expressed by the percentage of successful registrations from all registrations whose initial position is in a given interval from the “gold standard” position.

The accuracy and robustness of registering a 3D preoperative image with 2D intraoperative images depend on the registration method and the anatomical structures that are to be registered. However, it is expected that the quality of 3D/2D registration also depends on the number of 2D images and the projections under which these images are acquired. In Chapter 4 we have conducted a study on the impact of the number of X-ray views and the angle between image views used for registration. The study was conducted for the 3D/2D gradient-based registration method proposed in Chapter 3 and on the “gold standard” image data from Chapter 2. The *accuracy* and *reliability* of the 3D/2D registration method was evaluated for sixteen different X-ray image

sets used for registration of CT and MR images to X-ray images. The X-ray image sets differ in number of X-ray images and the angle between X-ray image views. The results for both, CT/X-ray and MR/X-ray, registrations indicate that the number of X-ray images has a rather small impact on accuracy but that the proportion of successful registrations (robustness) becomes significantly higher when more X-ray images are used for registration. The accuracy and proportion of successful registrations of CT/X-ray registrations does not significantly depend on the projections. However, the projections have an impact on the proportion of successful MR/X-ray registrations.

Chapter 5 addresses the problem of evaluation and comparison of different 3D/2D registration methods. In this chapter, we describe the generation of novel 3D/2D registration “gold standard” image data. The data comprises 2D fluoroscopic X-ray images and 3D 3DRX, CT and MR images of two defrosted segments of vertebral bodies. One hundred 2D X-ray fluoroscopic images were obtained with a clinical 3D rotational X-ray (3DRX) system and 3DRX image was reconstructed for each spinal segment. The gold standard registration between CT and MR images with 2D fluoroscopic images was obtained by 3D/3D rigid registration of CT and MR images to the corresponding 3DRX image using the mutual information maximization registration method (Maes *et al.*, 1997). Besides image data and gold standard registration, this chapter presents a novel evaluation methodology for 3D/2D registration evaluation. Similarly to the methodology described in Chapter 3, this methodology is comprised of various starting positions that have to be used in registration experiment and the definition of accuracy, capture range, and failure criteria. The methodology was used for the comparison of two different 3D/2D registration methods, the DRR based method of (Penney *et al.*, 1998) and gradient-based method proposed in Chapter 3.

So far, all multi-dimensional 3D/2D registration methods registering 3D images with multiple 2D X-ray projection images have treated each X-ray image involved in the registration process separately. The alignment criterion is generally calculated as a sum of all partial criteria between the 3D image and each 2D X-ray image. In Chapter 6, we are proposing a new and totally different approach to the 3D/2D registration problem where information from different 2D X-ray images is merged in a common 3D space, while registration is based on matching the 3D image with the merged 3D information. For merging image information from 2D X-ray images in the 3D image space, we are proposing the usage of iterative algebraic reconstruction technique (ART) (Kak and Slaney, 1988). For measuring the alignment between a 3D image and a 3D reconstructed counterpart we are proposing the usage of a novel asymmetric multi-feature mutual information similarity measure. The proposed 3D/2D registration has been evaluated by using the publicly available spine phantom image data and standardized evaluation methodology described in Chapter 5. By using the same methodology, the performance of the proposed method was compared with the performance of the 3D/2D registration method presented in Chapter 3, which implements the standard

approach of calculating the sum of partial alignment criteria. The comparisons indicate that the proposed registration approach outperforms the standard approach to 3D/2D registration.

In Chapter 7, different similarity measures for reconstruction-based 3D/2D registrations are evaluated. These similarity measures are the, asymmetric gradient based mutual information described in Chapter 6, symmetric gradient based mutual information (Tomažević *et al.*, 2006), mutual information (Maes *et al.*, 1997; Wells *et al.*, 1996), normalized mutual information (Studholme *et al.*, 1999), entropy correlation coefficient (Astola and Virtanen, 1982; Maes *et al.*, 1997), joint entropy (Collignon *et al.*, 1995; Studholme *et al.*, 1995), correlation ratio (Roche *et al.*, 1998), Woods criterion (Woods *et al.*, 1993) and Pearson's correlation coefficient. The measures are evaluated for different number of 2D projections and different number of ART reconstruction iterations. The evaluation is carried out by implementing a protocol for rigid registration similarity measure evaluation (Škerl *et al.*, 2006a) and the evaluation methodology proposed in Chapter 5. From the evaluation it can be concluded that using a higher number of 2D projections or reconstruction iterations results in higher accuracy and slightly lower robustness. The most appropriate similarity measure for reconstruction-based 3D/2D registrations proved to be the asymmetric multi-feature mutual information.

A novel 3D/2D registration method is proposed in Chapter 8. The method is based on robustly matching 3D preoperative image gradients and coarsely reconstructed 3D gradients from interventional 2D images. To improve the robustness of finding the correspondences between the two sets of gradients, hypothetical correspondences are searched for along normals to anatomical structures in 3D images, while the final correspondences are established in an iterative process, combining the robust random sample consensus algorithm (RANSAC) and a special gradient matching criterion function. The proposed method is evaluated on the publicly available standardized evaluation methodology for 3D/2D registration presented in Chapter 5. In this way, the proposed method may be objectively compared to three state-of-the-art registration methods, namely, the intensity (Penney *et al.*, 1998), gradient (Chapter 3), and reconstruction-based methods (Chapter 6). The obtained results indicate that the proposed method performs favorably both in term of registration accuracy and robustness. The method is especially superior when just a few X-ray images and when MR pre-interventional images are used for registration.



Where facts are few, experts are many.

*Donald R. Gannon*

## **Chapter 2**

# **“Gold standard” data for evaluation and comparison of 3D/2D registration methods**

Dejan Tomaževič, Boštjan Likar and Franjo Pernuš  
Computer Aided Surgery 9: 137-144 (2004)

### **Abstract**

*Evaluation and comparison of registration techniques needed for image-guided surgery is an important problem, which received little attention in the literature. In this paper we address the challenging problem of generating reliable “gold standard” data for evaluating the accuracy of 3D/2D registrations. We have devised a cadaveric lumbar spine phantom with fiducial markers and established highly accurate correspondences between 3D CT and MR images and 18 2D X-ray images. The expected target registration errors for target points on the pedicles are below 0.26 mm for CT to X-ray registration and below 0.42 mm for MR to X-ray registration. As such, the “gold standard” data, which has been made publicly available on the Internet (<http://lit.fe.uni-lj.si/Downloads/downloads.asp>), is useful for evaluation and comparison of 3D/2D image registration methods.*

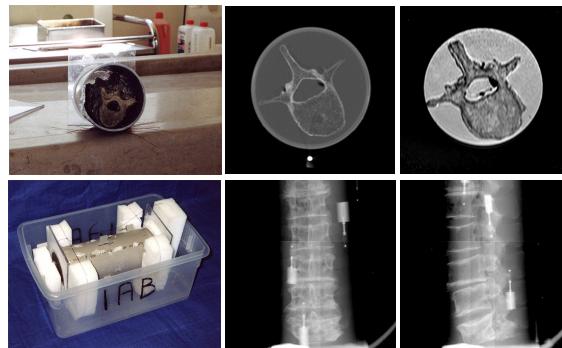
## 2.1 Introduction

In image-guided surgery, 3D preoperative medical data, such as computed tomography (CT) and magnetic resonance (MR) images, are commonly used to plan, simulate, guide, or otherwise assist a surgeon in performing a medical procedure. The plan, specifying how tasks are to be performed during surgery, is developed in the coordinate system of a preoperative image. To monitor and guide a surgical procedure, the preoperative image and plan need to be transformed into physical space, i.e. a patient-related coordinate system. The spatial transformation is obtained by acquiring intraoperative data and registering them to data extracted from preoperative images (Galloway, 2001). More recent and promising approaches to obtain the spatial transformation rely on intraoperative X-ray projections acquired with a calibrated X-ray device. The location and orientation of a structure in 3D CT or MR image with respect to the geometry of the X-ray device is determined by 3D/2D registration (Guéziec *et al.*, 1998; Lavallée and Szeliski, 1995; Lemieux *et al.*, 1994; Penney *et al.*, 1998).

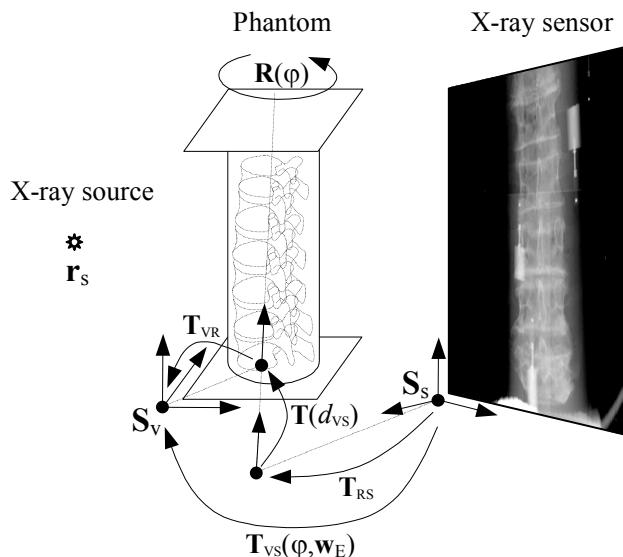
A necessary step, required before wide spread clinical use of any novel registration technique, is the evaluation and validation of the method (Jannin *et al.*, 2002). One difficulty in evaluating a registration technique is the need for highly accurate "gold standard". Motivated by the continuing need for data sets and "gold standard" to test, validate, and measure the accuracy of different 3D/2D registration techniques, we have devised a cadaveric lumbar spine phantom because it is very hard to establish "gold standard" with real patient data. In this paper we describe the construction of the phantom to which fiducial markers were attached. 3D CT, MR, and 2D X-ray images were acquired and accurate "gold standard" rigid registration between 3D and 2D images was established by means of fiducial markers. The accuracy of "gold standard" registration was assessed by target registration error (TRE) (Fitzpatrick *et al.*, 1998).

## 2.2 Phantom creation

A cadaveric lumbar spine, comprised of vertebra L1-L5 with intervertebral disks and several millimeters of soft tissue, of an 80 year-old female was placed into a plastic tube and tied with thin nylon strings (Figure 2.1, left). The tube was filled with water to simulate soft tissue and, therefore, to obtain more realistic MR, CT, and X-ray images. Six fiducial markers (Stryker Leibinger, Germany) were rigidly attached to the outside of plastic tube. Each fiducial marker had two parts, a base that could be screwed to a rigid body and a replaceable marker. Different markers were used for MR and CT and X-ray imaging. Markers, containing a metal ball (1.5 mm in diameter) were used for CT and X-ray imaging, while markers with a spherical cavity (2 mm in diameter) filled with water solution of Dotarem contrast agent produced by Gothia, Sweden, were used for MR.



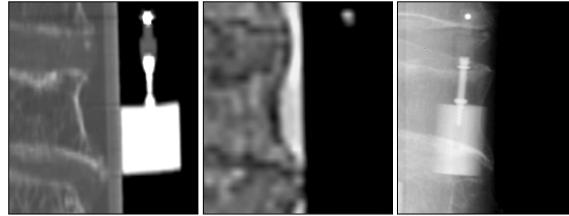
**Figure 2.1.** The spine fastened in a plastic tube (top-left), final phantom with fiducial markers attached to the outside of the plastic tube (bottom-left), CT image (top-center), MR image (top-right), and AP (bottom-center), and lateral X-ray image (bottom-right) of the phantom.



**Figure 2.2.** X-ray image acquisition.

## 2.3 Image acquisition

The CT image (Figure 2.1, top-center) was obtained with General Electric HiSpeed CT/i scanner. Axial slices were taken with intraslice resolution of  $0.27 \times 0.27$  mm and 1 mm interslice distance. For MR imaging, Philips Gyroscan NT Intera 1.5 T scanner and T1 protocol (flip angle 90°, TR=3220 ms, TE=11 ms) was used (Figure 2.1, top-right). Axial slices were obtained with  $0.39 \times 0.39$  mm intraslice resolution and 1.9 mm interslice distance. After



**Figure 2.3.** Close up views of fiducial markers in CT (left), MR (middle), and X-ray (left) images. The markers are indicated by arrows.

acquisition, the acquired MR image was retrospectively corrected for intensity inhomogeneity by the information minimization method (Likar *et al.*, 2001). X-ray images (Figure 2.1) were captured by PIXIUM 4600 (Trixell, France) digital X-ray detector. The detector had a 429x429 mm active surface, with 0.143x0.143 mm pixel size and 14-bit dynamic range. To simulate C-arm acquisition, the X-ray source and sensor plane were fixed while the spine phantom was rotated on a turntable (Figure 2.2). In this way, mechanical distortion due to gravitational force and other mechanical imperfections of C-arms were avoided. By rotating (step=20°) the spine phantom around its long axis, 18 X-ray images were acquired. The X-ray images were filtered by 3x3 median filter and then subsampled by the factor of two in order to remove dead pixel artifacts.

## 2.4 Finding centers of fiducial markers

Figure 2.3 shows close up views of fiducial markers in CT, MR, and X-ray images. In all 3D and 2D images a rough position  $\mathbf{p}_m$  of each fiducial marker was first defined manually. Next, an intensity threshold  $I_T$ , that separated a marker from surrounding tissues, was selected for each marker. Finally, the center  $\mathbf{p}_c$  of each marker was defined as:

$$\mathbf{p}_c = \frac{\sum_{\mathbf{p} \in \Omega} (I(\mathbf{p}) - I_T) \mathbf{p}}{\sum_{\mathbf{p} \in \Omega} (I(\mathbf{p}) - I_T)}, \quad (2.1)$$

where  $I(\mathbf{p})$  is the intensity at point  $\mathbf{p}$  and  $\Omega$  is a small neighborhood around point  $\mathbf{p}_m$ . By this method, centers of markers may be found to subpixel or subvoxel accuracy (Bose and Amir, 1990; Chiorboli and Vecchi, 1993).

Let  $\mathbf{X}_{CT}$  and  $\mathbf{X}_{MR}$  be 3x6 matrices, each containing six 3D vectors representing the centers of fiducial markers found in CT and MR, respectively:

$$\begin{aligned} \mathbf{X}_{MR} &= [\mathbf{r}_1^{MR}, \mathbf{r}_2^{MR}, \dots, \mathbf{r}_6^{MR}] \\ \mathbf{X}_{CT} &= [\mathbf{r}_1^{CT}, \mathbf{r}_2^{CT}, \dots, \mathbf{r}_6^{CT}] \end{aligned}, \quad (2.2)$$

where  $\mathbf{r}=(x,y,z)^T$ . Similarly, let  $\mathbf{X}_\varphi$  be a  $2 \times 6$  matrix containing six 2D vectors representing the centers of markers found in X-ray images obtained after rotating the phantom for  $\varphi$  degrees ( $\varphi=0^\circ, 20^\circ, \dots, 340^\circ$ ):

$$\mathbf{X}_\varphi = [\mathbf{p}_1^\varphi, \mathbf{p}_2^\varphi, \dots, \mathbf{p}_6^\varphi], \quad (2.3)$$

where  $\mathbf{p}=(x,y)^T$ .

## 2.5 Retrospective calibration of X-ray setup

To be able to reconstruct the positions of fiducial markers in 3D from their positions in 2D X-ray images, the acquisition setup (Figure 2.2) was retrospectively calibrated. Calibration required the determination of 12 geometrical parameters, 3 intrinsic  $\mathbf{w}_I$  and 9 extrinsic  $\mathbf{w}_E$ , denoted by calibration parameter vector  $\mathbf{w}$ ,  $\mathbf{w}=(\mathbf{w}_I, \mathbf{w}_E)$ . The intrinsic parameters  $\mathbf{w}_I=(x_s, y_s, z_s)$  define the position of the X-ray source  $\mathbf{r}_s$  in the coordinate systems  $\mathbf{S}_s$  of the sensor plane and, therefore, define the projection  $\mathbf{P}_S(\mathbf{w}_I)$  of any 3D point described in the sensor coordinate system  $\mathbf{S}_s$  to the 2D sensor plane. The extrinsic parameters  $\mathbf{w}_E$  describe the geometrical relation between the rotating phantom and the X-ray system. Four parameters define the axis of rotation in coordinate system  $\mathbf{S}_v$  of the phantom. We have chosen the coordinate system of the CT volume for  $\mathbf{S}_v$ . The axis of rotation is defined by point  $(tx_v, ty_v)$ , which is the intersection of the axis with  $x$ - $y$  coordinate plane of  $\mathbf{S}_v$  and by rotation  $(\omega x_v, \omega y_v)$  of the axis around  $x$  and  $y$  of  $\mathbf{S}_v$ . Similarly, four parameters  $(tx_s, ty_s)$  and  $(\omega x_s, \omega y_s)$  define the same axis of rotation in coordinate system  $\mathbf{S}_s$  of the X-ray sensor plane. The last parameter, needed to determine the relation between  $\mathbf{S}_s$  and  $\mathbf{S}_v$  on the rotation axis, is distance  $d_{vs}$  between the two points of intersection  $(tx_v, ty_v)$  and  $(tx_s, ty_s)$ . The extrinsic parameters  $\mathbf{w}_E=(tx_v, ty_v, \omega x_v, \omega y_v, d_{vs}, tx_s, ty_s, \omega x_s, \omega y_s)^T$  thus define the transformation  $\mathbf{T}_{VS}(\varphi, \mathbf{w}_E)$ :

$$\mathbf{T}_{VS}(\varphi, \mathbf{w}_E) = \mathbf{T}_{RS}(tx_s, ty_s, \omega x_s, \omega y_s) \cdot \mathbf{T}(d_{vs}) \cdot \mathbf{R}(\varphi) \cdot \mathbf{T}_{VR}(tx_v, ty_v, \omega x_v, \omega y_v) \quad (2.4)$$

that maps, for a given rotation  $\varphi$  of the phantom, any 3D point in coordinate system  $\mathbf{S}_v$  to a 3D point in coordinate system  $\mathbf{S}_s$  (Figure 2.2).  $\mathbf{T}_{VR}$  is the transformation from coordinate system  $\mathbf{S}_v$  to the axis of rotation,  $\mathbf{R}(\varphi)$  is the rotation around rotation axis,  $\mathbf{T}(d_{vs})$  is the translation along rotation axis, and  $\mathbf{T}_{RS}$  is the final transformation to the coordinate system  $\mathbf{S}_s$ . For any rotation  $\varphi$ , the projection  $\mathbf{P}_{VS}(\varphi, \mathbf{w})$  of a 3D point defined in the coordinate system  $\mathbf{S}_v$  to the 2D point lying in the sensor plane defined by  $\mathbf{S}_s$  is obtained by applying the projection  $\mathbf{P}_S(\mathbf{w}_I)$  and transformation  $\mathbf{T}_{VS}(\varphi, \mathbf{w}_E)$ :

$$\mathbf{P}_{VS}(\varphi, \mathbf{w}) = \mathbf{P}_S(\mathbf{w}_I) \mathbf{T}_{VS}(\varphi, \mathbf{w}_E). \quad (2.5)$$

To retrospectively calibrate the X-ray acquisition system, we thus need to define 12 geometrical parameters  $\mathbf{w}$  of the projection  $\mathbf{P}_{VS}(\varphi, \mathbf{w})$ . For this

purpose we have used centers  $\mathbf{X}_\varphi$  of fiducial markers found in X-ray images and the corresponding centers  $\mathbf{X}_{CT}$  of markers found in CT volume. The optimal calibration parameters  $\mathbf{w}$  are the ones that bring the fiducial markers  $\mathbf{X}_{CT}$  in CT volume to the best correspondence with the corresponding fiducial markers  $\mathbf{X}_\varphi$  in X-ray images. To find the optimal parameters we projected the centers of fiducial markers  $\mathbf{X}_{CT}$  in CT volume to the sensor plane and computed the root mean squared (RMS) distance  $E_{calib}$  to the corresponding centers of fiducial markers  $\mathbf{X}_\varphi$  in X-ray images:

$$E_{calib}(\mathbf{w}) = \sqrt{\frac{1}{M} \sum_{\varphi \in \Phi} \frac{1}{N} \sum_{i=1}^N \left( \mathbf{p}_i^\varphi - \mathbf{P}_{VS}(\varphi, \mathbf{w}) \mathbf{r}_i^{CT} \right)^2}, \quad (2.6)$$

where  $N$  and  $M$  stand for the number of fiducial markers and X-ray images, respectively, and  $\Phi = \{\varphi_1, \varphi_2, \dots, \varphi_M\}$  defines the X-ray images taken at different phantom rotations. To find the optimal calibration parameters  $\mathbf{w}$ , we used nine X-ray images  $\Phi = \{0^\circ, 40^\circ, \dots, 320^\circ\}$  and iterative optimization, which resulted in minimum RMS distance ( $E_{calib}$ ) of 0.31 mm. The small RMS indicates that calibration was performed well and reflects the uncertainty of fiducial marker localization in CT and X-ray images.

## 2.6 Reconstruction of 3D markers

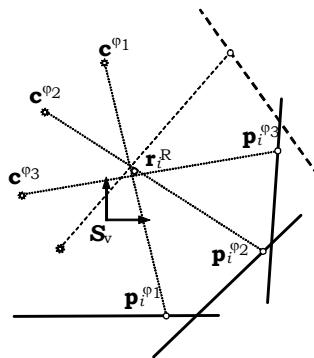
Once the X-ray acquisition system had been calibrated, the positions of X-ray fiducial markers in 3D could be reconstructed from their positions in 2D X-ray images. The least squares solution is used to obtain the 3D marker position (Siddon and Chin, 1985), as described below. Each point  $\mathbf{p}_i^\varphi$ , representing the center of  $i^{\text{th}}$  fiducial marker in X-ray image taken at rotation  $\varphi$  of the phantom, was back-projected to the X-ray source  $\mathbf{r}_s$ , which yielded the imaginary line  $L_i^\varphi$  (Figure 2.4). Line  $L_i^\varphi$ , which defines the perspective projection of a 3D marker to the 2D X-ray plane, can be expressed in the coordinate system  $\mathbf{S}_v$  of the phantom by mapping the X-ray source  $\mathbf{r}_s$  to point  $\mathbf{c}^\varphi$ :

$$\mathbf{c}^\varphi = \mathbf{T}_{VS}^{-1}(\varphi, \mathbf{w}_E) \mathbf{r}_s \quad (2.7)$$

and by expressing the line direction in  $\mathbf{S}_v$  coordinates as:

$$\mathbf{v}_i^\varphi = \frac{\mathbf{T}_{VS}^{-1}(\varphi, \mathbf{w}_E) (\mathbf{p}_i^\varphi - \mathbf{r}_s)}{\left| \mathbf{T}_{VS}^{-1}(\varphi, \mathbf{w}_E) (\mathbf{p}_i^\varphi - \mathbf{r}_s) \right|}, \quad (2.8)$$

where  $\mathbf{r}_s$  and  $\mathbf{p}_i^\varphi$  are points defined in the sensor coordinate system  $\mathbf{S}_s$ . A marker's 3D position  $\mathbf{r}_i^R$  in the coordinate system  $\mathbf{S}_v$  was reconstructed by minimizing the RMS distance  $E_{rec}$  from point  $\mathbf{r}_i^R$  to all lines  $L_i^\varphi$ :



**Figure 2.4.** Reconstruction of 3D marker position. Due to uncertainty of X-ray marker localization and X-ray setup calibration the projection lines are not crossing at the same point.

$$E_{rec}(\mathbf{r}_i^R) = \sqrt{\frac{1}{M} \sum_{\varphi \in \Phi} |(\mathbf{r}_i^R - \mathbf{c}^\varphi) \times \mathbf{v}_i^\varphi|^2}. \quad (2.9)$$

Reconstruction of 3D position of six fiducial markers from the nine X-ray images  $\Phi=\{20^\circ, 60^\circ, \dots, 340^\circ\}$ , which were not used for calibration, yielded RMS of less than 0.06 mm for each of the six fiducial markers. The positions of reconstructed fiducial markers were incorporated in the 3x6 matrix  $\mathbf{X}_R=[\mathbf{r}_1^R, \mathbf{r}_2^R, \dots, \mathbf{r}_6^R]$ .

By using different sets of X-ray images for reconstruction and calibration, we were able to validate the calibration procedure. Small RMS of 0.06 mm indicated that the uncertainty of fiducial marker localization in X-ray images was smaller than in CT images and that calibration had been performed well. Therefore, the major source of calibration uncertainty is the uncertainty of fiducial marker localization in CT images, however, its effect on calibration precision is obviously very small.

## 2.7 “Gold standard” registration and validation

After calibrating the X-ray acquisition system and reconstructing 3D markers  $\mathbf{X}_R$  from X-ray images, we were able to establish “gold standard” registration between X-ray and CT images and between X-ray and MR images in coordinate system  $\mathbf{S}_v$  of the phantom. This was achieved by rigid 3D/3D transformation  $\mathbf{T}$  that minimized the RMS distance  $E_{reg}$  between reconstructed fiducial markers  $\mathbf{X}_R$  from X-ray images and marker points  $\mathbf{X}_{CT}$  from CT or  $\mathbf{X}_{MR}$  from MR images:

$$E_{reg}(\mathbf{T}) = \sqrt{\frac{1}{N} \sum_{i=1}^N (\mathbf{r}_i^R - \mathbf{T} \mathbf{r}_i)^2}, \quad (2.10)$$

where  $\mathbf{r}_i$  stands for either  $\mathbf{r}_i^{CT}$  or  $\mathbf{r}_i^{MR}$ . The closed form solution of this minimal RMS problem is known (Arun *et al.*, 1987; Umeyama, 1991). Rigid transformation  $\mathbf{T}$  can be decomposed to the rotation component  $\mathbf{R}$ , represented by 3x3 matrix, and translation vector  $\mathbf{t}$ :

$$\mathbf{Tr} = \mathbf{R}\mathbf{r} + \mathbf{t}. \quad (2.11)$$

The optimal solution for the translation component is given as:

$$\mathbf{t} = \bar{\mathbf{r}}^R - \mathbf{R}\bar{\mathbf{r}}, \quad (2.12)$$

where  $\bar{\mathbf{r}}^R$  and  $\bar{\mathbf{r}}$  stand for mean positions of point sets  $\mathbf{X}_R$  and  $\mathbf{X}$ , respectively, where set  $\mathbf{X}$  is either  $\mathbf{X}_{CT}$  or  $\mathbf{X}_{MR}$ . The optimal solution for the rotation component is given as:

$$\mathbf{R} = \mathbf{B}\mathbf{A}^T \quad (2.13)$$

$\mathbf{A}$  and  $\mathbf{B}$  are two orthogonal matrices obtained by singular value decomposition (SVD) of the matrix:

$$\bar{\mathbf{X}}_R \bar{\mathbf{X}}^T = \mathbf{A} \mathbf{D} \mathbf{B}^T, \quad (2.14)$$

where  $\mathbf{D}$  is a diagonal matrix and  $\bar{\mathbf{X}}_R$  and  $\bar{\mathbf{X}}$  are the point sets  $\mathbf{X}_R$  and  $\mathbf{X}$ , centered at corresponding mean positions  $\bar{\mathbf{r}}^R$  and  $\bar{\mathbf{r}}$ , respectively.

Rigid registration of point sets ( $\mathbf{X}_{CT}, \mathbf{X}_R$ ) and ( $\mathbf{X}_{MR}, \mathbf{X}_R$ ) resulted in minimum RMS distance  $E_{reg}$  of 0.27 mm for CT and 0.44 mm for MR to X-ray registration. Higher RMS for MR than for CT can be attributed to three reasons. First, CT was used in calibration, second, intra and interslice resolutions of MR images were lower than of CT, which resulted in higher fiducial localization uncertainty, and third, MR images suffered from non-rigid spatial distortion.

The minimum RMS distance  $E_{reg}$  is also known as fiducial registration error (FRE) and can be used to evaluate the accuracy of point based rigid registration (Fitzpatrick *et al.*, 1998). By knowing FRE we can determine the target registration error (TRE), which is the distance between true, but unknown position of the target, and target position obtained by registration. The expected TRE of a target point  $\mathbf{r}$  can be estimated from FRE (Fitzpatrick *et al.*, 1998):

**Table 2.1.** Expected RMS TREs for “gold standard” registration in mm.

	Vertebrae				
	<b>L1</b>	<b>L2</b>	<b>L3</b>	<b>L4</b>	<b>L5</b>
CT	0.2	0.15	0.15	0.19	0.26
MR	0.33	0.24	0.24	0.31	0.42

$$\langle TRE^2(\mathbf{r}) \rangle = \frac{\langle FLE^2 \rangle}{N} \left( 1 + \sum_{k=1}^3 \frac{d_k}{f_k} \right), \quad (2.15)$$

where  $f_k$  is the RMS of the projections of fiducial markers to  $k^{\text{th}}$  principal axis of marker configuration,  $d_k$  is the projection of target point  $\mathbf{r}$  to principle axis  $k$ ,  $N$  is the number of fiducial markers, and FLE is the fiducial localization error obtained from FRE:

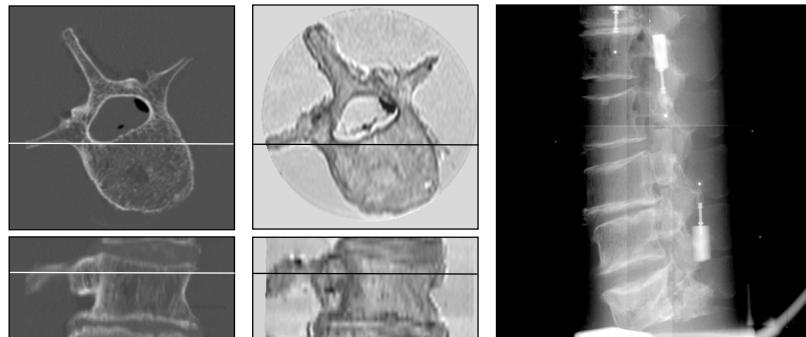
$$FLE^2 = \frac{N}{N-2} FRE^2. \quad (2.16)$$

Using the above formulation, we had validated the “gold standard” registration by manually defining eight target points, four per each pedicle, on each of the 5 vertebra and by computing the mean TRE for each vertebra. The results of “gold standard” validation for CT to X-ray and MR to X-ray registrations are given in Table 2.1. Expected TREs for the pedicles are below 0.26 mm for CT to X-ray registration and below 0.42 mm for MR to X-ray registration.

## 2.8 Publicly available “gold standard” data

The “gold standard” data and detailed information on how to use the data are publicly available from the Department of Electrical Engineering, University of Ljubljana, Laboratory of Imaging Technologies web site<sup>1</sup>. The image database consists of 18 X-ray images, and 5 CT and 5 MR subvolumes (Figure 2.5). In CT and MR images, cubic subvolumes, each containing a single vertebra, approximately one third of neighboring vertebrae, and no fiducial markers, were manually defined. In each subvolume the areas occupied by air and the plastic tube were manually masked and their intensities were replaced by the average intensity corresponding to water. The edges between water, the plastic tube, and air were thus almost eliminated. In this way, any rigid 3D/2D registration method that will be evaluated using our “gold standard” data can not take advantage of markers or edges that would not be present in actual clinical data. The markers in X-ray images, which were not excluded or “airbrushed”, may be considered as outliers, similarly as surgical tools would be in a clinical setting.

<sup>1</sup> <http://lit.fe.uni-lj.si/Downloads/downloads.asp>



**Figure 2.5.** Transversal and sagittal views of CT (left) and MR (middle) subvolumes, and a lateral X-ray image (right).

For each CT or MR subvolume, the "gold standard" registration position, coordinates of eight target points on pedicles, and 450 randomly chosen starting positions around the "gold standard" registration position are also provided in the database. Details about the generation of starting positions can be found in a very recent paper by Tomažević *et al.* (Tomažević *et al.*, 2003). Furthermore, the data on our web site contain the Matlab (MathWorks, Natick, Mass) source code for analyzing registration results. The code allows calculation of TREs and estimation of capturing ranges from the results of 450 registrations per modality and vertebra. If different registration methods are to be objectively compared, it is not only important that same image data sets are used but it is equally important that the evaluation protocol, criteria for successful registrations, and error metrics are the same. Otherwise it is difficult to place the results of a novel registration method in the context of previously published work.

## 2.9 Discussion and conclusion

Before an image guided therapy (IGT) system is put into clinical use, its individual components must undergo rigorous validation. Prerequisites for validation of 3D/2D image registration, which is the crucial part of an IGT system, is standardization of validation methodology, which includes design of validation data sets, definition of corresponding "gold standard" and its accuracy, validation protocol, and design of validation metrics (Tomažević *et al.*, 2003). Besides, a fair comparison of different registration techniques is possible only to a limited degree if standard validation methodology is not publicly available. Motivated by the lack of publicly available "gold standard" data to evaluate and compare different 3D/2D rigid registration methods, we have devised a lumbar spine phantom, obtained X-ray to CT and X-ray to MR "gold standard" registrations and established the accuracy of these registrations. We are aware that surgical interventions and radiosurgery

treatments of the spine exist for which markers are routinely inserted into the spine of the patient. Marker-based registered CT, MR, and X-ray images of the spine can serve as a truly clinical “gold standard” but they are very difficult to obtain. Recently, Russakoff *et al.* (Russakoff *et al.*, 2003a) reported on the evaluation of intensity-based 3D/2D spine image registration using clinical “gold standard data”. Unfortunately, to the best of our knowledge, their data are not publicly available. One problem with such a “gold standard” is that a huge effort has to be spent on removing all traces of the implanted markers from images to avoid bias, i.e. to make sure the evaluated registration method could not take advantage of markers that would usually not be present in images acquired for image-guided therapy. Another approach to obtain more realistic “gold standard” data is to overlay features segmented from clinical images on phantom images (Penney *et al.*, 1998). Soft tissue, surgical instruments, and other structures can be overlaid. Although our “gold standard” is not equivalent to the clinical “gold standard” and no structures have been overlaid on our phantom images, we believe it is realistic enough to allow fair evaluation and comparison of registration methods needed for image-guided therapy. Besides, and we believe this is very valuable, our database contains MR images as well.

The X-ray acquisition system was calibrated retrospectively by matching the projections of CT markers with corresponding markers in X-ray images. Calibration with CT markers is generally superior to calibration with MR markers because CT offers better resolution and spatial stability. This observation was confirmed experimentally, as CT-based calibration yielded smaller calibration error  $E_{calib}$  of 0.31 mm over 0.47 mm obtained by calibrating the X-ray system with MR data. CT-based calibration of the X-ray image acquisition setup already provides registration of CT to X-ray images but does not give any indices of the registration accuracy. Therefore, we have reconstructed the 3D positions of markers from calibrated 2D X-ray images, which allowed us to implement 3D/3D registration between the reconstructed markers and those found in CT and MR volumes. The result of such a registration reflected: a) uncertainty of marker localization in 2D X-ray images, b) uncertainty of marker localization in 3D CT or MR images, c) uncertainty of the X-ray acquisition calibration, and d) uncertainty of marker reconstruction. FRE of 3D/3D registration, which reflected all these uncertainties, was used to evaluate TRE of the “gold standard” CT to X-ray and MR to X-ray registration by the theory developed in (Fitzpatrick *et al.*, 1998).

The results in Table 2.1 indicate that “gold standard” registration is highly accurate and therefore useful for testing 3D/2D registration methods. However, it should be stressed that the expected TREs for CT to X-ray “gold standard” registration may possibly be a little larger than those presented in Table 2.1. This is because the same CT markers were used for X-ray system calibration and for CT to X-ray registration, which could have involved the same bias in calibration and registration. By acquiring a second CT scan at the time of image acquisitions and using one for calibration and the other for validation, bias could be eliminated. Nevertheless, if we assume that

localization errors for CT markers are much smaller than for MR markers, the expected TREs for CT to X-ray "gold standard" registration should be close to those given in Table 2.1 and are certainly not larger than TREs for MR to X-ray registration. We believe that, especially due to the lack of publicly available "gold standards" for 3D/2D rigid registration, the presented "gold standard" data will prove useful for evaluation of newly developed methods and comparison of different existing registration methods.

The incredible thing about the human  
mind is that it didn't come with an  
instruction book.

*Tony Riley*

## Chapter 3

# 3D/2D registration of CT and MR to X-ray images

Dejan Tomažević, Boštjan Likar, Tomaž Slivnik and Franjo Pernuš  
IEEE Transactions on Medical Imaging 22: 1407-1416 (2003)

### Abstract

*A crucial part of image-guided therapy is registration of preoperative and intraoperative images, by which the precise position and orientation of the patient's anatomy is determined in three dimensions. This paper presents a novel approach to register three-dimensional (3D) computed tomography (CT) or magnetic resonance (MR) images to one or more two-dimensional (2D) X-ray images. The registration is based solely on the information present in 2D and 3D images. It does not require fiducial markers, intraoperative X-ray image segmentation, or timely construction of digitally reconstructed radiographs. The originality of the approach is in using normals to bone surfaces, preoperatively defined in 3D MR or CT data, and gradients of intraoperative X-ray images at locations defined by the X-ray source and 3D surface points. The registration is concerned with finding the rigid transformation of a CT or MR volume, which provides the best match between surface normals and back projected gradients, considering their amplitudes and orientations. We have thoroughly validated our registration method by using MR, CT, and X-ray images of a cadaveric lumbar spine phantom for which "gold standard" registration was established by means of fiducial markers, and its accuracy assessed by target registration error. Volumes of interest, containing single vertebrae L1-L5, were registered to different pairs of X-ray images from different starting positions, chosen randomly and uniformly around the "gold standard" position. CT/X-ray (MR/X-ray) registration, which is fast, was successful in more than 91% (82% except*

for L1) of trials if started from the “gold standard” translated or rotated for less than 6 mm or 17° (3 mm or 8.6°), respectively. Root-mean-square target registration errors were below 0.5 mm for the CT to X-ray registration and below 1.4 mm for MR to X-ray registration.

### 3.1 Introduction

Medical imaging is currently undergoing rapid development with a strong emphasis being placed on the use of imaging technology to render surgical and therapeutic procedures less and less invasive and to improve the accuracy with which a given procedure can be performed compared with conventional methods (Galloway, 2001; Taylor *et al.*, 1995). In image-guided therapy (IGT) (surgery, radiotherapy, or radiological intervention), preoperative medical data, usually three-dimensional (3D) computed tomography (CT) or magnetic resonance (MR) images, are used to diagnose, plan, simulate, guide, or otherwise assist a surgeon, or possibly a robot, in performing a surgical or therapeutic procedure. The plan is constructed in the coordinate system relative to preoperative data, while the surgical procedure is performed in the coordinate system relative to the patient. The relationship or spatial transformation between preoperative data and plan and physical space occupied by the patient during treatment is established by registration. Registration, which is the crucial part of IGT, allows any 3D point defined in the preoperative image to be precisely located in a patient coordinate system and may thus provide the surgeon information about position of his instruments relative to the planned trajectory, nearby vulnerable structures, or the ultimate target. However, to be suitable for a clinical application, a registration algorithm must satisfy several requirements. These requirements concern registration accuracy, computation time, robustness to outliers, e.g. produced by surgical instruments, and complexity and invasiveness of intraoperative data acquisition.

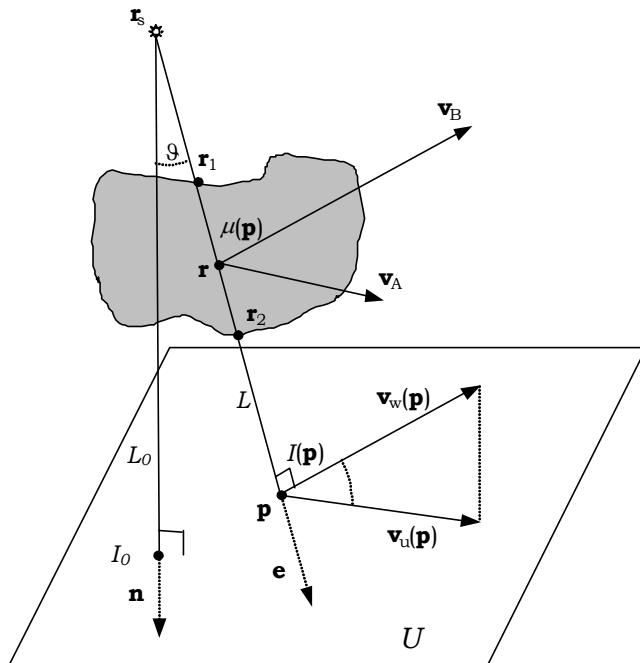
A variety of methods have been proposed in the last decade for image to physical space (patient) registration. The approaches differ in the information on which registration relies. The most accurate methods utilize points defined by fiducial markers that are preoperatively rigidly affixed to the patient’s anatomy, most often to bone (Jones *et al.*, 1993; Maurer *et al.*, 1997; Taylor *et al.*, 1994). Intraoperatively, the locations of markers, which can easily be detected in 3D CT or MR images, are measured using a positioning device. Point-based registration is achieved by finding the rigid transformation that brings the fiducial points in the two spaces into alignment. Although these methods are fast, accurate, and robust, they require fixation of fiducial markers to rigid structures, which is not always possible or is too invasive to be acceptable. The advantage of fiducial marker-based approaches is that registration error depends only on fiducial localization and is, thus, to a large extent independent of the particular structure being registered (Fitzpatrick and West, 2001). Because in both spaces fiducial marker localization may be very accurate, registration by this method is often used as reference or “gold

standard”, suitable for evaluating other registration methods (Guéziec *et al.*, 1998; Herring *et al.*, 1998; Penney *et al.*, 1998).

Another class of registration methods is based on surfaces (Colchester *et al.*, 1996; Grimson *et al.*, 1996; Herring *et al.*, 1998; Lavallée *et al.*, 1995; Lavallée and Szeliski, 1995; Maurer *et al.*, 1998; Simon *et al.*, 1995b). These methods rely on a description of the shape of an anatomical structure, usually skin or outer bone surface, which is extracted from the preoperative 3D image. Intraoperative modalities, used to acquire a similar representation of the shape of the same structure include digitizing probes, passive stereo, and structured light. A registration process, minimizing points-to-surface or surface-to-surface distances, is then used to derive the spatial transformation (Besl and McKay, 1992). The problems with these techniques are that skin is not a rigid structure and that bone surface, containing enough shape information, has to be exposed during the procedure.

The third class of registration methods relies on intraoperative X-ray projections acquired with a calibrated X-ray device. The location and orientation of a structure in 3D CT or MR image with respect to the geometry of the X-ray device is determined by 3D/2D registration. Two different 3D/2D registration approaches have been proposed. The first approach, which is a feature-based one, requires extracting bone surfaces from preoperative data and contours of the same structure from one or more intraoperative projection images (Guéziec *et al.*, 1998; Lavallée and Szeliski, 1995). Registration is performed by minimizing the distance between the surface model and the lines connecting contour points with the X-ray source. Once segmentation has been undertaken, the reduced amount of data makes the registration fast. Unfortunately, the intraoperative segmentation is hard to achieve automatically and errors in segmentation lead to errors in registration. The second approach to 3D/2D registration is based solely on voxel and pixel intensity information of 3D preoperative and 2D intraoperative image, respectively. Simulated X-ray projection images (Goitein *et al.*, 1983), called digitally reconstructed radiographs (DRRs), are produced from a CT image and the unknown pose of the CT volume relative to the X-ray image set is estimated by optimizing the similarity measure calculated from the DRR and X-ray image (Lemieux *et al.*, 1994; Penney *et al.*, 1998). These algorithms require little or no segmentation but are time consuming. This problem may be reduced to some extent by calculating DRRs containing only structures of interest (Weese *et al.*, 1997) or by implementing a faster DRR generation method (LaRose *et al.*, 2000a; Weese *et al.*, 1999). Nevertheless, these approaches are not suitable for registration of a preoperative MR image to X-ray images, because there is practically no correlation between MR-based DRRs and X-ray projections.

In this paper we present a novel method for registering 3D CT or MR images to 2D X-ray images with the final goal to estimate the position and orientation of the patient’s anatomy during surgery or therapy. The proposed method combines the advantages of both 3D/2D approaches, i.e. the reduced amount



**Figure 3.1.** X-ray image generation – geometrical considerations.

of data of the first approach, resulting in faster registration, and no need for intraoperative segmentation of the second approach. Like in the first 3D/2D approach, the method assumes that surfaces of bony structures are extracted preoperatively from CT or MR images. The method uses a novel criterion function (CF) to measure the match between normals to the surface and corresponding back-projected intensity gradients of X-ray images. Such matching avoids intraoperative segmentation of X-ray images and thus eliminates misregistration induced by intraoperative segmentation errors. We use rigid registration, which is appropriate when the imaged object is itself rigid, like vertebrae, femur, pelvis, other bones, and head, which may be considered rigid for many diagnostic and therapeutic purposes. With respect to the application, we focus on registration of lumbar vertebrae and provide results that can be used in the context of pedicle screw placement.

A necessary step, required before widespread clinical use of any novel registration technique, is the evaluation and validation of the method (Jannin *et al.*, 2002). One difficulty in evaluating a registration technique is the need for a highly accurate “gold standard”. Recently, we have devised a cadaveric lumbar spine phantom, acquired 3D MR and CT images and 2D X-ray images of the phantom, and established accurate “gold standard” rigid registration between 3D and 2D images by means of fiducial markers (Tomažević *et al.*,

2002). We have used these data to study the capturing range and registration accuracy of the proposed method.

## 3.2 Theory

In this section, we first formulate the X-ray image generation process and then give the relationship between X-ray image intensity gradients and attenuation coefficient gradients, which is later used in our registration method.

### 3.2.1 X-Ray image generation

The following model describes the generation of an X-ray image:

$$I(\mathbf{p}) = I_0 \cos^3 \theta e^{-\int_L^\mathbf{p} \mu(\mathbf{r}) dr}. \quad (3.1)$$

Image intensity  $I(\mathbf{p})$  at point  $\mathbf{p}$ , lying on detector plane  $U$ , is obtained by integrating the X-ray attenuation coefficient  $\mu(\mathbf{r})$ , from X-ray source  $\mathbf{r}_s$  to point  $\mathbf{p}$ , along the projection beam  $L$  (Figure 3.1). The factor  $\cos^3 \theta$  describes the effect of X-ray beam divergence. Angle  $\theta$  is the angle between a projection beam and line  $L_0$ , which originates in X-ray source and is perpendicular to detector plane.  $I_0$  is the reference image intensity at point where line  $L_0$  intersects the detector plane. If the X-ray imaging system corrects for beam divergence, a simpler model is obtained:

$$I(\mathbf{p}) = I_0 e^{-\int_L^\mathbf{p} \mu(\mathbf{r}) dr}. \quad (3.2)$$

For the X-ray sensor with a logarithmic static response, the relationship between intensity  $I(\mathbf{p})$  and attenuation coefficients  $\mu(\mathbf{r})$  can be expressed by a linear model:

$$I(\mathbf{p}) = a \int_L^\mathbf{p} \mu(\mathbf{r}) dr + b \quad (3.3)$$

or even simpler:

$$I(\mathbf{p}) = \int_L^\mathbf{p} \mu(\mathbf{r}) dr. \quad (3.4)$$

### 3.2.2 Image intensity and attenuation coefficient gradients

We use the X-ray image generation model defined in (3.4) to derive the relationship between gradients of attenuation coefficients and image intensity gradients. Let  $\mathbf{v}_u(\mathbf{p})$ ,  $\mathbf{v}_u(\mathbf{p}) = \text{grad}_U I(\mathbf{p})$ , be the image intensity gradient in

projection plane  $U$  and let  $\mathbf{v}_w(\mathbf{p})$ ,  $\mathbf{v}_w(\mathbf{p}) = \text{grad}_W I(\mathbf{p})$ , be the intensity gradient in plane  $W$ , which is perpendicular to projection beam  $L$  (Figure 3.1). Gradient  $\mathbf{v}_u(\mathbf{p})$  can be expressed by gradient  $\mathbf{v}_w(\mathbf{p})$  and vice versa (Appendix A):

$$\mathbf{v}_u(\mathbf{p}) = (\mathbf{n} \times \mathbf{v}_w(\mathbf{p})) \times \mathbf{n}, \quad (3.5)$$

$$\mathbf{v}_w(\mathbf{p}) = \frac{(\mathbf{n} \times \mathbf{v}_u(\mathbf{p})) \times \mathbf{e}}{\mathbf{n} \cdot \mathbf{e}}, \quad (3.6)$$

where  $\mathbf{n}$  is the unit normal to projection plane  $U$  and  $\mathbf{e}$  a unit vector defining the direction of projection beam  $L$ .

The attenuation coefficient gradient  $\mathbf{v}_A(\mathbf{r})$ ,  $\mathbf{v}_A(\mathbf{r}) = \text{grad } \mu(\mathbf{r})$ , can be divided into two components as (Appendix B):

$$\mathbf{v}_A(\mathbf{r}) = \text{grad } \mu(\mathbf{r}) = \mathbf{v}_{Ar}(\mathbf{r}) + \mathbf{v}_{At}(\mathbf{r}). \quad (3.7)$$

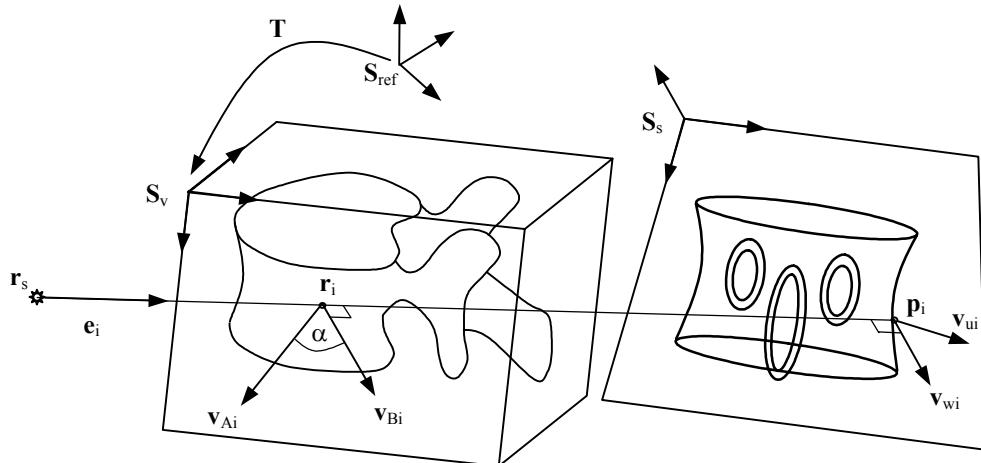
The first component  $\mathbf{v}_{Ar}(\mathbf{r})$  has the same direction as the projection beam  $L$ , while the second component  $\mathbf{v}_{At}(\mathbf{r})$  is perpendicular to the beam. The relationship between intensity gradient  $\mathbf{v}_w(\mathbf{p})$  and gradient  $\mathbf{v}_{At}(\mathbf{r})$  is given by (Appendix C):

$$\mathbf{v}_w(\mathbf{p}) = \frac{1}{|\mathbf{p} - \mathbf{r}_s|} \int_L |\mathbf{r} - \mathbf{r}_s| \mathbf{v}_{At}(\mathbf{r}) d\mathbf{r}, \quad (3.8)$$

where  $|\mathbf{r} - \mathbf{r}_s|$  and  $|\mathbf{p} - \mathbf{r}_s|$  are the distances of X-ray source  $\mathbf{r}_s$  to points  $\mathbf{r}$  and  $\mathbf{p}$ , respectively. This relation shows that intensity gradient  $\mathbf{v}_w(\mathbf{p})$  depends only on attenuation coefficient gradient component  $\mathbf{v}_{At}(\mathbf{r})$ , which is perpendicular to the projection beam  $L$ . The obtained relation is in accordance with the known fact that X-ray images depict structure changes perpendicular to the projection beam and not structure changes along the beam. From (3.8) it is also obvious that if point  $\mathbf{r}$  is closer to the X-ray source the component  $\mathbf{v}_{At}(\mathbf{r})$  has less impact on  $\mathbf{v}_w(\mathbf{p})$  and, consequently, a smaller X-ray image intensity gradient  $\mathbf{v}_u(\mathbf{p})$  is obtained. Knowing how the attenuation coefficient gradient  $\mathbf{v}_A(\mathbf{r})$  is projected to point  $\mathbf{p}$  (3.8), we can back-project the gradient  $\mathbf{v}_w(\mathbf{p})$  towards the X-ray source  $\mathbf{r}_s$  and obtain the back-projected gradient  $\mathbf{v}_B(\mathbf{r})$  for any location  $\mathbf{r}$  along the beam:

$$\mathbf{v}_B(\mathbf{r}) = \frac{|\mathbf{p} - \mathbf{r}_s|}{|\mathbf{r} - \mathbf{r}_s|} \mathbf{v}_w(\mathbf{p}). \quad (3.9)$$

Gradient  $\mathbf{v}_B(\mathbf{r})$  is the enlarged gradient  $\mathbf{v}_w(\mathbf{p})$ . The closer the point  $\mathbf{r}$  to the X-ray source  $\mathbf{r}_s$  the greater the enlargement. By inserting (3.6) into (3.9) we obtain the expression by which the image intensity gradient  $\mathbf{v}_u(\mathbf{p})$  is back-projected towards the X-ray source:



**Figure 3.2.** Geometrical setup with normals to surfaces and intensity gradients used in 3D/2D registration.

$$\mathbf{v}_B(\mathbf{r}) = \frac{|\mathbf{p} - \mathbf{r}_s|}{|\mathbf{r} - \mathbf{r}_s|} \cdot \frac{(\mathbf{n} \times \mathbf{v}_u(\mathbf{p})) \times \mathbf{e}}{\mathbf{n} \cdot \mathbf{e}}. \quad (3.10)$$

### 3.3 Method

Rigid 3D/2D registration is concerned with finding the rigid transformation that brings a 3D (CT or MR) image into the best possible spatial relation with the corresponding 2D projection (X-ray) image. Registration objectives may be met if the following two assumptions are fulfilled. First, the X-ray imaging device is calibrated, i.e. for every X-ray image the precise position of X-ray source and position and orientation of detector plane are known in a reference coordinate system, and second, CT, MR, and X-ray images are relatively accurate geometrically. In a clinical setting, calibration could be performed as described in (Livyatan *et al.*, 2002). Features that we use for rigid 3D/2D registration, are normals to surfaces of bony structures found in preoperative CT or MR volumes and intensity gradients of intraoperative X-ray images. Feature selection is founded on the assumption that strong intensity gradients in X-ray images correspond to boundaries of bony structures.

Let  $\mathbf{r}_i^{S_v}$  be a point, defined in the coordinate system  $\mathbf{S}_v$  of a 3D volume (CT or MR), lying on the surface of a 3D structure and let  $\mathbf{v}_{Ai}$  be the vector, normal to the surface at point  $\mathbf{r}_i^{S_v}$  (Figure 3.2). The position of the same point in reference coordinate system  $\mathbf{S}_{ref}$  is given by the rigid transformation  $\mathbf{T}$ :

$$\mathbf{r}_i = \mathbf{T}(\mathbf{r}_i^{S_v}) = \mathbf{R}\mathbf{r}_i^{S_v} + \mathbf{t}, \quad (3.11)$$

where  $\mathbf{R}$  and  $\mathbf{t}$  describe, respectively, the rotation and translation of coordinate system  $\mathbf{S}_v$  with respect to  $\mathbf{S}_{\text{ref}}$ . Matrix  $\mathbf{R}$  is defined by three angles  $\omega=(\omega_x, \omega_y, \omega_z)^T$  representing rotations around  $x$ ,  $y$  and  $z$  axes, while vector  $\mathbf{t}=(t_x, t_y, t_z)^T$  describes translations along  $x$ ,  $y$ , and  $z$  axes. Transformation  $\mathbf{T}$  is thus defined by six parameters  $\mathbf{q}=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)^T$ . Let  $\mathbf{r}_s$  be the position of the X-ray source in the reference coordinate system  $\mathbf{S}_{\text{ref}}$ . For a given position of a 3D image, defined by vector  $\mathbf{q}$ , line  $L_i$ , which connects  $\mathbf{r}_i$  and  $\mathbf{r}_s$  and has direction defined by unit vector  $\mathbf{e}_i$ , intersects the detector plane  $U$  at point  $\mathbf{p}=p(\mathbf{r}_i)$  (Figure 3.2). The intensity gradient at  $\mathbf{p}_i$  is  $\mathbf{v}_{ui}$ ,  $\mathbf{v}_{ui}=\text{grad}_U I(\mathbf{p}_i)$ .

