Motion Measurements in Low-Contrast X-ray Imagery

Martin Berger and Guido Gerig

Swiss Federal Institute of Technology
Communication Technology Lab, Image Science
8092 Zürich, Switzerland

{berger,gerig}@vision.ee.ethz.ch

Abstract. Measuring motion in medical imagery becomes more and more important, in particular for object tracking, image registration, and local displacement measurements. Such measurements are especially difficult in megavoltage X-ray images (*portal images*), which are used to control the position of patients in high precision radiotherapy. Low contrast, blur, and noise render accurate measurements difficult.

In this work we review the framework of a generic matching algorithm only based on the image signal and not on binary image features. Thus, the often unreliable step of feature extraction in such imagery is circumvented. Another major advantage is the possibility of self-diagnosis, which is used for restricting the transformation in motion measurements if the image quality is not sufficient.

The method of digitally reconstructed radiographs (DRR) allow for the computation of error free reference images, avoiding the additional step of therapy simulation. The multi-modal match between such DRRs and portal images lead to an estimate of the patient position during radiotherapy treatment. Results of generated data with known ground truth as well as results of a multi-modal match are presented.

1 Introduction

Accurate motion measurements in images are essential to solve numerous problems in computer vision. For medical imagery in particular, precise position measurements and registration of image series represent two important applications. Wherever feature extraction is difficult or high precision is required, the least squares template matching algorithm (LSM) reviewed in this paper has many advantages over other methods. LSM is a generic matching algorithm suitable for many applications including motion estimation in low-contrast megavoltage X-ray images, also called *portal images*.

The specific goal in this work is the exact positioning of patients during radiotherapy, which is essential for high precision treatment. This involves automatically measuring patient setup deviation between or even during treatment sessions. One possible sensor is an electronic portal imaging device (EPID),

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, parts of a rigid 3D motion must be estimated by evaluating projected 2D images.

Because portal images are inherently noisy and low in contrast, it is difficult to robustly extract features like edges, ridges or cores. This is the weakness of feature-based methods applied to portal images like chamfer matching [Gilhuijs and van Herk 1993] and core-based image registration [Fritsch *et al.* 1995]. Previous area-based methods include the greyvalue correlation techniques described in [Dong and Boyer 1996, Moseley and Munro 1994]. Their limitations lie in the restriction to a translation or in a coarse search grid for computational reasons.

In [Berger and Danuser 1997] we proposed the area-based LSM algorithm to find displacements between two portal images. The method of LSM with deformable templates meets the requirements of being an area-based approach with the possibility of self-diagnosis. Early work in this field was presented by [Lucas and Kanade 1981], who published an iterative image registration scheme base 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]. Following and extending the work of Grün, [Danuser and Mazza 1996] achieved highly accurate results at the resolution limit of a light microscope. Compared to these previous applications of LSM, additional problems arise in portal images from the higher complexity of the image scene and the out-of-plane rotations.

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 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 intensity inhomogeneity is reduced.

Extending the approach in [Berger and Danuser 1997], we further exploit the self-diagnosis capabilities. The variation of image quality of portal images inhibits the estimation of a 2D affine transformation for all cases. An adaptive scheme based on self-diagnostic measures allows for an automatic reduction of the parameter set where the full parameter set is not determinable. Furthermore we include the multi-modal matching of portal images against digitally reconstructed radiographs (DRR) computed from the CT volume data.

2 Least squares template matching

LSM is an area-based matching algorithm, thus does not depend on the extraction of binary image features. This is a very 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.

The following sections give a short review over the LSM framework. Further information can be found in [Berger 1998, Berger and Danuser 1997].

2.1 Unconstrained LSM

The LSM includes two observations, the template image $f[.]$ and the search image $g[.]$, called patch. 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 addition to the geometric transformation, the observations must be adjusted radiometrically. The simultaneous estimation of both types of transformations would lead to an overdetermined system, since it is not possible to distinguish them locally. 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. We use a linear transformation which can be written as $\bar{g}[u] = \alpha + \beta\, g[u]$, where u stands for the discrete image coordinates.

