Diss. ETH No.

Deformable Area-based Template Matching with Application to Low Contrast Imagery

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY ZURICH

for the degree of
Doctor of Technical Sciences

presented by
MARTIN BERGER
Dipl. El. Ing. ETH
ETH - Swiss Federal Institute of Technology
born the 21st of April 1969
citizen of Zürich

Prof. Dr. Guido Gerig, examiner
Prof. Dr. Luc van Gool, co-examiner
Dr. Urs Meier, co-examiner

April 1999

Abstract

The exact positioning of patients during radiotherapy is essential for high precision treatment. Before each session, the patient must be accurately placed within the treatment device. The analysis of megavoltage X-ray images, *portal images*, can help to control this patient positioning. The particular problem of electronic portal imaging devices however, is that they provide imagery with extremely low contrast. Thus, common feature extraction schemes, such as simple edge detection, do not produce reliable results. Moreover, it is a very challenging task to design any robust feature extractor for such images, since a method must be found that reliably transforms the image information into a more useful form, whilst remaining robust to a large degree of image quality.

To circumvent the step of feature extraction, the area-based method of least squares template matching (LSM) has been chosen. LSM is an iterative and area-based fitting method especially suitable for attaining very high precision, or for processing low-contrast, noisy, and blurred imagery. This thesis reviews the mathematical formalism of LSM and presents a framework for automatic quality control of the resulting match. This quality control—a component often missing in commonly used image matching methods—is achieved by self-diagnostic measures supervising the iterative procedure.

The application of LSM to the problem of patient positioning was thoroughly investigated. Three issues had to be tackled: the calibration of the newly acquired portal image; the checking of the field shape; and the displacement measurement of bony structures. Precise calibration has been achieved by matching the field edge region to a reference shape, where LSM has proven to yield robust and accurate results. Hence, it was feasible to check the field shape by simply computing the normalized cross correlation between the actual and the prescribed field.

In several test series, the suitability of LSM was assessed to measure patient displacement. The area-based matching of reference structures to the treatment image was carried out using an affine or congruent transformation model. A very promising success rate of over 90 % was achieved using clinical test data which consisted of roughly 60 image series with a total of 500 portal images.

The optimum sources for reference data are the actual treatment settings for the calibration, the planning data for checking the field shape, and digitally reconstructed radiographs (DRR) for the displacement measurements. Due to missing system integration however, an interim approach in terms of employed reference data had to be chosen. That is, a validated portal image served as reference image for both the field edge as well as the anatomy match.

Using megavoltage DRRs with known ground truth, the systematic error, introduced by neglecting the projective nature of the portal images, was investigated. Furthermore, promising results for the multi-modal match between a DRR reference image and clinical portal images are presented; an approach which might significantly increase the treatment accuracy.

The generic measuring tool developed in the course of this project has been installed at the department of radio-oncology at the University Hospital of Zürich, with additional installations planned for the near future. This enables further clinical studies to be carried out, which will give useful insight about the current accuracy of patient positioning, and eventually about the potential and the effects of improving the treatment accuracy.

Zusammenfassung

Die Positionierung des Patienten während der Bestrahlung ist ein wesentlicher Bestandteil in der präzisen Strahlentherapie. Bei jeder Sitzung muss der Patient im Bestrahlungsgerät genau positioniert werden. Diese Positionierung kann durch die Analyse von Portalbildern, während der Bestrahlung aufgenommenen Röntgenbildern, unterstützt werden.

Das Problem von Portalbildern ist, dass sie extrem kontrastarm sind. Deshalb funktionieren Standard-Algorithmen für die Merkmalsextraktion, wie zum Beispiel die Kantendetektion, nicht mehr zuverlässig. Es ist eine sehr anspruchsvolle Aufgabe, einen robusten Algorithmus zur Merkmalsextraktion zu entwickeln, der bei Portalbildern funktioniert. Es müsste eine Methode gefunden werden, welche die geringe Bildinformation in eine nützlichere Form bringt, ohne direkt von der Bildqualität abhängig zu sein.

Um den Schritt der Merkmalsextraktion zu umgehen, wurde die *least squares template matching* (LSM) Methode gewählt. LSM ist ein iterativer und auf Grauwerten basierender Algorithmus, mit dem hohe Präzision erreicht werden kann, und der für die Verarbeitung kontrastarmer und unscharfer Bilder geeignet ist. In dieser Dissertation wird der mathematische Formalismus von LSM zusammengefasst und eine Methode zur automatischen Qualitätskontrolle, die in vielen Bildanalyseverfahren vernachlässigt wird, vorgestellt.

Bei der Anwendung von LSM auf das Problem der Patienten-Positionierung mussten drei Teilprobleme gelöst werden: die Kalibrierung des eben aufgenommenen Portalbildes, die Überprüfung der Form des Bestrahlungsfeldes, und die Messung der Verschiebung von Knochenstrukturen. Die präzise Kalibrierung des Portalbildes wurde mittels Grauwert-Vergleich (Matching) der jeweiligen Feldform mit der Feldform des Referenzbildes erreicht. Die Überprüfung der Feldform war somit durch Berechnen der normierten Kreuzkorrelation von eigentlicher zu vom Arzt geplanter Form möglich.

In diversen Testserien konnte die Eignung von LSM zur Verschiebungsmessung von Knochenstrukturen bestätigt werden. Das Matching der Referenzstrukturen mit dem Behandlungsbild basierte auf einer affinen oder kongruenten Abbildung. Die klinischen Testdaten bestanden aus ungefähr 60 Serien von insgesamt 500 Portalbildern, worin eine sehr vielversprechende Erfolgsquote von über 90 % erzielt wurde.

Optimale Referenzdaten wären die momentane Kollimatoreinstellung für die Kalibrierung, die Planungsdaten für die Überprüfung der Feldform, und digital rekonstruierte Röntgenbilder für die Verschiebungsmessung. Wegen der fehlenden Systemintegration mussten jedoch Zwischenlösungen gewählt werden. Das heisst ein validiertes Portalbild diente als Referenzbild für das Matching der Feldgrenze und der Knochenstrukturen.

Mit Hilfe von digital rekonstruierten Röntgenbildern wurde eine Testserie generiert, bei der die wirklichen Verschiebungen bekannt sind. Damit konnte der systematische Fehler untersucht werden, der durch die Vernachlässigung der projektiven Eigenschaften von Röntgenbildern eingeführt wurde. Darüber hinaus lieferte LSM vielversprechende Resultate beim (multi-modalen) Matching zwischen einem digital rekonstruierten Referenzbild und klinischen Portalbildern. Die Verwendung eines digital rekonstruierten Referenzbildes könnte die Genauigkeit der Positionierung nochmals erhöhen.

Das generische Messprogramm, das im Verlauf dieses Projektes entwickelt wurde, ist an der Abteilung für Radio-Onkologie am Universitätsspital Zürich installiert worden. Ausserdem ist die Installation am Kantonsspital Winterthur noch in diesem Jahr geplant. Somit wird es Klinikern ermöglicht, weiterführende medizinische Studien durchzuführen. Diese Studien sollten nützliche Erkenntnisse über die gegenwärtige Genauigkeit der Patienten-Positionierung und schliesslich über das Potenzial und die Auswirkungen der Genauigkeitsverbesserung liefern.

Acknowledgments

First of all, I like to thank my examiner Prof. Dr. Guido Gerig for supporting me and this project even after he has left the ETH. I also thank my co-examiner and medical supervisor Dr. Urs Meier, who answered all my medical questions and showed me how to see more than just gray on portal images. I further appreciate the readiness of Prof. Dr. Luc Van Gool to be my co-examiner and am looking forward to his remarks on the final thesis.

This project would not have been possible without the continuous support of Prof. Dr. Urs M. Lütolf and his team at the department of radio-oncology at the University Hospital of Zürich. I also like to thank PD Dr. Gábor Székely for his efforts of keeping this project alive before I started, and of course for his ongoing support.

Many thanks go to Gaudenz Danuser, who introduced me to the mysteries of least squares template matching, and with whom I had many interesting discussions about computer vision topics and other aspects of life.

Special thanks go to Jonathan Oakley for proofreading parts of my thesis and for the many discussions about the subtlety of the English language. Then to Trini Burckhardt, who helped me writing correct German again, by sorting out all the anglicisms in the German abstract. I also appreciate the innumerable discussions with Dr. Christian Brechbühler about mathematical problems and their formalism, in particular his constructive comments on the theoretical part of this thesis. Not to forget all my other colleagues at the Image Science Lab for their contributions and for making my time at BIWI a pleasant experience.

I further like to thank Lukas Adam for his efforts to use my measuring software in the course of his medical dissertation, proposing many improvements. His work provided me with valuable data concerning the clinical validation.

Last but not least, my warmest thanks to my family and all my friends for bearing with me in the final rush and simply for making life exciting.

This project was partly founded by the Swiss Cancer League and the Walter Honegger Foundation.

Contents

List of Figures

List of Tables

List of Symbols

ℓ	ideal, unperturbed observations
l	actual observations
ζ, ξ	parameter vectors
n	number of observations
r	number of parameters
f, g	reference and search image
e	residual vector
e_L	residual vector of the combined observations
e_0	residual vector of the spring constraints
w	functional residual vector
I	identity matrix
A	$n \times r$ Jacobi matrix
P	augmented weight matrix of the constraint system
P_{LL}	weight matrix of the combined observations
P_0	weight matrix of the spring constraints
N	$r \times r$ normal matrix $N = A^T P A$
Q	cofactor matrix $Q = P^{-1}$
$Q_{\hat{\zeta}\hat{\zeta}}$	$r \times r$ cofactor matrix of the parameters $Q_{\hat{\zeta}\hat{\zeta}} = N^{-1}$
$\Sigma_{\hat{\zeta}\hat{\zeta}}$	$r \times r$ covariance matrix of the parameters $\Sigma_{\hat{\zeta}\hat{\zeta}} = \sigma_0^2 \, Q_{\hat{\zeta}\hat{\zeta}}$
σ_0	a priori noise
$\hat{\sigma}_0$	a posteriori noise estimate

1

Introduction

The accuracy of radiotherapy treatment became an important issue in the last decade. The advances in tumor localization and the possibility of three dimensional treatment planning has been a cornerstone to more precise dose calculation, and hence to more accurate irradiation. However, this still depends on the accurate positioning of the patient at each treatment session, and an addressal of such a fundamental aspect has been somewhat neglected in the development of irradiation systems.

Improving the precision and reliability of the patient's position provided the motivation for this work. In particular, the analysis of megavoltage X-ray images, *portal images*, acquired during therapy should help to check and improve the actual patient's position. Thus, positioning based on external features such as skin marks should be replaced by positioning based on internal structures as close to the tumor as possible.

The main problem to be tackled in order to make use of portal images is their inherently low contrast. For computer vision, this leads to the challenge of robustly analyzing low-contrast and noisy imagery.

1.1 High precision radiotherapy

Radiotherapy is, together with surgery and chemotherapy, one of the major cancer treatment methods. Radiotherapy itself encompasses two different types of treatment: palliative or curative. Palliative treatment is mainly to alleviate pain incurred by the patient, as opposed to seeking a cure. The requirements on accuracy are therefore rather low, and problems occurring due to errors in patient alignment can be accommodated. In this work however, I will concentrate on the curative treatment, where accuracy is an important issue. The more accurate the irradiation is, the better the dose distribution can be computed. This leads to narrower peaks at the location of the tumor in the dose distribution. Hence, it is possible to administer a higher radiation dose to the tumor while reducing the dosage in the surrounding healthy tissue.

The advances in treatment planning systems allow a precise computation of the theoretical three dimensional dose distribution. This computation is based on the patient's CT volume, acquired from the tumor region. The processing of this CT volume is not fully automated, however. It is still necessary, for instance, to manually optimize the segmentation of the CT volume, the procedure of separating cancer cells from healthy tissue. Regardless of possible additional errors introduced by the manual segmentation, planning systems remain an indispensable tool for dose calculation and the resulting precision is sufficient for high precision radiotherapy.

An extreme case in terms of requirements on precision is the field of stereotactic radiosurgery. The treatment procedures involved are based on the irradiation of a small target volume with very narrow beams and from many different directions. Precise positioning is required in order to ensure the concentration of radiation in the target volume. Stereotactic radiosurgery is used in the treatment of brain tumors and represents a valuable adjunct to surgical resection, radiotherapy, and chemotherapy.

In each case, the limiting factor for an accurate treatment becomes the positioning of the patient. During a typical treatment, the patient is irradiated several times over a the period of days or weeks. Before each session, the patient has to be accurately positioned in the treatment device, a linear accelerator. The problem then is twofold: before the first session, the patient's pose within the treatment device must be found to establish the connection between the planning data and the actual treatment. Secondly, this initial position has to be marked in order to be able to correctly reposition the patient before each session.

The initial position is always determined during a simulation phase, in which the treatment is performed using diagnostic X-rays. The problem of fixation, the marking of this position, has been tackled in several ways.

- Skin marks aligned with laser beams:
  Once correctly positioned in the treatment simulation, lines are drawn on the skin along two calibrated laser beams. In the treatment room, identical laser beams allow the patient to be repositioned more or less accurately in the linear accelerator.

- Moldings and other fixations:
  Devices to fix the position relative to the table play an important role for the treatment of certain regions. This is especially the case for brain tumors since the skull can be fixed in an accurate way. Other regions however, such as the pelvis, cause more difficulties in finding a reliable fixation.

- Video controlling:
  An interesting application is the three dimensional analysis of video sequences. The signals of two or more video cameras allow one to track the position in space of special reflective markers on the skin.

Although important, these three measures suffer from the drawback that the fixation and positioning is based on external and often non-rigid features. Furthermore, only the video controlling approach allows the position to be checked *during* the treatment. Thus, the goal is to find alternative means of checking the patient position based on robust features situated as close as possible to the tumor itself.

1.2 The potential of portal images

One promising type of sensor is the electronic portal imaging device (EPID), which has become more readily available in recent years. This device registers a two dimensional exit dose distribution during radiotherapy, similar to a diagnostic X-ray.[1] The main drawback of such portal images is their inherently low contrast and substantial noise, rendering reliable image analysis a difficult problem. A main goal of this project is to assess the suitability of portal images for an improved position control. That is, to design a reliable measurement method for the patient displacement.

Before exploiting the information of portal images, a few additional thoughts are necessary regarding the actual problem. The patient's position is defined in a three dimensional patient coordinate system. Under the assumption of a rigid motion—which I adopt throughout this thesis—the six parameters of translation and rotation must be found to fully correct any patient setup error. This will hardly be possible by analyzing one portal image, which only depicts the two dimensional projection of the bony structures. It is straightforward to show that trying to solve for the complete three dimensional parameter set with only one portal image would lead to an ill-posed problem.

Still, portal images are a promising source of information. On one hand, we can make assumptions on the three dimensional movement and use the two dimensional measurements to estimate parts of the patient motion. In this case, the problems to be dealt with include the, possibly neglected, projective nature of the portal images and the systematic errors introduced by the rather constraining assumptions. These shortcomings are partly compensated by introducing certain knowledge about the anatomy or about the range of possible movements.

On the other hand, portal images are best used as *additional* information, supplemented with the CT volume, the planning data and treatment settings, and maybe the simulation images. Within such a framework, portal images can deliver the missing amount of information to ensure accurate patient positioning. In order to fully exploit and combine all this data, a great deal of system integration on the side of the device software is essential. However, this necessary integration is not yet available at the time of writing. Nonetheless, suitable interim solutions can be found until further advances in terms of system integration will allow for an optimum solution.

In employing portal images, two additional problems have to be tackled. Firstly, the exact position of the EPID is unknown in general. Therefore, the portal images have to be registered to establish a common coordinate system. There are several

---

[1] Electronic portal imaging devices are discussed in more detail in section 2.3.

solutions to this problem, depending on the aforementioned system integration. Usually, the registration is achieved by comparing the edge of the treatment field with a gold standard of known position. Optimally, this gold standard is based on the treatment settings; that is, on the current setup of the field shaping device, the so-called multi-leaf collimator. If the information about treatment settings is not available, as is often the case, the field shape of a previously registered portal image must serve as gold standard.

Secondly, the shape of this treatment field must be checked if it matches the form defined in the planning step. Again, there is an interim and an optimum approach. The latter is given by comparing the actual field shape with a simulated shape computed from the planning data. If the planning system does not support such a simulation, the field shape of a previously checked portal image is used, as in the registration step.

1.3 Low-contrast imagery: Area-based versus feature-based methods

Noisy and low-contrast images pose uncommon problems to computer vision algorithms. Usually, one of the first steps in image analysis is to transform the large amount of information typically present in an image into a more practical representation. In most computer vision algorithms, this transformation also involves a significant reduction of the original information content.

A well-known example is applying a feature extraction operator, for instance an edge detector. Although there are more general approaches, edge detectors are mostly used to extract edges as a *binary* feature, that is, a gray level image signal is reduced to a binary mask.

In many cases, exactly this reduction is desired to further process the vast amount of input data, and hence the popularity of such approaches. However, many existing methods deal rather carelessly with the information content of images. Often, too much or the wrong part of the information is discarded during the various processing steps, leading to a loss of both accuracy and reliability.

Nevertheless, numerous applications do not depend on fully exploiting all available information. It is thus reasonable to accept some loss of valid information in exchange with a simplification of the algorithm. In the case of low-contrast imagery however, a loss of accuracy or reliability might inhibit a successful outcome. Thus, any information reduction applied to low-contrast imagery has to be carefully designed. But it remains an open issue how to design such a feature extractor. Most, if not all, feature extraction algorithms will not produce a reliable result operating on imagery with a low signal to noise ratio like portal images.

This is the justification for using an *area-based* method, as opposed to a *feature-based* method. Area-based methods do not require a specific feature extraction step, since they directly use the image signal as input data. That is, the complete image information is accessible. This may be both an advantage and a disadvantage,

	area-based	**feature-based**
methods	maximize cross correlation maximize mutual information minimize sum of squared errors (least squares template matching) ...	Canny edge detection Förstner's operator Gabor filters ...
information	full image information	reduced information
advantages	no feature extraction high precision	high separability
drawbacks	gray scale dependent	feature extraction necessary precision dependent on feature extraction

Table 1.1: Advantages and disadvantages of area-based and feature-based methods. The listed methods are arbitrarily chosen examples from the numerous available algorithms.

depending on the application. Table 1.1 summarizes the qualities of each type.

The main advantage of area-based methods—besides not having to extract any features—is their potential for high precision measurements, which is a direct result of using the full image information. For a correctly sampled image signal, the precision is limited only by the signal to noise ratio. On the other hand, care must be taken in designing area-based methods in order to reduce the gray scale dependence as much as possible. In contrast to feature-based methods, where this dependence is usually eliminated in the feature extraction step, area-based methods must inherently supply the desired level of gray scale independence.

The advantage of feature-based methods lies in the possibility of achieving high separability for certain features. With a well designed feature extractor, several instances of one feature can be separately detected even if they are only few pixels apart.[2] However, the step of feature extraction is crucial, since the subsequent matching procedure is dependent on its outcome.

Thus, area-based methods provide an elegant approach to circumvent the difficult step of feature extraction in low-contrast imagery. The costs are the higher complexity of the matching algorithm and an inherent gray scale dependence, both of which must be accommodated for. But as this thesis shows, both problems can be solved for this application by employing the least squares template matching method.

---

[2] For a more detailed discussion and interesting examples of different feature extractors please refer to [Danuser 1997, section 3.3].

2

Electronic portal imaging

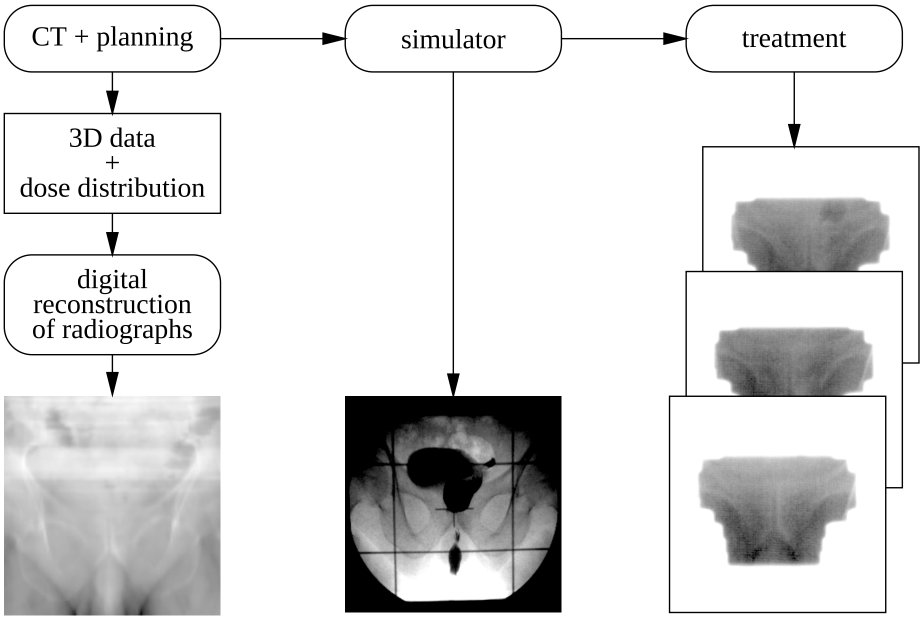
In the last years, electronic portal imaging devices (EPID) have become more readily available. The devices have definitively evolved from an experimental stage to a level, where they can be employed in daily hospital routine. What is missing however, is a convenient and reliable way of analyzing the thusly acquired images. In other words, methods have to be developed to make reliable measurements based on such portal images, in order to fully exploit the information contained herein.

Before I explain in more detail the potential and the drawbacks of the EPID in section 2.3, a short overview of the current hospital practice and of the involved devices is given in the first two sections. The important aspect of digitally reconstructing a portal image with known parameters—hence generating an optimum reference image from the CT volume—is outlined in section 2.4, followed by a description of the specific problem to solve. The chapter concludes with an overview of previous work on portal imaging.

2.1 Current procedure for radiotherapy treatment

In high precision conformal radiotherapy it is essential to accurately position the patient during the series of treatment sessions. Figure 2.1 gives an overview of the various steps of conformal radiotherapy. First, CT volume data is acquired. Based on this CT volume and the desired dose distribution, the physician defines the irradiation directions and the field shapes—seen as white outlines in the megavoltage X-ray images—in the planning step. At this stage, the theoretical dose distribution within the CT volume is computed. Furthermore, digitally reconstructed radiographs (DRR) can be computed from the CT data (figure 2.1a), which will be discussed in section 2.4.

In order to reach the theoretical dose distribution as exactly as possible, the patient must be positioned at each treatment session in accordance with the CT coordinate system. Two possibilities exist to find this position for the initial treatment. The conventional approach is to simulate the therapy using diagnostic X-rays instead of megavoltage X-rays, which allows standard X-ray images to be acquired (figure 2.1b). Based on these images, the patient is iteratively moved towards the optimum position. The coordinate systems of the simulator and of the linear accel-



a. DRR **b. diagnostic X-ray** **c. megavoltage X-ray**

Figure 2.1: Overview of the different steps of radiotherapy treatment. The three steps *CT+planning*, *simulator*, and *treatment* yield three types of imagery: digitally reconstructed radiographs (a), diagnostic X-ray images (b), and portal images (c).

erator are identical. As localization device serve two laser beams present at both devices, which define two perpendicular planes. Hence, the patient can be repositioned in the same location by marking the skin along these planes and align these line marks with the laser beams in the subsequent treatment.

The second possibility is to use virtual simulators. This is a rather new technology, where the localization device is directly integrated in the CT scanner. The treatment simulation is carried out virtually by computing DRRs, thus rendering an additional simulation step obsolete. Analogous to the conventional method, line marks on the skin are used to realign the patient in the treatment room. Since one source of error has thusly been eliminated, virtual simulation should be the preferred method from the computer vision point of view. However, the computed DRRs are not yet of the same quality as the diagnostic X-ray images obtained from the standard simulator, which still lowers the acceptance of this technology among physicians.

The drawback of these procedures is that the patient positioning is only based on reference points *outside* the body. This might be suitable for rigid structures as for tumors in the head region, but is error prone where the reference points are on soft tissue, for instance in the case of the pelvis region. The position of these reference points can significantly change over the period of treatment, which introduces another source of error. For a more reliable alignment, stable features close to the tumor should be employed.

2.2 Simulator and linear accelerator

Both devices, the simulator and the linear accelerator, are identical in terms of coordinate systems and geometry. As described in the previous section, two laser beams present at both devices enable marking and localization of the patient within this coordinate system.

The difference between the simulator and the linear accelerator is the different types of X-rays used. Whereas the simulator contains a standard diagnostic X-ray source of about 100 keV, the linear accelerator works with a megavoltage source ranging from 6 to over 20 MeV. This has a major influence on the image contrast as we will see in section 2.3. Otherwise, the two devices are very similar, which allows me to concentrate on the linear accelerator in this section.

The main parts of a linear accelerator are the gantry, the collimator, the table, and the portal imaging device. Figure 2.2 illustrates the various coordinate systems corresponding to these parts. The main reference point is given by the isocenter, in the figure marked by a thick black dot. The isocenter represents the intersection of the beam axis with the rotation axis of the gantry.

The gantry contains the radiation source and a collimator device. It can be rotated about the isocenter. The patient—or a CT volume during the digital reconstruction of a portal image—is placed on a table which usually has four degrees of freedom: translation in all three directions and rotation in the table plane. Therefore, the gantry position together with the table orientation defines the irradiation angle, that is, from which direction the beam enters the patient's body.

The collimator device may be a simple perspex plate with mounted lead blocks or, in newer installations, a multi-leaf collimator. Such a multi-leaf collimator device consists of an array of small, longish lead blocks which can be positioned along one axis according to the required beam shape. This results in characteristic steps in the field edge contour, for instance visible in figure 2.1c.

2.3 Electronic portal imaging device

The most promising sensor for improving patient alignment is the electronic portal imaging device (EPID). This device delivers images of the exit dose distribution during treatment similar to a standard X-ray image.

Various types of EPIDs are available. An interesting although little bit outdated review can be found in [Boyer *et al.* 1992]. Presently, most devices are either based on a liquid ionization chamber or a fluorescent phosphor screen viewed by a CCD camera. Examples of the latter system are the devices *iView* from Elekta Oncology Systems (EOS) and *BeamView* from Siemens Oncology Care Systems.

An example of a liquid ionization chamber was developed by [van Herk and Meertens 1988] and is available as *PortalVision* from Varian Oncology Systems. Their portal imaging system consists of a 256×256 matrix ionization chamber on an area of $32 \times 32\,\text{cm}^2$. This leads to a square pixel size of approximately 0.9 mm

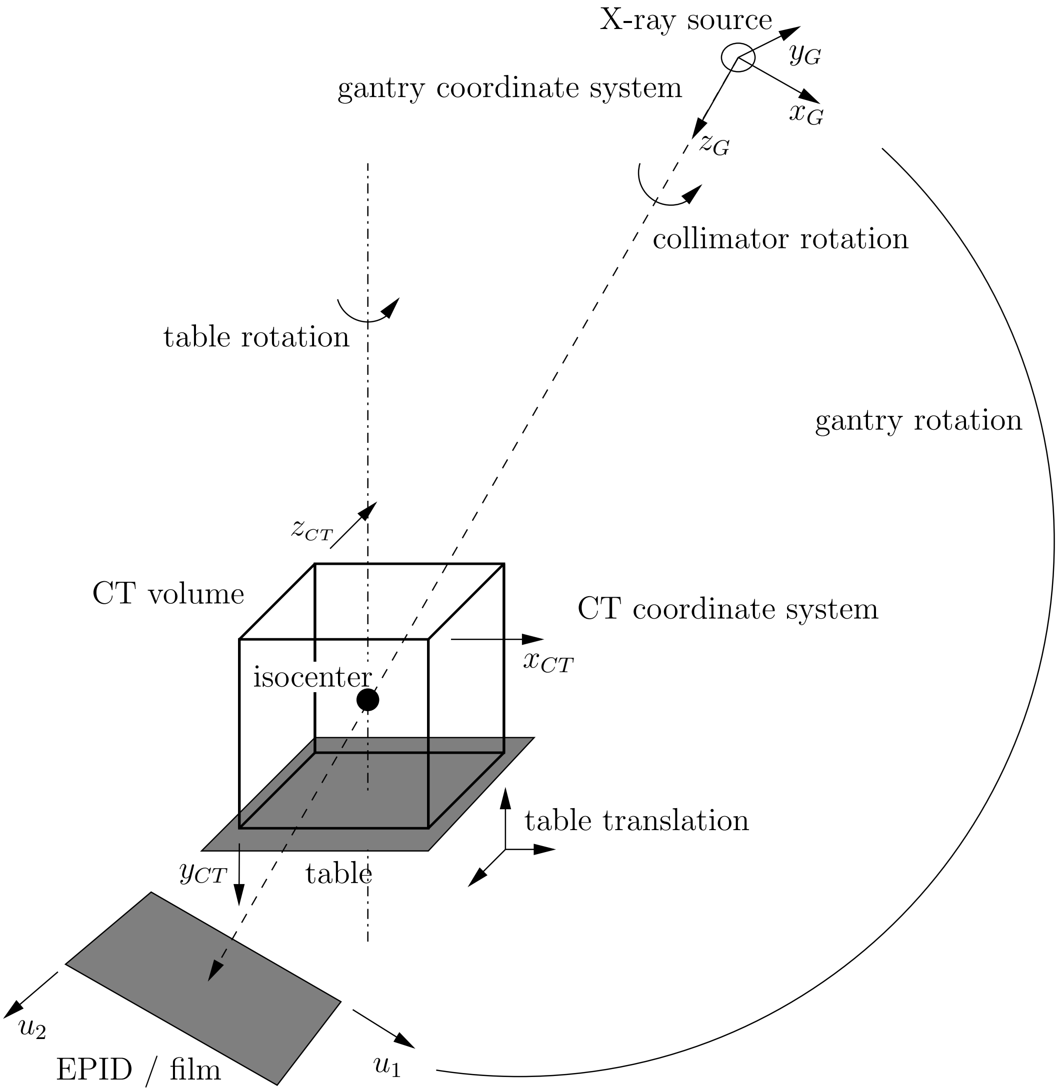


Figure 2.2: Coordinate systems of the linear accelerator.

in the plane of the isocenter for the standard setup as depicted in figure 2.3a. The gray value resolution is 12 bit and the nonlinear input/output characteristics of this device was further investigated by [Yin *et al.* 1994].

Most portal images used in this study were acquired using one of the Varian accelerators 600C or 2100C and their EPID PortalVision. A few images were obtained using the BeamView device from Siemens. In contrast to PortalVision, their device provides 8 bit gray value images with a spatial resolution of 512×480 pixels resulting in a square pixel size of about 0.5 mm in the isocenter plane. Both types of images, PortalVision and BeamView images, were acquired at the University Hospital of Zürich.

