Deformable Multi Template Matching with Application to Portal Images

Martin Berger and Gaudenz Danuser
Swiss Federal Institute of Technology
Communication Technology Lab, Image Science
8092 Zürich, Switzerland
{berger,danuser}@vision.ee.ethz.ch

Abstract

The exact positioning of patients during radiotherapy is essential for high precision treatment. The registration of portal image sequences can help to control the patient position. The particular problem of such megavoltage X-ray imagery is its extremely low contrast, rendering accurate feature extraction a difficult task. To circumvent the step of feature extraction, the algorithm presented in this paper relies on an area-based matching of the image signal using deformable templates. This strategy contrasts with most state of the art registration algorithms for portal imagery.

The paper includes the mathematical formalism of the least squares template matching method, as well as the framework for automated quality control, together yielding a fast, robust and very accurate image matching procedure. Tests on 17 portal image series with more than 100 images in total have shown very satisfying results. Artificially rotated and shifted images demonstrate the performance of the method with respect to a ground truth.

1. Introduction

The exact positioning of patients during radiotherapy is essential for high precision treatment. Hence, it is a major goal to automatically measure patient setup deviation between or even during treatment sessions. One possibility is to use a portal imaging device which delivers images of the exit dose distribution during treatment. Unfortunately, the contrast of these megavoltage X-ray images is very low due to the high energy beam, and, since we are dealing with projected images, there is the possibility of out-of-plane rotations of the patient resulting in complex distortions. Thus, the quantitative processing of portal images is a very challenging task.

So far, several methods have been implemented, all of which lacking either the robustness or the accuracy necessary to be reliably used in daily hospital routine. Furthermore, most state of the art techniques require too much user interaction. Algorithms like point-to-point (or landmark) matching greatly depend on the exact localization of the landmarks by the physician. This is not only time-consuming, but also varies for different operators. Less user interaction is required by the chamfer matching algorithm where significant ridges are manually outlined in a reference image and matched onto the detected features of the treatment image [?]. A similar approach in the sense that it also uses binary features, namely cores, is described in [?], where also a quite complete review about published algorithms for portal imaging can be found.

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 match is superior to a feature-based algorithm. Greyvalue correlation techniques are described in [?], [?], [?]. Their limitations lie in the restriction to a translation or in a coarse search grid for computational reasons.

A distinct approach has been proposed in [?]. Based on 3D CT or MR data, Brunie et al. match 2D projections of a segmented feature with its corresponding feature in the portal image. The strength of this technique is the rigorous fusion of 3D and 2D data. Results have been shown on a head phantom. However, we believe that the abovementioned drawbacks of feature-based matching will reduce the accuracy and robustness in practice.

None of the mentioned papers show an error propagation technique for the measured displacement. This would be crucial for any kind of automated quality control. In the current clinical practice, the result must be checked visually by a physician. The method of least squares matching (LSM) with deformable templates meets the requirements of being an area-based approach with rigorous error propagation, self-diagnosis and thus, minimized user interaction. It was first introduced by Grün for precise measurements in aerial and close range photogrammetry [?]. Basically the same algorithm is described in [?] for motion estimation, which was further developed using a multi-scale approach...
Following and extending the original work of Grün, highly accurate results at the resolution limit of a light microscope were achieved in [?]. The high accuracy of this technique even in the case of low-contrast imagery is extensively exploited in [?]. The paper reports high accuracy positional measurements of a calibration grid used to calibrate a stereo light microscope.

This paper presents the application of LSM to the low-contrast and noisy portal images. Additional problems compared to the light microscope images in [?, ?] arise from the higher complexity of the image scene and the out-of-plane rotations. On the other hand, the requirements on accuracy are not as high in portal imaging as in their application.

A similar technique for the registration of medical image series is reported by Unser et al. [?], where each image is matched to the reference image based on a global greyvalue difference measure. 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 greyvalue errors such as bias fields is reduced.

