



GPU-BASED RECONSTRUCTION AND VISUALIZATION OF NEEDLES IN X-RAY IMAGES

Matthias Karl Scharrer

*Inst. for Computer Graphics and Vision
Graz University of Technology, Austria*

Master's Thesis

Graz, January 28, 2010

Abstract

This work presents an algorithm to detect rigid straight biopsy needles in multi-view C-arm X-ray images and reconstruction of its tip and orientation in three dimensional space. Several well known computer vision techniques are applied to achieve this goal. The processing pipeline described consists of several stages including a denoising and preprocessing stage with image filtering, computing a norm for better distinction of needle and background and applying the Radon transform to determine the orientation, refinement of the latter by utilizing the random sample consensus paradigm, needle tip detection, reconstruction of the data to three dimensional space with Direct Linear Transform and improving robustness by determining deviation in back projection. Afterwards, the results are visualized by a recently developed application, which facilitates the display of volumetric data together with polygonal geometry and intersection thereof. The processing steps are described in detail, a short overview is given about the surrounding application and finally, the evaluation results of the experiments on real X-ray imaging data is presented and discussed.

Keywords: *medical image processing, multi-view, prostate biopsy, volume visualization*

Zusammenfassung

Diese Arbeit stellt einen Algorithmus zur Detektion von starren, geraden Biopsienadeln in C-Bogen Röntgenbildern aus mehreren Ansichten vor, sowie die Rekonstruktion ihrer Spitze und die Orientierung im dreidimensionalen Raum. Zu diesem Zweck werden mehrere bekannte Techniken aus dem Bereich der Computer Vision angewandt. Der Verarbeitungsablauf, der beschrieben wird, setzt sich aus mehreren Stufen zusammen bestehend aus Glättung und Vorverarbeitung, der Berechnung einer Norm zur besseren Unterscheidung zwischen Nadel und Hintergrund und der Radon Transformation zur Bestimmung der Orientierung, Verfeinerung der letzteren unter Zuhilfenahme des Random-Sample-Consensus-Paradigmas, der Nadelspitzenfindung, der Rekonstruktion der Daten in den dreidimensionalen Raum mittels Direkter linearer Transformation und Verbesserung der Robustheit mittels Feststellung der Abweichung in der Rückprojektion. Danach folgt die Visualisierung der Ergebnisse mittels einer kürzlich entwickelten Anwendung zur Anzeige von volumsbasierten Daten in Verbindung mit polygonaler Geometrie und Verschneidungen der selben. Die einzelnen Verarbeitungsschritte werden detailliert beschrieben, eine kurze Übersicht über das umgebende Anwendungsgerüst wird gegeben und abschließend werden die Versuchsergebnisse der Experimente mit tatsächlichen Röntgenbildern präsentiert und diskutiert.

Schlüsselwörter: *Medizinische Bildverarbeitung, Multi-View, Prostatabiopsie, Volumsvisualisierung*

Deutsche Fassung:
Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008
Genehmigung des Senates am 1.12.2008

EIDESSTÄTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommene Stellen als solche kenntlich gemacht habe.

Graz, am

.....
(Unterschrift)

Englische Fassung:

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

.....
date

.....
(signature)

Acknowledgements

First and foremost, I would like to thank my thesis advisor Dipl.-Ing. Dr.techn. Markus Grabner for making this thesis possible and always providing most valuable advice, supporting me throughout the thesis with patience and yet allowing me to work in my own way.

Many thanks to Univ.-Prof. Dipl.-Ing. Dr.techn. Dieter Schmalstieg. Completion of this thesis would not have been possible without his support, fairness and patience.

Additionally, I want to thank Dipl.-Ing. Bernhard Kainz for providing early support and his very special kind of motivation. Thanks to Dipl.-Ing. Stefan Hauswiesner, who did clarify a few visualization related problems, and Dipl.-Ing. Matthias R  ther, who did provide a nice and comprehensible implementation of multi view reconstruction.

I am heartily thankful to my girlfriend Petra, who did never turn away when confronted with yet another thesis related problem and for endless support during the final stage. I offer my sincerest gratitude to my parents for supporting me throughout all my studies at university and my brothers, who did always encourage and support me.

Lastly, I offer my regards and blessings to all of those who supported me in any way during the completion of this project.

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Motivation | 1 |
| 1.2 | Outline | 2 |
| 2 | Related work | 4 |
| 2.1 | Detection of a needle | 4 |
| 2.2 | Visualization of medical data | 7 |
| 2.3 | The Radon transform for line detection | 11 |
| 2.4 | Random Sample Consensus | 12 |
| 2.5 | The Hessian matrix for ridge detection | 13 |
| 2.6 | Reconstruction of a point to 3D-space | 15 |
| 3 | Processing pipeline | 16 |
| 3.1 | Application Framework | 17 |
| 3.2 | Filtering and masking images | 18 |
| 3.3 | Detecting the line the needle lies on | 18 |
| 3.4 | Detecting the needle's tip | 21 |
| 3.5 | 3D-Reconstruction of the needle | 24 |
| 3.5.1 | Reconstruction of the needle tip in 3D-space | 24 |
| 3.5.2 | Reconstruction of the needle orientation in 3D-space | 24 |
| 3.6 | Improving robustness | 27 |
| 4 | Visualization | 28 |
| 4.1 | Visualization of needle models | 29 |
| 4.2 | Visualization of samples | 31 |
| 5 | Experimental results | 33 |
| 6 | Conclusions | 38 |
| | Bibliography | 43 |

Chapter 1

Introduction

1.1 Motivation

According to analysis of epidemiological data in past decades such as described in [Börgermann et al., 2007], the early detection of a prostate carcinoma plays a role of increasing importance in medical healthcare, as curative treatment is most likely to be successful for tumors at stages T1 and T2, i.e., the tumor is limited to the gland. Preventive screening involves measuring of prostate specific antigen (PSA) values and/or digital rectal examination (DRE).

Since PSA values are not directly related to tumor presence and can always be measured in blood serum of men - it may even increase noticeably by mechanical influence such as riding a bicycle - and DRE results depend on the performing doctor's experience, the only method to affirm the diagnosis is prostate punch biopsy either transrectal or perineal. Transrectal ultrasound (TRUS) guided needle biopsies are current standard in prostate cancer diagnosis. It is common practice to take at least six samples – as shown in Figure 1.1 –, but the number may vary according to prostate and estimated tumor volume, as depicted in [Serfling et al., 2007].

During prostate biopsy a probe head and a core needle are inserted to excise a tissue sample from the prostate for histological examination to find out if cancer or other abnormal cells are present. Since this happens *in vivo* with least possible intrusion, the surgeon has no direct visual opportunity to get feedback on the needle's exact position apart from the sonogram. Due to this, the surgeon can not be certain about the exact position of the sample that was taken. Thus the histologist may only know about the approximate

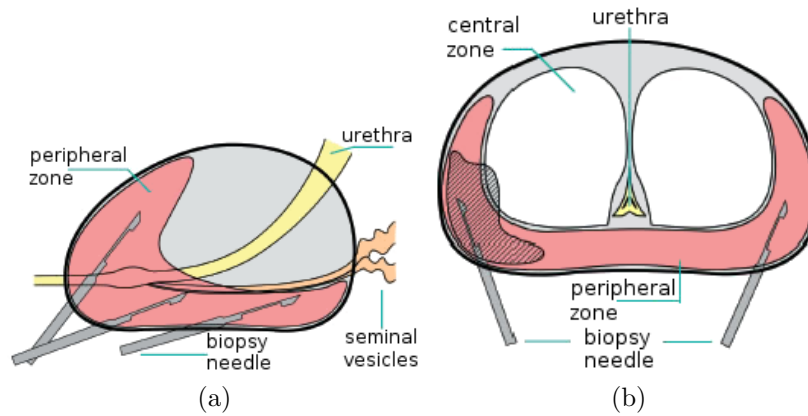


Figure 1.1: Schematic sagittal (a) and axial (b) view of sextant prostate biopsy including possible needle positions. [Pühse and Semjonow, 2009]

location the tissue comes from and positive samples may only lead to roughly estimated tumor location and extents.

To overcome this problem, during the biopsy several X-ray photographs are taken with a C-arm device from different angles so that later on position and orientation of the needle can be estimated very accurately. A good approach to locating cancerous tissue would be to determine a model of the patient's organs with a computed tomography as a preliminary stage to the biopsy, fit this model to the reconstruction from the inter-operative X-ray images [Maintz and Viergever, 1998], reconstruct the needle within this model and visualize the result. This will permit the surgeon to quickly receive feedback of the position of the needle and, if necessary, to adhere to a fixed predefined pattern of taking a number of samples.

Recent work has already covered the registration of the pre-operative CT volume dataset and inter-operative X-ray images [Gross, 2008]. In this work an algorithm to reconstruct the needle's position and orientation in a robust way based upon several views as well as a method for visualization of the result is presented.

1.2 Outline

The remainder of this work is organized as follows. In Section 2 we give a short overview of similar work that has been done considering the task of

INTRODUCTION

detecting a needle and basics of the algorithms and principles of the methods that we apply and software that is used.

The processing pipeline consists of the following stages: detecting the needle ridge in each image, finding its tip, reconstructing this information in 3D-space and rejecting outliers that would lead to misdetection. The stages of the algorithm are covered in Section 3 along with a short description of the framework that this algorithm does fit in. In Section 4 a brief overview of the visualization part is given and in Section 5 the results are discussed.

Finally, in Section 6 the main issues that have arisen during the experiments are pointed out and suggestions for future work are made.

Chapter 2

Related work

In this chapter a short overview of a few papers is given that have similar focus as the presented approach. Additionally, since in this work several well established methods are applied to accomplish the given task the principles of these are described shortly.

2.1 Detection of a needle

A lot of research has been put recently into robotic steerable needle insertion as this seems to be very promising to improve methods in minimal invasive therapy and surgery in near future. Though most of the papers focus on different aspects of this task, needle detection seems to play an increasingly important role.

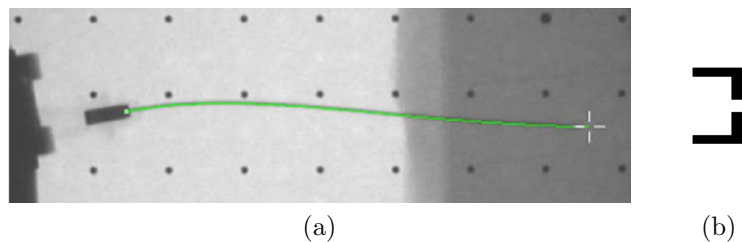


Figure 2.1: Needle detection on an X-ray image by [Glozman and Shoham, 2007]

In [Glozman and Shoham, 2007], an optically controlled flexible needle steering device is presented, where needle detection in X-ray images plays an

RELATED WORK

important role in the feedback loop of the steering control. It is based on detecting a characteristic shape at the basis of the needle with a normalized cross correlaton of a template image (see Figure 2.1b) and the X-ray image (see Figure 2.1a). The needle tip itself is found afterwards by following the low gradient area that starts at the detected basis feature. Knowing the length of the needle, the shape detected can easily be verified. Additionally, the shape of the needle is fitted by a polynomial that smoothes the line. This approach seems sufficiently robust for testing environments, but as the authors state it is subject to error due to noise, occlusion by beads and low contrast of image intensities. Unfortunately, there are no results stated regarding accuracy and error rates.