The exact location of the EPID is not known in general. Depending on the installation, the degrees of freedom of the EPID ranges from 0 to 3, that is, from more or less fixed to allowing translations in all three directions. Since rotational

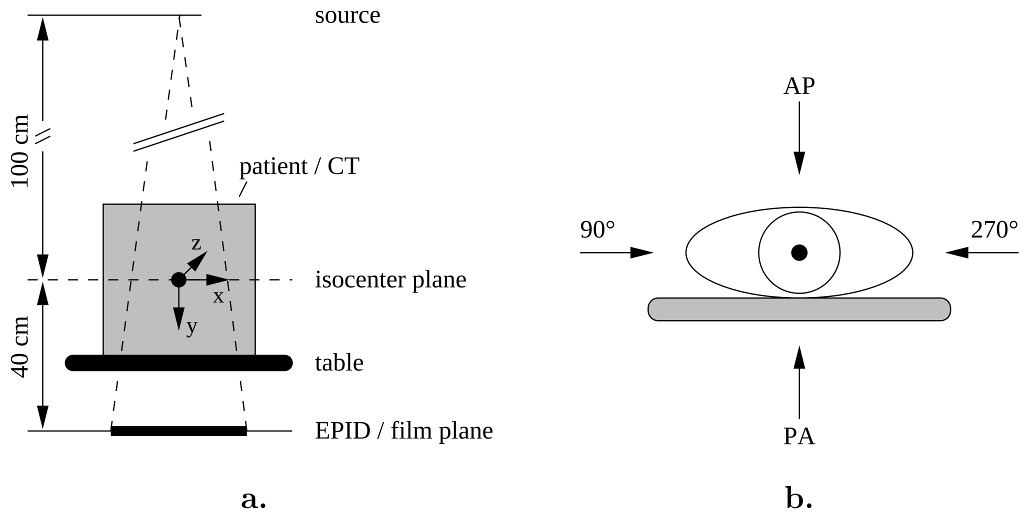


Figure 2.3: A standard treatment setup is illustrated in (a) including the patient coordinate system. The EPID is typically located about 40 cm below the isocenter plane. The four main beam directions are shown in (b).

movements are not possible, the beam axis always stays perpendicular to the image plane.

An EPID is normally mounted opposite the gantry on a retractable arm, which on some installations also reports the position of the EPID. In all cases however, the position is only known up to a few millimeters. This calls for an additional registration step before the connection between the image and the patient coordinate system is made and a portal image can be used for measurements.

In contrast to diagnostic X-ray, which is acquired at energies below 200 keV, megavoltage X-ray images are taken at treatment energy level which ranges from 6 to over 20 MeV. At these energy levels, the attenuation is governed by Compton effects and pair production. Figure 2.4 illustrates the influence of the various interaction types in the case of photons in water.

The contrast of a X-ray image is mainly given by the difference of total attenuation between soft tissue and bone. Qualitatively, this attenuation difference is determined by analyzing the dependency of the various effects on the atomic number Z. Since attenuation by photo effect is proportional to Z^3, high contrast images are obtained at low energies. Conversely, Compton absorption is only proportional to Z which leads to significantly lower contrast in megavoltage X-ray images. Moreover, there is typically a high amount of noise at high energies, thus further reducing the signal to noise ratio of portal images. Figure 2.5 depicts two typical portal images from the pelvis region.

Moreover, there is also the possibility of unstable features as shown in figure 2.5b. The dark blur in the center of the image originates from air in the rectum and may not be used for matching. Another type of unstable features are lines or edges only visible under a certain view angle.

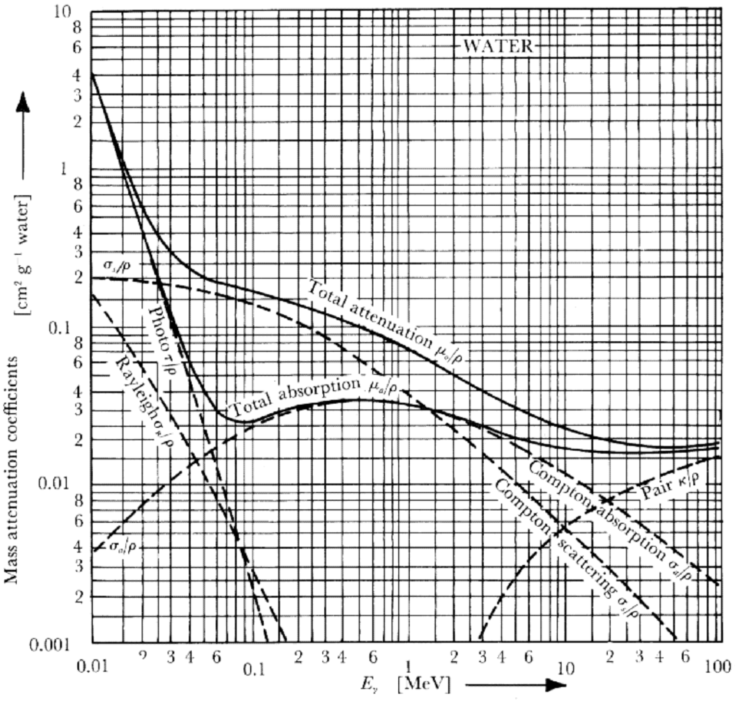
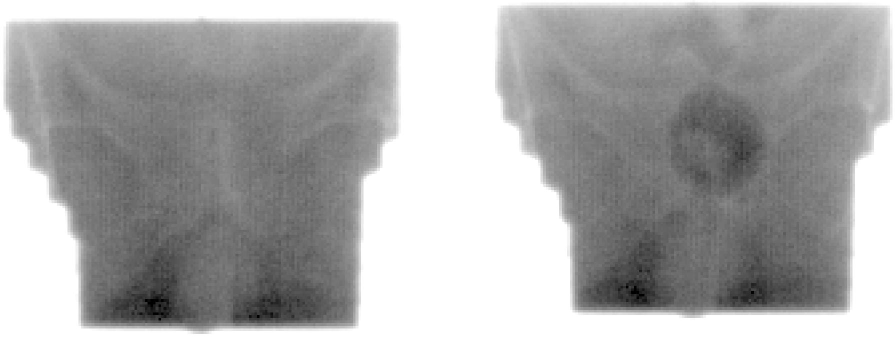
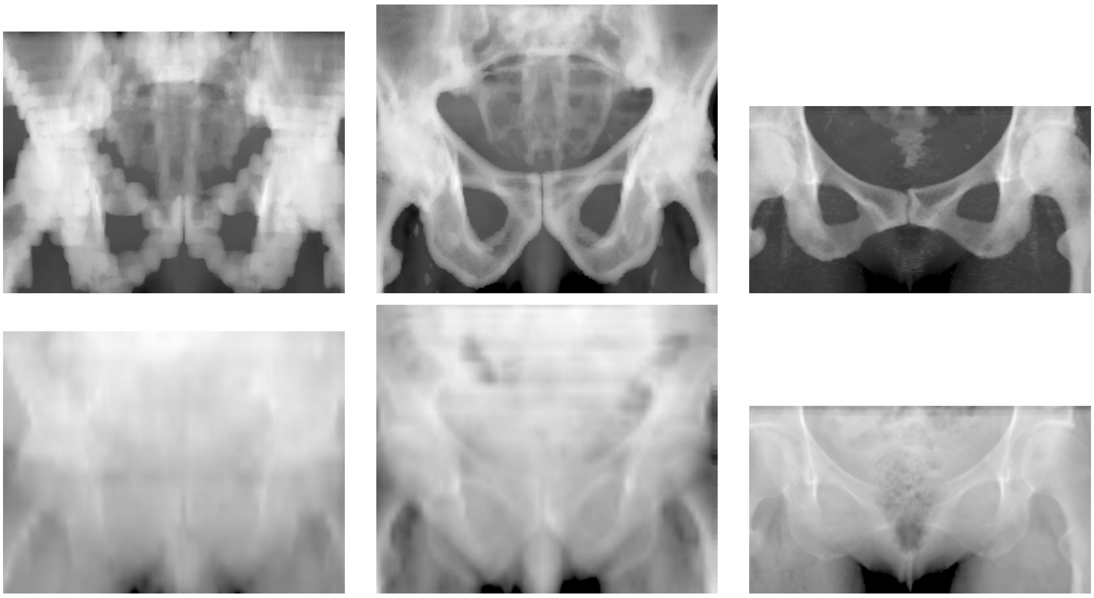


Figure 2.4: Mass attenuation coefficients for photons in water or biological tissue. Compton effects and pair production govern the total attenuation at high energies [Marmier and Sheldon 1969].



a. **b.**

Figure 2.5: Two typical portal images from the pelvis region. The images are bounded by the field edge and usually differ slightly in position and scale as in this example. Typical for portal images are also unstable features like the dark blur in (b), which originates from air in the rectum.



a. 10 mm slices **b. 3 mm slices** **c. 1 mm slices**

Figure 2.6: Diagnostic (top row) and megavoltage DRRs (bottom row) of three CT volumes with differing slice thickness. A CT volume with 10 mm slices is clearly not suitable to compute DRRs (a). The quality increases significantly when using 3 mm slices (b), which should be sufficient for measuring portal images. For comparison, a high resolution example was computed from the Visible Woman dataset (c).

2.4 Digitally reconstructed radiographs

As illustrated in figure 2.1a, artificial X-ray images or *digitally reconstructed radiographs* (DRR) can be computed from the CT volume [Sherouse *et al.* 1990, Chaney *et al.* 1995]. The treatment setup depicted in figure 2.2 is simulated by a ray-tracing based algorithm, correcting for the different absorption coefficients at different beam energies. Depending on this correction, either diagnostic or megavoltage DRRs are obtained. Presently, most treatment planning systems allow the computation of both types of DRRs. The simulation software used within this project was ported from the "portal software tool" described in [Chaney *et al.* 1995] and implemented in our lab during a master's thesis [Styner 1997].

Portal images acquired during treatment can be directly matched with a megavoltage DRR, since their gray value characteristics are similar. This provides an accurate link between the planning step and the actual treatment. Employing such DRRs as reference images is potentially very interesting for increasing the accuracy and efficiency of radiotherapy treatment. However, the slice thickness of the CT must be less then 5 mm in order to achieve a sufficient image quality. Figure 2.6 illustrates the effect of the CT slice thickness on the quality of the DRRs. Clearly, a CT volume with 10 mm slices is not suitable for computing a DRR, whereas 3 mm slicing should be sufficient to reconstruct a reference portal image.

2.5 Patient positioning using portal images

The accuracy achieved with the conventional procedure as described in section 2.1 is in the range of 10 to 20 mm. The goal is to find a method that increases this accuracy to a few millimeters by analyzing the portal images. Furthermore, the analysis must be fast enough to allow an early interruption of a treatment session. Table 2.1 summarizes these specifications.

size of search window	10–20 mm
required accuracy	2–5 mm
computation time	< 5 sec

Table 2.1: Specifications for the portal image analysis.

Before a portal image can be used for measurements, its coordinate system must be calibrated with respect to the patient or CT coordinate system (see also figure 2.2). I have discussed in section 2.3 that the exact position of the EPID is not know. In a first step, this position must be found, that is, the portal image must be calibrated based on known features.

An ideal feature is the radiation field edge. Its shape is defined by the physician in the planning step. During the treatment, the field shape is formed either by lead blocks or by a computer controlled multi-leaf collimator. In the latter case, the settings of the collimator can be used to generate a perfect standard for the calibration. If this data is not to hand, the field shape of a previously registered portal image serves as reference. In both approaches, the field edge of a newly acquired portal image must be aligned with a reference field edge. A simple but usually sufficient model for this *field edge alignment* is to allow for translation and scale, accounting for the three degrees of freedom of the EPID. When minor rotation should also be taken into account, a similarity transformation is used instead.

The exact shape of the field must also be checked based on the planned shape. Ideally, this *field edge check* is achieved by directly comparing the current field shape with a simulated field shape from the planning data. To date however, the field shape must often be compared to a previously registered portal image, since the planning data is not available. As similarity measure between the reference and the treatment field edge serves the normalized cross correlation of their gray values, which proves to be a robust measure for detecting field shape deviations.

Based on the field edge alignment, the displacement of bony structures must be measured in order to check the patient position. As already mentioned in section 2.3, the main difficulties arise from low contrast and unstable features. Also, the projective nature of the portal images should be taken into account when the three dimensional patient motion is to be estimated. This leads to the challenging task of estimating at least parts of the rigid patient motion based on two dimensional projections.

Similar to the previous steps of field edge alignment and check, there is an optimum and an interim approach in terms of choosing the reference image. An

	interim solution	optimum solution
field edge alignment (calibration)	• field edge extraction on reference image • field edge alignment	• field simulation based on collimator settings • field edge alignment
field edge check	• field edge extraction on reference image • field edge alignment • similarity test on aligned field edges	• field edge simulation based on planning data • field edge alignment • similarity test
anatomy alignment	• define portal image as reference image • portal alignment to portal image alignment	• computation of DRRs based on planning data • DRR to portal image alignment

Table 2.2: Interim and optimum solution of the three main stages. Interim solutions only differ in the choice of the reference data and of the algorithm.

optimum reference are megavoltage DRRs, computed from the CT volume and the planning data. However, such a DRR can only be computed when a high resolution CT was acquired (slice thickness smaller than 5 mm). Within this work, only a few such datasets were acquired. In all other image series, an interim approach had to be chosen, where the reference image is a previously registered portal image, usually the first portal image of a series.

Table 2.2 gives an overview of the three main stages and both their interim and optimum solution. It is important to note that the interim solutions differ only in the choice of the reference data and not in the choice of the algorithm. That is, the same methods can be applied once the optimum reference images are available. Thus, most examples in this thesis were computed by applying the framework of the interim solution. Figure 2.7 depicts an overview of the resulting algorithm parts (gray boxes) and their results.

2.6 Previous work on portal image analysis

The previous section explained the need to extract and align the field edges of two portal images before measuring the actual anatomy displacement. On all three steps of portal image analysis, previous work has been carried out, which is summarized in the following. The anatomy alignment is further subdivided into strict two dimensional methods and combined or three dimensional methods, which try to take into account the third dimension.

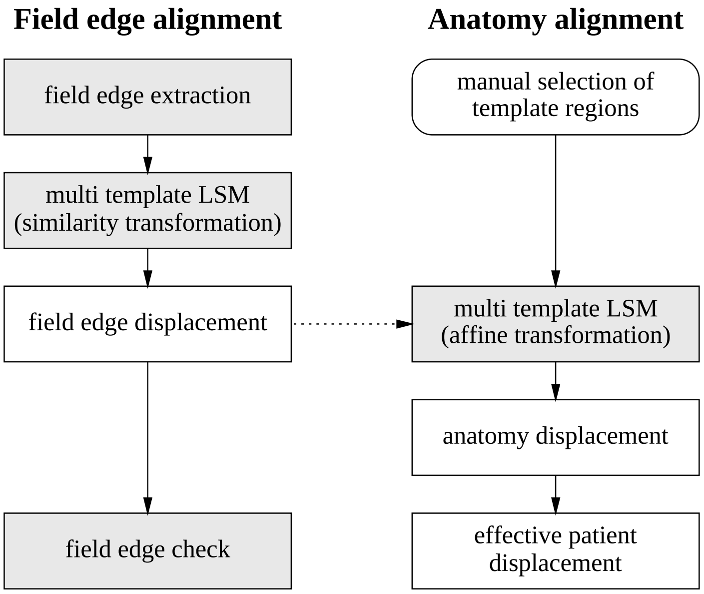


Figure 2.7: Flowchart of the interim alignment procedure (gray boxes represent algorithm parts). The result of the field edge match is used as initial guess for the anatomy alignment (dotted arrow).

2.6.1 Field edge extraction

A simple method for extracting the field edge was presented by [Bijhold *et al.* 1991a]. They first approximate the field edge position by thresholding the portal image at a suitable value which is obtained analyzing local minima in the gray value histogram. In a second phase, all edge pixels are shifted in the direction of local maximum gradient employing Sobel compass filters.

[Leszczynski *et al.* 1992] proposed the application of the Canny edge detection operator followed by a sequential contour following algorithm. Their method is based on the work of [Lacroix 1988] and [Kunt 1982]. Additionally, they apply a heuristic, regression-type strategy to segment the contour into straight line segments.

Completely based on line segments is the field edge extraction method published by [Eilertsen *et al.* 1994]. They use the *Laplace of Gaussian* operator as feature extractor and apply the Radon transformation to detect straight lines. In a second phase, these line segments are connected again using the Radon transformation.

A slightly different approach was described by [Wang and Fallone 1995]. They threshold the gradient image leading to a binary edge map, where the edge is a few pixels wide. Consequently, the region inside this binary edge is a few pixels smaller than the real size of the radiation field. Applying several dilation and erosion operations, a series of one pixel wide field edge candidates is generated. Their respective mean gray values represent the general form of the field edge slope, to which a hyperbolic tangent function with four degrees of freedom is fitted. The most probable field edge is then selected based on the resulting parameters. They claimed to have successful results on both, double and single exposure portal images.

2.6.2 Field edge alignment

For the field edge alignment, [Bijhold *et al.* 1992] proposed to apply the method of normalized or invariant moments. They computed the invariant moments of the one-pixel wide field edge contour and compared them to the moments of the reference contour. [Wang and Fallone 1994] applied the same method to the whole radiation field mask, not only to the contour. Their results have indicated that this approach leads to an improved registration of the field edges.

The method of invariant moments to compare image patterns was first introduced by [Hu 1962] and further extended by various authors. Moments can be made invariant to translation, rotation and scale and thus provide a measure to compare the shape of two objects. Refer to [Teh and Chin 1988] for an overview and evaluation of the different algorithms. However, employing invariant moments for measuring a transformation between two objects has various disadvantages. The main problem is that differences in shape may significantly influence translation and rotation measurements, especially if only applied to the contour. Thus, it is not surprising that [Wang and Fallone 1994] achieved better results, but they still suffer from the mentioned drawbacks. Moreover, both approaches strongly rely on a reliable and accurate field edge extraction algorithm, which is difficult to achieve for all kinds of portal images.

Another approach to align to outlines is given by polynomial warping as proposed by [Eilertsen *et al.* 1994]. The corner coordinates of the field edge polygons were chosen as *tie-points*, which can be mapped onto each other using a least squares estimation technique. They used the six affine parameters as transformation model. However, this method depends on the accuracy of the extracted field edge contour like the method of invariant moments.

2.6.3 Two dimensional anatomy alignment

The various methods for portal image registration can be divided into the two categories presented in section 1.3: area-based and feature-based methods. Probably the best known feature-based methods are point-to-point (or landmark) algorithms, where a certain number of landmarks have to be identified in both images (for instance [Meertens *et al.* 1990], [Ding *et al.* 1993], [McParland and Kumaradas 1995]). This approach was further developed by [Balter *et al.* 1992], who align open curve segments and points seen on two radiographs employing the Procrustes algorithm [Schonemann 1966]. Still, the features must be delineated in both the reference and the search image and thus requires a great deal of user interaction.

Such landmark methods greatly depend on the exact localization of the landmarks by the physician. This is not only time-consuming, but also varies for different operators. First attempts to reduce the workload and the variation lead to hybrid methods, where features are only delineated in the reference image. Even if these methods are still strongly operator dependent, the restriction to the reference image diminishes the workload substantially.

An example for this type of methods is the chamfer matching algorithm. Chamfer matching is a technique originating from the artificial intelligence community [Barrow *et al.* 1977] and was applied to edge matching by [Borgefors 1988]. The application to portal images was presented by [Gilhuijs and van Herk 1993]. In their approach, significant ridges have to be manually outlined in the reference image. These outlines are then matched onto the detected features of the treatment image. As feature detector served the morphological top-hat transform (see for instance [Serra 1982]). Hence, this method depends on the correct selection of ridges, that is, the operator has to know about the top-hat transform in order to extract the same features as the morphological feature detector. In the follow-up paper [Gilhuijs *et al.* 1995], they used a multi-scale medial axis transformation as feature detector and found slightly better results.

A similar approach is described in [Fritsch *et al.* 1995]. Instead of the top-hat transform, they introduce the notion of *cores*. Cores are skeleton-like structures, representing medialness of the object. The advantage of their method is that cores are extracted on the reference image as well. This reduces the operator variability, since reference structures are merely selected and do not have to be manually outlined.

[Eilertsen *et al.* 1994] presented the application of an unsharp masking technique. They implemented unsharp masking by subtracting a low-pass filtered version of the image with subsequent manual thresholding. The resulting binary ridge maps of the reference and treatment image were then correlated within the three dimensional search space of translation and rotation.

However, since portal images are inherently noisy and low in contrast, it is difficult to robustly detect features like edges, ridges or cores. Therefore, an area-based algorithm promises to be superior to a feature-based algorithm.

A well known area-based technique is minimizing the gray value correlation coefficient. An early paper on the application of gray value correlation to portal images was presented by [Jones and Boyer 1991]. They implemented a FFT-based search of one region of interest on the discrete pixel grid. Further speed up of the extensive search for the best fit was achieved by [Radcliffe *et al.* 1994], with a special type of “statistical” correlation. Such speed up is necessary in order to allow for more degrees of freedom than just translation, that is to include rotation or even scale in the search space. Additional parameters quickly lead to an explosion of the size of the search space, which renders an efficient implementation difficult. [Moseley and Munro 1994] therefore restricted their independent registration of several anatomic regions to translation with the drawback of being sensitive to rotation. From the resulting set of translations of the individual regions, they computed a similarity transformation. Due to the initial restriction to translation however, this method is sensitive to rotation. Still, they claim a precision within one pixel if the rotation was below 4°.

A two step procedure was proposed by [Dong and Boyer 1996], where a coarse result is first computed based on translation only. Then, a refined result is obtained by searching the three dimensional search space also including rotation. Even with this two step approach, they were forced to use rather large step size, for instance 3° for the rotation parameter.

In an earlier publication, [Dong and Boyer 1995] used a megavoltage DRR as reference image instead of an approved portal image as in the aforementioned area-based methods. They showed promising results on the gray value correlation of a megavoltage DRR with a treatment image.

However, the limitations of all these area-based methods for matching portal images lie in the restriction to translation or in a coarse search grid for computational reasons. Furthermore, none of the aforementioned methods, feature-based or area-based, show an error propagation analysis for the measured displacement. This would be crucial for any kind of automated quality control, which is necessary to achieve the reliability and accuracy required for daily hospital routine. In current clinical practice, the result must be visually checked and often manually corrected.

In [Berger and Danuser 1997], the application of least squares template matching (LSM) to portal images has been presented. As I will discuss in this thesis, LSM overcomes the problem of large search spaces since it is based on an optimization strategy. Moreover, the least squares framework allows one to compute the error propagation and thus to develop an automated quality control.

2.6.4 Three dimensional anatomy alignment

One of the first three dimensional approach was presented by [Brunie *et al.* 1993]. Based on CT or MR volume data, they matched the projections of a segmented feature with its corresponding feature in the portal image. The strength of this technique is the fusion of three and two dimensional data. Results were shown on a head phantom. However, I believe that the aforementioned drawbacks of feature-based matching will reduce the accuracy and robustness in practice.

A landmark based, three dimensional verification method was presented by [Bijhold 1993]. Based on the CT, he manually localized anatomical match points in both the simulator and the treatment image including the third dimension.

A different approach was proposed by [Gilhuijs *et al.* 1996]. Bony structures present in the CT data are mapped onto the projected portal images under different irradiation angles and the most similar view is chosen as the current patient position. They used the simplex method as optimization strategy. However, they only showed results on simulated data, which is simpler than the application to real portal images.

An interesting combination of area-based and feature-based methods has been presented by [Bansal *et al.* 1998], who applied a framework similar to the Expectation-Maximization algorithm (see for instance [Dempster *et al.* 1977]). The mutual information between the portal image and a computed projection is maximized, while estimating a rough classification into bone and background at the same time. They chose an annealing schedule as optimization strategy comparable to the method of simulated annealing.

In [Berger and Gerig 1998], we have shown the feasibility of comparing a megavoltage DRR with portal images by applying least squares template matching.

3

Least squares template matching

LSM is an area-based matching algorithm. It replaces the conventional multi-stage approach where feature extraction is followed by thresholding, binarization and a discrete search. The main idea of LSM is to fit the image signal of a template into the search image, minimizing the sum of squared errors between the gray values of the two regions.

Thus, LSM does not depend on the extraction of binary (also called *non-iconic*) image features. This is an important advantage in low contrast and blurred imagery, where feature extraction is mostly unreliable. Furthermore, unlike in most correlation methods, the optimum transformation is not searched on a discrete grid, but approached using an optimization scheme. Assuming that a fair initial guess can be supplied, this is not only faster but also more accurate.

3.1 Previous work on LSM

Early work in this field was presented by [Lucas and Kanade 1981], who published an iterative image registration scheme based on LSM. Among the first papers that discussed the concept of exploiting the full information of the statistical models for robust template matching are [Grün 1985] and [Förstner 1987]. [Bergen *et al.* 1992] describe basically the same algorithm for motion estimation. It was further developed by [Lindeberg 1995] using a multi-scale approach.

Following and extending the work of Grün, [Danuser and Mazza 1996] achieved highly accurate results at the resolution limit of a light microscope. The high accuracy of this technique even in the case of low-contrast imagery is extensively exploited in [Danuser 1996]. The paper reports on high accuracy positional measurements of a calibration grid used to calibrate a stereo light microscope. Compared to this application, additional problems arise in portal images from the higher complexity and variability of the image scene, and from the effects of out-of-plane rotations on projected images. On the other hand, the requirements on accuracy are not as high in portal imaging.

A similar technique for the registration of medical image series is reported by [Unser *et al.* 1995], where each image is matched to the reference image based on a global gray value difference measure using a pyramid scheme. It was further developed in [Thévenaz *et al.* 1998] to include volume registration using a three dimensional affine transformation. In contrast to their work, our framework does not rely on one global template, but on several small templates each containing a significant image structure. Thus, the inclusion of distinct but insignificant image features which vary between the data of one sequence is avoided and the impact of global gray value errors such as intensity inhomogeneity is reduced.

Least squares matching techniques have also been applied to tracking problems. [Jianbo Shi and Tomasi 1994] separately track a series of small templates using an affine transformation model and also select good features based on the normal matrix of the linear least squares system. An application of piecewise projective tracking was described by [Gleicher 1997]. Both approaches did not tackle the problem of illumination insensitivity, which was presented in a more recent paper by [Hager and Belhumeur 1998]. They also provide an overview of the application of LSM to tracking problems.

In the following three sections, the generic framework of solving least squares problems is reviewed including error propagation and self-diagnosis. The less general case of template matching is presented in detail starting with section 3.5.

3.2 General least squares framework

Least squares estimation is widely dealt with in standard literature. In this section, a short summary is given about the parts relevant to LSM. For more in-depth information the reader may for instance refer to [Höpcke 1980], where the emphasis is on adjustment theory, or [Koch 1988] for parameter estimation.

Least squares estimation problems are commonly described with a set of relations between idealized, unperturbed measurements, denoted by ℓ, and the unknown model parameters ζ. The most general model consists of a set of functions F_k, which we collect in a vector valued function

$$F(\zeta, \ell) = 0\,. \tag{3.1}$$

The unperturbed measurements ℓ are unknown. The actual measurements l, also called observations, are always subject to errors and only their expected value is equivalent to ℓ

$$\ell = \mathrm{E}[\,l\,]\,.$$

In order to account for these errors, the expected value ℓ is replaced by $\ell = l + e_l$. The vector e_l is called *residual vector*. Estimates of the measurements and parameters must not be confounded with the actual unknown values and are denoted by $\hat{e}_l$, $\hat{\ell}$, and $\hat{\zeta}$, respectively.

The measurements are often at least partly independent of the parameters. With regard to the image matching problem, equation (3.1) is formulated as a sum of the dependent and independent observations. Denoting the independent observations with l_2 and the others with l_1, we write

$$F_1(\zeta, l_1 + e_{l1}) + F_2(l_2 + e_{l2}) = 0 \tag{3.2}$$

without loss of generality, since the vector valued functions $F_{1,2}$ can both take on zero values.

The residuals can also be introduced in the space of the function vector F instead of the observation space of l. However, there are substantial disadvantages in some circumstances and a few thoughts are necessary before such simplification. Solving (3.2) in the least squares sense results in a *best linear unbiased estimator*[1] for ζ. If we further assume normally distributed observations, it turns into a maximum likelihood estimator (MLE). These important properties are lost when changing the residual model to functional residuals. Nevertheless, under certain assumptions about the functions F_i, the additional bias can be neglected. One important prerequisite is that F_i must be linear in l_i, which we will always assume within this work. For a more detailed discussion about differences between these two residual models I refer to [Danuser and Stricker 1998]. Exploiting these assumptions, equation (3.2) simplifies to

$$\begin{aligned} F_1(\zeta, l_1) + e'_{l1} + F_2(l_2) + e'_{l2} &= 0 \\ F_1(\zeta, l_1) + F_2(l_2) + e_L &= 0 \end{aligned} \tag{3.3}$$

where we combine the residuals

$$e_L := e'_{l1} + e'_{l2}\,. \tag{3.4}$$

At this point, I also introduce the notion of the weight matrix P_{ll} and the cofactor matrix Q_{ll}. Denoting the covariance matrix of the measurements l with Σ_{ll}, their weight matrix is defined as multiple of the inverse of Σ_{ll}.

$$P_{ll} = \sigma_0^2\,\Sigma_{ll}^{-1} \tag{3.5}$$

The factor σ_0^2 is called the *variance of unit weight*, since with $P_{ll} := I$ follows $\Sigma_{ll} = \sigma_0^2\,I$. The cofactor matrix Q_{ll} is defined as the inverse of P_{ll}

$$Q_{ll} = P_{ll}^{-1} = \frac{1}{\sigma_0^2}\Sigma_{ll}\,. \tag{3.6}$$

Defining the covariance matrix of the measurements is an important modeling step. Usually, it is not possible to fully define Σ_{ll}, as the factor σ_0^2 is often unknown a priori. Nevertheless, the weight matrix P_{ll} is specified after an a priori analysis of the estimation problem. Therefore, the weights of the optimization are based on physical properties instead of heuristics, which is an important advantage over other optimization methods. In particular, the knowledge of the relative uncertainty of the input data renders data fusion and error propagation possible.

The least squares objective function aims at minimizing the squared sum of the residuals considering the different weights of the observations. Along with the combination of the residual defined in equation (3.4), a weight matrix P_{LL} and an unit

---

[1] Strictly, the estimator is only unbiased for linear problems. Applied to nonlinear models, the random errors cause an additional statistical bias [Box 1971], which can be neglected in most cases.

variance $\sigma_{\mathrm{L}0}^2$ are introduced for the combined observations. Thus, we can formulate the goal function as $\Theta = e_L^T P_{LL}\, e_L$. Notice that for uncorrelated observations, P_{LL} is a diagonal matrix.

Minimizing this goal function leads together with equation (3.3) to an unconstrained nonlinear least squares (NLS) problem

$$\text{minimize} \left[e_L^T P_{LL}\, e_L \right] \quad \text{where } \; e_L = -F_1(\zeta, l_1) - F_2(l_2)\,.$$

This optimization problem can either be solved using the Levenberg–Marquardt method [Marquardt 1963] or by applying a Gauss–Newton scheme. Please consult for instance [Gill *et al.* 1981, section 4.7] or [Press *et al.* 1994, pp. 683] for details. Both methods iteratively solve this equation using a linear approximation for equation (3.3), but the Levenberg–Marquardt method combines Gauss–Newton with the steepest-descent method. The differences of these two optimization methods are discussed in detail in section A.2. The linearization of equation (3.3) is further described in section 3.3.