2. Least Squares Matching

LSM is an area-based matching algorithm, thus, not depending on extraction of binary image features. A transformation defines the geometric relation between the original template and the matched area (called patch) in the search image. Unlike in most correlation methods, the optimum transformation is not searched by testing all possible cases, but approached using an optimization scheme. Assuming that a fair initial guess can be supplied, this is not only faster but also more accurate.

The backbone of the LSM is the complete formalism that leads from a series of observation equations

\[ f(\xi) + e(\xi) = \tilde{g}(\xi_p) \]  

(1)

of a template image \( f(\cdot) \) and a radiometrically adjusted search image \( \tilde{g}(\cdot) \) to a linearized least squares problem. \( e(\cdot) \) denotes the radiometric error. The image coordinates \( \xi \) are transformed to \( \xi_p = \psi(u, \xi) \) to allow for displacement, rotation and deformation of the template using a parameter vector \( u \). For an affine transformation, the parameter vector consists of six variables \( u = [t_1, t_2, m_1, s_1, s_2, m_2]^T \) and the coordinate transformation can be written as

\[ \xi_p = \begin{bmatrix} t_1 \\ t_2 \\ m_1 \\ s_1 \\ s_2 \\ m_2 \end{bmatrix} \xi. \]  

(2)

For portal imaging, an affine transformation is suitable since it simultaneously handles patient displacement and in-plane rotation as well as it compensates a significant part of the geometric distortions originating from small out-of-plane rotations.

Equation (2) represents a relation between each grey-value within the template and its corresponding image intensity in the search image. Notice that the template grey-values \( f(\xi) \) are defined on the grid of the reference image while \( \tilde{g}(\xi_p) \) fall between the grid of the search image. Interpolating the greyvalues \( \tilde{g}(\psi(u, \xi)) \) we substitute \( \tilde{g}(u, \xi) := \tilde{g}(\psi(u, \xi)) \). Equation (2) is now reordered into a vector notation

\[ f + e = \tilde{g}, \]  

(3)

building a series of \( N \) equations, where \( N \) is the number of pixels included in the template. Following [?], the goal of the matching is to minimize the squared error \( e^T P e \) with an optional weighting matrix \( P \). The strategy for optimization follows the Gauss-Newton scheme, linearizing the observation equation around the current estimate \( u^0 \) or \( \xi^0_p = \psi(u^0, \xi) \), respectively:

\[ f(\xi) + e(\xi) = \tilde{g}(\xi^0_p) + \nabla_u \tilde{g}(\xi^0_p) \Delta u \]  

(4)

\[ f + e = \tilde{g}^0 + A \cdot \Delta u, \]  

(5)

where \( A \) is the Jacobian matrix with respect to the parameter vector \( u \). This linear problem can be solved analytically setting the first derivative of the squared linear error to 0 which yields the normal equation system

\[ A^T P A \cdot \Delta u = -A^T P \tilde{g}^0 - f. \]  

(6)

As long as \( A \) has no row deficiency, \( A^T P A \) is always positive definite and symmetric and hence the Cholesky decomposition can be applied to solve equation (6). After each iteration step, matrix \( A \) must be recomputed using the updated set of parameters \( u^{l+1} = u^l + \Delta u \). When the parameter change \( \Delta u \) is below a specified numerical resolution the iteration process is stopped.

The least squares formalism allows one to introduce additional constraints in a simple and intuitive way. Each constraint \( l \) can be formulated as a zero observation (pseudo observation)

\[ 0 + e_l = c_l(u). \]  

(7)

\( c_l(u) \) represents a function of the parameter vector \( u \). In particular, constraints can be used to suppress the estimation of a parameter \( u_i \) while fixing it to an a priori value. In this case, a constraint function \( c_l(\cdot) \) is formulated as

\[ c_l(u) : \ u_i - \bar{u}_i = 0, \]  

(8)

where \( \bar{u}_i \) is the a priori value. By introducing a large weight in \( P \) for the pseudo observation \( i \), the constraint \( c_l(\cdot) \) will hold through the optimization. This technique is used to stabilize numerically weak normal equations in case of low-contrast imagery (cf. section ??). The inclusion of the constraints in the normal equation system (6) strictly follows
the framework shown for regular observation equations of type (2).