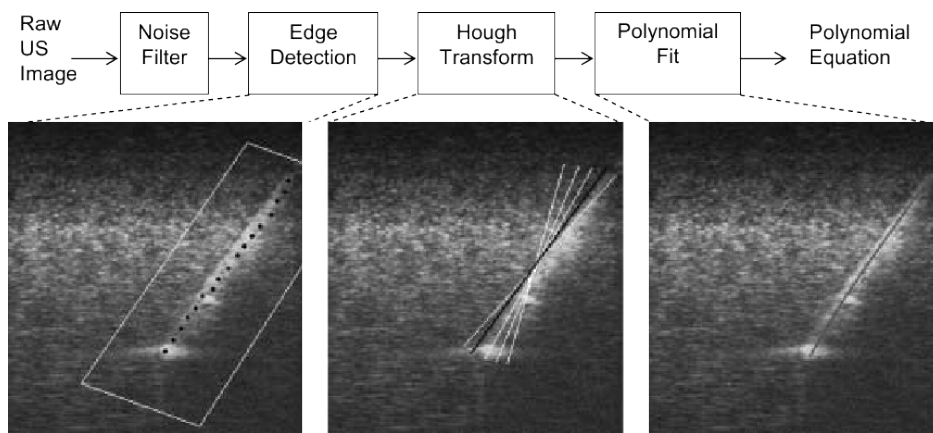


Figure 2.2: Needle segmentation described in [Ebrahimi et al., 2003]

In [Ebrahimi et al., 2003] a steering aid for flexible needles is developed that is intended to operate as an image enhancer to the physician or fully automatically. In contrast to the method described by Glozman and Shoham [Glozman and Shoham, 2007], the image of the needle is taken on an ultrasound basis, where the transducer has to be aligned to the image plane manually once. The needle detection algorithm takes two points on the approximate needle axis manually determined as an initial value and then performs the detection. The detection (see Figure 2.2) is based on noise filtering, gradient based edge detection, Hough transform and polynomial fitting of the needle. It is claimed that the needle detection is very robust, whereas the detection of the needle tip is not very accurate due to the noise in ultrasound images, e.g., speckles and signal drop-out, but there are no further results mentioned.

However, to further improve needle tip detection it is suggested to actively oscillate the tip with some higher frequency. This is not interfering with the

RELATED WORK

steering of the needle and permits the ultra sound device to detect the source of the frequency easier.

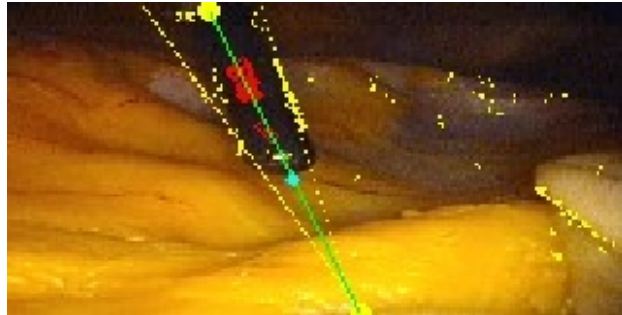


Figure 2.3: Laparoscopic instrument tracking for servoing a robotic camera holder described in [Voros et al., 2006]

The small yellow dots represent response to edge detection, the green line depicts the symmetry axis found and the cyan dot marks the needle tip found.

In [Voros et al., 2006] a new approach to laparoscopic instrument tracking for controlling a robotic camera holder is presented. Although it is based on visual feedback and its purpose is not to reconstruct the needle but rather give a rough estimation of the position of the tip in the image taken by an endoscopic camera, the principles applied are similar to the approach of this work.

Since the camera is calibrated and mounted on a calibrated camera holder, the initialisation is based on stereo vision techniques to reconstruct the insertion point of the instrument in three dimensional space. This is done by moving the camera to two positions and manually labeling the point. Based on this manual initialization, the point can be back projected to the image plane and further processing can be done automatically.

The detection of the instrument is based on four steps. First, all possible edges are computed by applying a simple gradient method constrained by the manually labeled insertion point P which is not further mentioned. Due to the high contrast between the instrument and the surrounding tissue – as can be seen in Figure 2.3 – this largely yields points at the edge of the instrument. Utilizing these points, the symmetry axis of the instrument is searched for by applying a specialized Hough transform centered at P taking only points attached to a satisfyingly oriented edge into account. After the symmetry axis is found, the instrument contour is computed by taking into account symmetry and by applying linear regression. Finally, thresholding the image profile along the previously found symmetry axis into two classes and

RELATED WORK

determining the class associated to the instrument leads to the instrument tip. This last step is done either by determining the longest zone for the first image or by taking the color of the instrument previously detected during the first evaluation into account.

Although the accuracy of the tip detection needed for servoing a camera robot is stated to be not very high with ≈ 11 pixels, the results of the method described above seem sufficiently accurate, as the error between the detected and a manually labeled tip is smaller than this in 87% of the tested frames. Detecting the axis of the instrument rarely fails – only one frame was reported where this was the case –, but detecting the needle seems to be subject to many error sources, such as specular reflections and low contrast along the axis.

2.2 Visualization of medical data

Visualization of medical data is a very important area of research, since it aids physicians in understanding, interpreting and communicating their data. One of the most challenging tasks in this field is the visualization of volumetric data, i.e., the projection of one or possibly several 3D datasets onto a 2D image plane. Usually the dataset is given as scalar values or vectors defined at the nodes on a regular grid. Typically, sources of volume datasets – as far as the medical domain is concerned – are three dimensional imaging techniques, e.g., Computer Tomography (CT) and Magnetic Resonance Imaging (MRI), which differ by what characteristics are being scanned. Visualizing a volume typically involves displaying significant parts such as skin, bones, organs or tumourous tissue. There exists a variety of methods for this task that can be categorized as either Surface-Fitting or Direct Volume Rendering [Elvins, 1992].

Surface-Fitting (SF) methods such as the popular Marching Cubes introduced in [Lorensen and Cline, 1987] and its further improved versions, e.g., [Cline et al., 1988, Nielson and Hamann, 1991, Lewiner et al., 2003], aim to generate an object representation based on planar geometric primitives such as polygons that mark the target object's boundary. Geometry generated by SF algorithms usually can be displayed fast and efficiently as there exist numerous means by which the rendering of large numbers of primitives is sped up and changes in viewing and lighting parameters are quickly passed on to the observer. It is very well suited to display tissue interfaces, but it is subject to problems due to incorrect handling of small features and branches. Additionally, low-contrast objects may not be presented very well and mod-

RELATED WORK

ifications of the parameters of SF methods require an update of the entire representation by rerunning the method.

Direct Volume Rendering (DVR) methods differ from SF methods in that they do not create an intermediate representation based on geometric primitives but map volume elements (*voxels*) directly onto the image plane [Engel et al., 2006]. There exist several techniques that can be categorized in image-space (also referred to as *backward mapping*) and object-space order (*forward mapping*) methods. With ray casting being an image order method, for every pixel in the image plane a color and opacity is computed by following an imaginary ray that is cast through the data starting at the projection center and sampling voxel values for color and opacity along its way. This straightforward approach results in both optically correctness and very high quality [Levoy, 1988], but it is computationally expensive as the data structure has to be traversed for every ray. Other DVR methods such as Splatting [Westover, 1990] are object order methods that assemble voxel values in object space and project the result back to image space. With Splatting, every voxel is convoluted with a three dimensional reconstruction filter, the so-called *reconstruction kernel* and accumulated onto the image plane. An example of a very efficient combination of both paradigms is Shear-Warp [Lacrout and Levoy, 1994]. Here, the object space is sheared in order to make viewing rays perpendicular to an intermediate image plane, i.e., all voxel values of a ray stack upon another. The application of DVR methods involves the specification of color and opacity transfer functions that map the values of the source volume to meaningful values, e.g., mapping the bones-specific intensity of a CT to opaque white, blood to a less opaque red, skin to transparent beige and so on.

However, perception of volumetric medical data on a 2D screen requires additional depth cues, such as perspective, shading and the ability to investigate the object from different angles. Since perspective results in a distorted image and still images are of rather illustrative use, interactivity in terms of manipulating view point or data representation is one of the key features of visualization systems. Since DVR methods involve big computational effort, the emergence of inexpensive programmable graphics processing units (GPUs) and efficiently programming the same resulted in very good frame-rates such as described in [Krüger and Westermann, 2003]. An illustrative description of DVR methods is given in [Pfister, 2004], more detailed information can be found in [Engel et al., 2006] especially w.r.t. hardware acceleration.

The benefits of visualizing of volumetric data can be further improved by

RELATED WORK

displaying additional information, e.g., combining CT and MRI data with arbitrary geometric objects such as models of surgical instruments. Therefore, efficient methods need to be applied to tackle this more comprehensive task.

Recent work presented in [Brecheisen et al., 2008] came up with a GPU accelerated ray casting solution with a *depth peeling* algorithm similar to the implementation described in [Everitt, 2001] that supports the task of correctly visualizing intersecting translucent objects by depth sorting polygonal surfaces. But as it is tightly coupled to the conventional shader model there exists a performance bottleneck in communication between the depth peeling shader and the ray casting shader by storing data in global memory. Additionally, each depth layer requires a further ray casting pass resulting in an increase in runtime as the number of intersecting objects increases.

Another very recent publication describes the utilization of the *Compute Unified Device Architecture* (CUDA) [NVIDIA, 2008] for implementing GPU accelerated ray casting [Kainz et al., 2009]. They implemented a volume rendering pipeline based on polygon tiling wholly in software, which bypasses the flaws inherent to the system’s architecture mentioned above. Geometry is only processed once without the need of recurring alternation of depth peeling and ray casting passes and data is passed on utilizing the GPU’s fast shared memory, the access speed of which clearly surpasses global memory access many times over.

The system described by [Kainz et al., 2009] supports the key features described above, including the visualization of any selection of the following:

- multiple intersecting grid based volume datasets at diverse resolution, spacing and orientation with arbitrary two-manifold boundaries or clipping geometry and individual transfer function
- multiple possibly concave or translucent polygonal objects intersecting any other object
- Constructive Solid Geometry (CSG) operations to classify intersecting regions and partitioning volumes by polyhedral boundaries – including application of another transfer function to each partition

Figure 2.4 shows an example scene and an illustrative representation of the underlying scene graph. Volume data is stored in a volume atlas which is a single large 3D texture, whereas cylinder and sphere geometry and volume bounds are kept in common geometry storage. The difference between volume

RELATED WORK

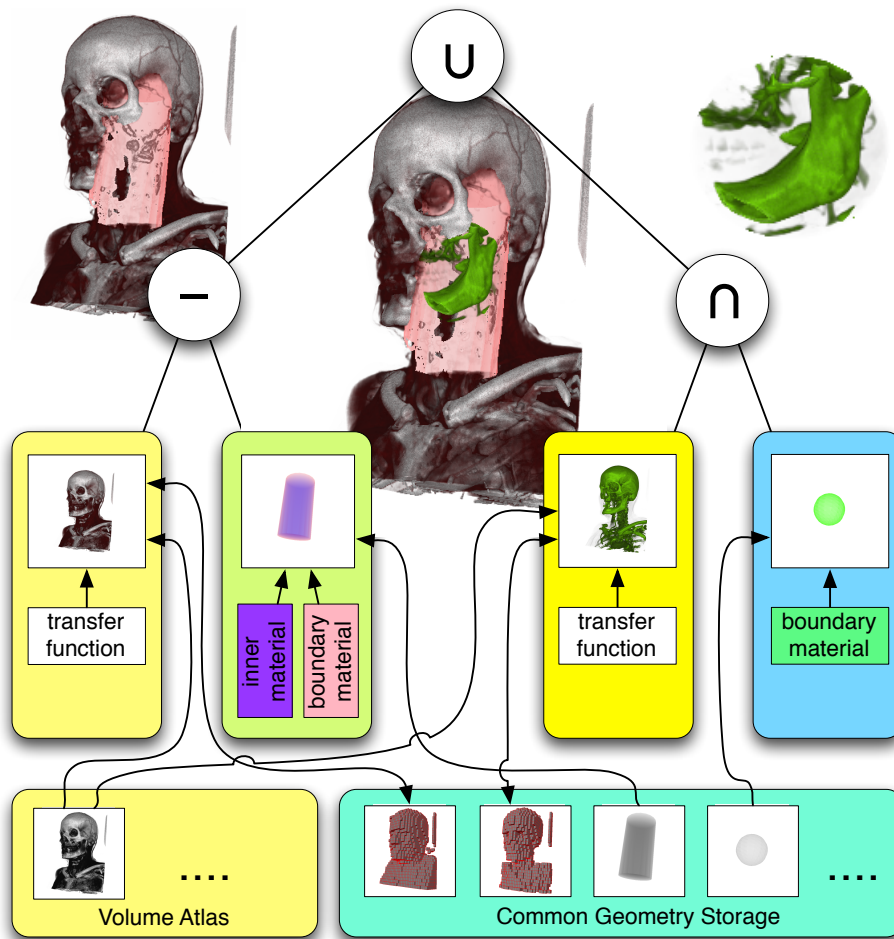


Figure 2.4: A scene graph example of the volume renderer described in [Kainz et al., 2009]. It shows the nodes and their relationship of a scene graph containing Boolean difference $-$, intersection \cap and union \cup CSG operations, polyhedral objects and different transfer functions on two instances of the same volume.