Constraints on the model parameters ζ are formulated in analogy to equation (3.1). Again, the constraint functions G_k are collected in a vector valued function G:

$$G(\zeta) = 0\,. \tag{3.7}$$

The constraint NLS problem defined by equations (3.1) and (3.7) is either solved using a Lagrange formalism or by introducing the constraints as so-called zero or pseudo observations. In this work, I will concentrate on the latter, interpreting the constraints as zero observations with very small variances.

$$G(\zeta) + e_0 = 0\,, \tag{3.8}$$

denoting the residuals by e_0 analogous to e_L. Such constraints are called *soft* or *spring constraints*, in contrast to *hard constraints* when applying the Lagrange formalism.

In particular, constraints can be used to suppress the estimation of a certain parameter by fixing it to an a priori value. In this case, a constraint function G_k is formulated as $G_k(\zeta) = \zeta_i - \bar{\zeta}_i$, where $\bar{\zeta}_i$ is the a priori value for parameter ζ_i. Inserting into equation (3.8) it follows

$$\zeta_i - \bar{\zeta}_i + e_{0_k} = 0\,. \tag{3.9}$$

Another important case is tying two parameters together. It is straightforward to derive the corresponding constraint equation, which will force parameter ζ_i and ζ_j to be equal after the optimization

$$\zeta_i - \zeta_j + e_{0_k} = 0\,. \tag{3.10}$$

Using such constraints, implementation problems due to the varying number of parameters are avoided. This of course at the cost of a slightly less stable equation system and larger matrices.

All constraints G_k are added to the original observations using large corresponding weights P_0 in the augmented weight matrix P. These weights correspond to the relative accuracy compared to the variances of the estimated parameters as will be discussed in section 3.3. Each constraint $G_k(\zeta) = 0$ will then hold through the optimization with this specified accuracy. The augmented function vectors and the residual vector are denoted by

$$F_1^* := \begin{bmatrix} F_1 \\ G \end{bmatrix} , \quad F_2^* := \begin{bmatrix} F_2 \\ 0 \end{bmatrix} , \quad e := \begin{bmatrix} e_L \\ e_0 \end{bmatrix} ,$$

leading to a formally unconstrained NLS problem

$$\text{minimize} \left[e^T P e \right] \quad \text{where } e = -F_1^*(\zeta, l_1) - F_2^*(l_2) \tag{3.11}$$

Matrices P and Q represent the augmented weight matrix and cofactor matrix, respectively:

$$P := \begin{bmatrix} P_{LL} & 0 \\ 0 & P_0 \end{bmatrix} , \quad Q := \begin{bmatrix} Q_{LL} & 0 \\ 0 & Q_0 \end{bmatrix} ,$$

with $Q_0 = P_0^{-1}$ being the cofactor matrix of the constraints.

3.3 Linearization and error propagation

One major advantage of the least squares framework is the possibility of rigorous error propagation. Equation (3.11) defines a NLS estimation problem which can be linearized around the current estimate $\hat{\zeta}^\circ$. Note that we already assumed F_i^* to be linear in l_i when the residual e was introduced. Thus, the derivatives $\nabla_\ell F_{1,2}^*$ reduce to constants and are included in e.

$$F_1^*(\hat{\zeta}^\circ, l_1) + A\,\Delta\zeta + F_2^*(l_2) + e = 0 \tag{3.12}$$

Matrix A represents the Jacobian $\nabla_\zeta F_1^*$ and $\Delta\zeta = \hat{\zeta} - \hat{\zeta}^\circ$. Denoting the functional residuals by $w = F_1^*(\hat{\zeta}^\circ, l_1) + F_2^*(l_2)$ and solving for the residuals e, the following expression for the goal function is found

$$e^T P e = \Delta\zeta^T A^T P A\,\Delta\zeta + 2\,\Delta\zeta^T A^T P\,w + w^T P w\,, \tag{3.13}$$

which is solved by setting the first derivative with respect to the variable $\Delta\zeta$ to zero

$$A^T P A\,\Delta\zeta + A^T P\,w = 0\,. \tag{3.14}$$

Thus, the estimates for the parameter change and the residuals are given by

$$\begin{aligned} \widehat{\Delta\zeta} &= -(A^T P A)^{-1}\,A^T P\,w \\ \hat{e} &= -A\,\widehat{\Delta\zeta} - w \end{aligned}$$

The statistics of the parameter estimate $\hat{\zeta} := \hat{\zeta}^\circ + \widehat{\Delta\zeta}$ in the adjusted system is the same as the statistics of the parameter update $\widehat{\Delta\zeta}$. Therefore, we can directly compute the cofactor matrix $Q_{\hat{\zeta}\hat{\zeta}}$. To propagate the covariances from the observations to the estimated parameters, the following error propagation theorem is used:

Theorem (error propagation). The covariance matrix Σ_{yy} of a linearly transformed random vector $y = Hx + c$ is given by $\Sigma_{yy} = H\,\Sigma_{xx}\,H^T$ (cf. for instance [Koch 1988, pp. 115]).

Thus, we derive the cofactor matrix $Q_{\hat{\varsigma}\hat{\varsigma}}$ using the properties $Q_{ww} = Q = P^{-1}$ and $Q^T = Q$

$$\begin{aligned} Q_{\hat{\varsigma}\hat{\varsigma}} &= (A^TPA)^{-1}\,A^TP\,Q\,((A^TPA)^{-1}\,A^TP)^T \\ Q_{\hat{\varsigma}\hat{\varsigma}} &= (A^TPA)^{-1} \end{aligned} \tag{3.15}$$

and a similar expression is found for the cofactor matrix of the residuals $Q_{\hat{e}\hat{e}}$:

$$Q_{\hat{e}\hat{e}} = Q - A\,(A^TPA)^{-1}\,A^T\,. \tag{3.16}$$

3.4 Diagnostic measures

The strict least squares framework allows for a number of diagnostic measures, for instance to judge the goodness of fit or the parameter determinability. Hence, various tests are applied to supervise the iteration progress and the final result, yielding self-diagnosis of the framework. The following section outlines the employed measures. Details and proofs about statistics and hypothesis testing in general are found in many textbooks, for example [Papoulis 1991] or [Koch 1988] for hypothesis testing in linear models.

3.4.1 A posteriori noise estimate

After the completed adjustment, the a posteriori noise estimate $\hat{\sigma}_{\mathrm{L}0}$ of the combined observations is computed using

$$\hat{\sigma}_{\mathrm{L}0}^2 = \frac{\hat{e}_L^T P_{LL}\,\hat{e}_L}{n-r}\,, \tag{3.17}$$

where n is the number of observations and r is the number of parameters or degrees of freedom. Please recall from equation (3.3) that $\hat{e}_L$ is an estimate of the combined residual of the observations l_1 and l_2. Hence, $\hat{\sigma}_{\mathrm{L}0}$ is the a posteriori noise estimate of the *combined* observations.

The noise estimate of the *original* observations can only be computed with knowledge of the original noise processes. Two simple but common cases are of particular interest:

1. Combination of exact measurements with zero variance, for instance artificial model data, with the actual measurements of variance σ_0. In this case the a posteriori noise estimate $\hat{\sigma}_0$ of the actual measurements is identical to $\hat{\sigma}_{\mathrm{L}0}$:

   $$\hat{\sigma}_0 \equiv \hat{\sigma}_{\mathrm{L}0}\,.$$

2. Addition or subtraction of two normally distributed measurements with different variances σ_0 and $\lambda\sigma_0$. Then, the a priori noise of the combined observations amounts to

   $$\sigma_{\mathrm{L}0} = \sqrt{1+\lambda}\,\sigma_0$$

   and the a posteriori estimate $\hat{\sigma}_0$ equals to

   $$\hat{\sigma}_0 = \frac{\hat{\sigma}_{\mathrm{L}0}}{\sqrt{1+\lambda}}\,.$$

3.4.2 Global model test

It can be shown that $\hat{\sigma}_0$ is an unbiased estimator of σ_0 if and only if the model corresponds to the input data [Koch 1988]. But in most practical cases $\hat{\sigma}_0$ is biased since it also includes errors caused by an inaccurate model. A standard procedure in parameter estimation theory is to test the hypothesis $\hat{\sigma}_0 = \sigma_0$ against $\hat{\sigma}_0 > \sigma_0$ in order to check the correctness of the model:

$$H_0\colon\ \hat{\sigma}_0 = \sigma_0 \quad \text{against} \quad H_1\colon\ \hat{\sigma}_0 > \sigma_0 \quad \text{based on} \quad \mathbf{q} = (n-r)\,\frac{\hat{\sigma}_0^2}{\sigma_0^2}\,.$$

Under the assumption that the observations are normally distributed, the test statistic $\mathbf{q}$ is chi-square distributed with $n-r$ degrees of freedom: $\mathbf{q} \sim \chi^2(n-r)$. The hypothesis H_0 is accepted with a significance level of α if and only if

$$\mathbf{q} < \chi^2_{1-\alpha}(n-r)\,.$$

In some cases when $\hat{\sigma}_0 < \sigma_0$ must be considered an error, the hypothesis

$$H_0\colon\ \hat{\sigma}_0 = \sigma_0 \quad \text{against} \quad H_1\colon\ \hat{\sigma}_0 \neq \sigma_0 \quad \text{based on} \quad \mathbf{q} = (n-r)\,\frac{\hat{\sigma}_0^2}{\sigma_0^2}\,.$$

shall be tested instead, which is accepted with a significance level of α if and only if

$$\chi^2_{\alpha/2}(n-r) < \mathbf{q} < \chi^2_{1-\alpha/2}(n-r)\,.$$

However, either of the two measures do not include any normalization of the model data. Therefore it depends on the gray value scaling of the model compared to the noise level, that is the signal to noise ratio. Additionally, the a priori noise σ_0 must be known. These two problems are serious drawbacks especially in the case of image matching. On the one hand, it is often difficult to gain a reliable a priori estimate of the noise level. A possible estimation method was presented by [Voorhess and Poggio 1987], but it starts to fail when too many features are present in the image. On the other hand, signal to noise ratios may vary significantly. For these reasons, the goodness of fit is tested using the cross correlation value instead, as described in the following section.

3.4.3 Cross correlation value

In the adjusted state, that is after completing the optimization, the cross correlation between the model and the observed data represents a criterion for the goodness of fit. It describes the linear correlation of two random variables and is mostly used in its normalized form. For the sake of completeness the formula is repeated here for the correlation of $-F_1$ with F_2:

$$\rho(-F_1, F_1) = -\frac{E[F_1 F_2] - E[F_1]E[F_2]}{\sqrt{E[F_1^2]\,E[F_2^2]}} \,.$$

At the correct estimate, this value is very close to 1.0 and therefore suitable to discriminate wrong results from correct estimates. The application to LSM is described in more detail in section 3.11.

3.4.4 Estimation of parameter accuracy

Parameter estimation in linear least squares problems is extensively discussed in standard literature on parameter estimation theory (for instance [Koch 1988]). The iterative solution of equation (3.3) is an unbiased estimate for the unknowns with a stochastic variance expressed by the diagonal elements of the covariance matrix

$$\Sigma_{\hat{\zeta}\hat{\zeta}} = \hat{\sigma}_0 \cdot Q_{\hat{\zeta}\hat{\zeta}} \,. \tag{3.18}$$

The value $\hat{\sigma}_0$ denotes the a posteriori noise estimate computed using equation (3.17) and $Q_{\hat{\zeta}\hat{\zeta}}$ is the cofactor matrix as defined in the equation (3.15). It is important to note that this formula strongly relies on a correct weight matrix P. Please refer to section 3.10 for a description of the problems arising from an incorrect weight matrix and how to avoid them.

3.4.5 Determinability analysis

Special attention has been paid to the analysis of the parameter determinability. [Danuser and Mazza 1996] proposed to test the relative contribution δ_i to the trace of the cofactor matrix $Q_{\hat{\zeta}\hat{\zeta}}$

$$\delta_i^* = \frac{\left|\mathrm{tr}[Q_{\hat{\zeta}\hat{\zeta}}] - \mathrm{tr}[Q_{\hat{\zeta}\hat{\zeta}}^i]\right|}{\mathrm{tr}[Q_{\hat{\zeta}\hat{\zeta}}]} \,, \tag{3.19}$$

where $Q_{\hat{\zeta}\hat{\zeta}}^i$ is the cofactor matrix with parameter i excluded.

This measure stems from the idea to use the trace of the cofactor matrix $Q_{\hat{\zeta}\hat{\zeta}}$ as basis for a measure of the overall precision of the least squares estimation problem [Grün 1986].[2] Since the difference of the traces represent a change in precision, δ_i describes the influence of each parameter to the overall precision and thus is

---

[2] In particular he describes the mean variance $\frac{1}{r}\,\mathrm{tr}[Q_{\hat{\zeta}\hat{\zeta}}]$ as a suitable precision measure.

coupled with the influence of each parameter upon the others. However, a few thoughts are necessary to describe this influence in more detail and to understand why equation (3.19) has to be adapted for our application.

In order to exclude one parameter, the constraint $\zeta_i - \bar{\zeta}_i = 0$ is added in the form of an additional zero observation as described in section 3.2, equation (3.9). We will assume a large weight of $p_{n+1} \rightarrow \infty$ for this constraint, which automatically leads to a zero variance for the excluded parameter i. This variance difference appears of course in the measure presented in equation (3.19), which is not desirable since δ_i would be dependent of the original cofactor $(q_{\hat{\zeta}\hat{\zeta}})_{ii}$ of parameter i. Thus, the corrected equation for the contribution is written as

$$\delta_i = \frac{\mathrm{tr}^{\mathrm{i}}[Q_{\hat{\zeta}\hat{\zeta}}] - \mathrm{tr}^{\mathrm{i}}[Q^i_{\hat{\zeta}\hat{\zeta}}]}{\mathrm{tr}[Q_{\hat{\zeta}\hat{\zeta}}]} \, , \tag{3.20}$$

where $\mathrm{tr}^{\mathrm{i}}[.]$ stands for the trace omitting element (i, i) and where we also took advantage of the fact that the trace of the cofactor matrix—which is the sum of the variances—will always decrease when excluding a parameter. This latter expression for the contribution δ_i describes the influence of each parameter upon the precision of the others. Weakly determinable parameters cause quasi-singular normal equations, thus a large change in the trace of the cofactor matrix.

Strictly, the contribution value is only meaningful in the adjusted state, that is after optimization. Still it is useful before and between iteration steps employing a multistage approach. The application of such a procedure to LSM is explained in section 3.11. More details on the contribution values are shown in the result section 3.12.

The influence of this additional uncorrelated observation to the original normal equation system is derived by applying the framework of the Kalman-Bucy filter technique for recursive parameter estimation (see for instance [Koch 1988]). Hence, the partial cofactor matrix $Q^i_{\hat{\zeta}\hat{\zeta}}$ can be directly computed from $Q_{\hat{\zeta}\hat{\zeta}}$:

$$Q^i_{\hat{\zeta}\hat{\zeta}} = Q_{\hat{\zeta}\hat{\zeta}} - Q_{\hat{\zeta}\hat{\zeta}} A^T_{n+1} \left(A_{n+1} Q_{\hat{\zeta}\hat{\zeta}} A^T_{n+1} \right)^{-1} A_{n+1} Q_{\hat{\zeta}\hat{\zeta}} \, , \tag{3.21}$$

under the assumption that the constraint was introduced with a weight $p_{n+1} \rightarrow \infty$. Applying this framework to (3.20) yields a simple expression for the contribution value

$$\delta_i = \frac{\sum_{j \neq i} {q_{ij}}^2}{q_{ii} \sum_j {q_{jj}}^2} \, , \tag{3.22}$$

where q_{ij} denotes the elements of the full cofactor matrix $Q_{\hat{\zeta}\hat{\zeta}}$.

If this contribution δ_i is high, this parameter strongly influences the estimation of one or more parameters. One might consider excluding parameter i by applying the formalism for parameter constraints described in section 3.2. However, three important aspects have to be considered when using the contribution δ_i as determinability measure.

1. It is only a suitable measure for parameter sets which are more or less uncorrelated by design. In the case of correlated parameters, the exclusion of one

single parameter obviously influences its correlated counterparts and probably leads to a high contribution value δ_i. Depending on the degree of correlation, excluding such a parameter might actually weaken the determinability or at least lead to the undesired result of quasi excluding the correlated parameters as well. More suitable would be to test the contribution when tying correlated parameters together, which can be achieved by a similar framework. However, the contribution value does not reveal any information about which parameters are correlated with each other. Thus, the parameter correlations must be investigated in more detail, which is explained in the following section.

2. The trace of the cofactor matrix is only then a meaningful measure, if the variances of the parameters are of the same magnitude. We will see in section 3.11 that this is not true for affine transformation parameters, for instance. In such cases, either parameter subgroups are tested or the correlation matrix $C_{\hat{\zeta}\hat{\zeta}}$ must be used instead of the cofactor matrix.

3. The contribution value δ_i is a sum of local measures and not a global measure. That is, no global determinability is tested. This may lead to a high contribution value for a parameter which would be determinable from a global point of view but not anywhere locally. An illustration of this fact is found in figure 3.11.

4. It might not be a good idea to exclude one single parameter after all. A good example is a set of transformation parameters which define an affine transformation. The exclusion of one of the scaling parameters for instance automatically inhibits the estimation of a rotation as well. Thus, it is important to define a suitable reduction of the parameter set if the full set is not determinable.

3.4.6 Parameter correlation analysis

The matrix of parameter correlation is computed by normalizing the elements of the cofactor matrix $Q_{\hat{\zeta}\hat{\zeta}}$ by the corresponding diagonal elements:

$$(c_{\hat{\zeta}\hat{\zeta}})_{ij} = \frac{q_{ij}}{\sqrt{q_{ii}\,q_{jj}}} .$$

The resulting correlation matrix $C_{\hat{\zeta}\hat{\zeta}}$ corresponds to the cofactor matrix of a transformed parameter set $\zeta^* = [\frac{\zeta_1}{q_{11}}, \dots, \frac{\zeta_i}{q_{ii}}, \dots, \frac{\zeta_r}{q_{rr}}]$.

The strategy how to react on high correlations strongly depends on the application. Often, it is possible to anticipate potentially correlated parameters, as in the case of transformation parameters. Please refer to section 3.11 for a possible strategy.

3.4.7 Residual analysis: local redundancies

The matrix $R = Q_{\hat{e}\hat{e}}\, Q^{-1}$ can be used to compute local redundancies [Förstner 1987]. Note that the cofactor matrix $Q_{\hat{e}\hat{e}}$ of the residuals is always of full rank, even if the observations are uncorrelated. Since the number of observation is typically large, the computation of the full matrix $Q_{\hat{e}\hat{e}}$ is not feasible. Nevertheless, only the diagonal elements of R have to be known in order to compute local redundancies. From equation (3.16)

$$\begin{aligned} Q_{\hat{e}\hat{e}}\, Q^{-1} &= I - A\,(A^T P A)^{-1}\, A^T Q^{-1} \\ R &= I - A\, Q_{\hat{\zeta}\hat{\zeta}}\, A^T Q^{-1}\ , \end{aligned}$$

the following expression for a diagonal element r_{kk} is found for uncorrelated observations, that is if Q is diagonal:

$$r_{kk} = 1 - A_k\, Q_{\hat{\zeta}\hat{\zeta}}\, A_k^T\, q_{kk}\ ,$$

where A_k is row k of matrix A. Thus, is is possible to compute the diagonal elements of R without having to build up matrix A, analogous to building the normal matrix N. The trace of R, that is the sum of all local redundancies, equals to the overall redundancy of the least squares problem $n-r$ where n is the number of observations and r the number of parameters.

$$\mathrm{tr}[Q_{\hat{e}\hat{e}}\, Q^{-1}] = n - r\ .$$

Thus, the values r_{kk} describe how the redundancy is distributed over the observations. Even if it is not meaningful to analyze the redundancy of each observation, it is often helpful to compute group redundancies of certain observation groups.

3.5 Unconstrained LSM

The LSM includes two observations, the template image $f[.]$ and the search image $g[.]$, called patch. While the template is independent of the model parameters ζ, the patch $g[.]$ cannot be separated. Applying strictly the least squares framework of equation (3.2), this leads to the equation

$$F(\zeta, g{+}e_g) + f + e_f = 0\ ,$$

where f correspond to l_2 and g to l_1. In the case of LSM, where the function F consists of a geometric and a linear radiometric transformation of g, the aforementioned simplification of applying the residuals in the transformed observation space does not considerably influence the result. Hence, equation (3.3) is used instead:

$$F(\zeta, g) + f + e = 0\ . \tag{3.23}$$

The main modeling step lies in the definition of the function F. The simultaneous estimation of both types of transformations (geometric and radiometric) would lead

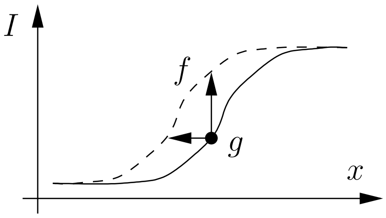


Figure 3.1: The ambiguity between radiometric and geometric transformation. Local operators cannot distinguish between the two types of transformations, hence a simultaneous estimation could lead to an ill-conditioned system.

to an overdetermined system, since it is not possible to distinguish them locally (cf. figure 3.1). In order to overcome this problem, the parameters of the radiometric transformation are estimated based on a global measure within the template region apart from the actual least squares optimization. The resulting radiometrically adjusted patch $\bar{g}[.]$ is then used for the next optimization step. In the case of a linear transformation this can be written as

$$\mathcal{T}_{\mathcal{R}} : g \longrightarrow \bar{g}\,, \quad \text{where} \quad \bar{g}[.] = \alpha + \beta\, g[.]\,. \tag{3.24}$$

The radiometric parameters α and β are then treated as known values and hence equation (3.23) becomes

$$\bar{F}(\xi, \bar{g}) + f + e = 0\,, \tag{3.25}$$

writing ξ for the remaining parameters.

The geometric relation between the original template and the matched area is defined by an arbitrary transformation. Depending on the type of the chosen transformation, this allows for displacement, rotation and/or deformation of the template. In the general case, the image coordinates u are transformed using the parameter vector ξ to

$$x = \psi(\xi, u)\,. \tag{3.26}$$

Thus, the function $\bar{F}$ is defined as

$$\bar{F}(\xi, \bar{g}) := -\,\bar{g}(\psi(\xi, u))$$

which leads together with equation (3.25) to the final form of the LSM observation equations

$$f[u] + e[u] = \bar{g}(x)\,. \tag{3.27}$$

Equation (3.27) represents a relation between each gray value within the template and its corresponding image intensity in the search image. Notice that square brackets denote functions defined on a discrete grid. The functions $g(.)$ and $\bar{g}(.)$ simply represent the continuous versions of $g[.]$ and $\bar{g}[.]$, respectively. Hence, the template gray values $f[u]$ are defined on the grid of the reference image while $\bar{g}(x) = \bar{g}(\psi(\xi, u))$ fall between the grid of the search image. Interpolating the gray values $\bar{g}(\psi(\xi, u))$ for a given ξ we substitute

$$\tilde{g}_\xi[u] := \bar{g}(\psi(\xi, u))\,.$$

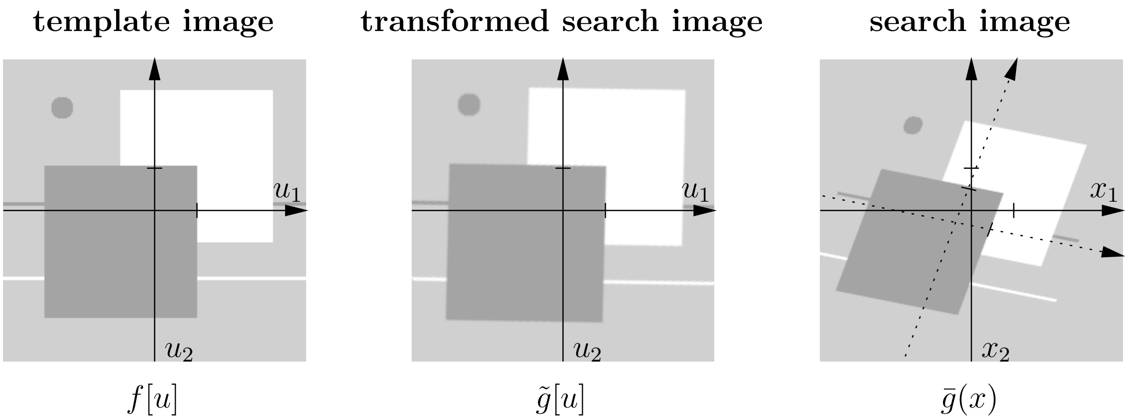


Figure 3.2: Coordinate system transformation between template f and radiometrically adjusted search image $\bar{g}$. The vector u stands for the discrete image coordinates, x denotes the transformed coordinates as defined in equation (3.26).

Based on a coordinate list $u[k]$, equation (3.27) is reordered into a vector notation

$$f + e = \tilde{g}\,, \tag{3.28}$$

building a series of n equations, where n is the number of pixels included in the template and $k = 1 \dots n$. Together with the least squares objective function $e^T P e$ this defines an unconstrained NLS problem. Following the framework described in section 3.3, equations (3.12) to (3.15), this nonlinear problem is iteratively solved using a Gauss–Newton scheme. A new estimate $\hat{\xi} = \xi^\circ + \Delta\xi$ is computed linearizing the observation equations around the current estimate ξ° or $x^\circ = \psi(\xi^\circ, u)$, respectively:

$$\begin{aligned} f[u] + e[u] &= \bar{g}(x^\circ) + \nabla_\xi\, \bar{g}(x^\circ)\, \Delta\xi && (3.29)\\ f + e &= \tilde{g}^\circ + A \cdot \Delta\xi\,. && (3.30) \end{aligned}$$

Matrix A is the $n \times r$ Jacobian matrix with respect to the parameter vector ξ where r denotes the number of parameters.

If we neglect the stochastic nature of the Jacobian matrix, this linear problem corresponds to a Gauss-Markov model with full rank. Note that only under this assumption, the observations are separated from the parameters. The linear problem (3.30) is solved analytically setting the first derivative of the least squares goal function $e^T P e$ to zero, which yields the normal equation system

$$\begin{aligned} A^T P A \cdot \Delta\xi &= -A^T P\, (\tilde{g}^\circ - f) && (3.31)\\ N \cdot \Delta\xi &= -A^T P\, w\,. && (3.32) \end{aligned}$$

Notice that if P is a diagonal or a band-diagonal matrix, the matrix A does not have to be computed and stored as a whole. The weight matrix P is diagonal if and only if the observations are independent of each other. This is usually assumed for the original observations f and g. However, this assumption is not fulfilled for the transformed and interpolated patch $\tilde{g}$ or for Gaussian filtered observations. The statistics of such transformed observations are investigated in more detail in section 3.10 for the case of two dimensional LSM.

An estimate of the influence of an incorrect weight matrix is given in [Koch 1988]. For the sake of completeness, the result of this analysis is described here; more details can be found in the original work. Denoting the correct weight matrix with P, we choose an incorrect weight matrix $P + \Delta P$. We further assume the elements of ΔP to be small such that products can be neglected where ΔP appears twice. This leads to the following expression for the actual parameter estimate

$$A^T P A \cdot \Delta\tilde{\xi} = -A^T P w - A^T \Delta P w + A^T \Delta P A \, (A^T P A)^{-1} A^T P w \, . \tag{3.33}$$

Nevertheless, it is common to choose a diagonal weight matrix P, since the influence of neglecting the statistical dependence becomes only significant for high precision measurements. Therefore, it is sufficient to compute A row by row in order to build the normal matrix N. This is an important property, since the size of A increases with the number of pixels included in the templates.

As long as A has no row deficiency, the $r \times r$ normal matrix $A^T P A$ is always positive definite and symmetric and hence the Cholesky decomposition can be applied to solve equation (3.31). After each iteration step, matrix A must be recomputed using the updated set of parameters $\xi^{t+1} = \xi^t + \Delta\xi$. When the parameter change $\Delta\xi$ falls below a specified numerical resolution the iteration process is stopped.

3.6 Multi template extension

Using multiple templates instead of one large template allows significant and stable regions to be selected without including regions unsuitable for matching. This is an important aspect for the application to megavoltage X-ray imagery and will be discussed in chapter 5.

There are several ways to extend the standard LSM to multiple templates. For all of them, the equation (3.27) has to be adapted to include multiple templates and their corresponding patches, indicated in the following with a superscript $K = 1 \ldots N$. Denoting the different areas of each template by u^K, the observation equations are thus written as

$$f[u^K] + e[u^K] = \bar{g}^K(x^K) \, , \tag{3.34}$$

with the radiometric corrections

$$\mathcal{T}_{\mathcal{R}}^K : g \longrightarrow \bar{g}^K \, , \quad \text{where} \quad \bar{g}^K[.] = \alpha^K + \beta^K g[.] \tag{3.35}$$

and the geometric transformations

$$x^K = \psi(\xi^K, u^K) \, , \tag{3.36}$$

analogous to equations (3.24) and (3.26). Of course it is also possible to define a different geometric transformation function ψ^K for each template. However, it will usually be sufficient to employ the same function with different parameter sets as written in equation (3.36).

The observations can be reordered into a vector notation in the same way as for the single template matching (3.28) using the coordinate lists $u^K[k]$. The vector notations $f^K + e^K = \tilde{g}^K$ for each template are combined into one equation

$$\begin{pmatrix} f^1 \\ f^2 \\ \vdots \\ f^N \end{pmatrix} + \begin{pmatrix} e^1 \\ e^2 \\ \vdots \\ e^N \end{pmatrix} = \begin{pmatrix} \tilde{g}^1 \\ \tilde{g}^2 \\ \vdots \\ \tilde{g}^N \end{pmatrix} \tag{3.37}$$

3.6.1 Single global transformation

The most straightforward extension is to keep one single transformation for all patches such that $\xi^K \equiv \xi$, which leads to

$$x^K = \psi(\xi, u^K) \,. \tag{3.38}$$

Formally, this procedure is equal to defining one large template with several scattered regions of interest. However, since the radiometric parameters α^K and β^K may vary between the templates, it is possible to compensate for global gray value differences like bias fields.

3.6.2 Global and local transformations

A more general approach is to assign a separate set of parameters ξ^K to each patch and add global soft constraints. This will result in a blockwise filled Jacobian matrix A as depicted in figure 3.3.

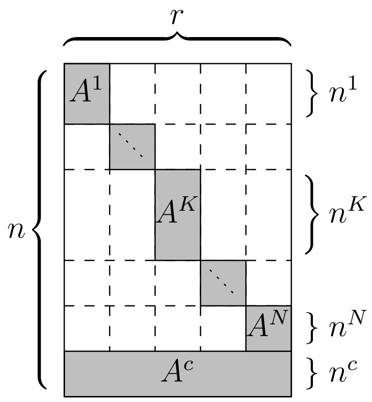


Figure 3.3: Illustration of the Jacobian matrix A when using different parameter sets for each template and soft constraints. A^K denote the submatrices of each template while additional soft constraints are represented by A^c.

For the following analysis on the sparsity of the normal matrix A^TPA, the weight matrix P is assumed to be a diagonal matrix, which is true for many practical cases. In order to further simplify the expressions, the weight matrix is set to the identity matrix, since the weight factors have no qualitative influence on the result.