To measure the correspondence between surface normals  $\mathbf{v}_{Ai}$  and back-projected gradients  $\mathbf{v}_{Bi}$  (3.10) we propose the following criterion function  $CF$ :

$$CF = \frac{\sum_{i=1}^N |\mathbf{v}_{Ai}| \cdot |\mathbf{v}_{Bi}| \cdot f(\alpha_i)}{\sum_{i=1}^N |\mathbf{v}_{Ai}| \cdot \sum_{i=1}^N |\mathbf{v}_{Bi}|}, \quad (3.12)$$

where  $f(\alpha_i)$  is the weighting function depending on angle  $\alpha_i$  between gradients  $\mathbf{v}_{Ai}$  and  $\mathbf{v}_{Bi}$ , and  $N$  is the number of surface points.

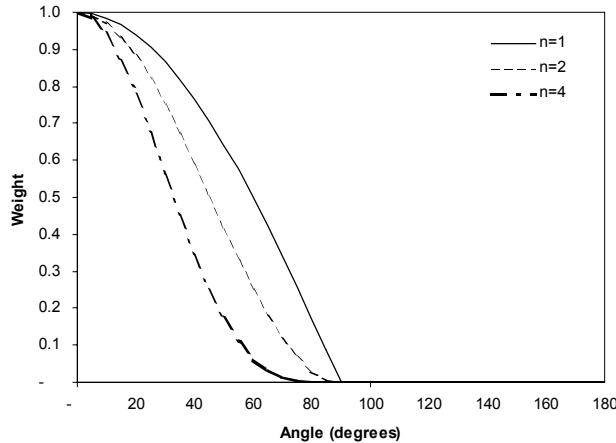
The function  $f(\alpha_i)$ , defined as:

$$f(\alpha_i) = \begin{cases} \cos^n \alpha_i & 0 < |\alpha_i| < 90^\circ \\ 0 & \text{otherwise} \end{cases} \quad (3.13)$$

favors gradients that have similar directions. Parameter  $n$  determines the sensitivity of weighting to angle  $\alpha_i$  (Figure 3.3). In practice, the weighting function, for  $\alpha_i$  smaller than  $90^\circ$ , is calculated as:

$$f(\alpha_i) = \cos^n \alpha_i = \frac{|\mathbf{v}_{Ai} \cdot \mathbf{v}_{Bi}|^n}{(|\mathbf{v}_{Ai}| \cdot |\mathbf{v}_{Bi}|)^n}. \quad (3.14)$$

When more than one X-ray image is used for registration, the criterion function is a sum of criteria functions of individual images. Rigid 3D/2D registration is performed by finding the set of parameters  $\mathbf{q}$  that optimize the criterion function.

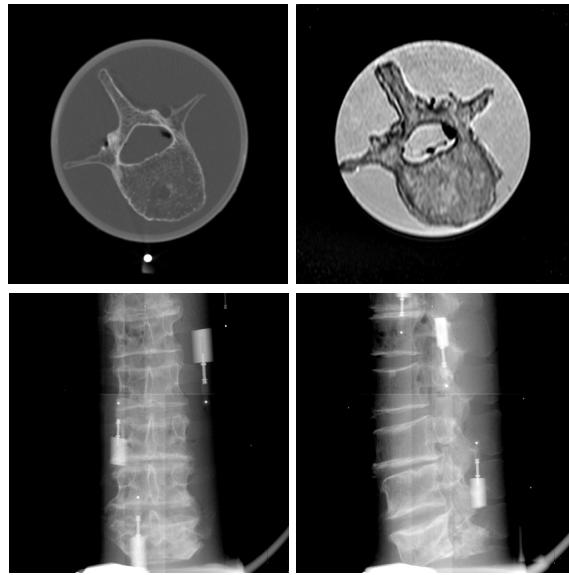


**Figure 3.3.** Weighting function  $f(\alpha_i)$  for different values of parameter  $n$ .

## 3.4 Evaluation

### 3.4.1 “Gold standard”

Because it is extremely difficult to obtain an accurate “gold standard” registration on clinical images, we have obtained a “gold standard” on a section of cadaveric lumbar spine, comprising vertebra L1-L5 with intervertebral disks and several millimeters of soft tissue, of an 80 year-old female. The spine was placed into a plastic tube and tied to it with thin nylon strings. Six fiducial markers were rigidly attached to the outside of the tube. The tube was filled with water to simulate soft tissue. In this way, realistic MR, CT, and X-ray images were obtained. The CT image was obtained using a General Electric HiSpeed CT/i scanner. Axial slices were taken with intraslice resolution of  $0.27 \times 0.27$  mm and 1 mm interslice distance. For magnetic resonance imaging (MRI), Philips Gyroscan NT Intera 1.5 T scanner and T1 protocol (flip angle 90°, TR=3220 ms, TE=11 ms) was used. Axial slices were obtained with  $0.39 \times 0.39$  mm intraslice resolution and 1.9 mm between-slice distance. The MR image was retrospectively corrected for intensity inhomogeneity by the information minimization method (Likar *et al.*, 2001). X-ray images were obtained by PIXIUM 4600 (Trixell) digital X-ray detector with a  $429 \times 429$  mm large active surface,  $0.143 \times 0.143$  mm pixel size and 14 bits of dynamic range. The X-ray source and detector-plane were fixed during image acquisition while the spine phantom was rotated on a turntable to simulate a setup with C-arm. By rotating (step=20°) the spine phantom around its long axis 18 X-ray images were acquired. CT, MR, and X-ray images are given in Figure 3.4.



**Figure 3.4.** Axial slices from a CT (top left) and MR (top right) image and anterior-posterior (bottom left) and lateral (bottom right) X-ray image of the spine phantom with fiducial markers.

“Gold standard” registrations of CT to X-ray and MR to X-ray images were obtained by rigid registration of CT and MR marker points to 3D marker points reconstructed from X-ray images, respectively. The accuracy of “gold standard” registration was obtained by estimating the target registration error TRE (Fitzpatrick *et al.*, 1998) for eight targets, four per each pedicle, which were manually defined on each of the 5 vertebrae. All root mean square (rms) TREs were below 0.29 mm for CT to X-ray registration and below 0.42 mm for MR to X-ray registration, which indicates that the “gold standard” is highly accurate. A detailed description of the spine phantom, X-ray system calibration, calculation of “gold standard” registration and “gold standard” validation can be found in (Tomažević *et al.*, 2002).

### 3.4.2 Implementation details

Cubic subvolumes that contained single vertebrae were defined manually in CT and MR images. Each subvolume was blurred using a Gaussian filter ( $\sigma = 0.5$  mm) and isotropically re-sampled to the resolution of 1 mm. The Canny edge detector and a threshold were applied to automatically extract locations of points corresponding to boundaries of bony structures and to estimate surface normal directions at these points. This resulted in approximately 19000 edge points per CT and 32000 points per MR image. The X-ray images were blurred with a Gaussian filter ( $\sigma = 0.5$  mm) and Roberts edge detector

was applied to calculate intensity gradients  $\text{grad}_{\text{ul}}(\mathbf{p})$ . Gradients perpendicular to the projection beam  $\text{grad}_{\text{wl}}(\mathbf{p})$  were then obtained using (3.6). Parameter  $n$  (3.14), needed to calculate the criterion function, was set to 4. Optimization of transformation parameters  $\mathbf{q} = (t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)^T$  was performed by Powell's method (Press *et al.*, 1992).

### 3.4.3 Experiments

To test the accuracy, speed, and capturing range of the method, registration was performed from a wide range of starting positions and orientations around the "gold standard" registration position. If the parametrical space is normalized, Euclidean metrics may be used to calculate the displacement (in parametrical space) of a starting position from "gold standard" position. Besides, the values of parameters defining the starting point may be chosen randomly around the "gold standard" position in such a way that a uniform distribution of displacements in the interval  $[0, D_e]$  is achieved. For single vertebrae registration the following assumption was used to normalize the parametrical space: rotation of a volume containing a single vertebra of size 80 mm around its center of 0.1 radians (rad) ( $5.7^\circ$ ) causes mean translation of volume points of 2 mm.

Eighteen X-ray image pairs were chosen from the set of 18 X-ray images, such that the rotation angle between two images forming a pair was  $80^\circ$ . For each pair, registration was run from 25 starting positions generated as described above, setting  $D_e$  to 18mm ( $51.6^\circ$ ) and to 9mm ( $25.8^\circ$ ) for CT and MR registration, respectively. This resulted in 450 registrations per modality and vertebra. By using a smaller value of  $D_e$  for MR, we were able to more precisely define the capturing ranges for MR/X-ray registrations, which were expected to be smaller than the capturing ranges for CT/X-ray registrations.

To measure the registration error before and after registration, TRE was calculated for eight target points (four on each pedicle) as the distance between target points in registered and "gold standard" position:

$$\text{TRE}(\mathbf{r}) = \|\mathbf{T}_r \mathbf{r} - \mathbf{T}_g \mathbf{r}\|, \quad (3.15)$$

where  $\mathbf{r}$  is a target point, and  $\mathbf{T}_r$  and  $\mathbf{T}_g$  are the transformations obtained by the proposed registration method and "gold standard" registration, respectively. The rotation error  $\gamma$  was determined by decomposing TRE to translation and rotation components (West *et al.*, 1997). A registration was treated as successful if TREs of all 8 target points were below 2 mm.

### 3.4.4 Results

Results of CT/X-ray and MR/X-ray registrations are presented in Table 3.1 and Table 3.2, respectively. Both tables give the root mean squares and

**Table 3.1.** The RMS and maximum target registration and rotation component errors before and after CT/X-ray registration and the proportion of successful registrations.

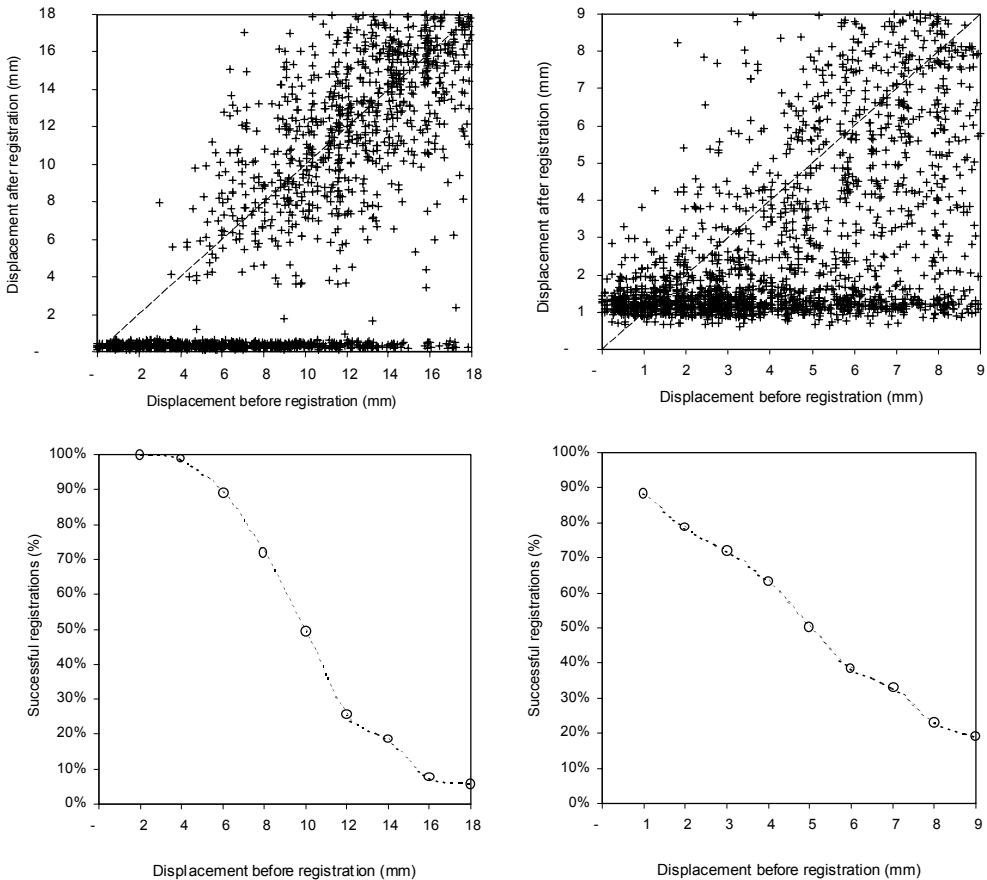
	Before registration				After registration				Successful registrations (%)		
	Rms (TRE) (mm)	Max (TRE) (°)	Rms (γ) (mm)	Rax (γ) (°)	Rms (TRE) (mm)	Max (TRE) (mm)	Rms (γ) (mm)	Rax (γ) (°)	0÷6mm 0÷17.2°	6÷12mm 17.2÷34.4°	12÷18mm 34.4÷51.7°
<b>L1</b>	6.0	21.7	13.7	51.6	0.36	0.70	0.35	0.86	98.6	55.0	10.9
<b>L2</b>	5.5	19.9	11.5	34.2	0.30	1.17	0.48	3.22	93.3	45.0	2.7
<b>L3</b>	5.8	21.1	13.2	35.4	0.22	0.62	0.37	1.23	91.3	56.3	6.8
<b>L4</b>	6.1	23.7	12.7	41.1	0.24	0.44	0.32	0.50	94.0	42.4	6.8
<b>L5</b>	6.6	25.7	14.2	52.0	0.49	1.05	0.34	1.92	96.7	42.0	12.9

**Table 3.2.** The RMS and maximum target registration and rotation component errors before and after MR/X-ray registration and the proportion of successful registrations.

	Before registration				After registration				Successful registrations (%)		
	Rms (TRE) (mm)	Max (TRE) (°)	Rms (γ) (mm)	Rax (γ) (°)	Rms (TRE) (mm)	Max (TRE) (mm)	Rms (γ) (mm)	Rax (γ) (°)	0÷3mm 0÷8.6°	3÷6mm 8.6÷17.2°	6÷9mm 17.2÷25.8°
<b>L1</b>	3.7	11.0	8.1	18.7	1.38	2.00	1.59	3.31	53.6	37.1	19.1
<b>L2</b>	3.5	10.2	7.4	25.8	1.35	2.00	1.60	3.50	82.4	58.0	16.9
<b>L3</b>	3.6	12.1	8.6	24.9	1.16	1.98	1.48	4.25	83.0	48.9	29.3
<b>L4</b>	3.6	10.8	7.9	25.8	1.07	1.67	1.12	2.07	89.4	48.2	22.7
<b>L5</b>	4.1	12.5	8.5	24.3	1.23	2.00	0.95	1.98	85.8	59.7	36.6

maximum values of TRE and rotation  $\gamma$  before and after registration. Only results of successful registrations are given. The tables also give the proportion of successful registrations for different intervals of displacements. Displacements (in parametrical space) before and after registration for 2250 registrations (450 registrations/5 vertebrae) are illustrated in Figure 3.5., which also show the relationship between initial displacement and rate of successful registrations.

Successful CT to X-ray registrations of single vertebrae show (Table 3.1), that the root mean squares of TREs range from 0.2 mm for L3 to 0.5 mm for L5, and that the maximal TREs are less than 1.2 mm. The rms angle error ranges from 0.35° for vertebra L4 to 0.48° for L2, with a maximum of 3.2° obtained for L2. More than 91% of registrations were successful if the starting position



**Figure 3.5.** Scatter diagrams of displacements before versus after registration (top) and the effect of initial displacement on proportion of successful registrations (bottom). Results for CT/X-ray registrations are given on the left and for MR/X-ray registrations on the right.

had not been more than 6 mm or 17.2° from the “gold standard” position. If the starting positions had been in the interval of 6÷12 mm (17.2÷34.4°) the registration succeeded in 42-56% cases, while less than 10% registrations were successful if the initial displacement had been in the interval of 12÷18 mm (34.4÷51.7°).

Successful MR to X-ray registrations of single vertebra (Table 3.2) show that root mean squares of TREs vary from 1.1 to 1.4 mm. The rms angle errors range from 0.95° to 1.6° and maxima are less than 4.5°. A successful registration was reached in more than 82% of cases if the displacement  $d_e$  had been between 0 and 3 mm (0° and 8.6°), except for vertebra L1 for which only

53% of registrations were successful. For the other two intervals of initial displacement, 3–6 mm ( $8.6^\circ$ – $17.2^\circ$ ) and 6–9 mm ( $17.2^\circ$ – $25.8^\circ$ ), registration succeeded in 37–59% and 19–36% of cases, respectively.

The average registration time on a Pentium III, 733 MHz PC was 20 seconds for X-ray to CT and 32 seconds for X-ray to MR registration.

### 3.5 Discussion and conclusion

If image-based registration is to be used for guiding a surgeon or surgical robot, accuracy requirements are particularly stringent. The accuracy with which a surgeon can achieve a preoperatively constructed plan is directly related to the accuracy with which registration can be performed. In general, the necessary accuracy varies significantly, depending on the type of the procedure and the patient's anatomy. Because we have used a spine phantom and because targets have been chosen on both pedicles, the accuracy of our method may be best judged in the context of spinal pedicle screw placement. Recently, the margin of error for pedicle screw placement was defined as the allowable translational and rotational deviation of the screw trajectory from an idealized path along the central axis of the pedicle (Rampersaud *et al.*, 2001). Rampersaud *et al.* (Rampersaud *et al.*, 2001) reported the following maximum permissible translational (rotational) errors: 0.65 mm ( $2.1^\circ$ ), 0.85 mm ( $2.7^\circ$ ), 1.5 mm ( $4.9^\circ$ ), 3 mm ( $9.8^\circ$ ), and 3.8 mm ( $12^\circ$ ) for vertebrae L1–L5, respectively. The proposed method is quite robust to outliers, i.e. edges not corresponding to bony surfaces, because a large number of edge points add to the criterion function. Registration results in Table 3.1, given for a variety of X-ray views and starting positions, indicate that the registration errors, although defined in a slightly different way as in (Rampersaud *et al.*, 2001) are within error margins given above. Registration was successful in more than 91% of trials if registration was performed from a starting position, which was the "gold standard" registration position translated or rotated for less than 6 mm or  $17^\circ$ , respectively. This capturing range is wide enough to be reached by a simple manual registration technique. In a clinical system using the present registration method, a simple visualization tool, by which bony edge points in CT or MR are superimposed on X-ray images, may help an operator in the process of initial registration as well as in estimation of registration accuracy. Whenever the proposed 3D/2D registration was not successful, the pose obtained after registration was far enough from the correct pose, so that all these cases could easily be detected using visual verification.

Importantly, the proposed registration method allows registration of MR to X-ray images. However, the registration error was larger than for CT to X-ray registration, 1.4 mm and  $1.6^\circ$  for registrations that were successful, which is comparable with the accuracy required for vertebrae L3–L5 but not for vertebrae L1 and L2 (Rampersaud *et al.*, 2001). Besides, fewer registrations were successful and in some cases the registration was unsuccessful even if the starting pose was close to the "gold standard". This was most probably

due to the quite porous spine immersed in water and the MRI sequence, which did not produce clear tissue contrast between the bone and surrounding tissue of the spine. We found that approximately 40% of edge points extracted from the MR image did not correspond to boundaries of bony structures. There are more possible solutions to this problem. One solution is to use a different MRI sequence, like the double echo fast gradient echo sequence recently proposed by Hoad *et al.* (Hoad *et al.*, 2001), to obtain good tissue contrast between bone and soft tissue. Another solution to this problem would be to reduce the number of edges that do not correspond to bony structures by some postprocessing of the MR edge image. As most registration methods, this method also gets trapped in local minima. To limit this behavior, a global instead a local optimization procedure could be applied (Guéziec *et al.*, 2000), however, this requires intraoperative time, which is constrained. Using the visual inspection, we have observed that it is much more difficult to assess the quality of MR to X-ray registration than the quality of CT to X-ray registrations. Extra edge points projected onto the X-ray image make visual verification harder and less precise.

The value of parameter  $n$  in the weighting function (3.14), affects smoothness of the criterion function and final accuracy, but the relations are little more complicated. We have observed that small  $n$  ( $n=0,1,2$ ) results in a criterion function which is smooth near the correct registration (closer than 0.5 mm and 0.5°) but contains more local optima further from the correct registration. When small values of  $n$  were used, up to 0.5 mm larger CT/X-ray registration errors were obtained in comparison to the errors obtained if  $n$  was set to 4. If  $n$  is larger than 5, only the points where the normal to the surface is almost perpendicular to the ray will add to the criterion function, which will result in a more ragged criterion function. After experimenting with different values of  $n$ , we have set  $n$  to 4. Such a choice was a trade off between smoothness of the criterion function close to and further away from the correct registration.

Before an image-guided therapy (IGT) system is put into clinical use, it must undergo rigorous validation, which includes validation of individual components, validation of the overall system and a study of how uncertainties propagate through the IGT process. Prerequisites for validation of an IGT system component, like 3D/2D registration is standardization of the validation methodology, which should include design of validation data sets, definition of corresponding “ground truth” and its accuracy, validation protocol, and design of validation metrics (Jannin *et al.*, 2002). We have thoroughly validated our registration method by using MR, CT, and X-ray images of a cadaveric lumbar spine phantom for which “gold standard” registration was established by means of fiducial markers, and its accuracy assessed by target registration error (Tomažević *et al.*, 2002). However, because other researchers have used different phantoms, validation data sets, validation protocols, criteria for successful registrations, and error metrics, it is difficult to place the results of the proposed 3D/2D registration method in the context of previously published work. The lack of standardized validation methodology is probably the reason that most authors have made no

comparative evaluations of their methods (Gottesfeld Brown and Boult, 1996; Guéziec *et al.*, 1998; LaRose *et al.*, 2000a; Penney *et al.*, 1998; Weese *et al.*, 1997). To show that it is indeed difficult to make quantitative comparisons, we discuss the work of (Gottesfeld Brown and Boult, 1996; Penney *et al.*, 1998; Weese *et al.*, 1997) and as far as it is possible make comparisons with regard to the capturing range and registration accuracy. The methods used in (Gottesfeld Brown and Boult, 1996; Penney *et al.*, 1998; Weese *et al.*, 1997) are based on DRRs, and similarly to ours, relies on intraoperative X-ray projections, which does not require segmentation. Similar to our method, methods in (Penney *et al.*, 1998) and (Weese *et al.*, 1997) have been validated on lumbar spine.

In their experiment, Gottesfeld Brown and Boult (Gottesfeld Brown and Boult, 1996) used a CT and planar X-ray image of a dry femur bone rigidly attached to a precisely machined calibration object, which contained a grid of embedded stainless steel balls. These balls served to establish the “gold standard” 3D/2D registration. They performed only 5 registrations from starting positions obtained by translating or rotating the “gold standard” position for a random value in the range of  $\pm 2\text{mm}$  or degrees. To evaluate the error in the registration they computed the mean TRE in 3D for 100 target points and obtained 2.8mm for one femur configuration and 0.8mm for the other. For initial estimates greater than  $\pm 4\text{mm}$  or degrees, they did not typically get close to the “gold standard” registration. We make the following observations. First, registration of a dry femur is an easier task than registration of vertebrae with soft tissue that is submerged in water. Second, it is not possible to fairly compare the capturing ranges if the origin of the rotation is not known and if the parametrical space is not normalized according to the size of the object on which the targets are defined. Third, it is important where the target points are placed. In (Gottesfeld Brown and Boult, 1996) the target points were placed randomly in a  $1\text{dm}^3$  cube, while we have placed the targets around pedicles, which is the structure image-guided surgery will be performed on.

Weese *et al.* (Weese *et al.*, 1997) and Penney *et al.* (Penney *et al.*, 1998) experimented with a CT and an anterior-posterior (AP) fluoroscopy image of a phantom that consisted of a lumbar spine and pelvis encased in acrylic. Accurate “gold standard” registration was derived from fiducial markers attached to the phantom prior to imaging. Starting estimates in (Weese *et al.*, 1997) were slightly different than the starting estimates in (Penney *et al.*, 1998). Weese *et al.* (Weese *et al.*, 1997) started with the “gold standard” registration and obtained 81 different starting estimates for the rotation parameters and the height above the projection plane from all the combinations of ( $\omega_{Gi} + \Delta\omega$ ,  $\omega_{Gi}$ ,  $\omega_{Gi} - \Delta\omega$ ;  $i = x, y, z$ ) and ( $t_{Gz} + \Delta t$ ,  $t_{Gz}$ ,  $t_{Gz} - \Delta t$ ), where  $\omega_G$  is the “gold standard” rotation parameter. If the rotations  $\Delta\omega$ , which were set to 0.1, 0.15, and 0.2 radians, are transformed to our normalized parametrical space they correspond to displacements of 3.4, 5.2, and 6.9mm, respectively. The 64 starting estimates for registration in (Penney *et al.*, 1998) were the “gold standard” value  $\pm \Delta P$ , which was 3.4, 7.6 and 7.8 degrees for x,

y, and z rotation parameters, 3.6 and 2.4mm for in-plane, and 50.8mm for out-of-plane translation parameters. If the out-of-plane translation is neglected these values translate to 6mm in our normalized parametrical space. The starting estimates in (Penney *et al.*, 1998; Weese *et al.*, 1997) and in our experiments are thus comparable, but it is difficult to compare the rate of successful registrations and accuracy of our method and the methods in (Penney *et al.*, 1998; Weese *et al.*, 1997). Weese *et al.* (Weese *et al.*, 1997) had classified a registration as successful if the similarity measure after registration was larger than 0.75 times the similarity measure's maximum. In (Penney *et al.*, 1998), a failure was defined as when the final registration position, in any of the six-degrees-of-freedom, was further from "gold standard" than  $\Delta P$ . We believe that both criteria are less rigorous than ours, which is that the TRE of each target point is below 2mm. Using the criterion in (Penney *et al.*, 1998), quite some of our misregistrations would be classified as successful. The similarity measure, as in (Weese *et al.*, 1997), is not a good criterion because it depends on the behavior of the measure. Small changes may correspond to large spatial displacements from the "gold standard" position. For the method based on DRRs, as well as for our method, to work, it is important that at the start of registration, some corresponding features are at least partially overlapping. Weese *et al.* (Weese *et al.*, 1997) and Penney *et al.* (Penney *et al.*, 1998) had facilitated registration by determining the shifts parallel to the projection plane from manually indicated corresponding points in the DRR and X-ray image. In (Penney *et al.*, 1998; Weese *et al.*, 1997) the registration accuracy is given in terms of the differences between "gold standard" and final rotation and translation parameters. The parameter accuracy or error is not a useful measure, what should be measured is TRE (Fitzpatrick *et al.*, 1998). For different origins and object sizes the parameter values may be very similar, but there can be significant differences in TRE. In (Weese *et al.*, 1997) the rotation parameters corresponded within 0.5 degrees, the shifts parallel to the projection plane agreed within 0.5mm, while the deviations of the height above the projection plane were up to 5mm large. If we neglect the 5mm out-of-plane translation, these values correspond to 1mm in our normalized parametrical space. The smallest error values in rigid-body parameters obtained by Penney *et al.* (Penney *et al.*, 1998) were for the spine phantom with no added structures. If we neglect the 4mm out-of-plane translation error, the error values translate to 0.7mm in our normalized parametrical space. This is higher than the rms(TRE) we have obtained for CT/X-ray registrations.

To summarize, a fair comparison of different registration methods is possible only to a limited degree. What is needed to conduct a just comparative evaluation are validation data sets, definition of corresponding "ground truth" and its accuracy, validation protocol, and design of validation metrics. We have made a step in this direction and obtained such data, which is publicly available (Tomažević *et al.*, 2002).

The proposed 3D/2D registration method is neither a pure intensity based method, such as the timely DRR-based methods developed for X-ray/CT

registration, nor a feature-based method, which requires segmentation of both preoperative and intraoperative images, which is a serious drawback. By actually combining the two generic approaches to 3D/2D registration, the proposed method inherits their advantages, namely speed, accuracy and no need for intraoperative segmentation, and thus gains in generality and applicability. The 3D/2D registration has been performed for images of a spine phantom and results for different vertebrae were given to get an indication of registration accuracy. Further experiments should concentrate on questions such as: how the introduction of soft-tissue structures and surgical instruments and implants into the phantom images affect registration, which projections and how many X-ray images to use in a particular procedure, and how to reduce the edges in MR which do not correspond to bony structures. However, the proposed method, which matches a large number (20000-30000) of 3D surface normals and back-projected 2D gradients should be robust to outliers produced by surgical instruments and implants.

### 3.6 Appendix A

Let the X-ray source be at the origin of a coordinate system and the detector plane  $U$  perpendicular to  $z$ -axis. In spherical coordinates  $(r, \Theta, \varphi)$ , the image intensity  $I(\mathbf{p})$  and image intensity gradient  $\text{grad } I(\mathbf{p})$  at point  $\mathbf{p}$  are given as:

$$I(\mathbf{p}) = I(\Theta, \varphi) \quad (3.A1)$$

and

$$\text{grad } I(\mathbf{p}) = \frac{\partial I}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial I}{\partial \Theta} \mathbf{e}_\Theta + \frac{1}{r \sin \Theta} \frac{\partial I}{\partial \varphi} \mathbf{e}_\varphi = \frac{1}{r} \frac{\partial I}{\partial \Theta} \mathbf{e}_\Theta + \frac{1}{r \sin \Theta} \frac{\partial I}{\partial \varphi} \mathbf{e}_\varphi. \quad (3.A2)$$

The intensity gradient is thus lying in plane  $W$ , the plane that is perpendicular to projection beam  $L$ :

$$\text{grad } I(\mathbf{p}) = \text{grad}_W I(\mathbf{p}). \quad (3.A3)$$

Taking into account the relations:

$$\begin{aligned} x &= r \sin \Theta \cos \varphi \\ y &= r \sin \Theta \sin \varphi, \end{aligned} \quad (3.A4)$$

the image intensity gradient in detector plane  $U$ :

$$\text{grad}_U I(\mathbf{p}) = \frac{\partial I}{\partial x} \mathbf{i} + \frac{\partial I}{\partial y} \mathbf{j}, \quad (3.A5)$$

is defined in spherical coordinate system as:

$$\text{grad}_U I(\mathbf{p}) = \frac{\partial I}{\partial \Theta} \left[ \frac{\sin \Theta \cos \Theta}{r} \mathbf{e}_r + \frac{\cos^2 \Theta}{r} \mathbf{e}_\Theta \right] + \frac{1}{r \sin \Theta} \frac{\partial I}{\partial \varphi} \mathbf{e}_\varphi \quad (3.A6)$$

or

$$\text{grad}_U I(\mathbf{p}) = \text{grad}_W I(\mathbf{p}) - (\text{grad}_W I(\mathbf{p}) \cdot \mathbf{n}) \mathbf{n} \quad (3.A7)$$

This relation can be written as:

$$\text{grad}_U I(\mathbf{p}) = (\mathbf{n} \times \text{grad}_W I(\mathbf{p})) \times \mathbf{n} \quad (3.A8)$$

or

$$\text{grad}_W I(\mathbf{p}) = \frac{(\mathbf{n} \times \text{grad}_U I(\mathbf{p})) \times \mathbf{e}_r}{\mathbf{n} \cdot \mathbf{e}_r}. \quad (3.A9)$$

## 3.7 Appendix B

In the spherical coordinate system, the attenuation coefficient gradient  $\text{grad} \mu(\mathbf{r})$  is given as:

$$\begin{aligned} \text{grad} \mu(\tilde{r}, \Theta, \varphi) &= \frac{\partial \mu(\tilde{r}, \Theta, \varphi)}{\partial \tilde{r}} \mathbf{e}_r + \left[ \frac{1}{\tilde{r}} \frac{\partial \mu(\tilde{r}, \Theta, \varphi)}{\partial \Theta} \mathbf{e}_\Theta + \frac{1}{\tilde{r} \sin \Theta} \frac{\partial \mu(\tilde{r}, \Theta, \varphi)}{\partial \varphi} \mathbf{e}_\varphi \right] = . \quad (3.B1) \\ &= \text{grad}_r \mu(\tilde{r}, \Theta, \varphi) + \text{grad}_t \mu(\tilde{r}, \Theta, \varphi) \end{aligned}$$

The gradient is composed of two parts,  $\text{grad}_r \mu(\mathbf{r})$  in the direction of projection beam  $L$  and  $\text{grad}_t \mu(\mathbf{r})$  perpendicular to the beam.

## 3.8 Appendix C

Let  $S$  be a closed surface surrounding the imaged object and let  $r_1$  and  $r_2$  be distances from the origin of coordinate system to points where a projection beam enters and leaves the closed volume, respectively. Equation (3.4) in spherical coordinates is given as:

$$I(\mathbf{p}) = I(\Theta, \varphi) = \int_{r_1(\Theta, \varphi)}^{r_2(\Theta, \varphi)} \mu(\tilde{r}, \Theta, \varphi) d\tilde{r} \quad (3.C1)$$

and the partial derivative of  $I(\Theta, \varphi)$  on  $\Theta$  is:

$$\frac{\partial I}{\partial \Theta} = \int_{r_1(\Theta, \varphi)}^{r_2(\Theta, \varphi)} \frac{\partial \mu(\tilde{r}, \Theta, \varphi)}{\partial \Theta} d\tilde{r} + \mu(\tilde{r}_2(\Theta, \varphi), \Theta, \varphi) \frac{\partial r_2}{\partial \Theta} - \mu(\tilde{r}_1(\Theta, \varphi), \Theta, \varphi) \frac{\partial r_1}{\partial \Theta}. \quad (3.C2)$$

By assuming that  $\mu(\mathbf{r})$  is a smooth function and that  $\mu(\mathbf{r})$  is zero for points lying on  $S$  we obtain:

$$\frac{\partial I}{\partial \Theta} = \int_{r_1(\Theta, \varphi)}^{r_2(\Theta, \varphi)} \frac{\partial \mu(\tilde{r}, \Theta, \varphi)}{\partial \Theta} d\tilde{r}. \quad (3.C3)$$

Similarly, the partial derivative of  $I(\Theta, \varphi)$  on  $\varphi$  is given as:

$$\frac{\partial I}{\partial \varphi} = \int_{r_1(\Theta, \varphi)}^{r_2(\Theta, \varphi)} \frac{\partial \mu(\tilde{r}, \Theta, \varphi)}{\partial \varphi} d\tilde{r}. \quad (3.C4)$$

By inserting the partial derivatives  $\partial I / \partial \Theta$  and  $\partial I / \partial \varphi$  in (3.A2), the following expression is obtained:

$$\text{grad}_W I(\mathbf{p}) = \frac{1}{r} \int_{r_1}^{r_2} \tilde{r} \text{grad}_t \mu(\tilde{r}, \Theta, \varphi) d\tilde{r}. \quad (3.C5)$$

Every experiment proves something. If it doesn't prove what you wanted it to prove, it proves something else.

*Anonymous*

## Chapter 4

# 3D/2D registration of CT and MR images to X-ray images: The impact of number and projection of X-ray images.

Dejan Tomažević, Boštjan Likar and Franjo Pernuš  
Submitted for journal publication.

### Abstract

A crucial part of accurate image-guided radiation therapy or surgery is registration of a three-dimensional (3D) preoperative image, together with the radiation therapy or surgery plan, to the patient in the treatment room. By registration, the patient setup is controlled which enables more accurate and minimal invasive interventions. Among the many existing registration methods we consider the ones based on matching intensities or gradients of 3D computed tomography or magnetic resonance preoperative images and intraoperative two-dimensional (2D) X-ray images the most promising. It is expected that the accuracy, capturing range and robustness of these 3D/2D image-based registration methods do not depend solely on the imaging modalities and the registration method itself but also on the specific projections (views) and the number of intraoperative images used for registration. As these influences have not been systematically studied before, we investigate them in this study by using image data, comprised of CT, MR, and 18 X-ray images of a cadaveric lumbar spine phantom (Tomažević et al., 2004a) and a recently proposed 3D/2D registration method (Tomažević et al., 2003). The results for both, CT/X-ray and MR/X-ray, registrations indicate that the number of X-ray

*images has a rather small impact on accuracy but that the proportion of successful registrations (robustness) becomes significantly higher when more X-ray images are used for registration. The accuracy and proportion of successful registrations of CT/X-ray registrations did not significantly depend on the projections. However, the projections had an impact on the proportion of successful MR/X-ray registrations.*

## 4.1 Introduction

In radiation therapy and surgery, there is a constant demand to render the therapeutic procedures less and less invasive and to improve the accuracy with which a given procedure can be performed compared to conventional methods. In radiation therapy, the treatment is divided into fractions. This means that the setup of the patient has to be reproduced a large number of times and it must repeatedly be verified that the treatment is in accordance with the radiation plan. This is a critical step because misregistrations may lead to massive irradiation of healthy tissue while the tumor may be partially or even entirely missed. In surgery, the requirements for minimally invasiveness and higher accuracy are contradicting. Minimally invasive procedures (MIPs) commonly use small incisions that spatially constrain the surgeons' sight. Without seeing the precise location, shape, and margins of the targeted anatomical structures, surgeons performing MIPs may damage healthy tissues, critical nerves, and blood vessels, causing excess bleeding, paralysis, or other complications. Besides, surgeons may leave parts of target tissues untreated, which may also lead to complications or even to reduction of patients' long-term survival. The emerging technology that has the potential to decrease the invasiveness and increase the accuracy of radiation therapy and surgery is image guidance (Galloway, 2001; Jolesz, 1997; Taylor *et al.*, 1995). The crucial part of image-guided systems that enable repeatable intratherapy patient setup or provide intraoperative navigation guidance is registration of a patient in the treatment room to preoperative images or to models obtained from these images. A correct registration allows any 3D point defined in the preoperative image to be precisely located in a patient or treatment room coordinate system. After registration, the radiotherapist is provided with information about the real patient position with respect to the virtual patient position in the treatment plan, while the properly displayed 3D images provide the surgeon information about position of his instruments relative to the planned trajectory, nearby vulnerable structures, and the ultimate target. Any registration algorithm, to be used in a clinical application, must satisfy several requirements. The most important are those that concern accuracy, robustness, computation time, and complexity and invasiveness of intraoperative data acquisition.

A variety of approaches that rigidly register preoperative data to patients in the treatment room directly or indirectly via intraoperatively acquired images have been proposed in the past. The methods may be classified according to the data that is used to compute the registration transformation (Livyatan *et*

*al.*, 2003; Tomažević *et al.*, 2003). Patient-based methods use markers attached to the patient (Jones *et al.*, 1993; Maurer *et al.*, 1997; Taylor *et al.*, 1994) or patient surfaces (Colchester *et al.*, 1996; Grimson *et al.*, 1996; Herring *et al.*, 1998; Lavallée *et al.*, 1995; Maurer *et al.*, 1998; Simon *et al.*, 1995b), while the image-based methods use surfaces and contours of corresponding anatomical structures obtained by segmentation of pre- and intraoperative images, respectively (Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Lavallée and Szeliski, 1995), or image voxel and pixel intensity or gradient information (Gilhuijs *et al.*, 1996a; Goitein *et al.*, 1983; Hipwell *et al.*, 2003; Lemieux *et al.*, 1994; Livyatan *et al.*, 2003; Murphy, 1997; Penney *et al.*, 1998; Russakoff *et al.*, 2005a; Tomažević *et al.*, 2006; Tomažević *et al.*, 2003; Weese *et al.*, 1997). Besides the methods that register preoperative images to intraoperative data, a few methods have been proposed for non-rigid registration of 3D statistical models of anatomical structures to intraoperative data (Benameur *et al.*, 2003; Fleute and Lavallée, 1999; Hamadeh *et al.*, 1998; Zheng *et al.*, 2006). The methods that register preoperative 3D images or models to one or more intraoperative 2D images are commonly referred to as 3D/2D or 2D/3D registration methods.

Before an image-guided therapy (IGT) system is put into clinical use, it must undergo rigorous validation, which includes validation of individual components, validation of the overall system and a study of how uncertainties propagate through the entire image-guided therapy process. A prerequisites for validation of an IGT system component, like 3D/2D registration, is standardization of the validation methodology, which should include design of validation data sets, definition of corresponding “ground truth” and its accuracy, validation protocol, and design of validation metrics (Jannin *et al.*, 2002; Tomažević *et al.*, 2003). For the purpose of 3D/2D registration validation, we have recently devised a cadaveric lumbar spine phantom, acquired corresponding 3D MR and CT images and 2D X-ray images of the phantom, and established accurate “gold standard” rigid registration between 3D and 2D images by means of fiducial markers (Tomažević *et al.*, 2004a).

The accuracy, capturing range and robustness of a 3D/2D registration method based on intraoperative X-ray images don't depend solely on the registration method but also on the specific anatomical structure, imaging modality, the number of X-ray images used for registration and the angle between these images. In the past, researchers have used one, two, or more intra-operative images but, to the best of our knowledge, we are not aware of any study that would report on the impact of X-ray views and their number on the performance of a 3D/2D registration method. In the present study, we have therefore used image data, comprising CT, MR, and 18 X-ray images of a cadaveric lumbar spine phantom (Tomažević *et al.*, 2004a) and the validation protocol and validation metrics (Tomažević *et al.*, 2003) to study the impact of the number of intraoperative X-ray images and the impact of selected projections (views) on the capture range and registration accuracy of our recently proposed method for registering 3D CT or MR images to 2D X-ray images (Tomažević *et al.*, 2003). This method has been selected because it is

probably the only intensity-based method that works on CT and MR images and because a recent study of van de Kraats *et al.* (van de Kraats *et al.*, 2005a) showed that this method outperformed the method based on digitally reconstructed radiographs (DRRs) (Lemieux *et al.*, 1994; Penney *et al.*, 1998).

## 4.2 3D/2D registration method

Features that we use for rigid 3D/2D registration are normals  $\mathbf{v}_A$  to surfaces of bony structures found in preoperative CT or MR volumes and back-projected intensity gradients  $\mathbf{v}_B$  of intraoperative X-ray images (Figure 4.1)(Tomažević *et al.*, 2003). Feature selection is based on the assumption that strong intensity gradients in X-ray images correspond to boundaries of bony structures. Let  $\mathbf{r}_i^{S_v}$ ,  $i=1,2,\dots,N$ , be one of the  $N$  points, defined in the coordinate system  $\mathbf{S}_v$  of a 3D volume (CT or MR), lying on the surface of a 3D structure (Figure 4.1). The position of the same point in reference (patient) coordinate system  $\mathbf{S}_{ref}$  is given by rigid transformation  $\mathbf{T}$ :

$$\mathbf{r}_i^{S_{ref}} = \mathbf{r}_i = \mathbf{T}(\mathbf{r}_i^{S_v}) = \mathbf{R}\mathbf{r}_i^{S_v} + \mathbf{t}, \quad (4.1)$$

where  $\mathbf{R}$  and  $\mathbf{t}$  describe, respectively, the rotation and translation of the coordinate system  $\mathbf{S}_v$  with respect to  $\mathbf{S}_{ref}$ . The matrix  $\mathbf{R}$  is defined by three angles  $\omega=(\omega_x, \omega_y, \omega_z)^T$  representing rotations around  $x$ ,  $y$  and  $z$  axes, while the vector  $\mathbf{t}=(t_x, t_y, t_z)^T$  describes translations along  $x$ ,  $y$ , and  $z$  axes. Transformation  $\mathbf{T}$  is thus defined by six parameters  $\mathbf{q}=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)^T$ . Let  $\mathbf{r}_s$  be the position of the X-ray source in the reference coordinate system  $\mathbf{S}_{ref}$ . For a given position of a 3D image, defined by vector  $\mathbf{q}$ , the line which connects  $\mathbf{r}_s$  and  $\mathbf{r}_i$ , and has direction defined by unit vector  $\mathbf{e}_i$ , intersects the X-ray image plane  $U$  at point  $\mathbf{p}=p(\mathbf{r}_i)$  (Figure 4.1). Rigid registration of a 3D preoperative image to  $M$  intraoperative 2D images is concerned with finding the set of parameters  $\mathbf{q}$  that optimizes the criterion function  $CF$ :

$$CF = \sum_{j=1}^M \frac{\sum_{i=1}^N |\mathbf{v}_{Ai}^j| \cdot |\mathbf{v}_{Bi}^j| \cdot f(\alpha_i)}{\sum_{i=1}^N |\mathbf{v}_{Ai}^j| \cdot \sum_{i=1}^N |\mathbf{v}_{Bi}^j|}, \quad (4.2)$$

where  $N$  is the number of surface points  $\mathbf{r}_i^{S_v}$ ,  $\mathbf{v}_{Ai}$  is the vector normal to the surface at point  $\mathbf{r}_i^{S_v}$ ,  $\mathbf{v}_{Bi}$  is the X-ray image gradient  $\mathbf{v}_{Ui}$  back-projected to the position defined by  $\mathbf{r}_i^{S_v}$ , and  $f(\alpha_i)$  is the weighting function depending on angle  $\alpha_i$  between  $\mathbf{v}_{Ai}$  and  $\mathbf{v}_{Bi}$ . The gradient  $\mathbf{v}_{Bi}$  is obtained by back-projecting the gradient  $\mathbf{v}_{Ui}(\mathbf{p}_i)$ ,  $\mathbf{v}_{Ui}(\mathbf{p}_i)=\text{grad}_U I(\mathbf{p}_i)$ , of X-ray image intensity  $I(\mathbf{p}_i)$ :

$$\mathbf{v}_{Bi}(\mathbf{r}) = \frac{|\mathbf{p}_i - \mathbf{r}_s|}{|\mathbf{r}_i - \mathbf{r}_s|} \cdot \mathbf{v}_{Wi} = \frac{|\mathbf{p}_i - \mathbf{r}_s|}{|\mathbf{r}_i - \mathbf{r}_s|} \cdot \frac{(\mathbf{n} \times \mathbf{v}_{Ui}(\mathbf{p})) \times \mathbf{e}_i}{\mathbf{n} \cdot \mathbf{e}_i}, \quad (4.3)$$

where  $|\mathbf{r}_i - \mathbf{r}_s|$  and  $|\mathbf{p}_i - \mathbf{r}_s|$  are the distances of X-ray source  $\mathbf{r}_s$  to points  $\mathbf{r}_i$  and  $\mathbf{p}_i$ , respectively,  $\mathbf{v}_w(\mathbf{p})$ ,  $\mathbf{v}_w(\mathbf{p}) = \text{grad}_w I(\mathbf{p})$ , is the intensity gradient in the plane perpendicular to projection beam defined by  $\mathbf{e}_i$ , and  $\mathbf{n}$  is the unit normal to projection plane  $U$ . The function  $f(\alpha_i)$  is chosen such that it favors gradients that have similar directions:

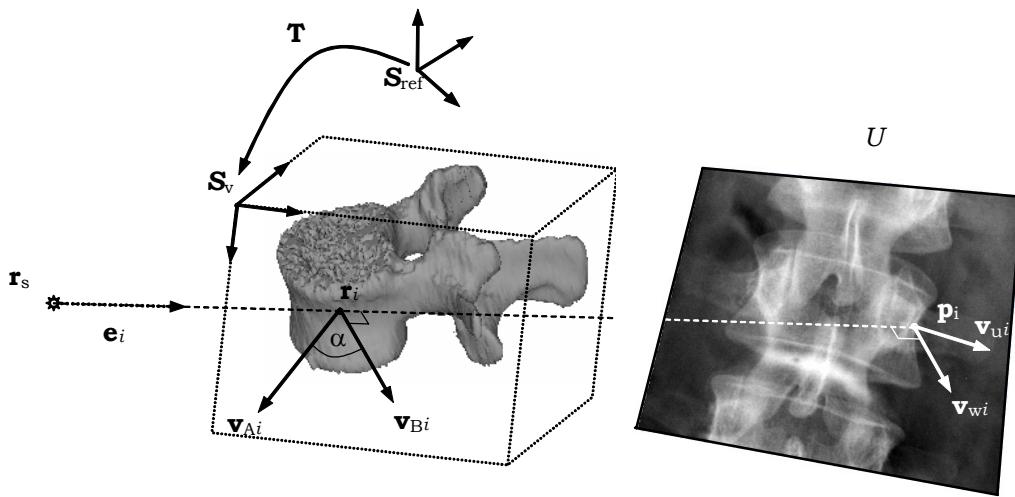
$$f(\alpha_i) = \begin{cases} \cos^4 \alpha_i = \left( \frac{|\mathbf{v}_{Ai} \cdot \mathbf{v}_{Bi}|}{|\mathbf{v}_{Ai}| \cdot |\mathbf{v}_{Bi}|} \right)^4 & 0^\circ < |\alpha_i| < 90^\circ \\ 0 & \text{otherwise} \end{cases} . \quad (4.4)$$

A detailed description of the registration method can be found in (Tomažević *et al.*, 2003).

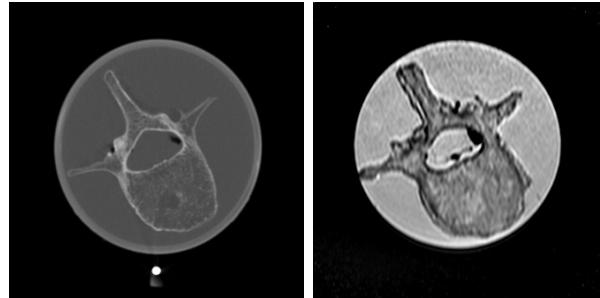
### 4.3 Images and “gold standard” registrations

Images used in the experiment were part of our 3D/2D registration “gold standard” lumbar spine phantom data set, which has been made publicly available<sup>2</sup> (Tomažević *et al.*, 2004a). A lumbar spine phantom had been constructed by placing a cadaveric lumbar spine, comprising vertebra L1-L5 with intervertebral disks and several millimeters of soft tissue, into a plastic tube filled with water. Six fiducial markers were rigidly attached to the surface of the tube. The CT image was obtained using a General Electric HiSpeed CT/i scanner. Axial slices were taken with intra-slice resolution of 0.27 x 0.27 mm and 1 mm inter-slice distance. For MR imaging, Philips Gyroscan NT Intera 1.5 T scanner and T1 protocol (flip angle 90°, TR=3220 ms, TE=11 ms) was used. Axial slices were obtained with 0.39 x 0.39 mm intra-slice resolution and 1.9 mm between-slice distance. Figure 4.2 shows CT and MR axial slices. The MR image was retrospectively corrected for intensity inhomogeneity by the information minimization method (Likar *et al.*, 2001). X-ray images were obtained by PIXIUM 4600 (Trixell) digital X-ray detector with a 429 x 429 mm large active surface, 0.143 x 0.143 mm pixel size, and 14 bits of dynamic range. The X-ray source and detector-plane were fixed during image acquisition while the spine phantom was rotated on a turntable to simulate a setup with C-arm. By rotating (step=20°) the spine phantom around its long axis, 18 X-ray images were acquired (Figure 4.3).

<sup>2</sup> <http://lit.fe.uni-lj.si/Downloads/downloads.asp>

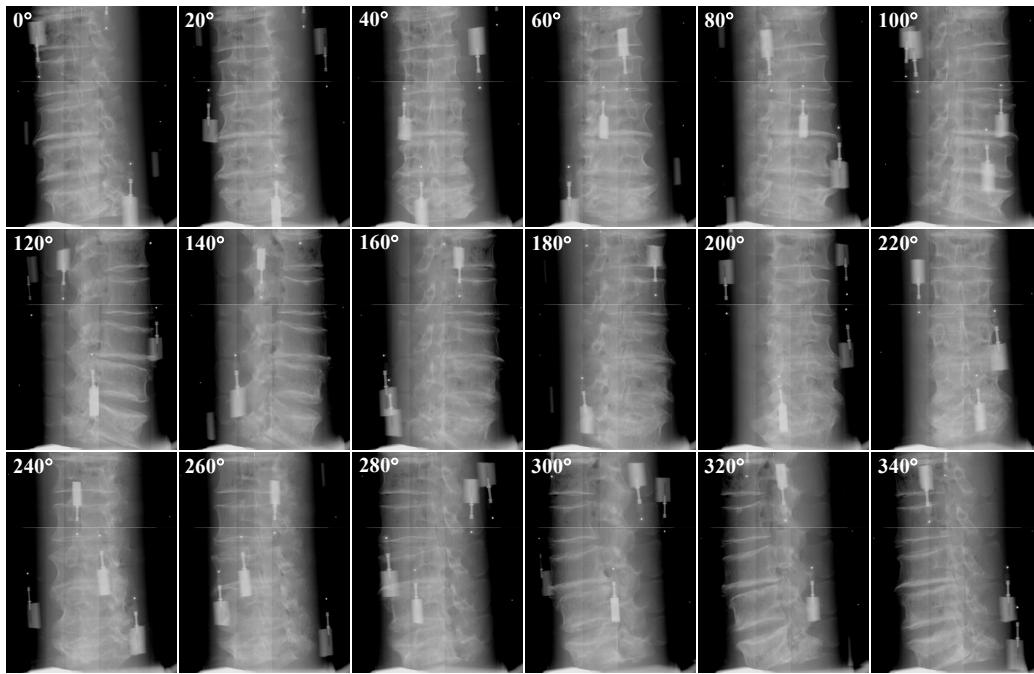


**Figure 4.1.** 3D/2D registration geometrical setup with normals to surfaces  $\mathbf{v}_{Ai}$  and back-projected intensity gradients  $\mathbf{v}_{Bi}$ .



**Figure 4.2.** An axial slice from a CT (left) and MR (right) volume.

“Gold standard” registrations of CT to X-ray and MR to X-ray images were obtained by rigid registration of CT and MR marker points to 3D marker points reconstructed from X-ray images, respectively. As such, the obtained “gold standard” registration defines the relative position of each of the X-ray images to CT and MR volumes and, consequently, the relative positions between X-ray images. The accuracy of “gold standard” registration was obtained by estimating the target registration error TRE (Fitzpatrick *et al.*, 1998) for eight targets, four per each pedicle, which were manually defined on each of the 5 vertebrae. The estimated root mean square (RMS) TREs for all vertebrae were below 0.29 mm for CT to X-ray registration and below 0.42 mm for MR to X-ray registration, which indicated that the “gold standard” was highly accurate. A detailed description of the spine phantom, X-ray system calibration, calculation of “gold standard” registration and “gold standard”



**Figure 4.3.** The 18 acquired X-ray images of the spine phantom.

validation can be found in (Tomažević *et al.*, 2004a). After registration, cubic sub-volumes that contained single vertebrae were defined manually in CT and MR images. Each sub-volume was blurred using a Gaussian filter ( $\sigma = 0.5$  mm) and isotropically re-sampled to the resolution of 1 mm. The Canny edge detector and a threshold were applied to automatically extract locations of points corresponding to boundaries of bony structures and to estimate surface normal directions at these points. This resulted in approximately 20000 points for the CT and 30000 points for the MR images. The X-ray images were blurred with a Gaussian filter ( $\sigma = 0.5$  mm) and the Roberts edge detector was applied to calculate intensity gradients  $\text{grad}I(\mathbf{p})$ .

## 4.4 Experiments

To test the impact of the number and projections (views) of X-ray images on the accuracy, capturing range, and robustness of the 3D/2D registration method, registrations were performed from a wide range of starting positions and orientations around the “gold standard” registration position using different projections and different numbers of X-ray images. Before registrations, the 6-dimensional parametrical space was normalized, so that a rotation of the volume containing a single vertebra of size 80 mm around its center for 0.1 radians (5.7°) was equivalent to mean translation of volume points for 2 mm. In this way, Euclidean metrics could be used to calculate the displacement (in parametrical space) of a starting position from the “gold

**Table 4.1.** The number and angle between two consecutive views in 16 sets of X-ray images registered to CT and MR images.