The general geometric transformation is denoted by $x = \psi(\xi, u)$, transforming the image coordinates u using the parameter vector ξ. Applying the least squares framework, this leads to the observation equation

$$f[u] + e[u] = \bar{g}(x) \,. \tag{1}$$

Equation (1) represents a relation between each greyvalue within the template and its corresponding image intensity in the search image.[1] Interpolating the greyvalues for a given ξ we substitute $\tilde{g}_\xi[u] := \bar{g}(\psi(\xi, u))$. Based on a coordinate list $u[k]$, equation (1) is reordered into a vector notation

$$f + e = \tilde{g} \,, \tag{2}$$

building a series of n equations, where n is the number of pixels included in the template and $k = 1 \ldots n$. Together with the least squares objective function $e^T P e$, with P as optional weight matrix, this defines an unconstrained nonlinear least squares (NLS) problem. This nonlinear problem is iteratively solved using a Newton-Raphson scheme. A new estimate $\hat{\xi} = \xi^\circ + \Delta\xi$ is computed linearizing the observation equations (1) around the current estimate ξ°

$$f + e = \bar{g}^\circ + A \cdot \Delta\xi \,. \tag{3}$$

Matrix A is the $n \times r$ Jacobian matrix $\nabla_\xi\, \bar{g}(x^\circ)$ with respect to the parameter vector ξ and r denotes the number of parameters. The linear problem (3) is

---

[1] 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.

solved analytically setting the first derivative of the least squares goal function $e^T P e$ to zero, which yields the normal equation system

$$A^T P A \cdot \Delta \xi = -A^T P \left(\tilde{g}^\circ - f \right) . \tag{4}$$

This linear equation system is then solved using Cholesky decomposition. 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.

2.2   Error propagation and parameter determinability

Parameter estimation in linear least squares problems are extensively discussed in standard literature on parameter estimation theory, for instance in [Koch 1988]. The iterative solution of equation (1) is an unbiased estimate for the unknowns with a stochastic variance expressed by the diagonal elements of the covariance matrix $\Sigma_{\xi\xi} = \hat{\sigma}_0 \cdot Q_{\xi\xi}$. The value $\hat{\sigma}_0$ denotes the a posteriori noise estimate and $Q_{\xi\xi} = (A^T P A)^{-1}$ is the cofactor matrix.

The determinability of a parameter i is tested using its relative contribution δ_i to the trace of the cofactor matrix $Q_{\xi\xi}$

$$\delta_i = \frac{\left| \text{tr}[Q_{\xi\xi}] - \text{tr}[Q_{\xi\xi}^i] \right|}{\text{tr}[Q_{\xi\xi}]} = \frac{\sum_j q_{ij}^2}{q_{ii} \sum_j q_{jj}^2} , \tag{5}$$

where $Q_{\xi\xi}^i$ is the cofactor matrix with parameter i excluded and q_{ij} denote the elements of the full cofactor matrix $Q_{\xi\xi}$. The efficient implementation on the right hand side is achieved employing the Kalman-Bucy filter technique (cf. [Koch 1988]), computing the partial cofactor matrix $Q_{\xi\xi}^i$ directly from $Q_{\xi\xi}$.

If a contribution δ_i of the parameter i is high, this parameter strongly correlates to one or more parameters. One should either exclude parameter i or combine it with the correlating parameters by applying parameter constraints as described in section 2.5.

2.3   Multi template extension

Using multiples template instead of one large template allows significant and stable regions to be selected without including regions unsuitable for matching. There are several ways to extend the standard LSM to multiple templates. For all of them, the equation (1) has to be adapted to include multiple templates and their corresponding patches.

The most straightforward extension is to keep one single transformation for all patches with the same parameter set shown in equation (7), leading to

$$f[u^K] + e[u^K] = \bar{g}^K(x^K) , \tag{6}$$

where $\bar{g}[.] = \alpha^K + \beta^K g[.]$ and $x^K = \psi(\xi, u^K)$. Formally, this procedure is similar 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 greyvalue differences like bias fields.