3. Material and methods

The steps before high precision conformal therapy include a CT, then a 3D planning of beam directions, field shape and dose distribution and finally the positioning of the patient using a simulator with the same geometry as the linear accelerator. To generate reference images for matching, it would be best to use the CT data calculating digitally reconstructed portal radiographs [?, ?]. Future implementations will rely on such data. However, in this work the tests were carried out using the initial or best portal image from a series as reference image. The diagnostic X-ray images from the simulator are not suitable for automated area-based matching. Although they provide high contrast imagery from the treated area, their greyvalue characteristics strongly differs from portal images.

The portal images were acquired at the University Hospital of Zürich using a Varian accelerator and their electronic portal imaging device (EPID). This device delivers a distortion free image with a resolution of 256×256 pixel on an area of 32×32 cm² [?].

3.1. Selecting suitable templates

A fully automated template selection would require a sophisticated expert system combining medical and computer vision knowledge. However, state of the art image processing algorithms are not able to discriminate stable structures from artifacts or other distinct but unstable features (e.g., originating from air in the rectum). Thus, the strategy is to implement an operator guided template selection. At the moment, the physician has to define regions containing significant structures that are known to be stable over a series of images. In contrast to the manual definition of landmarks or edges, which is necessary for feature-based algorithms, this is quickly accomplished without introducing the same variability between different operators. In order to further reduce the workload, standard configurations can be stored in a template database. The resulting regions are considered templates for the subsequent matching process.

Future versions will support this definition by an online analysis of the contrast and its distribution within the selected templates. From equations (1), (2) and (3) it is obvious that the normal equation matrix $A^TPA$ contains the structural information of the patch. This is used in section 3 where the inversion of $A^TPA$ together with the a posteriori greyvalue error yields an estimate of the covariances of the transformation parameters. During the template initialization, $A^TPA$ can be approximately evaluated based on the template information instead of the patch information. A method to find a priori instead of a posteriori greyvalue errors is described in [?]. The resulting a priori standard deviations can then be employed to control the selection and definition of the templates: After an approximate initialization based on medically relevant features, their exact position, shape and size are determined by maximizing the a priori localization while minimizing the number of pixels. Therefore, insignificant parts will be excluded while rather significant structures are included in a semi-automated way. The aim of this procedure is to optimally combine the expert knowledge of the physician with the quantitative image analysis.

3.2. Field edge displacement

In general, the EPID is not in a fixed position and therefore the image coordinate systems can differ between images. In a first step, the field edges are aligned with LSM using a small region along the edge as template. For the extraction of the field edge, we use a simple histogram analysis. Since the area flanking the extracted edge is used for the alignment, the exact position of this binary feature is not very crucial. The matching algorithm correctly determines the displacement even if the extracted binary edge is not exactly at the 50% dose level. The verification of the field edge shape is accomplished using the diagnostic tools outlined below.

3.3. Anatomy alignment

The current implementation of the multi template LSM is a straightforward extension to the algorithm described in section 3. The observations now include several templates and their corresponding patches $f^k(\xi) + e^k(\xi) = g^k(\xi^k_P)$, where we introduce one single transformation for all patches $\xi^k_P \equiv \xi_P$ with the parameter set shown in equation (1). The initial guess is derived from the transformation of the field edge match. The observations can be reordered into a vector notation analogous to the single template matching (2). Formally, this procedure is equal to the definition of one large template with several scattered regions of interest.

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.