X-ray images					
Set			Subsets		
No.	M	Δ	1	2	...
<b>1</b>	<b>1</b>		(0°)	(20°)	...
<b>2</b>	<b>2</b>	<b>20°</b>	(0°,20°)	(20°,40°)	...
<b>3</b>		<b>40°</b>	(0°,40°)	(20°,60°)	...
<b>4</b>		<b>60°</b>	(0°,60°)	(20°,80°)	...
<b>5</b>		<b>80°</b>	(0°,80°)	(20°,100°)	...
<b>6</b>		<b>100°</b>	(0°,100°)	(20°,120°)	...
<b>7</b>		<b>120°</b>	(0°,120°)	(20°,140°)	...
<b>8</b>		<b>140°</b>	(0°,140°)	(20°,160°)	...
<b>9</b>		<b>160°</b>	(0°,160°)	(20°,180°)	...
<b>10</b>	<b>3</b>	<b>20°</b>	(0°,20°,40°)	(20°,40°,60°)	...
<b>11</b>		<b>40°</b>	(0°,40°,80°)	(20°,60°,100°)	...
<b>12</b>		<b>60°</b>	(0°,60°,120°)	(20°,80°,140°)	...
<b>13</b>		<b>80°</b>	(0°,80°,160°)	(20°,100°,180°)	...
<b>14</b>	<b>4</b>	<b>20°</b>	(0°,20°,40°,60°)	(20°,40°,60°,80°)	...
<b>15</b>		<b>40°</b>	(0°,40°,80°,120°)	(20°,60°,100°,140°)	...
<b>16</b>	<b>8</b>	<b>20°</b>	(0°,20°,...,140°)	(20°,40°,...,160°)	...
					(340°,0°,...,120°)

**M - number of images in a subset, Δ - angle between consecutive images in a subset**

standard” position (Tomažević *et al.*, 2003). Values of parameters defining a starting point were chosen randomly around the “gold standard” position in such a way that a uniform distribution of displacements in the interval [0, D] was achieved. D was set to 18mm (51.6°) for CT and to 9mm (25.8°) for MR registrations. By using a smaller value of D for MR, we were able to more precisely define the capturing ranges for MR/X-ray registrations, which, because of larger differences in modalities, had been expected to be smaller than the capturing ranges for CT/X-ray registrations. Optimization of transformation parameters  $\mathbf{q}=(t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)^T$  was performed by Powell’s method (Press *et al.*, 1992). To measure the registration error before and after registration, TRE was calculated for eight target points (four on each pedicle) as the distance between target points in the registered and “gold standard” position:

$$TRE(\mathbf{r}) = \|\mathbf{T}_r \mathbf{r} - \mathbf{T}_g \mathbf{r}\|, \quad (4.5)$$

where  $\mathbf{r}$  is a target point, and  $\mathbf{T}_r$  and  $\mathbf{T}_g$  are the transformations obtained by the proposed registration method and “gold standard” registration, respectively. The rotation error  $\theta$  was determined by decomposing TRE into translation and rotation components (West *et al.*, 1997). A registration was treated successful if TREs of all 8 target points were below 2 mm.

Sixteen X-ray image sets, each one containing 18 image subsets were formed (Table 4.1). The first set contained the 18 single images. Subsets in the next eight sets were comprised of image pairs taken at views, which differed by 20°, 40°,..., 160°, respectively. Subsets in the next four sets contained three images. The angles between the three images in a subset were 20°, 40°, 60°, or 80°. In the next two subsets of set 4 there were four images, the angles between them were 20° and 40°. In the 18 subsets of the last set there were eight images taken at views 20° apart. The 5 CT and 5 MR sub-volumes, each containing one of the L1-L5 vertebra, were registered from 50 initial positions to each of the 18 subsets of X-ray images from all 16 image sets, which yielded 4500 (5x50x18) registrations for each set and each modality.

## 4.5 Results and discussion

Tables 4.2 and 4.3 show the registration results for CT to X-ray and MR to X-ray registrations, respectively. Both tables give the means and standard deviations of TREs and rotation components  $\theta$  before and after registration. Only results of successful registrations are given. Tables 4.2 and 4.3 also give the proportion of successful registrations for three intervals of displacements.

The results for X-ray to CT registrations (Table 4.2) show that the registration errors fell with the higher number of images utilized for registration. The mean TRE achieved with one image was 0.9 mm, with 2 images between 0.32 and 0.41, and with 8 images 0.3 mm. The accuracy of registrations did not significantly depend on the projections. For instance, when registering X-ray image pairs to the CT image, the best results were achieved with X-ray images being 80 and 100 degrees apart. However, the values for 80 and 100 degrees (0.32 mm) were not significantly smaller than the values for other angles between image pairs.

**Table 4.2.** Registration errors and proportion of successful registrations for CT to X-ray registration using different X-ray projections and number of projections.

Sets			Before registration		After registration		Successful registrations (%)		
No.	M	Δ	TRE [mm] Mean (Std)	θ [°] Mean (Std)	TRE [mm] Mean (Std)	θ [°] Mean (Std)	0÷6mm 0÷17.2°	6÷12mm 17.2÷34.4°	12÷18mm 34.4÷51.7°
<b>1</b>	<b>1</b>		3.3 (2.9)	7.3 (6.4)	0.90 (0.49)	0.52 (0.30)	61.8%	30.6%	13.5%
<b>2</b>	<b>2</b>	<b>20°</b>	5.2 (3.7)	11.0 (8.1)	0.41 (0.18)	0.41 (0.22)	96.4%	56.4%	13.7%
<b>3</b>		<b>40°</b>	5.1 (3.6)	11.0 (8.1)	0.34 (0.13)	0.38 (0.21)	95.6%	55.4%	13.1%
<b>4</b>		<b>60°</b>	5.0 (3.5)	10.7 (8.0)	0.33 (0.11)	0.36 (0.20)	95.2%	52.7%	11.3%
<b>5</b>		<b>80°</b>	4.9 (3.5)	10.6 (7.7)	0.32 (0.11)	0.34 (0.18)	94.7%	52.0%	10.5%
<b>6</b>		<b>100°</b>	4.9 (3.5)	10.6 (7.9)	0.32 (0.12)	0.35 (0.18)	94.2%	48.3%	11.2%
<b>7</b>		<b>120°</b>	5.0 (3.6)	10.8 (8.0)	0.34 (0.11)	0.36 (0.19)	93.8%	52.1%	12.3%
<b>8</b>		<b>140°</b>	5.1 (3.6)	11.1 (8.3)	0.34 (0.12)	0.38 (0.19)	95.8%	55.0%	13.6%
<b>9</b>		<b>160°</b>	5.3 (3.7)	11.2 (8.1)	0.40 (0.17)	0.38 (0.18)	95.3%	59.0%	15.1%
<b>10</b>	<b>3</b>	<b>20°</b>	5.5 (3.8)	11.7 (8.4)	0.34 (0.13)	0.34 (0.19)	98.4%	63.6%	18.6%
<b>11</b>		<b>40°</b>	5.4 (3.7)	11.5 (8.2)	0.31 (0.11)	0.29 (0.13)	98.6%	64.5%	16.4%
<b>12</b>		<b>60°</b>	5.5 (3.8)	11.9 (8.6)	0.31 (0.11)	0.28 (0.13)	97.9%	66.0%	19.4%
<b>13</b>		<b>80°</b>	5.4 (3.7)	11.7 (8.5)	0.31 (0.12)	0.30 (0.16)	97.7%	64.4%	18.2%
<b>14</b>	<b>4</b>	<b>20°</b>	5.8 (3.9)	12.4 (8.9)	0.32 (0.11)	0.31 (0.16)	98.9%	69.0%	23.8%
<b>15</b>		<b>40°</b>	5.7 (3.9)	12.3 (8.7)	0.30 (0.11)	0.26 (0.12)	99.2%	70.3%	23.1%
<b>16</b>	<b>8</b>	<b>20°</b>	6.3 (4.1)	13.6 (9.5)	0.30 (0.11)	0.24 (0.12)	99.5%	80.4%	36.7%

M - number of images in a subset, Δ - angle between consecutive images in a subset

The percentage of successful registrations also increased with the higher number of images used for registration. It increased from 96% for two images to 99.5% for eight images if the initial displacements were between 0 and 6 mm. For the initial displacements between 6 and 12 mm the increase was much more dramatic. The proportion of successful registrations increased from 30% (one image), to 59% (two images), to 66% (three images), to 79% (four images) and, finally, to 80% for eight images. The same trend could be observed for displacements between 12 and 18 mm. However, even with eight images the proportion of successful registrations was much smaller than it would be required in a clinical situation.

**Table 4.3.** Registration errors and proportion of successful registrations for MR to X-ray registration using different X-ray projections and number of projections.

Sets			Before registration		After registration		Successful registrations (%)		
No.	M	Δ	TRE [mm] Mean (Std)	θ [°] Mean (Std)	TRE [mm] Mean (Std)	θ [°] Mean (Std)	0±3mm 0÷8.6°	3±6mm 8.6÷17.2°	6±9mm 17.2÷25.8°
<b>1</b>	<b>1</b>		2.0 (1.8)	4.6 (4.2)	1.30 (0.27)	1.48 (0.82)	27.0%	7.2%	2.9%
<b>2</b>	<b>2</b>	<b>20°</b>	3.0 (2.1)	6.5 (4.8)	1.27 (0.26)	1.17 (0.74)	64.3%	43.2%	20.7%
<b>3</b>		<b>40°</b>	3.2 (2.2)	6.7 (4.9)	1.26 (0.25)	1.16 (0.68)	70.3%	49.7%	27.0%
<b>4</b>		<b>60°</b>	3.1 (2.1)	6.6 (5.0)	1.21 (0.23)	1.14 (0.71)	75.8%	52.7%	26.8%
<b>5</b>		<b>80°</b>	2.9 (2.1)	6.3 (4.9)	1.20 (0.22)	1.20 (0.72)	79.4%	47.7%	24.9%
<b>6</b>		<b>100°</b>	3.1 (2.2)	6.6 (5.0)	1.21 (0.25)	1.19 (0.69)	80.2%	52.5%	28.6%
<b>7</b>		<b>120°</b>	3.1 (2.2)	6.6 (5.0)	1.22 (0.24)	1.13 (0.58)	75.3%	51.7%	27.3%
<b>8</b>		<b>140°</b>	3.1 (2.1)	6.6 (4.9)	1.26 (0.26)	1.12 (0.60)	72.8%	53.2%	26.0%
<b>9</b>		<b>160°</b>	3.1 (2.1)	6.5 (4.8)	1.26 (0.24)	1.07 (0.56)	64.6%	42.7%	22.2%
<b>10</b>	<b>3</b>	<b>20°</b>	3.3 (2.2)	7.2 (5.1)	1.26 (0.25)	1.03 (0.60)	80.8%	64.3%	37.9%
<b>11</b>		<b>40°</b>	3.4 (2.2)	7.4 (5.2)	1.23 (0.24)	1.06 (0.64)	87.6%	71.2%	44.7%
<b>12</b>		<b>60°</b>	3.4 (2.2)	7.4 (5.3)	1.19 (0.23)	0.99 (0.49)	88.6%	70.8%	45.7%
<b>13</b>		<b>80°</b>	3.3 (2.2)	7.1 (5.1)	1.19 (0.21)	1.00 (0.56)	90.6%	70.8%	41.4%
<b>14</b>	<b>4</b>	<b>20°</b>	3.5 (2.3)	7.5 (5.2)	1.23 (0.25)	0.91 (0.50)	89.8%	77.4%	50.3%
<b>15</b>		<b>40°</b>	3.5 (2.3)	7.7 (5.3)	1.21 (0.21)	0.97 (0.48)	96.9%	85.3%	59.8%
<b>16</b>	<b>8</b>	<b>20°</b>	3.8 (2.3)	8.3 (5.6)	1.20 (0.20)	0.76 (0.35)	98.1%	94.6%	79.5%

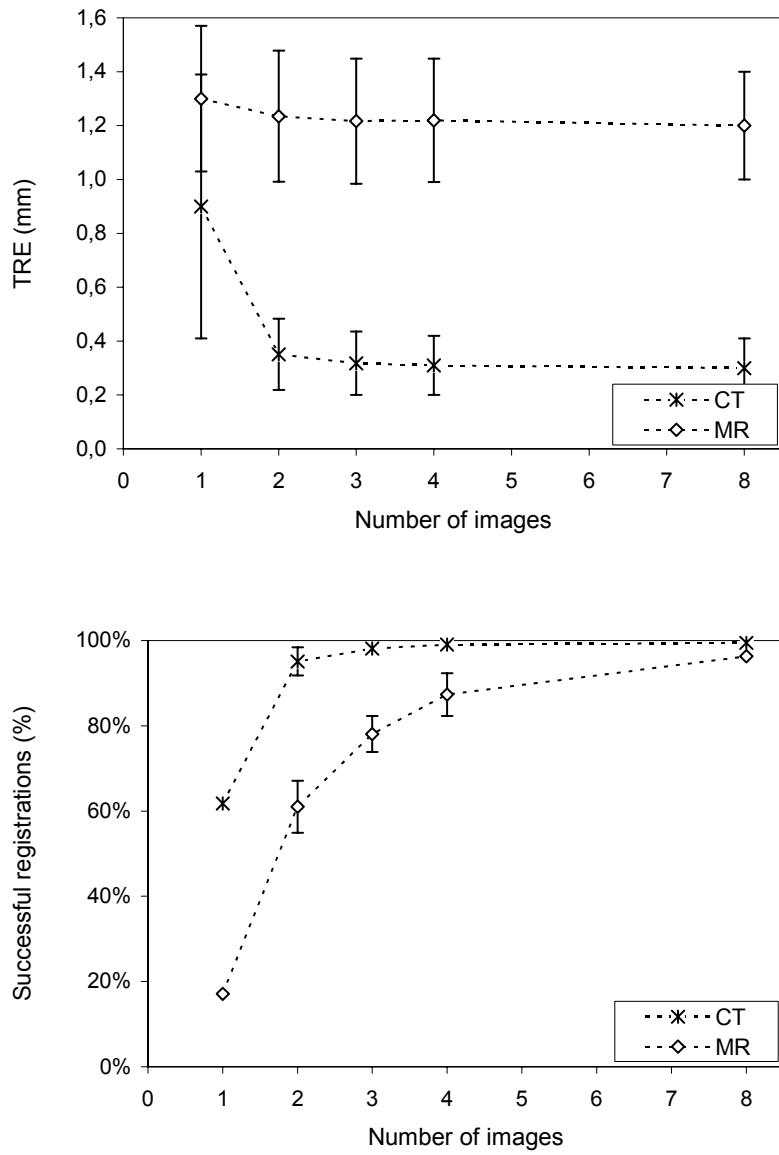
M - number of images in a subset, Δ - angle between consecutive images in a subset

Because MR and X-ray images have completely different physical backgrounds, worse registration results were expected for these modalities than for CT/X-ray registrations. Indeed, X-ray to MR registrations were not as accurate as X-ray to CT registrations (Table 4.3). Besides, the capture ranges, i.e. proportion of successful registrations for a given range of initial displacements, were smaller. The differences in mean TRE, achieved with one, two, three, four, and eight images were small. The smallest mean TRE (1.19mm) was achieved with three images which were 60 degrees apart. The impact of the higher number of images on the proportion of successful registrations was, however, significant. It increased from 27% (7%, 3%) for one image to 98% (95%, 80%) for eight images if the initial displacements were 0±3mm (3±6mm, 6±9mm), respectively. As with CT/X-ray registrations, the TREs and proportion of successful registrations also depended on the angles between the images. The best results for two, three, and four images were

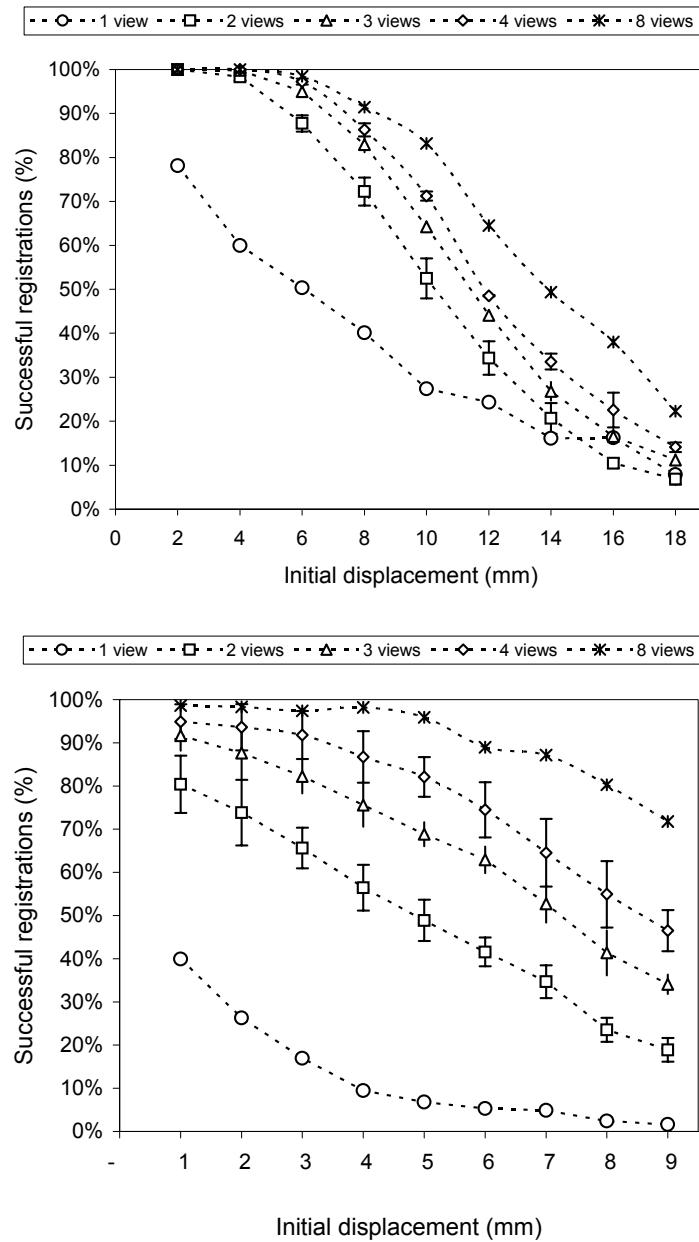
obtained when the images were 80 and 100, 60 and 80, and 40 degrees apart, respectively.

The impact of the number of intraoperative X-ray images on TREs and proportion of successful registrations is graphically illustrated in Figure 4.4, where the results for all image sets with same number of views were joined. Similarly, Figure 4.5 shows the proportion of successful CT to X-ray and MR to X-ray registrations with respect to the initial displacements and the number of X-ray images. The impact of the angle between two image views (results for sets 2-9, Table 4.1 and 4.2) on TREs and proportion of successful registrations is illustrated in Figure 4.6.

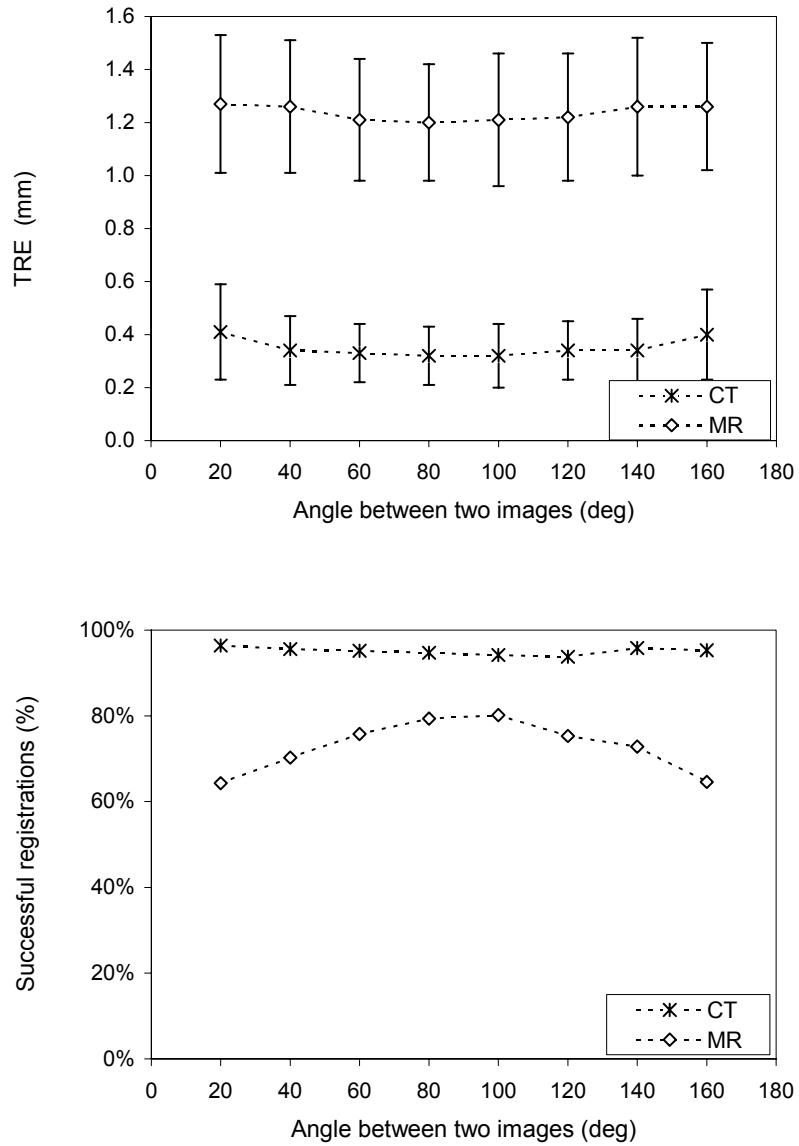
Results in Figure 4.4 (top) indicate that registration errors significantly dropped when two images were used instead of one. This was much more significant for CT/X-ray registration where using an additional X-ray image more than doubled the registration accuracy, i.e. from 0.9 mm to 0.4 mm, than for MR/X-ray registration. By taking more than two images the accuracy only slightly improved. The impact of the number of images on registration accuracy could be explained with a fact that, when using a single X-ray image, the registration is well-defined for two in-plane translations and one in-plane rotation and ill-posed for one out-of-plane translation and two out-of-plane rotations (van de Kraats *et al.*, 2005a; Weese *et al.*, 1997). By adding an additional X-ray image, with a significantly different angle of view, one out-of-plane translation and one out-of-plane rotation becomes in-plane and thereby well-defined. Regardless of the angle of view, one out-of-plane rotation still remains, if the X-ray images are acquired in a single plane, which is the case with a single rotation axis C-arm. However, even the out-of-plane rotation can be well-defined with two X-rays if the object of registration is not symmetric around one of its axis. The observation that images additional to two insignificantly contribute to registration accuracy (top of Figure 4.4) is thus explained with a fact that a vertebra body has no symmetrical rotation axes. However, for some other anatomical structure, for example for the shaft of the femur, which is close but not perfectly symmetric around one of the rotational axes, two X-ray images may not be sufficient to obtain the desired registration accuracy.



**Figure 4.4.** The impact of the number of X-ray images on the TRE (top) and proportion of successful registrations (bottom) for 0-6mm (CT) and 0-3mm (MR) initial displacements.



**Figure 4.5.** Means and standard deviations of successful X-ray to CT (top) and X-ray to MR (bottom) registrations as a function of initial displacement and the number of X-ray images.



**Figure 4.6.** The impact of the angle between images in image sets containing two images on the TRE (top) and proportion of successful registrations (bottom) for 0-6mm (CT) and 0-3mm (MR) initial displacements.

The number of images had, however, a much greater impact on the proportion of successful registrations for both CT/X-ray and MR/X-ray registrations (bottom of Figure 4.4). By utilizing more images for registration the proportion of successful registrations significantly increased, indicating that the capture ranges of CT/X-ray and MR/X-ray registrations increased and that the registrations became more robust. The impact of the number of images on the proportion of successful registrations was higher for MR/X-ray registrations than for CT/X-ray registrations. However, as the data in Tables 4.2 and 4.3 indicate, registration failures occurred even if 8 images were used. The main reasons for misregistrations are the local minima of the similarity measure into which the optimization gets trapped. The probability that optimization will be trapped in a local minima is higher when initial displacements are larger (Figure 4.5). This is a property of all automated registration methods and is a consequence of imaging modalities and associated image artifacts, imperfect similarity measures and local optimization procedures. By using more X-ray images the statistical power of the criterion function is increased and, consequently, the criterion function is less noisy. Moreover, the information from additional images reduces the effect of outliers, e.g. occlusions or image artifacts, because they are different in different X-ray images and in CT or MR volumes. At the end, both, the increased statistical power and the decreased effect of outliers, result in a smoother criterion function, which reduces the probability of the optimization algorithm to converge to some local optimum.

The angle between views has a small and insignificant impact on registration accuracy (Figure 4.6 top). The smallest TREs of 0.32 mm for CT and 1.2 mm for MR were achieved with two almost perpendicular views, while for views which were 20 degrees apart, the TREs were of 0.41 mm and 1.27 mm for CT and MR data, respectively. As mentioned above, in order that a 3D/2D rigid registration is well posed, it should use at least two X-ray images with significantly different views. The angle could be as small as 20 degrees for CT/X-ray registration, since the results show that the angle of view has a small impact on CT/X-ray registration robustness (Figure 4.6 bottom). The angle between two views had a much greater impact on the proportion of successful MR/X-ray registrations. The success rate varied from 64% to 80% for angles from 20 to 80 degrees, respectively, for a initial displacement of 0-3 mm (Table 4.3). However, for CT or MR registration to three or more X-rays, the accuracy and robustness was not affected by the angle of view as long as the angle between views was at least 20 degrees (Table 4.2 and 4.3).

As was expected, due to completely different imaging modalities, MR/X-ray registrations were not as accurate and robust as CT/X-ray registrations. The poorer performance of MR/X-ray registration lies in the fundamental assumption of the applied registration method (Tomažević *et al.*, 2003). The method assumes that strong intensity gradients in X-ray images correspond to boundaries of bony structures extracted from the 3D preoperative image. While, the extraction of bone boundaries from CT images is trivial, as it only requires detection of large intensity gradients, which have corresponding

gradients in X-ray images, this is not true for MR images. Usually MR imaging provides images with high contrast between different soft tissues, while the boundaries between bone and soft tissues do not exactly correspond to the bone boundaries that would be extracted from CT images. It has been shown in (Markelj *et al.*, 2007) bone boundaries in MR images can be displaced up to 1 mm in the positive or negative direction of the normal to the bone surface. Moreover, the contrast between different soft tissues can be of the same magnitude as the contrast between bone and its surrounding tissue. The consequence is that a high percentage of the edges extracted with the Canny edge detector does not correspond to bone boundaries but to boundaries of different soft tissues. These false bony edges represent outliers in the registration process. The displaced true bony edges yield lower correspondence with the gradients in X-ray images and result in lower accuracy of MR/X-ray registrations. Furthermore, the high percentage of outliers in conjunction with displaced edges makes MR/X-ray registrations less robust than CT/X-ray registrations. Since the accuracy and robustness of the applied 3D/2D registration method depends on true bone edges, extraction of such edges can be improved by carefully selecting the MR imaging protocol and a better algorithm for bone edge detection. Kraats *et al.* (van de Kraats *et al.*, 2005b) proposed the use of multi-spectral MR image acquisition which enhances bone edge detection and consequently yields better MR/X-ray registrations.

There are four important issues that have to be addressed before a 3D/2D registration based procedure can be considered clinically acceptable. Namely, the registration accuracy required for a specific clinical application, the robustness of the algorithm in the terms of success rates and capture range, the speed with which the method can be performed, and the clinical feasibility in terms of image acquisition, user interaction and intervention protocol requirements and/or acceptance. The selection of the preoperative 3D imaging modality, the number of X-ray image acquisitions and the angle between views depend on the issues above. The obtained accuracy of CT to more than one X-ray image was below 0.4 mm, regardless of the angle between image views, which is probably sufficient for orthopedic procedures. The accuracy of 1.3 mm for MR/X-ray single vertebrae body registration can still be adequate for some applications (Rampersaud *et al.*, 2001). For a registration to be acceptable in a clinical environment, the success rate should be above 95%. We have achieved such a rate in CT/X-ray registrations when two or more images were used and when the initial displacement was in the range of 0-6mm. For the MR/X-ray images a success rate higher than 95% was achieved only when the initial displacement was in the range of 0-3mm and 4 or 8 images were used. The initial displacement of the patient in the treatment room will generally be larger than the ranges that we used in the experiments. Therefore, some kind of simple user initialization is needed that will align the patient and his preoperative images to such an extent that automated registration will be successful in more than 95%. Because even if the initial displacement is small the possibility of a false registration still exists, the registration protocol should provide the user a tool for validation of the

outcome of the registration procedure (Guéziec *et al.*, 2000). The time required for the whole registration procedure can be divided into the time necessary for intraoperative image acquisition, the time required by the user to closely initialize the patient position, the actual registration time of the algorithm and the time needed by the user to validate the registration results. Acquisition and use of more images can improve the registration success rate but requires extra time. Besides, the patient is exposed to higher radiation. With regard to feasibility, a minimal angle between views can be useful for those clinical applications that cannot afford time and space for acquiring X-rays with perpendicular views. For CT/X-ray registrations the angle of 20 degrees between two X-ray image views proved sufficient, while for a reliable MR/X-ray registration at least four X-ray images with views at least 20 degrees apart were required. The preoperative imaging modality (CT or MR) strongly depends on the clinical application. Usually, the main aspect is the type of information that is needed for the specific procedure. For example, when exact dimensions of bone are needed for planning orthopedic procedures (Taylor *et al.*, 1994), the CT imaging is preferred, while for defining the position and size of a brain tumor for radiotherapy MR can be more appropriate (Yin *et al.*, 1998). Although accurate and robust MR/X-ray registration require more X-ray images the overall radiation dose that the patient receives is generally smaller than in CT/X-ray registrations (Wall and Hart, 1997).

## 4.6 Conclusion

The accuracy and robustness of registering a 3D preoperative image with 2D intraoperative images depends on the imaging modality, registration method and the anatomical structures that are to be registered. However, it is expected that the quality of 3D/2D registration also depends on the number of 2D intraoperative images and the projections under which these images are acquired. In the past, researchers have used one, two or more intra-operative images but have not systematically studied the impact of X-ray views and their number on the performance of an intensity-based 3D/2D registration method. To confirm the expectations that better registrations results can be achieved with more intraoperative images and with images taken at certain projections, we have conducted a study using the recently proposed 3D/2D intensity-based registration method (Tomažević *et al.*, 2003) and images and “gold standard” registration data (Tomažević *et al.*, 2004a). The results, presented in this paper confirmed the expectations. The accuracy and percentage of successful registrations for different initial displacements (capture range) of both, CT/X-ray and MR/X-ray, registrations increased when more X-ray images had been utilized. On the other hand, the projections had a significant impact only on the proportion of successful MR/X-ray registrations. These conclusions were made on the basis of registration results obtained by a specific 3D/2D registration method and anatomical structures. However, we believe that the present study provides general trends on how the number and the relative angle between X-ray image views affect a 3D/2D registration and gives some useful guidelines for further studies.

Theory guides. Experiment decides.

*Donald E. Simanek*

## **Chapter 5**

# **Standardized evaluation methodology for 2D-3D registration**

Everine B. van de Kraats, Graeme P. Penney,  
Dejan Tomažević, Theo van Walsum and Wiro J. Niessen  
IEEE Transactions on Medical Imaging 24: 1177-1189 (2005)

### **Abstract**

*In the past few years a number of 2D-3D registration algorithms have been introduced. However, these methods have been developed and evaluated for specific applications, and have not been directly compared. Understanding and evaluating their performance is therefore an open and important issue. To address this challenge we introduce a standardized evaluation methodology, which can be used for all types of 2D-3D registration methods and for different applications and anatomies. Our evaluation methodology uses the calibrated geometry of a 3D rotational X-ray (3DRX) imaging system in combination with image-based 3D-3D registration for attaining a highly accurate gold standard for 2D X-ray to 3D MR/CT/3DRX registration. Furthermore, we propose standardized starting positions and failure criteria to allow future researchers to directly compare their methods. As an illustration, the proposed methodology has been used to evaluate the performance of two 2D-3D registration techniques, viz. a gradient-based and an intensity-based method, for images of the spine. The data and gold standard transformations are available to other researchers.*

## 5.1 Introduction

Registration of 2D to 3D images has been proposed to help in a number of clinical areas, such as radiotherapy planning and treatment verification, spinal surgery, hip replacement, neurointerventions and aortic stenting. 2D-3D registration can be a means to noninvasively register the patient to the volume for image-guided navigation by finding the best match between one or more intraoperative X-ray projections of the patient and the preoperative 3D volume.

Several researchers have described and evaluated 2D-3D registration methods, which can roughly be subdivided in feature-based (Bansal *et al.*, 1999; Czopf *et al.*, 1999; Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Hamadeh *et al.*, 1998; Kita *et al.*, 1998; Lavallée and Szeliski, 1995; Murphy, 1997), in intensity-based (Hipwell *et al.*, 2003; Jonić *et al.*, 2001; Lemieux *et al.*, 1994; Penney *et al.*, 2001; Penney *et al.*, 1998; Rohlfing *et al.*, 2002; Russakoff *et al.*, 2003a; Sarrut and Clippe, 2000; Weese *et al.*, 1999; Weese *et al.*, 1997), hybrid (Russakoff *et al.*, 2003a; Vermandel *et al.*, 2003) and, most recently, gradient-based (Livyatan *et al.*, 2003; Tomažević *et al.*, 2003) methods. Their results are not directly comparable, owing to the use of different datasets, different starting positions, different failure criteria and different error calculation methods, as was also commented on in (Tomažević *et al.*, 2003). Only a few efforts have been aimed at comparing 2D-3D registration approaches (McLaughlin *et al.*, 2002; Penney *et al.*, 1998; Russakoff *et al.*, 2003a), mainly comparing similarity measures for the intensity-based approach. Moreover, many of the commonly used error measures do not adequately represent the accuracy of registration performance (LaRose *et al.*, 2000b).

Evaluating the performance of 2D-3D registration approaches requires a reference dataset with a known gold standard registration. Previously, gold standards for 2D-3D registration were either obtained manually (Hipwell *et al.*, 2003; McLaughlin *et al.*, 2002), by controlled motion (Lavallée and Szeliski, 1995; Murphy, 1997; Tomažević *et al.*, 2004a), by independent measuring devices (Feldmar *et al.*, 1997a; Lavallée and Szeliski, 1995; Lemieux *et al.*, 1994), by back-projection of images features (mainly implanted fiducials) (Brunie *et al.*, 1993; Guéziec *et al.*, 1998; Hipwell *et al.*, 2003; Livyatan *et al.*, 2003; Penney *et al.*, 1998; Russakoff *et al.*, 2003a; Tomažević *et al.*, 2004a; Vermandel *et al.*, 2003; Weese *et al.*, 1999; Weese *et al.*, 1997) or by using simulated data (Bansal *et al.*; Czopf *et al.*; Jonić *et al.*, 2001; Lavallée and Szeliski, 1995; Livyatan *et al.*, 2003; Murphy, 1997; Rohlfing *et al.*, 2002; Sarrut and Clippe, 2000). These methods all have their own drawbacks: manual methods are often subjective, and can require a large amount of user interaction. Methods which use controlled motion, independent measuring devices or fiducials usually require the use of rigidly attached frames or markers which often limits their use to phantom or cadaver studies. Simulated projection images often lack a number of characteristics of real X-rays (*e.g.* coarser resolution) and may introduce a bias as the projection

images are usually generated from the 3D volume, often using the same method as is employed by the registration algorithm.

To address the challenge of comparing 2D-3D registration algorithms, we introduce a standardized evaluation methodology that can be used for different algorithms and different applications, and that can easily be extended to a large database including multiple anatomies and patient data. The proposed methodology exploits the precalibrated relation that exists between acquired 2D X-ray projections and reconstructed 3D data in a 3DRX imaging system to generate a gold standard dataset. The motion of the 3DRX system is highly reproducible, and there is a fixed relationship between each of the X-ray projections used to reconstruct the 3DRX image and the resultant 3DRX volume. This means that the distortion corrected projections, the 3DRX volume and the calibrated projection geometry can be used as a gold standard for 2D-3D registration. This has the big advantage that no markers need to be rigidly attached to the object being imaged. Therefore, it is much easier to acquire gold standard data from patients compared to using conventional fiducial-based methods. The 2D-3D gold standard registration can be extended to other 3D modalities, such as CT or MR, by 3D-3D image registration of the CT or MR volume to the 3DRX data using established and well-validated intensity-based registration (Maes *et al.*, 1997). Our methodology also comprises a standardized procedure for generating starting positions, and appropriate definitions for error measures and failure criteria.

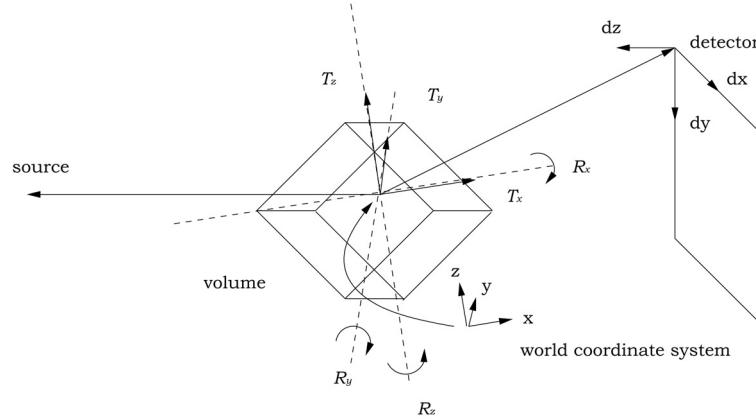
To show the usefulness of such an evaluation methodology we apply our methodology to the registration of 3DRX, CT and MR volumes of vertebral bodies to mono- and biplane X-ray projections using two registration algorithms: gradient based (Tomažević *et al.*, 2003) and intensity based (Penney *et al.*, 1998). These two approaches were selected because they have been developed for 2D-3D spine registration. From the information present in the papers it was not possible to compare them. Since the original implementations were available to us, these two methods served well in the demonstration of the standardized evaluation methodology.

## 5.2 Evaluation methodology

This section describes in detail: the generation of gold standard data, how registration accuracy was assessed, our definition of capture range and failure criteria, how starting positions were determined, and how to present the data to registration algorithms.

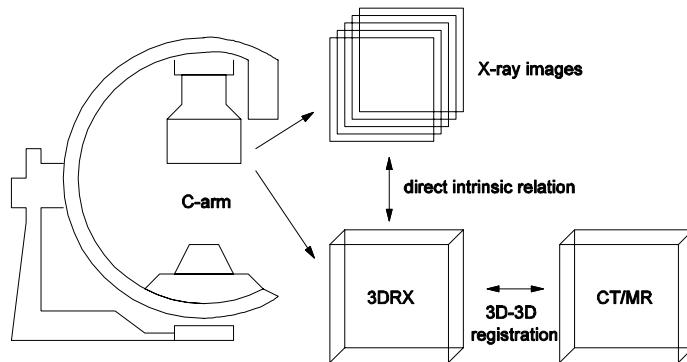
### 5.2.1 Generation of gold standard data

In the proposed evaluation methodology, gold standard data is acquired with a clinical floor-mounted 3D rotational X-ray (3DRX) C-arm system (Integris BV5000, Philips Medical Systems, Best, The Netherlands). During an



**Figure 5.1.** The six degrees of freedom ( $T_x, T_y, T_z, R_x, R_y, R_z$ ) in our rigid body 2D-3D registration problem. The intrinsic parameters, which are the source and detector positions and orientations, are known. The world coordinate system is intrinsically linked to the center of the initially reconstructed 3DRX volume, and it is completely determined by the 3DRX system calibration.

eight second run of 180 degrees around the patient the 3DRX system acquires 100 projection images, which are used to reconstruct a high resolution 3D volume using a filtered back-projection reconstruction technique (Grass *et al.*, 1999). The C-arm system is calibrated for 3DRX imaging (Koppe *et al.*, 1995) at installation. The standard 3DRX system calibration is fiducial based and the calibration is checked every six months. Two calibration runs are required. The first run determines the image intensifier distortion and the position of the focal spot of each projection with respect to the detector. The second run establishes the projection parameters (rotations around the x-, y-, and z-axis of the calibration phantom and shifts in the x, y, and z directions) for each projection. The most crucial aspect for our evaluation methodology is the accurate projection geometry calibration of the system, which has been calculated by a number of studies. The accuracy of the derived geometrical parameters has been demonstrated in (Rasche *et al.*, 2003), where the 3D image quality was evaluated in terms of spatial resolution and geometrical accuracy. It was shown that the system can obtain an almost isotropic 3D resolution of up to 22 lp/cm (line pairs per cm). Furthermore, Rasche *et al.* showed that the three-dimensional geometrical accuracy of an reconstructed volume was in the order of its spatial resolution. Using the same 3DRX system, Movassaghi *et al.* (Movassaghi *et al.*, 2004) investigated the geometrical accuracy of 3D point reconstruction in a bullet phantom from pairs of 2D projections using the calibrated projection geometry. Their results showed only a small error as a function of the projection pair used ( $0.26 \pm 0.08$  mm). The small standard deviation of 0.08 mm shows that the projection geometry produces spatially consistent volumes. Also, using the same system, Baert *et al.* (Baert *et al.*, 2004) performed guide wire reconstruction in 3D from 2D projections, achieving a high accuracy.



**Figure 5.2.** Gold standard data obtained with 3DRX C-arm. The X-ray images are used to reconstruct the 3DRX volume and they are thus directly related to the 3DRX volume. The CT and MR data are registered to the 3DRX data using 3D-3D registration, thereby obtaining the gold standard for CT/MR data to X-ray images.

Moreover, they showed that when introducing small errors into the projection geometry information, a large guidewire reconstruction error was obtained. This again indicates that the geometrical calibration process is highly accurate. As a result of the calibration, both the projection geometry of the X-ray images with respect to the reconstructed volume and the distortion of the X-ray images are known (Haaker *et al.*, 1990). Hence, the 3DRX volume and the 100 corresponding distortion corrected X-ray projections can be used as a gold standard dataset for 2D-3D registration. Figure 5.1 shows all the degrees of freedom that are known for 3DRX due to the calibration. For the 2D-3D registration, the intrinsic parameters are calculated using a calibration step and then the registration algorithms calculate the extrinsic parameters. The extrinsic parameters describe the position and orientation of the 3D volume in space (translations in x, y and z directions ( $T_x$ ,  $T_y$ ,  $T_z$ ) and rotations around x-, y- and z-axis ( $R_x$ ,  $R_y$ ,  $R_z$ )).

The gold standard can be extended to other 3D modalities (such as CT or MR) by registering corresponding CT or MR data to the 3DRX data; thereby indirectly obtaining the transformation to the X-ray images (Figure 5.2). An application based on the MIRIT (Multi-modality Image Registration using Information Theory) software by the Laboratory for Medical Imaging Research, Leuven, Belgium (Maes *et al.*, 1997) was used for 3D-3D registration. This algorithm optimizes mutual information. Voxel-based 3D-3D registration methods have been previously described in the literature and have been validated to be subvoxel accurate for certain applications (West *et al.*, 1997). In a previous study we showed that, for spine images, mutual information-based registration of MR data to 3DRX data, with an isotropic resolution of  $2 \times 2 \times 2 \text{ mm}^3$  and  $0.62 \times 0.62 \times 0.62 \text{ mm}^3$  respectively, yields a root mean squared error (RMSE) in the order of 0.8 mm at fiducial markers, and is at least as accurate as marker-based registration (van de Kraats *et al.*, 2004).

### 5.2.2 Method for assessment of 2D-3D registration accuracy

For our evaluation we introduce the following transformation matrices:  $T_{\text{gold}}$  is the  $4 \times 4$  gold standard transformation matrix from CT/MR image space to the 3D space of the 3DRX system obtained by registering the CT/MR volumes to the 3DRX volume;  $T_{\text{reg}}$  is the same transformation but obtained after having performed 2D-3D registration.  $M_{\text{gold}} = T_{\text{proj}}T_{\text{gold}}$  is the  $3 \times 4$  gold standard projection matrix from 3D space of the 3DRX system to 2D space of the X-ray projection, where  $T_{\text{proj}}$  is obtained from the system calibration, and  $M_{\text{reg}} = T_{\text{proj}}T_{\text{reg}}$  is the same projection matrix but obtained after 2D-3D registration.

Various performance measures have been used in view of different registration tasks. LaRose *et al.* (LaRose *et al.*, 2000b) categorized the measures as 2D image space error, pose parameter error and 3D registration error. 2D image space error is the error computed using 2D image space features (*e.g.* fiducials or contours), where small errors in 2D can be caused by large errors in 3D. The pose parameter error, which consists of the 6D vector ( $T_x$ ,  $T_y$ ,  $T_z$ ,  $R_x$ ,  $R_y$ ,  $R_z$ ), can give misleading results as it depends on the position of the center of rotation, because the rotational and translational components of the transformation are not independent. Lastly, the 3D registration error computes the pose position error in 3D.

Since the presented standardized evaluation methodology is meant for 2D-3D registration in general, different performance measures suited for different tasks are used.

A widely used 3D error measure is the target registration error (TRE) (Fitzpatrick and West, 2001), where the “targets” in the TRE calculation can be predefined locations (either fiducials or landmarks), surface points (Livyatan *et al.*, 2003), or arbitrary chosen points inside a region of interest. We propose to compute a mean TRE (mTRE) to determine the 3D error of a registration. For  $k$  points  $\mathbf{p}_i$  in a fixed set  $P$  of 3D points on a regular grid (*i.e.* uniformly distributed), the distance between the point  $\mathbf{p}_i$  transformed with  $T_{\text{gold}}$ , the gold standard registration, and the same point transformed with  $T_{\text{reg}}$ , the transformation determined by the registration algorithm, is computed. The average distance between the points defines the mTRE:

$$\text{mTRE}(P, T_{\text{reg}}, T_{\text{gold}}) = \frac{1}{k} \sum_{i=1}^k \|T_{\text{reg}} \mathbf{p}_i - T_{\text{gold}} \mathbf{p}_i\|. \quad (5.1)$$

The mTRE, and other errors described below, are depicted in Figure 5.3.

The 3D registration error can be best used when the application is image-to-patient registration for navigation, because it computes the error on specific points of interest within the registered volume where manipulations will take place.

However, when using a single X-ray image for assessing the 3D position of an object, determining the 3D spatial position in the projection direction is difficult. To assess the magnitude of the error in the projection direction, the mTRE in the projection direction, mTREproj, is computed. This is done by averaging the component of the displacement vector that is normal to the X-ray plane:

$$\text{mTREproj}(P, T_{\text{reg}}, T_{\text{gold}}, \hat{\mathbf{n}}) = \frac{1}{k} \sum_{i=1}^k \| (T_{\text{reg}} \mathbf{p}_i - T_{\text{gold}} \mathbf{p}_i) \cdot \hat{\mathbf{n}} \|, \quad (5.2)$$

where  $\hat{\mathbf{n}}$  is the normal to the projection plane.

In applications where 2D-3D registration is used for projecting 3D information onto a 2D image plane, 2D error measures are more appropriate than the mTRE. In this case, we use an error termed the mean projection distance (mPD) which is the error in the projection plane. More specifically, projection distance is the distance between the projection of a 3D point at registration position and the projection of the same point at gold standard position:

$$\text{mPD}(P, M_{\text{reg}}, M_{\text{gold}}) = \frac{1}{k} \sum_{i=1}^k \| M_{\text{reg}} \mathbf{p}_i - M_{\text{gold}} \mathbf{p}_i \|, \quad (5.3)$$

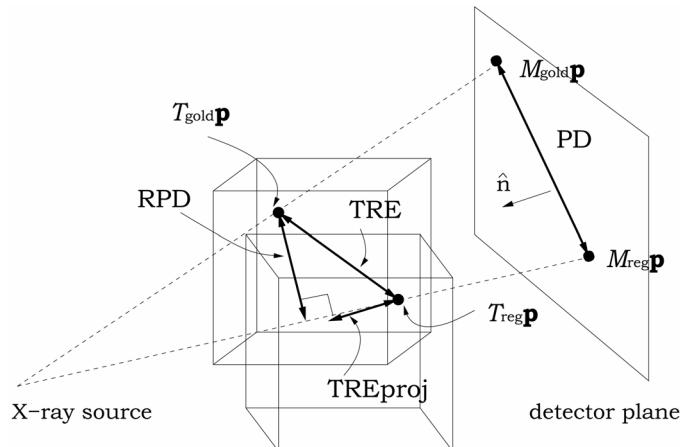
where  $M_{\text{reg}}$  and  $M_{\text{gold}}$  are the final perspective projection matrices for the registration and the gold standard position respectively, and where the distance is calculated after division by the homogeneous coordinate.

Rather than using the mean projection distance, some authors have used the reprojection distance (Masutani *et al.*, 1997), which computes the minimum distance between the line from the 2D projected point at the registered position to the X-ray source and the gold standard 3D position of that point. This line can also be described by the line through the 3D point at the registered position and the X-ray source. The mean reprojection distance (mRPD) is calculated as follows:

$$\text{mRPD}(P, M_{\text{reg}}, M_{\text{gold}}) = \frac{1}{k} \sum_{i=1}^k D(L_i(\text{source}, T_{\text{gold}} \mathbf{p}_i), T_{\text{gold}} \mathbf{p}_i). \quad (5.4)$$

Here  $D(L_i, T_{\text{gold}} \mathbf{p}_i)$  is the minimum distance between the 3D point  $\mathbf{p}_i$  at gold standard position and a line  $L_i$ , where  $L_i(\text{source}, T_{\text{gold}} \mathbf{p}_i)$  represents a 3D line through the 3D point at the registered position and the X-ray source. Unlike the mPD, the reprojection distance is independent of the distance to the projection plane.

The authors recommend that for evaluation of the performance of 2D-3D registration using a single X-ray image the mPD and mRPD are to be determined along with the more general mTRE.



**Figure 5.3.** Depiction of the error for each evaluation method displayed for a single point within the object, where TRE is the target registration error, TREproj the error in the projection direction, PD the projection distance, RPD the reprojection distance,  $T_{\text{gold}}$  the gold standard transformation,  $T_{\text{reg}}$  the registration transformation,  $M_{\text{gold}}$  the gold standard projection matrix,  $M_{\text{reg}}$  the registration projection matrix and  $\hat{\mathbf{n}}$  the normal to the projection plane.

### 5.2.3 Capture range and failure criteria

The capture range defines the range of starting positions from which an algorithm finds a sufficiently accurate transformation. As shown in the previous section, accuracy can be determined using various error measures. Starting ranges can also be determined using these error measures, describing the starting range as the initial accuracy from where registrations are still successful. Although the initial accuracy can be evaluated using different error measures, we always determine the capture range using the initial mTRE, which enables direct comparison of reported capture ranges. Moreover, the initial mTRE is most informative about the initial misalignment. Since the capture range naturally starts from zero, it is reported by stating the maximum start mTRE value of this range. Two factors are involved to calculate the capture range: the definition of a misregistration (using either mTRE, mPD or mRPD), and the fraction of allowed misregistrations. Both these factors depend on the application of the 2D-3D registration method.

### 5.2.4 Starting positions

Several approaches are possible for determining the initial distance to the gold standard of a starting position. Tomažević *et al.* (Tomažević *et al.*, 2003) use a “normalized distance”, that relates rotations to linear distances. Rohlfing *et al.* (Rohlfing *et al.*, 2002) use the effective motion in the projection plane

per parameter for normalization. We choose to use the mTRE at the starting position as a measure for the initial distance to the gold standard, similar to Russakoff *et al.* (Russakoff *et al.*, 2003a). The mTRE is determined over a fixed set of 3D points uniformly distributed in an average-sized region (which can differ per anatomy). By using the same quantity for initial distance and final result, the effect of the algorithm is directly measurable. Of course, the effect on the end mPD or end mRPD is not directly measurable when plotting them against the start mTRE.

The starting positions are offsets from the gold standard position. Given our definition of capture range, and the use of the initial mTRE as the distance measure for the starting positions, the starting parameter values ( $T_x$ ,  $T_y$ ,  $T_z$ ,  $R_x$ ,  $R_y$ ,  $R_z$ ) must be generated such that several resulting transformations are within several ranges of mTRE. This is achieved by:

1. Choosing intervals for the starting position distance, *e.g.* 0–1, 1–2 mm, and determining the center of rotation, which generally is the center of the fixed set of 3D points.
2. Then, for each interval, for each of the six transformation parameters, the range is determined that will yield an mTRE less than or equal to the interval upper bound, *e.g.* for a starting mTRE between 1 and 2 mm, the range for each translation is 0–2 mm. Since the region used to determine the starting mTREs is not necessarily cubic, rotations around different axes may not have the same effect on the mTRE. This is taken into account in the following way: The angle of rotation that results in an mTRE of 1 mm for each individual rotation is calculated. This angle is then linearly scaled to determine the maximum allowed rotation for the other intervals. In this calculation the small angle approximation is used, *i.e.* the effective displacement is linearly related to the angle.
3. Next, for each interval, transformations are generated, where the transformation parameters are chosen randomly (uniformly distributed) from their predetermined range. Subsequently, the mTRE of the composite transformation is determined, and if that mTRE is within the interval it is kept, otherwise the transformation is discarded.

The last step is repeated until each interval contains a predetermined number of starting positions.

### 5.2.5 Data preparation

Most registration algorithms only use a part of the available image data, to focus the registration on the object of interest and to reduce the amount of data, resulting in lower computational costs. Since some algorithms reduce data by using regions of interest (ROIs) in the 2D projection images and others

**Table 5.1.** Parameter values for the gradient-based registration method.

Parameter	MR	3DRX	CT
( $\sigma$ = Blur volume)	0.3	0.5	0.3
( $\sigma$ = Blur X-ray)	0.5 / 1.0	0.5 / 1.0	0.5 / 1.0
SampleStep	1.0 mm	1.0 mm	1.0 mm
Threshold	10	25	18

**Table 5.2.** Parameter values for the intensity-based registration method (HU = Hounsfield unit).

Parameter	3DRX	CT
MinStep	0.125	0.125
Threshold	13000 / 11000	7000 ( $\approx$ 400 HU)

by using volumes of interest (VOIs) in the 3D volume, it is not possible to have exactly the same data input for both types of algorithm. For the presented evaluation methodology the VOIs are manually determined in the 3DRX volume, by determining a rectangular region of interest around the anatomy. The VOIs should at least contain the anatomy that is used and relevant for the registration. When performing inter-modality comparison the VOIs are transformed to the corresponding CT or MR dataset using the gold standard image-based 3D-3D registration. Rectangular ROIs are also manually determined in the X-ray image so that they contain approximately the same anatomy as is contained in the VOIs. However, the ROIs have no direct relation to the VOIs in the 3D volumes. Corresponding ROIs could have been determined in the X-ray images by projecting the VOIs in the 3D volume to the 2D image using the gold standard calibration. However, the effect on the registration when using slightly different ROIs is expected to be minimal. Moreover, for algorithms that have both a data reduction in 2D and in 3D directly related VOIs and ROIs should not be used as this would bias the registration.

### 5.3 Experiment 1: Comparison of two methods

To illustrate the potential of the proposed evaluation framework, it has been applied to compare two previously published 2D-3D registration methods, *viz.* a gradient-based and an intensity-based method. For both algorithms the original implementations were available, though the intensity-based algorithm had to be adapted so it could take multiple X-ray images as input.

**Table 5.3.** Summary of image dimensions and voxel sizes for the images acquired of two spine segments using four different modalities.

Modality	Segment	Resolution	Size
MR	1	$1.00 \times 0.75 \times 0.75 \text{ mm}^3$	$100 \times 256 \times 256$
	2	$1.00 \times 0.88 \times 0.88 \text{ mm}^3$	$120 \times 256 \times 256$
CT	1	$0.31 \times 0.49 \times 0.31 \text{ mm}^3$	$320 \times 260 \times 320$
	2	$0.31 \times 0.49 \times 0.31 \text{ mm}^3$	$280 \times 300 \times 300$
3DRX	1	$0.87 \times 0.87 \times 0.87 \text{ mm}^3$	$256^3$
	2	$0.52 \times 0.52 \times 0.52 \text{ mm}^3$	$256^3$
X-ray	1	$0.63 \times 0.63 \text{ mm}^2$	$512^2$
	2	$0.53 \times 0.53 \text{ mm}^2$	$512^2$

The gradient-based 2D-3D registration method (Tomažević *et al.*, 2003) registers 3D volumes to 2D images by searching for the best match between surface normals in the 3D volume and back-projected gradients from the 2D X-ray images. The surface normals and their positions are extracted from the 3D volume in a preprocessing step. In this preprocessing step the volume is blurred with a Gaussian filter, isotropically resampled, and the locations with a gradient magnitude larger than a predefined threshold are extracted. The gradient-based registration method needs the following parameters: volume blur, X-ray image blur, sample step and a gradient threshold. Although the original authors have only used their method for registration with multiple X-ray images, we also applied it in registrations using only one X-ray image. See Table 5.1 for the parameter values used in this evaluation. The parameters were chosen such that correct points were extracted in the 3D volume, which required some tuning especially for the MR volume.

The intensity-based registration method (Penney *et al.*, 1998) compares digitally reconstructed radiographs (DRRs or simulated projections) to X-ray images using a similarity measure called gradient difference. The algorithm has two parameters: bone threshold and minimal step-size. The bone threshold is manually determined so that only voxels containing bone contribute to the DRRs. The minimal step-size is the smallest step that the algorithm takes in the optimization procedure, which is a factor by which the position parameter values can change in an iteration. The rotations, which are expressed in radians, need an extra factor of 0.02 to make the steps comparable to the translation steps, which are in millimeters. The optimization uses a multi-resolution approach, where initially coarse resolution  $256 \times 256$  pixel DRR images are generated, and then finer  $512 \times 512$  pixel images. See Table 5.2 for a summary of the parameter values used, which were considered typical.