Figure 3.4 shows the two parts of the normal matrix: the blockdiagonal part of the observations and the constraint part $(A^c)^T(A^c)$ which is in general a full matrix.

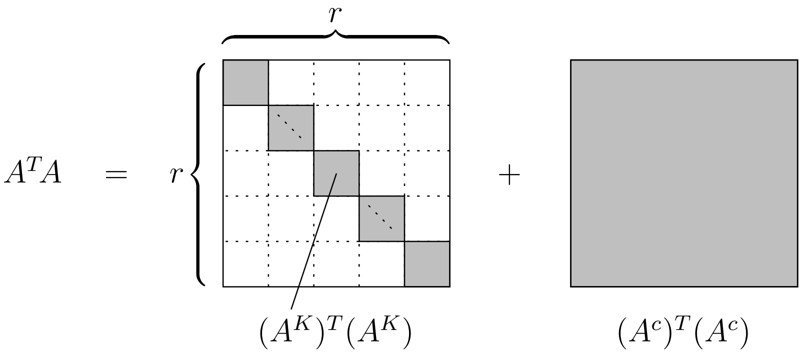


Figure 3.4: Illustration of the normal matrix $N = A^TA$ when using different parameter sets for each template and soft constraints. N can be split into an observation part (left) and a constraint part (right).

3.7 Affine transformation as geometric transformation

So far no assumptions have been made on the dimensionality of the problem and on what type of transformation is used. In the following, the case of a two dimensional affine transformation is presented. The corresponding parameter vector consists of six variables $\xi = [t_1, t_2, m_1, s_1, s_2, m_2]^T$ and the coordinate transformation is written as[3]

$$x = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} + \begin{bmatrix} m_1 & s_1 \\ s_2 & m_2 \end{bmatrix} u \, . \tag{3.39}$$

The derivative $\nabla_\xi \, \bar{g}(x)$ is then calculated explicitly using the chain rule:

$$\begin{aligned} \nabla_\xi \, \bar{g}(x) &= \left(\frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \frac{\partial}{\partial m_1}, \frac{\partial}{\partial s_1}, \frac{\partial}{\partial s_2}, \frac{\partial}{\partial m_2}\right) \bar{g}(x) \\ &= \left[\frac{\partial \bar{g}}{\partial x_1}, \frac{\partial \bar{g}}{\partial x_2}, \frac{\partial \bar{g}}{\partial x_1} u_1, \frac{\partial \bar{g}}{\partial x_1} u_2, \frac{\partial \bar{g}}{\partial x_2} u_1, \frac{\partial \bar{g}}{\partial x_2} u_2\right] \, . \end{aligned}$$

In vector notation, this leads to the $n \times 6$ Jacobian matrix A (cf. equation (3.30)), each row A_k representing the derivatives at $x_k = \psi(\xi, u[k])$. Denoting the derivatives by $\tilde{g}_{x_1}()$ and $\tilde{g}_{x_2}()$, we write

$$A_k = [\bar{g}_{x_1}, \, \bar{g}_{x_2}, \, \bar{g}_{x_1} u_1, \, \bar{g}_{x_1} u_2, \, \bar{g}_{x_2} u_1, \, \bar{g}_{x_2} u_2] \ ,$$

---

[3] The parameter names t_i, m_i and s_i were chosen based on their interpretation as translation, magnification and shear, respectively.

leaving out the parameter $x[k]$ of the functions $\bar{g}_{x_i}(.)$ for better readability. Using the resampled patch image $\tilde{g}[.]$ instead of $\bar{g}(.)$, the final form for the implementation is achieved

$$A_k = [\tilde{g}_{u_1},\, \tilde{g}_{u_2},\, \tilde{g}_{u_1}u_1,\, \tilde{g}_{u_1}u_2,\, \tilde{g}_{u_2}u_1,\, \tilde{g}_{u_2}u_2] \ ,$$

again omitting the parameter $u[k]$ of the functions $\tilde{g}_{u_i}[.]$.

As mentioned in the previous section, the normal matrix $N = A^T PA$ can be built computing A row by row, as long as the weight matrix P is diagonal. If the gray values of each pixel are considered independent this is fulfilled and N is computed without the need to multiply large matrices.

3.8 Linear constraints

As we have seen in section 3.2, the least squares formalism allows one to introduce additional constraints in a simple and intuitive way. In addition to the observation equations, zero observations are included in the framework, which results in soft or spring constraints. Equation (3.9) and (3.10) give two examples of such constraint equations.

In this section, I will apply this technique to LSM. As an examples serves the reduction of an affine to a similarity transformation. Instead of reparametrization, we still employ equation (3.39) as transformation equation and add the following constraints to the parameter vector ξ:

$$\begin{aligned} m_1 - m_2 + e_m &= 0 \\ s_1 + s_2 + e_s &= 0 \end{aligned} \tag{3.40}$$

The inclusion of the constraints in the normal equation system (3.32) strictly follows the framework shown in section 3.2. Analogous to the observation equations (3.27), the constraints are linearized around the current estimates m_i° and s_i°. Thus, the matrix A is augmented by the constraint vectors

$$\begin{aligned} A_m &= [0, 0, 1, 0, 0, -1] \\ A_s &= [0, 0, 0, 1, 1, \ \ \ 0] \end{aligned}$$

and the corresponding residuals are given by $w_m = -m_1^\circ + m_2^\circ$ and $w_s = -s_1^\circ - s_2^\circ$.

3.9 Nonlinear constraints

Nonlinear constraints must be linearized in order to include them in the least squares framework. This is not trivial in the general case and often increases the risk of oscillation in parameter space. Still, this problem can be solved for special cases, in particular for the constraint

$$m_1^2 + s_1^2 - 1 + e_c = 0 \ , \tag{3.41}$$

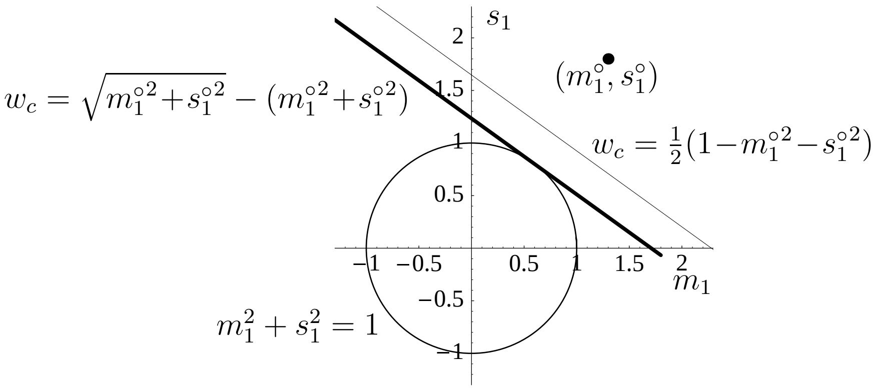


Figure 3.5: Graphical interpretation of the uncorrected linearized constraint (thin line) and corrected version (thick line). Oscillations are greatly reduced when using the corrected version.

which is used together with constraints (3.40) for estimating a congruent transformation. The straightforward linearization around the current estimates m_1° and s_1° yields the linear constraint equation

$$2\,\Delta m_1\, m_1^\circ + 2\,\Delta s_1\, s_1^\circ - (1-m_1^{\circ 2}-s_1^{\circ 2}) + e_c = 0\,,$$

leading to the constraint vector $A_c = [0, 0, m_1^\circ, s_1^\circ, 0, 0]$ with a residual of $w_c = \frac{1}{2}(1-m_1^{\circ 2}-s_1^{\circ 2})$.

However, this is not optimal. The thin line in figure 3.5 illustrates the constraint line defined by this equation. Clearly, the simple linearized constraint is biased and thus often causes oscillation during optimization. An improved version is achieved by adding a correction term to the residual, computed by using simple geometric relations (figure 3.5, thick line). This adjustment results in the improved residual

$$w_c = \sqrt{m_1^{\circ 2} + s_1^{\circ 2}} - (m_1^{\circ 2} + s_1^{\circ 2})\,. \tag{3.42}$$

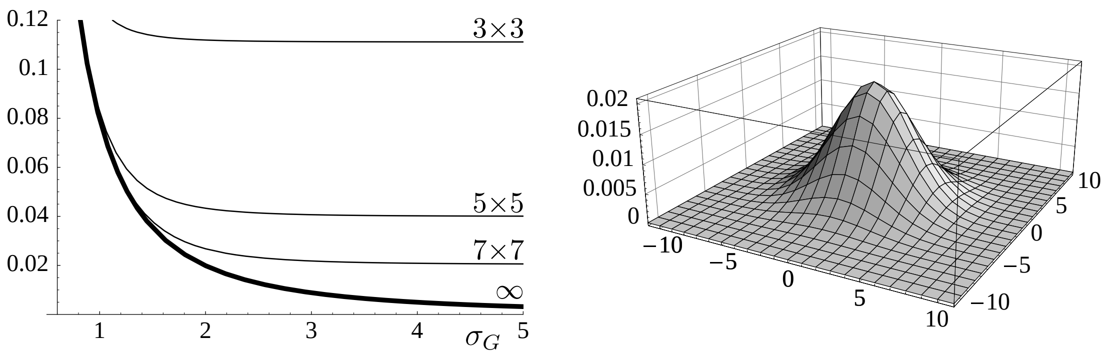
3.10 Correct design of the weight matrix

In order to investigate the actual variances and covariances of an interpolated or Gaussian filtered image, we will assume that all pixel values are independent and normally distributed with an identical variance σ_0^2 and varying mean values μ_k

$$f, g \sim \mathcal{N}(\mu, \sigma_0^2 I)\,. \tag{3.43}$$

3.10.1 Variance and covariance after Gaussian filtering

The main effect of Gaussian filtering is the reduction of the variance of the input signal. A detailed analysis of this noise reduction can be found in section A.3, while the most important findings are summarized in this section.



a. Variance **b.** Covariance

Figure 3.6: Figure (a) depicts the variance after Gaussian filtering depending on the filter width σ_G and with different filter support. In (b), the covariance values between two filtered values are shown for a 11×11 filter with $\sigma_G = 2$.

For a discrete Gaussian filter of width σ_G and limited support $(2m+1) \times (2m+1)$, the variance reduction can be written as

$$\sigma_0^{*2}(m, \sigma_G) = \sigma_0^2 \; \frac{\sum_{i=-m}^{m} \sum_{j=-m}^{m} e^{\frac{-(i^2+j^2)}{\sigma_G}}}{\left(\sum_{i=-m}^{m} \sum_{j=-m}^{m} e^{\frac{-(i^2+j^2)}{2\sigma_G}}\right)^2} \, ,$$

always under the assumption (3.43). The term in the denominator is a correction factor in order to avoid amplification of the image signal. Interesting limits of this expression and formulas for specific support sizes are discussed in section A.3.

A simple approximation can be computed under the assumption that $\sigma_G > 0.6$ and that m is sufficiently large (at least about $2\,\sigma_G$):

$$\sigma_0^{*2}(\sigma_G, m) \approx \frac{\sigma_0^2}{4\pi\sigma_G^2} \qquad \text{for } \sigma_G \geq 0.6\,. \tag{3.44}$$

Figure 3.6 illustrates the variance reduction for the common support sizes 3×3, 5×5, and 7×7 pixels. For each size, the maximum reduction corresponds to applying a box filter, that is $\sigma_0^{*2} = \sigma_0^2/(2m+1)^2$.

Analogous, the covariance of a discrete Gaussian filter is approximated by

$$\sigma_{xy}^{*\,2}(\sigma_G, d) \approx \frac{\sigma_0^2}{4\pi\,\sigma_G^2}\; e^{-\frac{d_1^2+d_2^2}{4\,\sigma_G^2}} \qquad \text{for } \sigma_G \geq 0.6\,, \tag{3.45}$$

where d is the distance between the two pixels. This distribution corresponds to a Gaussian distribution with a standard deviation of $\sqrt{2}\,\sigma_G$ and is illustrated in figure 3.6b for a 11×11 filter with $\sigma_G = 2$.

3.10.2 Variance and covariance after bilinear interpolation

Bilinear interpolated values are computed by a weighted sum of the four neighboring pixels. See equation (A.9) for the complete formula. Exploiting assumption (3.43),

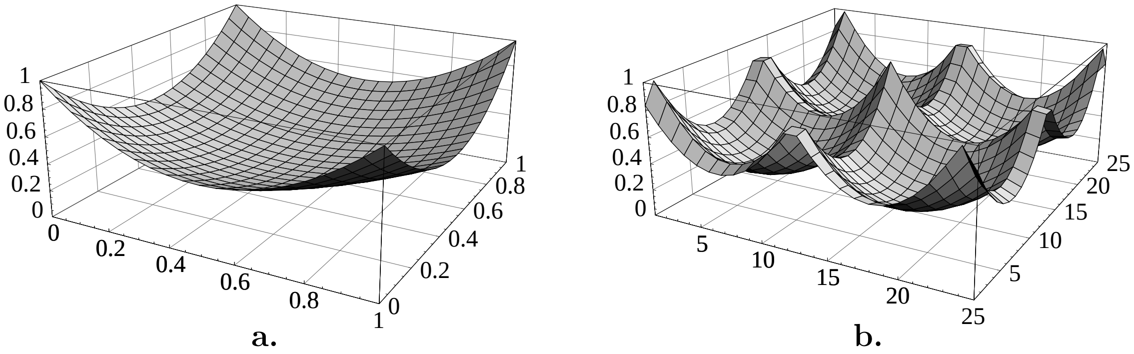


Figure 3.7: Variance of an interpolated pixel depending on the fractional part of its coordinate (a) and the distribution of variances within a resampled image under a rotation of 5° (b).

the following simple expression for the actual standard deviation σ_0^* of an interpolated value $f(x = \lfloor x \rfloor + d)$ is found:

$$\sigma_0^{*2}(d) = \sigma_0^2 \left(1 - 2\,d_1 + 2\,{d_1}^2\right) \left(1 - 2\,d_2 + 2\,{d_2}^2\right) . \tag{3.46}$$

Figure 3.7a illustrates this influence of bilinear interpolation to the pixel variance. At integer coordinates (the four corners in figure 3.7a), the variance of course equals the original variance σ_0^2, which was set to 1 for this illustration. As expected, the lowest value is attained in the center and equals $\sigma_0^2/4$. More interesting is the distribution of variances within an image after interpolation. Figure 3.7b depicts the situation after resampling a rotated image.

In the same manner, the covariance between two interpolated pixels is computed. Please refer to section A.4 for a detailed analysis. An interesting special case is the covariance after a pure translation, which is displayed in figure 3.8a. In this case, both coordinates have equal fractional parts d but they are one pixel apart. The resulting covariance is written as

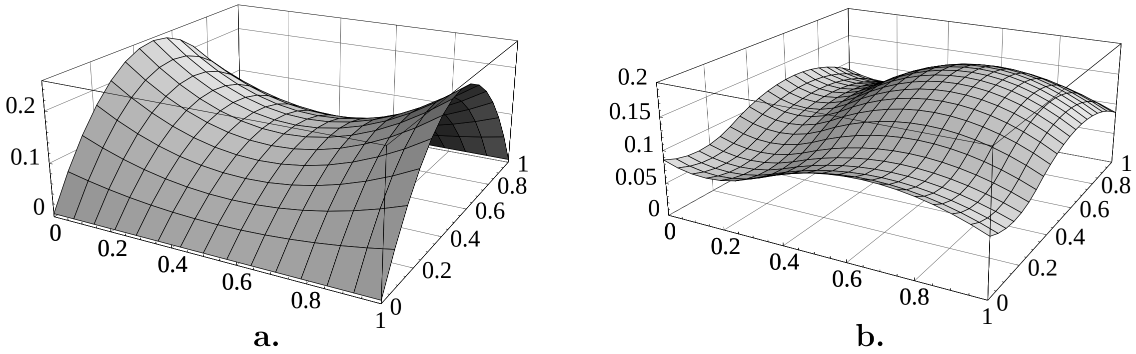
$$\sigma^{*\,2}_{xy,2}(d) = \sigma_0^2\,(-2{d_1}^2 + 2d_1 - 1)\,(d_2 - 1)\,d_2 . \tag{3.47}$$

Similar expressions can be found for other special cases. An example of the covariances after a rotation of 45° is given in figure 3.8b.

3.10.3 Combined weight matrix

The combined observations in the LSM observation equation 3.27 are given by $\tilde{g} - f$. In real applications, the original observations f and g are usually Gaussian filtered to reduce noise and to lower the effect of the uneven smoothing by the subsequent bilinear interpolation of g (see also figure 3.7b).

First of all, this leads to dependencies between adjacent observations. Thus, it would be necessary to introduce observation covariances, which leads to a non-



a. **b.**

Figure 3.8: Covariance values between two interpolated pixels depending on the fractional part of the translation. (a) depicts the situation without rotation, whereas in (b) the transformation included a rotation of 45°.

diagonal weight matrix. Since it is far easier to work with diagonal weight matrices, covariances are most often neglected. I will neglect them also, although a few thoughts are necessary before doing so.

We have seen from equation (3.45) that the covariance between two pixels in a Gaussian filtered image signal is approximated by

$$\sigma_{xy}^{*\,2}(\sigma_G, d) \approx \sigma_0^{*\,2}\, e^{-\frac{d_1^2+d_2^2}{4\,\sigma_G^2}} .$$

Assuming a filter width of $\sigma_G = 1$, the covariance of directly adjacent pixels amounts to $0.78\,\sigma_0^{*\,2}$. When neglecting such a high covariance, we have to expect large error in the subsequent statistics, even if the actual parameter estimate will not be influenced significantly. The reason is that correlated observations are added to the system as independent observations, which will result in a too optimistic estimate of parameter variances, for instance.

An easy but very effective solution is to leave a certain distance between employed observations. Already a minimum distance of 2 (every other pixel) leads to a maximum covariance of about $0.37\,\sigma_0^{*\,2}$. In many applications it is even possible to only use every third pixel in each dimension which lowers the covariance to little over $0.1\,\sigma_0^{*\,2}$. It is important to note that no loss of accuracy is to be expected, since almost no information is taken away, only correlated observations.

Having justified neglecting the covariances under certain conditions, a closer look to the observation variances is taken. Assuming the same variance σ_0^2 for both original observations, the variance of the smoothed observations are all equal and approximated by equation (3.44). However, the variance of the interpolated values $\tilde{g}$ is further reduced by the bilinear interpolation. Please refer to section A.4.1 for details on this reduction.

The result is a different variance for each pixel in $\tilde{g}$. Since this is a little awkward to deal with, the effect of bilinear interpolation is often neglected. Equation (3.33) provides an expression to estimate the error which has to be expected. In this formula, P stands for the correct but unknown weight matrix and ΔP for the perturbation. The magnitude of this perturbation is given by the largest possible error in the weight matrix, which is made for interpolated pixels lying exactly in the middle of the four neighboring original pixels. Equation (A.12) can be used to approximate this maximum variance reduction.

If we for instance assume a filter width of $\sigma_G = 1$, the relative error $(\sigma_0^* - \bar{\sigma}_0)/\bar{\sigma}_0$ is approximately 0.2. Thus, a feasible empirical test is to construct a diagonal matrix ΔP with diagonal elements randomly chosen from the interval $[-0.1, 0.1]$ and to define the matrix P as $I - \Delta P$, without loss of generality. It can be shown that the resulting error is very small for typical computer vision applications and may be neglected unless very high accuracy is required.

3.11 Self-diagnosis within LSM

The diagnostic measures presented in section 3.4 are integrated into the least squares template matching in order to stabilize the optimization and to reduce the risk of getting caught in a local minimum different from the global minimum. Figure 3.10 gives an overview of the complete matching algorithm including the self-diagnosis.

Whenever feasible, the initial parameter set is reduced such that the convergence radius of the optimization is as large as possible. Of course, this restricted parameter set must sufficiently approximate the final parameter set. A possible restriction might be for instance the use of a congruent instead of a full affine transformation. A complete reduction scheme employing the information of parameter correlation is given in figure 3.9.

The next step is to check the determinability of the chosen parameter set. Since the statistics behind the diagnostic measures is valid only in the adjusted state, they can not be applied directly to the initial system. However, an upper bound for the determinability is computed matching the templates onto themselves before the actual optimization and analyzing the parameter correlation and the contribution values δ_i of this perfect match. It is advisable not to use exactly the same template data to match onto, but a smoothed or noisy version, in order to avoid a too optimistic upper bound.

Based on this upper bound, a coarse result is computed using the restricted parameter set. This first estimate is then tested for plausibility by computing the normalized cross correlation

$$\rho(f, g) = \frac{\sum (f - \mu_f)(g - \mu_g)}{\left(\sum f^2 \cdot \sum g^2\right)^{\frac{1}{2}}} .$$

As we have discussed in section 3.4.2, this correlation measure is better suited than the global model test, because of its normalization. Depending on the outcome of the plausibility check, different actions are taken:

- If the estimate is considered plausible, the full affine parameter set is tested for

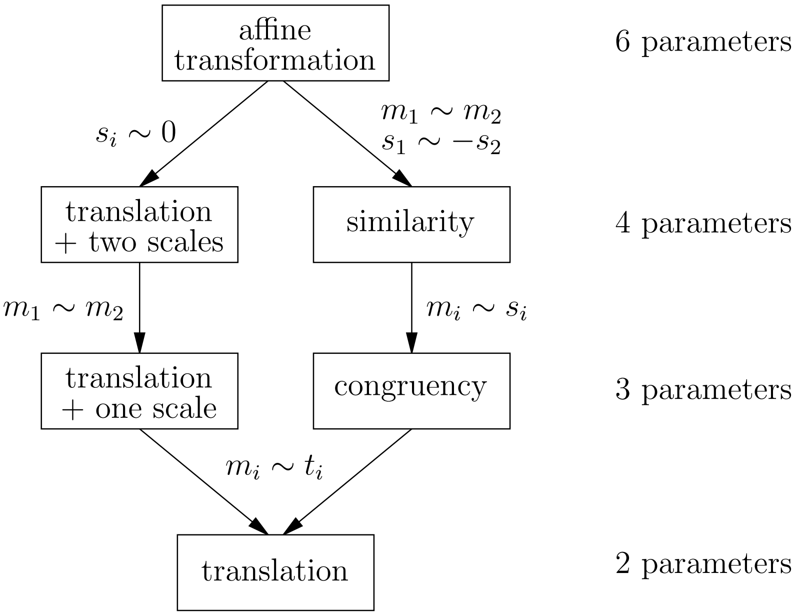


Figure 3.9: Reduction scheme for the estimation of transformation parameters. Depending on the parameter correlations denoted by '∼' the transformation model is restricted.

determinability. The optimization is continued with the full parameter set if none of the parameters show intolerably large correlations or contributions δ_i.

- If the estimate was rejected, the estimation may be restarted with an even more reduced parameter set if feasible or starting from a different initial guess. Alternatively, if multiple templates are used and the correlation value is computed for each template, specific templates can be excluded when their correlation value indicates a local mismatch.

3.12 Results of statistical tests

The meaning of the contribution values δ_i in equation (3.22) is explained in figure 3.11. The 63×63 test images contain a circle and a corner with a signal to noise ratio of 10. Both are Gaussian filtered with $\sigma = 1$.

The circle feature allows for the estimation of the translation and the scaling parameters but not the shear parameters. Building the normal equation system by matching the template onto itself and using an affine transformation as defined by equation (3.39), the contribution values shown in figure 3.11b result. They are once computed based on the correlation matrix $C_{\hat{\xi}\hat{\xi}}$ and also based on the shape part $Q^{(3-6)}_{\hat{\xi}\hat{\xi}}$ of the cofactor matrix. As expected, the contribution of the shear parameters $s_{1,2}$ is significantly higher than of the scaling parameters $m_{1,2}$, which reveals the

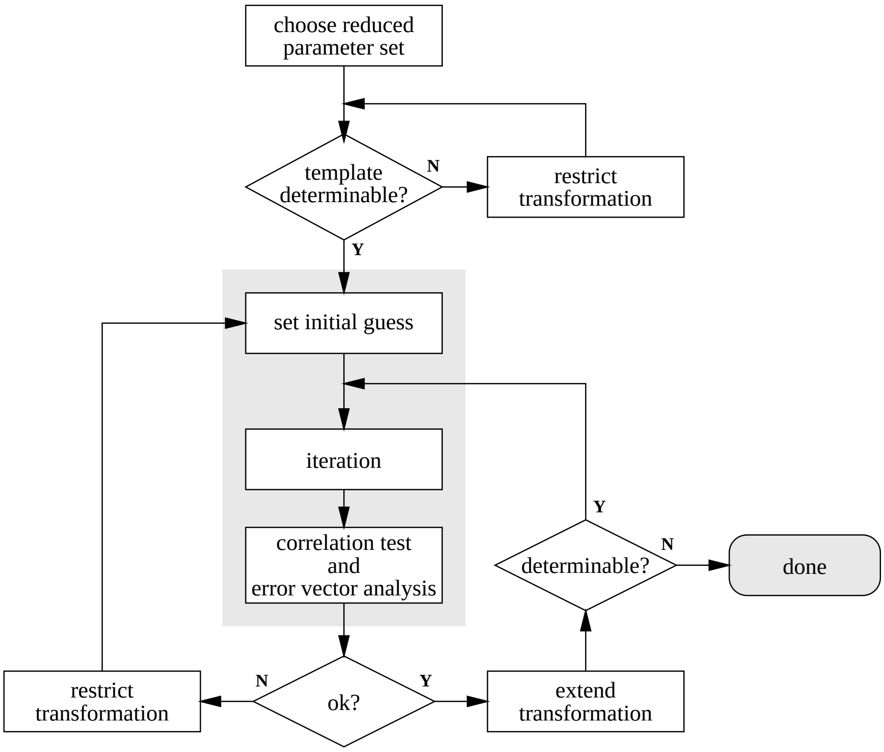


Figure 3.10: Flowchart of the diagnostic measures.

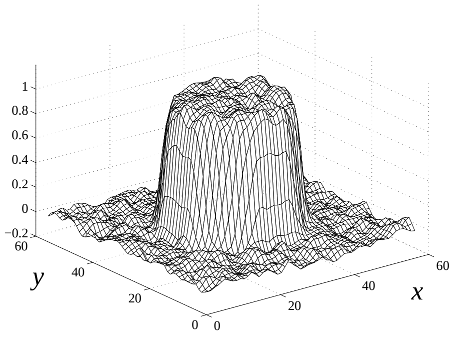
weak determinability of the two shear parameters. Both measures—based on the correlation and on the cofactor matrix—perform well.

The second example reveals one of the problems when this local measure is used as determinability test. Such a corner feature would allow for the estimation of translation and shear. However, the shear parameters s_i are locally correlated with the translation parameters although they are globally determinable, which leads to high contribution values (figure 3.11d).

These cases must identified by a parameter correlation analysis. In the example of this corner feature, the correlation matrix correctly points out the dependence between translation and shear by a correlation value over 0.7. Another example is given in figure 3.12 where the same corner feature was rotated by 45°. In other words, scaling and shear parameters are strongly correlated with themselves. The correlation matrix $C_{\hat{\xi}\hat{\xi}}$ amounts to

$$C_{\hat{\xi}\hat{\xi}} = \begin{pmatrix} 1.000 & -0.005 & -0.271 & 0.019 & 0.022 & 0.276 \\ -0.005 & 1.000 & 0.016 & 0.319 & -0.261 & 0.012 \\ -0.271 & 0.016 & 1.000 & -0.038 & -0.061 & \mathbf{0.772} \\ 0.019 & 0.319 & -0.038 & 1.000 & \mathbf{0.756} & -0.025 \\ 0.022 & -0.261 & -0.061 & \mathbf{0.756} & 1.000 & -0.039 \\ 0.276 & 0.012 & \mathbf{0.772} & -0.025 & -0.039 & 1.000 \end{pmatrix} \tag{3.48}$$

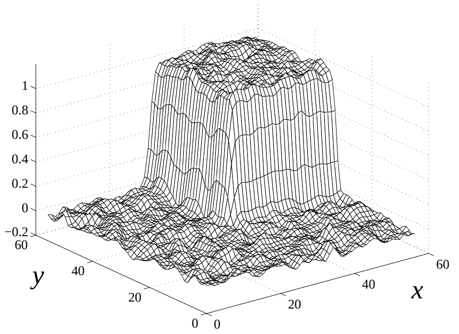
and reveals the high correlations as expected (printed in bold).



a.

	based on	
	$C_{\hat{\xi}\hat{\xi}}$	$Q_{\hat{\xi}\hat{\xi}}$
δ_{t_1}	0.0001	
δ_{t_2}	0.0001	
δ_{m_1}	0.0164	0.0092
δ_{s_1}	0.0951	0.2382
δ_{s_2}	0.0951	0.2376
δ_{m_2}	0.0161	0.0081

b.

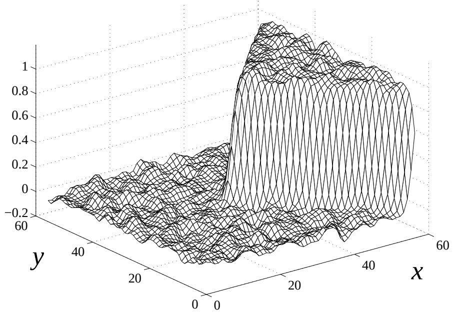


c.

	based on	
	$C_{\hat{\xi}\hat{\xi}}$	$Q_{\hat{\xi}\hat{\xi}}$
δ_{t_1}	0.1051	
δ_{t_2}	0.1052	
δ_{m_1}	0.0026	0.0013
δ_{s_1}	0.1058	0.0052
δ_{s_2}	0.1044	0.0026
δ_{m_2}	0.0015	0.0006

d.

Figure 3.11: Mesh views of the test images for the determinability analysis and the corresponding contribution values, based on a match onto themselves. In (a and b), the contribution values correctly identify the shear parameters as being weakly determinable. However, the second example points out one of the problems of this local measure. The shear parameters s_i are strongly correlated with the translation parameters although they are globally determinable (c and d).



a.

$$\begin{aligned}(c_{\hat{\xi}\hat{\xi}})_{m_1,m_2} &= 0.772\\ (c_{\hat{\xi}\hat{\xi}})_{s_1,s_2} &= 0.756\end{aligned}$$

b.

Figure 3.12: Correlation analysis of a corner feature. The correlation matrix correctly reveals the strong dependency of the scaling and the shear parameters with themselves. See equation 3.48 for the complete correlation matrix.

4

Field edge alignment and checking

The shape of the field edge for each beam direction is defined by the physician during the planning step. For the treatment, this field edge shape is approximated by either lead blocks or by a multi-leaf collimator device (see also section 2.2). One goal of this project is to automatically check the therapy setup for errors in the field edge shape, which currently must be done visually by the operator.