RELATED WORK

and cylinder and the intersection of volume and sphere are computed and different transfer functions are applied.

2.3 The Radon transform for line detection

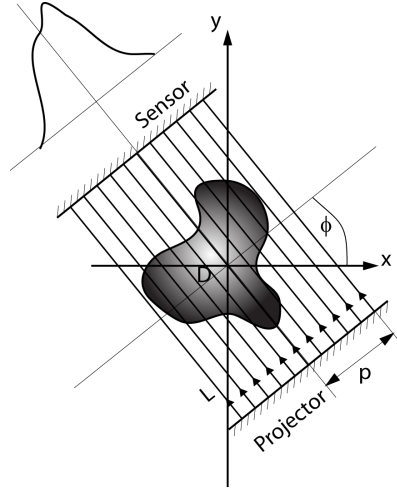


Figure 2.5: The beamer-sensor system the Radon transform can be thought of.

The Radon transform in general is defined as to integrate functions f on n -dimensional Euclidean space \mathbb{R}^n over all $n - 1$ -dimensional hyperplanes [Deans, 1983]. In our case, f , that can be considered a two dimensional distribution, is projected onto all lines by determining all line integrals of f on some domain D on \mathbb{R}^2 .

$$\hat{f} = \int_L f(x, y) ds \quad (2.1)$$

This can also be thought of rotating a parallel beam projector around a problem domain and simultaneously registering the incoming beams at the opposite site of the domain, as illustrated in Figure 2.5. A line L can be denoted by the normal form

$$p = x \cos \phi + y \sin \phi \quad (2.2)$$

A straight line becomes a single point in this transformed space, so the problem of detecting a line can be further reduced to detecting a peak in two dimensional parameter space, which can be found easily.

2.4 Random Sample Consensus

For processing experimental data it is often very useful to fit a suitable model to it by estimating the free parameters of the model and continue to work with the latter. A basic attempt would be to fit a model to the data in its entirety, e.g., by minimizing the sum of squared differences. Since this approach is in general not robust to outliers that may arise due to noisy data, it is very common to apply the Random Sample Consensus (RANSAC) paradigm first introduced in [Fischler and Bolles, 1981]. The principle of the approach is described in Algorithm 2.4.1. The application of it is based on

Algorithm 2.4.1 The RANSAC paradigm

Require:

- a data model \mathcal{M} with n free parameters
- a set of data points P , such that $|P| \geq n$
- a threshold t
- an error tolerance ϵ
- a maximum number of iterations N
- (randomly) select a subset S of n data points from P
- while** numberofiteration $< N$ **do**
 - instantiate a model M based on S
 - use the instantiated model M to determine the subset S^*
 - the consensus set - of points in P that are within some error tolerance ϵ of M .
 - if** $|S^*| > t$ **then**
 - use S^* to compute a new model M^* .
 - return** the newly instantiated model M^* .
 - end if**
 - randomly select a new subset S
- end while**
- return** best fitting M^* or failure in case of no suitable consensus set was found

three parameters: A threshold number t , an error tolerance ϵ and a maximum number of iterations N . All of these depend on the model and have to be determined ahead of execution.

The threshold number t is chosen such that the expected number of inliers in the data set P is met. It usually relates to a percentage of $|P|$. Although it is stated to terminate after a consensus set consisting of more than t elements was found returning M^* – marked by the dotted underline

RELATED WORK

in Algorithm 2.4.1 –, it would as well be possible to limit the execution only by N and finally return the model M^* fitting best.

In the majority of the cases the error tolerance ϵ – also referred to as *distance threshold* – is specified according empirical study such that with a probability α it is an inlier.

The number of iterations N is chosen, such that with a probability p at least one set S is found that contains no outliers. Given β that denotes the probability that N iterations in sequence result in failure and α^n the probability that a given set S results in a valid model, by rearranging it follows that:

$$N = \frac{\log \beta}{\log (1 - \alpha^n)} \quad (2.3)$$

2.5 The Hessian matrix for ridge detection

The ridge detector we used is based on the Hessian matrix defined in (2.4), as it is supposed to be very responsive at edges and blobs [Mikolajczyk et al., 2005]. The Hessian matrix is defined as:

$$\mathbf{H}(x, y) = (\nabla^T \cdot \nabla) I = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix}, \quad (2.4)$$

where ∇ denotes the operator $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right)$, I_{xx} and I_{yy} are the second derivatives of the image intensity I at position (x, y) and $I_{xy} = I_{yx}$ is the mixed derivative in both directions. After performing the eigenvalue decomposition of the Hessian matrix

$$\lambda_{1,2} = \frac{1}{2} \left[(\mathbf{H}_{11} + \mathbf{H}_{22}) \pm \sqrt{4\mathbf{H}_{12}\mathbf{H}_{21} + (\mathbf{H}_{11} - \mathbf{H}_{22})^2} \right], \quad (2.5)$$

where $|\lambda_1| > |\lambda_2|$ are the eigenvalues of $\mathbf{H}(x, y)$ and

$$\mathbf{H}(x, y) v_i = \lambda_i v_i, \quad i \in \{1, 2\}, \quad (2.6)$$

where $v_{1,2}$ are the eigenvectors of $\mathbf{H}(x, y)$. Decomposition into eigenvalues and eigenvectors give an indication about local curvature at a specific point. Simply put, the eigenvector associated to the strongest eigenvalue tends to point into the direction of the strongest change in data whereas the other vector points into the direction of least change if there is a big difference between λ_1 and λ_2 . Eigenvalues themselves provide information about the

RELATED WORK

strength of the change in the local neighborhood. Considering ridge detection, it can be stated, that if $\nabla I v_1 = 0$, the point belongs to a ridge ($\lambda < 0$) or a valley ($\lambda > 0$) [López et al., 1999].

In [Frangi et al., 1998] a new approach to vessel enhancement in medical images is introduced. Although not intended for straight ridge or line enhancement, the principles behind the algorithm apply very well to these simplified constraints while providing a very intuitive interpretation. It is based on the local information gathered by the Hessian matrix for every single point of I in 3D space, but can be transferred analogously to 2D space. By geometrically interpreting a generalized form of (2.6), it can be stated that a circular neighborhood centered at (x, y) will be mapped by \mathbf{H} to an ellipse, where eigenvectors define the axes and eigenvalues define the extents associated. For describing a perfectly linear structure, the eigenvalues should approximate $|\lambda_2| \approx 0$ and $|\lambda_1| \gg |\lambda_2|$ – the sign indicates the polarity. The measure defined by

$$\mathcal{R}_{\mathcal{B}} = \frac{|\lambda_2|}{|\lambda_1|} \quad (2.7)$$

directly relates to the area of the *second order ellipse*. With increasing $|\lambda_2|$, the ellipse becomes more and more circular or blob-like, whereas with increasing $|\lambda_1|$ the ellipse collapses towards a line. If both values approximate zero, i.e., when the local neighborhood is homogeneous, the measure rises towards infinity. This relation however does only take geometry information into account. To further distinguish between bright foreground and dark background or vice versa, the Frobenius matrix norm is used by [Frangi et al., 1998], which indicates the magnitude of the derivatives, which will be low for background values, where only small contrast is given.

$$\mathcal{S} = \|\mathbf{H}\|_F = \sqrt{\lambda_1^2 + \lambda_2^2} \quad (2.8)$$

Combining (2.7) and (2.8) gives the vesselness function for two-dimensional images,

$$\nu(s) = \begin{cases} 0 & \text{for } \lambda_1 > 0 \\ \exp\left(-\frac{\mathcal{R}_{\mathcal{B}}^2}{2\beta^2}\right) \left(1 - \exp\left(-\frac{\mathcal{S}^2}{2\gamma^2}\right)\right) & \end{cases} \quad (2.9)$$

where β and γ are thresholds that may be used to adjust the sensitivity to $\mathcal{R}_{\mathcal{B}}$ and \mathcal{S} , and the parameter s denotes the scale at which the function is evaluated. By evaluating the norm at different scales, it can be ensured that structure with approximately the size of the scale will be found. Eventually, by maximum projection of ν a structure enhanced image is obtained that

RELATED WORK

can be further processed:

$$\nu(x, y) = \max_{s_{min} \leq s \leq s_{max}} \nu(s), \quad (2.10)$$

where s_{min} and s_{max} should be chosen according to the expected size of the structure to be found, e.g., the width of lines.

2.6 Reconstruction of a point to 3D-space

A very simple linear triangulation method would be to combine the known coordinates of the points in the image planes $\mathbf{x}_i = \mathbf{P}_i \mathbf{X}_i$, which can also be written as

$$\mathbf{x}_i \times (\mathbf{P}_i \mathbf{X}_i) = 0, \quad (2.11)$$

where \mathbf{X}_i are the world coordinates of \mathbf{x}_i and $\mathbf{P}_i = (\mathbf{p}_{i,1}, \mathbf{p}_{i,2}, \mathbf{p}_{i,3})^T$ are the respective projection matrices. Reconstruction can be done by bringing the above into a form of homogeneous equations $\mathbf{A}\mathbf{X} = 0$, where \mathbf{A} is a composition of two Equations (2.11) for each image:

$$\begin{pmatrix} u_1 \mathbf{p}_{1,3}^T & - & \mathbf{p}_{1,1}^T \\ v_1 \mathbf{p}_{1,3}^T & - & \mathbf{p}_{1,2}^T \\ u_2 \mathbf{p}_{2,3}^T & - & \mathbf{p}_{2,1}^T \\ v_2 \mathbf{p}_{2,3}^T & - & \mathbf{p}_{2,2}^T \\ & & \vdots \\ u_n \mathbf{p}_{n,3}^T & - & \mathbf{p}_{n,1}^T \\ v_n \mathbf{p}_{n,3}^T & - & \mathbf{p}_{n,2}^T \end{pmatrix} \mathbf{X} = 0, \quad (2.12)$$

where u_i and v_i are the elements of \mathbf{x}_i .

However, it is not possible to reconstruct the world coordinates exactly from several views, but rather it is necessary to estimate them in an optimization function that minimizes some cost function, as the rays that can be backprojected from the image through their respective camera centers are skew and the system of Equations (2.12) is overdetermined.