To select the gradients in the volume of interest, *i.e.* at the bone interface, an interactive windowing step is required for the gradient-based method when using 3DRX and MR data. In order to make the evaluation comparable the same data was used for the intensity-based method. Both methods were directly applied to the CT data, whose intensity range was converted to comply with the input criteria of the methods.

### 5.3.1 Image data

Two defrosted segments of a spinal column, both fixed in foam, were imaged with three modalities: 3DRX, CT and MR. The 3DRX images were reconstructed using 100 projection images acquired with a clinical 3DRX system. As mentioned before, the geometric relation and distortion correction for each of these images are known from the standard 3DRX system calibration. Spinal segment 1 was obtained with an image intensifier II size of 38 cm and spinal segment 2 with an II size of 31 cm. The CT volumes were acquired with a clinical multi-slice CT scanner (MX8000, IDT 16, Philips Medical Systems). The MR volumes were acquired with a clinical 1.5-Tesla MR scanner (Gyroscan NT, Philips Medical Systems) using a sagittal 3D turbo spin echo acquisition, with a turbo factor of 29, TR/TE of 1500 ms/90 ms.

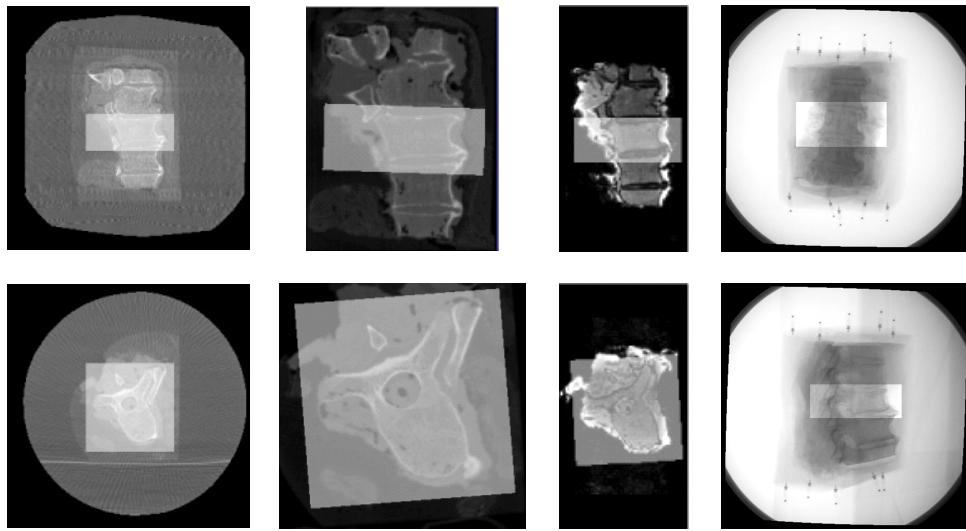
Spinal segment 1 consists of 3 thoracolumbal vertebral bodies and segment 2 consists of 5 thoracic vertebral bodies. Some soft tissue around the spinal segments was still present. See Table 5.3 for details on image sizes and resolutions.

The CT and MR volumes were registered to the corresponding 3DRX volumes to establish the gold standard for both CT-to-X-ray registration and MR-to-X-ray registration using the previously described maximization of mutual information (Collignon *et al.*, 1995; Viola and Wells III, 1997).

The rectangular VOIs used by the gradient-based method were manually determined in the 3DRX volumes for each vertebra. The size of these VOIs was approximately 100 x 50 x 100 voxels in segment 1 and 120 x 50 x 150 voxels in segment 2. The VOIs were transformed to the CT and MR data using the gold standard image-based 3D-3D registration. In the X-ray images the rectangular ROIs that were used by the intensity-based method were manually determined around the vertebral bodies (approximately 170 x 70 pixels). See Figure 5.4 for examples of the data with corresponding VOIs and ROIs. For each of the eight vertebral bodies, centers of rotation were determined in the world coordinate system (directly linked to the 3DRX data).

### 5.3.2 Experiments and evaluation

The starting positions were generated as described in Section 5.2, with 10 starting positions per 1 mm interval up to 15 mm mTRE, resulting in 150



**Figure 5.4.** Top row from left to right: Coronal planes with VOI for 3DRX, CT, and MR data, and an anterior-posterior X-ray image with ROI. Bottom row from left to right: Transversal planes with VOI for 3DRX, CT, and MR data, and a lateral X-ray image with ROI.

starting positions. Ten starting positions per bin were chosen because we considered 80 (10 times 8 vertebral bodies) registrations per 1 mm bin sufficient to assess the capture range. In this starting position generation process, the mTRE was determined over a fixed set of 3D points in an average vertebra-sized region of  $95 \times 45 \times 95$  voxels (size  $0.87^3$  mm $^3$ ) (as described in Section 5.2). For the accuracy evaluation this fixed grid was also used. The same starting positions (offsets from the gold standard) were used for each vertebra.

From the 100 acquired X-ray images per spinal segment two images were selected, one anterior-posterior (AP) and one lateral (LAT), which were used for the registration experiments for both 3DRX, CT, and MR to X-ray registration. For 3DRX and CT, we also performed experiments using only one X-ray image, which was AP for spinal segment 1 and LAT for spinal segment 2.

In this example evaluation, registrations were classified as successful when the end mTRE over the fixed set of 3D points (same as was used for the generation of the starting positions) was smaller than 2 mm. This threshold was also used in (Tomažević *et al.*, 2003). It is however application specific as the required accuracy for spine procedures varies from 0.0 mm to 3.8 mm (Rampersaud *et al.*, 2001) in different levels of the spine. The capture range was defined as the 95% success range. When using a single X-ray image the accuracy evaluation was extended to include end mPD and end mRPD, where

success was also defined as an error smaller than 2 mm to illustrate the effect of different accuracy measures on the observed performance of an algorithm.

## 5.4 Experiment 2: Evaluation of the gold standard

The results of the evaluation depend on the accuracy of the gold standard. Earlier studies (Baert *et al.*, 2004; Movassaghi *et al.*, 2004) as discussed in Section 5.2.1 indicated that the calibration is very accurate. Previous investigation shows that errors in multi-modality 3D-3D registration can be in the order of the voxel size of the images (West *et al.*, 1997). To investigate how errors in the calibration and in the 3D-3D registration propagate into 2D-3D registration errors, a number of experiments was performed.

Errors in the calibration were simulated by adding 0.5° and 1.0° to the rotation angle (roll direction of C-arm), and 0.5° and 1.0° to the angulation angle (tilt direction of C-arm) of the calibrated projection geometry determined in the standard system calibration. These errors are larger than the expected gold standard error, as has been shown in previous studies (Baert *et al.*, 2004).

Errors in the 3D-3D registration were simulated by adding random transformations to the gold standard that introduce an error of up to 1 mm mTRE (which is approximately the voxel size). These error transformations were generated in a similar fashion as the 2D-3D starting positions. Ten 3D-3D error transformations were produced per bin of size 0.2 mm.

For each of these specific simulated errors (both calibration and 3D-3D registration errors) 100 registrations for one vertebral body in spinal segment 2 were performed using the gradient-based method. For the calibration error experiments 3DRX was registered to two X-ray projections, and for the 3D-3D registration error experiments CT was registered to two X-ray projections. The average registration errors (mean mTRE) for successful registrations per bin were determined and compared to the mean mTRE when using the real gold standard.

## 5.5 Results and discussion

### 5.5.1 Experiment 1: Comparison of two methods

We performed 150 registrations for each of the eight vertebral bodies, for each modality and for a different number of X-ray images (either single plane or using two X-ray images simultaneously). The following registrations were carried out: 3DRX to one X-ray, 3DRX to two X-rays, CT to one X-ray, CT to two X-rays, and MR to two X-rays. For two X-ray images, the average

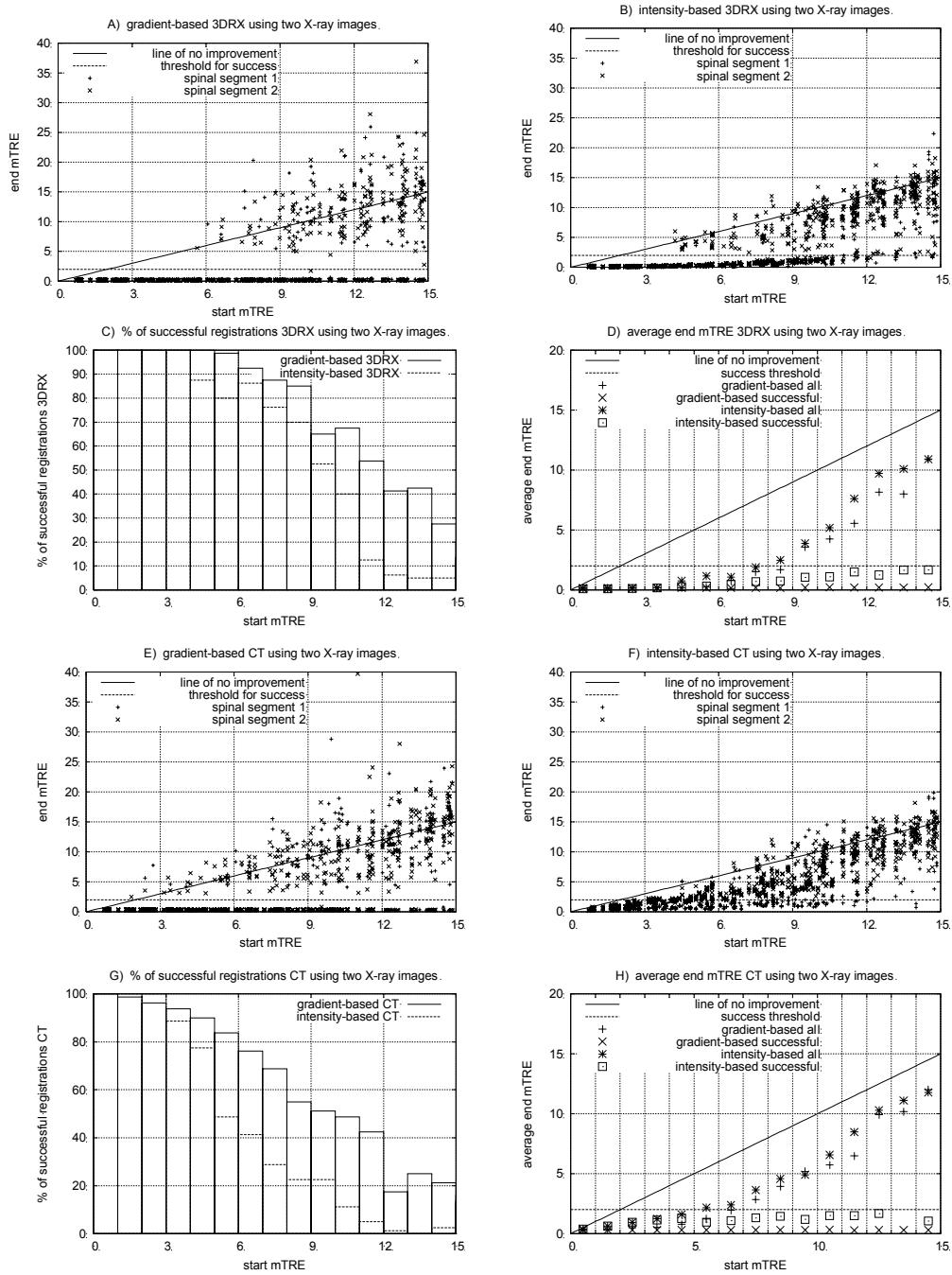
registration time was approximately 25 seconds for the gradient-based method, running on a Windows 2000 Dell Workstation PWS340 Intel Pentium 4, 1.7 GHz, 1.6 GB RAM, and approximately 9 minutes for the intensity-based method, running on a multi-user Linux Dell PowerEdge 1600, dual Xeon 2.8 GHz with 4 GB memory. The algorithms were not optimized for speed.

The results for registrations using two 2D X-ray images are presented in Figure 5.5 (for CT and 3DRX) and in Figure 5.6 (for MR). Figure 5.7 shows the results of registering CT and 3DRX to a single 2D X-ray image. Each of these figures presents the registration results for both algorithms, except for Figure 5.6, because only the gradient-based algorithm was used on MR. The intensity-based method was not performed on the MR dataset, as no DRRs can be generated from MR images. The end mTRE values for each registration, the percentage of successful registrations, the average end mTRE per bin for all registrations and for successful registrations are displayed. These evaluations are extended for the registrations based on a single X-ray image by also calculating end mTREproj (Figure 5.8) and end mPD and end mRPD (Figure 5.9).

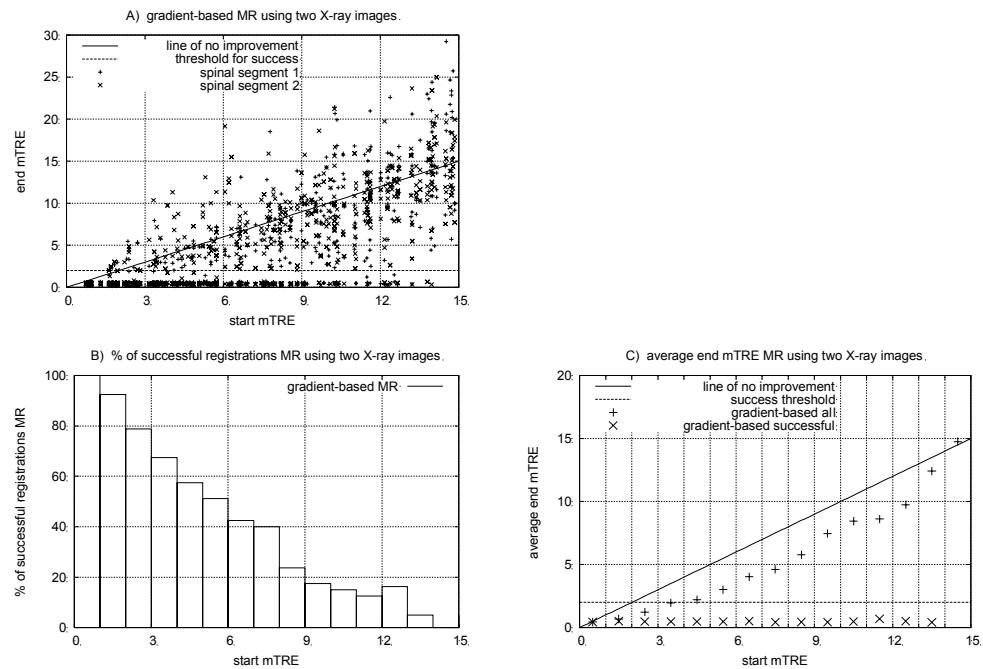
From the results for registrations performed using two X-ray images several conclusions can be drawn. The average error for successful registrations is stable for the gradient-based method, while the intensity-based method has increasing difficulty in finding the optimal position when the initial offset from the gold standard position increases (Figures 8.5D,H). Failure determination is most probably easier for the gradient-based method (Figures 8.5A,E) than for the intensity-based method (Figures 8.5B,F), because the gradient-based registration seems to be either correct or incorrect. In a clinical situation, a failure can be visually assessed by displaying the DRR belonging to the registration next to (or on top of) the X-ray image. For both methods the average error is larger for CT to X-ray registration than for 3DRX to X-ray registration.

The capture ranges for 3DRX and CT, respectively, end at 6 mm and 3 mm for the gradient-based method and at 4 mm and 3 mm for the intensity-based method (Figures 8.5C,G). Within the capture range the average end mTREs of successful registrations for 3DRX and CT respectively are 0.18 mm and 0.30 mm for the gradient-based method and 0.13 mm and 0.65 mm for the intensity-based method.

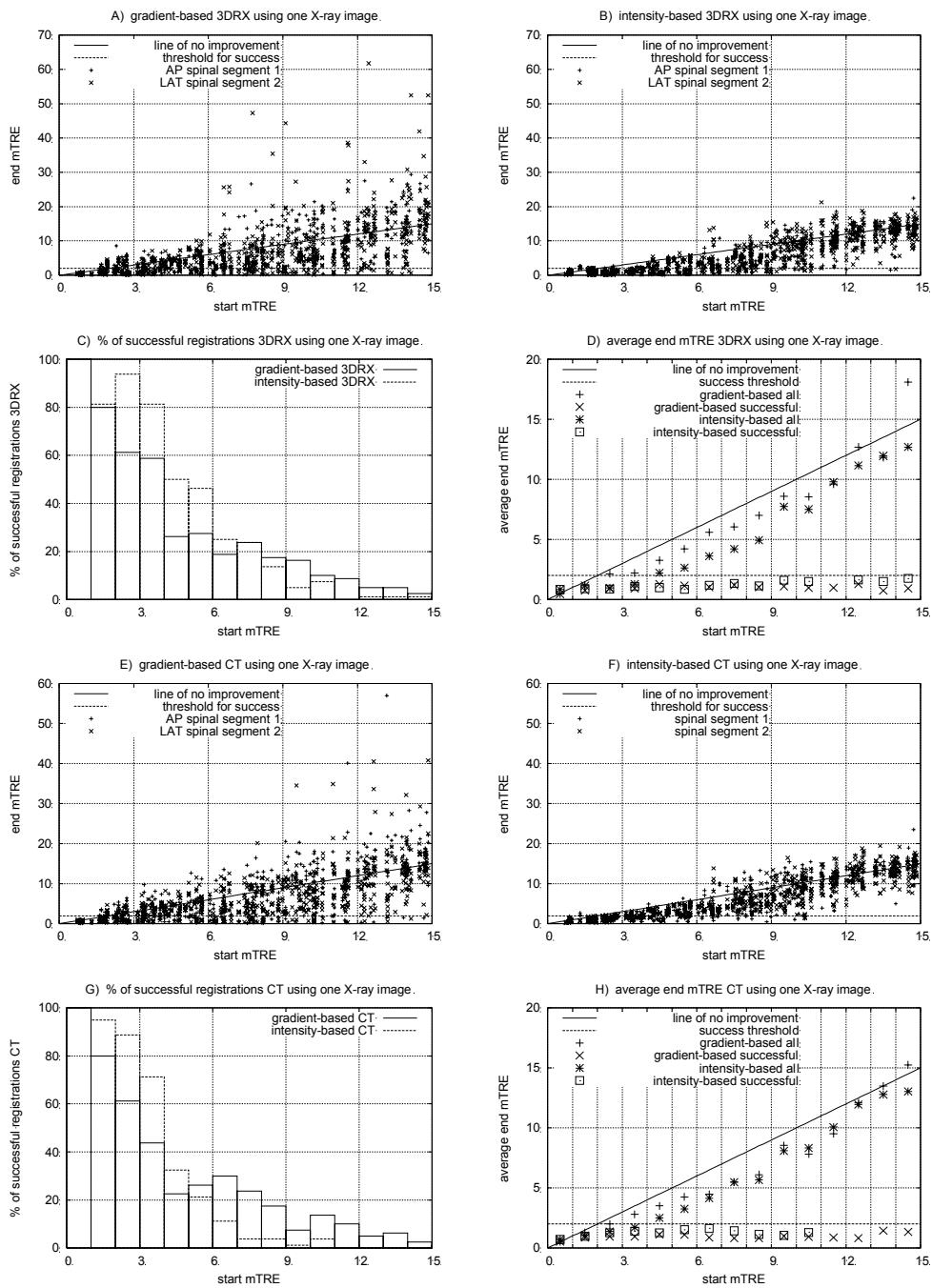
Better results for 3DRX to X-ray registration could be explained by greater similarity between the 2D and 3D modalities: both X-ray images and 3D volume are produced using X-rays at the same energy (around 60 keV), whereas CT uses 80 keV, and the 3DRX volume is reconstructed using the X-ray projections. Another reason could be the difference between the gold standard for 3DRX and the gold standard for CT and MR, where the 3D-3D registration could have introduced additional errors.



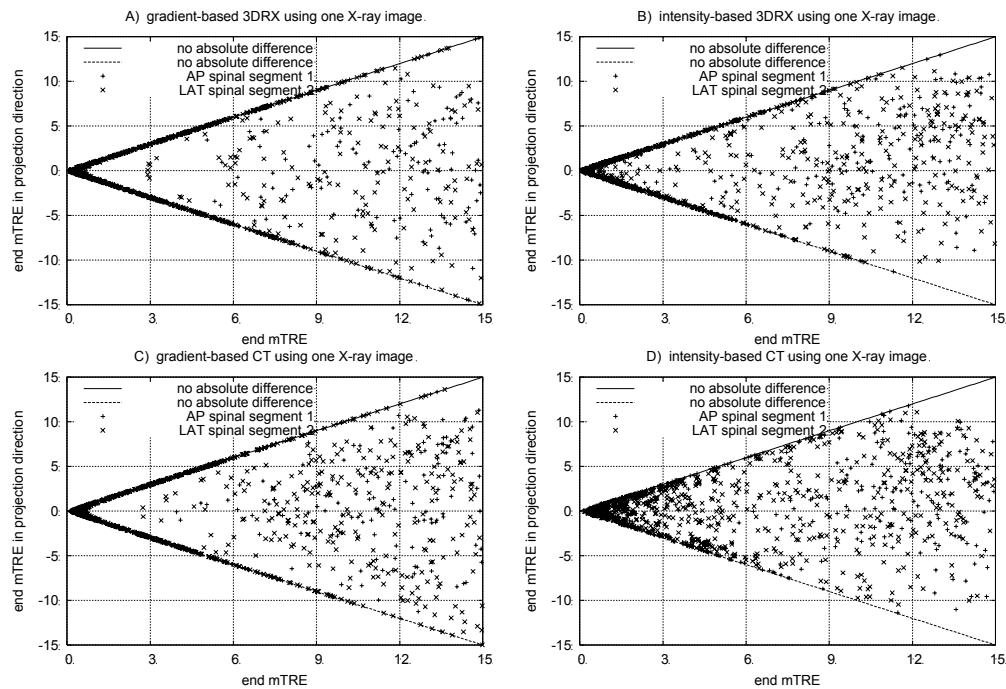
**Figure 5.5.** Results for 3DRX and CT to X-ray registration using two X-ray images. For brevity, all 450 registrations ( $150 \times 3$  vertebral bodies) and all 750 registrations ( $150 \times 5$  vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.



**Figure 5.6.** Results for MR to X-ray registration using the gradient-based method using two X-ray images. For brevity, all 450 registrations ( $150 \times 3$  vertebral bodies) and all 750 registrations ( $150 \times 5$  vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.



**Figure 5.7.** Results for 3DRX and CT to X-ray registration using one X-ray image. For brevity, all 450 registrations ( $150 \times 3$  vertebral bodies) and all 750 registrations ( $150 \times 5$  vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.



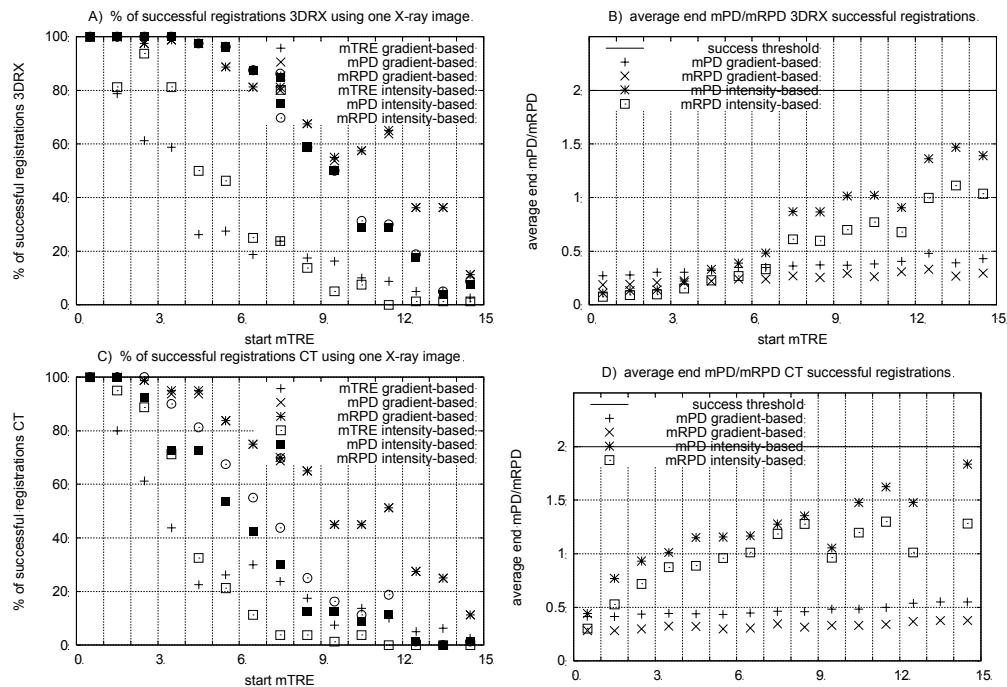
**Figure 5.8.** Mean TREs in the projection direction. For brevity, all 450 registrations ( $150 \times 3$  vertebral bodies) and all 750 registrations ( $150 \times 5$  vertebral bodies) for both spinal column segments were labeled as spinal segment 1 and spinal segment 2, respectively.

Results for MR to two X-ray image registration are displayed in Figure 5.6 for the gradient-based method. Since the gradient-based method is intended for more than one X-ray image, and since MR to X-ray registration is quite challenging, we decided to only perform MR to X-ray registration using two X-ray images. The capture range ends at 1 mm and within this range the average end mTRE of success is 0.45 mm. It must be noted that the gradient-based registration was not optimized for this MR protocol.

Results for the registrations based on one X-ray image are displayed in Figure 5.7. The capture ranges for 3DRX and CT respectively end at 1 mm and 1 mm for the gradient-based method and 1 mm and 2 mm for the intensity-based method (Figures 5.7C,G). Within the capture range the average end mTREs of successful registrations for 3DRX and CT respectively are 0.46 mm and 0.53 mm for the gradient-based method and 0.84 mm and 0.84 mm for the intensity-based method (Figures 5.7D,H). The intensity-based method seems to perform better (higher capture range) on one X-ray image than the gradient-based method. The results for both methods are significantly worse than when using two X-ray images. The main reason is difficulty in estimating the correct position of the 3D volume in the projection direction.

In order to assess whether the largest part of the error was in the direction of the normal to the projection plane, the end mTRE<sub>proj</sub> was calculated as explained in Section 5.2.2. In Figure 5.8 the end mTRE in the projection direction is plotted as a function of the total end mTRE for registrations based on one X-ray image. From the graphs in Figure 5.8 we can conclude that especially for small end mTREs (thus accurate registrations) the major part of the total end mTRE can be explained by the end mTRE in the projection direction only.

In order to quantify the in-plane error and the reprojection distance, the end mPD and end mRPD were also calculated for both registration methods for registrations using one X-ray image (as explained in Section 5.2.2). The results are displayed in Figure 5.9, where results for spinal segments 1 and 2 are combined and the end mTRE is shown as a reference. Only the averages per bin of the successful registrations were displayed to show the difference between the error measures more clearly. For the mPD, the 3DRX capture ranges for the gradient-based method and the intensity-based method respectively end at: 5 mm (0.29 mm average mPD for success) and 6 mm (0.22 mm average mPD for success), the CT capture ranges respectively end at: 3 mm (0.42 mm average mPD for success) and 2 mm (0.61 mm average mPD for success). For the mRPD, the 3DRX capture ranges for the gradient-based method and the intensity-based method respectively end at: 5 mm (0.20 mm average mRPD for success) and 6 mm (0.15 mm average mRPD for success), the CT capture ranges respectively end at: 5 mm (0.30 mm average mRPD for success) and 3 mm (0.52 mm average mRPD for success). The results show that the capture ranges are much larger when the error in the projection direction is not relevant. The most appropriate error measure to use depends on the application.



**Figure 5.9.** Mean projection distance (mPD), mean reprojection distance (mRPD), and mean target registration error (mTRE) for registrations for 3DRX and CT based on a single X-ray image.

**Table 5.4.** Summary of the results of all registration experiments. `GB` represents the gradient-based method and `IB` the intensity-based method, `One` and `Two` represent one X-ray image and two X-ray images respectively, and `E` and `C.R.` are abbreviations for error and capture range.

			3DRX		CT	
			E(mm)	C.R. (mm)	E(mm)	C.R. (mm)
GB	Two	mTRE	0.18	6	0.30	3
		mTRE	0.46	1	0.53	1
	One	mPD	0.29	5	0.42	3
		mRPD	0.20	5	0.30	5
IB	Two	mTRE	0.13	4	0.65	3
		mTRE	0.84	1	0.84	2
	One	mPD	0.22	6	0.61	2
		mRPD	0.15	6	0.52	3

Table 5.4 summarizes the main results for all 3DRX and CT experiments. The main results for MR to two X-ray projections experiments for the gradient-based method were a capture range of 1 mm with an average error for successful registrations of 0.45 mm. The evaluation method reveals that for both methods the capture ranges are limited. This could be a problem for fully automated 2D-3D registration. Manual initialization may be required to be within this capture range at the start of registration. By superimposing bony edge points in CT or MR on X-ray images, an operator can find an initial registration. The achievable accuracy of manual initial registration should be investigated further. The registration method may also be used in situations where despite measures to limit movement, some motion has occurred. In a previous study (Byrne *et al.*, 2004) it was shown that in many interventions the intraoperative movement is limited, which suggests that in these situations no further initialization is required. Nevertheless, the limited capture range remains a potential problem.

It has been shown that the capture range depends strongly on the evaluation measure used, and this in turn depends on the application. When the objective is to find the 3D position of an object, we recommend the use of more than one projection and evaluation using the mTRE. When the objective is to display 3D information in 2D, one projection is sufficient, and the performance can be adequately evaluated using the mPD or the mRPD. For such applications a larger capture range will be obtained.

The evaluation also reveals that MR to X-ray registration exhibited larger errors and a smaller capture range than 3DRX and CT based registration. This seems plausible as it is much more difficult to relate the intensities of MR data and X-ray images. Many more graphs and conclusions can be drawn from the

generated data; this chapter, however, only shows the potential of the presented evaluation methodology.

### 5.5.2 Experiment 2: Evaluation of the gold standard

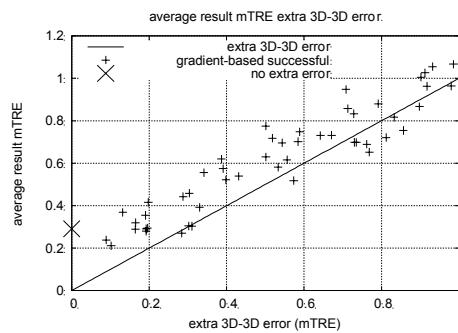
The obtained results depend on the gold standard which in turn depends on the calibrated geometry of the system and the 3D-3D registration (for MR and CT). As the reported mean errors are small (end mTRE) for successful registrations over a large number of experiments, there is high confidence in the accuracy of the gold standard. However, to investigate the influence of errors in the geometry calibration and 3D-3D registration on the evaluation, additional experiments were performed by determining the mTRE for a modified gold standard, as described in Section 5.4. The results of these experiments are displayed in Table 5.5 and Figure 5.10.

The error for successful registrations on vertebral body 1 of segment 2 was on average 0.11 mm when using our gold standard (with no additional errors) for registration of 3DRX to two X-ray projections. Extra geometry calibration errors increase the resulting mTREs for successful registrations significantly (*e.g.* a change of 0.5° angulation already increases the mTRE from 0.11 mm to 0.51 mm). Thus it can be concluded that the calibration is very accurate, and that accurate calibration is essential. The error for successful registrations on vertebral body 1 of segment 2 was on average 0.29 mm when using our gold standard (with no additional errors) for registration of CT to two X-ray images. As expected, the addition of an error in 3D-3D registration decreases the accuracy of the registration. It can be observed that even for small extra errors in 3D-3D registration, the calculated 2D-3D registration error significantly increases. For example, for an extra 3D-3D mTRE error of 0.50 mm the 2D-3D error increases from 0.29 mm to 0.78 mm. This indicates that the initial 3D-3D registration of CT to 3DRX was fairly accurate (less than a voxel). The extra error does not have a significant effect on the capture range, since the maximum extra 3D-3D error is 1 mm, and so it can only decrease the capture range by a maximum of 1 mm.

In supplement to these experiments we would like to make an additional remark. The error in the gold standard consists of projection calibration errors, distortion correction errors, and 3D-3D registration errors. Projection calibration and distortion calculation are performed separately. 3D-3D registration is performed on whole segments whereas 2D-3D registration is performed on each vertebra separately. Therefore, the errors in 2D-3D registration and the errors in the gold standard should be independent. Moreover, the two evaluated 2D-3D registration methods use different optimization techniques, similarity criteria, and either ROIs or VOIs. In the case of independent errors, the total error is the root of the sum of squared individual errors. Together with the fact that the errors reported in Section 5.5.1 are small, these assumptions give us confidence that the gold standard is accurate.

**Table 5.5.** Average registration errors (mean mTRE) for successful registrations when extra errors are introduced to the calibrated projection geometry. Registrations were performed for 3DRX to two X-ray projections. The original gold standard error was 0.11 mm.

	Extra calibration errors				
	0	Ang. 0.5°	Ang. 1.0°	Rot. 0.5°	Rot. 1.0°
Mean mTRE (mm)	0.11	0.51	0.94	0.36	0.66



**Figure 5.10.** Graph displaying the average result mTRE for successful registrations when an extra error is introduced in the 3D-3D registration. The registrations were performed for CT to two X-ray projections using the gradient-based method. The original gold standard error was 0.29 mm.

## 5.6 Conclusions

A standardized evaluation methodology for objective inter-method and inter-modality comparison of 2D-3D registration methods has been introduced. Objective and systematic comparison is facilitated by providing a gold standard and using the same datasets, starting positions, regions of interest, centers of rotation, error measures, capture range and failure criteria. Also, by measuring the error in the starting position and the registration error in the same way, the effect of an algorithm can be clearly seen.

The calibration procedure used in generating the gold standard is very accurate and the 3D-3D registration necessary to obtain the gold standard for CT/MR only appears to introduce subvoxel errors. Our experiments showed that the total error in the gold standard (which consists of a calibration error and a 3D-3D registration error) is similar to the error in 3D-3D registration.

We demonstrated the potential of the described evaluation methodology by comparing two 2D-3D registration methods. Not surprisingly, the results show that the performance of each method is significantly better when using two X-ray images, than when using only one X-ray image for finding the 3D position

of an object (quantified by the mTRE). Even for registrations based on two X-ray images the capture range of the algorithms is limited, but for successful registrations the attained accuracy is quite high (0.18 mm mTRE and 0.30 mm mTRE for 3DRX and CT respectively with the gradient-based method and 0.13 mm mTRE and 0.65 mm mTRE for 3DRX and CT respectively with the intensity-based method). For registrations based on one X-ray image, the capture range of the intensity-based method is larger than for the gradient-based method, but the gradient-based method has a larger capture range when using two X-ray images.

A possible extension of the evaluation method is to investigate how the capture range is influenced by different choices of the parameters for each method, and thereby it can help in parameter tuning. The effect of a different success threshold on the capture range can also be assessed, and another extension is to evaluate the performance of registration methods when using more than two X-ray images, since more X-ray images are available.

The 3DRX, CT, MR and X-ray data, together with gold standard transformations, starting positions, ROIs, VOIs, centers of rotation, and evaluation criteria as written in this chapter are available to other researchers. Since the described methodology uses a calibrated 3DRX system, extension to other anatomies which have high contrast on X-ray (*e.g.* bone) is straightforward and this data as well as patient data can be included in the database. In the field of rigid 3D-3D registration, the availability of the Vanderbilt dataset has already shown the importance of a common database, validation statistics, and error measure for comparison of multiple algorithms (West *et al.*, 1997). Similarly, our aim is to create a standardized dataset for 2D-3D registration which can be used for future evaluations by other researchers.



Discovery consists of seeing what everybody has seen and thinking what nobody has thought.

*Albert von Szent-Gyorgyi*

## Chapter 6

# 3D/2D registration by integrating 2D information in 3D

Dejan Tomažević, Boštjan Likar and Franjo Pernuš  
IEEE Transactions on Medical Imaging 25: 17-27 (2006)

### Abstract

*In image-guided therapy, high quality preoperative images serve for planning and simulation, and intraoperatively as “background”, onto which models of surgical instruments or radiation beams are projected. The link between a preoperative image and intraoperative physical space of the patient is established by image-to-patient registration. In this paper we present a novel 3D/2D registration method. First, a 3D image is reconstructed from a few 2D X-ray images and next, the preoperative 3D image is brought into the best possible spatial correspondence with the reconstructed image by optimizing a similarity measure. Because the quality of the reconstructed image is generally low, we introduce a novel similarity measure, which is able to cope with low image quality as well as with different imaging modalities. The novel 3D/2D registration method has been evaluated and compared to the gradient-based method (Tomažević et al., 2003) using standardized evaluation methodology and publicly available 3D CT, 3DRX, and MR and 2D X-ray images of two spine phantoms, for which gold standard registrations were known. For each of the 3DRX, CT, or MR images and each set of X-ray images, 1600 registrations were performed from starting positions, defined as the mean target registration error (mTRE), randomly generated and uniformly distributed in the interval of 0-20 mm around the gold standard. The capture range was defined as the distance from gold standard for which the final TRE was less than 2 mm in at least 95% of all cases. In terms of success rate as the function of initial misalignment and*

capture range the proposed method outperformed the gradient-based method. TREs of the novel method and the gradient-based method were approximately the same. For the registration of 3DRX and CT images to X-ray images as few as 2-3 X-ray views were sufficient to obtain approximately 0.4 mm TREs, 7-9 mm capture range, and 80-90% of successful registrations. To obtain similar results for MR to X-ray registrations, an image, reconstructed from at least 11 X-ray images was required. Reconstructions from more than 11 images had no effect on the registration results.

## 6.1 Introduction

For many years, medical imaging has been essential for medical diagnosis, treatment planning, and monitoring the progress of disease or results of treatment. More recently, images have been used to guide minimally invasive procedures. In image-guided therapy, raw preoperative three-dimensional (3D) computed tomography (CT) or magnetic resonance (MR) images and models of anatomical structures, obtained by image segmentation, serve for preoperative planning and simulation, and for “background” shown on the monitor in the treatment room onto which models of surgical instruments or radiation beams are projected. The link between preoperative images and intraoperative physical space of the patient is established by image-to-patient registration. Registration may be carried out either by registering the preoperative image directly to the patient or to one or more intraoperative 3D or two-dimensional (2D) images of the patient. Registration allows any 3D point defined in the patient (treatment room) coordinate system to be precisely located in the preoperative image coordinate system. Correct registration is a prerequisite for presenting the physician with valuable 3D information about the position of instruments (radiation beams) relative to the planned trajectory, nearby vulnerable structures, and the ultimate target, which is essential for performing less invasive and more accurate procedures. Different medical treatment procedures may benefit from registration-based image guidance. One of them is radiotherapy, where the exact position of the anatomy that is to be treated is crucial to maximize the results of therapy and, at the same time, minimize the damage of radiation to healthy tissues (Bijhold, 1993; Clippe *et al.*, 2003; Gilhuijs *et al.*, 1996a; Murphy, 1997). Other possible areas are neurointervention (Feldmar *et al.*, 1997b; Hipwell *et al.*, 2003; Kita *et al.*, 1998), spinal surgery (Lavallée, 1995; Weese *et al.*, 1997), hip replacement (Guéziec *et al.*, 1998), and all those minimally invasive interventions where the surgeon's view of the anatomy is restricted by small incisions.

In this paper we concentrate on image registration of one 3D rotational X-ray (3DRX), CT, or MR image to a set of two or more intraoperative 2D X-ray images. In the last decade, different 3D/2D image registration methods have been proposed. They can be roughly divided into segmentation-based (Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Guéziec *et al.*, 2000; Hamadeh *et al.*, 1998; Lavallée and Szeliski, 1995), intensity-based (Hipwell *et al.*, 2003; Lemieux *et al.*, 1994; Penney *et al.*, 1998; Weese *et al.*, 1997), and hybrid (Gilhuijs *et al.*, 1996a; Livyatan *et al.*, 2003; Tomažević *et al.*, 2003). A segmentation-based

3D/2D registration tries to find such a transformation between images which minimizes the spatial distance between positions of corresponding geometrical features that have previously been extracted from both, the pre- and intraoperative images. The nature of segmented geometrical features may be *extrinsic*, like positions of fiducial markers attached to the patient (Russakoff *et al.*, 2003a). Fiducial markers are designed in such a way to be able to accurately segment them on images obtained by different imaging modalities. The drawback of fiducial markers is that they have to be either rigidly attached to the patient's anatomy, which is usually invasive and inconvenient to the patient, or glued to the skin, which does not guarantee, due to the elasticity of the skin, that they will remain at the same position during treatment. *Intrinsic* geometrical features may also be points, i.e. distinctive points on anatomical structures, or curves and surfaces that model anatomical structures in 2D and 3D, respectively (Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Guéziec *et al.*, 2000; Hamadeh *et al.*, 1998; Lavallée and Szeliski, 1995). Registration of surfaces to curves is performed by minimizing the distance between the surface model derived from the preoperative image and a set of lines connecting points on the intraoperatively segmented contour (curve) with the X-ray source. Accurate automatic segmentation is a challenging problem, while manual segmentation is time consuming and depends on the skills of the human operator. The reduced amount of segmented data makes the registration fast but, unfortunately, errors in segmentation propagate to errors in registration.

The alternative to segmentation-based registration methods, are intensity-based 3D/2D registration methods, which rely on image information included in image intensities or intensity gradients of pixels and voxels. Most of the intensity-based 3D/2D registration methods, designed to register CT images and X-ray images, are based on simulated X-ray projection images (Goitein *et al.*, 1983), called digitally reconstructed radiographs (DRRs), produced from a CT image. The unknown pose of the CT volume relative to the X-ray image set is estimated by optimizing the similarity measure calculated from a DRR and X-ray image (Hipwell *et al.*, 2003; Lemieux *et al.*, 1994; Penney *et al.*, 1998; Weese *et al.*, 1997). Since intensity-based 3D/2D registration methods employ intensities as features they are potentially more accurate than segmentation-based methods, yielding registration errors below the spatial resolution of images involved in registration. Unfortunately, intensity-based registration methods are slow due to time consuming calculation of DRRs. The problem may be reduced to some extent by calculating DRRs containing only structures of interest (Weese *et al.*, 1997) or by implementing a faster DRR generation method (LaRose *et al.*, 2000a; Russakoff *et al.*, 2003b; Weese *et al.*, 1999). Nevertheless, by projecting a high quality 3D image into 2D, valuable information, needed for accurate and robust registration, may be lost. A further drawback of this method is that it is not suitable for registration of MR images to X-ray images, as there is practically no correlation between MR-based DRRs and X-ray images.

3D/2D registration methods, which we name hybrid, combine elements of segmentation- and intensity-based methods with the purpose to achieve the speed of a segmentation-based method and the accuracy of an intensity-based method (Gilhuijs *et al.*, 1996a; Livyatan *et al.*, 2003; Tomažević *et al.*, 2003). Accuracy, comparable to the accuracy of intensity-based methods, is achieved by using intensities or intensity gradients as registration features, while the speed, comparable to the speed of segmentation-based methods, is obtained by reducing the amount of data that has to be processed during registration. Data reduction is carried out by automatic segmentation of either the intraoperative 2D images (Gilhuijs *et al.*, 1996a; Livyatan *et al.*, 2003) or the preoperative 3D image (Tomažević *et al.*, 2003). By segmentation, pixels or voxels that carry relevant information are extracted. Because, in general, many more points, i.e. information, are extracted from the 2D or 3D images than in the case of intrinsic segmentation-based methods, the adverse impact of outliers on registration is smaller.

The majority of methods, proposed in recent years have been intensity-based or hybrid. From the validation results provided by the authors of these methods and from a recent comparison study of van de Kraats *et al.* (van de Kraats *et al.*, 2005a), it is obvious that some of these methods are highly accurate when registering a CT image to two or more X-ray images, but that their capture range and robustness need to be improved. It is also clear, that 3D/2D registration of very different imaging modalities, e.g. registration of a 3D MR image to 2D X-ray images, is still a challenging problem.

Most 3D/2D registration methods adjust the parameters of an appropriate spatial transformation model until the similarity measure, describing the match between images or image features, reaches an optimum. If a 3D image is registered to multiple 2D images, the similarity measure is usually simply the sum of similarity measures, each corresponding to the match of the 3D image to one of the 2D X-ray images. In this paper we propose a novel and general approach to 3D/2D registration. The proposed 3D/2D image registration method is based on integrating the information from a few 2D projection images into 3D image space and finding that geometrical transformation that brings the preoperative 3D image into the best possible spatial correspondence with the integrated 3D information. We thus follow our previous approach (Tomažević *et al.*, 2003) of transforming information from one or more 2D images into 3D and not the approaches, like the DRR-based approach, that optimize a similarity measure between 2D images, i.e. the 3D preoperative image transformed and projected into 2D and the 2D intraoperative image. A natural solution for integrating information from multiple 2D projection images into 3D image space is 3D reconstruction from 2D projections. The quality of a 3D image, reconstructed from a small number of 2D images is expected to be low. The similarity measure applied in such a registration method should therefore be able to cope with low image quality. The similarity measure that nowadays enjoys the reputation of an accurate and robust multimodality image registration criterion is mutual information (Collignon *et al.*, 1995; Viola and Wells, 1995). However, even registrations

based on mutual information can fail if the statistical relationship between features of the two images is weak, which is the case when one of the images is the image reconstructed from a few X-ray images. This is even more likely if an MR image is to be registered to image reconstructed from the X-ray views. We believe the only way to cope with the insufficient information is to add spatial information into the similarity measure. In this paper we therefore propose a novel similarity measure, which we call the asymmetric multi-feature mutual information measure. The measure is based on the multi-feature mutual information measure, recently proposed by the same authors (Tomažević *et al.*, 2004b).

The paper is organized into six sections. In Section 6.2 we introduce a novel similarity measure capable of registering preoperative 3D 3DRX, CT, and MR images to images reconstructed from a few X-ray projections. Experiments are given in Section 6.3 and results in Section 6.4. The reconstruction-based 3D/2D registration method and results are discussed in Section 6.5, conclusions are drawn in Section 6.6.

## 6.2 Registration of preoperative and reconstructed image

We denote the two images to be registered as floating image  $A$  and reference image  $B$ , and represent them as vector functions  $\mathbf{z}_a(x)$  and  $\mathbf{z}_b(x)$  of position  $x$  in image space, respectively. Each vector function  $\mathbf{z}(x)$ ,  $\mathbf{z}(x)=(z_1(x), \dots, z_K(x))$  is comprised of values of  $K$  image features  $z_k(x)$ ,  $k=1\dots K$ . For a given spatial transformation  $\mathfrak{J}$ ,  $SM$  denote the criterion function or similarity measure between corresponding feature sets  $\mathbf{z}_a(x)$  and  $\mathbf{z}_b(\mathfrak{J}(x))$ . Registration then seeks the spatial transformation  $\hat{\mathfrak{J}}$  that maximizes the similarity measure  $SM$ :

$$\hat{\mathfrak{J}} = \arg \max_{\mathfrak{J}} SM(\mathbf{z}_a(x), \mathbf{z}_b(\mathfrak{J}(x))). \quad (6.1)$$

To be able to successfully match 3DRX, CT and MR images to images reconstructed from a few X-ray images we have devised a novel similarity measure, which we called the asymmetric multi-feature mutual information (AMMI). The AMMI similarity measure is a modification of a recently proposed multi-feature mutual information (MMI) similarity measure (Tomažević *et al.*, 2004b). Both similarity measures, which represent a generalization of the widely used single-feature mutual information criteria (Collignon *et al.*, 1995; Viola and Wells, 1995), require efficient estimation of multivariate probability distributions.

### 6.2.1 Asymmetric multi-feature mutual information

Let the values of  $\mathbf{z}_a(x)$  and  $\mathbf{z}_b(x)$  be the observed values of vectors of random variables  $\mathbf{Z}_a$  and  $\mathbf{Z}_b$ , respectively. In terms of entropy, multi-feature mutual information (MMI) is defined as (Tomažević *et al.*, 2004b):

$$MMI(\mathbf{Z}_a, \mathbf{Z}_b) = H(\mathbf{Z}_a) + H(\mathbf{Z}_b) - H(\mathbf{Z}_a, \mathbf{Z}_b), \quad (6.2)$$

where  $H(\mathbf{Z}_a)$ ,  $H(\mathbf{Z}_b)$  and  $H(\mathbf{Z}_a, \mathbf{Z}_b)$  are entropies of vectors of random variables  $\mathbf{Z}_a = (Z_{a1}, \dots, Z_{aK})$ ,  $\mathbf{Z}_b = (Z_{b1}, \dots, Z_{bK})$  and  $(\mathbf{Z}_a, \mathbf{Z}_b) = (Z_{a1}, \dots, Z_{aK}, Z_{b1}, \dots, Z_{bK})$ , respectively. In general, the entropy of a  $K$ -dimensional random variable  $\mathbf{Z}$ ,  $\mathbf{Z} = (Z_1, \dots, Z_K)$  is defined as:

$$H(\mathbf{Z}) = - \int p(\mathbf{z}) \log p(\mathbf{z}) d\mathbf{z} = - \int \dots \int p(z_1, \dots, z_K) \log p(z_1, \dots, z_K) dz_1 \dots dz_K. \quad (6.3)$$

When  $Z_1, \dots, Z_K$  are discrete random variables, the entropy  $H(\mathbf{Z})$  is obtained as:

$$H(\mathbf{Z}) = - \sum_{\mathbf{z}} p(\mathbf{z}) \log p(\mathbf{z}) = - \sum_{z_1} \dots \sum_{z_K} p(z_1, \dots, z_K) \log p(z_1, \dots, z_K). \quad (6.4)$$

To assess the entropies  $H(\mathbf{Z}_a)$ ,  $H(\mathbf{Z}_b)$  and  $H(\mathbf{Z}_a, \mathbf{Z}_b)$ , and consequently the multi-feature mutual information (6.2), multivariate probability distributions  $p(\mathbf{z}_a)$ ,  $p(\mathbf{z}_b)$  and  $p(\mathbf{z}_a, \mathbf{z}_b)$  have to be known. As in single-feature mutual information, the multivariate probability distributions can be estimated from joint histograms (Maes *et al.*, 1997). Unfortunately, even in the case of two features, the four-dimensional histogram  $h(\mathbf{z}_a, \mathbf{z}_b)$ , will probably be so sparse that a meaningful estimation of  $p(\mathbf{z}_a, \mathbf{z}_b)$  will become practically impossible.

By calculating MMI in other ways than the one in (6.2), the estimation of entropies via high-dimensional joint histograms may be avoided. Tomažević *et al.* (Tomažević *et al.*, 2004b) proposed to decompose the vector function  $\mathbf{z}(x)$  into a scalar function  $i(x)$ , describing a basic feature, and a vector function  $\mathbf{v}(x)$ , describing additional features. Separating features of both, the floating and the reference image, into a basic feature  $i(x)$ , and additional features  $\mathbf{v}(x)$ , i.e.  $\mathbf{z}_a(x) = (i_a(x), \mathbf{v}_a(x))$  and  $\mathbf{z}_b(x) = (i_b(x), \mathbf{v}_b(x))$ , and using the known property of entropy (Cover and Thomas, 1991) that:

$$H(\mathbf{Z}) = H(Z_k) + H((Z_1, \dots, Z_{k-1}, Z_{k+1}, \dots, Z_K) | Z_k), \quad (6.5)$$

MMI (6.2) is obtained as:

$$\begin{aligned}
MMI(\mathbf{Z}_a, \mathbf{Z}_b) &= H(I_a, \mathbf{V}_a) + H(I_b, \mathbf{V}_b) - H(I_a, I_b, \mathbf{V}_a, \mathbf{V}_b) = \\
&= H(I_a) + H(\mathbf{V}_a | I_a) + H(I_b) + H(\mathbf{V}_b | I_b) - H(I_a, I_b) - H(\mathbf{V}_a, \mathbf{V}_b | I_a, I_b) = . \quad (6.6) \\
&= H(I_a) + H(I_b) - H(I_a, I_b) + H(\mathbf{V}_a | I_a) + H(\mathbf{V}_b | I_b) - H(\mathbf{V}_a, \mathbf{V}_b | I_a, I_b)
\end{aligned}$$

Tomažević *et al.* (Tomažević *et al.*, 2004b) obtained entropies  $H(I_a)$ ,  $H(I_b)$ , and  $H(I_a, I_b)$  by estimating probability distributions  $p(i_a)$ ,  $p(i_b)$ , and  $p(i_a, i_b)$  from corresponding histograms, respectively, while they got the estimations of conditional entropies  $H(\mathbf{V}_a | I_a)$ ,  $H(\mathbf{V}_b | I_b)$ , and  $H(\mathbf{V}_a, \mathbf{V}_b | I_a, I_b)$  from:

$$H(\mathbf{Y} | I) = \sum_i p(i)H(\mathbf{Y} | i), \quad (6.7)$$

$$H(\mathbf{Y} | i) = - \int p(\mathbf{y} | i) \log p(\mathbf{y} | i) d\mathbf{y}, \quad (6.8)$$

where  $p(i)$  stands for probability distribution  $p(i_a)$ ,  $p(i_b)$ , or  $p(i_a, i_b)$ ,  $H(\mathbf{Y} | i)$  for entropies  $H(\mathbf{V}_a | i_a)$ ,  $H(\mathbf{V}_b | i_b)$ , or  $H(\mathbf{V}_a, \mathbf{V}_b, \mathbf{Z}_b | i_a, i_b)$ , and  $p(\mathbf{y} | i)$  for conditional multivariable probability distributions  $p(\mathbf{v}_a | i_a)$ ,  $p(\mathbf{v}_b | i_b)$ , or  $p(\mathbf{v}_a, \mathbf{v}_b | i_a, i_b)$ . Assuming that the conditional multivariable probability distributions were normal, each distribution was defined by its mean  $\mu_y$  and  $n$ -dimensional covariance matrix  $\Sigma_y$ . Knowing the covariance matrix  $\Sigma_y$  of distribution  $\mathbf{Y}$ , entropy  $H(\mathbf{Y})$  can be easily assessed from the determinant of the covariance matrix  $\Sigma_y$  as (Cover and Thomas, 1991):

$$H(\mathbf{Y}) = \frac{1}{2} \log |\Sigma_y| + \frac{n}{2} \log(2\pi e). \quad (6.9)$$

The MMI, obtained in this way was called symmetric MMI (SMMI) (Tomažević *et al.*, 2004b).

MMI can, however, also be defined if features of only one image and not both are separated into basic and additional features. We call this approach asymmetric MMI (AMMI) and propose to use it for registering a preoperative image to a reconstructed image. Let, therefore, only one vector function, say  $\mathbf{z}_a(x)$ , be divided into a basic feature  $i_a(x)$  and vector of additional features  $\mathbf{v}_a(x)$ , i.e.  $\mathbf{z}_a(x) = (i_a(x), \mathbf{v}_a(x))$ . Using (6.5), MMI is:

$$\begin{aligned}
MMI(\mathbf{Z}_a, \mathbf{Z}_b) &= H(\mathbf{Z}_a) + H(\mathbf{Z}_b) - H(\mathbf{Z}_a, \mathbf{Z}_b) = \\
&= H(I_a, \mathbf{V}_a) + H(\mathbf{Z}_b) - H(I_a, \mathbf{V}_a, \mathbf{Z}_b) = \\
&= H(I_a) + H(\mathbf{V}_a | I_a) + H(\mathbf{Z}_b) - H(I_a) - H(\mathbf{V}_a, \mathbf{Z}_b | I_a) = , \quad (6.10) \\
&= H(\mathbf{Z}_b) + H(\mathbf{V}_a | I_a) - H(\mathbf{V}_a, \mathbf{Z}_b | I_a)
\end{aligned}$$

where  $H(\mathbf{Z}_b)$  is the entropy of  $\mathbf{Z}_b$ , and  $H(\mathbf{V}_a|I_a)$  and  $H(\mathbf{V}_a, \mathbf{Z}_b|I_a)$  are entropies of  $\mathbf{V}_a$  and  $\mathbf{V}_a\mathbf{Z}_b$  under the condition  $I_a$ , respectively. Assuming that distributions of  $\mathbf{Z}_b$ ,  $\mathbf{V}_a|I_a$  and  $\mathbf{V}_a\mathbf{Z}_b|I_a$  are normal, entropy  $H(\mathbf{Z}_b)$  and conditional entropies  $H(\mathbf{V}_a|I_a)$  and  $H(\mathbf{V}_a, \mathbf{Z}_b|I_a)$  may be defined by (6.7) and (6.9), knowing the distribution  $p(i_a)$  and covariance matrices  $\Sigma_{\mathbf{z}_b}$ ,  $\Sigma_{\mathbf{v}_a|i_a}$  and  $\Sigma_{\mathbf{v}_a\mathbf{z}_b|i_a}$ . Covariance matrices  $\Sigma_{\mathbf{v}_a|i_a}$  and  $\Sigma_{\mathbf{v}_a\mathbf{z}_b|i_a}$  are estimated for every feature value  $i_a$ . This approach requires many fewer samples than the estimation through high dimensional histograms. Moreover, in the case of AMMI, the size of histogram needed to estimate the probability distribution of the basic feature is reduced from a two-dimensional histogram  $h(i_a, i_b)$ , as for SMMI, to a one-dimensional histogram  $h(i_a)$ . By reducing the size of histogram, the number of conditional covariance matrices that have to be estimated is reduced. In this way, more samples become available for estimation of individual covariance matrix.