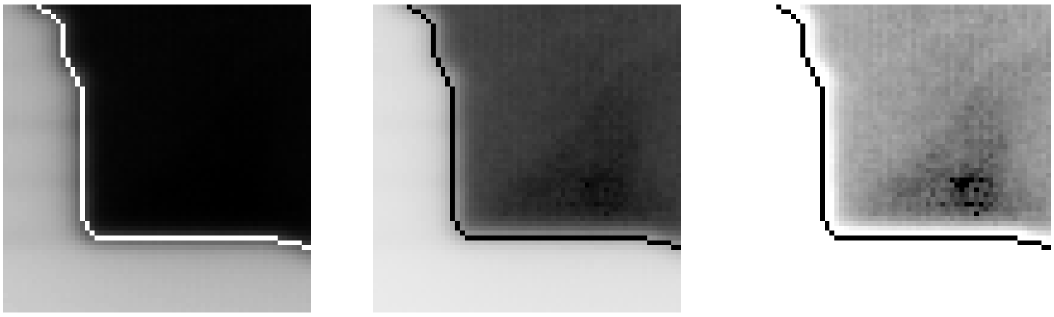
Moreover, the exact position of the EPID is not known in general, which leads to different coordinate systems between images. Thus, each field edge must be aligned to a reference field edge in a first step of the matching procedure in order to establish a common coordinate system. This reference shape can either be extracted from the reference image or better directly taken from the planning data. Within this work, the former approach was chosen merely because of missing software integration. Thus, the alignment step must be preceded by a *field edge extraction*.

The following sections describe first the method for extracting the region of the field edge and for computing an approximated field edge contour. Secondly, the approach taken in this work to align the field edges is presented, which inherently contains a reliable field edge check. For a detailed overview of previous work please refer to section 2.6.

4.1 Field edge extraction

The edge of the physical field is usually defined as the 50 % isodose curve at the depth of maximum dose or at the isocenter [ICRU 1976]. However, this definition is not useful for the analysis of portal images, since it is not possible to obtain this kind of information. Instead, the criterion of maximum gradient is chosen as in existing approaches. Due to this necessary assumption and due to the wide variety of field shapes in portal images, it is difficult to design a reliable and accurate extraction method for the field edge.

However, an automatic field edge extraction with subsequent accurate alignment is still an important prerequisite for patient displacement measurements. Manual extraction of the field edge contour is not only time-consuming, but simply not suitable, since the human observer depends on the chosen gray scale characteristics. That is, the position of field edge can not be visually extracted or checked in a



a. $\gamma = 0.8$ **b.** $\gamma = 2.2$ **c.** $\gamma = 3.0$

Figure 4.1: A section of the same portal image with different gamma values for the gray scale mapping. The position of the field edge contour is identical in all images. A human observer is mislead by the different gray scale characteristics.

reliable way, as illustrated in figure 4.1. Although the field edge contour is at the same position in all three images, it appears to be too far inside in the dark image and too far outside in the lighter image.

Automatic extraction based on the criterion of maximum gradient performs better although not perfect. The location of the maximum gradient changes as well with different gray scale characteristics. This shift is smaller than the expected error for a human observer but still about one to two pixels for the above range of different gamma values (figure 4.2). Thus, the exact location of the field edge contour can not be reliably extracted from portal images. The remaining uncertainty will be the limiting factor for the precision of feature-based alignment methods which rely on the binary contour feature.

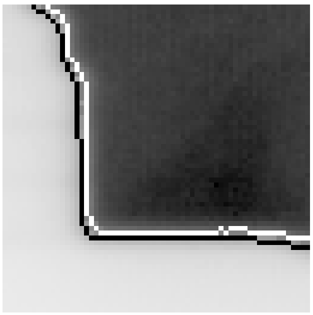


Figure 4.2: Location of maximum gradient depending on the gamma value. The shift from $\gamma = 0.8$ to $\gamma = 3.0$ is about one pixel.

At the same time, this is a substantial argument to employ an area-based method like LSM for the alignment of field edges. Within an area-based approach, the field edge extraction is only necessary to build up an area-based model. The precise position of the field edge contour is not needed, as the subsequent measuring step takes into account the whole region flanking the field edge. Furthermore, even this model building step will become obsolete as soon as the field edge model can be computed directly from the planning data (see also section 2.1). As this is merely a problem of system and software integration, this should be feasible in the near future.

Two approaches were compared: a simple threshold method based on the work of [Bijhold *et al.* 1991a] and a novel approach based on image gradients. The first method aims at selecting an optimum threshold by analyzing the smoothed histogram. Figure 4.3 shows three examples, one typical and two more complex cases. In a typical portal image, the histogram is nearly bimodal. One peak represents the blocked region, the other the significantly darker treatment region (figure 4.3a and b). By selecting the rightmost local minimum as threshold value—that is, the local minimum corresponding to the highest gray value—the treatment region can be extracted. However, in more complex images, a third or forth mode appears in the histogram. If these special cases are not taken into account, the method most likely fails as shown in figure 4.3c through f. Moreover, in order to fulfill the criterion of maximum gradient, an additional non-trivial processing step is required which gradually moves each pixel of the thresholded field edge towards the nearest location of the maximum gradient. Please refer to [Bijhold *et al.* 1991a] for an example of such an algorithm.

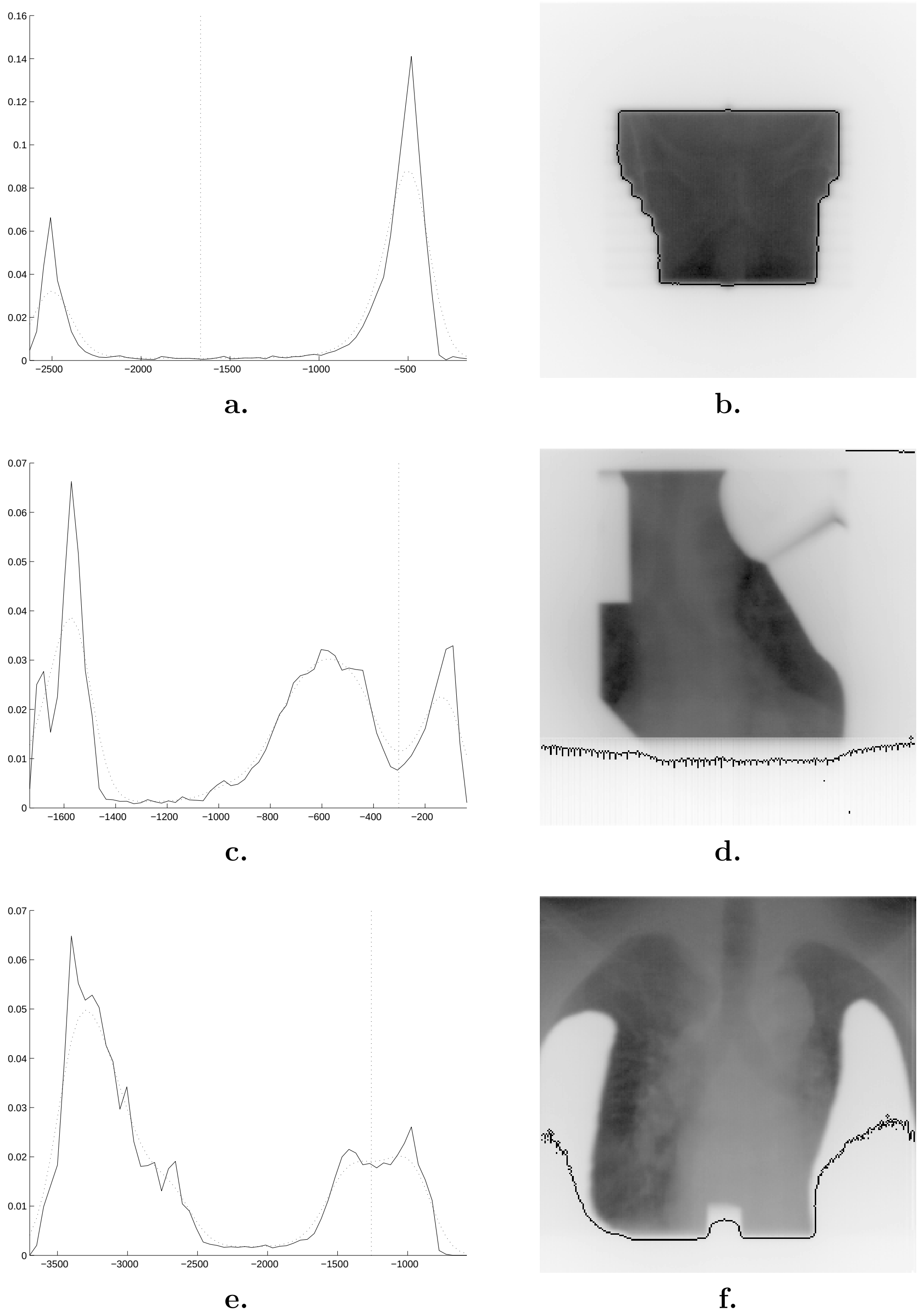
In order to overcome these drawbacks, a novel approach based solely on the image gradients was developed. Besides not being directly dependent on the histogram, it has the further advantage of inherently yielding a field edge region in addition to the contour. This is a requirement for the area-based alignment method using LSM, in contrast to feature-based alignment methods.

Thus, the algorithm was designed to yield a reliable mask of this *field edge region*, which is used to build an area-based model of the field edge. The location of the field edge contour is computed in a second phase but only for the purpose of visual checking since it is not needed within the alignment and automatic checking procedure.

The following sections describe the details of my approach which is summarized in figure 4.4. Example images of intermediate results are depicted in figure 4.5. It is important to note that all parameters involved in the presented method were kept at the same value for the complete dataset. Hence, no manual parameter tuning is necessary and the success rate lies at 99 % as it only failed in 24 cases out of all 2200 portal images.

4.1.1 Preliminary steps

It is possible that the portal image does not contain the complete field edge. The right column in figure 4.5 depicts such a case. Without special treatment, the field edge would fall apart which makes a robust extraction difficult. Therefore, an additional margin is defined to simulate a blocked field around the image. The exact value of this margin is not crucial, as long as it lies near the value of other pixels in a blocked region. The following heuristic but simple approach finds a suitable background value in almost all portal images in the dataset:



a. **b.**

c. **d.**

e. **f.**

Figure 4.3: The histogram method for field edge extraction. The rightmost local minimum is selected as threshold value. This method works well on typical portal images as shown in (a, b). However, it is bound to fail in more difficult cases where for instance an additional wedge is included (c, d) or where the blocked field is not homogeneous (e, f).

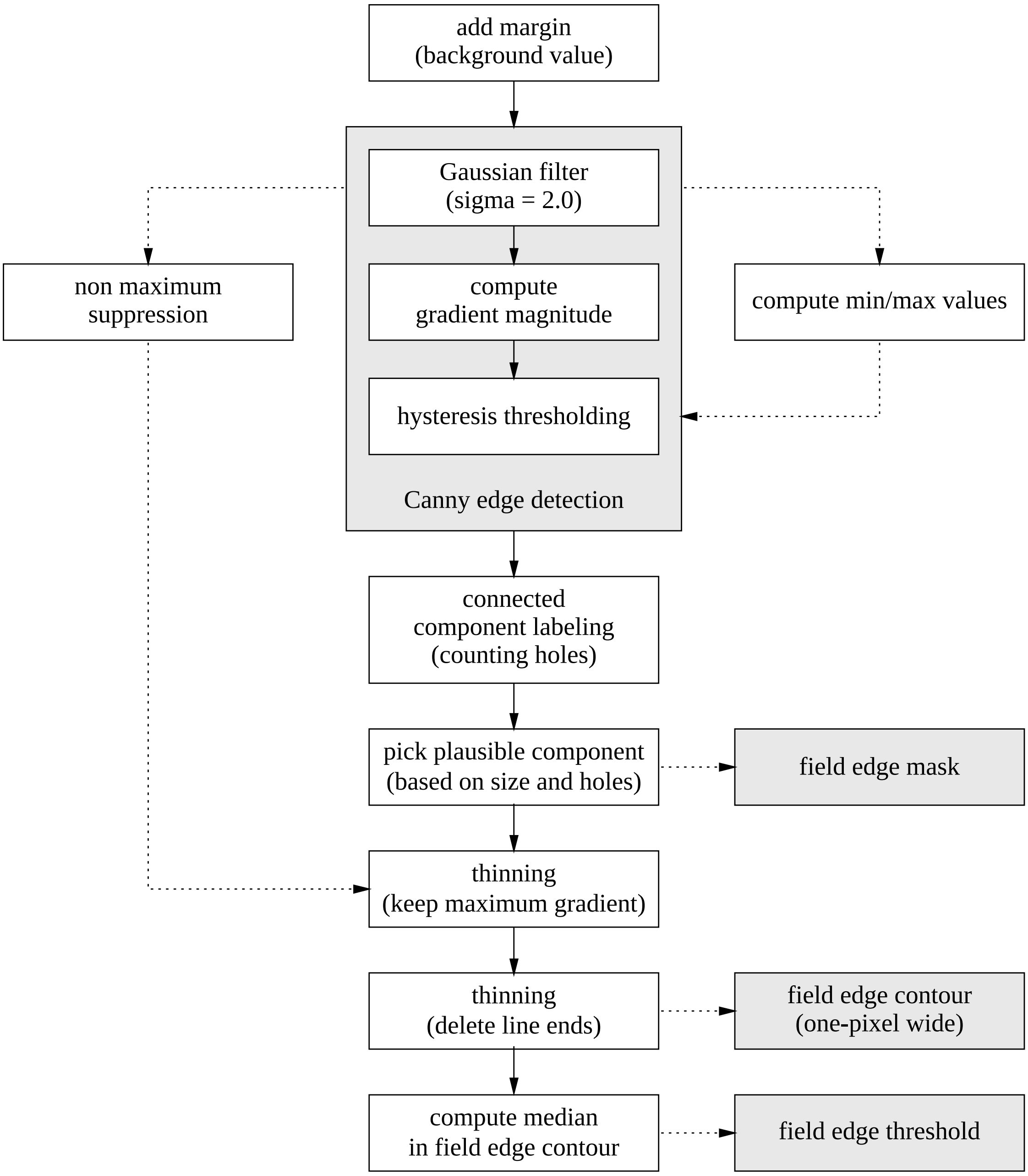


Figure 4.4: Flowchart of the field edge extraction algorithm.

1. collect all border pixels (1020 pixels in the case of a 256×256 image)
2. sort them and pick the 75 % percentile.

4.1.2 Edge detection

The following steps are based on the Canny edge detection algorithm [Canny 1983]. After Gaussian smoothing using a fixed sigma of 2.0 pixel, the gradient magnitude is computed without applying the non-maximum-suppression method described by Canny. Using two relative threshold values based on the minimum and maximum value of the smoothed image, hysteresis thresholding is performed as proposed by [Canny 1983] (figure 4.5b and c). This double threshold operation leads to a binary image where regions including high gradients are marked.

4.1.3 Region selection

If multiple regions are found, one plausible region is selected after labeling the connected components (refer for instance to [Rosenfeld and Pfaltz 1966]). The employed labeling algorithm returns the size and the number of holes for each object. Thus, the final field edge region is picked under the assumption that large regions with few holes are more plausible than others. This heuristic rule was true in more than 99 % of all portal images in the clinical test data (over 2000 images). The final region is then slightly smoothed by dilating the binary image using a 3×3 square kernel.

4.1.4 Compute field edge contour

As last step, an approximate field edge contour is computed. The goal is to apply a topological invariant thinning operation such that a connected field edge region yields a connected contour. However, the field edge must also lie at the position of the maximum gradient. These two requirements are fulfilled by combining the thinning operation with the non-maximum-suppression: pixels with local maximum gradient are marked as non-deletable. Thus, the thinning algorithm keeps Canny edge pixels and connects them to preserve topology.

4.2 Field edge template generation

The subsequent alignment procedure relies on the definition of template regions of significant structures. After the extraction of the field edge region, it is thus a simple task to define such templates. First, the edge region is enlarged by applying another dilation step with a larger kernel of 5×5 pixels. The size of this kernel is not crucial and is chosen rather small for performance reasons.

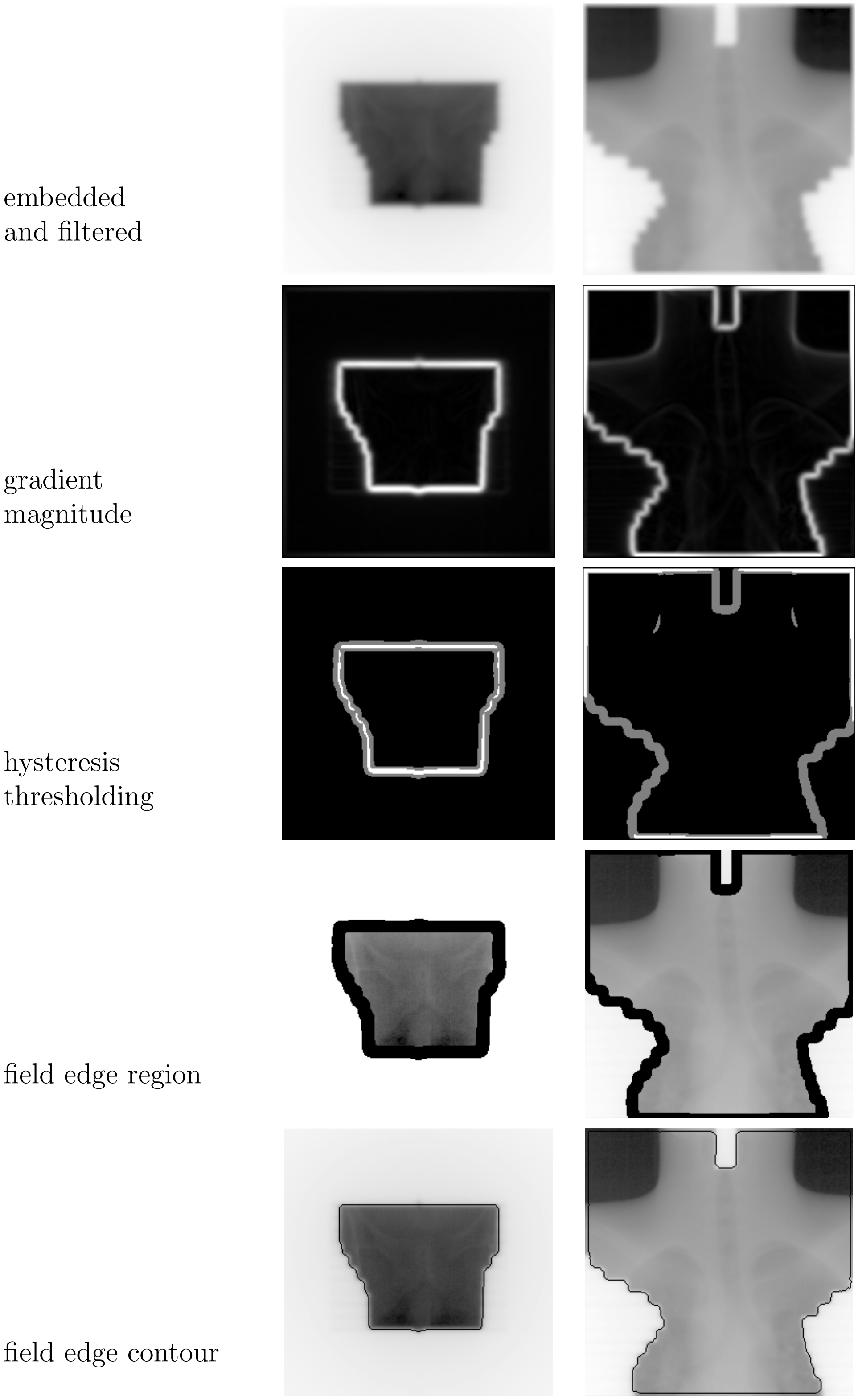


Figure 4.5: Intermediate steps of the field edge extraction algorithm.

Secondly, the region is subdivided as long as it contains holes or the coefficient ‘area over bounding box area’ is smaller than one third. This is a heuristic strategy to generate more or less rectangular shaped templates. The reason to aim for rectangular templates lies in implementation details and is not a prerequisite for the algorithm. In the current implementation, some computation parts as for instance resampling are carried out within the full bounding box area and not only on the template pixels. This lead to simplifications within the program code which justified the additional computational overhead. Similarly, the elimination of holes is not necessary because of the algorithm but simply because the prototype version of the user interface only handles templates defined by a single outline.

Further processing of these binary regions is required for the interaction with the user interface. The operator outlines regions by defining polygons. Thus, the representation of the field edge templates must be changed from a binary map to a polygonal outline. This is achieved using a simple contour following algorithm with subsequent point reduction as proposed by [Douglas and Peucker 1973]. Examples of such generated template configurations are depicted in figure 4.6 for various types of portal images.

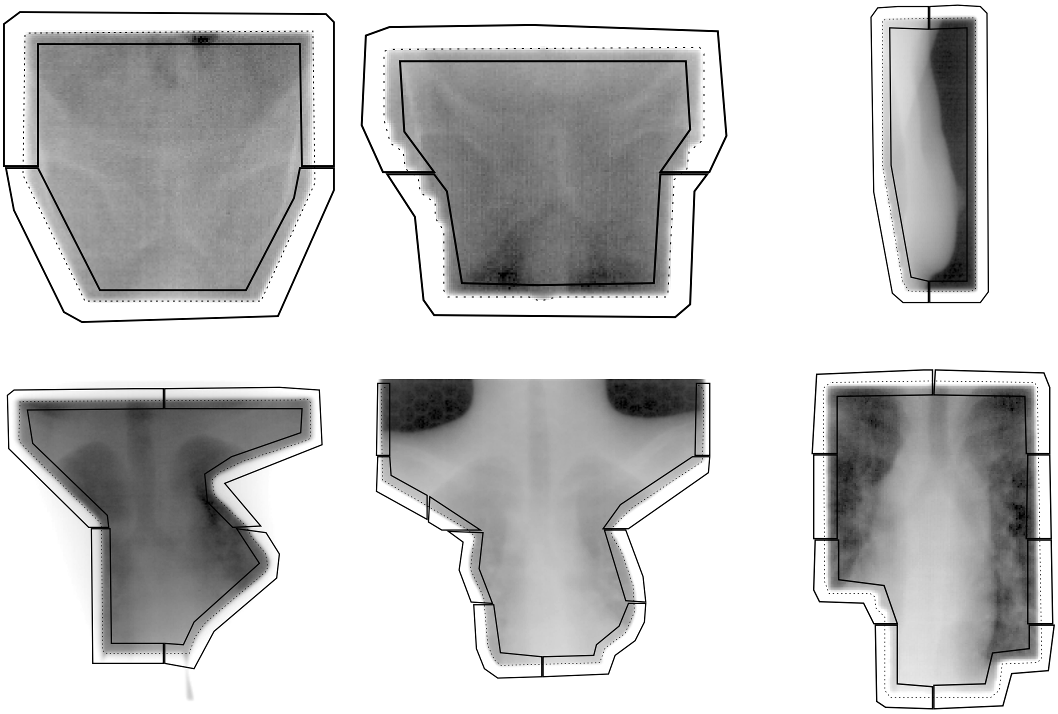
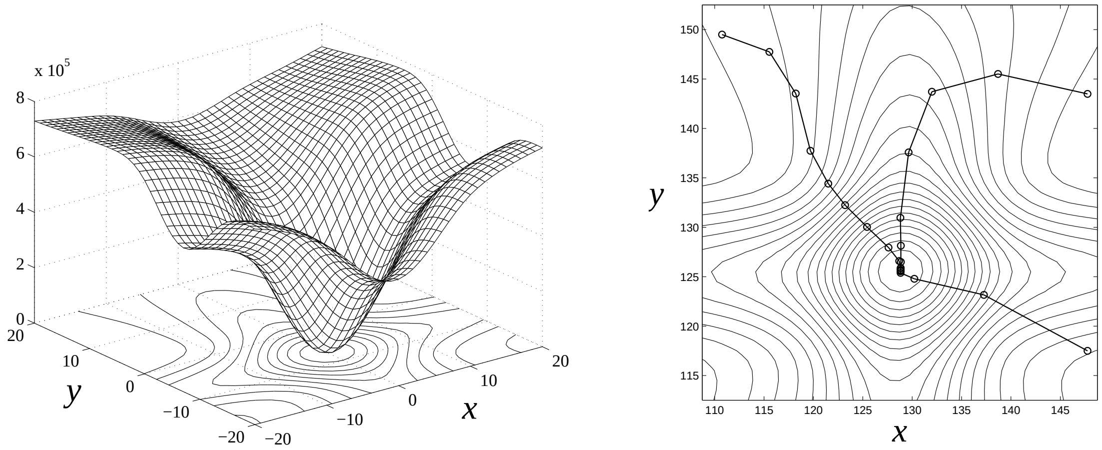


Figure 4.6: Examples of the automatic definition of the field edge templates. The black lines outline the template regions, the dotted line represents the estimated field edge position.



a. **b.**

Figure 4.7: Example of a typical error surface (a) and its contour plot for the field edge match in dependence of the translation parameters. Overlayed on the contour plot are the optimization paths for three different starting points (b).

4.3 Measuring field edge displacement

Although the position of the EPID is not known exactly, the device is assumed to be perpendicular to the beam axis. Simple geometric considerations lead to the conclusion that it is sufficient to estimate a two dimensional similarity transformation between the projected field edges—that is, the four parameters defining translation, rotation and scale—in order to establish a common coordinate system.

This field edge alignment is achieved by applying the LSM framework as described in chapter 3. The template regions for the area-based match are given by the result of the extraction of the field edge region. Although manual optimization of these automatically generated templates is possible, it is only necessary in rare special cases.

The field edge match itself is very robust, due to the restriction to four parameters in the transformation model and the fact that the field edge is a distinct and unambiguous feature. Together with the least squares goal function, this leads to a well behaved error surface during optimization. Thus, the risk that the optimization stops in a local minimum away from the global minimum is very small. Figure 4.7 depicts an example of a two dimensional error surface and its contour plot in dependence of the translation parameters. The optimization paths for three different starting points are overlayed on the contour plot.

This robustness of the field edge match is an important step in checking the patient position. Firstly, the resulting transformation can be safely used as initial guess for the anatomy match. Secondly, a reliable detection of field edge shape deviation is rendered possible. This is achieved by applying the statistical tests outlined in section 3.11 to the field edge match. Especially the cross correlation value gives a reliable feed back as will be demonstrated in section 4.6 below.

4.4 Position of the projected isocenter

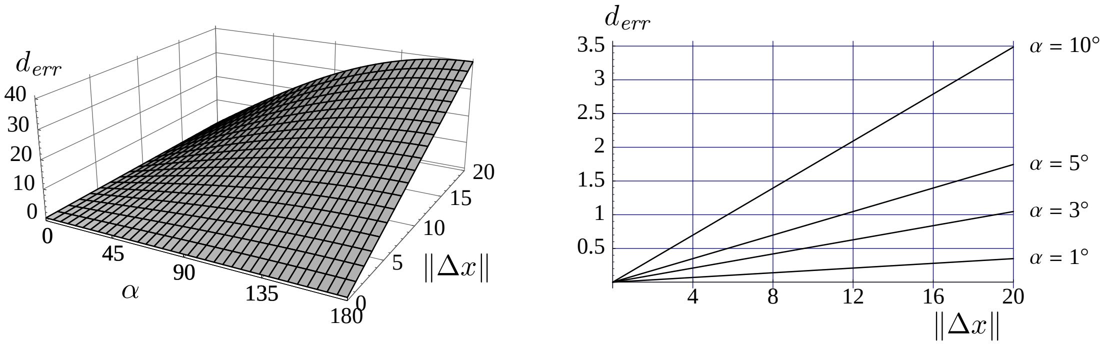
Since the position of the EPID is unknown, so are the coordinates of the projected isocenter in standard portal images. Without further knowledge about the actual field edge shape, it is not possible to directly estimate the isocenter coordinates. However, the EPID was positioned approximately in central position in all portal images within the dataset. That is, the projection of the isocenter—which is at the same time the intersection of the beam axis with the EPID—lies near the center of the portal image.

When correcting the patient position, the axis of rotation is the beam axis. Hence, if this position is not known exactly, an additional translation error is introduced. Simple geometric considerations lead to the following formula for this translation error:

$$d_{err}(\|\Delta x\|, \alpha) = \|\Delta x\| \sqrt{2\,(1 - \cos\alpha)}\,,$$

where $\|\Delta x\|$ is the distance between the assumed and the correct position of the beam axis. Nevertheless, small errors in the position of the beam axis will only lead to a minor translation error after the patient position correction. A qualitative visualization of this error is shown in figure 4.8a, whereas figure 4.8b gives a closer view on the relevant range of up to 10 degrees rotation. Even with such a large rotation correction, the error stays below 2 pixel as long as $\|\Delta x\|$ is less then 11 pixel. In typical portal images this amounts to an additional error of less than 2 mm if the image center is within 10 mm of the beam axis.

Similarly to the field edge extraction, this problem will become obsolete once it is possible to gain knowledge about the field edge from the planning data or even more accurate from the collimator device. This will be feasible in the near future with increased integration of the multi-leaf collimator technique. Due to this fact and to the above finding that the typical error is far below 2 mm, no attempt was made to estimate the position of the isocenter within this work.



a. **b.**

Figure 4.8: Additional translation error introduced by incorrect position of the isocenter. The error depends on the estimated rotation (x axis) and on the distance between assumed and correct position of the isocenter (y axis). The complete range up to 180 degrees is depicted in (a), a detailed view of the relevant range in (b).

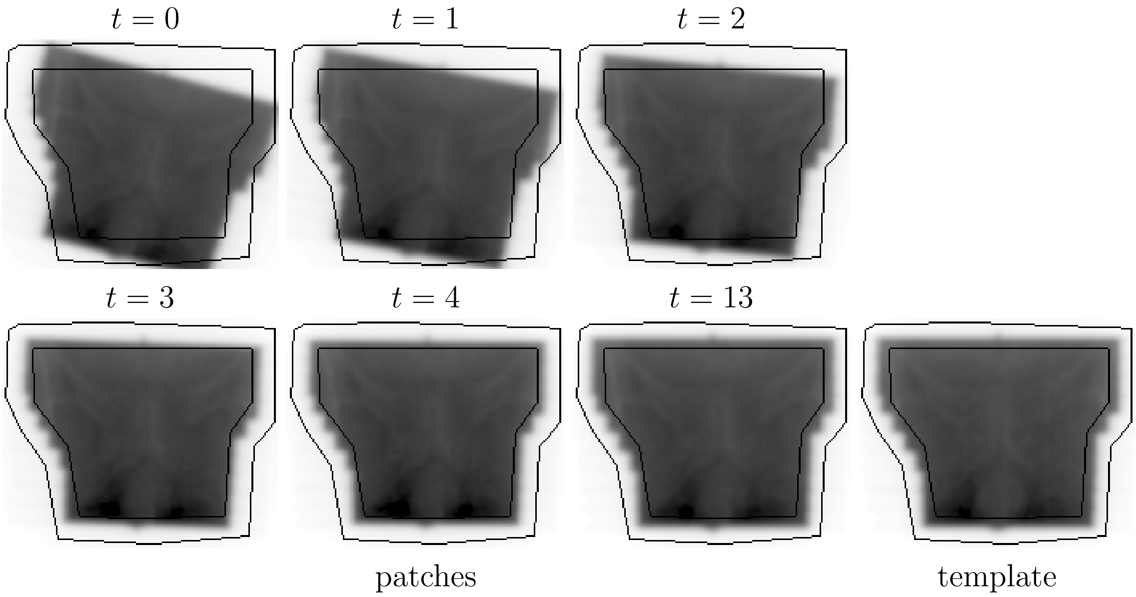


Figure 4.9: Iteration series of an artificially rotated field edge of an anterior-posterior pelvis image. The search image ($t=0$) is rotated by -15°. The black lines enclose the automatically extracted template placed on the search image. After 13 iterations, a perfect match to the template is found.