3.4. Estimation of parameter accuracy

Parameter estimation in linear least squares problems are extensively discussed in standard literature on parameter es-
timation theory, e.g., [??]. The iterative solution of equation (??) is an unbiased estimate for the unknowns with a stochastic variance expressed by the diagonal elements of the covariance matrix $S_{uu} = \hat{\sigma} \cdot Q_{uu}$. The value $\hat{\sigma}$ denotes the a posteriori noise estimate $\hat{\sigma} = (e^T P e)/N$ and $Q_{uu}$ is the inverse of the normal equation matrix

$$Q_{uu} = (A^T P A)^{-1},$$

also called cofactor matrix. For an in-depth analysis and proof of these properties the reader may refer to [?].

3.5. Matching diagnosis

Various tests are applied to supervise the matching, yielding self-diagnosis of the framework. These tests can be divided into two categories. The first category includes tests carried out during the iteration, the second those applied after completed optimization.

**Tests during optimization.** The most important test is based on a determinability measure of the parameters. As in [?], the influence of each shape parameter upon the others is estimated by computing the contribution $\delta_i$ to the trace of the cofactor matrix $Q_{uu}$

$$\delta_i = \text{tr}[Q_{uu'}] - \text{tr}[Q_{uu}],$$

(10)

where $Q_{uu'}$ is the cofactor matrix with parameter $i$ excluded. An efficient implementation of (??) is achieved by using the Kalman-Bucy filter technique (cf. [?]) elegantly propagating $Q_{uu'}$ from $Q_{uu}$. Applying this framework to (??) yields an expression for the contribution $\delta_i = \sum_{k \neq i} q_{ik} q_{ik}/q_{ii}$, where $q_{ik}$ denotes the elements of the full cofactor matrix $Q_{uu}$.

If a contribution $\delta_i$ of the shape parameter $i$ is high, this parameter strongly correlates to one or more parameters. One should either exclude this parameter $i$ by applying the formalism for parameter constraints described in section ?? or restrict the transformation to similarity.

Since the contribution value can significantly change during the iteration, it should be reevaluated at the end of the optimization. A suitable strategy is to estimate a similarity transformation first and then to test the full affine parameter set for determinability. If none of them shows intolerably large contributions $\delta$, the optimization is continued with the full parameter set (cf. figure ??).

**Tests after optimization.** After completing the optimization, two tests check for the goodness of fit. The cross-correlation between the template $f(.)$ and the patch $\hat{g}(.)$ — which is interpolated from the search image using the final parameter set $u$ — should be very close to 1.0. If the correlation coefficient is not above a certain threshold, the analysis of the final error vector $e$ might point at an incorrectly aligned template. Large values of $e$ indicate a mismatch for this region. Excluding such a template and reoptimizing the transformation may result in a better parameter set.

4. Results

The test data consists of 17 image series with totally 106 images. About 30 artificially transformed images are also generated by taking existing portal images and resampling them with an additional translation and/or rotation using bilinear interpolation.

**Field edge displacement.** The field edge is a very distinct image feature and therefore easy to match. As expected, field edge alignment is very robust and LSM finds even larger displacements and rotations as should occur in radiotherapy treatment.

Figure ?? shows the iterative computation of the optimum position and orientation starting from an initial guess, in this case an identity transformation. The search image is artificially rotated by -15° (clockwise) and the algorithm estimates a rotation of -14.95°. Together with the rotation between the original search image and the reference image of 0.05°, perfect performance with respect to a ground truth is demonstrated (cf. table ??).

**Effective patient displacement.** Due to much weaker contrast, the anatomy alignment is not as robust as the field edge match. Applied to artificially rotated images, the algorithm starts to fail when more than about 15 mm shift (17 pixel) or 10° rotation are introduced. However, in the practical test series patient displacements go only up to
Figure 2. Iteration series of an artificially rotated field edge of an AP pelvis image. The search image is rotated by -15 degrees. After 11 iterations a perfect match to the template is found.