Entropies  $H(\mathbf{Z}_b)$ ,  $H(\mathbf{V}_a|I_a)$  and  $H(\mathbf{V}_a, \mathbf{Z}_b|I_a)$ , estimated by (6.7) and (6.9) are monotonically increasing functions of variances of random variables  $\mathbf{Z}_b$ ,  $\mathbf{V}_a|I_a$  and  $\mathbf{V}_a\mathbf{Z}_b|I_a$ , respectively, which are assumed to be normally distributed. The condition that random variable  $\mathbf{Z}_b$ , and conditional random variables  $\mathbf{V}_a|I_a$  and  $\mathbf{V}_a\mathbf{Z}_b|I_a$  are normally distributed will, generally, not be fulfilled in practice. Nevertheless, (6.7) and (6.9) may be used to estimate entropies needed to calculate AMMI as long as the real, but unknown entropies are also monotonically increasing functions of variances of  $\mathbf{Z}_b$ ,  $\mathbf{V}_a|I_a$  and  $\mathbf{V}_a\mathbf{Z}_b|I_a$ . The crucial part of calculation of mutual information is the estimation of marginal and joint feature probability distributions that are needed for assessing corresponding entropies. A good probability estimation method should efficiently solve the problem of feature interpolation and the problem of weak statistics. The estimation of probability distribution is given in the Appendix.

## 6.3 Experiments

The purpose of the experimental part of the work was to evaluate the performances of the proposed 3D/2D registration method based on 3D integration of 2D information (the reconstruction based method) and compare this with the previously published gradient-based method of Tomažević *et al.* (Tomažević *et al.*, 2003). The latter method is based on the sum of similarity measures between normals to bony surfaces extracted from the 3D image and gradients of individual X-ray images. The methods were compared using data and standardized evaluation methodology for 3D/2D registration evaluation described by van de Kraats *et al.* (van de Kraats *et al.*, 2005a).

### 6.3.1 Experimental data sets

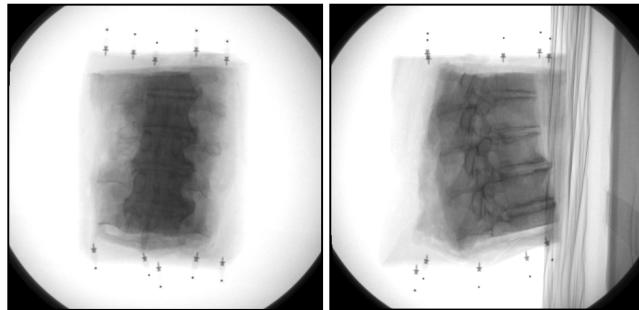
The publicly available<sup>3</sup> image data used in the registration experiment is part of the standardized evaluation methodology for testing 3D/2D registration

---

<sup>3</sup> <http://www.isi.uu.nl/Research/Databases/>

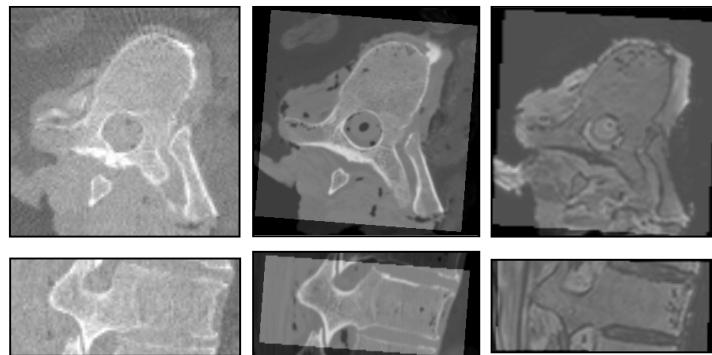
**Table 6.1.** Spatial resolutions and sizes of image data.

Modality	Segment	Resolution	Size
MR	1	$1.00 \times 0.75 \times 0.75 \text{ mm}^3$	$100 \times 256 \times 256$
	2	$1.00 \times 0.88 \times 0.88 \text{ mm}^3$	$120 \times 256 \times 256$
CT	1	$0.31 \times 0.49 \times 0.31 \text{ mm}^3$	$320 \times 260 \times 320$
	2	$0.31 \times 0.49 \times 0.31 \text{ mm}^3$	$280 \times 300 \times 300$
3DRX	1	$0.87 \times 0.87 \times 0.87 \text{ mm}^3$	$256^3$
	2	$0.52 \times 0.52 \times 0.52 \text{ mm}^3$	$256^3$
X-ray	1	$0.63 \times 0.63 \text{ mm}^2$	$512^2$
	2	$0.53 \times 0.53 \text{ mm}^2$	$512^2$

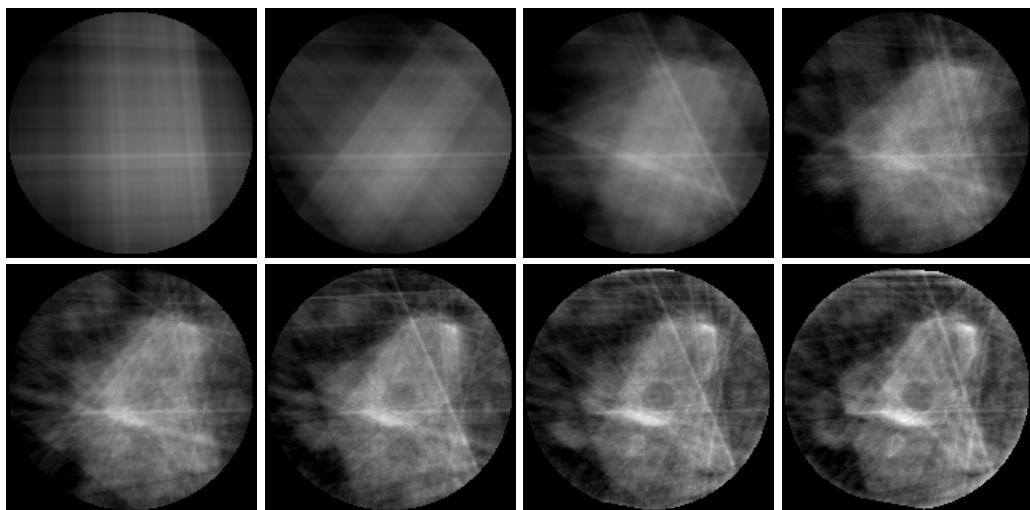
**Figure 6.1.** Anterior-posterior (left) and lateral X-ray images of a spine phantom.

methods proposed by van de Kraats *et al.* (van de Kraats *et al.*, 2005a). The data comprises 2D fluoroscopic X-ray images and 3D 3DRX, CT and MR images of two defrosted segments of vertebral bodies. The first vertebral column consists of three thoracolumbar vertebrae bodies and the second segment consist of five thoracic vertebrae bodies with some soft tissue still present around both segments.

The 2D fluoroscopic images were obtained with a clinical 3D rotational X-ray (3DRX) system (Integris BV5000, Philips Medical System, Best, The Netherlands). A set of 100 X-ray images was acquired for each spinal segment in 8 second runs of 180 degrees rotation around the imaged object. The images were acquired with image intensifier size of 38 cm and 31 cm for the first and second column, respectively. Two examples of acquired X-ray images are shown in Figure 6.1. For each vertebral column, the 3DRX image was reconstructed from a set of 100 X-ray images using a filtered back-projection reconstruction technique (Grass *et al.*, 1999). The CT-images were acquired with a clinical 16-detector-row multi-slice CT scanner (MSCT, Philips Medical System, Best, The Netherlands). The MR images were obtained with a clinical



**Figure 6.2.** Transversal (top row) and lateral planes (bottom row) of corresponding VOIs taken from 3DRX (left column), CT (middle column) and MR (right column) volumes.



**Figure 6.3.** Transversal planes of 3D images reconstructed from 2, 3, 5, and 7 projection images (top row, left to right) and 9, 11, 13 and 15 projection images (bottom row, left to right) using the SART reconstruction method.

1.5 Tesla MR scanner (Gyroscan NT, Philips Medical System, Best, The Netherlands) using a sagittal 3D turbo spin echo acquisition and turbo factor of 29, TR/TE of 1500 ms/90 ms. The sizes and resolution of acquired images are given in Table 6.1.

The gold standard registration between 3DRX images and 2D projection images was already established in the process of creating 3DRX images, while the gold standard registration between CT and MR images with 2D fluoroscopic images was obtained by 3D/3D rigid registration of CT and MR

images, to corresponding 3DRX image using the mutual information maximization registration method (Maes *et al.*, 1997). Volumes of interest (VOIs), containing a whole vertebra body and less than a quarter of neighboring vertebrae, were manually determined on 3DRX image volumes. By using the results of gold standard registration, the VOIs positions on 3DRX images were transformed to corresponding positions on CT and MR images. Examples of 3DRX, CT, and MR images of a VOI are shown in Figure 6.2.

### 6.3.2 Image preprocessing and implementation details

To reduce severe intensity inhomogeneity of MR images due to non-uniform coil energy distribution, the MR images were corrected with the retrospective intensity inhomogeneity correction method based on information minimization (Likar *et al.*, 2001). The gradient-based method was implemented as described in (Tomažević *et al.*, 2003). Two-dimensional X-ray images and VOIs were blurred with a Gaussian kernel of 0.5 mm, the sample step for isotropic resampling of 3D images was set to 0.95493 mm, while the threshold to extract bone edges from VOIs was set to 18 for 3DRX and CT images and 15 for MR images. The 3D image, containing the whole spinal segment was reconstructed from a small subset of the X-ray images by the SART reconstruction method (Andersen and Kak, 1984; Kak and Slaney, 1988). The sizes of reconstructed 3D images were 128x128x128 and 128x188x128 image elements for the first and second spinal segment, respectively, with isotropic spatial resolution of 0.6271 mm. Figure 6.3 shows examples of 3D images reconstructed from a few X-ray images.

The preoperative 3D image was taken as the floating image  $A$  and the reconstructed image as the reference image  $B$ . Both feature sets  $\mathbf{z}_a(x)$  and  $\mathbf{z}_b(x)$ , characterizing the preoperative and reconstructed image, respectively, consisted of image intensity  $i(x)$  and image intensity gradient  $\mathbf{v}(x)$  features, i.e.  $\mathbf{z}_a(x) = (i_a(x), \mathbf{v}_a(x))$  and  $\mathbf{z}_b(x) = (i_b(x), \mathbf{v}_b(x))$ . The AMMI similarity measure (6.10) was used to measure the match between  $\mathbf{z}_a(x)$  and  $\mathbf{z}_b(\mathfrak{I}(x))$ . Image intensity of the preoperative (floating) image  $i_a(x)$  was the only feature whose probability distribution was estimated by using a one-dimensional histogram  $h(i_a)$  of intensity values. A histogram, having 64 bins was used to insure statistical power. Assuming normal distribution of  $\mathbf{Z}_b$ ,  $\mathbf{V}_a | i_a$  and  $\mathbf{V}_a \mathbf{Z}_b | i_a$ , the multivariate probability distribution  $p(\mathbf{z}_b)$  was estimated through the covariance matrix  $\Sigma_{\mathbf{z}_b}$ , while the distributions  $p(\mathbf{v}_a | i_a)$  and  $p(\mathbf{v}_a, \mathbf{z}_b | i_a)$  were estimated through conditional covariance matrices  $\Sigma_{\mathbf{v}_a | i_a}$  and  $\Sigma_{\mathbf{v}_a \mathbf{z}_b | i_a}$ , respectively, for every intensity value  $i_a$ .

Before extracting gradients, 3DRX, CT, and MR images were isotropically resampled to 0.95493 mm voxel sizes by linear interpolation. To reduce sensitivity to image noise and non-isotropic image acquisition, the gradients of 3D images were obtained after convolving 3D intensities with a Gaussian kernel of scale  $\sigma$ :

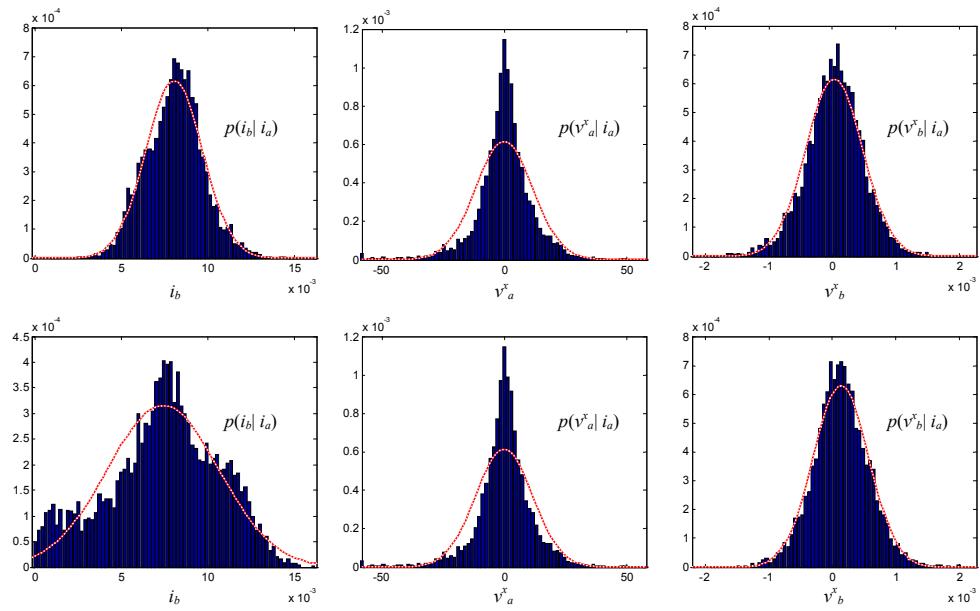
$$\mathbf{v}(x) = \nabla(i(x) * g(x, \sigma)). \quad (6.11)$$

Scales of 0.5 mm and 0.35 mm were applied to the original 3D images and 3D reconstructed images, respectively. Both, the gradient-based and the proposed reconstruction-based 3D/2D registration methods used the same Powell's optimization method (Press *et al.*, 1992) to optimize the given similarity measure for six rigid-body transformation parameters ( $t_x, t_y, t_z, \omega_x, \omega_y, \omega_z$ ).

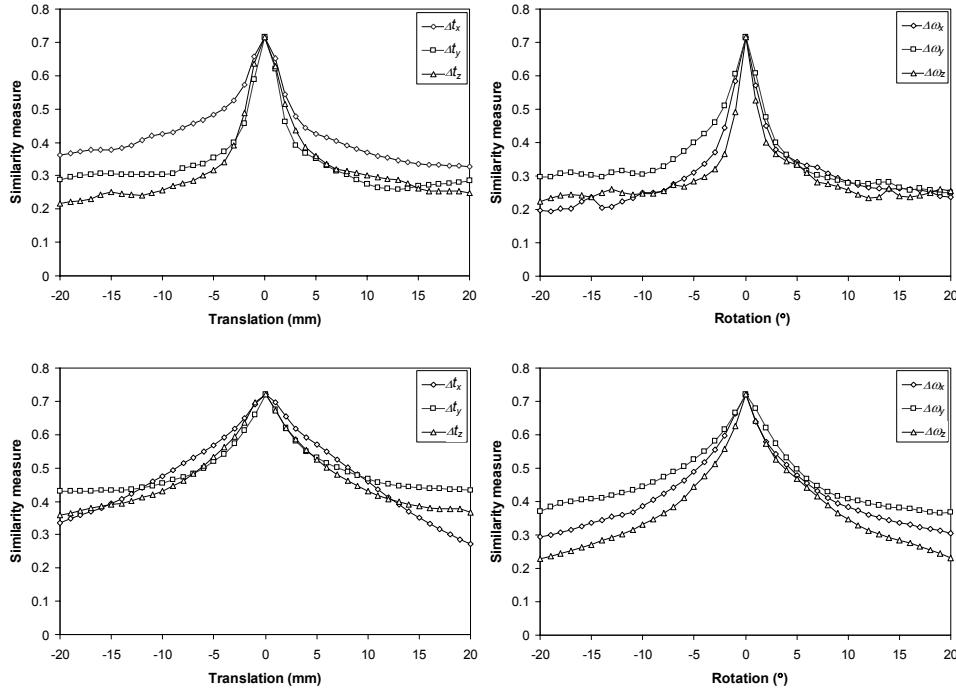
### 6.3.3 Registration experiment

Both, gradient-based and reconstruction-based 3D/2D registration methods were evaluated using the standardized methodology for 3D/2D registration evaluation described by van de Kraats *et al.* (van de Kraats *et al.*, 2005a). The evaluation methodology uses the mean target registration error (mTRE) to measure the distance of VOI position from gold standard before and after registration. The positions of all image elements in a VOI were used as target points. For evaluating the capture range and robustness of the 3D/2D registration method, the methodology requires that a large number of registrations are performed from different starting positions for each VOI and each X-ray image data set. Van de Kraats *et al.* (van de Kraats *et al.*, 2005a) provided 200 starting positions for each VOI. The 200 starting position were randomly generated around the gold standard position in such a way that the distance from gold standard measured by mTRE was uniformly distributed in the interval of 0-20 mm, with 10 positions in each of the 1 mm wide subintervals. Each registration was considered successful if mTRE after registration was lower than the predefined clinically relevant threshold, which had been set to 2 mm in the case of vertebra data. The registration error was defined as mTRE of all successful registrations, while the capture range was defined as the distance from gold standard for which the registration had proved to be successful in at least 95% of all cases.

Both methods were tested for the registration of VOIs containing a single vertebra to a set of X-ray images. The VOIs from 3DRX and CT images were registered to image sets containing 2, 3, 5, or 7 X-ray images and VOIs from the MR image to image sets containing 9, 11, 13, and 15 X-ray images. Unlike in (van de Kraats *et al.*, 2005a), the first image in each set was the first image acquired with the 3DRX system. The angle between image views in each set was approximately 90°, 60°, 35°, 24° for 3DRX and CT registrations and 18°, 15°, 13° and 11° for MR registrations, respectively. For each of 3DRX, CT or MR modalities and each set of X-ray images 1600 registrations were performed, 200 per each of the eight VOIs.



**Figure 6.4.** Typical histograms of conditional probabilities  $p(i_b|i_a)$ ,  $p(v^x_a|i_a)$  and  $p(v^x_b|i_a)$  with overlaid normal distributions of the MR image and image reconstructed from 9 X-ray images in gold standard (top) and displaced positions (bottom).



**Figure 6.5.** Behavior of the similarity measures of the gradient-based method (top row) and the reconstruction-based method (bottom row), when matching a CT volume of interest with three X-ray images and varying one translation (left column) and one rotation (right column) parameter around gold standard position.

## 6.4 Results

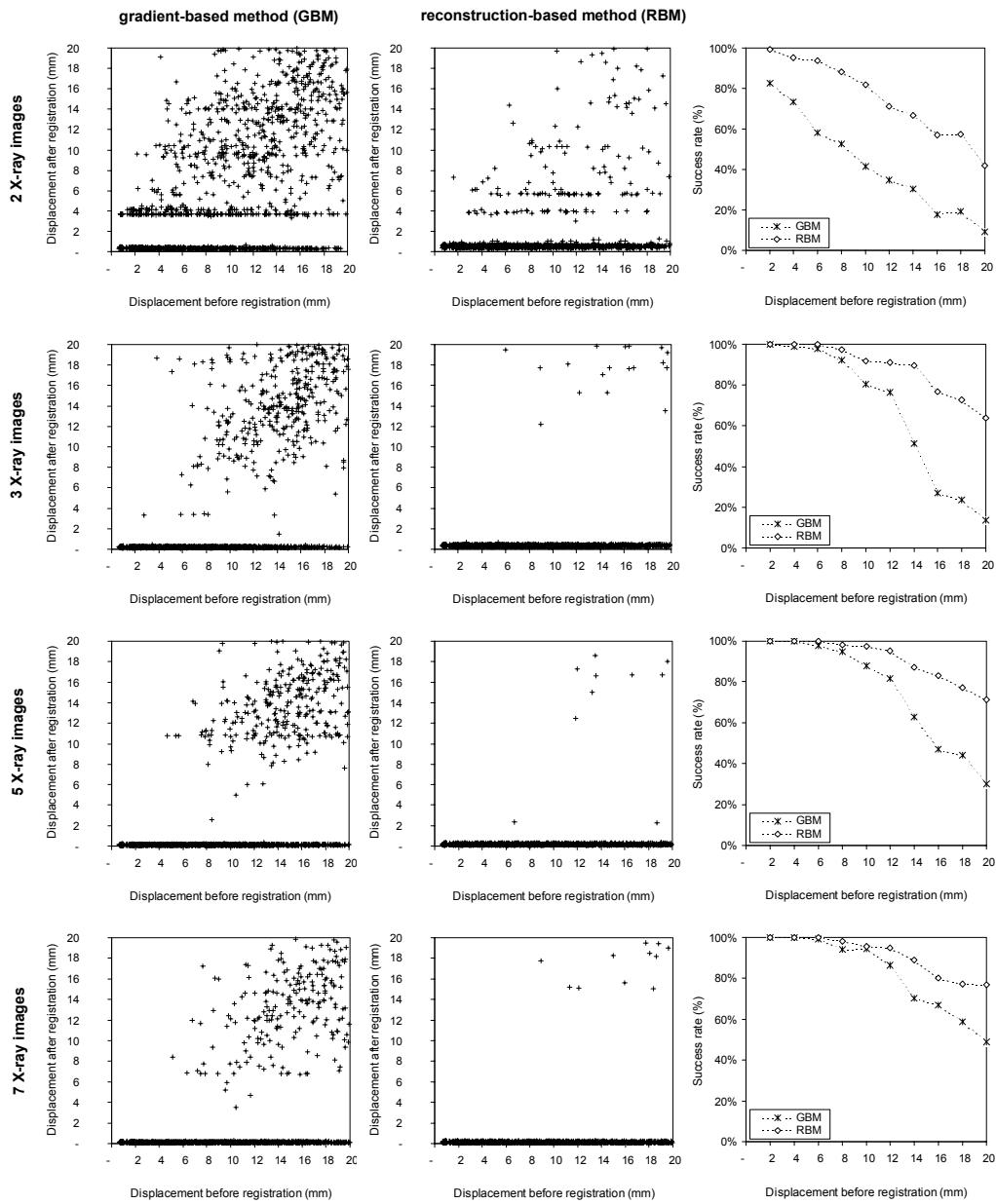
To demonstrate the assumption that conditional probabilities can be estimated from covariance matrices (6.9) we show typical histograms of conditional probabilities  $p(i_b | i_a)$ ,  $p(v^x_a | i_a)$  and  $p(v^x_b | i_a)$  with overlaid normal distributions in gold standard and displaced positions (Figure 6.4).  $v^x_a$  and  $v^x_b$  are components of vectors  $\mathbf{v}_a$  and  $\mathbf{v}_b$ , respectively, in direction  $x$ . All histograms are more or less monomodal and can be approximated by normal distributions. Figure 6.5 shows the behavior of the similarity measures when matching a CT to three X-ray images. The gradient-based method's similarity measure has a typical behavior of a measure that is based on image gradients. It has a narrow and very steep optimum near the true registered position, while the number of local optima increases when going away from the true registration position. On the other hand, the reconstruction-based method's similarity measure shows a smoother behavior on a larger distance from the true registration, which is probably the contribution of the image intensity feature, and has a significant and steep global optimum due to the contribution of the intensity gradient feature.

**Table 6.2.** Mean TREs, capture ranges, and success rate registrations for gradient-based (GBM) and reconstruction-based (RBM) methods using different modalities and number of X-rays.

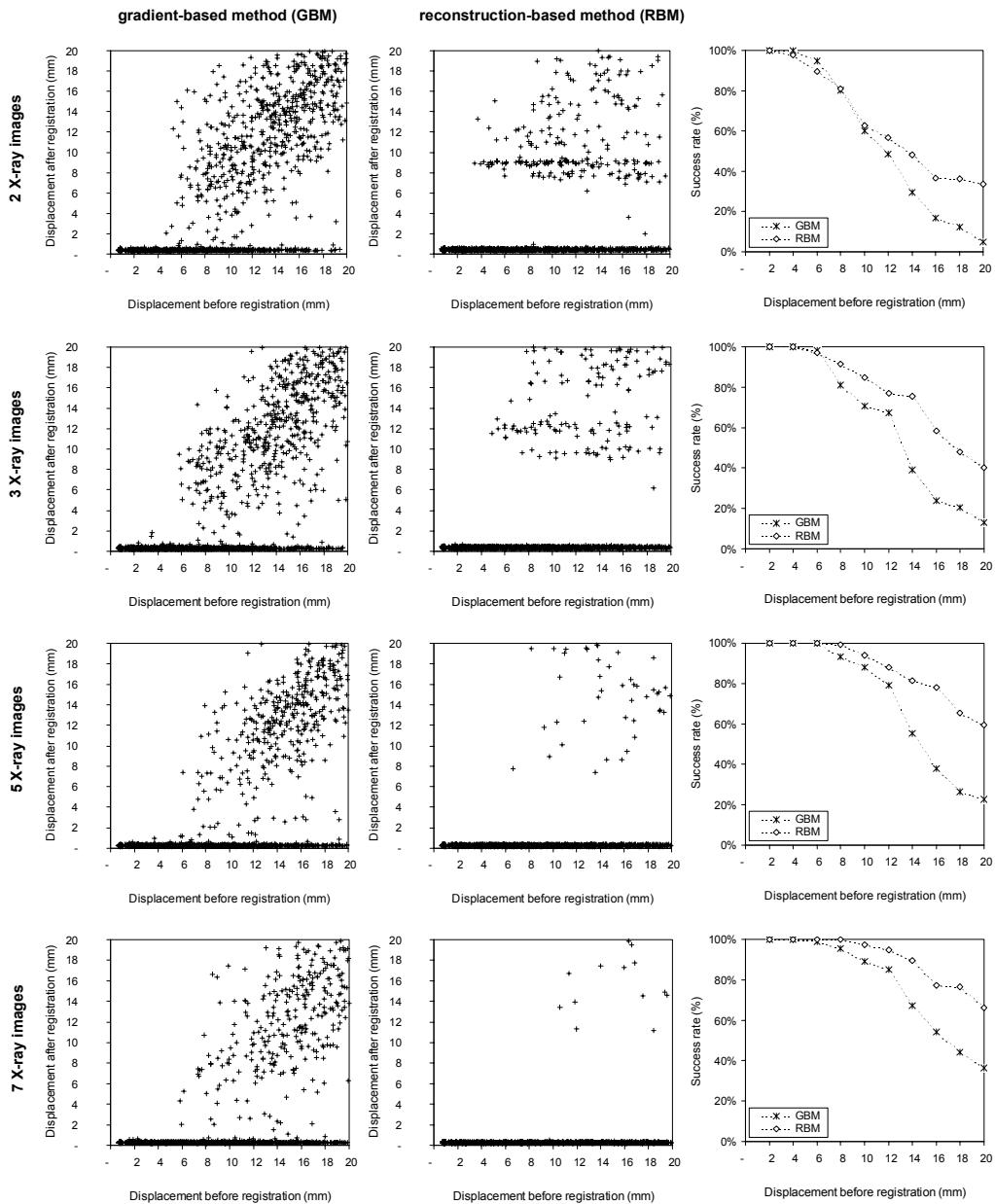
<b>Modality</b>	<b>X-rays</b>	<b>mTRE (mm)</b>		<b>Capture range (mm)</b>		<b>Success rate (%)</b>	
		GBM	RBM	GBM	RBM	GBM	RBM
<b>3DRX</b>	<b>2</b>	0.31*	0.52	0	4	43%	76%
	<b>3</b>	0.19*	0.33	7	9	68%	89%
	<b>5</b>	0.12	0.17	8	12	76%	91%
	<b>7</b>	0.11	0.12	8	12	83%	92%
<b>CT</b>	<b>2</b>	0.38	0.43	6	5	56%	65%
	<b>3</b>	0.32	0.37	7	7	63%	78%
	<b>5</b>	0.27	0.27	8	10	72%	87%
	<b>7</b>	0.27	0.26	9	12	78%	91%
<b>MR</b>	<b>9</b>	0.54*	0.98	2	0	24%	69%
	<b>11</b>	0.50*	0.67	2	7	23%	84%
	<b>13</b>	0.53	0.54	2	7	24%	84%
	<b>15</b>	0.50	0.46	2	7	27%	84%

\* GBM significantly more accurate ( $P<0.01$ ) than RBM

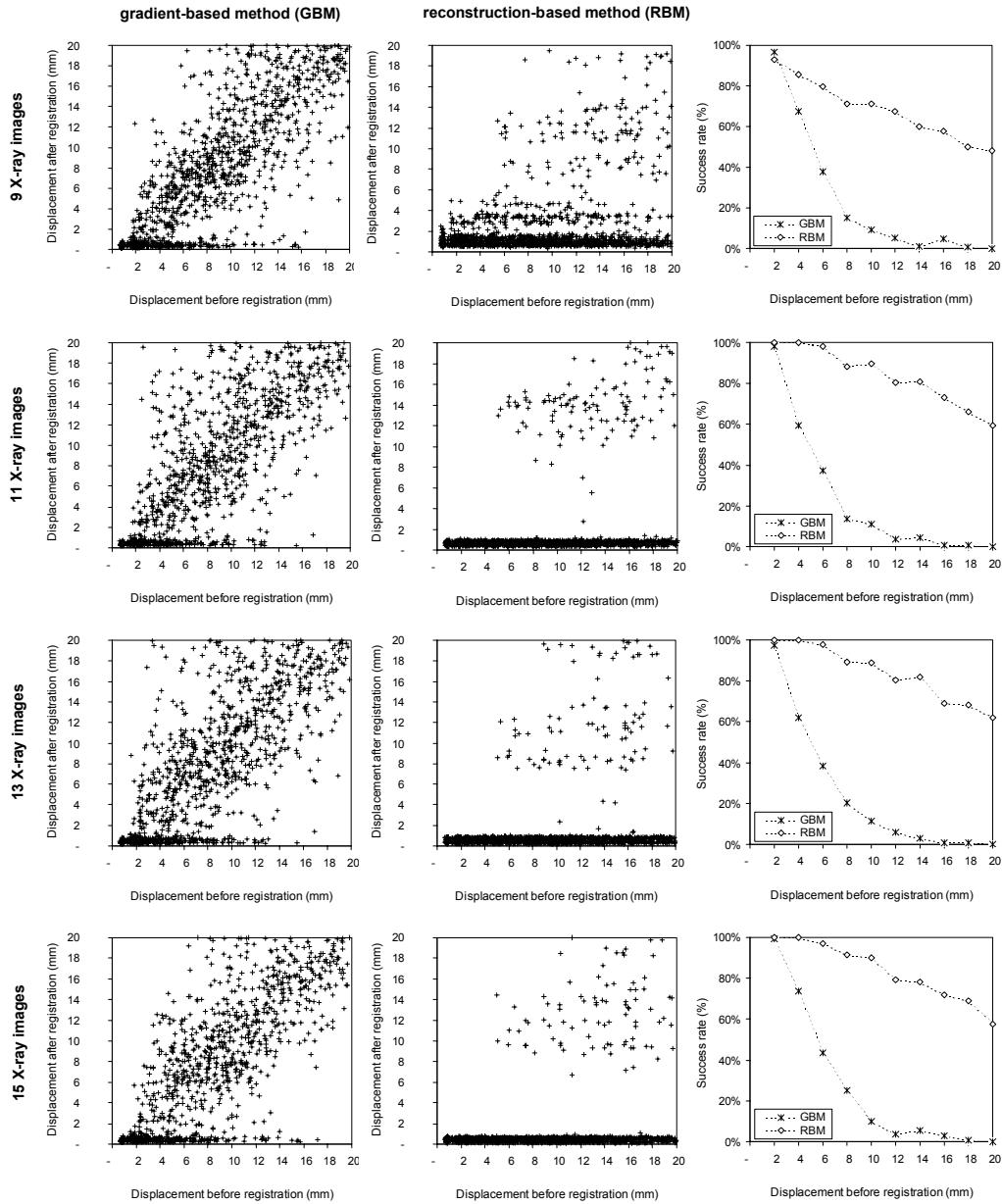
Table 6.2 presents registration errors, capture range and success rate (percentages of successful registrations) for each combination of preoperative imaging modality and set of X-ray images. To test the hypothesis that the GBM method is significantly more accurate than the RBM method, a paired t-test on the mean TRE of the eight VOIs obtained by both methods was performed. Figures 6.6, 6.7, and 6.8 show results of registering 3DRX, CT and MR images to X-ray image sets, respectively. Results are in the form of scatter diagrams of displacements before and after registration and success rate with respect to the initial displacement. The mTRE was used as a measure of displacement before and after registration.



**Figure 6.6.** Registration results of 3DRX VOIs to 2, 3, 5 and 7 X-ray images (rows, respectively). Scatter diagrams of displacements before and after registration (only displacements smaller than 20mm are shown) for the gradient-based method (GBM) (left column) and the reconstruction-based method (RBM) (middle column), and success rate (right column).



**Figure 6.7.** Registration results of CT VOIs to 2, 3, 5 and 7 X-ray images (rows, respectively). Scatter diagrams of displacements before and after registration (only displacements smaller than 20mm are shown) for the gradient-based method (GBM) (left column) and the reconstruction-based method (RBM) (middle column), and success rate (right column).



**Figure 6.8.** Registration results of MR VOIs to 9, 11, 13 and 15 X-ray images (rows, respectively). Scatter diagrams of displacements before and after registration (only displacements smaller than 20mm are shown) for the gradient-based method (GBM) (left column) and the reconstruction-based method (RBM) (middle column), and success rate (right column).

## 6.5 Discussion

In the last decade several 3D/2D registration approaches have been proposed and different registration methods have been developed. However, only a few of the published methods had undergone extensive experimental evaluation. Published comparisons are even rarer. It is difficult to make a fair comparison of methods because different researchers used different data sets, evaluation protocols, criteria for successful registrations, and error metrics. Recently, there have been some efforts to provide publicly available 3D and 2D image data, gold standard registrations, and evaluation methodology to make comparison of different registration methods possible and fair (Tomažević *et al.*, 2004a; van de Kraats *et al.*, 2005a). In their recent research, van de Kraats *et al.* (van de Kraats *et al.*, 2005a) compared two different 3DRX and CT to X-ray registration approaches, the intensity-based approach (Penney *et al.*, 1998) and the gradient-based (hybrid) approach of Tomažević *et al.* (Tomažević *et al.*, 2003). The method of Penney *et al.* (Penney *et al.*, 1998) compares 2D DRRs to X-ray images using a similarity measure based on image gradient difference. The gradient-based approach of Tomažević *et al.* (Tomažević *et al.*, 2003) matches two 3D feature sets, i.e. normals to bony surfaces extracted from the 3D preoperative image and gradients of X-ray images projected into 3D. The study of van de Kraats *et al.* showed that the capture ranges for 3DRX and CT were, respectively, 0-6 mm and 0-3 mm for the gradient-based and 0-4 mm and 0-3 mm for the intensity-based method. Within the capture range the average mean TREs for 3DRX and CT were, respectively, 0.19 mm and 0.38 mm for the gradient-based and 0.39 mm and 0.75 mm for the intensity-based. We believe that the dimensionality of the space in which registrations are performed significantly contributes to the better performance of the gradient-based method. By creating a DRR the transformed preoperative image is projected into 2D and the applied similarity measure thus estimates the match between two 2D images. By projecting a 3D image into 2D, some valuable information needed for accurate and robust registration is most probably lost. The 3D preoperative image, which is usually of high quality, should thus be fully exploited by leaving it intact. This is the case with the gradient-based method. Besides, this method is probably the only method that is capable of registering MR images to X-ray images (Tomažević *et al.*, 2003). It should, however be noted that this method is much more sensitive to the initial misalignment if an MR instead of a CT image is registered.

By analyzing the results of the comparative study of van de Kraats *et al.* (van de Kraats *et al.*, 2005a) and other registration approaches and their results we came to the following conclusions. Registrations based solely on image intensities may not yield optimal results. Penney *et al.* (Penney *et al.*, 1998) obtained better results when they used a similarity measure based on 2D intensity gradients or local intensity patterns than when they used intensities themselves. Creation of a DRR by projecting 3D intensities into 2D results in a loss of information and additional artifacts (Sherouse *et al.*, 1990) due to

interpolation needed to calculate the integral of intensities along X-ray paths. The recently proposed gradient-based method (Tomažević *et al.*, 2003) does not encounter problems associated with DRRs. 3D gradient vectors that represent the surface of a 3D object are directly compared with X-ray gradients, without the need to calculate the integral of intensities or intensity gradients along the X-ray path. By employing intensity gradients and not image intensities, the gradient-based method proved to be accurate (Tomažević *et al.*, 2003; van de Kraats *et al.*, 2004). Unfortunately, the capture range of the gradient-based method is rather small due to the local nature of intensity gradient features used in the similarity measure. Based on these observations and motivated by the constant need to improve the accuracy and especially the capture range and robustness of image registration, we devised a reconstruction-based method, which first integrates the information from more 2D X-ray images into a single 3D image space and then matches two 3D images. Three-dimensional reconstruction from 2D projections represents a natural solution for integrating information from multiple 2D projection images into 3D image space provided sufficient views are available. Different iterative methods, called algebraic reconstruction techniques (ARTs) exist for the reconstruction from projections (Gordon *et al.*, 1970). These methods differ in how they define updates at each iteration. The best results are usually obtained with the simultaneous algebraic reconstruction technique (SART) (Andersen and Kak, 1984; Kak and Slaney, 1988). Algebraic reconstruction techniques are simple to implement but much slower than transformation-based reconstruction techniques. However, algebraic reconstruction techniques can provide better results in situations when a small number of projection measurements are available and/or when projections are not uniformly distributed (Andersen and Kak, 1984). This is exactly the case with 3D/2D registrations where the preoperative 3D image has to be registered to a small number of 2D projections acquired at arbitrary positions. Convergence is an important property of iterative techniques (Kak and Slaney, 1988). Usually, just a few iterations are enough to obtain useful reconstruction results, while increasing the number of iterations may increase reconstruction artifacts, which are caused by computation errors that are amplified in each iteration.

For the purpose of 3D/2D registration, usually only a few X-ray images are acquired. It is, therefore, highly unlikely that the reconstruction from a limited number of projections will result in an image that, in terms of its quality or information content, could match a CT or 3DRX image. Because distinct anatomical structures exhibited large intensity variations in the image reconstructed only from a few X-ray images, the symmetric MMI similarity measure (Tomažević *et al.*, 2004b) was not appropriate for such images. To be able to register low quality reconstructed images to preoperative images of similar modalities as well as to images of different modalities, a novel similarity measure had to be designed. The AMMI similarity measure proposed in this paper is based on image intensities and intensity gradient features of the preoperative 3DRX, CT, or MR image and the reconstructed image. We believe that image intensities ensure larger capture ranges while image

intensity gradients preserve the accuracy of methods that use gradient features. The preoperative images, like 3DRX, CT, and MR are generally of high quality, clearly showing different tissue classes. Intensities of such images form clusters (modes) in intensity histograms and therefore it cannot be assumed that the intensities of these images are normally distributed. For this reason we have estimated the probability distribution of preoperative image intensities from a one-dimensional intensity histogram. We have, however, assumed normal distribution for the multivariate probability distribution of intensity and gradients  $p(\mathbf{z}_b)$  of the reconstructed image as well as for multivariate conditional probability distributions  $p(\mathbf{v}_a | i_a)$  and  $p(\mathbf{v}_a, \mathbf{z}_b | i_a)$ . The assumption was based on the observation that in practice gradients of preoperative 3DRX, CT, and MR images and intensities and gradients of images reconstructed from a few 2D images do not form distinct clusters but exhibit large variations. By assuming normal distribution of these features, (6.7) and (6.9) may be used to efficiently estimate the entropies. Such estimation is meaningful, as long as the real, but unknown entropies, like the entropies obtained by (6.7) and (6.9) are monotonically increasing functions of variances of  $\mathbf{z}_b$ ,  $\mathbf{v}_a | i_a$  and  $\mathbf{v}_a \mathbf{z}_b | i_a$ .

Registration results for 3DRX and CT images (Table 6.2) showed that for the gradient-based and the proposed method the registration accuracy and success rate increased with a larger number of X-ray images used. However, regardless of the number of X-ray views, the proposed reconstruction-based method was successful in a significantly larger number of registrations than the gradient-based method. On the other hand, when registering the 3DRX image to X-ray images, the gradient-based method yielded significantly ( $P<0.01$ ) smaller registration errors for two out of four X-ray setups. However, the differences between mTREs obtained by both methods were small. The difference was 0.2 mm for two X-rays and it became less and less significant with a larger number of images. No significant differences in accuracy were found when registering the CT image to X-ray images. Except for two X-rays, the estimated capture ranges of 3DRX to X-ray image registrations were larger than for CT to X-ray registrations (Table 6.2). Better results of 3DRX registration were expected because 3DRX images and X-ray images were more closely related.

It is not only the number of final misregistrations, i.e. registration errors above 2 mm, and the influence of the initial misregistration that is important, but also the extent of misregistrations. In practice, if a misregistration occurs, it should be as large as possible to be able to detect it immediately. Misregistrations that are rather close to true registration are difficult to detect and undetected misregistrations can have serious consequences for the outcome of an image-guided procedure. This and other information about the performance of the methods can be obtained from scatter diagrams (Figures 6.6, 6.7 and 6.8) showing initial and final misregistrations (displacements). Scatter diagrams for 3DRX and CT images to X-ray images registrations are shown in Figures 6.6 and 6.7, respectively. The gradient-based method resulted in a significantly larger number of final misregistrations between 2

and 8 mm, while the reconstruction-based method, except for registration of a 3DRX to two X-ray views, yielded only a few misregistrations smaller than 8 mm. The scatter diagrams in Figure 6.6 show a large number of 4 mm misregistrations for the case of two X-ray images. The analysis of the results for individual VOIs revealed that for one of the eight VOIs the similarity measure of the gradient-based method did not have the global optimum at the gold standard. For this particular VOI, the registration failed regardless of the initial position. Similarly, for the reconstruction-based method, for two out of the eight VOIs the similarity measure had a large local optimum causing many misregistrations. Because this was not the case for CT images (Figure 6.7) and because CT images were of a much higher quality than 3DRX images, we believe that the false optima causing 3DRX to X-ray misregistrations were the consequence of 3DRX reconstruction artifacts. Some of the artifacts can be seen in Figure 6.2. However, reconstruction artifacts hampered the registrations only when intraoperative information was scarce, like in the case of reconstruction from two X-ray images.

From the curves (Figures 6.6 and 6.7 right) showing the success rate as a function of initial misregistration it is obvious that the success rate decreases with increasing initial displacements. It is also obvious that the proposed reconstruction-based registration method is less sensitive to the extent of initial displacement. These curves were also used to estimate the capture range of the registration methods, defined as the initial displacement at which at least 95% of registrations were successful. In this way estimated capture ranges, presented in Table 6.2, increased with the increasing number of X-ray images. The capture range of the proposed reconstruction-based method was 2 to 4 mm larger than the capture range of the gradient-based method. Since both methods used the same optimization method, the capture range directly depended on image features and the similarity measure used by each of the registration methods.

The results of registering MR VOIs to X-ray images (Table 6.2) show that the reconstruction-based method achieved a significantly larger success rate than the gradient-based method. The accuracy of the reconstruction-based method increased with a larger number of X-ray views, but only up to 11 X-ray images. Unlike the reconstruction-based method, the gradient-based method showed no improvement in accuracy and success rate, if the number of X-ray images was increased from 9 to 15. Although the GBM method was significantly ( $P<0.01$ ) more accurate than the RGB method when 9 or 11 X-rays were used, the differences were small. From the scatter diagram in Figure 6.8 it can be observed that the gradient-based method resulted in a significant number of misregistrations, many of them near the true registration, regardless of the number of X-ray images used. The reconstruction-based method yielded a large number of misregistrations close to the true registration only for 9 X-ray images. Besides, the number of all misregistrations decreased drastically if the number of X-ray images had been increased from 9 to 11. It seems that 11 X-ray images is a threshold, at which the reconstructed image contains enough information for successful

registration to the MR image. Since X-ray and MR images have completely different physical backgrounds, it was expected that more, in our case 11, X-ray images were needed for a successful MR/X-ray registration than for 3DRX/X-ray or CT/X-ray registrations. For the 3DRX and CT images two or three X-ray views were sufficient. The minimum of 11 X-ray images, needed for successful MR to X-ray registrations, is still much less than the 100 X-ray images that are required for a complete 3DRX reconstruction and MR/3DRX registration, which had been proposed in (van de Kraats *et al.*, 2004). An important consequence of a much smaller number of X-ray images required for successful registration to an MR image is the reduction of radiation dose delivered to the patient. We imagine that most clinical applications involving preoperative MR images will be limited to less than 11 X-ray images, some probably to only two. Although we have not tested our method on MR images other than of the spine, for which gold standard data was available, we believe that similar results could be obtained with fewer images if the anatomical structures that are clearly seen in X-ray images are also of a high contrast in MR images. One potential clinical application in which fewer than 11 X-ray images would most probably be required is the registration of magnetic resonance angiograms (MRA) and 2D X-ray digital subtraction angiograms (DSA)(Hipwell *et al.*, 2003).

## 6.6 Conclusion

The proposed novel approach to 3D/2D registration based on 3D integration of 2D information is general and does not make any constraints on the modalities and anatomies involved in the registration. The experimental results show that the proposed method outperforms our gradient-based method with respect to the capture range and success rate. The proposed method presents an important contribution to solving and understanding the 3D/2D registration problem because it increases the capture range of these algorithms, and essentially transforms the registration problem from being 3D/2D to being a 3D/3D registration.

## 6.7 Appendix

Probability distribution of the basic feature of the floating image  $i_a(x)$  is estimated via one dimensional histogram while joint probability distributions of other features, which are assumed to be normal, are estimated through covariance matrices. For a given transformation  $\mathfrak{I}$ , the probability distribution  $p(i_a)$  of intensity feature of floating image is estimated by normalizing the histogram  $h(i_a)$  of image feature values  $i_a(x)$  on the grid of positions  $x_j$  in floating image  $A$  ( $x_j \in A$ ) that overlaps reference image  $B$  ( $\mathfrak{I}(x_j) \in B$ ). Using the known expression for computing the covariance matrix  $\Sigma_y$  of distribution  $\mathbf{Y}$ :

$$\boldsymbol{\Sigma}_y = \frac{1}{N} \sum_{\mathbf{y}} \mathbf{y} \cdot \mathbf{y}^T - \frac{1}{N^2} \sum_{\mathbf{y}} \mathbf{y} \cdot \left( \sum_{\mathbf{y}} \mathbf{y} \right)^T. \quad (6.12)$$

The covariance matrix  $\boldsymbol{\Sigma}_{\mathbf{z}_b}$  of conditional distribution  $p(\mathbf{z}_b)$  for a given transformation  $\mathfrak{T}$  is then obtained from:

$$\boldsymbol{\Sigma}_{\mathbf{z}_b} = \frac{1}{N_a} \sum_j \mathbf{z}_b(t_j) \cdot \mathbf{z}_b(t_j)^T - \frac{1}{N_a^2} \sum_j \mathbf{z}_b(t_j) \cdot \left( \sum_j \mathbf{z}_b(t_j) \right)^T, \quad (6.13)$$

where  $t_j = \mathfrak{T}(x_j)$  are the positions in reference image  $B$  corresponding to the grid  $x_j$  of floating image  $A$ , and  $N_a$  is the number of floating image elements on the overlap with reference image. Since the transformed grid of positions  $\mathfrak{T}(x_j)$  in floating image  $A$  generally does not coincide with the grid of positions of the reference image  $B$ , trilinear interpolation of each feature value  $\mathbf{z}_b(t_j)$  from neighboring voxel elements should be applied.

Similarly, the covariance matrices  $\boldsymbol{\Sigma}_{\mathbf{v}_a|i_a}$  and  $\boldsymbol{\Sigma}_{\mathbf{v}_a \mathbf{z}_b | i_a}$  are calculated from vectors  $\mathbf{s}_{\mathbf{v}_a}(i_a)$  and  $\mathbf{s}_{\mathbf{v}_a \mathbf{z}_b}(i_a)$ , and matrices  $\mathbf{C}_{\mathbf{v}_a}(i_a)$  and  $\mathbf{C}_{\mathbf{v}_a \mathbf{z}_b}(i_a)$  as follows:

$$\boldsymbol{\Sigma}_{\mathbf{v}_a|i_a} = \frac{1}{h(i_a)} \mathbf{C}_{\mathbf{v}_a}(i_a) - \frac{1}{h(i_a)^2} \mathbf{s}_{\mathbf{v}_a}(i_a) \cdot \mathbf{s}_{\mathbf{v}_a}(i_a)^T, \quad (6.14)$$

$$\boldsymbol{\Sigma}_{\mathbf{v}_a \mathbf{z}_b | i_a} = \frac{1}{h(i_a)} \mathbf{C}_{\mathbf{v}_a \mathbf{z}_b}(i_a) - \frac{1}{h(i_a)^2} \mathbf{s}_{\mathbf{v}_a \mathbf{z}_b}(i_a) \cdot \mathbf{s}_{\mathbf{v}_a \mathbf{z}_b}(i_a)^T, \quad (6.15)$$

where  $\mathbf{s}_{\mathbf{v}_a}(i_a)$ ,  $\mathbf{s}_{\mathbf{v}_a \mathbf{z}_b}(i_a)$ ,  $\mathbf{C}_{\mathbf{v}_a}(i_a)$  and  $\mathbf{C}_{\mathbf{v}_a \mathbf{z}_b}(i_a)$  are obtained by summing<sup>4</sup> the contributions of feature values at corresponding positions  $x_j$  and  $t_j = \mathfrak{T}(x_j)$ :

$$\mathbf{s}_{\mathbf{v}_a}(i_a(x_j)) += \mathbf{v}_a(x_j), \quad (6.16)$$

$$\mathbf{s}_{\mathbf{v}_a \mathbf{z}_b}(i_a(x_j)) += (\mathbf{v}_a(x_j), \mathbf{z}_b(t_j)), \quad (6.17)$$

$$\mathbf{C}_{\mathbf{v}_a}(i_a(x_j)) += \mathbf{v}_a(x_j) \cdot \mathbf{v}_a(x_j)^T, \quad (6.18)$$

$$\mathbf{C}_{\mathbf{v}_a \mathbf{z}_b}(i_a(x_j)) += (\mathbf{v}_a(x_j), \mathbf{z}_b(t_j)) \cdot (\mathbf{v}_a(x_j), \mathbf{z}_b(t_j))^T. \quad (6.19)$$

Once the covariance matrices are defined, entropies are obtained by (6.7) and (6.9).

---

<sup>4</sup> *value*  $+=$  *increase* is equivalent to *value*  $=$  *value*  $+$  *increase*

The more alternatives, the more difficult  
the choice.

*Abbe' D'Allanival*

## **Chapter 7**

# **Evaluation of similarity measures for reconstruction- based registration in image- guided radiotherapy and surgery**

Darko Škerl, Dejan Tomažević, Boštjan Likar and Franjo Pernuš  
International Journal of Radiation Oncology Biology Physics 65: 943-953  
(2006)

### **Abstract**

*A promising patient positioning technique is based on registering CT or MR images to cone-beam CT images (CBCT). The extra radiation dose delivered to patient can be substantially reduced by using fewer projections. This approach results in lower quality CBCT images. The purpose of this study is to evaluate a number of similarity measures (SMs) suitable for registration of CT or MR images to low-quality CBCTs. Using the recently proposed evaluation protocol we evaluate nine SMs with respect to pretreatment imaging modalities, number of 2D images used for reconstruction and number of reconstruction iterations. The image database consisted of 100 X-ray and corresponding CT and MR images of two vertebral columns. Using a higher number of 2D projections or reconstruction iterations results in higher accuracy and slightly lower robustness. The similarity measures that behaved the best also yielded the best registration results. The most appropriate similarity measure was the asymmetric multi-feature mutual information (AMMI). The evaluation protocol proved to be a valuable tool for selecting the best similarity measure for the reconstruction-based registration. The results indicate that accurate and robust*

CT/CBCT or even MR/CBCT registrations are possible if the AMMI similarity measure is used.

## 7.1 Introduction

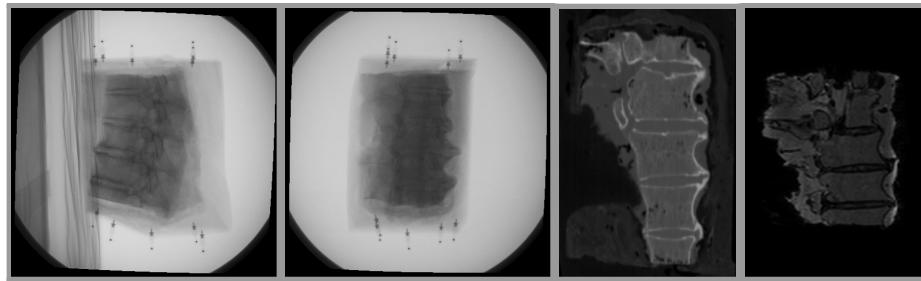
The precise localization of target and sensitive surrounding anatomical structures, at both the pre- and intra-treatment stages represent the most significant challenges to the implementation of three-dimensional (3D) conformal radiation therapy, intensity-modulated radiotherapy (IMRT), and minimally invasive surgery in routine clinical use. With the tremendous development of 3D imaging modalities such as computed tomography (CT), magnetic resonance (MR), and positron emission tomography (PET), finer and finer pre- and intra-treatment localization of different anatomical structures has been made possible. While pre-treatment localization is based on image segmentation, intra-treatment localization or verification of patient setup is based on registration of pre- and intra-treatment information, which is usually the information encountered in images. Registration allows any 3D point defined in the pre-treatment image to be located in a patient, treatment machine, or treatment room coordinate system. Modern planning and execution of radiotherapy or minimally invasive surgery are critically dependent on image guidance, which comprises image acquisition, segmentation, registration, and visualization (Jaffray, 2005; Mackie *et al.*, 2003; Peters, 2000).