4.5 Results of field edge alignment

As expected, field edge alignment is very robust and LSM even finds larger displacements and rotations as occur in radiotherapy treatment. The number of necessary iterations typically lies between 5 and 15. The undersampling factor was set to 3, that is every third pixel was picked as an observation.

The iterative nature of the LSM algorithm is illustrated in figure 4.9. Starting from an initial guess—in this case an identity transformation— the optimum position and orientation is iteratively estimated. In this example, the search image was artificially rotated by -15° (clockwise), in order to demonstrate the robustness of the field edge alignment and to better visualize the iteration. After 13 iterations, the template matches perfectly the corresponding area in the search image.

Figure 4.10 visualizes the actual gray level information used by the LSM algorithm. The single template from the last example was split into eight small templates such that the reference image is visible between the resampled template regions. Each iteration step of LSM includes the resampling of the search image within the template regions.

4.5.1 Self consistency test

In order to estimate the overall accuracy of the field edge match, a self consistency test was carried out. Its setup is described in figure 4.11. After estimating the field edge displacement between two portal images (figure 4.11a), the search image was

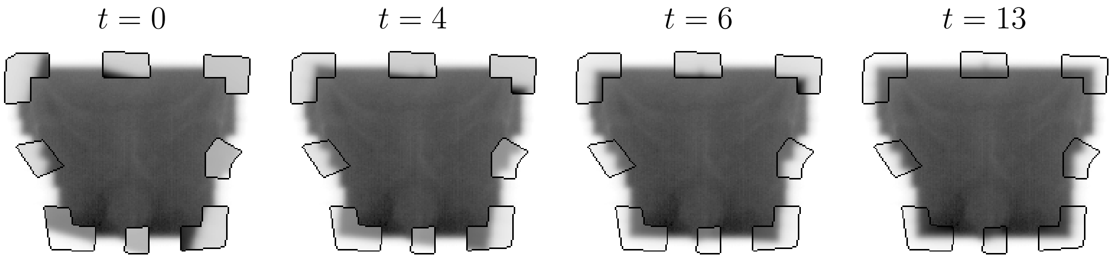


Figure 4.10: Resampled template regions overlayed on the reference image. The gray level difference and the gradient within each template region serve as information to the LSM algorithm.

artificially rotated by -15° and the displacement was estimated again. Table 4.1 shows a detailed comparison of the results. The rotation measurements of 0.047° for the original images and -14.954° after the artificial rotation demonstrate perfect performance with respect to a ground truth.

4.5.2 Parameter precision

The parameter precision is given by the diagonal elements of the covariance matrix $\Sigma_{\hat{\xi}\hat{\xi}}$ as defined in equation (3.18). In contrast to the overall accuracy, the parameter precision does not take into account external error sources.

Furthermore, they are only meaningful if the applied model was correct. In particular, the weight matrix must have been designed according to the correlation of the observations (see also section 3.10), that is the gray values. In this work, the observations were always assumed to be independent, which is of course not true for adjacent pixels. Therefore, only every n-th pixel is taken as an observation where n is typically 3. Even with this subsampling, there are often remaining correlations, which leads to a too optimistic estimate for the parameter precision.

Thus, the precision represents an upper bound of the overall accuracy. Table 4.2 summarizes the range of typical parameter precision for the field edge and also for the anatomy match.

	translation [pixel]	rotation [deg]
field edge match	0.005 – 0.02	0.005 – 0.02
anatomy match	0.09 – 0.12	0.1 – 0.3

Table 4.2: Typical range of parameter precision for the field edge and the anatomy match. The parameter precision does not include external errors and thus represents an upper bound of the overall accuracy.

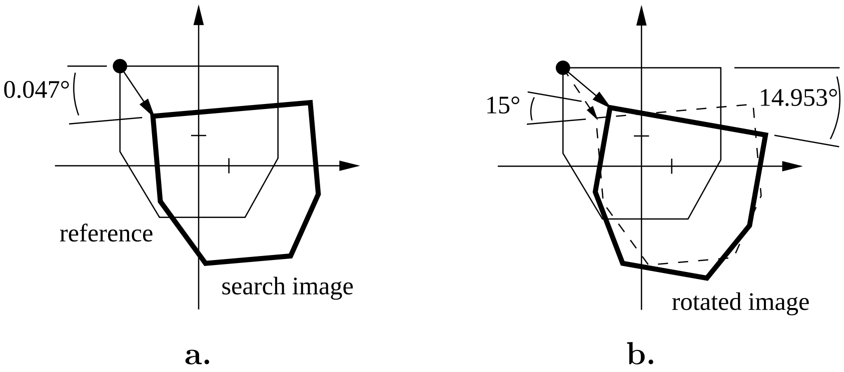


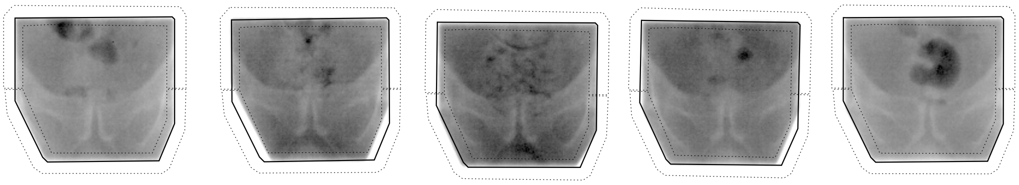
Figure 4.11: Setup of the self consistency test. In (a), the original relation of two portal images (reference and search image) is shown. Diagram (b) depicts the situation after rotating the search image by -15°. The measurements on this rotated image are then compared with the expected values (see table below).

	dx [mm]		dy [mm]		rot [deg]	
original (estimated)	-0.072	±0.0073	-6.883	±0.0059	0.041	±0.0057
rotated (estimated)	-1.850	±0.0080	-6.633	±0.0064	-14.959	±0.0062
rotated (ground truth)	-1.851		-6.630		-14.959	

Table 4.1: Results of the self consistency test, computed by estimating a similarity transformation. The estimated parameters (second row) almost perfectly agree with the expected values (third row). The cross correlation value for both measurements was 0.9996.

4.6 Results of field edge check

The normalized cross correlation value gives a reliable feedback on the result. A few examples are given in figure 4.12, showing that a match of field edges with identical shape results in a cross correlation value well above 0.99, whereas small differences in the field edge shape already lower this value significantly.



reference $\rho = 0.907$ $\rho = 0.976$ $\rho = 0.963$ $\rho = 0.991$

reference $\rho = 1.000$ reference $\rho = 0.996$

Figure 4.12: Field edge alignment and its cross correlation values ρ for three typical cases. The top row shows image series with slight variation of the field edge shape. A match of field edges with identical shape results in a cross correlation value well above 0.99 (bottom row).

5

Displacement measurements in portal images

In the last chapter, the comparably easy task of matching the radiation field edge was presented. The results showed the general suitability of the matching method and its diagnostic measures. Yet to discuss is the application to the low-contrast anatomy part of portal images.

As mentioned earlier, area-based methods have the potential of yielding very precise results. This can be exploited when measuring the field edge displacement. However, this precision is not the main reason to apply the method to the anatomy regions, since the overall accuracy is governed by external errors. As we will see, neglecting the projective and transparent nature of the imaging process already leads to variations in the magnitude of one to three millimeter in a typical portal image.

The main reason is rather to avoid the difficult step of a robust feature extraction in such low-contrast imagery. Thus, the emphasis in this chapter will be less on the precision of the method itself, but on the reliability and on an acceptable overall accuracy. As specified in section 2.5, high precision radiotherapy requires an overall accuracy of 2 to 5 mm depending on the type of treatment. In the special case of stereotactic radiosurgery, the desired accuracy is in the submillimeter range.

The first section deals with the kind of reference image to be employed. The choice of an area-based matching method already poses constraints on this selection. Then, I discuss the problem of selecting suitable areas, templates, for robust matching in the reference image. This is the stage where a minor amount of expert knowledge is required. Section 5.3 explains the details on how the LSM algorithm was applied to this matching problem. The further sections present results from artificial test data, from phantom measurements and from real patient image series. Also included at the end of this chapter is a feasibility test of a so-called multi-modal match between a computed reference image and the corresponding portal images.

5.1 Choosing a reference image

In order to define an accurate reference image for checking patient position, there are mainly three possible choices.

1. The first or best portal image of a series validated by a physician.

   Employing a portal image as reference image assures very similar gray value characteristics, hence yields the best match using an area-based matching algorithm. However, this portal image must be validated by other means, for instance manually matched to the simulator image, in order to establish a reference to the planning step.

2. The diagnostic X-ray image from the simulator.

   When measuring patient motion visually using portal images, often the diagnostic X-ray image (see also figure 2.1b) is used as reference. Bony structures are easier to identify in this imagery, since the contrast is much higher. Nevertheless, it is not suitable for an area-based algorithm, since the different gray value characteristics strongly influence the quality of the measurements.

3. A digitally reconstructed radiograph.

   As described in section 2.4, the most accurate reference is a megavoltage DRR (see also figure 2.1a), which is directly computed from the planning CT. Thus, additional sources of positioning errors are avoided, for instance the step of therapy simulation. However, the remaining gray value differences pose additional problems to an area-based algorithm like LSM compared to simply match two portal images.

5.2 Selecting suitable anatomy templates

Suitable templates must consist of structures that are known to be stable over a series of images. However, due to artifacts and the presence of distinct but unstable features—for instance originating from air in the rectum—a fully automated template selection is beyond the possibilities of computer vision. Thus, the strategy is to implement an operator guided template selection.

Structures must meet two requirements to be considered *stable*, a global and a local requirement. Global stability is given when any probable movement of the selected structures can be described by the three dimensional rigid body transformation. This is easily achieved in an anterior-posterior (AP) pelvis image, for instance, where all structures are selected from the pelvis bone. However, it is more difficult in other situations, for example in a lateral pelvis image, where the distinct structure of the hip bone can not be used.

The local stability criterion is somewhat more vague but equally important. Since we are dealing with transparent objects, edges and in particular lines can significantly change their relative position under a slightly different viewing angle, or even disappear completely. A good example are the thin bone structures in the pelvis region, which project onto a bright line when the beam axis lies exactly in the bone plane, but are almost featureless for any other angle. Thus, the operator must choose structures that are known to change little in the expected patient movement range and over a series of images.

Another important aspect is the unknown 3D position of the selected structures. Since the patient motion is estimated based on the projected portal image, the position along the beam axis of each structure is unknown and may introduce an additional systematic error, which is explained in more detail in the result section 5.6. This error can be reduced if mostly structures from one plane parallel to the image plane are selected.

All requirements call for a certain amount of expert knowledge. Nevertheless, this knowledge is already present, since a physician needs the same type of structures when visually validating a patient's position based on X-ray images. In the current implementation, the physician has to define regions containing such stable significant structures. These regions are considered templates for the subsequent matching process. Figure 5.1 shows typical template configurations.

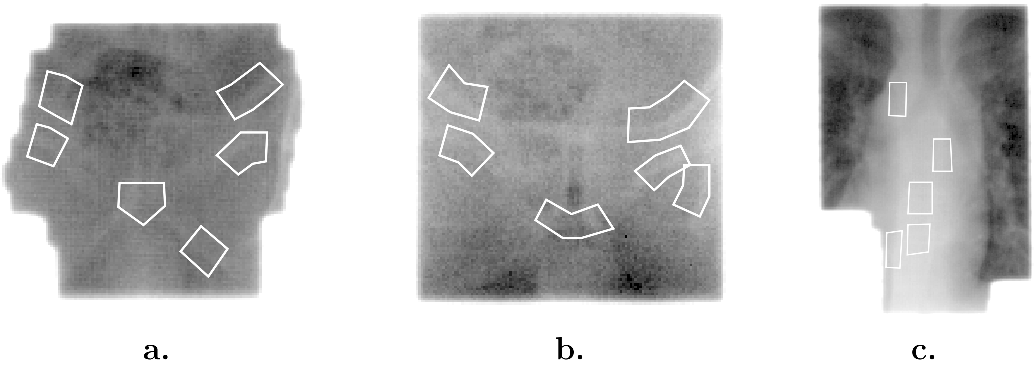


Figure 5.1: Images (a) and (b) show two typical template selections for AP pelvis fields. An example for a thorax image is given in (c).

In feature-based algorithms developed for this type of imagery, usually landmarks or contours must be manually defined. This is not only time-consuming, but also introduces a great variability between different operators. Since LSM is area-based, these problems do not arise and the physician only has to outline several regions where significant structures are found. This is accomplished with a few mouse clicks, drawing several polygons onto the reference image. In order to further reduce the workload, standard configurations can be stored in a template database.

Defining only small regions as templates has various advantages. Besides the lower computational costs, problematic zones can be avoided. Including insignificant structures or artifacts would impede a robust match. If one of the selected regions causes gross errors during the matching procedures, this can be detected by a posteriori self-diagnosis. Such templates are labeled unmatchable and are eliminated from the estimation.

5.3 Measuring anatomy displacement

In order to estimate the anatomy displacement, a few assumptions are made about the patient movement. First of all, the movement is assumed to be a three dimensional rigid body transformation. This is a necessary assumption, since a non-rigid movement would make meaningful measurement virtually impossible. Nevertheless, it is not a strong assumption if the templates were correctly selected as discussed in the previous section.

Moreover, the patient's displacement is assumed to be small: one to two centimeters translation and a few degrees rotation. This is a valid assumption in hospital practice, since the patient has to be aligned by other means before each treatment session, before a portal image is acquired. Existing methods for such prealignment were discussed in section 2.1. Since the portal image is only available *after* the start of the treatment, its main use is for checking and adjusting the patient's position, as opposed to ensuring a correct initial position.

Thus, the displacement of the projected anatomy is approximated by a two dimensional affine transformation. This accounts for parts of the projective distortion without introducing too many degrees of freedom. This approach leads to an implementation of the multi template LSM using a single affine transformation as described in section 3.6.1. One global transformation is hence used for all templates.

The radiometric scaling parameter β is also kept the same for all templates. Only the constant parts of the radiometric transformation α^K may differ between templates. Therefore, the total number of parameters using an affine transformation and N templates amounts to $r = 6+N+1$. This strategy allows for inherent bias correction without adding too many degrees of freedom.

As initial guess for the anatomy alignment serves the similarity transformation which resulted from the field edge match (see also figure 2.7). The diagnostic measures are applied as described in section 3.11.

The typical run time for the complete matching procedure including Gauss filtering of the search image, matching the field edge and matching the anatomy templates is about 3 seconds on a Sun Ultra 1 (167 MHz UltraSPARC CPU).

5.4 Results of headrest test series

A preliminary test to show the suitability of the method for stereotactic radiosurgery was carried out measuring images of a displaced headrest. The repositioning accuracy of this fixation device is specified as 0.2 mm. In 11 different positions, 14 images were acquired using the BeamView System. Figure 5.2 illustrates the image series of this fixation device. The measured displacements were compared with the ground truth and resulted in an overall point error of 0.2 mm (figure 5.3).

In this example, I will take a closer look at the validity of the estimated parameter precision as presented in section 3.4.4. Please note that pixels are assumed to be independent, which is mostly a wrong assumption and leads to an incorrect weight

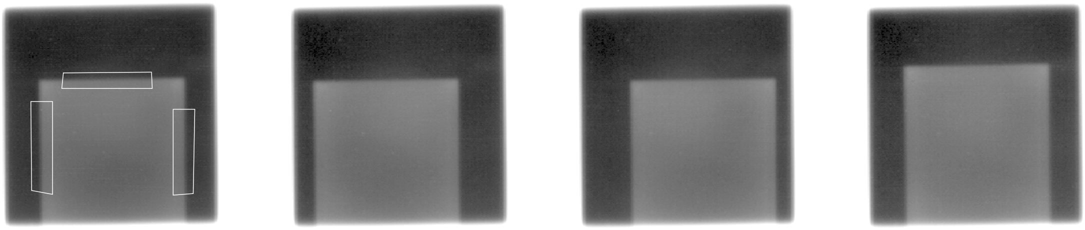


Figure 5.2: Headrest displacement test series. The first image was used as reference image using the regions within the white polygons as templates.

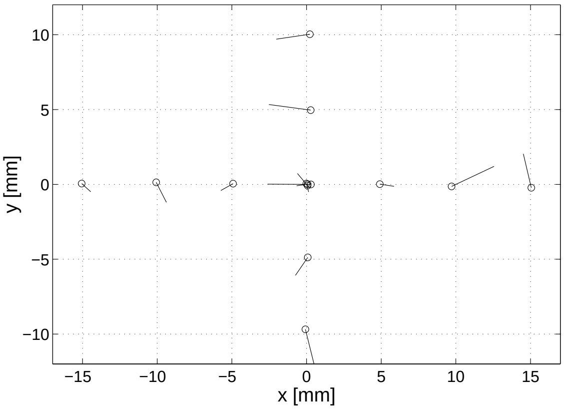


Figure 5.3: Error analysis of the headrest test series. The ground truth coordinates lie on the grid and the error vectors are enlarged by a factor 10. The overall point error of the measurement amounts to 0.2 mm (see also table 5.1).

			point estimate	interval estimate ($\alpha = 0.05$)	remarks
x	mean	[mm]	0.039	$-0.051 \ldots 0.130$	statistically equal to 0
	std dev	[mm]	0.158	$0.114 \ldots 0.254$	
	std dev	[mm]	0.157	$0.115 \ldots 0.247$	assumption: mean = 0
y	mean	[mm]	0.022	$-0.052 \ldots 0.095$	statistically equal to 0
	std dev	[mm]	0.128	$0.093 \ldots 0.206$	
	std dev	[mm]	0.125	$0.092 \ldots 0.198$	assumption: mean = 0
$\sqrt{x^2+y^2}$	std dev	[mm]	0.20	$0.16 \ldots 0.27$	assumption: mean = 0
rotation	std dev	[deg]	0.25	$0.18 \ldots 0.39$	assumption: mean = 0

Table 5.1: Statistical analysis of the headrest test series. The standard deviation in x and y was computed both using the computed mean and under the assumption of zero mean. The significance level α was set to 0.05.

matrix. Refer to section 3.10 for more details. As expected, the estimated parameter precision is far to optimistic when introducing every pixel within the templates as an independent observation, yielding a standard deviation between 0.005 and 0.01 mm. When picking every 25th pixel—evenly distributed in the templates—this value goes up to around 0.04 mm. This seems to be a reasonable value for the standard deviation. The remaining error is perfectly explained with the repositioning uncertainty of the fixation device.

5.5 Results of two dimensional patient displacement measurements

Due to much weaker contrast, the anatomy alignment is not as robust as the field edge match. Moreover, an affine transformation model is usually used for the anatomy match, which introduces two additional degrees of freedom. The iterative optimization scheme is illustrated in figure 5.4 and 5.5 with an AP pelvis image. During the optimization, the resampled patch gradually rotates and translates to fit the gray values in the template.

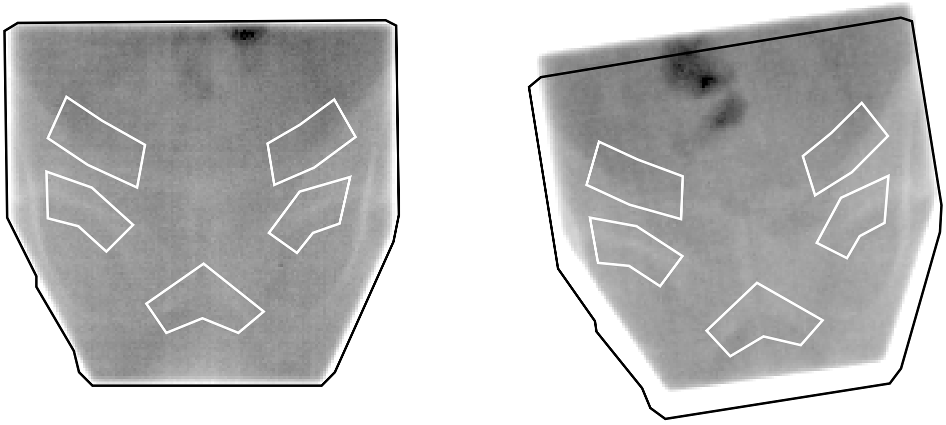
A first test series consisted of roughly 100 clinical portal images in 17 series, mostly from the pelvis region. The templates were defined by myself with the advice of a physician. In all measurements, an affine transformation was used for matching and the initial position was set to the field edge position.

The algorithm performed best on AP pelvis images, where it found a correct match in all but one image. A match result was defined correct when there were no visually noticeable errors. The overall success rate of LSM was over 90 %. That is, in less then 10 % of the displacement measurements, the result was either automatically or manually rejected, by diagnostic measures or after visual validation.

The overall accuracy is difficult to estimate, since no ground truth is available for real portal images. Various external errors add to the parameter precision presented in table 4.2, in particular the projective nature of the imagery. However, visual validation and comparison with a few manual measurements indicate an accuracy of about 1 or 2 pixel for translation and below 1° for rotation measurements. A more detailed comparison of manual and automatic measurements was carried out in a second test series, described in section 5.8. In the following, various tests are presented in order to assess the reliability of the measurements.

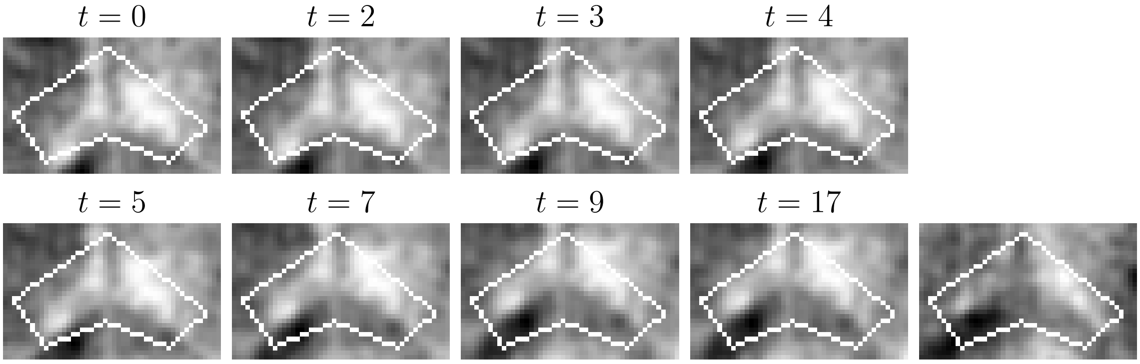
5.5.1 Sensitivity to template selection

In order to estimate the accuracy of the patient displacement measurements, the same images are matched using slightly varying set of templates. Figure 5.6 summarizes the results using 10 different sets of templates on a series of 6 images. The standard deviations of the x and y position are both about 0.3 pixel and of the orientation 0.3°.



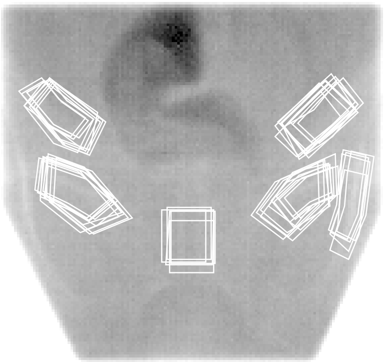
reference image search image

Figure 5.4: Anatomy match example of an AP pelvis image. The search image is additionally rotated by 8° to test the iterative matching procedure (see the iteration series below). White polygons outline the template regions and the patches found, respectively. The black line represents the field edge of the reference image and is overlayed onto the search image to visualize the patient displacement.



patches template

Figure 5.5: Iteration series of the middle template in the above pelvis image. Notice that the contrast of these images is manually enhanced for printing purpose.



a.

$$
\begin{aligned}
\sigma_x &= 0.30 \text{ pixel} = 0.26 \text{ mm}\\
\sigma_y &= 0.24 \text{ pixel} = 0.20 \text{ mm}\\
\sigma_{\sqrt{x^2+y^2}} &= 0.38 \text{ pixel} = 0.33 \text{ mm}\\
\sigma_{rot} &= 0.314 \text{ deg}
\end{aligned}
$$

b.

Figure 5.6: The sensitivity to slight variation of the template regions is tested using 10 different template configurations. They are all overlayed onto the reference image in (a). The resulting standard deviations of the measurements are shown on the right and represent an estimate of the overall accuracy (b).

5.5.2 Convergence radius

An important aspect in optimization schemes is the radius of convergence. The initial guess or starting point of the iterative optimization must be within this convergence area. The LSM algorithm starts to fail when the difference between the initial guess and the correct position amounts to more than half the template size.

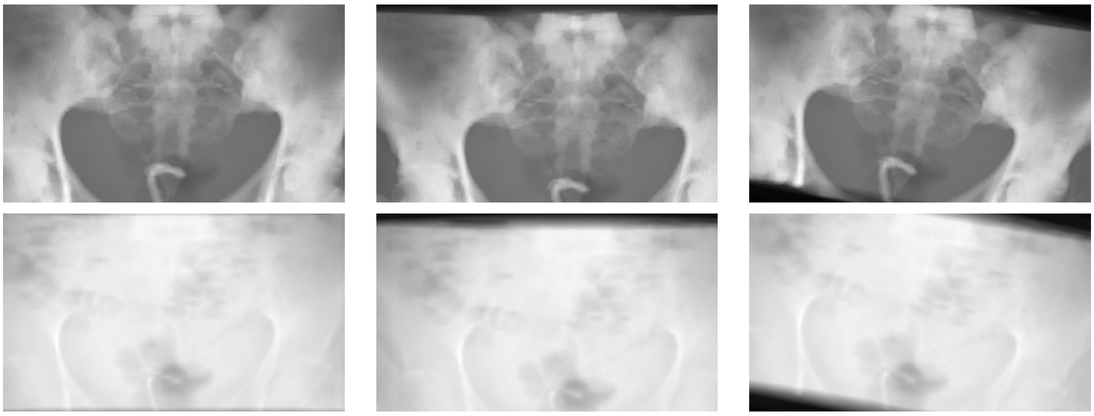
For a typical template configuration in an AP pelvis image, this corresponds to about 20 mm shift (22 pixels) or 10° rotation. However, these are at the same time reasonable upper limits for alignment errors in daily hospital routine.

5.6 Results on artificially generated data

In real datasets of portal images, the ground truth is always unknown. In order to test the algorithm on datasets with known ground truth, two artificial portal image series are generated by computing various megavoltage DRRs from the same CT volume.

5.6.1 In-plane translation and rotations

The following test series consisted of 35 simulated portal images with a maximum patient displacement of 20 mm in x and z direction and a maximum rotation of 10° (figure 5.7). The results are summarized in table 5.2 and figure 5.9a. The standard deviations of the translation measurements are 0.25 pixel (0.23 mm) in x direction and 0.37 pixel (0.33 mm) in z direction. These systematic errors are caused by the unknown y position in the CT coordinate system of the template features (see figure 2.3 for an illustration of the CT coordinate system). Within the rotation



a. reference image **b.** translated image **c.** rotated image

Figure 5.7: In-plane test series computed from CT volume. The CT is translated and rotated in the image plane yielding a test series of 24 translated and 10 rotated images. The top row depicts diagnostic, the bottom row megavoltage DRRs.

measurement, these systematic error do not occur and the standard deviations are below 0.01°.

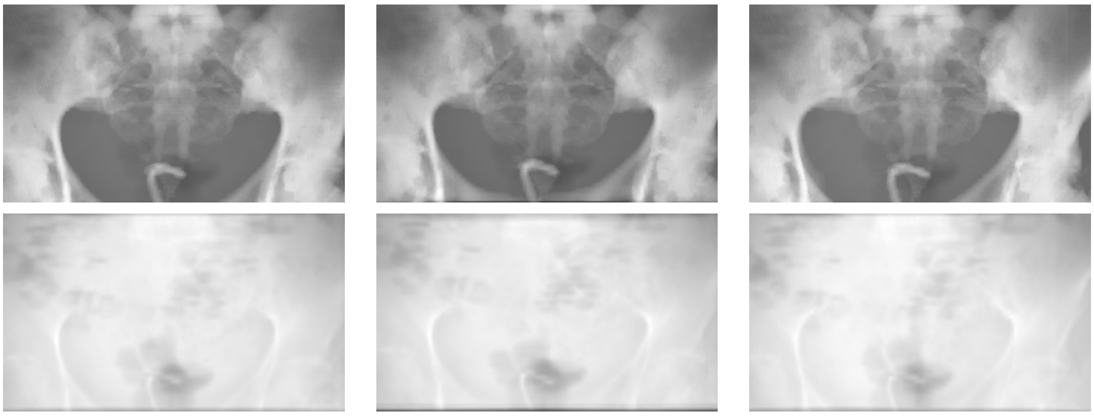
5.6.2 Including out-of-plane rotation

In order to test under more realistic conditions, a test series with small out-of-plane rotations is generated (figure 5.8). The area-based match still finds the corresponding regions with a correlation well above 0.9. The systematic errors already encountered in the example above, which are an inherent problem of using projected images, are of course higher in this example. But the total point errors of 1.24 mm for 2° rotation and 3 mm for 5° still are promising results (see figure 5.9b).

5.7 Results of multi-modal match

Matching portal images directly with a megavoltage DRR computed from the planning CT provides an accurate link between planning and treatment (cf. section 2.4 and figure 2.1a). Additional sources of positioning errors are avoided, thus it is a very interesting approach for increasing the accuracy of radiotherapy treatment.

Figure 5.10a shows a DRR and one set of templates including four validation lines, which are not used for matching. The DRR was computed from a CT volume with a voxel size of $2\times2\times3\,\mathrm{mm}^3$. The corresponding portal image series contained 22 images, two of which are depicted in 5.10b and c. All images were matched three times using different sets of templates. Of these 66 measurements, only one optimization failed, that is, did not find a minimum after 100 iterations. This was due to a rather large patient displacement (over 10 mm), for which one of the template sets was unsuitable. The remaining results were visually validated and all were accepted to be correct, which indicates a success rate of over 95 %.



a. reference image

b. rotated image around x axis

c. rotated image around z axis

Figure 5.8: Out-of-plane rotation test series computed from CT volume with a total of 200 images rotated around the x and z axis, respectively. The top row depicts diagnostic, the bottom row megavoltage DRRs.

	standard deviations			point
	σ_x	σ_z	σ_{rot}	error
in-plane translation	0.23 mm	0.33 mm	0.19°	0.29 mm
in-plane rotation	0.05 mm	0.05 mm	0.01°	0.05 mm
out-of-plane rotation (2°)	1.4 mm	1.1 mm	0.28°	1.2 mm
out-of-plane rotation (5°)	3.3 mm	3.0 mm	0.59°	3.2 mm

Table 5.2: Standard deviations of the displacement measurements in the artificially generated test series.