<table>
<thead>
<tr>
<th>dx [mm]</th>
<th>dy [mm]</th>
<th>rot [deg]</th>
</tr>
</thead>
<tbody>
<tr>
<td>original</td>
<td>-0.08</td>
<td>-6.88</td>
</tr>
<tr>
<td>rotated (ground truth)</td>
<td>-1.85</td>
<td>-6.63</td>
</tr>
<tr>
<td>rotated (estimated)</td>
<td>-1.86</td>
<td>-6.63</td>
</tr>
</tbody>
</table>

Table 1. Consistency test of field edge displacement with respect to a ground truth.

12 mm translation and 4.5° rotation, which is reasonable as an upper limit for alignment errors in daily hospital routine.

The matching results are validated by visual inspection using combined cursors and verification lines. For all but one AP pelvis image the algorithm finds the correct displacement, i.e., there are no visually noticeable errors. In two cases, the correct displacement is achieved after the manual exclusion or redefinition of a template. All others yield robust matches with a set of unoptimized templates. An example for a successful match is shown in figures ?? and ??.

Lateral pelvis fields are inherently more difficult to match. Still, from 25 images, only five pose problems in the sense that the templates had to be manually adjusted once. Two images needed more than one attempt. Figures ?? and ?? illustrate the potential of LSM. Even on this low-contrast and blurred imagery, a stable transformation is found.

In any case where an incorrect displacement is estimated, a low correlation value reliably indicates the mismatch.

5. Discussion and outlook

This work presents a novel application of least squares matching to portal images. The major difference between LSM and existing algorithms for portal imaging is the area-based matching approach in contrast to feature-based methods. The inherently low-contrast and noisy portal images pose problems on robust feature extraction, which is circumvented in this algorithm. Apart from avoiding the extraction of binary features, the area-based matching has the
further advantage of being more accurate. User interaction is also minimized. Neither landmarks nor contours must be drawn by the physician, but only regions of significant and stable structures. This is quickly accomplished by drawing polygons on the reference image with a few mouse clicks. In the current implementation, only little expert knowledge is required to define templates for a robust match. This paper also outlines various methods supporting the operator in this definition process.

The achieved results are very promising. Above 90% of all test images are robustly matched, i.e., no adjustments on templates are necessary. After one-time manual optimization of the template region, only two very bad images are still misaligned. It is important to notice that a misalignment always is automatically detected during the diagnostic phase due to a low cross-correlation value (cf. section ??).

Future investigations will focus firstly on the problem of template selection as outlined in section ?? and secondly on the multi-template matching itself. The limitation of our current implementation is that only one global geometric transformation is estimated based on an ensemble of local template–patch relations. This approach relies on the assumption that the patient’s body undergoes a rigid in-plane motion and an out-of-plane rotation, which can be compensated with one single affine transformation. Furthermore, perspective and higher order imaging distortions are neglected.

Although these simplifications are suitable for many cases, they can lead to systematic errors in the displacement estimate. The next version of our system will therefore contain not only one global transformation, but also local transformations which are able to deal with the abovementioned local effects. Between each template and its corresponding patch, a local affinity is estimated whose translative part is connected to the global affinity. This estimation problem can be solved with constrained least squares optimization. From this model refinement we expect an increased accuracy of the matching and, in particular, even better robustness which is the fundamental requirement for the application to the clinical practice.

Another important step needed for the daily use of the method in radiotherapy is the interface between the algorithm’s kernel and the operator. An ideal user interface consists of a comprehensible report about the success of the matching process. It also should provide the physician with an efficient tool that becomes active when additional information is needed for a successful match.

Acknowledgments. The authors would like to acknowledge Prof. Dr. G. Geric and Dr. G. Székely for the support of this project. We further appreciate the contribution of Dr. U. R. Meier and Prof. Dr. U. M. Lütholf providing the portal images and useful information concerning medical topics. This work is granted by the Swiss Cancer League and the Walter Honegger Foundation.

References