Different registration strategies for on-line verification of patient setup in the treatment room have been used in the past and new ones are constantly being proposed. A more recent and popular registration method in image-guided radiotherapy and surgery is based on registration of two-dimensional (2D) electronic portal images or X-ray images to digitally reconstructed radiographs (DRRs) (Gilhuijs *et al.*, 1996a; Lemieux *et al.*, 1994). DRRs are X-ray projection images computed for arbitrary angles of projection through the CT image (Goitein *et al.*, 1983; Sherouse *et al.*, 1990). The unknown pose of the pre-treatment CT volume relative to the portal or fluoroscopic X-ray image set, defining the patient in the treatment room, may be automatically estimated by optimizing the similarity measure calculated from DRR and corresponding portal or fluoroscopic X-ray images (Clippe *et al.*, 2003; Lemieux *et al.*, 1994; Penney *et al.*, 1998; Sirois *et al.*, 1999; Weese *et al.*, 1997). Similarity measure is a function that associates a numeric value with a pair of images with the idea that a higher, or for some similarity measures lower, value indicates greater similarity. Similarity measure can be looked upon as a K-dimensional function, with K the number of parameters of the spatial transformation model. In the ideal case, the K-dimensional parametrical space, in which the value at each location corresponds to the value of the similarity measure for that transformation estimate, would contain a sharp maximum (minimum) with monotonically decreasing (increasing) values with distance away from the optimum. The range of transformations around the optimum for which the value of the similarity measure is a monotonic function of misregistration is called the capture range. The distance from an optimum, to which registration

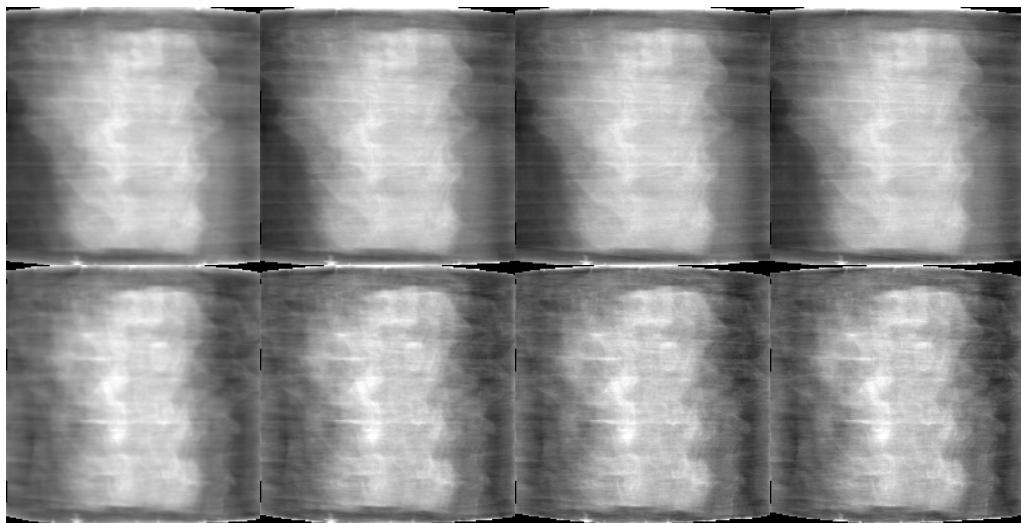
converged, to the true registration position is the registration accuracy. Some of the local optima may be very small, caused by either interpolation artefacts and/or by local good matches between voxel intensities. Robustness is the ability of a registration method to produce similar results on all trials, i.e. regardless of the starting position, implementation details, and differences in image contents. Ideally, a similarity measure, together with the optimization method, should provide accurate and robust registrations. Registrations that are based on similarity measures adjust the parameters of an appropriate spatial transformation model until the similarity measure reaches an optimum.

Unfortunately, the intensity-based registration of DRRs and corresponding portal or fluoroscopic X-ray images is slow due to time consuming calculation of DRRs. To some extent the problem may be reduced by calculating DRRs containing only structures of interest (Weese *et al.*, 1997) or by implementing a faster DRR generation method (Rohlfing *et al.*, 2005b). The other drawback of this method is that it is not best suitable for registration of MR to portal or fluoroscopic X-ray images, as there is practically no correlation between MR-based DRRs and X-ray projections (Tomažević *et al.*, 2003). Probably the most important drawback of the DRR-based registration method is that by projecting a high quality 3D CT image into 2D to generate a DRR image, valuable information, needed for accurate and robust registration, is lost.

These drawbacks have motivated the development of novel registration methods, based on reconstruction of a 3D volume, called cone-beam CT (CBCT), from a series of 2D projection images, and registration of pre-treatment CT or MR to CBCT images (Berbeco *et al.*, 2004; Ford *et al.*, 2002; Jaffray *et al.*, 1999; Pouliot *et al.*, 2005; Tomažević *et al.*, 2006). By this approach, the similarity is computed between a pair of 3D images instead of between numerous 2D image pairs, which is the case for DRR-based registration. In radiation therapy, two approaches have been considered based on whether kilo-voltage (kV) or mega-voltage (MV) beams are used to form the image (Groh *et al.*, 2002; Jaffray, 2005). For kV CBCT, a separate kV source and a second dedicated electronic portal imaging device (EPID), attached perpendicular to the treatment beam, have been proposed to acquire the projection images (Jaffray *et al.*, 1999; Letourneau *et al.*, 2005). For MV CBCT, the MV source of the linac and EPID currently employed to acquire portal images are used for imaging (Ford *et al.*, 2002; Midgley *et al.*, 1998; Pouliot *et al.*, 2005). In image-guided minimally invasive surgery, a C-arm has been used to acquire fluoroscopic X-ray images for CBCT formation (Siewerdsen *et al.*, 2005; Tomažević *et al.*, 2006). An important clinical issue of these techniques is the extra radiation dose delivered to the patient with each setup verification scan. To keep radiation to which the patient is exposed as low as possible, it is desired that CBCT images are reconstructed from as few projection images as possible. As a consequence, the quality of a 3D image reconstructed from a small number of projection images will be low. The similarity measure applied to automatically register the pre-treatment CT or



**Figure 7.1.** Two orthogonal X-ray projections and one slice of CT and MR images.



**Figure 7.2.** Images reconstructed from 4 (top) and 8 (bottom) projections. The number of iterations of the reconstruction algorithm was 1, 3, 5 and 7 (left to right).

MR image to the CBCT image must therefore be able to cope with low image quality.

The similarity measure that nowadays enjoys the reputation of an accurate and robust multimodality image registration criterion is mutual information (Maes *et al.*, 1997; Pouliot *et al.*, 2005; Wells *et al.*, 1996). However, even registrations based on mutual information will often fail if images contain insufficient information. Failure of registration is even more likely if an MR image is to be registered to the CBCT image. Probably the only way to cope with the insufficient information is to add spatial information, for instance in the form of intensity gradients, into the registration process (Tomažević *et al.*, 2006).

We have conducted two experiments. To be able to choose the best similarity measure for a reconstruction-based registration using a small number of projection images, we have first analyzed the behavior of nine similarity measures for rigid registration of CT and MR images to CBCT images. The behavior of each similarity function was analyzed using publicly available spine phantom image data (van de Kraats *et al.*, 2005a) and a recently proposed protocol for evaluation of similarity measures (Škerl *et al.*, 2006a). The protocol allows estimation of a similarity measure's capture range, number and extent of local optima, and the accuracy and distinctiveness of the global optimum. Second, to show that the information obtained by the evaluation protocol in the first experiment is valuable for choosing the best similarity measure for reconstruction-based registration, we have tested the performance of the reconstruction-based registration method (Tomažević *et al.*, 2006) using each of the nine similarity measures. The reconstruction-based registration method has been tested on the same data using standardized evaluation methodology (van de Kraats *et al.*, 2005a) by which registration accuracy and robustness were assessed.

## 7.2 Materials and methods

### 7.2.1 Experimental data

The publicly available image database<sup>5</sup> consisted of 2D X-ray images and corresponding 3D 3DRX, CT and MR images of two defrosted segments of vertebral bodies (van de Kraats *et al.*, 2005a). The first vertebral column consisted of three thoracolumbar vertebrae bodies while the second segment consisted of five thoracic vertebrae bodies. The X-ray images were acquired by a clinical floor-mounted 3D rotational X-ray (3DRX) C-arm system (Integris BV5000, Philips Medical Systems, Best, The Netherlands). During an 8-s run of 180° around each phantom two sets of 100 projection images were acquired and used to reconstruct two high resolution 3D volumes using a filtered back-projection reconstruction technique (Grass *et al.*, 1999). As the C-arm was calibrated, the projection geometry of the X-ray images with respect to the reconstructed volume was known. Therefore, the geometric relationship between a 3DRX volume and the 100 corresponding X-ray images was known. The CT-images of the two vertebral columns were acquired with a clinical 16-detector-row multi-slice CT scanner (MSCT, Philips Medical System, Best, The Netherlands). The MR images were obtained with a clinical 1.5 Tesla MR scanner (Gyroscan NT, Philips Medical System, Best, The Netherlands) using a sagittal 3D turbo spin echo acquisition and turbo factor of 29, TR/TE of 1500 ms/90 ms. The “gold standard” 3D/3D registrations of 3DRX to corresponding CT or MR data were obtained by maximization of mutual information (Maes *et al.*, 1997). Figure 7.1 shows two orthogonal X-ray image projections and a projection of CT and MR images of the first spine phantom. Figure 7.2 shows

<sup>5</sup> <http://www.isi.uu.nl/Research/Databases/>

slices of CBCT images reconstructed with different parameters. Eight volumes of interest (VOIs) were manually defined in the 3DRX volume, by determining a rectangular region of interest around each of the 8 vertebra. The VOIs were transformed from the 3DRX to corresponding CT or MR data sets using the “gold standard” transformations. The VOIs were not present in CBCT images.

### 7.2.2 Cone-beam CT

To study the influence of the number of X-ray images on the quality of the CBCT image and, consequently, on the behavior of the similarity measures and quality of a reconstruction-based registration we have reconstructed a number of CBCTs using different subsets of X-ray images out of the 2 sets of 100 X-ray images. Generally, the more X-ray images are used for reconstruction the better is the reconstructed image. The iterative simultaneous algebraic reconstruction technique (SART) was used for reconstruction (Andersen and Kak, 1984). SART, which is slow but simple to implement generally provides good results in situations when a small number of projections is available and/or when projections are not uniformly distributed. Usually, just a few iterations are enough to obtain useful reconstruction results, while increasing the number of iterations may increase reconstruction artifacts, which are caused by computation errors that are amplified in each iteration. Increasing the number of reconstruction iterations also results in a more time consuming reconstruction. Besides the number of images used for reconstruction we have therefore also analyzed the impact of the number of iterations on the behavior of similarity measures.

### 7.2.3 Similarity measure evaluation protocol

The evaluation protocol requires that images, typical for a specific registration task, and “gold standard” registration of these images are given, and that a spatial transformation model is selected (Škerl *et al.*, 2006a). The spatial transformation  $T$  that is supposed to bring two images, a floating and a target image, into correspondence is assumed to be rigid and is therefore composed of three translational ( $t_x, t_y, t_z$ ) and three rotational ( $\Theta_x, \Theta_y, \Theta_z$ ) parameters. The six-dimensional parametrical space is first normalized so that equal changes of each of the 6 parameters in the normalized parametrical space will have approximately equal impact on the transformation magnitude. By normalizing the parametrical space Euclidean metrics may be used to determine distances from the position at which the images are “best” aligned and where a similarity function should have its optimum. Let the origin  $\mathbf{X}_0$  of the 6-dimensional parametrical space be at the known “gold standard” position and let  $SM(\mathbf{X})$  be the value of a similarity measure for the spatial transformation defined by location  $\mathbf{X}; \mathbf{X}=[x_1, \dots, x_6]$  in this space. Similarity measure values  $SM(\mathbf{X}_{n,m})$  are defined for image pairs, with the target image at the origin  $\mathbf{X}_0$  and the floating image transformed from the origin to location  $\mathbf{X}_{n,m}$ . Values  $SM(\mathbf{X}_{n,m}), n=1,2,\dots,N; m=-M/2,\dots,M/2$ , of a similarity measure are defined on  $N$

lines probing the six dimensional parametrical space and at  $M+1$  points evenly spaced along each line. Each of the  $N$  lines is defined by a randomly selected starting position  $\mathbf{X}_{n,M/2}$  at a distance  $R$ ,  $R = \|\mathbf{X}_{n,M/2}\|$ , from the origin and its mirror point  $\mathbf{X}_{n,M/2}$ . To make the similarity measure invariant to the absolute scale, each original similarity measure  $SM_0(\mathbf{X}_{n,m})$  is normalized to the interval  $[0, 1]$ :

$$SM(\mathbf{X}_{n,m}) = \frac{SM_o(\mathbf{X}_{n,m}) - SM_{0min}}{SM_{0max} - SM_{0min}}, \quad (7.1)$$

where  $SM_{0min}$  and  $SM_{0max}$  are the minimal and maximal values of  $NM+1$  similarity measure values before normalization, respectively. If a similarity function is such that its minimum is sought for by optimization, the similarity measure value  $SM(\mathbf{X}_{n,m})$  is changed to  $-SM(\mathbf{X}_{n,m})$ . Let  $\mathbf{X}_{n,opt}$  be the position and  $SM(\mathbf{X}_{n,opt})$  the value of the global optimum of the similarity measure along line  $n$ , and let  $\mathbf{X}_{n,loc}$  be the position of the local optimum closest to  $\mathbf{X}_{n,opt}$ .

The behavior of a similarity measure is assessed by five properties. All properties are statistical estimations, derived from the “gold standard” position, similarity measure values  $SM(\mathbf{X}_{n,m})$  and positive gradients  $d_{n,m}$ :

$$d_{n,m} = \begin{cases} SM(\mathbf{X}_{n,m-1}) - SM(\mathbf{X}_{n,m}) & \text{if } m < opt \text{ \& } SM(\mathbf{X}_{n,m-1}) > SM(\mathbf{X}_{n,m}) \\ SM(\mathbf{X}_{n,m+1}) - SM(\mathbf{X}_{n,m}) & \text{if } m > opt \text{ \& } SM(\mathbf{X}_{n,m+1}) > SM(\mathbf{X}_{n,m}) \\ 0 & \text{otherwise} \end{cases}. \quad (7.2)$$

The five properties are:

1. Accuracy  $ACC$  of a similarity measure is defined as the root mean square (RMS) of distances  $\|\mathbf{X}_{n,opt} - \mathbf{X}_0\|$  between the origin  $\mathbf{X}_0$  and global optima  $\mathbf{X}_{n,opt}$ ,  $n=1,2,\dots,N$ :

$$ACC = \sqrt{\frac{1}{N} \sum_{n=1}^N \|\mathbf{X}_{n,opt} - \mathbf{X}_0\|^2} [\text{mm}]. \quad (7.3)$$

2. Distinctiveness of optimum  $DO$  is defined as a function of distance  $r$ ,  $r=s\cdot\delta_x$ , from the optimum, where  $\delta_x$ ,  $\delta_x = 2R/M$ , is the distance between two consecutive points along a line and  $s$ ,  $s=1,2,\dots$  is the number of steps:

$$DO(r) = \frac{1}{2rN} \sum_{n=1}^N [2 * SM(\mathbf{X}_{n,opt}) - SM(\mathbf{X}_{n,opt-s}) - SM(\mathbf{X}_{n,opt+s})] [10 - 3 / \text{mm}]. \quad (7.4)$$

3. Capture range  $CR$  is defined as the minimal distance between the position of optima  $\mathbf{X}_{n,opt}$  and the closest minimum  $\mathbf{X}_{n,loc}$ :

$$CR = \min_n(\|\mathbf{X}_{n,opt} - \mathbf{X}_{n,loc}\|) [\text{mm}]. \quad (7.5)$$

4. Number of minima  $NOM(r)$  is the sum of minima of the similarity measure within distance  $r$  from each of the  $N$  global optima, i.e. a cumulative number of minima as a function of distance  $r$ . The average number of minima per line  $NOM(r)/N$  is denoted by  $NOM$ .
5. Risk of nonconvergence  $RON(r)$  is the property that describes the behavior of a similarity measure around the  $N$  global optima. It is defined as the average of positive gradients  $d_{n,m}$  within distance  $r$  from each of the  $N$  global optima:

$$RON(r) = \frac{1}{2rN} \sum_{n=1}^N \sum_{m=opt-s}^{opt+s} d_{n,m} [10^{-6}/\text{mm}]. \quad (7.6)$$

A large value of  $RON(r)$  indicates that a similarity measure has distinctive (deep) and/or broader local optima in which optimization may get trapped.  $RON(r)$  is shorter denoted as  $RON$ .

The better a similarity measure is, the smaller are the values of the accuracy, number of minima, and risk of nonconvergence and the larger the capture range and distinctiveness of optimum values. CR, NOM and RON are the three properties that describe robustness. The better these values are the more robust is the similarity measure. More details on the evaluation protocol can be found in (Škerl *et al.*, 2006a) and in the online version at <http://lit.fe.uni-lj.si/Evaluation>.

#### 7.2.4 Similarity measures

We have implemented and evaluated the following 9 similarity measures:

1. Asymmetric gradient based mutual information AMMI (Tomažević *et al.*, 2006) allows to incorporate multi-variable image features (image intensity, image intensity gradients, ...) into a single information-based similarity measure. AMMI is based on the assumption that just one of the floating image features has an arbitrary distribution, which is estimated by intensity histogram, while the distributions of all other floating and all target image features are normal.
2. Symmetric gradient based mutual information SMMI (Tomažević *et al.*, 2006) also allows to incorporate multi-variable image features into a single information-based similarity measure, by assuming that one floating and one target image feature are arbitrary distributed and can be estimated via histograms.

3. Mutual information MI (Maes *et al.*, 1997; Wells *et al.*, 1996) is a single feature similarity measure which measures the amount of information that one image contains about the other. The influence of the size of the overlapping region between the images during registration is taken into account.
4. Normalized mutual information NMI (Studholme *et al.*, 1999) also measures the amount of information between the images. The influence of the size of the overlapping region between the images during registration is smaller than for MI.
5. Entropy correlation coefficient ECC (Astola and Virtanen, 1982; Maes *et al.*, 1997) is a measure of statistical dependence between the images.
6. Joint entropy H (Collignon *et al.*, 1995; Studholme *et al.*, 1995) measures the amount of information that we have in the combined images. The more similar the images are, the lower is the joint entropy compared to the sum of individual entropies. The important limitation of this measure is that it does not deal with the overlap problem.
7. Correlation ratio COR (Roche *et al.*, 1998) is an asymmetric measure of functional dependence between the images.
8. Woods criterion WC (Woods *et al.*, 1993), or partitioned intensity uniformity, first partitions the target image into isointensity sets (or histogram bins) based on voxel values. Then it optimizes the uniformity of the floating image's voxel values within each bin.
9. Pearson's correlation coefficient PCC is a symmetric measure of linear dependence between the images. The optimization of linear dependence limits the use of this measure to registration of mono-modal images.

### 7.3 Experiments and results

In all the following experiments the number of lines  $N$  was set to 50,  $R$  to 35 mm,  $M$  to 140,  $\delta_x$  to 0.5 mm, and  $s$  to 1. All similarity measures were applied to overlapping voxels of the floating and target images. In all experiments the CBCT image served as the target image and CT or MR VOIs as the floating images. Similarity measures were formulated on the 2D joint histogram or joint probability distribution of the intensities of the two images. The joint histogram was obtained using 64 bins and partial volume interpolation (Maes *et al.*, 1997). As similar trends were observed for each of the eight vertebral images we present results only for the first vertebrae of the first vertebral column.

**Table 7.1.** Properties of 9 similarity measures in CT/CBCT and MR/CBCT registrations.

	CT/CBCT					MR/CBCT				
	ACC	DO	CR	NOM	RON	ACC	DO	CR	NOM	RON
AMMI	<b>0.11</b>	<b>26.68</b>	25.6	0.14	<b>6.8</b>	0.96	7.00	14.41	0.26	<b>31.1</b>
SMMI	0.24	26.01	<b>18.1</b>	1.58	659.3	<b>0.43</b>	<b>14.17</b>	10.43	1.72	1486.0
MI	0.79	8.43	25.6	0.12	26.8	1.24	3.15	<b>24.84</b>	0.18	86.2
NMI	0.82	9.07	25.6	0.12	27.1	1.20	3.37	<b>24.84</b>	<b>0.14</b>	80.6
ECC	0.82	5.82	25.6	0.14	36.4	1.20	1.93	<b>24.84</b>	0.18	137.4
H	0.84	7.04	23.3	0.28	110.8	0.85	2.59	0.50	0.34	142.7
COR	1.16	1.36	<b>27.9</b>	<b>0.10</b>	14.2	1.24	0.9	0.50	0.24	50.0
WC	<b>1.17</b>	2.15	23.3	0.50	437.3	0.96	1.3	0.50	0.66	827.1
PCC	0.61	4.92	27.1	<b>0.10</b>	15.0	4.52	1.4	0.50	0.48	113.5

### 7.3.1 The impact of pre-treatment imaging modalities

First, we have analyzed how the similarity measures behaved if CT or MR images had been registered to the CBCT image. To analyze the similarity measures in CT/CBCT image registrations, the CBCT image was reconstructed from 4 X-ray images with one reconstruction iteration while for MR/CBCT registrations the CBCT image was reconstructed from 8 X-ray images with three reconstruction iterations. More images and iterations were used for reconstructing the CBCT to be registered with MR because these modalities differ more than the CT and CBCT. Table 7.1 shows the values of five properties of the 9 similarity measures. The numbers in bold and italic correspond to the best and the worst similarity measure, respectively, for a specific criterion. In CT/CBCT registrations, AMMI was the most accurate, had the most distinctive optima, the smallest RON, high CR, and small NOM. PCC had the smallest NOM indicating that it is robust but was not as accurate as AMMI and SMMI, which on the other hand was not robust as indicated by the large values of NOM and RON. For MR/CBCT registrations AMMI again had the smallest RON but was outperformed in terms of accuracy and distinctiveness of optimum by the SMMI similarity measure. The other measures that proved to be robust were MI and NMI. As expected, because of differences in modalities almost all properties of all similarity measures were worse in cases of MR/CBCT image registrations.

To further confirm the properties of similarity measures we have conducted a registration-based experiment. The number of images used for reconstruction and the number of iterations were the same as in the first experiment. Registrations incorporating different similarity measures were evaluated using the standardized methodology for 3D/2D registration evaluation (van de Kraats *et al.*, 2005a). The evaluation methodology uses the mean target registration error (mTRE) to measure the distance of a VOI's position from

**Table 7.2.** Mean TREs and success rates (SR) of registrations using different similarity measures and modalities.

	CT/CBCT		MR/CBCT	
	mTRE[mm]	SR[%]	mTRE[mm]	SR[%]
AMMI	<b>0.3347</b>	96.0	1.2897	<b>91.5</b>
SMMI	0.6643	88.0	<b>0.4070</b>	76.5
MI	1.3413	94.5	1.8480	82.0
NMI	1.4069	94.5	1.8303	47.5
ECC	1.3517	96.0	1.8369	83.0
H	1.6328	89.5	1.8200	33.5
COR	1.7421	<b>96.5</b>	>	0
WC	<b>1.9676</b>	39.5	>	0
PCC	1.2692	92.5	>	0
Combination	0.6471	93.0	0.4149	93.0

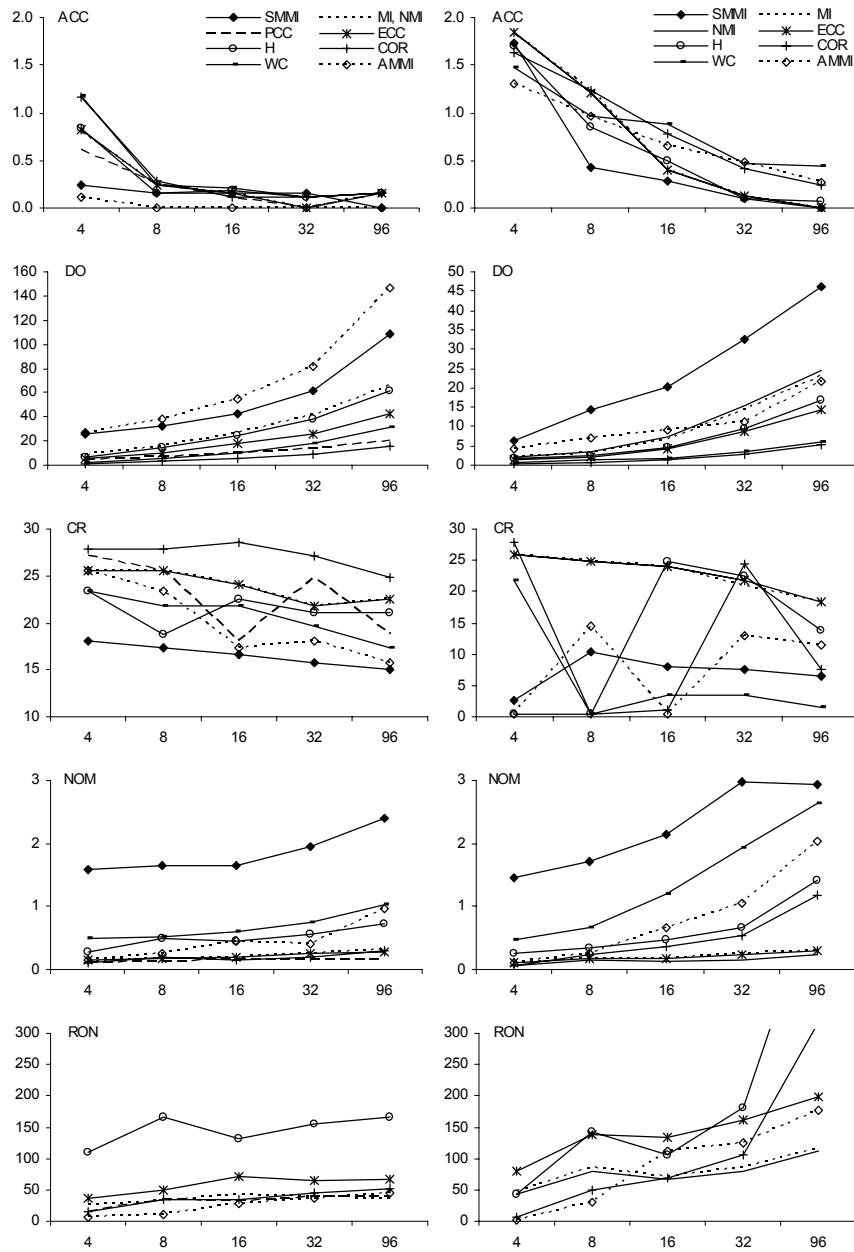
“gold standard” before and after registration. The positions of all image elements in a VOI were used as target points. To evaluate the robustness of the CT or MR to CBCT image registration method, the methodology requires that for each VOI and corresponding CBCT image a large number of registrations are performed from different starting positions. Van de Kraats *et al.* (van de Kraats *et al.*, 2005a) provided 200 starting positions for each of the 8 VOI. The starting positions were randomly generated around the “gold standard” position in such a way that the distance from “gold standard” measured by mTRE was uniformly distributed in the interval of 0 to 20 mm, with 10 positions in each of the 1 mm wide subintervals. Powell optimization scheme was used as the search strategy. Each registration was considered successful if mTRE after registration was lower than the predefined clinically relevant threshold, which had been set to 2 mm. The registration error was defined as mTRE of all successful registrations. The results in Table 7.2 are highly correlated with the results in Table 7.1. The success rate of COR, WC and PCC similarity measures applied in MR/CBCT registration was 0% as TREs were above 2 mm in all 200 registrations. This is indicated by a symbol ‘>’ in Table 7.2. This could be predicted from results in Table 7.1, as COR, WC and PCC were neither accurate nor robust. The results in Table 7.1 also well predicted that AMMI was accurate and robust in CT/CBCT registrations and that SMMI was the most accurate measure for MR/CBCT image registrations.

We have also analyzed the outcome of CT and MR to CBCT registrations utilizing a combination of two similarity measures. We have started the CT/CBCT registrations with the PCC similarity measure, which, as Table 7.1 indicates, is supposed to be very robust. After reaching the optimum, PCC was replaced by the SMMI similarity measure. The same experiment was performed for MR/CBCT registrations, except that COR was used instead of

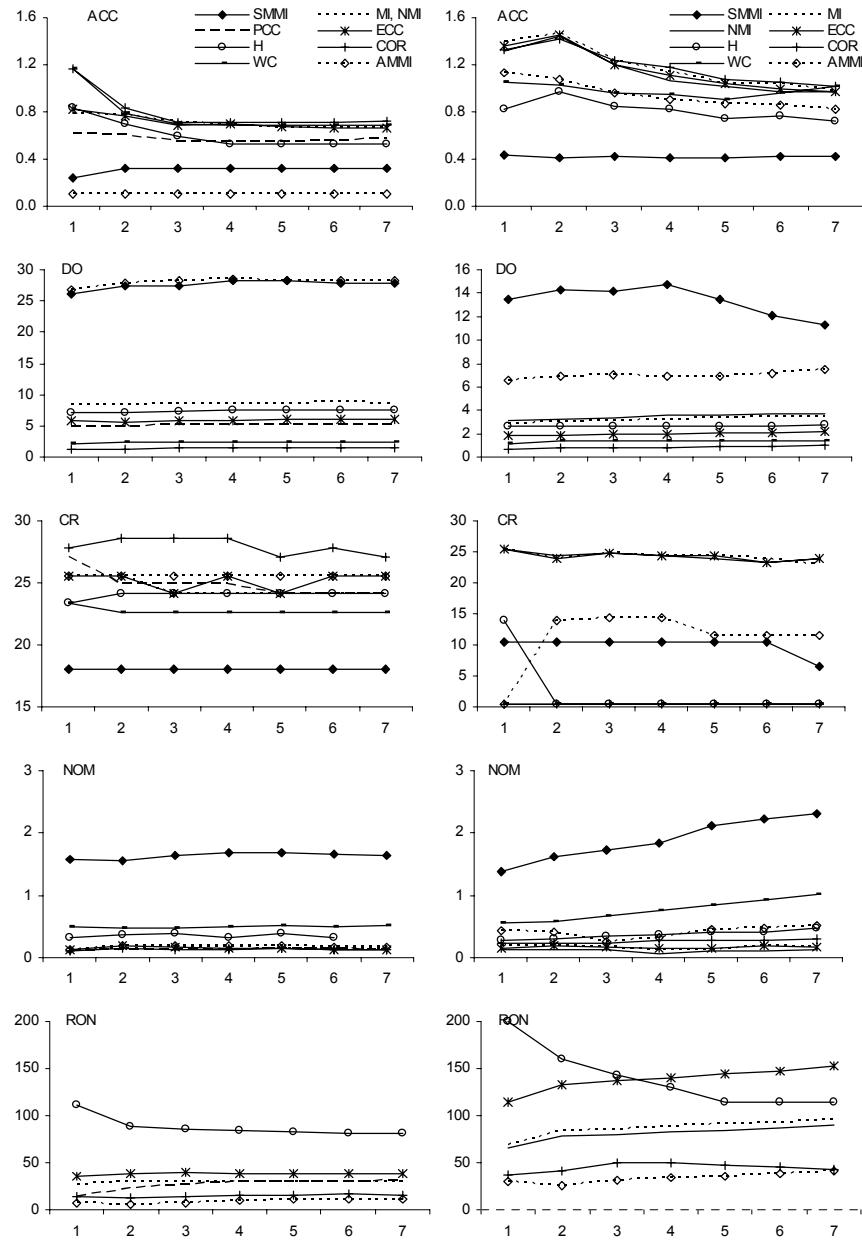
PCC as the initial similarity measure as suggested by results in Table 7.1. Optimization of the COR similarity measure was not very successful in terms of the mTRE (success rate) because in all 200 cases the optima had been further away from the “gold standard” than the predefined threshold of a successful registration. Because, according to the results in Table 7.1, COR was robust but very inaccurate it could be used first to find the initial position from which the registration with the SMMI similarity measure, which is very accurate but not so robust, is run. Indeed, the combination of COR and SMMI, as Table 7.2 indicates, resulted in a very accurate and robust registration. Results of this experiment, shown in the last row of Table 7.2, indicate that with a combination of similarity measures a much higher success rate can be achieved with SMMI without compromising the accuracy of registrations obtained with SMMI.

### **7.3.2 The impact of the number of 2D images used for reconstruction**

In this experiment the properties of the similarity measures were evaluated with respect to the number of projections used for reconstruction. CBCT images were reconstructed from 4, 8, 16, 32 and 96 projections using one, if intended for registration with CT, and 3 reconstruction iterations for registration with MR images. The results shown in Figure 7.3 indicate that the accuracy and distinctiveness of optimum of all similarity measures improved when more images had been used for reconstruction. However, with more images the measures became slightly less robust, as indicated by the increase of NOM and RON values. This can be explained with more details presented in 3D images reconstructed from more 2D projections. Because with more images the accuracy improved and the robustness slightly deteriorated, the optimal number of images used for reconstruction is between 8 and 16 for CT and around 16 for MR images. Again, the similarity measures behaved better in CT/CBCT than in MR/CBCT registrations. The number of 2D projections had a greater impact on MR/CBCT than on CT/CBCT registrations. This was most probably due to larger differences between the MR and CBCT modalities.



**Figure 7.3.** Five properties of the similarity measures (top to bottom) as a function of the number of 2D images used for CT/CBCT (left column) and MR/CBCT (right column) registrations. The RON for SMMI and WC is out of scale for both CT and MR to CBCT registrations. For MR to CBCT registration the ACC of PCC was out of scale so it is not shown in the charts.



**Figure 7.4.** Five properties (top to bottom) of similarity measures as a function of the number of iterations for CT/CBCT (left column) and MR/CBCT (right column) registrations. The RON for SMMI and WC is out of scale for both CT/CBCT and MR/CBCT registrations. For MR to CBCT registration the ACC of PCC was out of scale so it is not shown in the charts.

### 7.3.3 The impact of the number of reconstruction iterations

In this experiment we varied the number of iterations from 1 to 7 and analyzed the impact of iterations on the behavior of similarity measures. The CBCT image was reconstructed from 4 and 8 projection images for registration with CT and MR, respectively. Results in Figure 7.4 indicate that the behavior of the similarity measures did not improve when more than 3 iterations had been used to reconstruct the CBCT image. The number of iterations had almost no impact on the robustness and *DO* while the accuracy of some of the measures, like H, WC, and COR, improved if three instead of one or two iterations had been used. For MR/CBCT registration the similarity measures were the most accurate if three or more iterations of SART had been used while the robustness slightly deteriorated with more iterations. Considering the obtained results and corresponding computational demands, between 3 and 5 reconstruction iterations seem to be the most appropriate.

## 7.4 Discussion

The reconstruction of a CBCT from kV or MV beams and its registration to a CT or MR image has been recently proposed in radiation therapy and minimally invasive image-guided surgery for patient setup verification (Berbeco *et al.*, 2004; Ford *et al.*, 2002; Jaffray *et al.*, 1999; Pouliot *et al.*, 2005; Tomažević *et al.*, 2006). The advantages of CBCT over planar megavoltage radiographs are numerous. CBCT images are three-dimensional, they have better soft tissue contrast and they are easier to compare with planning CT images (Mackie *et al.*, 2003). By registering two 3D images it is avoided that valuable 3D information, needed for accurate and robust registration, is lost. This is the case with DRR-based registrations where a high quality 3D CT image is projected into 2D to generate a DRR image. An important clinical issue of reconstruction-based registration is the extra radiation dose delivered to the patient. The total dose required to perform kV or MV CBCT is approximately the product of the number of projection images and the dose delivered for each image. In the study of Pouliot *et al.* (Pouliot *et al.*, 2005) the MV CBCT image was reconstructed from 180 projection images and the radiation delivered to the patient was 15cGy. Jaffray *et al.* (Jaffray *et al.*, 1999) reconstructed a MV and kV CBCT from 90 MV and 195 kV projections of a head phantom, respectively. The central axis doses delivered in the MV and kV CBCT imaging procedures were 138 cGy and 3.1 cGy, respectively. In (Jaffray *et al.*, 2002) a kV CBCT of an anthropomorphic phantom was reconstructed from 321 projections. By decreasing the dose below 2cGy, CBCT could be performed daily and the dose to the patient would not significantly increase over the current standard. The dose can be reduced either by developing more sensitive flat-panel X-ray detectors or by reconstructing images from fewer projections. The later approach would result in CBCT images of lower quality. Therefore, if a similarity measure based registration is to be used to register a low quality CBCT image with the treatment planning CT image, the similarity measure should be able to cope

with low image quality. Given the images to be registered and the spatial transformation model, the quality of registration mainly depends on the similarity measure and the optimization method.

Each of the numerous similarity measures introduced in the past has different properties, and is thus differently sensitive to the imaging modality, image content, sampling, interpolation, histogram binning, partial image overlap, image degradation, such as noise, intensity inhomogeneities and geometrical distortions. To be able to choose the most appropriate similarity measure and its implementation, as well as the optimization method, it is often desirable to have some a-priori information on the behavior of the similarity function with regard to the above mentioned factors. A thorough analysis of the similarity function at every point of the parametrical space is impossible because the parametrical space is too large to be analyzed even in the simplest case of rigid registration of 2D images. Limited information on the similarity measure may be obtained either indirectly from the accuracy and robustness of the registration itself or more directly, by drawing plots, showing their behavior when one image is systematically translated from and/or rotated around the “gold standard” registration position. Unfortunately, the information obtained by the latter approach is qualitative. To obtain more a-priori and quantitative information on the behavior of a similarity measure, we have recently developed an optimization-independent similarity measure evaluation protocol (Škerl *et al.*, 2006a). By this protocol a similarity measure’s capture range, number and extent of local optima, and the accuracy and distinctiveness of the global optimum are estimated independently of optimization. We believe that the proposed properties provide valuable quantitative information on the behaviour of a similarity measure that can help researchers in deciding which similarity measure and which optimization protocol to use in a given application. However, it is clear that all possible behaviours of a similarity measure cannot be assessed neither by these five features nor by any other ones. The first three properties characterize the robustness of a similarity measure. If within the space of possible geometrical transformations a similarity function is smooth, then it is less likely that registration will be influenced by the choice of the optimization scheme. The number of minima (*NOM*) and risk of nonconvergence (*RON*) provide complementary information on convergence. A high value of *NOM* alone does not necessarily mean that a cost function is bad. The local optima can be very small and the optimization may thus not end in one of them. However, *NOM* and *RON* together estimate not only the number but also the extent of local optima and therefore provide more information on convergence. The number of local optima and risk of nonconvergence might thus prove useful in selecting an appropriate optimization scheme. The third property that is potentially useful for optimization implementation is distinctiveness of the optimum (*DO*).  $DO(r)$  estimates the average change of a similarity measure near the global optimum, i.e. the behavior of a similarity measure in the neighborhood, defined by distance  $r$ , of the global optimum. When approaching the global optimum, a similarity measure may rise (fall) steeply and then, after reaching the optimum fall (rise) quickly as well, causing a distinctive (sharp) peak at

the global optimum. Alternatively, it can rise and fall slowly, indicating an indistinctive (flat) optimum.  $DO(r)$  might help in selecting the stopping criterion of the optimization method. Our definition of the capture range ( $CR$ ) is rather strict. Because it is defined as the distance from the global optimum to the closest local optimum, regardless of its extent, it is actually the worst-case capture range. In practice, the optimization will probably not converge to a small local optimum. Besides, an optimization procedure will most probably not proceed along a line and the capture range may thus be larger than the  $CR$  obtained by the evaluation protocol. The capture range could be defined in many other ways. For instance,  $CR$  could be related to the distance  $r$  where  $RON(r)$  reaches a certain value or to the distance  $r$  where  $RON(r)$  abruptly raises.

In this paper, we have evaluated nine similarity measures used for rigid body registration of pre-treatment CT and MR to inter-treatment CBCT images. To be as fair as possible, all similarity measures were formulated on the 2D histogram or joint probability distribution of the two images to be registered. In this way the impact of interpolation and histogram creation did not bias the results. We have evaluated the behavior of each of the 9 similarity measures using the publicly available spine phantom CT, MR, and X-ray projection images for which “gold standard” registrations were available (van de Kraats *et al.*, 2005a). The results of the analysis of similarity measures are thus important in the context of CT or MR to kV CBCT registrations (Jaffray *et al.*, 1999; Jaffray *et al.*, 2002; Tomažević *et al.*, 2006).

Because CT images have excellent geometrical accuracy and provide electron density information directly, CT is nowadays the main imaging modality for treatment planning and patient setup verification. Results in Table 7.1 indicate that for CT to kV CBCT registrations the recently introduced AMMI similarity measure (Tomažević *et al.*, 2006) is the most suitable as it was the most accurate and robust and had the most distinctive optimum of all the 9 similarity measures analyzed. The AMMI similarity measure is based on image intensities and image intensity gradients of both images that are to be registered. The authors of AMMI believe that image intensities ensure larger capturing ranges while image intensity gradients preserve the accuracy of registration methods that use gradient features only (Tomažević *et al.*, 2006). The information on the behavior of similarity measures provided by the evaluation protocol correlated well with the accuracy and robustness of CT/CBCT registrations ran from 200 starting positions. The starting positions were uniformly distributed in the interval of 0 to 20 mm around the “gold standard” registration position. With AMMI a registration accuracy of 0.33 mm and 96% success rate has been achieved if the CBCT had been reconstructed from only 4 projections and with one reconstruction iteration.

Magnetic resonance imaging (MRI) provides superior image quality for soft tissue segmentation over CT and is used, for instance, for tumor and surrounding risk tissue segmentation in radiotherapy planning of brain tumors. However, the lack of electron density information, image distortion

leading to geometrical inaccuracies, and the differences between intensities of MR and portal images have precluded the more widespread use of MR images in radiotherapy treatment planning and patient setup verification. A solution for both accurate MR-based segmentation and CT-based dose calculation is acquisition of both CT and MR images and their registration. The use of MRI alone for treatment planning and patient setup verification would remove any MR/CT registration errors and would reduce treatment cost by avoiding redundant CT scans and save patient, staff, and machine time (Chen *et al.*, 2004). Beavis *et al.* (Beavis *et al.*, 1998) reported on brain tumor radiotherapy plans by assigning homogeneous electron density to the entire MR image. Chen *et al.* (Chen *et al.*, 2004) showed that no clinically significant differences were found between the MRI- and CT-based treatment plans using the same beam arrangements, dose constraints, and optimization parameters. Lee *et al.* (Lee *et al.*, 2003) compared dose distributions created on the homogeneous density and bulk-density assigned MR images to original CT images of the prostate. They have observed negligible differences in dose distribution between radiotherapy plans using bone+water CT number bulk-assigned image and original CT. Bone segmentation which is required for MR planning can be done manually or automatically. For patient setup verification using MR images registration methods have been proposed that are based on MR DRRs (Chen *et al.*, 2004; Ramsey *et al.*, 1999). In this paper, we have analyzed the similarity measures that would be appropriate for MR/CBCT registrations for patient setup verification. Results in Table 7.1 indicate that SMMI was the most accurate but, on the other hand, the least robust similarity measure. As in CT/CBCT registrations AMMI was the most robust because it had the smallest risk of nonconvergence and a small number of local optima. The information on the behavior of similarity measures provided by the evaluation protocol again well correlated with the accuracy and robustness of MR/CBCT registrations ran from 200 starting positions. In MR/CBCT registration the highest success rate (93%) and high accuracy (0.4 mm) were achieved with a combination of similarity measures. The registration was started with COR and, after reaching the optimum, continued with the SMMI similarity measure. That CT/CBCT registrations are more accurate and robust than MR/CBCT registrations was expected because of the larger differences between MR and CBCT modalities.

We have also analyzed the impact of the number of 2D projections that CBCT was reconstructed from on the behavior of the similarity measures. The accuracy and distinctiveness of the optimum improved significantly when CBCT had been reconstructed from more projections. This has been observed in both CT and MR to CBCT registrations. Unfortunately, the properties that describe robustness did not improve, they even slightly deteriorated. According to the results presented in Figure 7.3, the optimal number of projections for CT and MR to CBCT registrations is around 16. Even better registration results can be expected if the CT or MR images are first registered to the CBCT image reconstructed from 4 or 8 projections. The optimum of this registration will not be accurate but it can be used to initialize further CT or MR registrations with the CBCT image reconstructed from 16 projections.

More reconstruction iterations led to higher accuracy. The other properties of the 9 similarity measures were only slightly affected by the higher number of iterations. Unfortunately, increasing the number of reconstruction iterations also results in a more time consuming reconstruction procedure. The time needed for reconstruction depends on the size of the reconstructed volume, resolution of the projected image, number of projections and number of reconstruction iterations. The time needed to reconstruct an image increases linearly with the number of iterations and projections. In our case the reconstruction took around 5s per projection and iteration on a 2.8 GHz Pentium IV PC. Using two iterations would thus increase the time to about 10s.

## 7.5 Conclusion

Reconstruction-based registration for patient setup verification and the use of MR images for radiotherapy planning and setup verification are rather new and little explored. To keep the additional dose that a patient receives in such a scenario as low as possible the CBCT has to be reconstructed from a small number of projection images, either using kV or MV beams. We have analyzed the behavior of several similarity measures used for registration of CT and MR to CBCT images reconstructed with different reconstruction parameters. With the evaluation protocol we were able to find the best reconstruction setting and similarity measure. We have shown which of the presented similarity measures performs best for this kind of registration and for the images studied. It is, however, not clear that the similarity measure that performed best on images of the spine would also be the best for other body sites, like the pelvis, head, etc., as the behavior of a similarity measure depends not only on the imaging modality but also on image content and image degrading effects. We have also shown that a combination of different similarity measures can lead to a more robust registration. Considering the high accuracy and robustness of CT/CBCT and even MR/CBCT registrations using an appropriate similarity measure, the clinical use of such a registration for patient setup verification is worth of further investigation.



Ideas are like rabbits. You get a couple  
and learn how to handle them, and  
pretty soon you have a dozen.

*John Steinbeck*

## Chapter 8

# Robust gradient-based 3D/2D registration of CT and MR to X-ray images

Primož Markelj, Dejan Tomažević, Franjo Pernuš and Boštjan Likar  
Accepted for publication in IEEE Transactions on Medical Imaging.

### Abstract

*One of the most important technical challenges in image-guided intervention is to obtain a precise transformation between the intra-interventional patient's anatomy and corresponding pre-interventional 3D image on which the intervention was planned. This goal can be achieved by acquiring intra-interventional 2D images and matching them to the pre-interventional 3D image via 3D/2D image registration. A novel 3D/2D registration method is proposed in this paper. The method is based on robustly matching 3D pre-interventional image gradients and coarsely reconstructed 3D gradients from the intra-interventional 2D images. To improve the robustness of finding the correspondences between the two sets of gradients, hypothetical correspondences are searched for along normals to anatomical structures in 3D images, while the final correspondences are established in an iterative process, combining the robust random sample consensus algorithm (RANSAC) and a special gradient matching criterion function. The proposed method was evaluated using the publicly available standardized evaluation methodology for 3D/2D registration, consisting of 3D rotational X-ray, computed tomography, magnetic resonance (MR), and 2D X-ray images of two spine segments, and standardized evaluation criteria. In this way, the proposed method could be objectively compared to the intensity, gradient, and reconstruction-based*

registration methods. The obtained results indicate that the proposed method performs favorably both in terms of registration accuracy and robustness. The method is especially superior when just a few X-ray images and when MR pre-interventional images are used for registration, which are important advantages for many clinical applications.

## 8.1 Introduction

Medical imaging has long been crucial in providing the clinician with information about the internal anatomy of the human body, facilitating medical diagnosis, monitoring of disease or treatment progress, and planning of medical interventions. A successful intervention commonly requires that the clinician, using knowledge accumulated over years of training, in his mind establishes the correspondence between the patient and the information in his pre-operative images. The desire to facilitate the establishment of correspondence and the constant need to make the interventions even more accurate and less invasive gave rise to the discipline of image-guided intervention (IGI). The essence of IGI is to help to diagnose, plan, simulate, guide or otherwise assist an interventionalist or possibly a robot in performing an interventional procedure (surgery, radiation therapy, chemotherapy, biopsy, etc.) by using high quality pre-interventional medical data, usually computed tomography (CT) or magnetic resonance (MR) images in the intervention room. One of the most important technical challenges of an IGI system is to obtain a precise transformation between the patient's anatomy in the interventional coordinate system and any 3D point in the pre-interventional image coordinate system. This can be achieved by registering the pre-interventional data either directly to the patient or indirectly to one or more three-dimensional (3D) or two-dimensional (2D) images of the patient acquired during intervention. The correct transformation allows precise 3D positioning of the interventional instruments in pre-interventional images relative to the final target and nearby vulnerable structures that are to be avoided during the intervention. IGI is therefore most beneficial for surgical applications (neurosurgery (Hipwell *et al.*, 2003; McLaughlin *et al.*, 2005), spinal surgery (Weese *et al.*, 1997), orthopedics (Guéziec *et al.*, 1998; Livyatan *et al.*, 2003; Simon *et al.*, 1995a) and radiotherapy (Adler *et al.*, 1999; Gilhuijs *et al.*, 1996a; Khamene *et al.*, 2006; Murphy, 1997), where the target must be effectively treated (radiated) while minimizing damage to the healthy tissue.

With the widespread use of numerous 3D and 2D imaging modalities, registration of images of different modalities and dimensions has become an important issue of IGI. Commonly, X-ray or ultrasound imaging is used to acquire images during intervention, while CT and/or MR images are typically the pre-interventional modalities of choice, yielding detailed 3D anatomical information that can be efficiently exploited if registered to the intra-interventional images. A vast number of 3D/2D registration methods proposed in the past can be classified according to many criteria, such as the nature of registration basis, the nature and domain of transformation, optimization procedure, modalities involved etc. (Maintz and Viergever, 1998). Most

traditionally the 3D/2D methods can be coarsely classified into feature-based and intensity-based methods. However, to further discriminate and survey the major characteristics of the existing methods, we classify the methods into four categories, namely into the feature-based, intensity-based, gradient-based, and reconstruction-based methods.

Feature-based 3D/2D registration methods (Benameur *et al.*, 2003; Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Hamadeh *et al.*, 1998; Lavallée and Szeliski, 1995; Maurer *et al.*, 1998; Zheng *et al.*, 2007) are concerned with finding the transformation that minimizes the distances between corresponding features extracted from the pre-interventional image or anatomical model and the intra-interventional image. The features can be either extrinsic, requiring attachment of fiducial markers, or intrinsic, obtained by extracting features of anatomical structures. The fiducial markers are either implanted into bone of the patient or skin-affixed. Because implanted fiducials yield very accurate registrations, they are commonly used to define the reference or “gold” standard registration (Maurer *et al.*, 1997; Tomažević *et al.*, 2004a). However, such implantation is invasive and may be inconvenient to the patient. On the other hand, due to skin elasticity, fiducials attached to skin produce less accurate registration results. The less invasive but more sophisticated intrinsic approach uses points, contours, or surfaces extracted from image data sets (Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Hamadeh *et al.*, 1998; Lavallée and Szeliski, 1995) or from statistical models of anatomy (Benameur *et al.*, 2003; Zheng *et al.*, 2007). Some form of outlier removal is commonly used to refine the features used for registration. Extraction of geometrical features greatly reduces the amount of data, which in turn makes the registration fast. However, the accuracy of the procedure is directly dependent upon the accuracy of the segmentation, which by itself is a non-trivial procedure to perform automatically, while manual segmentation is time consuming and subjective.

Intensity-based registration methods use all the image data available by matching image intensities (Gilhuijs *et al.*, 1996a; Hipwell *et al.*, 2003; Khamene *et al.*, 2006; Kim *et al.*, 2007; Lemieux *et al.*, 1994; Murphy, 1997; Penney *et al.*, 2001; Penney *et al.*, 1998; Russakoff *et al.*, 2003a; Weese *et al.*, 1997). To compare intensities of 3D CT and 2D X-ray images in the 2D domain, projection images called digitally reconstructed radiographs (DRRs) are created from the CT image. The matching process is carried out by optimizing the measure of similarity between the DRR(s) and the X-ray image(s). A variety of similarity measures was proposed in the past (Hipwell *et al.*, 2003; Penney *et al.*, 1998; Russakoff *et al.*, 2003a; Weese *et al.*, 1997), among which the optimal similarity measure seems to be application specific, although gradient and correlation based similarity measures yield the most accurate results (Khamene *et al.*, 2006; Kim *et al.*, 2007; Penney *et al.*, 2001; Penney *et al.*, 1998). By using all the available information, the intensity-based methods generally outperform the intrinsic feature-based methods in terms of accuracy but the large amount of data makes intensity-based methods rather slow. This drawback can be reduced by using faster methods

for DRR generation (Birkfellner *et al.*, 2005; Russakoff *et al.*, 2005b; Weese *et al.*, 1999) or by reducing the matching to a region of interest (Hipwell *et al.*, 2003; Khamene *et al.*, 2006; Murphy, 1997; Penney *et al.*, 2001; Penney *et al.*, 1998; Russakoff *et al.*, 2003a; Weese *et al.*, 1997), all of which can be further accelerated using hardware-acceleration methods (Khamene *et al.*, 2006). However, by projecting a high-quality 3D image into 2D, valuable 3D information may be lost. Furthermore, since there is practically no correspondence between MR-based DRRs and X-ray images, the intensity-based methods are generally not suitable for registration of MR to X-ray images.

In gradient-based registration methods a subset of 3D intensity gradients is compared to X-ray gradients (Tomažević *et al.*, 2003) or a subset of 2D X-ray gradients is compared to 3D intensity gradients (Livyatan *et al.*, 2003). Gradient-based methods are accurate and the reduced set of significant gradients makes the gradient-based methods fast. Furthermore, as 3D gradients can also be extracted from MR images, MR to X-ray 3D/2D registration is also possible (Tomažević *et al.*, 2003). However, registration of intensity gradients typically results in poor registration convergence, if the initial misregistration is large (Livyatan *et al.*, 2003; Tomažević *et al.*, 2003).

Most recently, reconstruction-based methods have been proposed, that first reconstruct a 3D image from a few fluoroscopic or X-ray images (Prümmer *et al.*, 2006; Tomažević *et al.*, 2006), and then perform 3D/3D registration. To overcome the problem of poor quality of a 3D image reconstructed from only a few fluoroscopic X-ray images, a robust mutual information based similarity measure, called asymmetric multifeature mutual information, has recently been proposed (Škerl *et al.*, 2006b; Tomažević *et al.*, 2006). The measure uses additional spatial features in the form of intensity gradients. An extensive evaluation has shown that the 3D/2D reconstruction-based method (Tomažević *et al.*, 2006) outperforms the gradient-based method (Tomažević *et al.*, 2003) in terms of capture range and success rate. However, reconstruction-based methods are slower and typically require more intra-interventional images, especially for MR to X-ray registration, which may be a limitation for clinical use.

The aim of this work is to combine the advantages of gradient-based and reconstruction-based methods so as to obtain a robust 3D/2D registration method for both 3D CT and MR pre-interventional images that will require only a small number of 2D intra-interventional images. The proposed method is based on matching 3D gradients of 3D images to 3D gradients coarsely reconstructed from 2D images. To improve the process of finding the correspondences between the two sets of gradients and thereby the robustness of the registration method, hypothetical correspondences are searched for along normals to anatomical structures in 3D images, while the final correspondences are established in an iterative process, combining the robust random sample consensus algorithm (RANSAC) (Fischler and Bolles,

1981) and special gradient matching criterion function. The novel method has been evaluated using the standardized evaluation methodology for 3D/2D registration (van de Kraats *et al.*, 2005a), which enables objective comparison to some of the well-known registration methods (Penney *et al.*, 1998; Tomažević *et al.*, 2006; Tomažević *et al.*, 2003).

## 8.2 Method

The goal of rigid 3D/2D registration is to find a transformation  $\mathbf{T}$  that relates the coordinate system  $\mathbf{S}_{pre}$  of the 3D pre-interventional image to the world or patient coordinate system  $\mathbf{S}_w$  (Figure 8.1). To estimate this transformation by using 2D intra-interventional images, the X-ray acquisition system needs to be calibrated so that the transformation  $\mathbf{T}_c$  between the intra-interventional image coordinate system  $\mathbf{S}_{intra}$  and the world coordinate system  $\mathbf{S}_w$  is established (van de Kraats *et al.*, 2006). We focus on the calculation of the six parameters  $\mathbf{q} = (t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$  that define the rigid transformation  $\mathbf{T}$  and relate the coordinate system  $\mathbf{S}_{pre}$  of the pre-interventional image to the world coordinate system  $\mathbf{S}_w$ . The description of the proposed method is divided into three subsections. The first describes the calculation of gradient fields, the second addresses the gradient correspondence maximization, while the third outlines the gradient-reconstruction based registration.

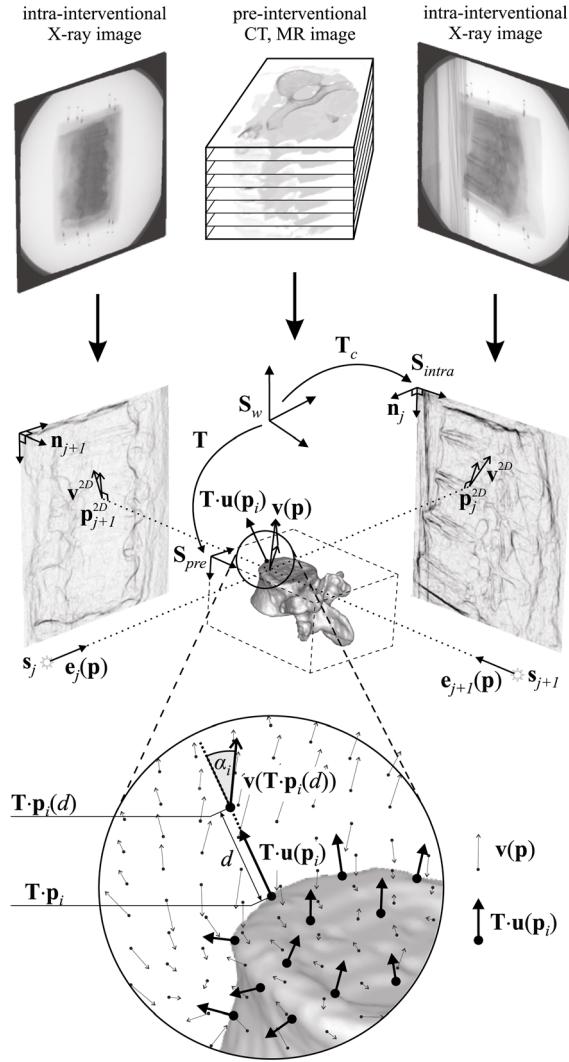
### 8.2.1 Calculation of gradient fields

The proposed method is based on the assumption that strong intensity gradients in 2D X-ray images correspond to distinctive boundaries of anatomical structures in 3D CT or MR images (Tomažević *et al.*, 2003). Using this idea, the 3D pre-interventional image is preprocessed to extract a set of strong intensity gradients  $\mathbf{u}(\mathbf{p}_i)$ ;  $i = 1, 2, \dots, I$ , that generally represent the surfaces and boundaries of anatomical structures and estimate the corresponding surface normals at 3D points  $\mathbf{p}_i$  defined in the coordinate system  $\mathbf{S}_{pre}$  of the pre-interventional image.