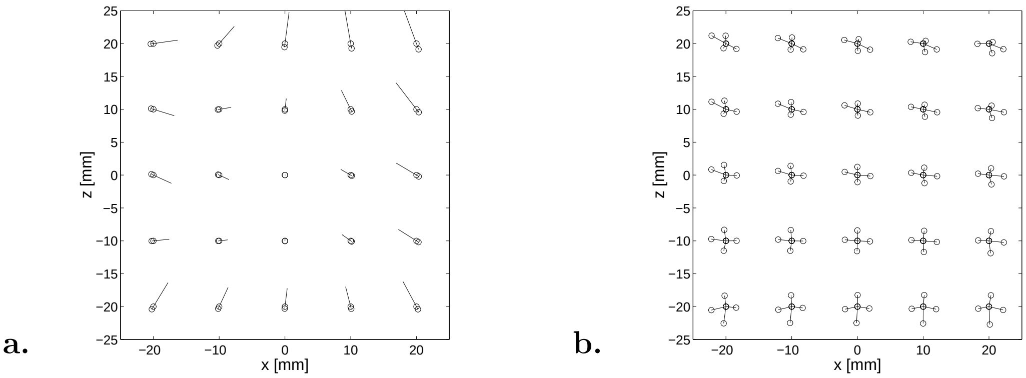
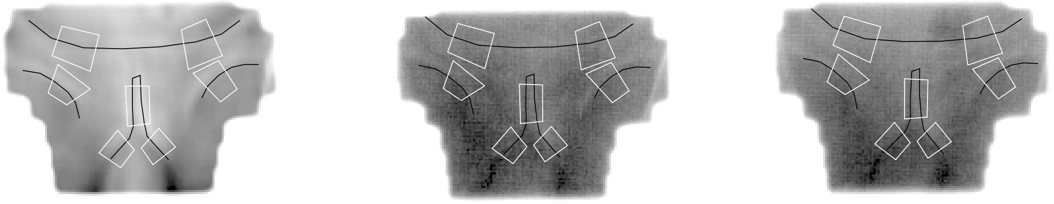


Figure 5.9: Displacement measurement errors in test series without (a) and including out-of-plane rotations of 2° (b). In (a) the error vectors are enlarged by a factor 10. The total point error amounts to 0.3 pixel (0.28 mm) without and 1.4 pixel (1.2 mm) with out-of-plane rotations.



a. reference image (DRR) **b.** translation (0,-4) mm rotation 1.5° **c.** translation (3,-3) mm rotation 2.6°

Figure 5.10: Multi-modal match between a DRR image computed from CT data (a) and a portal image series (b,c). The match is based on an affine transformation model. White polygons outline the template regions and patches respectively, black lines represent validation lines which are not used for matching.

5.8 Clinical validation

In order to clinically validate the suitability of the LSM method, a large test series of 500 portal images and 41 patients have been analyzed both manually and using LSM. This study has been carried out in the course of a medical dissertation in collaboration with this project. More details on the medically relevant outcomes can be found in [Adam 1999]. In the following section, I will summarize the most important findings concerning the performance of the LSM method.

Numerous other medical studies on the use of portal images have been published. The following references are thus merely starting points into this field of work and is not a representative selection. An analysis of the reproducibility of patient positioning was carried out by [Gildersleve *et al.* 1995]. They manually measured translation and rotation of over 200 pelvis fields. The results indicate a 95 % confidence limit in the range of ± 4 mm to ± 8 mm. Recent papers on the application of the feature-based chamfer matching method have been presented by [Leszczynski *et al.* 1998, Cai *et al.* 1998].

[Bijhold *et al.* 1991b] describe a mostly manual method for quantification of patient set-up errors. They claim that displacements can be determined with a variance of 2 mm within one minute using their method, provided that out-of-plane rotations are smaller than 2°. Also, they presented a preliminary study concerning the inter-observer variability. Two observers were asked to measure 15 lateral pelvis images, which resulted in a systematic difference of 4 mm.

An analysis of the prostate motion during during radiotherapy treatment has been presented by [Vigneault *et al.* 1997]. Since the position of the prostate itself can not be determined from portal images, radio-opaque markers were implanted under ultrasound guidance. They found a standard deviation 2.9 mm for the position in lateral and anterior-posterior direction (along x and y axis), as well as 4.6 mm in cephalo-caudal direction (along z axis of the CT coordinate system).

5.8.1 Datasets

In 500 images, patient translation and rotation has been manually measured at least once by an experienced radio-oncologist. Figure 5.11 depicts a few samples, including the line drawings defined for the manual measurements and the template regions used for the automated match. Of this complete dataset, 106 AP pelvis images in 16 series have been measured three times by the same physician, in order to obtain a rough estimate of the variability. The analysis of these manual measurements is presented in the following section.

The same 500 images were then matched using the LSM tool developed during this project. The template regions were defined by a medical student already having some experience with the matching algorithm. To ensure the comparability to the manual measurements, a congruent transformation model was applied. Optimization was aborted after 50 iterations. The evaluation of these LSM results are summarized in section 5.8.3.

5.8.2 Manual measurement results

The AP pelvis image series was manually measured three times. Although the measurements were all carried out by the same physician, an estimate of the reproducibility can be estimated. It is important to note however, that these results should not be interpreted as an estimate of accuracy.

The standard deviations lie in the range of 1 mm for translation and less than 1° for rotation. Table 5.3 gives a detailed overview of the point and 95 % interval estimates of these standard deviations.

			point estimate	interval estimate ($\alpha = 0.05$)
x	std dev	[mm]	0.8	0.69 . . . 0.83
y	std dev	[mm]	0.8	0.75 . . . 0.91
$\sqrt{x^2+y^2}$	std dev	[mm]	1.1	1.05 . . . 1.20
rotation	std dev	[deg]	0.7	0.68 . . . 0.82

Table 5.3: Evaluation of multiple manual measurements on AP pelvis fields. A total of 106 images were matched three times by the same physician. These standard deviations describe the level of reproducibility of the measurements rather than the measurement precision.

5.8.3 Evaluation of automatic measurements

In the following, the match result of the pelvis subset from the test data is presented. Both AP and lateral pelvis fields were included in the evaluation. Whereas AP pelvis

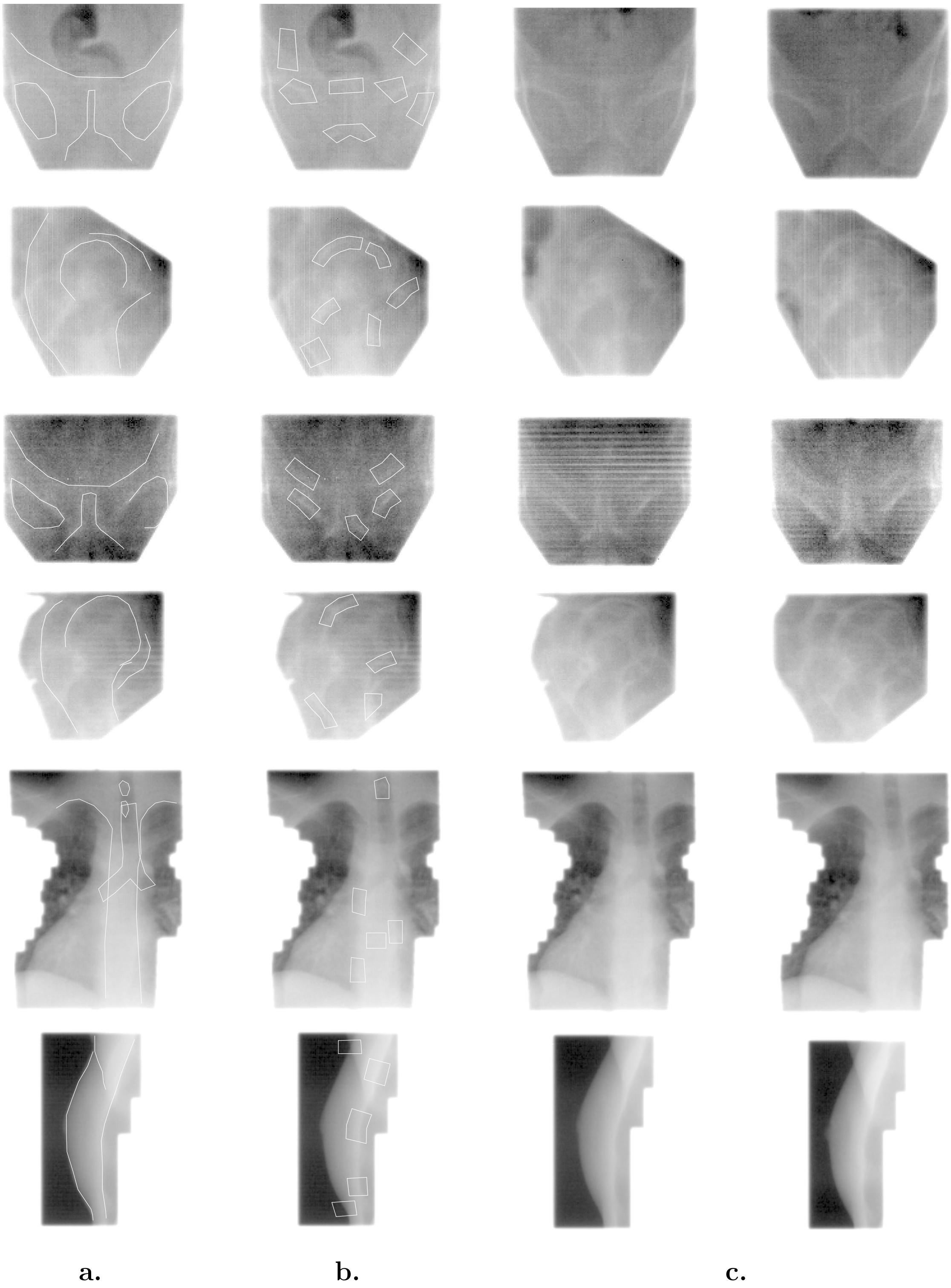


Figure 5.11: Sample of the complete dataset containing 500 images. Column (a) includes the line drawings defined for the manual measurements, whereas column (b) illustrates the choice of template regions for the LSM method. A few additional examples of varying quality are displayed in column (c).

images generally pose the least problems to matching, lateral pelvis images represent the most difficult type of images, since only little useful structure is visible. Thus, both the most suited and the most challenging images were evaluated.

Table 5.4 lists the standard deviations within the 16 image series. The match results were compared to the mean of the manual measurements. No outliers were removed, which is the reason for the relatively large error in series 69337, for instance. Nonetheless, an excellent result was achieved with a point error of 1 mm. The overall errors are summarized in table 5.5, including the 95 % interval estimates of the standard deviations.

ID	images	samples	standard deviation x [mm]	 y [mm]	 $\sqrt{x^2+y^2}$ [mm]	 rot [deg]
63922	5	10	0.5	0.6	0.8	0.3
69337	5	10	1.6	0.7	1.7	1.1
70160	8	16	0.5	0.3	0.6	0.3
70257	7	14	0.6	0.4	0.8	0.5
70298	5	10	0.5	0.5	0.8	0.3
70306	6	12	0.5	0.6	0.8	0.3
70391	7	14	0.3	0.4	0.5	0.6
70439	9	18	0.5	0.9	1.0	0.6
70475	13	26	0.6	0.8	1.0	0.6
70506	7	14	1.1	0.4	1.1	0.7
72340	4	8	0.3	0.2	0.4	0.4
74258	9	18	0.5	0.5	0.7	0.6
74394	13	26	0.6	1.1	1.3	0.8
74426	8	16	0.6	1.0	1.2	1.0
overall	106	212	0.67	0.72	0.98	0.65

Table 5.4: Standard deviations between manual and automatic measurements of all AP pelvis image series. (See table 5.5 for more details on overall statistics.)

As expected, the errors were large in lateral pelvis images. One reason is the little structural information present in such imagery, another are the higher number of outliers. Table 5.6 summarizes the overall errors without any outliers removed, which lead to a point error of 3.5 mm and a standard deviation of 2° for rotation.

By applying an automatic removal of outliers based on the normalized cross correlation, the results improved to 2.9 mm for the point error and 1.7° for rotation. Excluded were only obvious outliers with correlation values below 0.8, which applied to 10 out of 105 match results).

			point estimate	interval estimate ($\alpha = 0.05$)
x	std dev	[mm]	0.67	$0.59 \ldots 0.77$
y	std dev	[mm]	0.72	$0.63 \ldots 0.83$
$\sqrt{x^2+y^2}$	std dev	[mm]	0.98	$0.89 \ldots 1.08$
rotation	std dev	[deg]	0.65	$0.57 \ldots 0.75$

Table 5.5: Statistical comparison of the mean manual measurement with the automatic displacement measurement in an AP pelvis field.

			point estimate	interval estimate ($\alpha = 0.05$)
x	std dev	[mm]	2.1	$1.83 \ldots 2.40$
y	std dev	[mm]	2.8	$2.47 \ldots 3.25$
$\sqrt{x^2+y^2}$	std dev	[mm]	3.5	$3.19 \ldots 3.86$
rotation	std dev	[deg]	2.0	$1.74 \ldots 2.28$

Table 5.6: Statistical comparison of the manual with the automatic displacement measurement in lateral pelvis fields. The significance level α for the interval estimate was set to 0.05.

			point estimate	interval estimate ($\alpha = 0.05$)
x	std dev	[mm]	1.5	$1.30 \ldots 1.76$
y	std dev	[mm]	2.4	$2.13 \ldots 2.88$
$\sqrt{x^2+y^2}$	std dev	[mm]	2.9	$2.59 \ldots 3.21$
rotation	std dev	[deg]	1.7	$1.45 \ldots 1.96$

Table 5.7: The same statistical comparison as above after exclusion of outliers. 20 out of 105 automatic match results were excluded because of a correlation values below 0.8.

6

Conclusion

The method of LSM has been investigated with regard to its application to low contrast imagery. The main goal was to improve the precision and reliability of patient position during high precision radiotherapy. In order to attain this higher accuracy, *electronic portal images* are the most promising source of information, though they inherently suffer from low contrast.

6.1 Analysis of low contrast imagery

The main advantage of LSM, applied to low contrast images, is that no feature extraction step is necessary. The design of a robust and reliable feature extractor for such images is a very challenging task. The difficulty is to find a method that reliably transforms the image information into a more useful form, in particular into a set of features. In images with sufficient contrast to noise ratio, this is usually feasible if the requirements on accuracy is rather low, within a few pixels. This is the case in many typical computer vision applications and hence the popularity of, for instance, edge detectors.

However, to find a suitable extraction scheme for low contrast imagery is by far more difficult. Since only little information is available, the robustness of a feature extractor becomes crucial. Simple approaches such as applying standard edge detectors are bound to fail. Thus, instead of searching for an optimal feature extractor, I have chosen the area-based LSM method, circumventing this difficult preprocessing step.

In this work, I have shown that the application of LSM to this type of imagery is a suitable approach. Being an area-based method, LSM takes into account the full image information, which is an important aspect for analyzing portal images. Furthermore, in contrast to most correlation methods, it allows for the estimation of, for instance, an affine transformation or more complex transformation models.

Compared to previous applications, I thoroughly investigated the diagnostic capabilities within the LSM framework. Self-diagnosis is an important concept and the cornerstone of robust and reliable image analysis. Within LSM, correct error propagation is possible due to the least squares approach. Hence, many standard statistical measures can be computed, leading to the field of parameter estimation theory. It is, however, not always straightforward to apply general methods from parameter estimation theory to LSM. I have discussed that one problem is the estimation of the a priori errors present in images. Without precise knowledge about this a priori noise level, a comparison of the a priori with the a posteriori noise is usually not reasonable. Nevertheless, this measure can often be replaced by the normalized cross correlation.

Another problem is the unknown gray value dependency of adjacent pixels. Such covariances have to be known up to a constant factor to correctly formulate the LSM problem. A common assumption is that pixel values are independent, neglecting any covariances. I took the same approach, mainly due to computational constraints, but not before analyzing the effects of Gaussian filtering and bilinear interpolation onto the gray value covariance. Neglected covariances of adjacent pixels mainly influence the computation of parameter precision, yielding too optimistic estimates.

Since LSM is based on an optimization scheme, the determinability of a parameter set is an important aspect. An existing determinability measure has been slightly modified to render it useful for LSM. This measure is a sum of local, pixel-wise determinability, neglecting their respective locations. In particular, it is not a global measure, which might lead to unexpected results in certain cases as discussed in section 3.4. I also discussed that this measure is best applied in combination with an analysis of the parameter correlation. Correlated parameters strongly influence this determinability measure and should be detected beforehand.

In case of a weakly determinable parameter set, additional constraints are introduced within the LSM framework. This is straightforward for linear constraints, however nonlinear constraints must be linearized first. The linearization of a two dimensional distance constraint has been investigated in more detail in section 3.9. I found that the simple linearization by Taylor expansion may lead to oscillation and can be improved based on geometric considerations.

Besides detailed insight into the self-diagnostic measures, another contribution of this work is the extension of LSM to multiple templates. This allows local geometric distortions to be accounted for without losing the global context. However, the large increase in degrees of freedom is difficult to cope with in low contrast imagery. Therefore, this extension was not used within the application to portal images.

The generic LSM tool developed during this project is hopefully not only useful for portal imaging, but applicable to other types of imagery as well. It allows the precise estimation of two dimensional affine transformations using a selection of template regions, with the option of applying various diagnostic measures.

6.2 Patient positioning in radiotherapy

Electronic portal images are a valuable source of information for improving the reliability and accuracy of patient positioning. Since portal images are acquired during the actual radiotherapy treatment, they are an ideal basis for checking the treatment setup. However, this basis can only be fully exploited when other sources of information are included in the analysis:

- Multi-leaf collimator settings

  In order to calibrate a portal image, the distinct structure of the field edge serves as calibration standard. The optimum reference for this field edge shape is a simulated edge based on the settings of the multi-leaf collimator (MLC). Hence, possible setup errors do not influence the image calibration.

- Planning data

  One goal of the portal image analysis is checking the field setup. Since the planning data prescribes the beam directions and field shapes, it would be optimal to directly employ this data as reference data. In contrast to the calibration step, the reference field edge is therefore not simulated based on the MLC settings but on the prescribed field shape.

- CT volume

  When a medium to high resolution CT volume is available, optimum reference images can be computed, megavoltage DRRs. The requirement is a slice thickness of less than 5 mm.

Alas, these three sources have not been readily accessible within this project, due to missing system integration of the various devices. Nonetheless, suitable interim solutions have been found as I will present in the following. It is important to note that these interim solutions only differ in the choice of the reference data and not of the method. Once the optimum reference data is available, it will be possible to apply the same algorithm.

The first step in portal image analysis is the calibration of the input image. As mentioned above, the MLC settings are often not available, and an interim solution had to be chosen. That is, the reference edge is extracted from a validated portal image. For both, the optimum and the interim reference data, the area along the reference edge is used as template region and matched onto this distinct structure in the newly acquired treatment image. This match using LSM proved to be a robust and accurate calibration method.

The second step is to check the shape of the field edge, comparing it to the prescribed field shape. If the planning data is not accessible, as in this project, the field edge of a validated portal image must be used instead. As similarity measure served the normalized cross correlation of the gray values. Experiments have shown that this measure robustly detects even minor shape deviations given that the portal image has been accurately calibrated. Due to the close relation of calibration based on the field edge and checking the field shape, these two steps have often been combined, rendering reliable checking impossible. By separating these steps and by employing LSM for precise calibration, I am able to provide a reliable checking procedure.

The last and crucial step is the displacement measurement of bony structures. The goal is to estimate parts of the rigid, three dimensional patient movement. The results obtained within this project indicate that the contrast to noise ratio of portal images is still too low to allow for an estimation of the complete patient motion. Even in portal images of relatively good quality, it occurred that a two dimensional affine transformation model was not determinable. Therefore, the problem had to be reduced to only measuring the patient displacement parallel to the image plane.

The systematic error, introduced by this problem reduction to two dimensions, has been investigated using a computed test series. This series consisted of over 200 megavoltage DRRs (AP pelvis) with varying out-of-plane rotations. The error analysis yielded systematic point errors of 1.2 mm for 2°, and 3.2 mm for 5° out-of-plane rotation. This is a satisfactory result, considering that rather large displacements of up to ± 20 mm had to be measured in this test series.

An important aspect is the choice of template regions for the anatomy match. At this point, knowledge about the anatomy can help to improve the match results. Choosing structures, for instance, which are instable under small out-of-plane rotations, will have a negative effect on the measurements.

The influence of minor variation in position of the template regions was investigated in another experiment. One portal image series was matched several times, each time moving the templates by a few pixels. The resulting point error of 0.3 pixel or 0.25 mm indicates robustness against slight template variations and represents at the same time an estimate of the measurement precision. Thus, the operator dependence of the algorithm could be kept to a minimum.

The drawback in using LSM as matching algorithm is that the reference image must contain similar gray value characteristics. In particular, the diagnostic X-ray image acquired during treatment simulation is not a suitable reference image. Such simulator images are used in most conventional procedures. Nevertheless, there are two possible choices of reference images: a megavoltage DRR or a validated portal image. The former is the optimum reference in terms of providing an accurate link between the planning step and the actual treatment. However, a high resolution CT, necessary to compute such a DRR, was mostly not available within this project. Thus, a validated portal image has been used as reference image. Still, the feasibility of matching a portal image to a DRR has been shown in the case of an AP pelvis series, yielding promising results.

High precision is required in the field of stereotactic radiosurgery, dealing with the treatment of brain tumors. A test series with known displacements of the fixation device, the headrest, lead to an estimate of the overall point error of 0.2 mm, which is in the same range as the repositioning accuracy of the headrest itself. Although a rather simple setup, this is an important finding since LSM thusly proves to have potential for high precision measurements in X-ray imagery.

6.3 Clinical validation and application

The predominant existing technique for analyzing portal images remains the manual, two dimensional measurement of patient displacement. Thus, the performance of

LSM was compared to manual measurements in a large test series of over 500 portal images. All images were measured at least once by an experienced radio-oncologist. In order to estimate the measurement variability, about 100 AP pelvis images were matched three times by the same physician. The resulting standard deviation of 1.1 mm represents an estimate for the reproducibility, and hence an upper bound for the accuracy of manual methods.

The comparison of these results with the outcome of the LSM match showed varying agreement. In AP pelvis images, for instance, the total point error is about 1 mm, which is almost perfect performance, considering the aforementioned standard deviation of the manual measurements. In lateral pelvis images however, this point error reaches 3.5 mm. Nonetheless, this is in direct correspondence to the low information content of lateral pelvis images. It has been shown in earlier work that the inter-observer variability on such imagery amounts to 4 mm.

Importantly, the algorithm found a reasonable match in over 90 % of all tested portal images, without supplying an initial guess. Moreover, the examination of the test results indicates that the normalized cross correlation coefficient together with plausibility checks are a suitable measure to detect the remaining mismatches. Hence, LSM is definitely an improvement compared to existing semi-automated methods, which attain about 70 % success rate and yield less precise measurements.

This leads to the conclusion that the LSM method applied to portal images has the potential of improving patient positioning and thus quality assurance in radiotherapy. In order to introduce the novel method into the daily hospital routine, the LSM tool has been installed at the department of radio-oncology at the University Hospital of Zürich, with further installations planned for the near future.

6.4 Outlook

A first improvement relates to the remaining manual part of the matching algorithm, the definition of the template regions. This operator interaction could be further reduced by offering computer assistance. For instance by automatically suggesting a set of templates, out of which the operator may select and possibly modify the relevant regions.

In order to fully exploit the potential of electronic portal images, the next step is to integrate the various parts of radiotherapy treatment to a complete system. Planning data, CT volume, simulation images, and treatment settings should all be accessible when analyzing portal images. This will render it possible to use the optimum instead of interim solutions, as outlined in the previous sections. Thus, several additional sources of errors can be avoided, increasing the overall reliability of the treatment.

Furthermore, this integration is a prerequisite for solving the complete three dimensional problem. However, another requirement are improvements in the quality of portal images. Significant enhancements in terms of contrast are promised by the relatively new technology of amorphous silicon panels. Once this technology is readily available, it should be feasible to measure more complex transformations and thus to estimate the complete patient movement.

The most promising approach is the comparison of the current portal image to a series of DRRs, optimizing the pose of the CT volume until the most similar DRR is found. Again, the area-based LSM method could be employed to help the optimization and ensure highly accurate results. Thus, tight system integration together with an improvement of contrast in portal images will lead the way to accurate, three dimensional patient positioning.

Bibliography

[Adam 1999] Lukas Adam. MD thesis, University of Zürich, to be submitted, 1999.

[Balter *et al.* 1992] J. M. Balter, Ch. A. Pelizzari, and G. T. Y. Chen. Correlation of projection radiographs in radiation therapy using open curve segments and points. *Med. Phys.*, 19(2):329–334, March 1992.

[Bansal *et al.* 1998] R. Bansal, L. H. Staib, Zhe Chen, A. Rangarajan, J. Knisely, R. Nath, and J. S. Duncan. A novel approach for the registration of 2D portal and 3D CT images for treatment setup verification in radiotherapy. In *Medical Image Computing and Computer–Assisted Intervention — MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 1075–1086. Springer, 1998.

[Barrow *et al.* 1977] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching. In *Proc. of the fifth Int'l Joint Conf. on Artif. Intelligence*, pages 659–663, 1977.

[Bergen *et al.* 1992] J. R. Bergen, P. Anandan, K. J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *Computer Vision – ECCV '92*, volume 588 of *Lecture Notes in Computer Science*, pages 237–252. Springer-Verlag, May 1992.

[Berger and Danuser 1997] Martin Berger and Gaudenz Danuser. Deformable multi template matching with application to portal images. In *Proceedings Computer Vision and Pattern Recognition '97*, pages 374–379. IEEE Computer Society Press, June 1997.

[Berger and Gerig 1998] Martin Berger and Guido Gerig. Motion measurements in low-contrast X-ray imagery. In *Medical Image Computing and Computer–Assisted Intervention — MICCAI'98*, number 1496 in Lecture Notes in Computer Science, pages 832–841. Springer, 1998. Longer version available as BIWI TR 183 (http://www.vision.ee.ethz.ch).

[Bijhold *et al.* 1991a] J. Bijhold, K. G. A. Gilhuijs, M. van Herk, and H. Meertens. Radiation field edge detection in portal images. *Phys. Med. Biol.*, 36(12):1705–1710, December 1991.

[Bijhold *et al.* 1991b] J. Bijhold, M. van Herk, R. Vijlbrief, and J. V. Lebesque. Fast evaluation of patient set-up during radiotherapy by aligning features in portal and simulator images. *Phys. Med. Biol.*, 36(12):1665–1679, December 1991.

[Bijhold *et al.* 1992] J. Bijhold, K. G. A. Gilhuijs, and M. van Herk. Automatic verification of radiation field shape using digital portal images. *Med. Phys.*, 19(4):1007–1014, 1992.

[Bijhold 1993] J. Bijhold. Three-dimensional verification of patient placement during radiotherapy using portal images. *Med. Phys.*, 20(2):347–356, 1993.

[Borgefors 1988] G. Borgefors. Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(6):849–865, November 1988.

[Box 1971] M. J. Box. Bias in nonlinear estimation. *Journal of the Royal Statistical Society*, 33(2):171–202, 1971.

[Boyer *et al.* 1992] A. L. Boyer, L. Antonuk, A. Fenster, M. van Herk, H. Meertens, P. Munro, L. E. Reinstein, and J. Wong. A review of electronic portal imaging devices (EPIDs). *Med. Phys.*, 19(1):1–16, 1992.

[Brunie *et al.* 1993] L. Brunie, S. Lavallée, J. Troccaz, P. Cinquin, and M. Bolla. Pre- and intro-irradiation multimodal image registration: principles and first experiments. *Radiother. Oncol.*, 29(2):244–252, November 1993.

[Cai *et al.* 1998] J. Cai, J. C. Chu, V. A. Saxena, and L. H. Lanzl. A simple algorithm for planar image registration in radiation therapy. *Med. Phys.*, 25(6):824–829, June 1998.

[Canny 1983] J. F. Canny. Finding Edges and Lines in Images. Technical Report 720, MIT Artificial Intelligence Laboratory, Dept. of Electrical Engineering and Computer Science, M.I.T., Cambridge Mass., 1983.

[Chaney *et al.* 1995] E. L. Chaney, J. S. Thorn, G. Tracton, T. Cullip, J. G. Rosenman, and J. E. Tepper. A portable software tool for computing digitally reconstructed radiographs. *Int'l J. Radiation Oncology Biol. Phys.*, 32(2):491–497, May 1995.

[Danuser and Mazza 1996] G. Danuser and E. Mazza. Observing deformations of 20 nanometer with a low numerical aperture light microscope. In C. Gorecki, editor, *Optical Inspection and Micromeasurements*, volume 2782 of *Proceedings SPIE*, pages 180–191. European Optical Society, 1996.

[Danuser and Stricker 1998] G. Danuser and M. Stricker. Parametric model fitting: From inlier characterization to outlier detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(2):263–280, March 1998.

[Danuser 1996] G. Danuser. Stereo light microscope calibration for 3D submicron vision. In K. Kraus and P. Waldhäusl, editors, *Proceedings of the 18th ISPRS Congress*, volume 31/B5, pages 101–108, Vienna, Austria, July 1996. ISPRS.

[Danuser 1997] G. Danuser. *Quantitative Stereo Vision for the Stereo Light Microscope: An Attempt to Provide Control Feedback for a Nanorobot System.* PhD thesis, Swiss Federal Institute of Technology, Zürich, April 1997. Diss. ETH No. 12191. ISBN 3-905588-00-5.

[Dempster *et al.* 1977] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society, ser. B*, 39:1–38, 1977.

[Dennis *et al.* 1981] J. E. Dennis, D. M. Gay, and R. E. Welsch. An adaptive nonlinear least-squares algorithm. *ACM Transactions on Mathematical Software*, 7(3):348–368, September 1981.

[Ding *et al.* 1993] G. X. Ding, S. Shalev, and G. Gluchev. A ρ–θ technique for treatment verification in radiotherapy and its clinical applications. *Med. Phys.*, 20(4):1135–1143, July 1993.

[Dong and Boyer 1995] L. Dong and A. L. Boyer. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int'l J. Radiation Oncology Biol. Phys.*, 33(5):1053–1060, December 1995.

[Dong and Boyer 1996] L. Dong and A. L. Boyer. A portal image alignment and patient setup verification procedure using moments and correlation techniques. *Phys. Med. Biol.*, 41(4):697–723, April 1996.

[Douglas and Peucker 1973] D. H. Douglas and T. K. Peucker. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 10(2):112–122, 1973.

[Eilertsen *et al.* 1994] K. Eilertsen, A. Skretting, and T. L. Tennvassås. Methods for fully automated verification of patient set-up in external beam radiotherapy with polygon shaped fields. *Phys. Med. Biol.*, 39:993–1012, 1994.

[Förstner 1987] W. Förstner. Reliability analysis of parameter estimation in linear models with applications to mensuration problems in computer vision. *Computer Vision, Graphics, and Image Processing*, 40:273–310, 1987.

[Fritsch *et al.* 1995] D. S. Fritsch, E. L. Chaney, A. Boxwala, M. McAuliffe, S. Raghavan, A. Thall, and J. R. D. Earnhart. Core-based portal image registration for automatic radiotherapy treatment verification. *Int'l J. Radiation Oncology Biol. Phys.*, 33(5):1287–1300, December 1995.

[Gildersleve *et al.* 1995] J. Gildersleve, D. P. Dearnaley, P. M. Evans, and W. Swindell. Reproducibility of patient positioning during routine radiotherapy, as assessed by an integrated megavoltage imaging system. *Radiother. Oncol.*, 35(2):151–160, May 1995.

[Gilhuijs and van Herk 1993] K. G. A. Gilhuijs and M. van Herk. Automatic on-line inspection of patient setup in radiation therapy using digital portal images. *Med. Phys.*, 20(3):667–677, 1993.