As stated in [Hartley and Zisserman, 2004] this kind of reconstructional problems can be treated as finding a least-squares solution to Equation (2.12), i.e. finding \mathbf{X} that minimizes $\mathbf{A}\mathbf{X}$ subject to $\|\mathbf{X}\| = 1$

Chapter 3

Processing pipeline

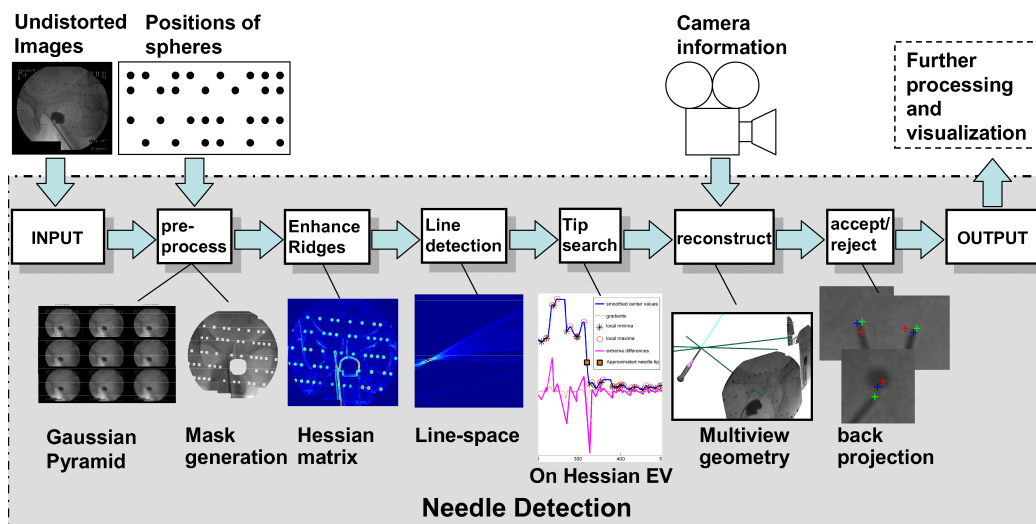


Figure 3.1: A schematic overview of the processing pipeline with its stages and its interaction with the entire application.

The chosen approach can be partitioned into several smaller stages, which can be treated as individual modules as depicted in Figure 3.1. There are a few components developed by others that perform calculations in advance. A short overview about them is given in Section 3.1. The implemented stages are detailed in their respective subsections. In Section 3.2 preprocessing of images is described including noise reduction and masking of the images. The approach how the line the needle is aligned to is found is presented in Section 3.3, followed by a description of how the tip is found along this line

in Section 3.4. Combining this data to find its representation in world coordinates is outlined in Section 3.5. Finally, in Section 3.6 a short description is given of the idea of how to filter out wrong results.

3.1 Application Framework

As already stated, this module is part of a more complex application supposed to assist the surgeon during biopsy of the prostate, i.e., exciting tissue for histological examination, the ablation of cancerous tissue or the implantation of low radiating beads that would destroy the former. The knowledge of the exact location of where the needle intersected with the prostate gland after taking a sample instead of a coarse description might prove very useful to further treatment planning.

In [Kainz et al., 2008] a new approach is considered to take several images with a calibrated mobile C-arm, each of which with maximal mechanical offset – $\pm 45^\circ$ due to limitations in operational range during prostate biopsies – but minimal femoral influence. For the purpose of reconstructing camera positions and orientations, a special target with small globes attached in a specific pattern visible to X-ray imaging has been designed by [Kainz et al., 2008] and used for reconstruction of the camera orientation and position relative to the patient. Ahead of the operation, calibration of the X-ray device is performed by taking a sufficiently large number of images of the defined structure from different view points and computing the intrinsic camera parameters thereafter, i.e., the principal point coordinates, focal length and distortion coefficients. During the biopsy, the target is placed underneath the patient and images are taken. After undistortion and filtering, the image positions of the globes are retrieved and used to estimate the extrinsic camera parameters, i.e., rotation and translation in three dimensional space.

In [Gross, 2008] a fast implementation of a rigid body intensity based 2D/3D registration was introduced, that by utilizing the power of the GPU computes the parameters of a Euclidean transform that maps a preoperative CT volume from model space to the operating room space defined by X-ray images taken during intervention. This is done by generating a digitally reconstructed radiograph (DRR), measuring the similarity to the X-ray images and optimizing the result iteratively until convergence. Eventually, the knowledge of this rigid transform facilitates intraoperative visualization of patient data acquired ahead of operation such as CT or MR images combined with intervention-based data such as the location of surgical instruments.

Visualization and interaction methods should finally allow the surgeon to easily access the desired information. The Coin3D¹ based visualization framework described in Section 2.2 is used for visualization of the resulting data.

3.2 Filtering and masking images

As shown in Figure 3.1, there is work done a priori [Kainz et al., 2008], the results of which ease the task of detecting the needle in the image, since the cameras are already calibrated and the images are undistorted. This plays a very important role, as X-ray images are subject to a very strong pincushion distortion. Without undistortion the needle located in the image would be bent depending on the location within the image. Furthermore, the positions of the small globes are computed during camera pose estimation in an earlier step, so this information can be used to mask out the small beads.

Due to the fact that all images during a biopsy are taken by the same machine, the geometry of the viewport showing X-ray data along with meta data of the patient, which is returned by the image acquisition device, always stays the same. This implies that a constant mask can be used to mask out the additional information that is written to the image, such as the name of the patient and others, so that the focus of all further processing lies on the actual image area, which is circular in this case.

However, prior to processing the data, a median filter is applied to the images, since X-ray image sources tend to be subject to so called *salt and pepper noise*, i.e., randomly occurring extreme value speckles, and to compute second order derivatives later on which are very sensitive to speckles in images, rank ordered filtering is a very popular approach to remove such.

3.3 Detecting the line the needle lies on

The core needle which is used to extract the tissue samples is a hollow needle which contains the tissue sample after a successful shot. It can be considered an elongated rigid object with a small outer diameter of $1.27mm$ in case of a *18-gauge* needle, which is commonly used for prostate biopsies. When viewing the image it can be considered a thick line, so the problem of detecting the needle orientation in the image is reduced to detecting a line in the image.

¹A free scene graph library under GPL - <http://www.coin3d.org>

PROCESSING PIPELINE

In [Deans, 1983] it is stated, that the Radon transform is strongly related to the Hough transform, which is commonly used to detect lines in images, as described in Section 2.3. But applying the Radon transform on the pre-processed grey value image will most probably not end up in a single point in parameter space, but rather in very unpredictable results, since the background will also contribute with a sine function for every point in the image. Therefore, the image will not be used directly for the Radon transform but it will be transformed to a different representation to enhance discrimination of lines against inhomogeneous background.

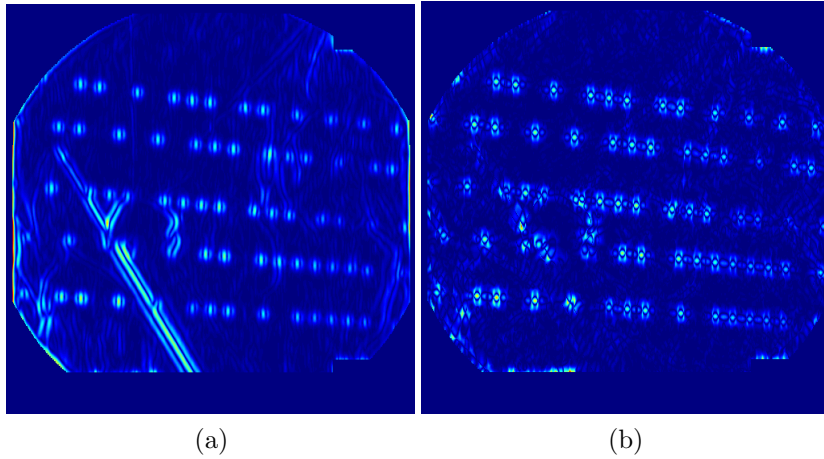


Figure 3.2: Eigenvalues $|\lambda_1|$ (a) and $|\lambda_2|$ (b) of the computed tensorfield at scale $\sigma = 2.5$. Values are coded by a *jet* color map, i.e., a blue (lowest value) to red (highest) spectrum.

As shown in Figure 3.2, the needle is a very prominent feature in the image of the stronger eigenvalues – but so are the small globes and the probehead. By applying Frangi’s norm ν at different scales on the eigenvalues of the Hessian transform and integrating the results as described in Section 2.5 a high distinction between the needle and all other objects within the image can be obtained. What remains is the needle and some very low value ridges of the patient’s bones and the probehead, as shown in Figure 3.3. Different scales are essential, since the *elongated rigid object* includes image features of variable width. The needle guidance – showing mainly high (red) response in Figure 3.3 – is obviously much thicker than the needle itself and the thin feature slightly below the tip – the notch where the tissue specimen is cut into – gives high response at a lower scale. Thresholding the objectness norm ν described in Section 2.5 to suppress all values below 30% of the range of ν and filtering out the response of the small globes by applying

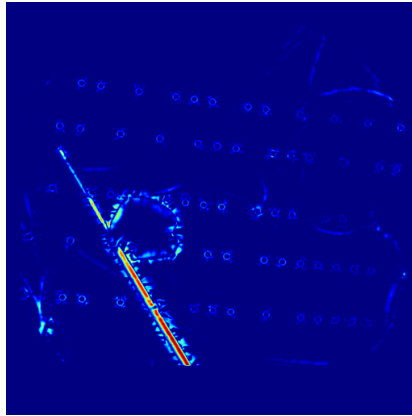


Figure 3.3: Result after applying Frangi's norm

the mask generated during preprocessing leaves an almost clutter free image. This is used along with the conditions for ridge detection defined above as input to the Radon transform. The result (see Figure 3.4) is then searched for local peaks after non-maximum suppression. The strongest of which is most likely to be the ridge of the needle. After the initial orientation of the needle is found, the image is rotated to align the needle orientation to the horizontal axis and cropped to small height centered at the needle axis. Afterwards, every column of this cropped image is examined for local maxima and presence of masked out area. If the latter is detected, the whole column is discarded, whereas if only one local maximum is detected along the profile, the position is considered a needle ridge support point and saved to a list for further use.

After having assembled all support points in the list, the line best fitting to these points is determined by applying the RANSAC paradigm described in Section 2.4. By choosing the error tolerance ϵ to not exceed one pixel and setting the maximum number of iterations N to 17 which represents the number of samples required to find at least one set that contains no outliers with a probability of 99% while assuming that half the data can be considered outliers [Hartley and Zisserman, 2004] and setting the threshold number t to approximately a tenth of the smaller image dimension, most probably the best consensus set is found given the support points list described above.