To coarsely reconstruct the gradients from 2D intra-interventional images, each 2D intra-interventional image  $j$ ;  $j = 1, 2, \dots, J$ , is preprocessed to obtain its 2D gradient field  $\mathbf{v}^{2D}(\mathbf{p}^{2D})$ . The extracted gradients  $\mathbf{v}^{2D}(\mathbf{p}_j^{2D})$  from all  $J$  X-ray images are then back-projected into 3D (Tomažević *et al.*, 2003) and their information integrated in 3D by simple summation so as to obtain a coarsely reconstructed 3D gradient field  $\mathbf{v}(\mathbf{p})$ :

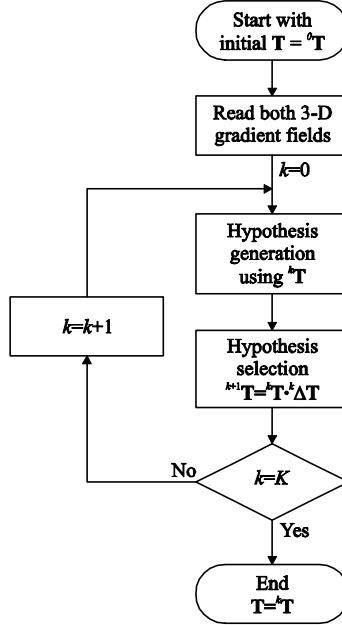
$$\mathbf{v}(\mathbf{p}) = \sum_{j=1}^J \frac{(\mathbf{n}_j \times \mathbf{v}^{2D}(\mathbf{p}_j^{2D})) \times \mathbf{e}_j(\mathbf{p})}{\mathbf{n}_j \cdot \mathbf{e}_j(\mathbf{p})} \cdot \frac{|\mathbf{p}_j^{2D} - \mathbf{s}_j|}{|\mathbf{p} - \mathbf{s}_j|}, \quad (8.1)$$

where  $\mathbf{p}$  is a 3D position,  $\mathbf{s}_j$  the position of the X-ray source of the  $j^{\text{th}}$  2D image and  $\mathbf{p}_j^{2D}$  the corresponding projection of  $\mathbf{p}$  to the  $j^{\text{th}}$  2D image,  $\mathbf{n}_j$  the unit



**Figure 8.1.** 3D/2D registration procedure and corresponding geometrical setup for the registration of gradients  $\mathbf{u}(\mathbf{p}_i)$  representing the boundaries of anatomical structures in a 3D image and the coarsely reconstructed gradient field  $\mathbf{v}(\mathbf{p})$  from the 2D gradient images  $\mathbf{v}^{2D}$ .

normal to the  $j^{\text{th}}$  2D image,  $\mathbf{e}_j(\mathbf{p})$  a unit vector defining the projection of point  $\mathbf{p}$  to the  $j^{\text{th}}$  2D image, and  $\mathbf{v}^{2D}$  the extracted gradient in the  $j^{\text{th}}$  2D image, all defined in the world coordinate system  $\mathbf{S}_w$  (Figure 8.1). The first term in (8.1) represents the component of the gradient  $\mathbf{v}^{2D}(\mathbf{p}_i^{2D})$  that is perpendicular to the projection beam  $\mathbf{e}_j(\mathbf{p})$ , while the second term represents the back-projection and corresponding scaling, which depends on the 3D position of  $\mathbf{p}$  (Tomažević *et al.*, 2003).



**Figure 8.2.** A flowchart of the robust gradient reconstruction-based (RGRB) 3D/2D registration method.

In this way, the 3D/2D registration problem is translated to the problem of 3D/3D registration of a set of gradients  $\mathbf{u}(\mathbf{p}_i)$  obtained from the 3D pre-interventional image and the coarsely reconstructed gradient field  $\mathbf{v}(\mathbf{p})$  from the 2D intra-interventional images.

### 8.2.2 Gradient correspondence maximization

Registration of the gradient sets  $\mathbf{u}(\mathbf{p}_i)$  to  $\mathbf{v}(\mathbf{p})$  is achieved by applying a transformation  $\mathbf{T}$  to gradients  $\mathbf{u}(\mathbf{p}_i)$  and maximizing the correspondences with the gradient field  $\mathbf{v}(\mathbf{p})$ , using a robust gradient-matching criterion function ( $CF$ ), like the one proposed in (Tomažević *et al.*, 2003):

$$CF(\mathbf{T}, \mathbf{u}(\mathbf{p}_i), \mathbf{v}(\mathbf{p})) = \frac{\sum_{i=1}^I |\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)| \cdot |\mathbf{v}(\mathbf{T} \cdot \mathbf{p}_i)| \cdot f(\alpha_i)}{\sum_{i=1}^I |\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)| \cdot \sum_{i=1}^I |\mathbf{v}(\mathbf{T} \cdot \mathbf{p}_i)|}, \quad (8.2)$$

where  $\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  denotes rigid transformation  $\mathbf{T}$  of the gradients  $\mathbf{u}(\mathbf{p}_i)$ , i.e. rigid transformation  $\mathbf{T} \cdot \mathbf{p}_i$  of the points  $\mathbf{p}_i$  and corresponding rotations of gradients  $\mathbf{u}(\mathbf{p}_i)$ , and where  $\mathbf{v}(\mathbf{T} \cdot \mathbf{p}_i)$  denotes the reconstructed gradients that correspond to the transformed gradients  $\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  at positions  $\mathbf{T} \cdot \mathbf{p}_i$ . The angle weighting

function  $f(\alpha_i)$  penalizes the angle difference  $\alpha_i$  between the corresponding gradients  $\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  and  $\mathbf{v}(\mathbf{T} \cdot \mathbf{p}_i)$ :

$$f(\alpha_i) = \begin{cases} \cos^n(\alpha_i - m \cdot 180^\circ) & ; \quad |\alpha_i - m \cdot 180^\circ| < 90^\circ \\ 0 & ; \quad \text{otherwise} \end{cases}. \quad (8.3)$$

The parameter  $n$  determines the sensitivity to angle difference  $\alpha_i$  and the parameter  $m$ ;  $m = 0$  or  $1$ , depends on the gradient directional correspondence between the imaging modalities. For example,  $m = 0$  for CT or 3DRX to X-ray registration in which the corresponding registered intensity gradients are supposed to have the same directions, while  $m = 1$  for MR to X-ray registration in which the corresponding registered intensity gradients have the opposite directions.

### 8.2.3 Gradient reconstruction-based registration

Optimizing the  $CF$  in (8.2) yields a gradient reconstruction-based (GRB) registration method, which is similar to the gradient-based (GB) method (Tomažević *et al.*, 2003), but with the criterion function with the coarsely reconstructed 3D gradients from all the 2D images and not by summing the criterion functions of individual 2D images as in (Tomažević *et al.*, 2003). Similarly to GB, the GRB method should be fast, robust to outliers, and potentially accurate. However, because directly relating image gradients usually results in poor registration convergence, the GRB method should be useful for fine registration of coarsely pre-registered images. To increase the convergence, the method is further sophisticated into the so-called robust gradient reconstruction-based (RGRB) method. The RGRB method combines the similarity measure (8.2) of the GRB method and a robust iterative registration scheme, interleaving the generation and selection of hypothetical gradient correspondences. In the hypothesis generation step, hypothetical point-to-point geometrical correspondences between each gradient  $\mathbf{u}(\mathbf{p}_i)$  and the corresponding gradient field  $\mathbf{v}(\mathbf{p})$  is established along a line defined by gradient  $\mathbf{u}(\mathbf{p}_i)$ . In the hypothesis selection step, the three point-to-point correspondences, defining the 3D rigid transformation  $\mathbf{T}$ , that maximize the similarity (8.2) between all gradients  $\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  and the coarsely reconstructed gradient field  $\mathbf{v}(\mathbf{p})$  are robustly selected among all pre-generated hypothetical point-to-point correspondences. The flowchart of the RGRB method is shown in Figure 8.2.

First, in the hypothesis generation step of the  $k^{th}$  iteration;  $k = 1, 2, \dots, K$ , for the current registration transformation  ${}^k\mathbf{T}$ , hypothetical point-to-point geometrical correspondences  $(\mathbf{p}_i, \mathbf{p}_i(d))$  between each gradient  $\mathbf{u}(\mathbf{p}_i)$  and the coarsely reconstructed gradient field  $\mathbf{v}(\mathbf{p})$  are established by finding the points  $\mathbf{p}_i(d)$ :

$$\mathbf{p}_i(d) = \mathbf{p}_i + \mathbf{u}_e(\mathbf{p}_i) \cdot d, \quad (8.4)$$

at a distance  $d$  along the direction  $\mathbf{u}_e(\mathbf{p}_i)$  of the gradient  $\mathbf{u}(\mathbf{p}_i)$ ;  $\mathbf{u}_e(\mathbf{p}_i) = \mathbf{u}(\mathbf{p}_i) / |\mathbf{u}(\mathbf{p}_i)|$ , using fixed sized increments (Markelj *et al.*, 2007). The optimal point  $\mathbf{p}_i(d^*)$  maximizes the point-to-point gradient similarity between the gradient  ${}^k\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  and corresponding gradient  $\mathbf{v}({}^k\mathbf{T} \cdot \mathbf{p}_i(d))$ :

$$\mathbf{p}_i(d^*) = \arg \max_{\mathbf{p}_i(d)} \left( |{}^k\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)| \cdot |\mathbf{v}({}^k\mathbf{T} \cdot \mathbf{p}_i(d))| \cdot f(\alpha_i) \cdot {}^k g(d) \right), \quad (8.5)$$

where  ${}^k g(d)$  is a monotonously decreasing weighting function that penalizes larger distances  $d$  from  $\mathbf{p}_i$ . The gradient  $\mathbf{v}({}^k\mathbf{T} \cdot \mathbf{p}_i(d))$  is obtained by interpolation of the gradient field  $\mathbf{v}(\mathbf{p})$ . For a current registration transformation  ${}^k\mathbf{T}$  and for each gradient  ${}^k\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$ , the obtained hypothetical point-to-point geometrical correspondence  $(\mathbf{p}_i, \mathbf{p}_i(d^*))$ ,  $i = 1, 2, \dots, I$ , thus represents the strongest and most colinear reconstructed gradient  $\mathbf{v}({}^k\mathbf{T} \cdot \mathbf{p}_i(d^*))$  along the line defined by  $\mathbf{u}(\mathbf{p}_i)$ .

Second, in the hypothesis selection step, the current registration transformation  ${}^k\mathbf{T}$  is refined:

$${}^{k+1}\mathbf{T} = {}^k\mathbf{T} \cdot {}^k\Delta\mathbf{T}, \quad (8.6)$$

by an incremental transformation  ${}^k\Delta\mathbf{T}$  defined by the pre-selected hypothetical point-to-point geometrical correspondences  $(\mathbf{p}_i, \mathbf{p}_i(d^*))$ . This is carried out in a robust manner by adapting a random sample consensus algorithm (RANSAC) (Fischler and Bolles, 1981). First, three, a minimal number for 3D rigid transformation, point-to-point correspondences  $(\mathbf{p}_i, \mathbf{p}_i(d^*))$  are repeatedly randomly selected  $L$  times;  $l = 1, 2, \dots, L$ , from all  $I$  hypothetical correspondences, least-squares fitted to obtain the hypothetical incremental transformation  ${}^{k,l}\Delta\mathbf{T}$  and hypothetical transformation  ${}^{k+1,l}\mathbf{T} = {}^k\mathbf{T} \cdot {}^{k,l}\Delta\mathbf{T}$ . Next, all gradients  $\mathbf{u}(\mathbf{p}_i)$  are transformed by  ${}^{k+1,l}\mathbf{T}$ , yielding  ${}^{k+1,l}\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$ . Finally, the transformation that maximizes the sample consensus, i.e. the robust gradient-matching function (8.2) between all transformed gradients  ${}^{k+1,l}\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  and the coarsely reconstructed gradient field  $\mathbf{v}(\mathbf{p})$ , is selected:

$${}^{k+1}\mathbf{T} = \arg \max_{{}^{k+1,l}\mathbf{T}} CF({}^{k+1,l}\mathbf{T}, \mathbf{u}(\mathbf{p}_i), \mathbf{v}(\mathbf{p})). \quad (8.7)$$

The registration transformation  ${}^k\mathbf{T}$  is iteratively;  $k = 1, 2, \dots, K$ , refined by interleaving the hypothesis generation and the hypothesis selection step and changing (annealing) the distance weighting function  ${}^k g(d)$  in (8.5) by the square root of the iteration index  $k$ :

$${}^k g(d) = g(d \cdot \sqrt{k}). \quad (8.8)$$

Considering the Gaussian weighted function  $g(d)$  and (8.8), results in the distance weighting function:

$$^k g(d) = e^{-\frac{1}{2} \left( \frac{d}{\sigma_d / \sqrt{k}} \right)^2}, \quad (8.9)$$

where  $\sigma_d$  is the distance weighting parameter, defining the point-to-point correspondence-finding capture range  $p_{CR}$ ;  $p_{CR} = \sigma^d / k - 1/2$ , that decreases by the square root of the iteration index  $k$ . Large initial  $p_{CR}$  ( $k=1, 2, 3, \dots$ ) enables the selection of hypothetical point-to-point correspondences at larger distances  $d$  and thereby increases the registration capture range. On the other hand, a smaller  $p_{CR}$  in the final iterations ( $k \rightarrow K$ ) generates more hypothetical point-to-point correspondences at smaller distances and thereby increases the registration accuracy.

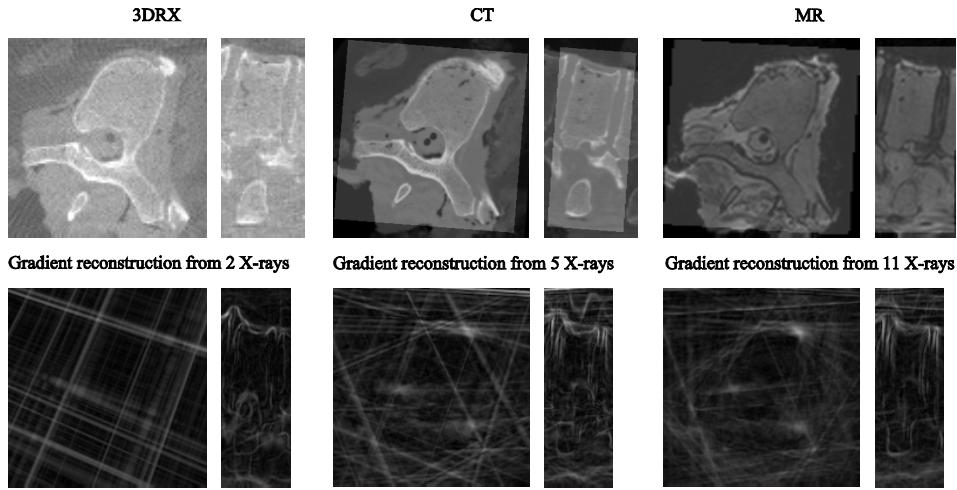
Since the RGB method should be fast and accurate, while the RGRB method was designed to increase the robustness of the registration, we also introduce an extension of the RGRB by the GRB method, yielding the so-called RGRB extended method (RGRB<sub>e</sub>) as it might prove the most optimal solution in terms of speed, accuracy and robustness of registration. For implementation details the reader is referred to Section 8.3.3.

### 8.3 Registration evaluation

Evaluating the performance of a 3D/2D registration method requires a reference data set with known reference registration (Jannin *et al.*, 2006) also referred to as “gold” standard registration (Tomažević *et al.*, 2004a; van de Kraats *et al.*, 2005a). A variety of validation approaches were used in the past, but to the best of our knowledge, currently there are only two reference data sets for 3D/2D registration publicly available (Tomažević *et al.*, 2004a; van de Kraats *et al.*, 2005a). The publicly available<sup>6</sup> standardized evaluation methodology (van de Kraats *et al.*, 2005a), consists of image data sets, reference transformations, regions of interest with corresponding centers of rotations, and measures for registration error, capture range, and success rate. Using this methodology in the current study, we were able to objectively compare the proposed method to three well-known 3D/2D registration methods. Namely, the intensity-based method (IB) (Penney *et al.*, 1998) that compares 2D digitally reconstructed radiographs (DRRs) and X-ray images using the gradient difference criterion function, the gradient-based method (GB) (Tomažević *et al.*, 2003) that matches 3D normals to bony surfaces and 2D X-ray gradients back-projected into 3D, and the reconstruction-based method (RB) (Tomažević *et al.*, 2006) that integrates 2D X-ray image information in 3D and then performs 3D/3D image registration by a special multifeature mutual information similarity measure.

---

<sup>6</sup> <http://www.isi.uu.nl/Research/Databases/GS/>



**Figure 8.3.** Axial and lateral cross-sections taken from 3DRX, CT, and MR VOIs (top row from left to right), and coarsely reconstructed 3-D gradients from 2, 5, and 11 X-ray images (bottom row from left to right).

**Table 8.1.** The sizes and resolutions of the X-ray, 3DRX, CT, and MR images.

Modality	Segment	Resolution	Size
X-ray	1	$0.63 \times 0.63 \text{ mm}^2$	$512^2$
	2	$0.53 \times 0.53 \text{ mm}^2$	$512^2$
3DRX	1	$0.87 \times 0.87 \times 0.87 \text{ mm}^3$	$256^3$
	2	$0.52 \times 0.52 \times 0.52 \text{ mm}^3$	$256^3$
CT	1	$0.31 \times 0.49 \times 0.31 \text{ mm}^3$	$320 \times 260 \times 320$
	2	$0.31 \times 0.49 \times 0.31 \text{ mm}^3$	$280 \times 300 \times 300$
MR	1	$1.00 \times 0.75 \times 0.75 \text{ mm}^3$	$100 \times 256 \times 256$
	2	$1.00 \times 0.88 \times 0.88 \text{ mm}^3$	$120 \times 256 \times 256$

### 8.3.1 Experimental data

The experimental image data consisted of 2D X-ray images, CT and MR images, and 3DRX images of two defrosted segments of a spinal column with some soft tissue around both segments. One spinal segment consisted of three thoracolumbar vertebral bodies while the other of five thoracic vertebral bodies. The X-ray images were obtained by a clinical 3DRX system (Integrus BV5000, Philips Medical Systems). A set of 100 X-ray images was acquired for

each spinal segment in 8 s runs of 180° rotation around the imaged object. The images of the first and the second segment were acquired with image intensifier size of 380 and 310 mm, respectively. The 3DRX images were reconstructed from 100 X-ray projection images. The CT images were acquired with a clinical multi-slice CT scanner (MX8000, IDT 16, Philips Medical Systems). The MR images were obtained with a 1.5-T MR scanner (Gyrosan NT, Philips Medical Systems) using a sagittal 3D turbo spin echo acquisition (turbo factor 29, TR = 1500 ms, TE = 90 ms). Examples of the acquired 3DRX, CT, and MR images are shown in Figure 8.3 (top row), while the sizes and resolutions of all image data are given in Table 8.1.

The reference registration between 3DRX images and 2D projection images was already established in the process of creating 3DRX images, while the reference registrations of CT and MR images to 2D X-ray images were obtained by a 3D/3D rigid registration of CT and MR images to the corresponding 3DRX images using the mutual information based registration method (Maes *et al.*, 1997). The error of the reference registration was evaluated to be similar to the error of 3D/3D registration, which was considered to be subvoxel and therefore at least as accurate as marker-based registration (van de Kraats *et al.*, 2004).

### 8.3.2 Evaluation methodology

The mean target registration error (mTRE) was used to measure the distance of a vertebra from the reference position before and after registration. For the calculation of mTRE, positions of all voxels in a volume of interest (VOI), containing a whole vertebra, were used as target points (van de Kraats *et al.*, 2005a):

$$\text{mTRE} = \frac{1}{M} \sum_{m=1}^M \left\| \mathbf{T}_{\text{reg}} \mathbf{p}_m - \mathbf{T}_{\text{ref}} \mathbf{p}_m \right\|, \quad (8.10)$$

where  $M$  is the number of target points  $\mathbf{p}_m$ ,  $\mathbf{T}_{\text{reg}}$  the registration determined by the registration algorithm and  $\mathbf{T}_{\text{ref}}$  the reference or “gold” standard registration. To assess the accuracy, success rate, and capture range, registrations from 200 starting positions defined by randomly generated translations and rotations were performed for each VOI. The translations and rotations were chosen to yield mTRE values of the starting positions uniformly distributed in an interval from 0 to 20 mm, with 10 starting positions in each of the 1 mm wide subintervals, as prepared by van de Kraats *et al* (van de Kraats *et al.*, 2005a). Each registration was classified as successful if the mTRE after registration was smaller than a pre-specified threshold of 2 mm. The success rate was defined as the number of successful registrations against the number of all registrations. The registration accuracy was computed as the average mTRE value of all successful registrations, while the capture range was defined as the distance from the reference position to the

first 1 mm subinterval for which the registration was successful in less than 95% of all cases (van de Kraats *et al.*, 2005a).

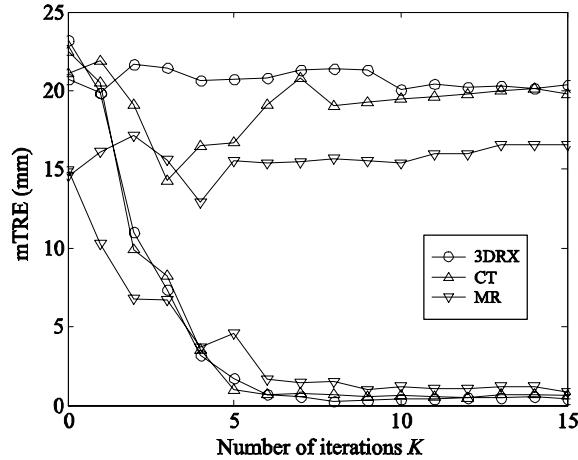
### 8.3.3 Implementation details

The 3D images were preprocessed using the 3D Gaussian filter with the standard deviation of 0.3 mm for the CT and MR images, and of 0.5 mm for the 3DRX images. Next, the images were isotropically resampled to the resolution of 1 mm and subjected to the 3D Canny edge detector. Finally, the resulting gradient images were thresholded as in (Tomažević *et al.*, 2006) and (van de Kraats *et al.*, 2005a) so as to extract the boundaries of anatomical structures and the corresponding set of gradients  $\mathbf{u}(\mathbf{p})$ . Before preprocessing and gradient extraction, the MR images were corrected for intensity inhomogeneity by the information minimization method (Likar *et al.*, 2001). Spurious strong gradients corresponding to the spine phantom-to-background transition, which are not present in real MR spine images, were suppressed by simple morphological operations.

The X-ray images were blurred by the Gaussian filter with the standard deviation of 0.5 mm and subjected to the Roberts edge detector so as to calculate the intensity gradients  $\mathbf{v}^{2D}(\mathbf{p}^{2D})$ . The 3D gradient fields  $\mathbf{v}(\mathbf{p})$  were coarsely reconstructed in volumes of  $137 \times 150 \times 145$  and  $140 \times 200 \times 134$  image elements for the first and second spinal segment, respectively, with the isotropic resolution of 0.63 mm. To calculate the gradient-matching criterion function (8.2), the gradient  $\mathbf{v}(\mathbf{T} \cdot \mathbf{p}_i)$  corresponding to each gradient  $\mathbf{T} \cdot \mathbf{u}(\mathbf{p}_i)$  was determined using trilinear interpolation of the gradient field  $\mathbf{v}(\mathbf{p})$ . The value of the parameter  $n$  in the angle weighting function (8.3) was set to 4 as in (Tomažević *et al.*, 2003).

In the hypothesis generation step, the distance weighting parameter  $\sigma_d$  was set to 10 mm. The point-to-point geometrical correspondences  $(\mathbf{p}_i, \mathbf{p}_i(d^*))$  were established in discrete steps of  $p_{CR}/10$  mm;  $p_{CR} = \sigma_d/k - 1/2$ , along directions  $\mathbf{u}_e(\mathbf{p}_i)$  in the range of  $\pm 2 \cdot p_{CR}$  mm, requiring a total of 41 evaluations of point-to-point gradient similarities (8.5) for each point  $\mathbf{p}_i$ .

In the hypothesis selection step, the number  $L$  of random selections of three point-to-point geometrical correspondences  $(\mathbf{p}_i, \mathbf{p}_i(d^*))$  has to be selected according to the rate of expected outliers, i.e. the rate of expected non-corresponding boundaries of anatomical structures that are due to different imaging modalities and corresponding artifacts in the two images undergoing registration. The parameter  $L$  was determined (Fischler and Bolles, 1981) for the 70% rate of expected outliers ( $w = 1 - 0.7 = 0.3$ ) and with 95% probability ( $z = 0.95$ ) of randomly selecting at least one set of three ( $c = 3$ ) outlier-free point-to-point geometrical correspondences  $(\mathbf{p}_i, \mathbf{p}_i(d^*))$ :



**Figure 8.4.** Typical convergence curves of the RGRB method for the registration of 3DRX, CT and MR images to two X-ray images showing the mTRE value as a function of iterations  $k$ . For each 3D modality, one successful registration (three bottom curves) and one unsuccessful registration (three top curves) are given.

$$L = \frac{\log(1 - z)}{\log(1 - w^c)} = \frac{\log(1 - 0.95)}{\log(1 - 0.3^3)} \approx 100 . \quad (8.11)$$

The value of the number of iterations  $K$  was determined experimentally by observing typical convergence curves, i.e. by observing the mTRE as a function of iteration index  $k$  (Figure 8.4). With respect to good convergence properties of the RGRB method, the value of parameter  $K$  was set to 10, although such a small number of iterations may yield sub-optimal registration accuracy. To increase the registration accuracy, a considerably larger number of iterations would be required. However, a computationally more attractive solution is to use the RGRBe method that performs a coarse registration by the RGRB method with a small  $K$ , the result of which is then refined by the GRB registration method that optimizes the gradient-matching criterion function (2) by the Powell's method (Press *et al.*, 1992).

### 8.3.4 Experiments

The performances of all three variations of the gradient reconstruction-based registration method (GRB, RGRB, and RGRBe method) were assessed by the standardized evaluation methodology (van de Kraats *et al.*, 2005a) and compared to the IB (Penney *et al.*, 1998), GB (Tomažević *et al.*, 2003), and RB (Tomažević *et al.*, 2006) methods. The VOIs, containing a single vertebra, from 3DRX, CT, and MR images were registered to sets of 2, 3, 5, and 7 X-ray images, while the MR images were also registered to sets of 9 and 11 X-ray

**Table 8.2.** The average processing times for the GRB, RGRB ( $K = 10$ ), and RGRBe registration methods, on an Intel Core 2 Duo, 2.13 GHz computer running Windows XP.

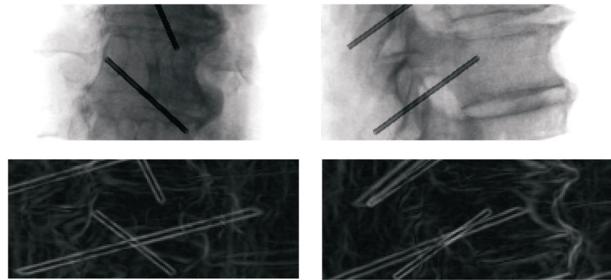
Modality	Time to process (s)		
	GRB	RGRB	RGRBe
3DRX	9	18	27
CT	12	30	42
MR	15	38	53

images. The same as the authors of the standardized evaluation methodology (van de Kraats *et al.*, 2005a), we have selected the anterior-posterior (AP) and the lateral (LAT) X-ray image from the 100 acquired X-ray images for the registration of two X-ray images to 3DRX, CT and MR images. When more than two X-ray images were used for registration, the first X-ray image was always the first image acquired with the 3DRX system, while the other images were chosen so as to yield an angle between the X-ray views of approximately 60°, 35°, 24°, 18°, and 15° for the 3, 5, 7, 9, and 11 X-ray images, respectively. For each of 3DRX, CT, and MR modalities and each set of X-ray images 1600 registrations were performed by each registration method, 200 per each of the eight VOIs, resulting in a total of more than 60.000 registrations.

Three additional experiments were performed. First, the performance of the RGRB method was assessed for more iterations in an experiment using registration of CT VOIs to two X-ray images with 30 iterations ( $K = 30$ ).

The second additional experiment was performed to compare the proposed approach for estimating the 3D gradients from 2D projection images to an alternative approach by reconstructing a 3D image from two X-ray images using the algebraic reconstruction technique (ART) followed by gradient calculation (Tomažević *et al.*, 2006). The performances of the GRB, RGRB, and RGRBe methods using the alternative gradient estimation approach were also assessed by the standardized evaluation methodology (van de Kraats *et al.*, 2005a).

The final additional experiment was performed to assess the algorithm's robustness to outliers, i.e. to additional structures that may appear in the projection images due to medical tools in clinical settings. Since to the best of our knowledge such data sets with reference registrations are not publicly available, we simulated a tube in 3D for each of the VOIs and projected it to the X-ray images. This was performed by a simplified model of the X-ray image generation:



**Figure 8.5.** Anterior-posterior (top left) and lateral (top right) X-ray images of a single vertebra with a simulated tube imitating medical tools and corresponding anterior-posterior (bottom left) and lateral (bottom right) cross-sections of the coarsely reconstructed 3D gradient images.

$$I(\mathbf{p}^{2D}) = I_0 e^{-\int \mu(\mathbf{p}) d\mathbf{p}}, \quad (8.12)$$

where  $I(\mathbf{p}^{2D})$  is the intensity of the X-ray image at point  $\mathbf{p}^{2D}$ ,  $I_0$  is the reference intensity,  $\mathbf{p}$  is a 3D point and  $\mu$  the attenuation coefficient of the anatomy (Tomažević *et al.*, 2003). The intensities of the X-ray image with simulated medical tool  $I_s(\mathbf{p}^{2D})$  were obtained as:

$$\begin{aligned} I_s(\mathbf{p}^{2D}) &= I_0 e^{-\int \mu(\mathbf{p}) d\mathbf{p} - \int \mu_t(\mathbf{p}) d\mathbf{p}} \\ &= I(\mathbf{p}^{2D}) e^{-\int \mu_t(\mathbf{p}) d\mathbf{p}}, \end{aligned} \quad (8.13)$$

where  $I_s(\mathbf{p}^{2D})$  is the simulated intensity at point  $\mathbf{p}^{2D}$  and  $\mu_t$  the attenuation coefficient of the simulated medical tool. Figure 8.5 shows the AP and LAT X-ray images and corresponding cross-sections of the coarsely reconstructed 3D gradient images. Registration experiments using clinically relevant CT and MR modalities and 2, 3, and 5 X-ray images were performed.

### 8.3.5 Results

Coarse 3D reconstructions of 2D gradients  $\mathbf{v}^{2D}(\mathbf{p}^{2D})$  extracted from 2, 5, and 11 X-ray images are shown in Figure 8.3. Reconstruction artifacts are notable in the axial cross-sections (left images). In the lateral cross-sections (right images) boundaries of anatomical structures are better depicted by the reconstructed 3D gradients even when gradients are reconstructed from only two X-ray images.

According to the standardized evaluation methodology (van de Kraats *et al.*, 2005a), the registration results are presented in Table 8.3 by the mTREs, capture ranges, and success rates. Besides the registration results obtained for the three proposed variations of the gradient reconstruction-based method,

**Table 8.3.** Mean target registration errors (TREs), capture ranges, and success rates for different modalities, number of X-ray images, and for the IB, GB, and RB methods and the three variations of the gradient reconstruction-based methods (GRB, RGRB, RGRBe). The numbers in bold are the results of experiments performed in this study, the numbers in plain text were taken from (Tomažević *et al.*, 2006), while the numbers in italic were acquired from (van de Kraats *et al.*, 2005a).

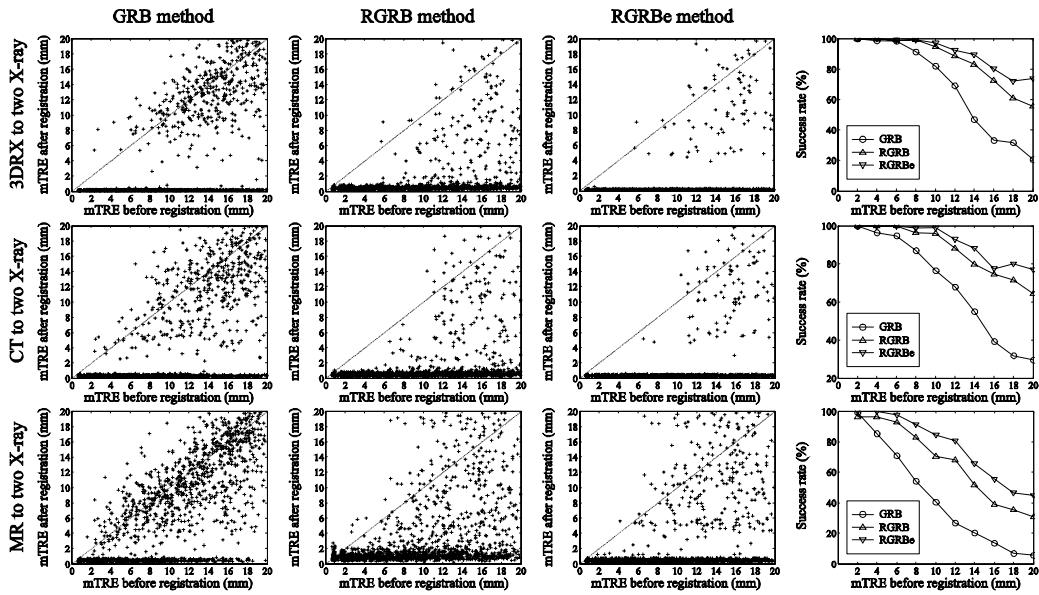
Modality	X-rays	mTRE (mm)						Capture range (mm)						Success rate (%)							
		IB	GB	RB	GRB	RGRB	RGRBe	IB	GB	RB	GRB	RGRB	RGRBe	IB	GB	RB	GRB	RGRB	RGRBe		
3DRX	2	0.13	0.31	0.52	<b>0.19</b>	<b>0.54</b>	<b>0.17</b>	4	0	4	<b>6</b>	<b>9</b>	<b>11</b>	43	76	<b>69</b>	<b>86</b>	<b>91</b>			
	3				0.19	0.33	<b>0.26</b>	<b>0.47</b>	<b>0.24</b>		7	9	<b>7</b>	<b>12</b>	<b>14</b>	68	89	<b>75</b>	<b>91</b>	<b>95</b>	
	5				0.12	0.17	<b>0.18</b>	<b>0.37</b>	<b>0.17</b>		8	12	<b>9</b>	<b>12</b>	<b>13</b>	76	91	<b>84</b>	<b>94</b>	<b>96</b>	
	7				0.11	0.12	<b>0.16</b>	<b>0.33</b>	<b>0.16</b>		8	12	<b>9</b>	<b>14</b>	<b>14</b>	83	92	<b>86</b>	<b>95</b>	<b>97</b>	
CT	2	0.65	0.38	0.43	<b>0.32</b>	<b>0.58</b>	<b>0.32</b>	3	6	5	<b>5</b>	<b>9</b>	<b>11</b>	56	65	<b>69</b>	<b>88</b>	<b>92</b>			
	3				0.32	0.37	<b>0.33</b>	<b>0.55</b>	<b>0.32</b>		7	7	<b>6</b>	<b>8</b>	<b>13</b>	63	78	<b>74</b>	<b>89</b>	<b>94</b>	
	5				0.27	0.27	<b>0.26</b>	<b>0.43</b>	<b>0.26</b>		8	10	<b>10</b>	<b>12</b>	<b>13</b>	72	87	<b>84</b>	<b>94</b>	<b>96</b>	
	7				0.27	0.26	<b>0.26</b>	<b>0.38</b>	<b>0.26</b>		9	12	<b>10</b>	<b>14</b>	<b>14</b>	78	91	<b>87</b>	<b>96</b>	<b>97</b>	
MR	2				0.45	<b>0.50</b>	<b>1.06</b>	<b>0.48</b>		1		2	4	6			<b>44</b>	<b>68</b>	<b>69</b>		
	3					<b>0.59</b>	<b>1.04</b>	<b>0.55</b>				<b>2</b>	<b>3</b>	<b>6</b>			<b>39</b>	<b>63</b>	<b>69</b>		
	5					<b>0.44</b>	<b>0.94</b>	<b>0.42</b>				<b>2</b>	<b>5</b>	<b>6</b>			<b>46</b>	<b>70</b>	<b>75</b>		
	7					<b>0.40</b>	<b>0.83</b>	<b>0.39</b>				<b>3</b>	<b>5</b>	<b>6</b>			<b>51</b>	<b>73</b>	<b>77</b>		
	9					0.54	0.98	<b>0.42</b>	<b>0.79</b>	<b>0.37</b>		2	0	<b>2</b>	<b>8</b>	<b>9</b>	24	69	<b>54</b>	<b>74</b>	<b>78</b>
	11					0.50	0.67	<b>0.38</b>	<b>0.74</b>	<b>0.37</b>		2	7	<b>3</b>	<b>7</b>	<b>8</b>	23	84	<b>56</b>	<b>74</b>	<b>78</b>

**Table 8.4.** Mean target registration errors (TREs), capture ranges, and success rates for GRB, RGRB, and RGRBe methods using the ART reconstruction from two X-ray images (Tomažević *et al.*, 2006).

Modality	mTRE (mm)			Capture range (mm)			Success rate (%)		
	GRB	RGRB	RGRBe	GRB	RGRB	RGRBe	GRB	RGRB	RGRBe
3DRX	0.21	0.51	0.20	5	8	10	70	88	92
CT	0.28	0.53	0.27	5	9	9	68	87	90
MR	0.67	1.10	0.65	2	0	5	44	62	74

GRB, RGRB, and RGRBe, the results of IB, GB, and RB methods acquired from the corresponding publications (Tomažević *et al.*, 2006; van de Kraats *et al.*, 2005a) are also given. The RGRBe method yielded the best accuracies, capture ranges, and success rates for all three pre-interventional modalities and especially when only two X-ray images were used for registration. The RGRBe method inherits the robustness of the RGRB method and the accuracy of the GRB method, which is demonstrated by scatter diagrams and convergence curves in Figure 8.6.

The processing times for the GRB, RGRB ( $K = 10$ ), and RGRBe registration methods, are presented in Table 8.2. Processing times are independent of the number of X-ray images because the reported data do not include the times needed for coarsely reconstructing the gradient fields  $\mathbf{v}(\mathbf{p})$ . These times ranged from 3 to 14 s for the coarse reconstruction from 2 to 11 X-ray images,



**Figure 8.6.** The results for registration of 3DRX, CT, and MR VOIs to two X-ray images (rows, respectively) for the GRB, RGRB, and RGRBe methods (first three columns, respectively) in the form of scatter diagrams, indicating the mTRE value before and after registration. The corresponding success rates are illustrated as a function of initial mTRE (rightmost column).

**Table 8.5.** Mean target registration errors (TREs), capture ranges, and success rates for the GRB and RGRBe methods using CT and MR, and 2, 3, and 5 X-ray images with simulated medical tools.

Modality	X-rays	mTRE (mm)		Capture range (mm)		Success rate (%)	
		GRB	RGRBe	GRB	RGRBe	GRB	RGRBe
CT	2	0.34	0.34	5	10	64	88
	3	0.35	0.34	6	11	67	91
	5	0.27	0.27	8	12	63	94
MR	2	0.67	0.60	2	5	35	67
	3	0.61	0.60	2	5	37	61
	5	0.47	0.45	2	4	40	65

respectively. All processing times were obtained on an Intel Core 2 Duo, 2.13 GHz computer running Windows XP.

The additional experiment using 30 iterations  $K$  of the RGRB method for the CT to two X-ray registration yielded an almost identical capture range and success rate with respect to the RGRBe method with just 10 iterations (Table 8.3, Figure 8.6). However, the registration accuracy was smaller and

the registration time was more than doubled, which justifies the proposed extension of the RGRB method by the optimization based GRB method, yielding better performance and higher speed.

The results of the registration methods using the gradients extracted from the reconstructed image, instead of coarsely reconstructed back-projected gradients, are presented in Table 8.4. By relating these results to the results in Table 8.3, no major differences can be identified. The CT to X-ray registration with ART reconstruction was slightly more accurate but less robust, while a slight reduction of accuracy can be observed for the 3DRX and MR to X-ray registrations. The performances of the two approaches for estimating 3D gradients from 2D X-ray images are therefore similar but the proposed coarse 3D gradient reconstruction is approximately 20 times faster.

The results of registration experiments using the X-ray images with the simulated medical tools are presented in Table 8.5. In comparison to the results in Table 8.3, a slight degradation of all registration performance criteria can be observed. For the CT to X-ray registration, the differences are very small, with the exception of the success rate of the GRB method. These results indicate that in terms of success rate the GRB method could not cope well with the induced outliers in the X-ray images, while in combination with the RGRB method the success rate is not seriously impaired. MR to X-ray registration, on the other hand, was slightly more affected by the simulated outliers. While the degradation of registration performance was minor when only two X-rays were used for registration, using more X-ray images did not yield better registration results in terms of capture range and success rate. This is most likely due to the fact that when more X-ray images are used, the tools are better reconstructed and thereby represent strong outliers, which adversely affect the registration methods. This phenomenon was expectedly much less prominent for the CT images as these typically exhibit stronger bone edges.

## 8.4 Discussion

There are four important issues that need to be addressed before a 3D/2D registration method can be considered clinically acceptable. Namely, the registration accuracy required for a specific clinical application, the robustness of the alignment in terms of success rate and capture range, the speed with which a method can be performed, and the clinical feasibility in terms of image acquisition, user interaction, and interventional protocol requirements and/or acceptance. The proposed methods were designed with respect to these four fundamental requirements, combining the advantages of the methods proposed in the past and proposing novel solutions for overcoming their drawbacks. The performances of the proposed solutions were objectively evaluated and compared to IB (Penney *et al.*, 1998), GB (Tomažević *et al.*, 2003), and RB (Tomažević *et al.*, 2006) methods via the standardized evaluation methodology (van de Kraats *et al.*, 2005a).

Image intensity gradients proved beneficial for increasing the registration accuracy not only of the gradient-based (Livyatan *et al.*, 2003; Tomažević *et al.*, 2003) but also of the intensity and reconstruction-based methods (Hipwell *et al.*, 2003; Khamene *et al.*, 2006; Murphy, 1997; Penney *et al.*, 2001; Penney *et al.*, 1998; Russakoff *et al.*, 2003a; Tomažević *et al.*, 2006). However, the local nature of intensity gradients greatly reduces the robustness of the gradient-based registration. While this is not the case for the reconstruction-based method (Tomažević *et al.*, 2006), in which gradients are used as additional features, these methods typically require more intra-interventional X-ray images, which may be a serious clinical limitation both in terms of the image availability and of high computational demands required for intra-interventional reconstruction. The computational demands are also high for the intensity DRR-based methods (Gilhuijs *et al.*, 1996a; Hipwell *et al.*, 2003; Khamene *et al.*, 2006; Lemieux *et al.*, 1994; Murphy, 1997; Penney *et al.*, 2001; Penney *et al.*, 1998; Russakoff *et al.*, 2003a; Weese *et al.*, 1997), although several solutions for acceleration were proposed (Birkfellner *et al.*, 2005; Russakoff *et al.*, 2005b; Weese *et al.*, 1999). Another drawback of the DRR-based methods is the projection of 3D image data into 2D, because of which some valuable 3D information is inherently lost. Feature-based methods seem appealing, using point distribution or statistical anatomical models (Benameur *et al.*, 2003; Zheng *et al.*, 2007), but the demanding model construction makes them less so. Furthermore, some feature-based methods require segmentation of pre-interventional images (Feldmar *et al.*, 1997a; Guéziec *et al.*, 1998; Hamadeh *et al.*, 1998; Lavallée and Szeliski, 1995), which is a very demanding task that usually requires user interaction.

To fulfill the abovementioned four clinical requirements for a 3D/2D registration method, a new gradient-based method has been proposed that coarsely reconstructs the 2D gradients in 3D. Calculated once before registration, the coarse gradient reconstruction is straightforward, computationally undemanding, and produces results comparable to those achieved by gradient extraction from reconstructed images (Table 8.4). The registration results (Table 8.3) showed that the proposed optimization based GRB method outperforms the IB and GB methods in terms of success rate and capture range, which is likely due to the coarse gradient reconstruction that integrates gradient information from all the available X-ray images in 3D. However, since medical interventions require high robustness, e.g. for the purpose of surgical guidance a success rate of 95% is typically required (Livyatan *et al.*, 2003), the robust iterative RGRB registration method was proposed. In this way, the success rate and capture range were improved significantly (Table 8.3), but because only a small number of iterations was used to keep the registration time short, the registration accuracy was not optimal. The obvious solution at hand was to extend the coarse iterative RGRB method by the fine optimization based GRB method into the coarse-to-fine RGRBe registration method. As a result, the obtained RGRBe method inherits the accuracy of the GRB method and even further increases the high robustness of the RGRB method (Figure 8.6).

Finally, we compare the proposed RGRBe method with the IB, GB, and RB methods with respect to accuracy, robustness, speed, and clinical feasibility. Because it is desirable that the number of intra-interventional images is kept low, we focus on registration results when only two X-ray images were used. First, the registration accuracy of the RGRBe method was on par with the IB, GB and RB methods, except for the 3DRX to two X-ray image registration for which the IB method was slightly more accurate. However, significant improvements in comparison with the GB and RB methods are notable when just two X-ray images were used. Furthermore, the RGRBe method was superior in MR to X-ray image registrations, yielding better accuracy with two X-ray images than the GB and RB methods with nine X-ray images.

Second, considering the robustness, i.e. the capture range and success rate, the RGRBe method has performed far the best. This was especially true for registering CT and 3DRX to two X-ray images for which the capture ranges and success rates were approximately 100% better than those of the IB, GB, and RB methods, not to mention MR to X-ray image registrations. However, when more X-ray images had been used for registration the improvements of the capture ranges and success rates were less prominent.

Third, the registration speed is a serious drawback of the IB (9 min on a multi-user Linux Dell PowerEdge 1600, dual Xeon 2.8 GHz with 4 GB memory) and RB methods for the reasons stated above, while the GB method is rather fast (25s on a Windows 2000 Dell Workstation PWS340 Intel Pentium 4, 1.7 GHz, 1.6 GB RAM). However, it should be noted that in the recent years the intensity-based methods were a subject of intensive research that led to much more efficient implementations. The processing times of 180 s were reported for hardware acceleration by NVidia FX 5600 graphics card with 256 megabytes of texture memory (Khamene *et al.*, 2006), while the processing times of 25-100 s were reported for software acceleration using a 2.2 GHz Intel Xeon processor (Russakoff *et al.*, 2005b). These accelerations make the intensity-based methods feasible for clinical use. The speed of the RGRBe method depends on the number  $I$  of the extracted 3D gradients that correspond to distinctive boundaries of anatomical structures and ranges from approximately 30 to 60 s on a Windows XP Intel Core 2 Duo, 2.13 GHz (Table 8.2). The method was implemented in C++ and was not optimized for speed, although this could easily be achieved by parallelization of finding the point-to-point geometrical correspondences in the hypothesis generation step or by a more optimal determination of the number  $L$  of random selections in the hypothesis selection step, to mention but a few.

Fourth, considering the clinical feasibility, the proposed methods, as well as the IB, GB and RB methods, do not require intra-interventional segmentation of 2D X-ray images, which is a drawback of feature-based methods. The extraction of the set of 3D gradients is accomplished by simple thresholding and since this is done prior to intervention, other more sophisticated methods can be used to extract less but more representative surface points of the 3D anatomical structures of interest and thereby further improve the registration

accuracy and speed. Nevertheless, the registration results demonstrate that the proposed RGRBe method outperforms the IB, GB and RB methods in terms of registration accuracy and robustness. This is especially true when only two X-ray images are used for registration and for MR to X-ray registration, which are important advantages for numerous clinical applications. Furthermore, additional experiments (Table 8.5) showed that for the CT to X-ray registration the proposed RGRBe method is robust to outliers like medical tools introduced in the intra-interventional X-ray images. This was not the case for the MR to X-ray registration, where outliers hampered the success rate when more than two X-rays were used for registration. Nevertheless, when only two X-rays were used, the registration accuracy and robustness were similar to the cases with no outliers also in MR to X-ray registration.

## 8.5 Conclusion

The advantages of gradient and reconstruction-based methods were efficiently combined into a novel robust 3D/2D registration method. The proposed method can register both 3D CT and MR pre-interventional images to only a few 2D X-ray intra-interventional images. The method was thoroughly tested using the standardized evaluation methodology for 3D/2D registration (van de Kraats *et al.*, 2005a) and objectively compared to some of the well-known registration methods (Penney *et al.*, 1998; Tomažević *et al.*, 2006; Tomažević *et al.*, 2003). The obtained results indicated better registration accuracy and robustness, especially when just a few X-ray images and when MR images were used for registration. As such, the proposed method might prove valuable in various image-guided interventions.

# Bibliography

- Adler, J.R., Murphy, M.J., Chang, S.D., Hancock, S.L. (1999). Image-guided robotic radiosurgery. *Neurosurgery* 44, 1299-1306.
- Andersen, A.H., Kak, A.C. (1984). Simultaneous algebraic reconstruction technique (SART): a superior implementation of the art algorithm. *Ultrason Imaging* 6, 81-94.
- Arun, K.S., Huang, T.S., Blostein, S.D. (1987). Least-squares fitting of two 3-D point sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 9, 698-700.
- Astola, J., Virtanen, I. (1982). Entropy correlation coefficient, a measure of statistical dependence for categorized data. *Proc. University Vaasa, Discussion Papers*. University of Vaasa, p. 12p.
- Aubry, J.F., Beaulieu, L., Girouard, L.M., Aubin, S., Tremblay, D., Laverdiere, J., Vigneault, E. (2004). Measurements of intrafraction motion and interfraction and intrafraction rotation of prostate by three-dimensional analysis of daily portal imaging with radiopaque markers. *International Journal of Radiation Oncology Biology Physics* 60, 30-39.
- Baert, S.A., van de Kraats, E.B., van Walsum, T., Viergever, M.A., Niessen, W.J. (2003). Three-dimensional guide-wire reconstruction from biplane image sequences for integrated display in 3-D vasculature. *IEEE Transactions on Medical Imaging* 22, 1252-1258.
- Baert, S.A.M., Penney, G.P., van Walsum, T., Niessen, W.J. (2004). Precalibration Versus 2D-3D Registration for 3D Guide Wire Display in Endovascular Interventions. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2004, Part 2*. Springer, pp. 577-584.
- Bansal, R., Staib, L., Chen, Z., Rangarajan, A., Knisely, J., Nath, R., Duncan, J.S. (1999). A Minimax Entropy Registration Framework for Patient Setup Verification in Radiotherapy. *Computer Aided Surgery* 4, 287-304.
- Beavis, A.W., P.Gibbs, Dealey, R.A., Whitton, V.J. (1998). Radiotherapy treatment planning of brain tumours using MRI alone. *British Journal of Radiology* 71, 544-548.
- Behrenbruch, C.P., Marias, K., Armitage, P.A., Yam, M., Moore, N., English, R.E., Brady, J.M. (2000). MRI - Mammography 2D/3D data fusion for

- breast pathology assessment. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2000*. Springer, pp. 307-316.
- Benameur, S., Mignotte, M., Destrempe, F., De Guise, J.A. (2005a). Three-dimensional biplanar reconstruction of scoliotic rib cage using the estimation of a mixture of probabilistic prior models. *IEEE Transactions on Biomedical Engineering* 52, 1713-1728.
- Benameur, S., Mignotte, M., Labelle, H., De Guise, J.A. (2005b). A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine. *IEEE Transactions on Biomedical Engineering* 52, 2041-2057.
- Benameur, S., Mignotte, M., Parent, S., Labelle, H., Skalli, W., de Guise, J. (2003). 3D/2D registration and segmentation of scoliotic vertebrae using statistical models. *Computerized Medical Imaging and Graphics* 27, 321-337.
- Berbeco, R.I., Jiang, S.B., Sharp, G.C., Chen, G.T., Mostafavi, H., Shirato, H. (2004). Integrated radiotherapy imaging system (IRIS): Design considerations of tumour tracking with linac gantry-mounted diagnostic x-ray systems with flat-panel detectors. *Physics in Medicine and Biology* 49, 243-255.
- Besl, P.J., McKay, N.D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14, 239-256.
- Bijhold, J. (1993). Three-dimensional verification of patient placement during radiotherapy using portal images. *Medical Physics* 20, 347-356.
- Birkfellner, W., Seemann, R., Figl, M., Hummel, J., Ede, C., Homolka, P., Yang, X.H., Niederer, P., Bergmann, H. (2005). Wobbled splatting - a fast perspective volume rendering method for simulation of x-ray images from CT. *Physics in Medicine and Biology* 50, N73-N84.
- Birkfellner, W., Wirth, J., Burgstaller, W., Baumann, B., Staedele, H., Hammer, B., Gellrich, N.C., Jacob, A.L., Regazzoni, P., Messmer, P. (2003). A faster method for 3D/2D medical image registration - a simulation study. *Physics in Medicine and Biology* 48, 2665-2679.
- Boda-Heggemann, J., Walter, C., Rahn, A., Wertz, H., Loeb, I., Lohr, F., Wenz, F. (2006). Repositioning accuracy of two different mask systems - 3D revisited: Comparison using true 3D/3D matching with cone-beam CT. *International Journal of Radiation Oncology Biology Physics* 66, 1568-1575.

- Bose, C.B., Amir, I. (1990). Design of Fiducials for Accurate Registration Using Machine Vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, 1196-1200.
- Brunie, L., Lavallée, S., Troccaz, J., Cinquin, P., Bolla, M. (1993). Pre- and intra-irradiation multimodal image registration: principles and first experiments. *Radiotherapy and Oncology* 29, 244-252.
- Byrne, J.V., Colominas, C., Hipwell, J., Cox, T., Noble, J.A., Penney, G.P., Hawkes, D.J. (2004). Assessment of a technique for 2D-3D registration of cerebral intra-arterial angiography. *British Journal of Radiology* 77, 123-128.
- Chen, L., Price, R.A., Jr., Wang, L., Li, J., Qin, L., McNeely, S., Ma, C.M., Freedman, G.M., Pollack, A. (2004). MRI-based treatment planning for radiotherapy: dosimetric verification for prostate IMRT. *International Journal of Radiation Oncology Biology Physics* 60, 636-647.
- Chiòrboli, G., Vecchi, G.P. (1993). Design of Fiducials for Accurate Registration Using Machine Vision - Comment. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15, 1330-1332.
- Clippe, S., Sarrut, D., Malet, C., Miguet, S., Ginestet, C., Carrie, C. (2003). Patient setup error measurement using 3D intensity-based image registration techniques. *International Journal of Radiation Oncology Biology Physics* 56, 259-265.
- Colchester, A.C.F., Zhao, J., Holton-Tainter, K.S., Henri, C.J., Maitland, N., Roberts, P.T.H., Harris, C.G., Evans, R.J. (1996). Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging. *Medical Image Analysis* 1, 73-90.
- Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., Marchal, G. (1995). Automated multi-modality image registration based on information theory. *Information Processing in Medical Imaging*. Kluwer, Norwell, MA, pp. 263-274.
- Cover, T.M., Thomas, J.A. (1991). *Elements of Information Theory*. John Wiley & Sons, New York, USA.
- Czopf, Á., Brack, C., Roth, M., Schweikard, A. (1999). 2D-3D Registration of Curved Objects. *Periodica Polytechnica-Electrical Engineering* 43, 19-41.
- De Buck, S., Maes, F., Ector, J., Bogaert, J., Dymarkowski, S., Heidbuchel, H., Suetens, P. (2005). An augmented reality system for patient-specific guidance of cardiac catheter ablation procedures. *IEEE Transactions on Biomedical Engineering* 24, 1512-1524.