2.4   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 $\boldsymbol{\xi} = [t_1, t_2, m_1, s_1, s_2, m_2]^T$ and the coordinate transformation is written as

$$x = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} + \begin{bmatrix} m_1 & s_1 \\ s_2 & m_2 \end{bmatrix} \boldsymbol{u} \ . \tag{7}$$

The derivative $\nabla_{\boldsymbol{\xi}} \, \bar{g}(\boldsymbol{x})$ is then calculated explicitly using the chain rule. In vector notation, this leads to the $n \times 6$ Jacobian matrix $\boldsymbol{A}$ (cf. equation (3)), each line $\boldsymbol{A_k}$ representing the derivatives at $\boldsymbol{x_k} = \boldsymbol{\psi}(\boldsymbol{\xi}, \boldsymbol{u}[k])$.

2.5   Employing constraints

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 using large weights in the weight matrix $\boldsymbol{P}$. In this section, we 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 (7) as transformation equation and add the following constraints to the parameter vector $\boldsymbol{\xi}$:

$$\begin{aligned} m_1 - m_2 + e_m &= 0 & \boldsymbol{A_m} &= [0, 0, 1, 0, 0, -1] \\ s_1 + s_2 + e_s &= 0 & \boldsymbol{A_s} &= [0, 0, 0, 1, 1, \ \ 0] \ . \end{aligned}$$

Analogous to the observation equations (1), the constraints are linearized around the current estimates m_i° and s_i°. Thus, the matrix $\boldsymbol{A}$ is augmented by the constraint vectors $\boldsymbol{A_m}$ and $\boldsymbol{A_s}$.

Employing constraints instead of reparametrization is easier to implement and more versatile. On the one hand, constraints can easily be changed during the iteration, quickly switching from a similarity to an affine transformation. On the other hand, only a few types of constraints can actually be expressed by reparametrization, which makes constraints more flexible to use.

3   Controlling patient position in radiotherapy

The steps before high precision conformal therapy include the acquisition of 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 (figure 1). During radiotherapy treatment, either portal films or electronic portal images are acquired for quality control.

The portal images in this work were acquired at the University Hospital of Zürich using a Varian accelerator and their electronic portal imaging device

a. DRR   b. diagnostic X-ray   c. megavoltage X-ray
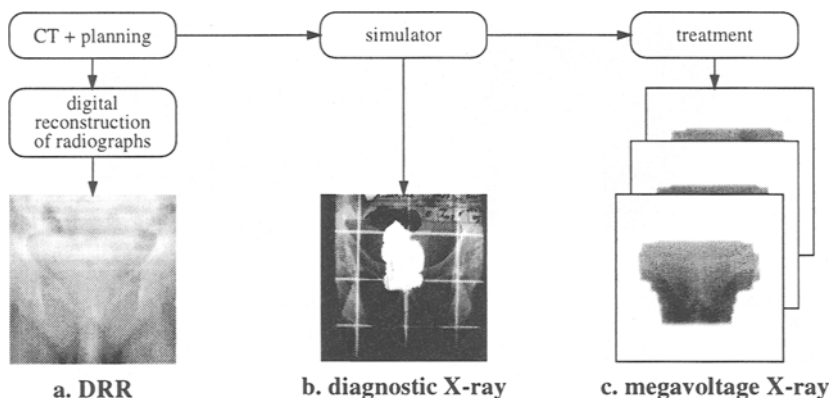
Fig. 1. Overview of the different steps of radiotherapy treatment.

(EPID). This device delivers a distortion free image with a resolution of 256×256 pixel on an area of $32 \times 32 \, \mathrm{cm}^2$ [van Herk and Meertens 1988].

The diagnostic X-ray images from the simulator are often used as a reference image for measuring the patient motion in subsequent portal images. Besides the fact that simulator images are not suitable for automated area-based matching due to their different greyvalue characteristics, they also represent an additional source of error with respect to the original planning data. Hence, one goal of this project is to eliminate the need for simulator X-ray images for motion measurements by direct comparison with a megavoltage DRR (figure 1a).

DRRs are computed simulating the therapy setup by a ray-tracing based algorithm, also correcting for the different absorption coefficients at different beam energies. Thus, a DRR representing the correct patient position serves as reference image. Subsequent portal images are compared to this reference image and the estimate of a 2D affine transformation leads to a correction of the patient position. However, in order to compute DRRs with sufficient quality, the CT slice thickness should be no larger than $5 \, \mathrm{mm}$.