To take the possibility into account that the above step results in a slightly differently oriented ridge caused by, for example, the initial window for examining the local ridge maxima only partly contained values of the needle, the steps above describing the assembly and the fitting of the ridge support points are repeated until convergence, i.e., the line fitting results were nearly

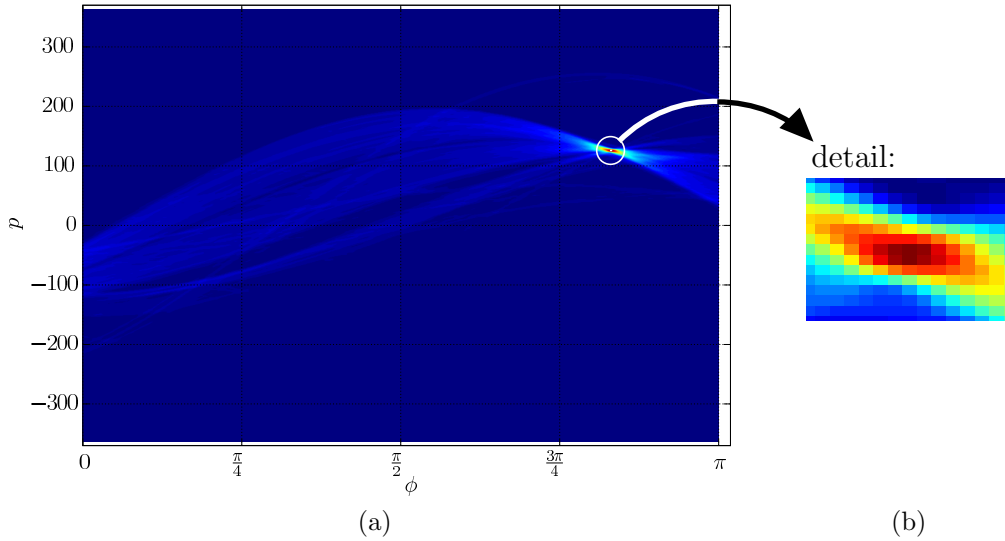


Figure 3.4: (a) shows the Radon transform of an enhanced image. (b) shows the local neighbourhood of the maximum peak. Values are color coded as in Figure 3.2

the same, or a maximum iteration number is reached.

3.4 Detecting the needle's tip

After finding the line that is most probably aligned to the needle, the position of the tip of the needle is searched for. All support points that were identified during the previous step are taken into account. Additional support points are determined by examining the profile of ridge enhanced values perpendicular to the ridge. Within a small window the maximum value is looked up and its position is inserted to the ridge support point list. If no maximum can be found due to homogeneity of the image, the support point under test is discarded. Of course, the region of interest is restricted by the iris mask defined earlier, so that the algorithm would not respond to the values outside. By contrast with Section 3.3, the mask defined for filtering out the globes is not applied, as the needle tip could have wrongly be classified as globe there or the tip could be in the vicinity of such a globe. Eventually, by defining the upmost support point as reference, a one dimensional cross section of the needle ridge $f(x)$ is obtained (see blue curve in Figure 3.5) depending on the distance x from the reference point.

PROCESSING PIPELINE

Since after the tip there should be a significant step among the values registered, the needle's tip recognition is based on these values. It is now supposed that the step, that marks the end of the needle provides a very steep slope with a very low value at the end. Since we would not want to threshold these values as the variations of the data are too high and occasional peaks due to high response at the edges of the globes used for pose reconstruction, we look for the steepest change in the mean level of the assembled ridge responses. This is done by fitting a unit step function to every point in the data and computing the sum of squared differences as an error measure. The function to fit is defined as

$$g(x) = \alpha + \beta u(x - \gamma), \quad (3.1)$$

where α and β denote the linear parameters, i.e., offset and height of the step function, and $u(x)$ denotes the unit step function defined as

$$u(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases}. \quad (3.2)$$

The parameters α and β of $g(x)$ can be derived by finding the least squares fit of a simple matrix \mathbf{A} consisting of a column of ones – representing the constant offset – and a column of zeros and ones – representing the step function – to the values of $f(x)$ by utilizing the *singular value decomposition* (SVD). Best fitting parameters α and β are estimated ahead of the final error estimation for every possible value of γ to obtain the maximum parameter range and further on α and β are set to their minimum and maximum, respectively.

Now the error measure to be minimized is defined as

$$e(\gamma) = \sum_x (f(x) - g(x))^2 \quad (3.3)$$

subject to the non-linear parameter γ only. By evaluation of $e(\gamma)$ for every possible γ that is restricted to the number of distances x from the reference point, the global minimum of the error can be found and its associated γ can be translated to the final image coordinates of the tip.

$$\mathbf{x} = \begin{pmatrix} u \\ v \end{pmatrix} \quad (3.4)$$

PROCESSING PIPELINE

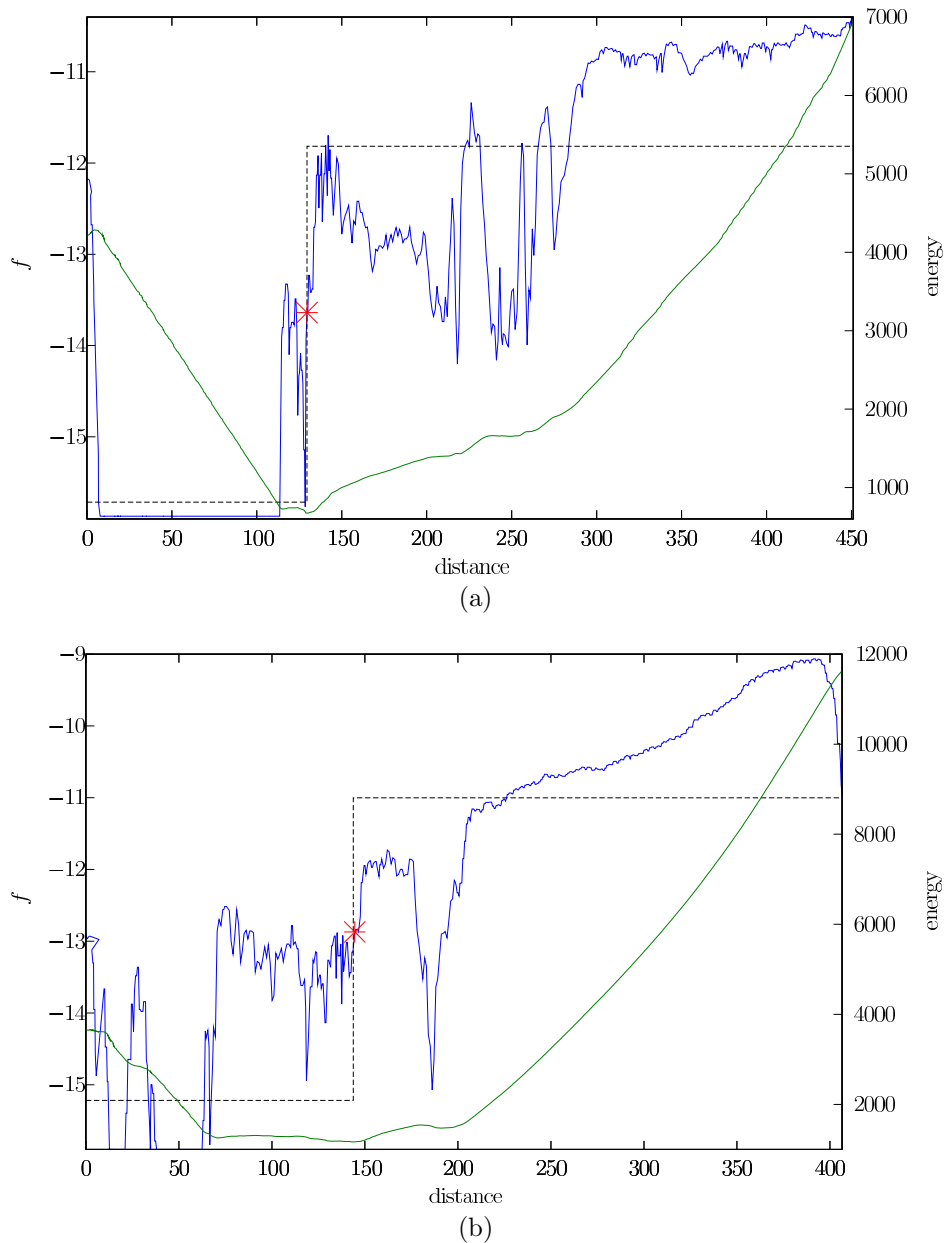


Figure 3.5: Needle tip detection based on assembled ridge-enhanced values f (blue) at the needle ridge, fitted step function (black) and error function (green).

(a) shows the basis of Figure 5.1a, where the smaller lobe to the left of the step indicates the presence of a small globe close to the tip in the image.

(b) shows the basis of Figure 5.1g, where detection of the correct tip at $distance \approx 70$ has failed – possibly due to offset being too high. This is the worst result obtained.

3.5 3D-Reconstruction of the needle

After having performed the previous steps to the images it is now possible to reconstruct the needle (i.e., estimating the position and orientation of the needle in three dimensional space). As already mentioned earlier, this work is part of a framework and a lot of work is done and information is gathered before needle detection is started. The camera calibration matrix and the external camera matrix (comprising the rotation matrix and the translation vector) are available from [Kainz et al., 2008] and used for reconstruction. Of course, this implies the projection matrices that can be computed out of the internal and external camera matrices are exact, but as their mean re-projection error implies very high accuracy, it seems sufficient to consider the cause of errors to be the coordinates obtained by the previous step described in Section 3.4.

3.5.1 Reconstruction of the needle tip in 3D-space

The solution to this problem is acquired applying the Direct Linear Transform (DLT) – see Algorithm 3.5.1 for the way a single point is reconstructed. [Hartley and Zisserman, 2004] present a procedure for the general case of reconstruction of a set of point correspondences in each image, but this algorithm describes a slightly modified form, since there only is a single point in each image to be reconstructed.

3.5.2 Reconstruction of the needle orientation in 3D-space

In Section 3.3, a fairly robust solution to detecting the line that supports the needle is presented. In this Section, another reconstruction method is described that deals with how the orientation of the needle can be recovered.

For the reconstruction of a line two points need to be known on that line in every image. If these point pairs happen to correspond to every other pair of the other images, they can be used to easily recover the line by reconstruction of each point with the method described above and compute the line vector with basic vector algebra.

After the previous step, one point in every image and its reconstructed 3D-point are given so far, but finding a second point could turn out difficult. By observing that the guidance of the needle is thicker than the needle,

PROCESSING PIPELINE

Algorithm 3.5.1 Direct Linear Transform triangulation for one point

Require:

image coordinates $\{\mathbf{x}_1 = (u_1 v_1)^T, \dots, \mathbf{x}_n = (u_n v_n)^T\}$,
projection matrices $\{P_1, \dots, P_n\}$

for $i = 1, \dots, n$ **do**

- compute the point normalizing transform matrices

$$\mathbf{T}_i = \begin{pmatrix} 1 & 0 & -u_i \\ 0 & 1 & -v_i \\ 0 & 0 & 1 \end{pmatrix}$$

- transform and normalize the projection matrices

$$\tilde{\mathbf{P}}_i = \mathbf{T}\mathbf{P}_i, \hat{\mathbf{P}}_i = \frac{\tilde{\mathbf{P}}_i}{\sqrt{\text{tr}(\tilde{\mathbf{P}}_i^T \tilde{\mathbf{P}}_i)}}$$

end for

- stack each normalized projection matrix' first and second row vector in a $2n \times 4$ -dimensional matrix

$$\mathbf{A} = -(\hat{p}_{1,1}\hat{p}_{1,2} \dots \hat{p}_{n,1}\hat{p}_{n,2})^T$$

- compute the singular value decomposition of A

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$$

return with $\mathbf{D} = \text{diag}(d_1 \dots d_n)$ and $d_i \geq d_{i+1}$

$$\mathbf{X} = \mathbf{V}(0 \dots 1)^T$$

PROCESSING PIPELINE

detecting the obviously characteristic position of where the needle leaves the guidance leads to several other point correspondences that would be sufficient for the task.

However, this approach seems far too vulnerable to misdetections, and there exists an easier way, which is less prone to error due to introduction of additional detection steps. Since we only need to reconstruct the orientation of the needle, as the tip position has already been computed, the problem can be stated as to find the intersection of planes that result in a line, which is always the case if the planes are not parallel or coincide.