[Gilhuijs *et al.* 1995] K. G. A. Gilhuijs, A. Touw, M. van Herk, and R. E. Vijlbrief. Optimization of automatic portal image analysis. *Med. Phys.*, 22(7):1089–1099, July 1995.

[Gilhuijs *et al.* 1996] K. G. A. Gilhuijs, P. J. H. van de Ven, and M. van Herk. Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data. *Med. Phys.*, 23(3):389–399, March 1996.

[Gill and Murray 1978] P. E. Gill and W. Murray. Algorithm for the solution of the nonlinear least-squares problem. *SIAM J. Numer. Anal.*, 15(0):977–992, 1978.

[Gill *et al.* 1981] P. E. Gill, W. Murray, and M. H. Wright. *Practical optimization.* Academic Press, 1981.

[Gleicher 1997] M. Gleicher. Projective registration with difference decomposition. In *Proceedings Computer Vision and Pattern Recognition '97*, pages 331–337. IEEE Computer Society Press, June 1997.

[Grün 1985] A. Grün. Adaptive least squares correlation: A powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing & Cartography*, 14(3):175–187, 1985.

[Grün 1986] A. Grün. Photogrammetrische Punktbestimmung mit der Bündelmethode. Technical Report 40, Institut für Geodäsie und Photogrammetrie, ETH Zürich, March 1986.

[Hager and Belhumeur 1998] G. D. Hager and P. N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(10):1025–1039, October 1998.

[Höpcke 1980] Walter Höpcke. *Fehlerlehre und Ausgleichsrechnung.* de Gruyter, 1980.

[Hu 1962] M.-K. Hu. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, IT-8(2):179–187, February 1962.

[ICRU 1976] Determination of absorbed dose in a patient irradiated by beams of X or gamma rays in radiotherapy procedures. International Commission on Radiation Units and Measurements. ICRU report 24, September 1976.

[Jianbo Shi and Tomasi 1994] Jianbo Shi and Carlo Tomasi. Good features to track. In *Proceedings Computer Vision and Pattern Recognition '94*, pages 593–600. IEEE Computer Society Press, June 1994.

[Jones and Boyer 1991] S. M. Jones and A. L. Boyer. Investigation of an FFT-based correlation technique for verification of radiation treatment setup. *Med. Phys.*, 18(6):1116–1125, 1991.

[Koch 1988] K. R. Koch. *Parameter Estimation and Hypothesis Testing in Linear Models.* Springer-Verlag, 1988.

[Kunt 1982] M. Kunt. Edge detection: A tutorial review. In *Proceedings of ICASSP 82*, volume 2, pages 1172–1175, 1982.

[Lacroix 1988] V. Lacroix. A three-module strategy for edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(6):803–810, November 1988.

[Leszczynski *et al.* 1992] K. W. Leszczynski, S. Shalev, and N. S. Cosby. The enhancement of radiotherapy verification images by an automated edge detection technique. *Med. Phys.*, 19:611–621, 1992.

[Leszczynski *et al.* 1998] K. W. Leszczynski, S. Loose, and S. Boyko. An image registration scheme applied to verification of radiation therapy. *Br. J. Radiol.*, 71(844):413–426, April 1998.

[Lindeberg 1995] T. Lindeberg. Direct estimation of affine image deformations using visual front-end operations with automatic scale selection. In *Fifth International Conference on Computer Vision ICCV'95*, pages 134–141, Cambridge, MA, 1995. IEEE Computer Society Press.

[Lucas and Kanade 1981] Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *International joint conference on artificial intelligence*, pages 674–679, 1981.

[Marmier and Sheldon 1969] P. Marmier and E. Sheldon. *Physics of Nuclei and Particles.* Academic Press, 1969.

[Marquardt 1963] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *J. of the Society for Industrial and Applied Mathematics*, 11:431–441, 1963.

[McParland and Kumaradas 1995] B. J. McParland and J. C. Kumaradas. Digital portal image registration by sequential anatomical matchpoint and image correlations for real-time continuous field alignment verification. *Med. Phys.*, 22(7):1063–1075, July 1995.

[Meertens *et al.* 1990] H. Meertens, J. Bijhold, and J. Strackee. A method for the measurement of field placement errors in digital portal images. *Phys. Med. Biol.*, 35(3):299–323, 1990.

[Moseley and Munro 1994] J. Moseley and P. Munro. A semiautomatic method for registration of portal images. *Med. Phys.*, 21(4):551–558, April 1994.

[Papoulis 1991] Athanasios Papoulis. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, 3rd edition, 1991.

[Press *et al.* 1994] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C.* Cambridge University Press, 2nd edition, 1994.

[Radcliffe *et al.* 1994] T. Radcliffe, R. Rajapakshe, and S. Shalev. Pseudocorrelation: A fast, robust, absolute, grey-level image alignment algorithm. *Med. Phys.*, 21(6):761–769, June 1994.

[Rosenfeld and Pfaltz 1966] A. Rosenfeld and J. L. Pfaltz. Sequential operations in digital picture processing. *Journal of the Association for Computing Machinery*, 13(4):471–494, October 1966.

[Schonemann 1966] P. H. Schonemann. A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, 31:1–10, 1966.

[Serra 1982] J. Serra. *Image Analysis and Mathematical Morphology*, volume 1. Academic Press, 1982.

[Sherouse *et al.* 1990] G. W. Sherouse, K. Novins, and E. L. Chaney. Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *Int'l J. Radiation Oncology Biol. Phys.*, 18(3):651–658, March 1990.

[Styner 1997] Martin Styner. Digital simulation of X-ray images from CT volumes. Master's thesis, ETH Zürich, February 1997. Original German title: *Digitale Simulation von Röntgenbildern aus CT-Daten.*

[Teh and Chin 1988] Ch-H. Teh and R. T. Chin. On image analysis by the methods of moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):496–513, July 1988.

[Thévenaz *et al.* 1998] P. Thévenaz, U. E. Ruttimann, and M. Unser. A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing*, 7(1):27–41, January 1998.

[Unser *et al.* 1995] M. Unser, P. Thévenaz, L. Chulhee, and U. Ruttimann. Registration and statistical analysis of PET images using the wavelet transform. *IEEE Engineering in Medicine and Biology*, September/October 1995.

[van Herk and Meertens 1988] M. van Herk and H. Meertens. A matrix ionisation chamber imaging device for on-line patient setup verification during radiotherapy. *Radiother. Oncol.*, 11(4):369–378, April 1988.

[Vigneault *et al.* 1997] E. Vigneault, J. Pouliot, J. Laverdière, J. Roy, and M. Dorion. Electronic portal imaging device detection of radioopaque markers for the evaluation of prostate position during megavoltage irradiation: A clinical study. *Int'l J. Radiation Oncology Biol. Phys.*, 37(1):205–212, January 1997.

[Voorhess and Poggio 1987] H. Voorhess and T. Poggio. Detecting blobs as textons in natural images. In *Image Understanding Workshop*, volume 2, pages 892–899. DARPA, February 1987.

[Wang and Fallone 1994] H. Wang and B. G. Fallone. A robust morphological algorithm for automatic radiation field extraction and correlation of portal images. *Med. Phys.*, 21(2):237–244, February 1994.

[Wang and Fallone 1995] H. Wang and B. G. Fallone. A mathematical model of radiation field edge localization. *Med. Phys.*, 22(7):1107–1110, July 1995.

[Yin *et al.* 1994] F. F. Yin, M. C. Schell, and P. Rubin. Input/output characteristics of a matrix ion-chamber electronic portal imaging device. *Med. Phys.*, 21(9):1447–1454, September 1994.

A

Algorithmic Details

A.1 Comparison of least squares error and normalized cross correlation

This section deals with the comparison between finding the minimum least squares error and finding the maximum correlation of two signals f and g, depending on an arbitrary set of parameters. The least squares error is defined as

$$\begin{aligned} e^2 &= \sum (f-g)^2 \\ &= \sum f^2 + \sum g^2 - 2\sum f \cdot g \,, \end{aligned}$$

opposed to the normalized cross correlation

$$\rho = \frac{\sum f \cdot g}{(\sum f^2 \cdot \sum g^2)^{\frac{1}{2}}} \,,$$

where we assumed without loss of generality that f and g have zero mean.

For large enough samples, it is usually safe to assume the terms $\hat{v}_f = \sum f^2$ and $\hat{v}_g = \sum g^2$ to be constant over the parameter search space. Under this assumption, the equations can be rewritten as

$$\begin{aligned} e^2 &\approx c_1 - 2\sum f \cdot g \\ \rho &\approx c_2 \cdot \sum f \cdot g \,, \end{aligned}$$

where $c_1 = \hat{v}_f + \hat{v}_g$ and $c_2 = (\hat{v}_f \, \hat{v}_g)^{-\frac{1}{2}}$ represent constant values. Thus minimizing e^2 is equivalent to maximizing ρ. Furthermore, if the two estimated variances are equal, that is $\hat{v}_f = \hat{v}_g = \hat{v}$, the following simple linear expression is found for the correlation coefficient in dependence of the least squares error

$$\rho = 1 - \frac{e^2}{2\hat{v}} \,,$$

where $e^2 \in [0..2\sqrt{\hat{v}}\,]$.

A.2 Comparison of Levenberg–Marquardt and Gauss–Newton methods

The unconstrained NLS problem defined by equation (3.11) is usually solved either by the Gauss–Newton [Gill *et al.* 1981, Dennis *et al.* 1981, Gill and Murray 1978] or Levenberg–Marquardt method [Press *et al.* 1994, Marquardt 1963]. The Gauss–Newton method is based on the linearization of the observation equation, which was outlined in the previous section. The Levenberg–Marquardt method on the other hand is often described as taking into account the first and second derivatives of the goal function e^TPe. In the following, the two approaches are compared and it is shown that they are very similar in most implementations.

The goal function of problem (3.11) is given by

$$e^TPe = \left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right)^T P \left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right) .$$

The first derivatives with respect to the parameters ζ at the current estimate $\hat{\zeta}^\circ$ can thus be written as

$$\begin{aligned}
\frac{\partial}{\partial \zeta_i}\left(e^TPe\right) &= 2\,\frac{\partial}{\partial \zeta_i}\left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right)^T P \left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right) \\
\frac{\partial}{\partial \zeta_i}\left(e^TPe\right)\Big|_{\zeta=\hat{\zeta}^\circ} &= 2\,\frac{\partial}{\partial \zeta_i} F_1^*(\zeta, l_1)^T\Big|_{\zeta=\hat{\zeta}^\circ} P \left(F_1^*(\hat{\zeta}^\circ, l_1) + F_2^*(l_2)\right) \\
&= 2\,A_i^T P\,w\,, \qquad \text{(A.1)}
\end{aligned}$$

where A_i is column i of the Jacobian $\nabla_\zeta F_1^*$ and $w = -F_1^*(\hat{\zeta}^\circ, l_1) - F_2^*(l_2)$ as defined in the previous section. The second derivatives are defined analogous by

$$\begin{aligned}
\frac{\partial^2}{\partial \zeta_i \partial \zeta_j}\left(e^TPe\right) &= 2\,\frac{\partial}{\partial \zeta_i}\left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right)^T P\,\frac{\partial}{\partial \zeta_j}\left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right) \\
&\quad + 2\frac{\partial^2}{\partial \zeta_i \partial \zeta_j}\left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right)^T P \left(F_1^*(\zeta, l_1) + F_2^*(l_2)\right) \\
\frac{\partial^2}{\partial \zeta_i \partial \zeta_j}\left(e^TPe\right)\Big|_{\zeta=\hat{\zeta}^\circ} &= 2\,A_i^T P A_j + 2\,\frac{\partial^2}{\partial \zeta_i \partial \zeta_j} F_1^*(\zeta, l_1)^T\Big|_{\zeta=\hat{\zeta}^\circ} P\,w\,.
\end{aligned}$$

However, the term including the second derivatives is neglected in the Levenberg–Marquardt method. This is justified by the fact that for a successful model and in the neighborhood of the solution, $\|w\|$ is often small compared to $\|A_i^T A_j\|$. Please refer to [Gill *et al.* 1981, section 4.7] or [Press *et al.* 1994, p. 683] for more details. This assumption leads to the approximation

$$\frac{\partial^2}{\partial \zeta_i \partial \zeta_j}\left(e^TPe\right)\Big|_{\zeta=\hat{\zeta}^\circ} \approx 2\,A_i^T P A_j\,. \qquad \text{(A.2)}$$

These expressions for the first (A.1) and second derivatives (A.2) are used to form the Gauss–Newton part of the Levenberg–Marquardt approach

$$A^TPA\,\Delta\zeta = -A^TP\,w\,, \qquad \text{(A.3)}$$

which is identical to the equation (3.14).

Hence, the only difference between the Gauss–Newton and the Levenberg–Marquardt method lies in the additional steepest-descent part within the latter, which is introduced in the following way

$$\left(A^TPA + \lambda I\right) \, \Delta\zeta = -A^TP\,w\,. \tag{A.4}$$

The factor λ defines the dominant method. If $\lambda = 0$, this equation is identical to equation (A.3) and and $\Delta\zeta$ is the Gauss–Newton direction. As $\lambda \to \infty$, $\|\Delta\zeta\|$ and $\Delta\zeta$ becomes parallel to the steepest-descent direction.

A.3 Variance and covariance after Gaussian filtering

The design of the weight matrix was outlined in section 3.10. In the following, several aspects are described in more detail. Please note that all pixel values are assumed to be independent and normally distributed with an identical variance σ_0^2 as defined by equation (3.43).

It is common practice to filter the original image signal with a Gaussian kernel to reduce noise. A two dimensional and isotropic Gaussian filter is given by sampling the continuous form

$$G(\sigma_G, x) = \frac{1}{2\pi\,\sigma_G^2}\,e^{-\frac{x_1^2+x_2^2}{2\,\sigma_G^2}}$$

at integer coordinates. Please note the continuous form is normalized such that the definite integral

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G(\sigma_G, x)\,dx_1\,dx_2 = 1 \tag{A.5}$$

equals to one. However, this is not true in general for the infinite sum of a sampled Gaussian. The filter coefficients have to be corrected such that the sum of the coefficients still yields unity, in order to avoid amplification of the image signal. Including this correction factor, the following expression is found for the discrete Gaussian filter:

$$G[\sigma_G, u] = \frac{e^{-\frac{u_1^2+u_2^2}{2\,\sigma_G^2}}}{\sum_{u_1}\sum_{u_2} e^{-\frac{u_1^2+u_2^2}{2\,\sigma_G^2}}}\,. \tag{A.6}$$

The reduction of the signal variance depends on the chosen filter width σ_G and on the filter support. Exploiting assumption (3.43), we derive the following expression for the variance reduction of a discrete Gaussian filter with the limited support $(2m+1) \times (2m+1)$:

$$\sigma_0^{*2}(\sigma_G, m) = \sigma_0^2 \;\frac{\sum_{i=-m}^{m}\sum_{j=-m}^{m} e^{-\frac{i^2+j^2}{\sigma_G^2}}}{\left(\sum_{i=-m}^{m}\sum_{j=-m}^{m} e^{-\frac{i^2+j^2}{2\sigma_G^2}}\right)^2}\,. \tag{A.7}$$

For the typical filter support sizes 5×5 and 7×7, equation (A.7) is further simplified to:

$$\begin{aligned}
\sigma_0^{*2}(\sigma_G, m{=}2) &= \sigma_0^2 \, \frac{\left(2 + 2e^{\frac{3}{\sigma_G^2}} + e^{\frac{4}{\sigma_G^2}}\right)^2}{\left(2 + 2e^{\frac{3}{2\sigma_G^2}} + e^{\frac{4}{2\sigma_G^2}}\right)^4} \\
\sigma_0^{*2}(\sigma_G, m{=}3) &= \sigma_0^2 \, \frac{\left(2 + 2e^{\frac{5}{\sigma_G^2}} + 2e^{\frac{8}{\sigma_G^2}} + e^{\frac{8}{\sigma_G^2}}\right)^2}{\left(2 + 2e^{\frac{5}{2\sigma_G^2}} + 2e^{\frac{8}{2\sigma_G^2}} + e^{\frac{9}{2\sigma_G^2}}\right)^4} \, .
\end{aligned}$$

When we further choose $\sigma_G{=}1$ the resulting pixel variance amounts to

$$\begin{aligned}
\sigma_0^{*2}(1.0, 2) &\approx 0.0825\, \sigma_0 \\
\sigma_0^{*2}(1.0, 3) &\approx 0.0797\, \sigma_0 \, .
\end{aligned}$$

An approximation for equation (A.7) can be given for $\sigma_G{>}0.6$ and for sufficiently large m (about $2\,\sigma_G$). Under these circumstances, the sums are approximated by an integration from minus infinity to infinity:

$$\sigma_0^{*2}(\sigma_G, m) \approx \frac{\sigma_0^2}{4\pi^2\sigma_G^4} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{x_1^2+x_2^2}{\sigma_G^2}}\, dx_1\, dx_2 \, .$$

Using relation (A.5), we solve the definite integral and find

$$\sigma_0^{*2}(\sigma_G, m) \approx \frac{\sigma_0^2}{4\pi\sigma_G^2} \qquad\qquad \text{for } \sigma_G \geq 0.6$$

Figure A.1 illustrates the validity of this approximation. The condition $\sigma_G{>}0.6$ was estimated by solving the equation

$$e^{-\frac{0.5^2}{\sigma_G^2}} = 0.5 \, ,$$

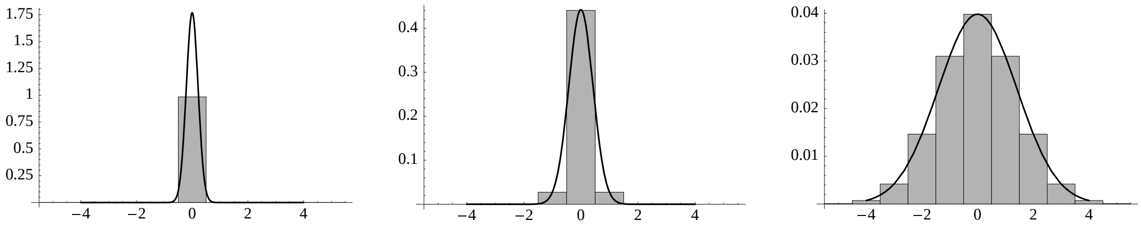
which is the situation where the continuous Gaussian intersects the central rectangle at half its height (figure A.1b). Below this limit, the sampling is too coarse. However, the limit $\sigma_G \to 0$ is the trivial case of filtering with a Dirac peak, which is the identity filter:

$$\lim_{\sigma_G \to 0} \sigma_0^{*2}(\sigma_G, m) = \sigma_0 \, .$$

Another interesting limit is $\sigma_G \to \infty$ for a given support size m:

$$\lim_{\sigma_G \to \infty} \sigma_0^{*2}(\sigma_G, m) = \frac{\sigma_0}{(2m{+}1)^2} \, ,$$

corresponding to a box filter of the same size.



a. $\sigma_G = 0.3$ **b.** $\sigma_G = 0.6$ **c.** $\sigma_G = 2.0$

Figure A.1: Approximation of the sum of Gaussian filter coefficients by the definite integral from minus infinity to infinity.

The covariance between two filtered values after applying a discrete Gaussian filter is given by an expression similar to (A.7), but is omitted here for simplicity. Analogous to the variance, the covariance is well approximated for $\sigma_G > 0.6$ by the equation

$$\sigma^{*\,2}_{xy}(\sigma_G, d) \approx \frac{\sigma_0^2}{4\pi^2\sigma_G^4} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{-(x_1^2+x_2^2)}{2\sigma_G^2}}\, e^{\frac{-((x_1-d_1)^2+(x_2-d_2)^2)}{2\sigma_G^2}}\, dx_1\, dx_2 \,. \tag{A.8}$$

where d is the distance between the two pixels. This definite integral can be simplified to a Gaussian distribution with a standard deviation of $\sqrt{2}\,\sigma_G$, leading to

$$\sigma^{*\,2}_{xy}(\sigma_G, d) \approx \frac{\sigma_0^2}{4\pi\sigma_G^2}\, e^{-\frac{d_1^2+d_2^2}{4\,\sigma_G^2}} \qquad \text{for } \sigma_G \geq 0.6\,.$$

A.4 Variance and covariance after bilinear interpolation

Bilinear interpolation is described by equation (A.9). The interpolated value is computed by a weighted sum of the four neighboring pixels. The weights depend on the fractional part $d = x - \lfloor x \rfloor$ of the coordinate x:

$$f(x) = \begin{bmatrix} 1 - d_1 - d_2 + d_1 d_2 \\ d_2 - d_1 d_2 \\ d_1 - d_1 d_2 \\ d_1 d_2 \end{bmatrix} \bullet \begin{bmatrix} f(\lfloor x \rfloor) \\ f(\lfloor x \rfloor + [0,1]^T) \\ f(\lfloor x \rfloor + [1,0]^T) \\ f(\lfloor x \rfloor + [1,1]^T) \end{bmatrix} . \tag{A.9}$$

In order to derive an expression for the variance reduction, we will assume that all pixel values are independent and normally distributed with an identical variance σ_0^2 as defined by equation (3.43). Exploiting this assumption, simple expressions are found for the actual standard deviation σ_0^* of an interpolated value $f(\lfloor x \rfloor + d)$ and also for the actual covariances between two interpolated pixels. The formula for the variance σ_0^{*2} was already given in equation (3.46) but is repeated here for completeness:

$$\sigma_0^{*2}(d) = \sigma_0^2 \left(1 - 2\,d_1 + 2\,{d_1}^2\right) \left(1 - 2\,d_2 + 2\,{d_2}^2\right) \,.$$

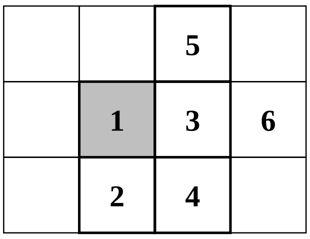


Figure A.2: Six different cases are distinguished for the analysis of the covariances. Case 6 is the trivial case of zero covariance. The expressions for the covariance in each case are summarized in equation (A.10).

In order to compute the covariances, six different cases were distinguished depending on relative position of the two coordinates (figure A.2). One of these cases is the trivial case of zero covariance, which is numbered case 6 in figure A.2. The five remaining expressions for the covariances are summarized in equation (A.10).

$$\sigma_{xy}^{*\,2}(d_1, d_2) = \sigma_0^2 \cdot \begin{cases} (1-d_{21}-d_{11}\,(1-2d_{21}))\;(1-d_{22}-d_{12}\,(1-2d_{22})) & [1] \\ (1-d_{21}-d_{11}\,(1-2d_{21}))\;(1-d_{22})\;d_{12} & [2] \\ (1-d_{22}-d_{12}\,(1-2d_{22}))\;(1-d_{21})\;d_{11} & [3] \\ (1-d_{21})\;(1-d_{22})\;d_{11}\,d_{12} & [4] \\ (1-d_{21})\;(1-d_{12})\;d_{11}\,d_{22} & [5] \\ 0 & [6] \end{cases} \tag{A.10}$$

A.4.1 Bilinear interpolation of Gaussian filtered signal

When the bilinear interpolation is applied to a Gaussian filtered signal, the assumption (3.43) does not hold anymore. However, it is possible to approximate the covariance matrix of the filtered image as we have seen in section A.3. The covariances of two neighboring pixels amount to

$$\bar{\sigma}_{xy}^2\Big|_{d^2=1} \approx \frac{\sigma_0^2}{4\pi\sigma_G^2}\; e^{\frac{-1}{4\,\sigma_G^2}} \qquad \text{and} \qquad \bar{\sigma}_{xy}^2\Big|_{d^2=2} \approx \frac{\sigma_0^2}{4\pi\sigma_G^2}\; e^{\frac{-2}{4\,\sigma_G^2}}$$

depending if the squared distance between the pixels is 1 or 2. Using these covariances, the following expression results for the variance of a filtered and interpolated pixel:

$$\sigma_0^{*2}(\sigma_G, d) = \frac{\sigma_0^2}{4\pi\sigma_G^2}\; e^{-\frac{1}{2\sigma_G^2}} \prod_{i=1}^{2} \left(e^{\frac{1}{4\sigma_G^2}} - 2d_i(e^{\frac{1}{4\sigma_G^2}} - 1) + 2d_i^2(e^{\frac{1}{4\sigma_G^2}} - 1) \right) \tag{A.11}$$

We are now interested in the additional change in pixel variance. The maximum reduction will be at $d=(\frac{1}{2},\frac{1}{2})$ and amounts to

$$\sigma_0^{*2}(\sigma_G,d)\Big|_{d=(\frac{1}{2},\frac{1}{2})} = \frac{\sigma_0^2}{16\pi\sigma_G^2}\left(1+e^{-\frac{1}{2\sigma_G^2}}+2\,e^{-\frac{1}{4\sigma_G^2}}\right) . \qquad \text{(A.12)}$$

Figure A.3 depicts the relative change $(\sigma_0^* - \bar{\sigma}_0)/\bar{\sigma}_0$ at its maximum $d = (\frac{1}{2},\frac{1}{2})$ depending on the filter width σ_G. Choosing the common filter width $\sigma_G = 1$, this maximum additional reduction equals to

$$\frac{\sigma_0^* - \bar{\sigma}_0}{\bar{\sigma}_0}\Bigg|_{\sigma_G=1} \approx 0.209 \, .$$

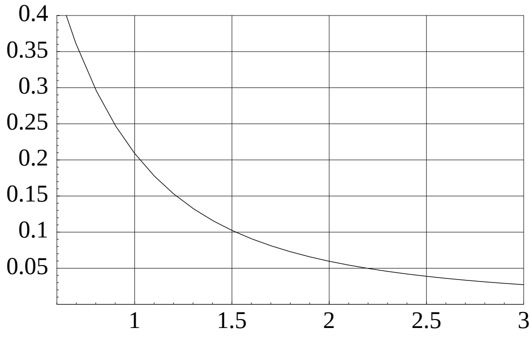


Figure A.3: Maximum additional variance reduction of bilinear interpolation of a Gaussian filtered signal.

A.5 Available software for solving NLS problems

Many software packages for finding the minimum of a sum of squares of nonlinear functions are available. The following section describes a few of these algorithms. A more complete list can be obtained through using the *Guide to Available Mathematical Software* (GAMS) at `http://gams.nist.gov`. The corresponding problem class is `K1b1a`.

A.5.1 DNL2S1 in CMLIB library

```
Minimizes a nonlinear sum of squares using both residual and gradient values
supplied by the user. (Double precision version of NL2S1.)

Classes:   K1b1a2. Unconstrained nonlinear least squares approximation by
                   smooth functions, user provides first derivatives
           L8e1b2. Parameter estimation in nonlinear least squares
                   regression using unweighted data, user provides
```

derivatives
L8e1b4. Parameter estimation in nonlinear least squares regression using weighted data, user provides derivatives

Type: Fortran subroutine in CMLIB library (DNL2SN sublibrary).
Access: Public domain. Portable.
Precision: Double.

Usage: CALL DNL2S1(N,P,X,CALCR,CALCJ,IV,LIV,LV,V,UIPARM,URPARM,UFPARM)

A.5.2 DUNLSJ in IMSLM library

Solve a nonlinear least squares problem using a modified Levenberg-Marquardt algorithm and a user-supplied Jacobian.

Classes: G1b1b. Unconstrained optimization of a smooth multivariate function, user provides first derivatives
K1b1a2. Unconstrained nonlinear least squares approximation by smooth functions, user provides first derivatives

Type: Fortran subroutine in IMSLM library.
Access: Proprietary. Many implementations available.
Precision: Double.

Usage: CALL DUNLSJ (FCN, JAC, M, N, XGUESS, XSCALE, FSCALE, IPARAM, RPARAM, X, FVEC, FJAC, LDFJAC)

A.5.3 E04GBF in NAG library

A comprehensive quasi-Newton algorithm for finding an unconstrained minimum of a sum of squares of M non-linear functions in N variables (m .ge. n). First derivatives are required.

Classes: K1b1a2. Unconstrained nonlinear least squares approximation by smooth functions, user provides first derivatives
L8e1b2. Parameter estimation in nonlinear least squares regression using unweighted data, user provides derivatives

Type: Fortran subroutine in NAG library (E04 sublibrary).
Access: Proprietary. Many implementations available.
Precision: Double.

Usage: CALL E04GBF (M, N, LSQLIN, LSQFUN, LSQMON, IPRINT, MAXCAL, ETA, XTOL, STEPMX, X, FSUMSQ, FVEC, FJAC, LJ, S, V, LV, NITER, NF, IW, LIW, W, LW, IFAIL)

See also: E04YCF

A.5.4 E04GDF in NAG library

A comprehensive modified Gauss-Newton algorithm for finding an unconstrained minimum of a sum of squares of M non-linear functions in N variables (m .ge. n). First derivatives are required. The routine is intended for functions which have continuous first and second derivatives (although it will usually work even if the derivatives have occasional discontinuities).

```
Classes:   K1b1a2. Unconstrained nonlinear least squares approximation by
                   smooth functions, user provides first derivatives
           L8e1b2. Parameter estimation in nonlinear least squares
                   regression using unweighted data, user provides
                   derivatives

Type:      Fortran subroutine in NAG library (E04 sublibrary).
Access:    Proprietary. Many implementations available.
Precision: Double.

Usage:     CALL E04GDF (M, N, LSQMON, IPRINT, MAXCAL, ETA, XTOL, STEPMX, X,
           FSUMSQ, FVEC, FJAC, LJ, S, V, LV, NITER, NF, IW, LIW, W, LW, IFAIL)

See also:  E04YCF
```

A.5.5 E04HEF in NAG library

A comprehensive modified Gauss-Newton algorithm for finding an unconstrained minimum of a sum of squares of M non-linear functions in N variables (m .ge. n). First and second derivatives are required.

```
Classes:   K1b1a3. Unconstrained nonlinear least squares approximation by
                   smooth functions, user provides first and second
                   derivatives

Type:      Fortran subroutine in NAG library (E04 sublibrary).
Access:    Proprietary. Many implementations available.
Precision: Double.

Usage:     CALL E04HEF (M, N, LSQFUN, LSQHES, LSQMON, IPRINT, MAXCAL, ETA,
           XTOL, STEPMS, X, FSUMSQ, FVEC, FJAC, LJ, S, V, LV, NITER, NF,
           IW, LIW, W, LW, IFAIL)

See also:  E04YCF
```

A.5.6 LMDER in MINPACK package

Finds a minimum of the sum of the squares of m nonlinear functions in n variables by a modification of the Levenberg-Marquardt algorithm. The user must provide a subroutine which calculates the functions and the Jacobian.

```
Classes:   K1b1a2. Unconstrained nonlinear least squares approximation by
                   smooth functions, user provides first derivatives
```

Type: Fortran subroutine in MINPACK package.
Access: Public domain. Portable.
Precision: Double.
Note: Single precision version available in NETLIB package SMINPACK.

Usage: CALL LMDER(FCN, M, N, X, FVEC, FJAC, LDFJAC, FTOL, XTOL, GTOL, MAXFEV, DIAG, MODE, FACTOR, NPRINT, INFO, NFEV, NJEV, IPVT, QTF, WA1, WA2, WA3, WA4)

See also: CHKDER checks user-supplied Jacobian