- Deligianni, F., Chung, A.J., Yang, G.Z. (2006). Nonrigid 2-D/3-D registration for patient specific bronchoscopy simulation with statistical shape modeling: Phantom validation. *IEEE Transactions on Medical Imaging* 25, 1462-1471.
- Dey, J., Napel, S. (2006). Targeted 2D/3D registration using ray normalization and a hybrid optimizer. *Medical Physics* 33, 4730-4738.
- Feldmar, J., Ayache, N., Betting, F. (1997a). 3D-2D projective registration of free-form curves and surfaces. *Computer Vision and Image Understanding* 65, 403-424.
- Feldmar, Y., Malandain, G., Ayache, N., Fernandez-Vidal, S., Maurincommé, E., Trousset, Y. (1997b). Matching 3D MR angiography data and 2D X-ray angiograms. *Virtual Reality and Robotics in Medicine and Medical Robotics and Computer-Assisted Surgery, CVRMed-MRCAS'97*. Springer, pp. 129-140.
- Fischler, M.A., Bolles, R.C. (1981). Random Sample Consensus - A Paradigm For Model-Fitting With Applications To Image-Analysis And Automated Cartography. *Communications of the ACM* 24, 381-395.
- Fitzpatrick, J.M., West, J.B. (2001). The distribution of target registration error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging* 20, 917-927.
- Fitzpatrick, J.M., West, J.B., Maurer, C.R., Jr. (1998). Predicting error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging* 17, 694-702.
- Fleute, M., Lavallée, S. (1999). Nonrigid 3-D/2-D registration of images using statistical models. *Medical Image Computing and Computer-Assisted Intervention, MICCAI '99*. Springer, pp. 138-147.
- Ford, E.C., Chang, J., Mueller, K., Sidhu, K., Todor, D., Mageras, G., Yorke, E., Ling, C.C., Amols, H. (2002). Cone-beam CT with megavoltage beams and an amorphous silicon electronic portal imaging device: potential for verification of radiotherapy of lung cancer. *Medical Physics* 29, 2913-2924.
- Gall, K.P., Verhey, L.J., Wagner, M. (1993). Computer-Assisted Positioning Of Radiotherapy Patients Using Implanted Radiopaque Fiducials. *Medical Physics* 20, 1153-1159.
- Galloway, R.L. (2001). The process and development of image-guided procedures. *Annual Review of Biomedical Engineering* 3, 83-108.

- Gilhuijs, K.G., van de Ven, P.J., van Herk, M. (1996a). Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Medical Physics* 23, 389-399.
- Gilhuijs, K.G.A., Drukker, K., Touw, A., vandeVen, P.J.H., vanHerk, M. (1996b). Interactive three dimensional inspection of patient setup in radiation therapy using digital portal images and computed tomography data. *International Journal of Radiation Oncology Biology Physics* 34, 873-885.
- Goitein, M., Abrams, M., Rowell, D., Pollari, H., J., W. (1983). Multidimensional treatment planning. 2. Beam eye-view, back projection, and projection through CT sections. *International Journal of Radiation Oncology Biology Physics* 9, 789-797.
- Gong, R.H., Stewart, A.J., Abolmaesumi, P. (2006). A new method for CT to fluoroscope registration based on unscented Kalman filter. *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 9, 891-898.
- Gordon, R., Bender, R., Herman, G.T. (1970). Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and x-ray photography. *Journal of Theoretical Biology* 29, 471-481.
- Gottesfeld Brown, L.M., Boult, T.E. (1996). Registration of planar film radiographs with computed tomography. *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, pp. 42-51.
- Grass, M., Koppe, R., Klotz, E., Proksa, R., Kuhn, M., Aerst, H., Beek, J.O.d.B. (1999). Three-dimensional reconstruction of high contrast objects using C-arm image intensifier projection data. *Computerized Medical Imaging and Graphics* 23, 311-321.
- Grimson, W.E.L., Ettinger, G.J., White, S.J., Lozano-Perez, T., Wells III, W.M., Kikinis, R. (1996). Automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization. *IEEE Transactions on Medical Imaging* 15, 129-140.
- Groh, B.A., Siewerdse, J.H., Drake, D.G., Wong, J.W., Jaffray, D.A. (2002). A performance comparison of flat-panel imager-based MV and kV cone-beam CT. *Medical Physics* 29, 967-975.
- Groher, M., Padoy, N., Jakobs, T.F., Navab, N. (2006). New CTA protocol and 2D-3D registration method for liver catheterization. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2006, Pt 1*. Springer, pp. 873-881.

- Guéziec, A., Kazanzides, P., Williamson, B., Taylor, R.H. (1998). Anatomy-based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot. *IEEE Transactions on Medical Imaging* 17, 715-728.
- Guéziec, A., Wu, K., Kalvin, A., Williamson, B., Kazanzides, P., Van Vorhis, R. (2000). Providing visual information to validate 2-D to 3-D registration. *Medical Image Analysis* 4, 357-374.
- Haaker, P., Klotz, E., Koppe, R., Linde, R. (1990). Real-time distortion correction of digital X-ray II/TV-systems: an application example for Digital Flashing Tomosynthesis (DFTS). *International Journal of Cardiac Imaging* 6, 39-45.
- Hamadeh, A., Lavallée, S., Cinquin, P. (1998). Automated 3-dimensional computed tomographic and fluoroscopic image registration. *Computer Aided Surgery* 3, 11-19.
- Hawkes, D.J. (1998). Algorithms for radiological image registration and their clinical application. *Journal of Anatomy* 193 (Pt 3), 347-361.
- Herring, J.L., Dawant, B.M., Maurer, C.R., Jr., Muratore, D.M., Galloway, R.L., Fitzpatrick, J.M. (1998). Surface-based registration of CT images to physical space for image-guided surgery of the spine: a sensitivity study. *IEEE Transactions on Medical Imaging* 17, 743-752.
- Hipwell, J.H., Penney, G.P., McLaughlin, R.A., Rhode, K., Summers, P., Cox, T.C., Byrne, J.V., Noble, J.A., Hawkes, D.J. (2003). Intensity-based 2-D-3-D registration of cerebral angiograms. *IEEE Transactions on Medical Imaging* 22, 1417-1426.
- Hoad, C.L., Martel, A.L., Kerslake, R., Grevitt, M. (2001). A 3D MRI sequence for computer assisted surgery of the lumbar spine. *Physics in Medicine and Biology* 46, 213-220.
- Hong, H., Kim, K., Park, S. (2006). Fast 2D-3D point-based registration using GPU-based preprocessing for image-guided surgery. *Progress In Pattern Recognition, Image Analysis And Applications, Proceedings*, pp. 218-226.
- Hristov, D.H., Fallone, B.G. (1996). A grey-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs. *Medical Physics* 23, 75-84.
- Jaffray, D.A. (2005). Emergent technologies for 3-dimensional image-guided radiation delivery. *Seminars in Radiation Oncology* 15, 208-216.
- Jaffray, D.A., Drake, D.G., Moreau, M., Martinez, A.A., Wong, J.W. (1999). A radiographic and tomographic imaging system integrated into a medical

- linear accelerator for localization of bone and soft-tissue targets. *International Journal of Radiation Oncology Biology Physics* 45, 773-789.
- Jaffray, D.A., Siewerdsen, J.H., Wong, J.W., Martinez, A.A. (2002). Flat-panel cone-beam computed tomography for image-guided radiation therapy. *International Journal of Radiation Oncology Biology Physics* 53, 1337-1349.
- Jannin, P., Fitzpatrick, J.M., Hawkes, D.J., Pennec, X., Shahidi, R., Vannier, M.W. (2002). White paper: Validation of Medical Image Processing in Image-guided Therapy. In: Lemke, H.U., Vannier, M.W., Inamura, K., Farman, A.G., Doi, K., Reiber, J.H.C. (Eds.), *Proceedings of CARS 2002*. Springer, Amsterdam, pp. 299-305.
- Jannin, P., Grova, C., Maurer, C. (2006). Model for defining and reporting reference-based validation protocols in medical image processing. *International Journal of Computer Assisted Radiology and Surgery* 1, 63.
- Jans, H.S., Syme, A.M., Rathee, S., Fallone, B.G. (2006). 3D interfractional patient position verification using 2D-3D registration of orthogonal images. *Medical Physics* 33, 1420-1439.
- Jaramaz, B., Eckman, K. (2006). 2D/3D registration for measurement of implant alignment after total hip replacement. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2006, Pt 2*. Springer, pp. 653-661.
- Jin, J.Y., Ryu, S., Faber, K., Mikkelsen, T., Chen, Q., Li, S.D., Movsas, B. (2006). 2D/3D image fusion for accurate target localization and evaluation of a mask based stereotactic system in fractionated stereotactic radiotherapy of cranial lesions. *Medical Physics* 33, 4557-4566.
- Jolesz, F.A. (1997). Image-guided procedures and the operating room of the future. *Radiology* 204, 601-612.
- Jomier, J., Bullitt, E., Van Horn, M., Pathak, C., Aylward, S.R. (2006). 3D/2D model-to-image registration applied to TIPS surgery. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2006, Pt 2*. Springer, pp. 662-669.
- Jones, D., Christopherson, D.A., Washington, J.T., Hafermann, M.D., Rieke, J.W., Travaglini, J.J., Vermeulen, S.S. (1993). A frameless method for stereotactic radiotherapy. *British Journal of Radiology* 66, 1142-1150.
- Jonić, S., Thévenaz, P.E., Unser, M. (2001). Multiresolution spline-based 3D/2D registration of CT volume and C-arm images for computer-

- assisted surgery. In: Sonka, M., Hanson, K.M. (Eds.), *Medical Imaging 2001: Image Processing*. SPIE, pp. 1101-1109.
- Kak, A.C., Slaney, M. (1988). *Principles of Computerized Tomographic Imaging*. IEEE Press, New York.
- Karger, C.P., Jakel, O., Debus, J., Kuhn, S., Hartmann, G.H. (2001). Three-dimensional accuracy and interfractional reproducibility of patient fixation and positioning using a stereotactic head mask system. *International Journal of Radiation Oncology Biology Physics* 49, 1493-1504.
- Khamene, A., Bloch, P., Wein, W., Svatos, M., Sauer, F. (2006). Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. *Medical Image Analysis* 10, 96-112.
- Kim, J., Fessler, J.A., Lam, K.L., Balter, J.M., Ten Haken, R.K. (2001). A feasibility study of mutual information based setup error estimation for radiotherapy. *Medical Physics* 28, 2507-2517.
- Kim, J., Li, S.D., Pradhan, D., Hammoud, R., Chen, Q., Yin, F.F., Zhao, Y., Kim, J.H., Movsas, B. (2007). Comparison of similarity measures for rigid-body CT/Dual x-ray image registrations. *Technology in Cancer Research & Treatment* 6, 337-345.
- Kita, Y., Wilson, D.L., Noble, J.A. (1998). Real-Time Registration of 3D Cerebral Vessels to X-ray Angiograms. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*. Springer, pp. 1125-1133.
- Koppe, R., Klotz, E., Op de Beeck, J., Aerts, H. (1995). 3D Vessel Reconstruction Based on Rotational Angiography. *Computer Assisted Radiology*, 101-107.
- LaRose, D., Bayouth, J., Kanade, T. (2000a). Transgraph: interactive intensity-based 2D/3D registration of X-ray and CT data. In: Hanson, K.M. (Ed.), *Medical Imaging 2000: Image Processing*. SPIE, pp. 385-396.
- LaRose, D.A., Bayouth, J., Kanade, T. (2000b). Evaluating 2D/3D registration accuracy. *Computer Assisted Radiology and Surgery*, 147-152.
- Lavallée, S. (1995). Registration for Computed Integrated Surgery: Methodology, state of the art. In: Taylor, R.H., Lavallée, S., Burdea, G.C., Mosges, R.W. (Eds.), *Computer-Integrated Surgery: Technology and Clinical Applications*. MIT Press, Cambridge, MA, pp. 425-449.
- Lavallée, S., Sautot, P., Troccaz, J., Cinquin, P., Merloz, P. (1995). Computer-assisted spine surgery: a technique for accurate transpedicular screw

- fixation using CT data and a 3-D optical localizer. *Journal of Image Guided Surgery* 1, 65-73.
- Lavallée, S., Szeliski, R. (1995). Recovering the Position and Orientation of Free-Form Objects from Image Contours Using 3d Distance Maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17, 378-390.
- Lee, Y.K., Bollet, M., Charles-Edwards, G., Flower, M.A., Leach, M.O., McNair, H., Moore, E., Rowbottom, C., Webb, S. (2003). Radiotherapy treatment planning of prostate cancer using magnetic resonance imaging alone. *Radiotherapy and Oncology* 66, 203-216.
- Lemieux, L., Jagoe, R., Fish, D.R., Kitchen, N.D., Thomas, D.G. (1994). A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Medical Physics* 21, 1749-1760.
- Letourneau, D., Wong, J.W., Oldham, M., Gulam, M., Watt, L., Jaffray, D.A., Siewerdsen, J.H., Martinez, A.A. (2005). Cone-beam-CT guided radiation therapy: Technical implementation. *Radiotherapy and Oncology* 75, 279-286.
- Li, X.L., Yang, J., Zhu, Y.M. (2006). Digitally reconstructed radiograph generation by an adaptive Monte Carlo method. *Physics in Medicine and Biology* 51, 2745-2752.
- Likar, B., Viergever, M.A., Pernuš, F. (2001). Retrospective correction of MR intensity inhomogeneity by information minimization. *IEEE Transactions on Medical Imaging* 20, 1398-1410.
- Litzenberg, D., Dawson, L.A., Sandler, H., Sanda, M.G., McShan, D., Ten Haken, R.K., Lam, K.L., Brock, K.K., Balter, J.M. (2002). Daily prostate targeting using implanted radiopaque markers. *International Journal of Radiation Oncology Biology Physics* 52, 699-703.
- Livyatan, H., Yaniv, Z., Joskowicz, L. (2002). Robust Automatic C-Arm Calibration for Fluoroscopy-Based Navigation: A Practical Approach. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002*. Springer, pp. 60-68.
- Livyatan, H., Yaniv, Z., Joskowicz, L. (2003). Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. *IEEE Transactions on Medical Imaging* 22, 1395-1406.
- Lujan, A.E., Balter, J.M., Ten Haken, R.K. (1998). Determination of rotations in three dimensions using two-dimensional portal image registration. *Medical Physics* 25, 703-708.

- Mackie, T.R., Kapatoes, J., Ruchala, K., Lu, W., Wu, C., Olivera, G., Forrest, L., Tome, W., Welsh, J., Jeraj, R., Harari, P., Reckwerdt, P., Paliwal, B., Ritter, M., Keller, H., Fowler, J., Mehta, M. (2003). Image guidance for precise conformal radiotherapy. *International Journal of Radiation Oncology Biology Physics* 56, 89-105.
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P. (1997). Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging* 16, 187-198.
- Mahfouz, M.R., Hoff, W.A., Komistek, R.D., Dennis, D.A. (2003). A robust method for registration of three-dimensional knee implant models to two-dimensional fluoroscopy images. *IEEE Transactions on Medical Imaging* 22, 1561-1574.
- Maintz, J.B., Viergever, M.A. (1998). A survey of medical image registration. *Medical Image Analysis* 2, 1-36.
- Markelj, P., Tomaževič, D., Pernuš, F., Likar, B. (2007). Optimizing bone extraction in MR images for 3D/2D gradient based registration of MR and X-ray images. In: Reinhardt, J.M., Pluim, J.P. (Eds.), *Medical Imaging 2007: Image Processing*. SPIE.
- Masutani, Y., Dohi, T., Yamane, F., Iseki, H., Takakura, K. (1997). Interactive virtualized display system for intravascular neurosurgery. *CVRMed-MRCAS '97* 1205, 427-435.
- Maurer, C.R., Jr., Fitzpatrick, J.M., Wang, M.Y., Galloway, R.L., Jr., Maciunas, R.J., Allen, G.S. (1997). Registration of head volume images using implantable fiducial markers. *IEEE Transactions on Medical Imaging* 16, 447-462.
- Maurer, C.R., Jr., Maciunas, R.J., Fitzpatrick, J.M. (1998). Registration of head CT images to physical space using a weighted combination of points and surfaces. *IEEE Transactions on Medical Imaging* 17, 753-761.
- Maurer, C.R.J., Fitzpatrick, J.M. (1993). A review of medical image registration. In: Maciunas, R.J. (Ed.), *Interactive Image-Guided Neurosurgery*. American Association of Neurological Surgeons, Park Ridge, pp. 17-44.
- McLaughlin, R.A., Hipwell, J., Hawkes, D.J., Noble, J.A., Byrne, J.V., Cox, T. (2002). A Comparison of 2D-3D Intensity-Based Registration and Feature-Based Registration for Neurointerventions. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002, Part 2*. Springer, pp. 517-524.

- McLaughlin, R.A., Hipwell, J., Hawkes, D.J., Noble, J.A., Byrne, J.V., Cox, T.C. (2005). A comparison of a similarity-based and a feature-based 2-D-3-D registration method for neurointerventional use. *IEEE Transactions on Medical Imaging* 24, 1058-1066.
- Midgley, S., Millar, R.M., Dudson, J. (1998). A feasibility study for megavoltage cone beam CT using a commercial EPID. *Physics in Medicine and Biology* 43, 155-169.
- Movassaghi, B., Rasche, V., Grass, M., Viergever, M.A., Niessen, W.J. (2004). A Quantitative Analysis of 3-D Coronary Modeling From Two or More Projection Images. *IEEE Transactions on Medical Imaging* 23, 1517-1531.
- Munbodh, R., Chen, Z., Jaffray, D.A., Moseley, D.J., Knisely, J.P., Duncan, J.S. (2007). A frequency-based approach to locate common structure for 2D-3D intensity-based registration of setup images in prostate radiotherapy. *Medical Physics* 34, 3005-3017.
- Murphy, M.J. (1997). An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery. *Medical Physics* 24, 857-866.
- Nakajima, Y., Tashiro, T., Sugano, N., Yonenobu, K., Koyama, T., Maeda, Y., Tamura, Y., Saito, M., Tamura, S., Mitsuishi, M., Sugita, N., Sakuma, I., Ochi, T., Matsumoto, Y. (2007). Fluoroscopic bone fragment tracking for surgical navigation in femur fracture reduction by incorporating optical tracking of hip joint rotation center. *IEEE Transactions on Biomedical Engineering* 54, 1703-1706.
- Penney, G.P., Batchelor, P.G., Hill, D.L., Hawkes, D.J., Weese, J. (2001). Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images. *Medical Physics* 28, 1024-1032.
- Penney, G.P., Edwards, P.J., Hipwell, J.H., Slomczykowski, M., Revie, I., Hawkes, D.J. (2007). Postoperative calculation of acetabular cup position using 2-D-3-D registration. *IEEE Transactions on Biomedical Engineering* 54, 1342-1348.
- Penney, G.P., Weese, J., Little, J.A., Desmedt, P., Hill, D.L., Hawkes, D.J. (1998). A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Transactions on Medical Imaging* 17, 586-595.
- Peters, T.M. (2000). Image-guided surgery: from X-rays to virtual reality. *Computer Methods in Biomechanics and Biomedical Engineering* 4, 27-57.
- Pouliot, J., Bani-Hashemi, A., Chen, J., Svatos, M., Ghelmansarai, F., Mitschke, M., Aubin, M., Xia, P., Morin, O., Bucci, K., Roach, M., 3rd,

- Hernandez, P., Zheng, Z., Hristov, D., Verhey, L. (2005). Low-dose megavoltage cone-beam CT for radiation therapy. *Int J Radiat Oncol Biol Phys* 61, 552-560.
- Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1992). *Numerical Recipes in C*, Second ed. Cambridge University Press, Cambridge, UK.
- Prümmer, M., Hornegger, J., Pfister, M., Dörfler, A. (2006). Multi-modal 2D-3D non-rigid registration. In: Reinhardt, J.M., Pluim, J.P. (Eds.), *Medical Imaging 2006: Image Processing*. SPIE.
- Rampersaud, Y.R., Simon, D.A., Foley, K.T. (2001). Accuracy requirements for image-guided spinal pedicle screw placement. *Spine* 26, 352-359.
- Ramsey, C.R., Arwood, D., Scaperoth, D., Oliver, A.L. (1999). Clinical application of digitally-reconstructed radiographs generated from magnetic resonance imaging for intracranial lesions. *International Journal of Radiation Oncology Biology Physics* 45, 797-802.
- Rasche, V., Schreiber, B., Graeff, C., Istel, T., Schomberg, H., Grass, M., Koppe, R., Klotz, E., Rose, G. (2003). Performance of image intensifier-equipped X-ray systems for three-dimensional imaging. *Computer Assisted Radiology and Surgery*. Elsevier Science B.V., pp. 187-192.
- Rhode, K.S., Hill, D.L., Edwards, P.J., Hipwell, J., Rueckert, D., Sanchez-Ortiz, G., Hegde, S., Rahunathan, V., Razavi, R. (2003). Registration and tracking to integrate X-ray and MR images in an XMR facility. *IEEE Transactions on Medical Imaging* 22, 1369-1378.
- Rhode, K.S., Sermesant, M., Brogan, D., Hegde, S., Hipwell, J., Lambiase, P., Rosenthal, E., Bucknall, C., Qureshi, S.A., Gill, J.S., Razavi, R., Hill, D.L. (2005). A system for real-time XMR guided cardiovascular intervention. *IEEE Transactions on Medical Imaging* 24, 1428-1440.
- Roche, A., Malandain, G., Pennec, X., Ayache, N. (1998). The correlation ratio as a new similarity measure for multimodality image registration. *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98*. Springer, pp. 1115-1124.
- Rohlfing, T., Denzler, J., Grassl, C., Russakoff, D.B., Maurer, C.R. (2005a). Markerless real-time 3-D target region tracking by motion backprojection from projection images. *IEEE Transactions on Medical Imaging* 24, 1455-1468.
- Rohlfing, T., Russakoff, D.B., Denzler, J., Mori, K., Maurer Jr., C.R. (2005b). Progressive attenuation fields: Fast 2D-3D image registration without precomputation. *Medical Physics* 32, 2870-2880.

- Rohlfing, T., Russakoff, D.B., Murphy, M.J., Maurer Jr., C.R. (2002). An intensity-based registration algorithm for probabilistic images and its application for 2-D to 3-D image registration. *Medical Imaging 2002: Image Processing* 4683, 581-591.
- Russakoff, D.B., Rohlfing, T., Adler, J.R., Maurer, C.R. (2005a). Intensity-based 2D-3D spine image registration incorporating a single fiducial marker. *Academic Radiology* 12, 37-50.
- Russakoff, D.B., Rohlfing, T., Ho, A., Kim, D.H., Shahidi, R., Adler, J.R., Maurer, C.R. (2003a). Evaluation of Intensity-Based 2D-3D Spine Image Registration Using Clinical Gold-Standard Data. In: Gee, J.C., Maintz, J.B., Vannier, M.W. (Eds.), *Biomedical Image Registration, Second International Workshop, WBIR 2003*. Springer, pp. 151-160.
- Russakoff, D.B., Rohlfing, T., Mori, K., Rueckert, D., Ho, A., Adler, J.R., Maurer, C.R. (2005b). Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration. *IEEE Transactions on Medical Imaging* 24, 1441-1454.
- Russakoff, D.B., Rohlfing, T., Rueckert, D., Shahidi, R., Kim, D., Maurer Jr., C.R. (2003b). Fast calculation of digitally reconstructed radiographs using light fields. In: Sonka, M., Fitzpatrick, J.M. (Eds.), *Medical Imaging 2003: Image Processing*. SPIE, pp. 684-695.
- Sadowsky, O., Cohen, J.D., Taylor, R.H. (2006). Projected tetrahedra revisited: a barycentric formulation applied to digital radiograph reconstruction using higher-order attenuation functions. *IEEE Transactions on Visualization and Computer Graphics* 12, 461-473.
- Salter, B.J., Fuss, M., Vollmer, D.G., Sadeghi, A., Bogaev, C.A., Cheek, D.A., Herman, T.S., Hevezi, J.M. (2001). The TALON removable head frame system for stereotactic radiosurgery/radiotherapy: measurement of the repositioning accuracy. *International Journal of Radiation Oncology Biology Physics* 51, 555-562.
- Sarrut, D., Clippe, S. (2000). Patient positioning in radiotherapy by registration of 2D portal to 3D CT images by a content-based research with similarity measures. *Computer Assisted Radiology and Surgery*, 707-712.
- Schweikard, A., Shiomi, H., Adler, J. (2005). Respiration tracking in radiosurgery without fiducials. *International Journal of Medical Robotics and Computer Assisted Surgery* 1, 19-27.
- Sherouse, G.W., Novins, K., Chaney, E.L. (1990). Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *International Journal of Radiation Oncology Biology Physics* 18, 651-658.

- Siddon, R.L., Chin, L.M. (1985). 2-film brachytherapy reconstruction algorithm. *Medical Physics* 12, 77-83.
- Siewerdsen, J.H., Moseley, D.J., Burch, S., Bisland, S.K., Bogaards, A., Wilson, B.C., Jaffray, D.A. (2005). Volume CT with a flat-panel detector on a mobile, isocentric C-arm: Pre-clinical investigation in guidance of minimally invasive surgery. *Medical Physics* 32, 241-254.
- Simon, D., O'Toole, R.V., Blackwell, M., Morgan, F., Di Gioia, A.M., Kanade, T. (1995a). Accuracy Validation In Image-Guided Orthopaedic Surgery. *Proceedings of the Second International Symposium on Medical Robotics and Computer Assisted Surgery*, pp. 185-192.
- Simon, D.A., Hebert, M., Kanade, T. (1995b). Techniques for fast and accurate intrasurgical registration. *Journal of Image Guided Surgery* 1, 17-29.
- Sirois, L.M., Hristov, D.H., Fallone, B.G. (1999). Three-dimensional anatomy setup verification by correlation of orthogonal portal images and digitally reconstructed radiographs. *Medical Physics* 26, 2422-2428.
- Soete, G., Van de Steene, J., Verellen, D., Vinh-Hung, V., Van den Berge, D., Michielsen, D., Keuppens, F., De Roover, P., Storme, G. (2002). Initial clinical experience with infrared-reflecting skin markers in the positioning of patients treated by conformal radiotherapy for prostate cancer. *International Journal of Radiation Oncology Biology Physics* 52, 694-698.
- Studholme, C., Hill, D.L., Hawkes, D.J. (1995). Multiresolution voxel similarity measures for MR-PET registration. In: Bizais, Y. (Ed.), *Proceedings of the Information Processing in Medical Imaging*. Kluwer Academic, Ile de Berder, France, pp. 287-298.
- Studholme, C., Hill, D.L., Hawkes, D.J. (1999). An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition* 32, 71-86.
- Škerl, D., Likar, B., Pernuš, F. (2006a). A Protocol for Evaluation of Similarity Measures for Rigid Registration. *IEEE Transactions on Medical Imaging* 25, 779-791.
- Škerl, D., Tomažević, D., Likar, B., Pernus, F. (2006b). Evaluation of similarity measures for reconstruction-based registration in image-guided radiotherapy and surgery. *International Journal of Radiation Oncology Biology Physics* 65, 943-953.
- Tang, T.S.Y., Ellis, R.E., Fichtinger, G. (2000). Fiducial registration from a single X-ray image: A new technique for fluoroscopic guidance and radiotherapy. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2000*. Springer, pp. 502-511.

- Taylor, R.H., Lavallée, S., Burdea, G.C., Mosges, R.W. (1995). *Computer-Integrated Surgery: Technology and Clinical Applications*. MIT Press, Cambridge, MA.
- Taylor, R.H., Mittlestadt, B.D., Paul, H.A., Hanson, W., Kazanzides, P., Zuhars, J.F., Williamson, B., Musits, B.L., Glassman, E., Bargar, W.L. (1994). An image-directed robotic system for precise orthopaedic surgery. *IEEE Transactions on Robotics and Automation* 10, 261-275.
- Tomažević, D., Likar, B., Pernuš, F. (2002). "Gold standard" 2D/3D registration of X-ray to CT and MR images., *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2002*. Springer, pp. 461-468.
- Tomažević, D., Likar, B., Pernuš, F. (2004a). "Gold standard" data for evaluation and comparison of 3D/2D registration methods. *Computer Aided Surgery* 9, 137-144.
- Tomažević, D., Likar, B., Pernuš, F. (2004b). Multi-feature mutual information. In: Fitzpatrick, J.M., Sonka, M. (Eds.), *Medical Imaging 2004: Image Processing*. SPIE, pp. 143-154.
- Tomažević, D., Likar, B., Pernuš, F. (2006). 3-D/2-D registration by integrating 2-D information in 3-D. *IEEE Transactions on Medical Imaging* 25, 17-27.
- Tomažević, D., Likar, B., Slivnik, T., Pernuš, F. (2003). 3-D/2-D registration of CT and MR to X-ray images. *IEEE Transactions on Medical Imaging* 22, 1407-1416.
- Umeyama, S. (1991). Least-squares estimation of transformation parameters between two point patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13, 376-380.
- van de Kraats, E.B., Penney, G.P., Tomažević, D., van Walsum, T., Niessen, W.J. (2005a). Standardized evaluation methodology for 2-D-3-D registration. *IEEE Transactions on Medical Imaging* 24, 1177-1189.
- van de Kraats, E.B., Penney, G.P., van Walsum, T., Niessen, W.J. (2005b). Multispectral MR to x-ray registration of vertebral bodies by generating CT-Like data. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2005, Pt 2*. Springer, pp. 911-918.
- van de Kraats, E.B., van Walsum, T., Kendrick, L., Noordhoek, N.J., Niessen, W.J. (2006). Accuracy evaluation of direct navigation with an isocentric 3D rotational X-ray system. *Medical Image Analysis* 10, 113.

- van de Kraats, E.B., van Walsum, T., Verlaan, J.J., Oner, F.C., Viergever, M.A., Niessen, W.J. (2004). Noninvasive magnetic resonance to three-dimensional rotational x-ray registration of vertebral bodies for image-guided spine surgery. *Spine* 29, 293-297.
- Vermandel, M., Betrouni, N., Palos, G., Gauvrit, J.Y., Vasseur, C., Rousseau, J. (2003). Registration, Matching, and Data Fusion in 2D/3D Medical Imaging: Application to DSA and MRA. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2003*. Springer, pp. 778-785.
- Viola, P., Wells III, W.M. (1997). Alignment by Maximization of Mutual Information. *International Journal of Computer Vision* 24, 137-154.
- Viola, P., Wells, W.M., 3rd (1995). Alignment by maximization of mutual information. *Proc. 5th Int. Conf. Computer Vision*, Boston, MA, pp. 16-23.
- Wall, B.F., Hart, D. (1997). Revised radiation doses for typical X-ray examinations. *British Journal of Radiology* 70, 437-439.
- Weese, J., Göcke, R., Penny, G.P., Desmedt, P., Buzug, T.M., Schumann, H. (1999). Fast voxel-based 2D/3D registration algorithm using a volume rendering method based on the shear-warp factorization. In: Hanson, K.M. (Ed.), *Medical Imaging 1999: Image Processing*. SPIE, pp. 802-810.
- Weese, J., Penney, G.P., Desmedt, P., Buzug, T.M., Hill, D.L., Hawkes, D.J. (1997). Voxel-based 2-D/3-D registration of fluoroscopy images and CT scans for image-guided surgery. *IEEE Transactions on Information Technology in Biomedicine* 1, 284-293.
- Wells, W.M., 3rd, Viola, P., Atsumi, H., Nakajima, S., Kikinis, R. (1996). Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis* 1, 35-51.
- West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer, C.R., Jr., Kessler, R.M., Maciunas, R.J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., van den Elsen, P.A., Napel, S., Sumanaweera, T.S., Harkness, B., Hemler, P.F., Hill, D.L., Hawkes, D.J., Studholme, C., Maintz, J.B., Viergever, M.A., Malandain, G., Woods, R.P., et al. (1997). Comparison and evaluation of retrospective intermodality brain image registration techniques. *Journal of Computer Assisted Tomography* 21, 554-566.
- Woods, R.P., Mazziotta, J.C., Cherry, S.R. (1993). MRI-PET registration with automated algorithm. *Journal of Computer Assisted Tomography* 17, 536-546.
- Yamazaki, T., Watanabe, T., Nakajima, Y., Sugamoto, K., Tomita, T., Yoshikawa, H., Tamura, S. (2004). Improvement of depth position in 2-

- D/3-D registration of knee implants using single-plane fluoroscopy. *IEEE Transactions on Medical Imaging* 23, 602-612.
- Yin, F.F., Gao, Q.H., Xie, H.C., Nelson, D.F., Yu, Y., Kwok, W.E., Totterman, S., Schell, M.C., Rubin, P. (1998). MR image-guided portal verification for brain treatment field. *International Journal of Radiation Oncology Biology Physics* 40, 703-711.
- You, B.M., Siy, P., Anderst, W., Tashman, S. (2001). In vivo measurement of 3-D skeletal kinematics from sequences of biplane radiographs: application to knee kinematics. *IEEE Transactions on Medical Imaging* 20, 514-525.
- Zhang, X., Zheng, G., Langlotz, F., Nolte, L.P. (2006). Assessment of spline-based 2D-3D registration for image-guided spine surgery. *Minim Invasive Ther Allied Technol* 15, 193-199.
- Zheng, G., Dong, X., Rajamani, K.T., Zhang, X., Styner, M., Thoranaghatte, R.U., Nolte, L.P., Ballester, M.A. (2007). Accurate and robust reconstruction of a surface model of the proximal femur from sparse-point data and a dense-point distribution model for surgical navigation. *IEEE Transactions on Biomedical Engineering* 54, 2109-2122.
- Zheng, G.Y., Ballester, M.A.G., Styner, M., Nolte, L.P. (2006). Reconstruction of patient-specific 3D bone surface from 2D calibrated fluoroscopic images and point distribution model. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006, Pt 1*. Springer, pp. 25-32.



# **Samenvatting**

Het onderwerp van dit proefschrift is registratie van driedimensionale (3D) medische beelden met overeenkomstige tweedimensionale (2D) projectiebeelden, verder aangeduid door 3D/2D registratie. Het doel van 3D/2D registratie is het vinden van spatiële relaties tussen een 3D beeld en een of meer 2D projectiebeelden. Er is een groot aantal mogelijke toepassingen van 3D/2D registratie in beeldgeleide diagnostiek en behandeling. In beeldgeleide chirurgie kan registratie van preoperatieve 3D en intraoperatieve 2D beelden de chirurg voorzien van informatie over de positie van operatie-instrumenten ten opzichte van het geplande traject, wat vooral van belang is in de nabijheid van kwetsbare structuren en van het doelvolume. Ook kan 3D/2D registratie het interpreteren van 3D informatie uit een of meer 2D projectiebeelden ondersteunen. Het onderhavige proefschrift tracht nieuwe oplossingen aan te dragen voor registratie van 3D (CT, MR) beelden met 2D radiografische (i.e., met behulp van Röntgenstralen gegenereerde) projectiebeelden, maar behandelt ook belangrijke thema's als validatie van 3D/2D registratie en vergelijking van verschillende registratiemethoden.

Een belangrijke voorwaarde voor klinisch gebruik van een nieuwe registratietechniek is dat de methode is geëvalueerd en gevalideerd. Evaluatie van registratiemethoden heeft tot nog toe weinig aandacht gekregen in de literatuur. Hoofdstuk 2 van het proefschrift behandelt het definiëren van een gouden standaard om de nauwkeurigheid en de betrouwbaarheid van rigide 3D/2D registratiemethoden te evalueren. Belangrijke stappen naar de constructie van een gouden standaard zijn: 1. Definitie van de registratiebasis, als zeer nauwkeurige referentie voor registratiestudies. 2. Selectie van een patiënt voor een humane studie, of van een representatief fysisch object voor een fantoomstudie. 3. Indien noodzakelijk, kalibratie van het acquisitiesysteem. 4. Beeldacquisitie. 5. Beeldregistratie. 6. Evaluatie van de registratienuwkeurigheid. Hoofdstuk 2 beschrijft de constructie van een lumbaal wervelkolom fantoom en de acquisitie van CT, MRI en digitale Röntgenbeelden van het fantoom. Met behulp van aan het fantoom bevestigde referentiemarkers is de 2D projectie gekalibreerd en is de registratiebasis vastgelegd. De nauwkeurigheid van registratiemethoden kan vervolgens worden bepaald door de registratieresultaten te vergelijken met de referentie.

In Hoofdstuk 3 introduceren we een nieuwe aanpak voor registratie van 3D CT of MR beelden met 2D Röntgenbeelden. De methode maakt gebruik van de normaal op het botoppervlak, dat verkregen is door segmentatie van de preoperatieve 3D MR of CT data, en gradiënten in de intraoperatieve Röntgenbeelden op plaatsen bepaald door de Röntgenbron en punten op het 3D oppervlak. Het verschil tussen de oppervlaktenormalen en de gradiënten is een maat voor misregistratie van de beelden. Door deze aanpak wordt de

noodzaak van tijdrovende en foutgevoelige intraoperatieve segmentatie van Röntgenbeelden vermeden, terwijl de preoperatieve segmentatie van de 3D data aanvullende tijdwinst geeft. Ook de methode voor validatie van de voorgestelde 3D/2D registratiemethode is nieuw. Evaluatie vereist een groot bereik aan initiële registratieposities, translaties en rotaties, met een verdeling die normaal is en uniform in een gegeven interval rond de referentie registratie. De nauwkeurigheid is gedefinieerd als de 'root mean square' (RMS) 'target registration error' (TRE) van alle succesvolle registraties.

Het criterium voor succesvolle registratie is dat de maximum TRE kleiner is dan de klinisch toelaatbare fout. De betrouwbaarheid van de methode wordt uitgedrukt in het percentage succesvolle registraties uit alle registraties wier beginwaarde ligt binnen een gegeven interval rond de gouden standaard.

De nauwkeurigheid en de robuustheid van het registreren van een 3D preoperatief beeld met een 2D intraoperatief beeld hangt af van de gekozen registratiemethode en van de anatomische structuren in het beeld. Te verwachten is dat de kwaliteit van de registratie ook afhangt van het aantal 2D projectiebeelden en van de hoeken onder welke deze geacquireerd zijn. In Hoofdstuk 4 beschrijven we een studie naar de invloed op de registratie van het aantal projectiebeelden en van de hoeken tussen de projectiebeelden. De studie is uitgevoerd voor de 3D/2D gradiëntgebaseerde registratiemethode van Hoofdstuk 3 en de gouden standaard beelddata van Hoofdstuk 2. De nauwkeurigheid en betrouwbaarheid van de 3D/2D registratiemethode is geëvalueerd met behulp van preoperatieve CT en MR beelden en zestien verschillende projectiesets Röntgenbeelden. De sets Röntgenbeelden verschillen in het aantal van de beelden en in de hoeken tussen de beelden. De resultaten voor CT/radiografie en MRI/radiografie registratie wijzen uit dat het aantal projectiebeelden een kleine invloed heeft op de nauwkeurigheid, maar een grote invloed op de robuustheid: het percentage succesvolle registraties neemt significant toe met het aantal projectiebeelden. De keuze van de projectiehoeken is niet belangrijk voor CT/radiografie registratie, maar wel voor de robuustheid van MRI/radiografie registratie.

Hoofdstuk 5 behandelt het probleem van evaluatie en vergelijking van de verschillende 3D/2D registratiemethoden. In dit hoofdstuk introduceren we een nieuwe dataset als gouden standaard voor 3D/2D registratie. De data betreffen 2D fluoroscopische Röntgenbeelden en 3DRX (rotatieradiografie), CT en MR beelden van twee wervelkolom segmenten. Voor elk segment zijn met een klinisch rotatieradiografie systeem 100 fluoroscopische Röntgenbeelden opgenomen, waaruit vervolgens een 3DRX beeld gereconstrueerd is. Door 3D/3D registratie van dit beeld met de corresponderende CT en MR beelden, is de gouden standaard voor CT/fluoroscopie en MRI/fluoroscopie registratie vastgelegd. De 3D/3D registratie is uitgevoerd door de wederzijdse informatie van de beelden te maximaliseren. De methodologie voor evaluatie van de 3D/2D registratie resultaten is analoog aan de methode die beschreven is in Hoofdstuk 3, en gaat uit van verschillende beginposities voor de registratie experimenten, en van welgedefinieerde criteria voor nauwkeurigheid, convergentiebereik en misregistratie. De evaluatiemethode is gebruikt om twee 3D/2D registratiemethoden, de methode gebaseerd op digitaal

gereconstrueerde radiogrammen en de gradiëntgebaseerde methode van Hoofdstuk 3, met elkaar te vergelijken.

Tot nog toe hebben alle 3D/2D registratiemethoden die 3D beelden registreren met verscheidene 2D Röntgenprojectiebeelden, elk projectiebeeld afzonderlijk in het registratieproces betrokken. De mate van overeenkomst wordt gewoonlijk berekend als de som van alle deelregistraties tussen het 3D beeld en elk 2D projectiebeeld. In Hoofdstuk 6 introduceren we een volledig nieuwe aanpak van dit 3D/2D registratieprobleem, waarbij informatie van verscheidene 2D Röntgenbeelden wordt samengevoegd in een 3D beeldruimte en registratie plaatsvindt door dit gecombineerde projectiebeeld in overeenstemming te brengen met het 3D beeld. Het combineren van informatie uit de 2D projectiebeelden in de 3D beeldruimte gebeurt door middel van de iteratieve algebraïsche reconstructietechniek (ART). Om de overeenkomst tussen een 3D beeld en een gereconstrueerd 3D analogon te meten, stellen we een nieuw registratiecriterium voor, gebaseerd op asymmetrische multidimensionale wederzijdse informatie. De in dit hoofdstuk beschreven 3D/2D registratiemethode is geëvalueerd met behulp van de wervelkolom data en de gestandaardiseerde evaluatiemethodologie uit Hoofdstuk 5. Dezelfde evaluatieprocedure is gebruikt om de resultaten van de voorgestelde methode te vergelijken met die van de 3D/2D registratiemethode uit Hoofdstuk 3, waarbij de mate van overeenkomst wordt berekend op conventionele wijze door de deelregistraties van het 3D beeld met elk van de projectiebeelden te sommeren. De evaluatie geeft aan dat de methode uit dit hoofdstuk beter presteert dan de conventionele methode.

In Hoofdstuk 7 worden verschillende registratiematen voor op reconstructie gebaseerde 3D/2D registraties geëvalueerd. Deze maten zijn: asymmetrische gradiëntgebaseerde wederzijdse informatie, beschreven in Hoofdstuk 6, en voorts symmetrische gradiëntgebaseerde wederzijdse informatie, wederzijdse informatie, genormaliseerde wederzijdse informatie, de entropie correlatie coëfficiënt, gezamenlijke entropie, de correlatie ratio, het criterium van Woods, en Pearson's correlatie coëfficiënt. De maten zijn geëvalueerd voor een verschillend aantal 2D projecties en een verschillend aantal ART reconstructie iteraties. De evaluatie is uitgevoerd voor rigide transformaties, met behulp van het protocol dat is voorgesteld in Hoofdstuk 5. De evaluatie wijst uit dat zowel een groter aantal 2D projecties als een groter aantal reconstructie iteraties leidt tot hogere nauwkeurigheid en iets lagere robuustheid. De meest geschikte maat voor op reconstructie gebaseerde 3D/2D registratie blijkt asymmetrische multidimensionale wederzijdse informatie.

Een nieuwe 3D/2D registratiemethode wordt voorgesteld in Hoofdstuk 8. De methode is gebaseerd op het vergelijken van 3D preoperatieve beeldgradiënten en grof gereconstrueerde 3D gradiënten uit 2D interventie beelden. Ten einde de robuustheid van het vinden van de correspondenties tussen de twee sets gradiënten te verbeteren, worden hypothetische correspondenties gezocht langs normalen op anatomische structuren in de 3D beelden. Daarna volgen de uiteindelijke correspondenties uit een iteratief proces, waarin het robuuste

random sample consensus (RANSAC) algoritme en een gradiëntgebaseerd registratiecriterium worden gecombineerd. De voorgestelde methode is geëvalueerd met de evaluatiemethodologie voor 3D/2D registratie uit Hoofdstuk 5. Zo kan de methode objectief worden vergeleken met de registratiemethoden uit Hoofdstukken 3 en 6. De resultaten laten zien dat de methode uit dit hoofdstuk leidt tot een hogere registratienuwkeurigheid en een grotere robuustheid.

# Publications

## Papers in journals

Tomažević, D., Likar, B., Pernuš, F. (2002). Comparative evaluation of retrospective shading correction methods. *Journal of Microscopy* 208, 212-223.

Derganc, J., Likar, B., Bernard, R., Pernuš, F. (2003). Real-time automated visual inspection of color tablets in pharmaceutical blisters. *Real-Time Imaging* 9, 113-124.

Tomažević, D., Likar, B., Slivnik, T., Pernuš, F. (2003). 3-D/2-D registration of CT and MR to X-ray images. *IEEE Transactions on Medical Imaging* 22, 1407-1416.

Tomažević, D., Likar, B., Pernuš, F. (2004). "Gold standard" data for evaluation and comparison of 3D/2D registration methods. *Computer Aided Surgery* 9, 137-144.

Vrtovec, T., Tomažević, D., Likar, B., Travnik, L., Pernuš, F. (2004). Automated construction of 3D statistical shape models. *Image Analysis and Stereology* 23, 111-120.

van de Kraats, E.B., Penney, G.P., Tomažević, D., van Walsum, T., Niessen, W.J. (2005). Standardized evaluation methodology for 2-D-3-D registration. *IEEE Transactions on Medical Imaging* 24, 1177-1189.

Škerl, D., Tomažević, D., Likar, B., Pernuš, F. (2006). Evaluation of similarity measures for reconstruction-based registration in image-guided radiotherapy and surgery. *International Journal of Radiation Oncology Biology Physics* 65, 943-953.

Tomažević, D., Likar, B., Pernuš, F. (2006). 3-D/2-D registration by integrating 2-D information in 3-D. *IEEE Transactions on Medical Imaging* 25, 17-27.

Tomažević, D., Likar, B., Pernuš, F. Multi-Feature Mutual Information Image Registration. *Submitted for publication.*

Tomaževič, D., Likar, B., Pernuš, F. 3D/2D Registration of CT and MR Images to X-ray Images: The Impact of Number and Projection of X-ray Images. *Submitted for publication.*

Markelj, P., Tomaževič, D., Pernuš, F., Likar, B. Robust gradient-based 3D/2D registration of CT and MR to X-ray images. Accepted for publication in *IEEE Transactions on Medical Imaging*.

Možina, M., Tomaževič, D., Pernuš, F., Likar, B. Real-time segmentation of tablet images. *Submitted for publication.*

## Papers in conference proceedings

Krašna, A., Tomaževič, D., Likar, B., Pernuš, F. (1998). Analiza slike v podvzorčenem prostoru in slikovne piramide. In: Zajc, B. (Ed.), *Zbornik osme Elektrotehniške in računalniške konference ERK '98*. Slovenska sekcija IEEE, Portorož, Slovenija, pp. 293-296.

Tomaževič, D. (1999). Analiza slike v podvzorčenem prostoru in slikovne piramide. In: Zajc, B. (Ed.), *Zbornik osme Elektrotehniške in računalniške konference ERK '99*. Slovenska sekcija IEEE, Portorož, Slovenija, pp. 467-468.

Tomaževič, D., Cigale, B., Zazula, D., Pernuš, F. (1999). Primerjava petih filtrov pri razgradnji ultrazvočnih posnetkov jajčnikov z aktivnimi konturami. In: Zajc, B. (Ed.), *Zbornik osme Elektrotehniške in računalniške konference ERK '99*. Slovenska sekcija IEEE, Portorož, Slovenija, pp. 265-268.

Dobravec, Š., Tasić, J., Tomaževič, D., Pernuš, F. (2000). Simulacija rentgenske slike iz digitalnega 3D modela pacienta. In: Zajc, B. (Ed.), *Zbornik devete Elektrotehniške in računalniške konference ERK 2000*. IEEE Computer Society, Portorož, Slovenija, pp. 241-244.

Tomaževič, D., Likar, B., Pernuš, F. (2000). A comparison of retrospective shading correction techniques. In: Sanfeliu, A. (Ed.), *15th International Conference on Pattern Recognition*. IEEE Computer Society, Barcelona, Spain, pp. 568-571.

Tomaževič, D., Likar, B., Pernuš, F. (2000). Retrospective shading correction: problem, methods, and evaluation. In: Svoboda, T. (Ed.), *Proceedings of the Czech Pattern Recognition Workshop 2000*. Czech Pattern Recognition Society, Peršlák, Czech Republic, pp. 15-20.

Derganc, J., Likar, B., Bernard, R., Tomaževič, D., Pernuš, F. (2001a). Automated visual inspection of color tablets in pharmaceutical blisters.

- 2001 International Conference on Quality Control by Artificial Vision, Le Creusot, France, pp. 415-420.
- Derganc, J., Likar, B., Bernard, R., Tomažević, D., Pernuš, F. (2001). Machine vision system for quality control of color tablets in pharmaceutical blisters. In: Zajc, B. (Ed.), *Zbornik desete Elektrotehniške in računalniške konference ERK 2001*. Slovenska sekcija IEEE, Portorož, Slovenija, pp. 305-308.
- Tomažević, D., Likar, B., Pernuš, F. (2001). Gradient - based registration of 3D MR and 2D X-ray images. In: Lemke, H.U. (Ed.), *CARS 2001: Computer Assisted Radiology and Surgery*. Elsevier, Berlin, pp. 327-332.
- Tomažević, D., Likar, B., Pernuš, F. (2001). Registration of MR and x-ray images. In: Likar, B. (Ed.), *Computer Vision Winter Workshop 2001*. Slovenian Pattern Recognition Society, pp. 92-103.
- Tomažević, D., Likar, B., Pernuš, F. (2002). "Gold standard" 2D/3D registration of X-ray to CT and MR images., *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2002*. Springer, pp. 461-468.
- Tomažević, D., Likar, B., Pernuš, F. (2002). Rigid 2D/3D registration of intraoperative digital X-ray images and preoperative CT and MR images. In: Sonka, M. (Ed.), *Medical imaging 2002, Image Processing*. SPIE, San Diego, USA, pp. 507-517.
- Tomažević, D., Likar, B., Pernuš, F. (2003). Validation data for 2D/3D registration of medical images. In: Drbohlav, O. (Ed.), *Computer vision - CVWW'03: proceedings of the 8th Computer Vision Winter Workshop*. Czech Pattern Recognition Society, Valtice, Czech Republic, pp. 15-19.
- Vrtovec, T., Tomažević, D., Likar, B., Travnik, L., Pernuš, F. (2003). Statistical shape deformable model of a lumbar vertebra. In: Drbohlav, O. (Ed.), *Computer vision - CVWW'03: proceedings of the 8th Computer Vision Winter Workshop*. Czech Pattern Recognition Society, Valtice, Czech Republic, pp. 91-96.
- Tomažević, D., Likar, B., Pernuš, F. (2004). Multi-feature mutual information. In: Fitzpatrick, J.M. (Ed.), *Medical imaging 2004, Image Processing*. SPIE, San Diego, California, USA, pp. 143-154.
- van de Kraats, E.B., Penney, G.P., Tomažević, D., van Walsum, T., Niessen, W.J. (2004). Standardized evaluation of 2D-3D registration. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2004, Pt 1, Proceedings*. Springer, pp. 574-581.

- Vrtovec, T., Tomažević, D., Likar, B., Pernuš, F. (2004). Automated robust generation of compact 3D statistical shape model. In: Fitzpatrick, J.M. (Ed.), *Medical imaging 2004, Image Processing*. SPIE, San Diego, California, USA, pp. 1312-1323.
- Tomažević, D., Likar, B., Pernuš, F. (2005). Reconstruction-based 3D/2D image registration. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2005, Pt 2*. Springer, pp. 231-238.
- Markelj, P., Tomažević, D., Likar, B., Pernuš, F. (2007). Optimizing bone extraction in MR images for 3D/2D gradient based registration of MR and X-ray images. In: Pluim, J.P. (Ed.), *Medical imaging 2007, Image processing*. SPIE, San Diego, California, USA, pp. 1-12.
- Tomažević, D., Likar, B., Pernuš, F. (2007). 3D/2D image registration: the impact of X-ray views and their number. *Medical Image Computing And Computer-Assisted Intervention - MICCAI 2007*. Springer, pp. 450-457.

# Acknowledgements

I am especially grateful to my mentors, Prof Dr Franjo Pernuš for the overall support and ten years of close and effective cooperation in the Laboratory of Imaging Technologies, University of Ljubljana, and Prof Dr Max A Viergever for broadening my research horizons by kindly inviting me to the Image Sciences Institute, University Medical Center Utrecht.

In addition, I would like to thank Prof Dr Boštjan Likar for numerous suggestions and fruitful scientific cooperation. Many thanks also to Dr Darko Škerl, Dr Everine van de Kraats and Primož Markelj for their contributions to some of the chapters of this thesis. I would also like to thank Prof Dr Aleš Iglič for hosting my postgraduate study at the University of Ljubljana.

Many thanks to my coworkers Dr Marko Bukovec, Dr Tomaž Vrtovec, Dr Darko Škerl, Dr Uroš Vovk, Mario Medved, Miran Büermen, Primož Markelj, Žiga Špiclin and Jaka Katrašnik from the Laboratory of Imaging Technologies and Dr Rok Bernard, Damijan Ipavec, Lidija Zajec, Miha Možina, Gregor Babič and Luka Golob from Sensum, Computer Vision Systems, for establishing a relaxing, friendly and inspirational working atmosphere. I would also like to thank Drs Wiro Niessen, Everine van de Kraats, Graeme Penney, Theo van Walsum and Shirley Baert for fruitful discussions on our weekly meetings of the IGT group. I am also grateful to all the other members of the Image Sciences Institute that made my visit in Utrecht interesting and enjoyable.

I am thankful to Dr Laurent Desbat, Dr Markus Fleute, Raphael Martin, Dr François Estève, and Dr Uroš Vovk for their generous help and support in acquisitions of “gold-standard” images used in Chapters 2, 3, and 4. I would like to thank Dr Ronald L.A.W. Bleys from the Department of Anatomy at the University Medical Center Utrecht for making the spinal segments presented in Chapters 5, 6, 7 and 8 available.

I would like to acknowledge the financial support provided by the Ministry of Higher Education, Science and Technology, Republic of Slovenia, under the grants P2-0232, L2-7381, L2-9758 and Z2-9366, and by the European Commission under grant IST-1999-12338.

Above all, I would like to thank my parents, sister, relatives, and friends for supporting me on my way. And last but not least, sincere thanks to my love Vesna and our children Zarja and Andraž for standing by my side and giving my life another important meaning.

Dejan



## About the author

Dejan Tomaževič was born in Novo mesto, Slovenia, on January 18, 1976. After finishing his secondary education at the Electrotechnical School in Novo mesto in 1994, he registered as a full-time student at the Faculty of Electrical Engineering, University of Ljubljana, Slovenia, where he graduated in 1999 with a diploma thesis on image analysis in scaled and subsampled image space. The same year he started his postgraduate study at the Faculty of Electrical Engineering, University of Ljubljana, as a junior researcher in the Laboratory of Imaging Technologies. His research was focused on image processing and analysis. In 2002 he obtained a Master's degree from the University of Ljubljana with a thesis on 2D-3D registration of medical images. At the same university he received a PhD degree with a thesis on medical image registration in the year 2005. In addition to his research in the Laboratory of Imaging Technologies he spent a six-week period in 1998 at the Image Analysis and Communications Laboratory, Johns Hopkins University, Baltimore, MD, USA. In 2003 he did a three month research fellowship at the Image Sciences Institute, University Medical Center Utrecht, The Netherlands. After having completed his PhD studies at the University of Ljubljana, he joined the company Sensum, Computer Vision Systems, where his current research activities are concerned with automatic visual quality inspection of pharmaceutical products. Since 2007, he has been Assistant Professor at the Faculty of Electrical Engineering, University of Ljubljana.



—

—

—

—