3.1 Selecting suitable templates

In the particular problem of portal images, two displacements must be computed. Since the EPID is in general not in a fixed position, a common coordinate system must be established using the edges of the radiation field. These edges are very distinct features and pose no problems to the matching algorithm. We refer to [Berger and Danuser 1997] for a more detailed description of the fieldedge match.

In the following, we will concentrate on the anatomy match. The selection of the template regions follows the previous work. 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 physician has to position predefined standard templates onto the significant structures in the reference image (figure 2).
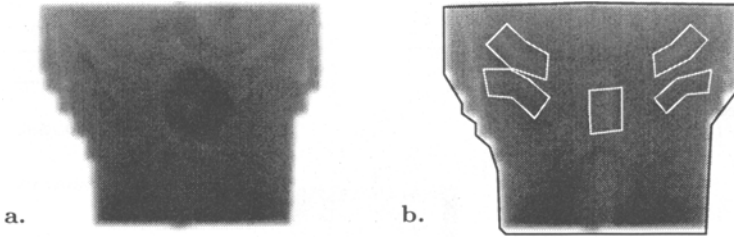
Fig. 2. Image (a) depicts a portal image with a common unstable features that originates from air in the rectum (dark blur in the center of the image). A typical template selection for a AP pelvis field is shown in (b).
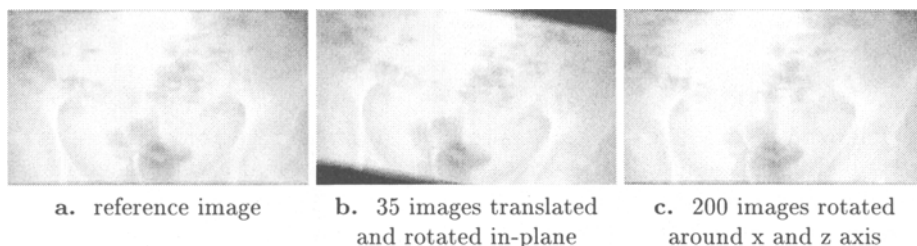
3.2 Self-diagnosis within LSM

The statistics behind the self-diagnosis outlined in section 2.2 is only valid in the adjusted state. Therefore, this measure can not be applied directly to the initial system. However, an upper bound for the determinability is computed matching the templates onto themselves. Based on this upper bound, a coarse result is computed using a restricted parameter set which still approximates the final parameter set sufficiently, usually a congruent transformation.

At this first estimate, the full affine parameter set is tested for determinability. If none of the parameters show large contributions δ_i, the optimization is continued with the full parameter set. A general flowchart is depicted in figure 3.
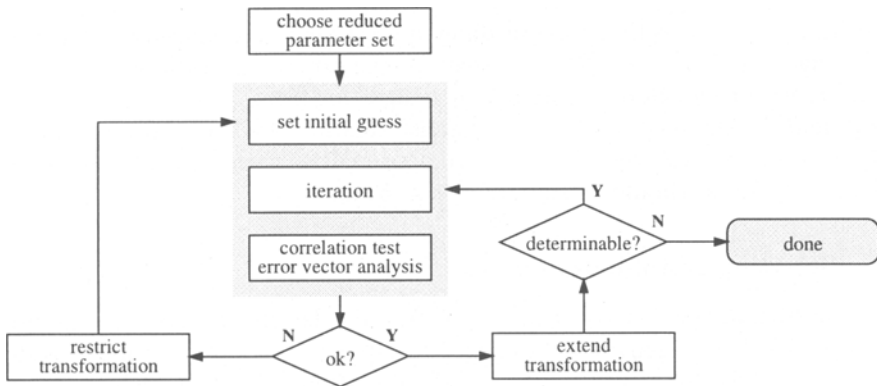
Fig. 3. Flowchart of the diagnostic measures.

The cross-correlation between the template $f[.]$ and the patch $\tilde{g}[.]$—which is interpolated from the search image using the final parameter set ξ—should be very close to 1.0. For multiple templates, the correlation value is computed for each template. Since a low correlation value indicates a mismatch for this

a. reference image   b. 35 images translated   c. 200 images rotated
and rotated in-plane   around x and z axis

Fig. 4. Example DRR images from the in-plane (b) and out-of-plane (c) test series computed from CT volume.

region, excluding such a template and reoptimizing the geometric transformation usually results in a better parameter set.