Basically, if the planes are given in Hessian normal form,

$$\mathbf{n} \cdot \mathbf{x} = -p, \quad (3.5)$$

where \mathbf{n} is the unit normal vector and p is a constant defining the distance of the plane from origin, for the case of two intersecting planes, the line of intersection has to be normal to either plane normal, so that the direction of this line can be computed by the cross product

$$\mathbf{l} = \mathbf{n}_1 \times \mathbf{n}_2, \quad (3.6)$$

As the planes defined by the lines in three or more images are not likely to form a sheaf of planes, i.e., a set of planes having the same line in common, due to deficiency of numerical precision and measurement errors in image based line detection and camera pose reconstruction, a general approach can be formulated as

$$\mathbf{A} = \begin{pmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \\ \vdots \\ \mathbf{n}_m^T \end{pmatrix}, \quad (3.7)$$

where the *null space* of \mathbf{A} gives the optimal orientation of the intersection line in a least-squares sense. Applying these relations to the problem of finding the needle orientation leads to the following procedure, again utilizing the SVD to compute the null space of \mathbf{A} :

- Extend the calibration matrix to a homogeneous 4×4 -matrix and recompute the projection matrix - which will now be invertible - for each image.
- Compute the camera center for each image.

PROCESSING PIPELINE

- Compute the 3D-coordinates for the two points of the needle's line crossing the border of their respective image with the inverted projection matrix. An arbitrary z -value can be chosen.
- Compute the normals of the planes that are spanned by these two position vectors and the camera projection center.
- Stack the normals into a new matrix \mathbf{A} and compute the SVD $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$ of this matrix.
- The column vector \mathbf{v} of \mathbf{V} that corresponds to the smallest eigenvalue of matrix \mathbf{D} will be the best fit to the orientation of intersection of any two planes w.r.t. geometrical error.

3.6 Improving robustness

After reconstructing the three dimensional point of the needle tip, it is easy to project this point back onto the images so that it is possible to compute the distance between the initially detected needle tip in an image and the back projected needle tip. If only images with this distance below a previously specified upper limit are passed on and groups of images, where this value is outside on at least one image, are rejected this makes the procedure very robust against outliers. If needle detection fails at any stage, it does not seem very likely that it fails on all three images in a similar way, i.e., accidentally detecting an entirely different correspondence than the needle tip. Furthermore, this did not happen during the experiments at any time, so this assumption is considered safe.

Chapter 4

Visualization

For visualization purposes, the renderer described in Section 2.2 was used. To further ease the usage of the latter, a small user interface was created that encapsulates the back-end implementation of the needle detection by providing convenient means of selecting the images used for needle reconstruction, processing the data of the back-end and creation of an Open Inventor compliant output file. There are three use cases that are considered and described in their subsequent sections. In Section 4.1, a very simple visualization example is given, showing the CT volume and six geometric models representing the biopsy position and orientation as reconstructed. In Section 4.2 a more sophisticated approach is shown, where the volume and its intersection with each respective biopsy core location is shown. The original CT volume's dimensions shown in the following section was scanned with a resolution of $512 \times 512 \times 278$ voxels with $0.6875mm \times 0.6875mm \times 1mm$ spacing in every direction, respectively. This was resampled for visualization to $256 \times 256 \times 256$ voxels to obtain framerates of approximately 3 – 5 frames per second.

Unfortunately, visualizing the prostate gland did not work very well due to several reasons. This task would have required either segmentation of CT data – which was far beyond the scope of this work –, MR or contrast enhanced images – both were not provided – to apply suitable transfer functions to.

4.1 Visualization of needle models

The first visualization type is supposed to give qualitative feedback of the results that were obtained by applying the processing pipeline described in Section 3,

$$\mathbf{X} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \quad (4.1)$$

where \mathbf{X} is the reconstructed point and \mathbf{v} is the reconstructed orientation of the line aligned to the needle in 3D space. As the model of the biopsy needle is aligned to the x -Axis with its tip pointing to $\mathbf{0}$, the rotation of the model given by Euler axis \mathbf{u} and angle θ can easily be obtained by computing the cross product and the dot product

$$\mathbf{u} = \mathbf{e}_x \times \mathbf{v}, \quad \theta = \cos^{-1}(\mathbf{e}_x \cdot \mathbf{v}) \quad (4.2)$$

By applying Rodrigues' rotation formula a rotation matrix – also known as *direct cosine matrix* – representing the rotation in matrix form can be obtained:

$$\mathbf{R} = \mathbf{I} + \mathbf{u} \sin(\theta) + \mathbf{u} \cdot \mathbf{u}^T (1 - \cos(\theta)), \quad (4.3)$$

where \mathbf{I} denotes the identity matrix. Incorporating the translation of \mathbf{X} and the rotation of \mathbf{R} into a homogeneous transformation matrix \mathbf{A}

$$\mathbf{A} = \begin{pmatrix} \mathbf{R} & \mathbf{X} \\ \mathbf{0} & 1 \end{pmatrix}, \quad (4.4)$$

facilitates further processing in that any combination with further transformation matrices can be constituted by simple multiplication.

The results of 2D/3D registration as detailed in [Gross, 2008] include

$$\begin{aligned} \mathbf{T} &= \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}, & \mathbf{R}_x &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta_x) & \sin(\theta_x) & 0 \\ 0 & -\sin(\theta_x) & \cos(\theta_x) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ \mathbf{R}_y &= \begin{pmatrix} \cos(\theta_y) & 0 & -\sin(\theta_y) & 0 \\ 0 & 1 & 0 & 0 \\ \sin(\theta_y) & 0 & \cos(\theta_y) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, & \mathbf{R}_z &= \begin{pmatrix} \cos(\theta_z) & \sin(\theta_z) & 0 & 0 \\ -\sin(\theta_z) & \cos(\theta_z) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \\ \mathbf{R} &= \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x, \end{aligned} \quad (4.5)$$

VISUALIZATION

5.8 fps

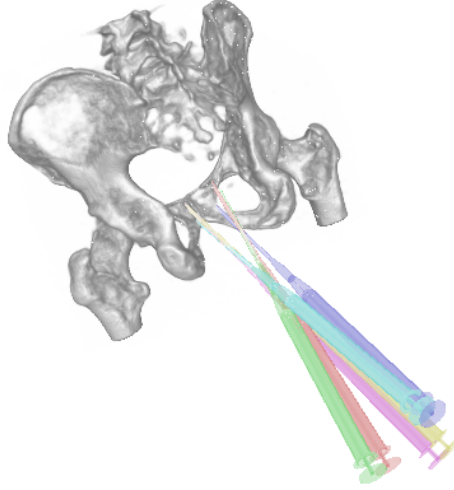


Figure 4.1: Visualization of reconstructed position and orientation of six biopsy samples (Patient A).

The image shows the patient's volume – located at the pelvis – smoothed by a $5 \times 5 \times 5$ -median filter with a transfer function restricted to the bone's values applied to and needle models located at the biopsy samples' positions.

where \mathbf{T} refers to translation and \mathbf{R} denotes the rotations around x -, y - and z -axis of the volume's coordinate system from model space to operating room space defined by the camera pose estimation in [Kainz et al., 2008]. Incorporating a translation matrix \mathbf{C} to center the volume ahead of the rotation and its inverse \mathbf{C}^{-1} prior to translation is commonly referred to as *model transform*, by which any homogeneous position vector $\mathbf{p}_{model} \in \mathbb{R}^4$ in model coordinates is mapped to another homogeneous position vector $\mathbf{p}_{world} \in \mathbb{R}^4$ in world coordinates:

$$\mathbf{p}_{world} = \mathbf{TC}^{-1}\mathbf{RCp}_{model} = \mathbf{Mp}_{model}, \quad (4.6)$$

where \mathbf{M} denotes the model transform matrix.

By taking the inverse of \mathbf{M} , the world coordinate system is mapped to the model space of the CT volume. Combining the previously obtained transform of the needle model \mathbf{A} with \mathbf{M}^{-1} gives the transform that maps the needle model into the coordinate space of the volume. Figure 4.1 shows the volume obtained by a preceding CT scan and six biopsy needle models that are

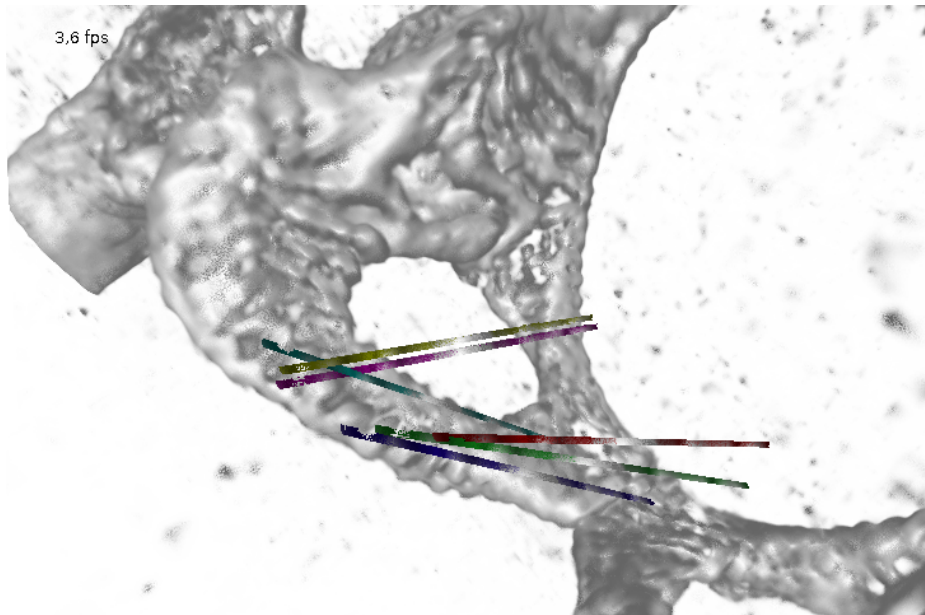


Figure 4.2: Visualization of intersection with CT volume (Patient A).

This image shows the same volume as in Figure 4.1 – without smoothing – and biopsy core data located at the biopsy samples' positions. Speckles are caused by high value speckle noise in the dataset.

oriented and positioned as implied by the results of needle detection. As already mentioned in Section 2.2, every geometrical object can have its own material properties rendered appropriately. As the reconstruction step does not give hints about the rotation of the needle around its axis, all models have their longest extent of the bevel located at the anterior side, i.e., it is pointing towards the head.

4.2 Visualization of samples

In Figure 4.2 another use case is shown, where the volume is intersected with small cylinders, that represent the biopsy cores. Again, the same geometrical transformations apply as described in the previous section. It shows the CT volume of patient A's pelvis with focus on bone tissue. Each displayed sample core is an intersection of a properly translated and oriented cylinder of $2mm$ radius with the CT data. With the simple approach of loading the volume for each object, the total memory usage of the graphics card would exceed

VISUALIZATION

any reasonable dimension, but as the renderer is capable of reusing the same volume with respect to different region bounding geometry, this was not an issue.

Whereas the models described in Section 4.1 each have their own materials, in this case every core object as well as the global volume have their own transfer function assigned to. This enables the operator to interactively change the representation of the data on inspection according to personal preference and optimal output of suspicious tissue properties. The transfer functions of the cylinders in Figure 4.2 are set to cover the range of values in the vicinity of the prostate, which are very low due to the soft tissue properties w.r.t. image acquisition type. In addition, the lowest values were assigned similar colors as in Figure 4.1 to improve visual discernability of the needle cores and facilitate comparison between the figures, whereas higher values gradually map to brighter color with white highlighting the highest values.

Although the ventral section of the cores displayed, i.e., the rightmost part in Figure 4.2, approximately reflect the extracted biopsy core, it is worth noting that the intersected data displayed further away from the tip does not relate to any extracted specimen, as the soft tissue close to the biopsy gun is subject to mechanical deformation, which is not simulated in the preoperative volume data.

Chapter 5

Experimental results

The data available to test consisted of 90 images of three patients with shots from three different views. To evaluate the errors made the tip of the needle has been marked by hand on every image. The three available coordinates are shown in Figure 5.1. Images (a)–(c) of this Figure show the best result obtained (Patient B), (d)–(f) show a slightly worse result (Patient A) and (g)–(i) show the worst result (Patient C) of the dataset w.r.t. the mean squared error between detected tip and labeled tip. For evaluating the accuracy it is interesting to look at the distances between these coordinates in the images, as the ground truth is not available in 3D-coordinates. Following the initial needle tip detection, the distance can be computed between the detected coordinate $\hat{\mathbf{x}}$ and the labeled coordinate \mathbf{x} . To evaluate the results of the reconstructional step back projection of the reconstructed needle tip from 3D-space coordinate \mathbf{X} to the respective image coordinates $\hat{\hat{\mathbf{x}}}$ is performed and the distance to the labeled coordinate is computed again. As already mentioned, the distance between $\hat{\mathbf{x}}$ and $\hat{\hat{\mathbf{x}}}$ is crucial to the robustness of the algorithm. It is used as an estimate of the confidence in the reconstructed data.

In Table 5.1 the underlying values are already restricted to results that passed outlier rejection. Although the maximum deviation of the reconstructed tip is very high, 68.35% of the reconstructed data lies within five pixels of the needle tip, which is roughly the thickness of the needle, which is most probably accurate enough for application. In Figure 5.2 the results of the needle tip detection with respect to the back projection classification of the needle tip error is shown. Although the worst results are rejected on basis of the distance between the back projected needle tip and the automatically detected tip, such as the result shown in Figures 5.1g and 3.5b, there

EXPERIMENTAL RESULTS

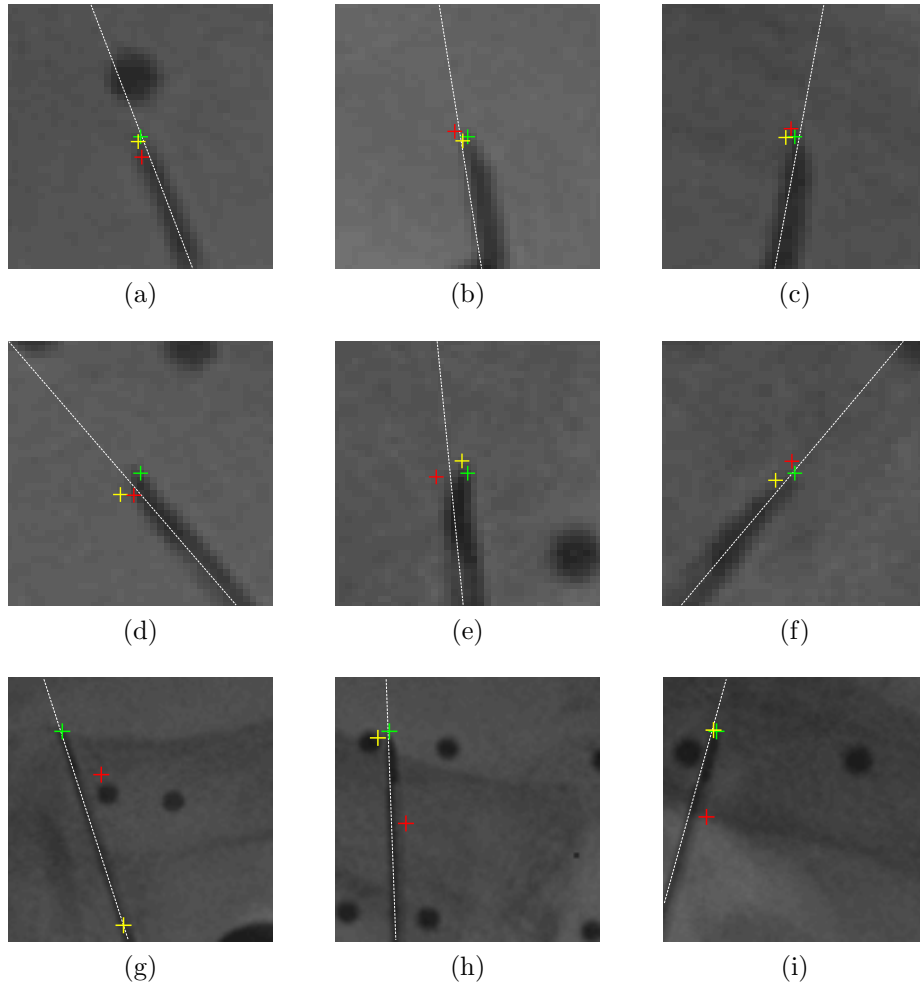


Figure 5.1: The needle tip labeled, detected and back projected from 3D reconstruction. The tip positions are colored *green* for manually labeled, *yellow* for image based detection and *red* for back projected, the detected needle ridge is represented by the thin white dashed line. Images (a)–(f) show the local pixel intensities centered at the manually labeled tip with a radius of 20 pixels, whereas (g)–(i) show a 100×100 pixels wide area to cover the result.

EXPERIMENTAL RESULTS

| | $d(\mathbf{x}, \hat{\mathbf{x}})$ | $d(\mathbf{x}, \hat{\hat{\mathbf{x}}})$ | $d(\hat{\mathbf{x}}, \hat{\hat{\mathbf{x}}})$ |
|----------------------------|-----------------------------------|---|---|
| Patient A – 24 images | | | |
| minimum | 0.480 | 0.230 | 0.191 |
| maximum | 7.289 | 14.538 | 7.931 |
| mean | 2.476 | 3.907 | 3.448 |
| std deviation | 1.734 | 2.750 | 1.777 |
| Patient B – 33 images | | | |
| minimum | 0.618 | 1.160 | 0.845 |
| maximum | 10.717 | 8.150 | 7.068 |
| mean | 2.450 | 3.456 | 2.983 |
| std deviation | 2.082 | 1.935 | 1.800 |
| Patient C – 33 images | | | |
| minimum | 0.832 | 0.613 | 0.778 |
| maximum | 23.295 | 21.180 | 8.181 |
| mean | 4.531 | 6.407 | 3.838 |
| std deviation | 4.073 | 4.051 | 2.022 |
| Entire dataset – 90 images | | | |
| minimum | 0.480 | 0.613 | 0.191 |
| maximum | 23.295 | 21.180 | 8.181 |
| mean | 3.248 | 4.702 | 3.437 |
| std deviation | 3.073 | 3.280 | 1.893 |

Table 5.1: Distances between hand labeled needle tip and image based detection in the first column and the back projection of the reconstructed needle tip in the second. The third column shows the difference between first detection and back projection. All values are measured in pixels.

are many *good* results discarded indicated by the red bars at $d(\mathbf{x}, \hat{\mathbf{x}}) \approx 2$ due to their dependency on each other. If one of the detected tips completely fails, as indicated in Figure 5.1 in the bottom row, all the other results are discarded as well, since the wrongly detected tip results in a far off back projection.

On the other hand, there are a few *bad* results that are not discarded due to the reconstruction resulting in a back projection closer to the detected tips than the range specified. This behaviour may arise by reason of the fitting behaviour of the singular value decomposition. In this situation this can be thought of a kind of weighted mean of the points in the images. While the detected points of similar quality *drag* the back projection towards their position, the outlier counteracts this effect in proportion to the amount of

EXPERIMENTAL RESULTS

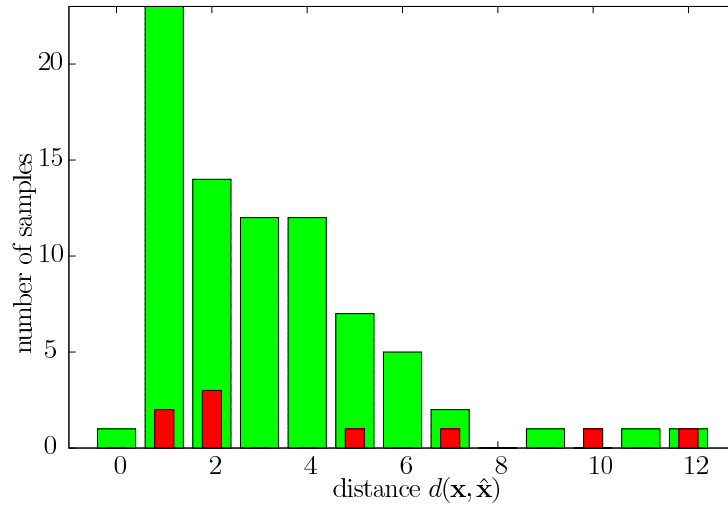


Figure 5.2: Distribution of the absolute detection error.

Green bars represent the number of positive detection results – Red bars represent the number of negative detection results, w.r.t. back projection error. Values higher than the limits are clamped to the maximum value.

its offset compared to the other points. This results in the back projection being closer to the mass of the detected points and a single bad detection has to be very far off to be correctly classified by the reprojection step.

In Figure 5.3 all results of the needle tip detection step are visualized at their position with respect to the manually labeled tip oriented to the upright position. It is clearly visible that the main proportion of the correctly classified tips can be found to the right of the manually labeled tip. This is assumed to be due to the bevel of the needle tip which cannot be clearly distinguished from the background tissue of the X-ray image during the course of manually labeling the needle tips.

As shown in Figure 5.4, it is possible, that the ultra sonic device overlaps the needle in a way that makes needle detection and thus needle reconstruction impossible. At the moment this is not handled separately and simply rejected.

EXPERIMENTAL RESULTS

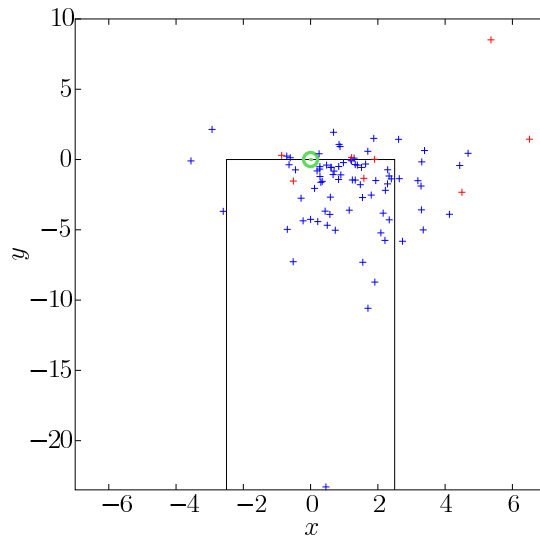


Figure 5.3: Detected needle tips compared to manually labeled ones.

The blue crosses show the detected needle tips with positive back projection, the red crosses show the detected needle tips that were negatively classified. The green circle depicts the manually labelled needle tip position. The black lines indicate the needle boundaries in the vicinity of the tip. Only the result of Figure 5.1g was omitted for scale reasons.