4 Results

The high precision capabilities of LSM are extensively shown in [Berger and Danuser 1997, Danuser and Mazza 1996]. In the following sections, the emphasis is on the estimation of the in-plane part of the 3D patient motion based on the evaluation of projected 2D images. We first apply the algorithm to generated datasets to test the potential of LSM. In the last section, we present results of the multi-modal match.

In all examples, an affine transformation was used for matching and the initial position was set to the fieldedge position. The typical run time including Gauss filtering of the search image, matching the fieldedge and matching the anatomy templates is about 4 seconds on a Sun Ultra 1 (167 Mhz UltraSPARC CPU).

4.1 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 y direction and a maximum rotation of $10°$ (figure 4b). The standard deviations of the translation measurements were 0.25 pixel (0.23 mm) in x direction and 0.37 pixel (0.33 mm) in y direction. These systematic errors are caused by the unknown y position in the CT coordinate system of the template features. Within the rotation measurement, these systematic error do not occur and the standard deviations are below $0.01°$.

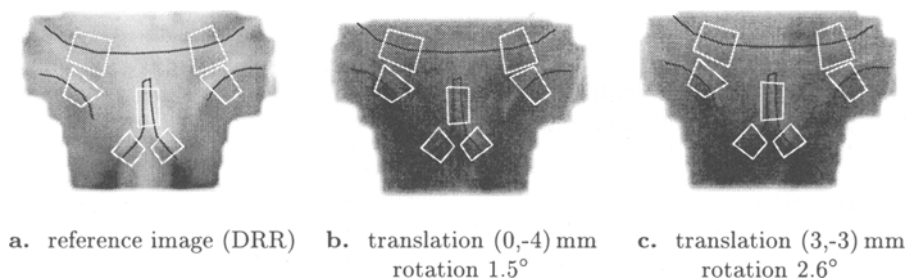## 4.2 Including out-of-plane rotation

In order to test under more realistic conditions, a test series of 200 images with small out-of-plane rotations is generated (figure 4c). The LSM still found 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, were 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.

4.3 Multi-modal match with portal images

Figure 5a shows a DRR and the chosen 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 5b and c. The results were visually validated and all but one were accepted to be correct, which indicates a success rate of well over 90 %.

a. reference image (DRR)   b. translation $(0,-4)\,\mathrm{mm}$ c. translation $(3,-3)\,\mathrm{mm}$
rotation $1.5°$ rotation $2.6°$

Fig. 5. 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 Conclusion and Outlook

The LSM method with deformable templates is a versatile matching algorithm. Since it is an area-based method, the often unreliable step of feature extraction is circumvented. Especially in low-contrast imagery like megavoltage X-ray images (portal images), this is an important feature.

In earlier work, LSM has been successfully applied to matching megavoltage X-ray images of the same modality. The results presented in this paper show the suitability for a multi-modal match between digitally reconstructed radiographs as reference images and the portal images acquired during radiotherapy. The area-based method LSM proved robustness even when matching slightly distorted patterns.

In two test series with generated data, the systematic error caused by estimating a 3D motion from projected 2D images is examined. Furthermore, tests with a DRR matched to real portal images show promising results which will be validated more thoroughly in the near future. Another goal is to combine multiple 2D measurements from different directions to better estimate the 3D motion and to reduce the influence of the systematic error occurring in the 2D measurements.

The presented methods are a step further in the direction of automatically controlling the patient position in radiotherapy. Applied in daily hospital routine, this should lead to an improved quality assurance in radiotherapy.

References

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

[Berger 1998] Martin Berger. The framework of least squares template matching. Technical Report 180, Image Science Lab, ETH Zürich, 1998. Available at http://www.vision.ee.ethz.ch/.

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

[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, 1996.

[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. J. Rad. Onc. Biol. Phys.*, 33(5):1287–1300, 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.

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

[Koch 1988] K. R. Koch. *Parameter estimation and hypothesis testing in linear models.* Springer, 1988.

[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.

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

[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*, Sep./Oct. 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.