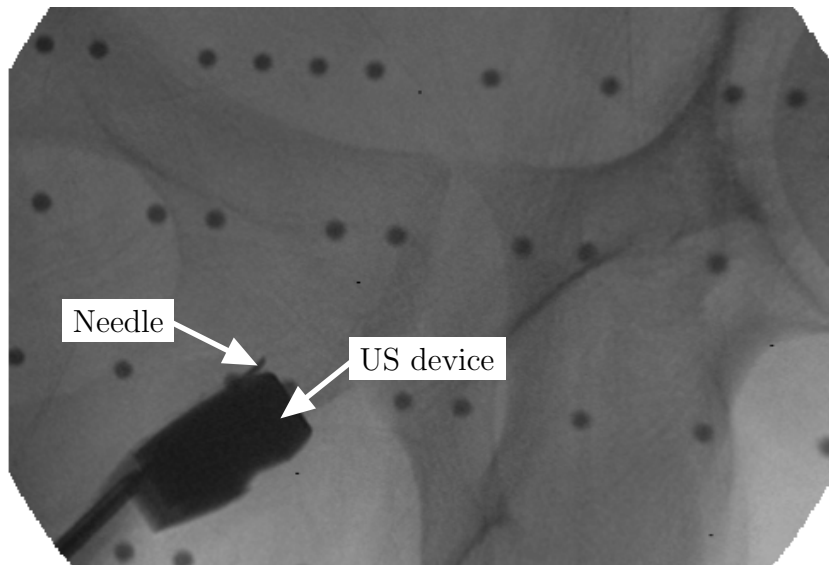


Figure 5.4: Needle occluded by the ultrasonic device

Chapter 6

Conclusions

This work presents the design and implementation of the processing pipeline of a needle detection and reconstruction system intended to be used as one component in a framework for computer aided biopsy and visualization of the results incorporating preoperative volumetric imaging data. It takes several established techniques to accomplish this task, such as the rank order filtering to preprocess X-ray data, the Radon transform to detect line features in the image and the robust fitting of line parameters to support points by applying the RANSAC paradigm. The parts involved to reconstruct the needle's tip and orientation are well known nowadays and a lot of research has already been put into it. Along with the results it can be said that the methods are very reliable and easy to use. Applying Frangi's norm to enhance the images for ridge and tip detection turned out to be very intuitive as well as accurate for the given task.

Although experimental results show acceptable accuracy in reconstruction and robustness to noise and clutter in the image there are still many problems to solve, such as further improvements to the detection of the needle tip so that the reconstruction can be done with higher accuracy or detecting special situations such as the occlusion of the needle by the ultrasonic probe head.

Different optimization strategies may lead to a more accurate tip detection, incorporating additional data from the ultrasound device which is already available during the biopsy or tracking the biopsy gun may further improve the detection and reconstruction by providing additional initialization values, other views and pose parameters. Applying different paradigms of detecting the needle, such as shape based detection may lead to very good results, as well, although this may result in much longer evaluation times.

However, the system is supposed to be running in a closed loop, which

CONCLUSIONS

causes the results to be available very shortly after taking the X-rays, so it would be possible to take another set of X-rays if the results are not convincing due to a lack of accuracy because of the needle tip not being found during the tip detection.

Since the result may suffer from inaccuracies caused by previous stages, such as camera calibration and pose reconstruction errors, another approach would be to reconstruct the needle directly from ultrasound devices and register a three dimensional sonogram to the preoperative data.

Visualization of the biopsy results offers even more room for improvements. Clearly, visualizing the histology result of positive samples along with a probability map of the tumor localization according to a tumor location probability atlas would be the best result to be received. Furthermore, this kind of visualization could include a multimodal visualization of the prostate, its context and the reconstruction results. Based upon this, treatment planning could be integrated into visualization by modeling and incorporating treatment results, such as hyperthermal temperature distribution or dose density of brachytherapy, which is a form of internal radiotherapy.

List of Figures

| | | |
|-----|--|----|
| 1.1 | Schematic saggital and axial view of sextant prostate biopsy [Pühse and Semjonow, 2009] | 2 |
| 2.1 | Needle detection in [Glozman and Shoham, 2007] | 4 |
| 2.2 | Needle segmentation described in [Ebrahimi et al., 2003] | 5 |
| 2.3 | Laparoscopic instrument tracking for servoing a robotic camera holder described in [Voros et al., 2006] | 6 |
| 2.4 | A scene graph example of the volume renderer described in [Kainz et al., 2009] | 10 |
| 2.5 | The beamer-sensor principle of the Radon transform | 11 |
| 3.1 | A schematic overview of the processing pipeline with its stages and its interaction with the entire application. | 16 |
| 3.2 | Absolute eigenvalues of the Hessian at scale $\sigma = 2.5$ | 19 |
| 3.3 | Result after applying Frangi’s norm | 20 |
| 3.4 | Radon transform of an enhanced image | 21 |
| 3.5 | Needle tip detection | 23 |
| 4.1 | Visualization of reconstructed position and orientation of six biopsy sampels (Patient A) | 30 |
| 4.2 | Visualization of intersection with CT volume (Patient A) | 31 |
| 5.1 | Needle tip labeled, detected and back projected | 34 |
| 5.2 | Detection error distribution | 36 |
| 5.3 | Detected needle tips compared to manually labeled | 37 |

LIST OF FIGURES

| | |
|--|----|
| 5.4 Needle occluded by the ultrasonic device | 37 |
|--|----|

List of Tables

| | | |
|-----|--------------------------------|----|
| 5.1 | Experimental results | 35 |
|-----|--------------------------------|----|

Bibliography

- [Brecheisen et al., 2008] Brecheisen, R., Platel, B., Vilanova, A., and ter Haar Romenij, B. (2008). Flexible GPU-based multi-volume ray-casting. In *Proceedings of Vision, Modelling and Visualization*, pages 1–6. 9
- [Börgermann et al., 2007] Börgermann, C., Luboldt, H., Schröder, F., and Fornara, P. (2007). Prostatakarzinom - Screening und Früherkennung. In Rübber, H., editor, *Uroonkologie*, chapter Prostatakarzinom, pages 406–413. Springer Medizin Verlag, 4 edition. 1
- [Cline et al., 1988] Cline, H., Lorensen, W., and Ludke, S. (1988). Two algorithms for the three-dimensional reconstruction of tomograms. *Medical Physics*, 15(3):320–327. 7
- [Deans, 1983] Deans, S. (1983). *The Radon Transform and Some of its Applications*. Wiley. 11, 19
- [Ebrahimi et al., 2003] Ebrahimi, R., Okazawa, S., Rohling, R., and Salcudean, S. (2003). Hand-Held Steerable Needle Device. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2879/2003:223–230. 5, 40
- [Elvins, 1992] Elvins, T. T. (1992). A survey of algorithms for volume visualization. *SIGGRAPH Comput. Graph.*, 26(3):194–201. 7
- [Engel et al., 2006] Engel, K., Hadwiger, M., Kniss, J., Rezk-Salama, C., and Weiskopf, D. (2006). *Real-time Volume Graphics*. A. K. Peters. 8
- [Everitt, 2001] Everitt, C. (2001). Interactive order-independent transparency. Tech. rep., NVIDIA. 9
- [Fischler and Bolles, 1981] Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395. 12

BIBLIOGRAPHY

- [Frangi et al., 1998] Frangi, A., Niessen, W., Vincken, K., and Viergever, M. (1998). Multiscale vessel enhancement filtering. pages 130–137. 14
- [Glozman and Shoham, 2007] Glozman, D. and Shoham, M. (2007). Image-Guided Robotic Flexible Needle Steering. *IEEE Transactions on Robotics*, 23(3):459–467. 4, 5, 40
- [Gross, 2008] Gross, T. (2008). GPU-based 2D/3D registration of X-Ray and CT data. Master’s thesis, Graz University of Technology, Institute for Computer Graphics and Vision, Graz, Austria. 2, 17, 29
- [Hartley and Zisserman, 2004] Hartley, R. and Zisserman, A. (2004). *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition. 15, 20, 24
- [Kainz et al., 2009] Kainz, B., Grabner, M., Bornik, A., Hauswiesner, S., Muehl, J., and Schmalstieg, D. (2009). Ray casting of multiple volumetric datasets with polyhedral boundaries on manycore gpus. *ACM Trans. Graph.*, 28(5):1–9. 9, 10, 40
- [Kainz et al., 2008] Kainz, B., Grabner, M., and R  ther, M. (2008). Fast marker based c-arm pose estimation. In *MICCAI ’08: Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II*, pages 652–659, Berlin, Heidelberg. Springer-Verlag. 17, 18, 24, 30
- [Kr  ger and Westermann, 2003] Kr  ger, J. and Westermann, R. (2003). Acceleration techniques for GPU-based volume rendering. In *Proceedings of IEEE Visualization*, pages 287–292, Washington, DC, USA. IEEE Computer Society. 8
- [Lacroute and Levoy, 1994] Lacroute, P. and Levoy, M. (1994). Fast volume rendering using a shear-warp factorization of the viewing transformation. In *SIGGRAPH ’94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*, pages 451–458, New York, NY, USA. ACM. 8
- [Levoy, 1988] Levoy, M. (1988). Display of surfaces from volume data. *IEEE Computer Graphics and Applications*, 8(3):29–37. 8
- [Lewiner et al., 2003] Lewiner, T., Lopes, H., Vieira, A., and Tavares, G. (2003). Efficient implementation of marching cubes’ cases with topological guarantees. *Journal of Graphics Tools*, 8:2003. 7

BIBLIOGRAPHY

- [López et al., 1999] López, A., Lumbreras, F., Serrat, J., and Villanueva, J. (1999). Evaluation of Methods for Ridge and Valley Detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(4):327–335. 14
- [Lorensen and Cline, 1987] Lorensen, W. and Cline, H. (1987). Marching cubes: A high resolution 3d surface construction algorithm. *SIGGRAPH Comput. Graph.*, 21(4):163–169. 7
- [Maintz and Viergever, 1998] Maintz, J. A. and Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36. 2
- [Mikolajczyk et al., 2005] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., and Gool, L. V. (2005). A comparison of affine region detectors. *International Journal of Computer Vision*, 65(1–2):43–72. 13
- [Nielson and Hamann, 1991] Nielson, G. and Hamann, B. (1991). The asymptotic decider: resolving the ambiguity in marching cubes. In *VIS '91: Proceedings of the 2nd conference on Visualization '91*, pages 83–91, Los Alamitos, CA, USA. IEEE Computer Society Press. 7
- [NVIDIA, 2008] NVIDIA (2008). *NVIDIA CUDA Programming Guide 2.0*. NVIDIA Corporation. 9
- [Pfister, 2004] Pfister, H. (2004). Moderne Volumenvisualisierung. *Information Technology*, 46(3):117–122. 8
- [Pühse and Semjonow, 2009] Pühse, G. and Semjonow, A. (2009). *Die Prostatastanzbiopsie - Gewebeentnahme aus der Prostata*. Prostatazentrum am Universitätsklinikum Münster, D-48129 Münster, 2nd edition. 2, 40
- [Serfling et al., 2007] Serfling, R., Shulman, M., Thompson, G., Xiao, Z., Benaim, E., Roehrborn, C., and Rittmaster, R. (2007). Quantifying the Impact of Prostate Volumes, Number of Biopsy Cores and 5 α -Reductase Inhibitor Therapy on the Probability of Prostate Cancer Detection Using Mathematical Modeling. *Journal of Urology*, 177:2352–2356. 1
- [Voros et al., 2006] Voros, S., Long, J.-A., and Cinquin, P. (2006). Automatic localization of laparoscopic instruments for the visual servoing of an endoscopic camera holder. pages 535–542. 6, 40

BIBLIOGRAPHY

- [Westover, 1990] Westover, L. (1990). Footprint evaluation for volume rendering. In *SIGGRAPH '90: Proceedings of the 17th annual conference on Computer graphics and interactive techniques*, pages 367–376, New York, NY, USA. ACM. 8