Evaluation of an Automatic System for Simulator/Portal Image Matching

F.M. Vos[1], J. Stoeckel[1,3], P.H. Vos[2], A.M. Vossepoel[1]

[1] Pattern Recognition Group, Delft University of Technology
Lorentzweg 1, 2628 CJ Delft, The Netherlands
{frans,jonathan,albert}@ph.tn.tudelft.nl

[2] Bernard Verbeeten Institution
Brugstraat 10, 5000 LA Tilburg, The Netherlands
vos.ph@bvi.nl

[3] EPIDAURE Project, INRIA
2004 route des Lucioles, BP 9306902, Sophia Antipolis Cedex, France
Jonathan.Stoeckel@sophia.inria.fr

Abstract. User guided systems for X-ray/portal image matching are commonly employed in radiation therapy for position verification and correction. To enable extended clinical support a fully acceptable solution must be generic, accurate, fast and automatic. In this paper a new registration system using the structure tensor is compared with the performance of experienced radiographers. It is shown that local extremum lines (emanating from bone ridges) are invariant with respect to the photon energy and that they can be detected without a priori knowledge. Registration based on such creases is validated with a large database of images from anterior irradiation fields for prostate treatments. It appears that the precision of the structure tensor approach is significantly better than human observers. The execution time is less than 10 seconds while human supervision can be restricted to detect outliers only.

1 Introduction

In the European Union, cancer is the second leading cause of death. Radiotherapy is one of the most effective means for its treatment [1]. The objective of radiation therapy is to deliver a prescribed radiation dose to the tumor site to eradicate the cancer cells. Unfortunately, incorrect dose delivery may easily result in complications or recurrence of the tumor. For this reason it is of paramount importance to verify the patient's position during treatment.

The boundary of a treatment area is often prescribed with the aid of a classic X-ray image (henceforth we will refer to it as the *simulator image,* see Figure 1). Alternatively, such a 'reference image' may be reconstructed from CT-data (through ray-tracing) resulting in a digitally reconstructed radiograph, DRR [2]. During treatments a so-called *portal image* is acquired with the treatment beam using an electronic portal imaging device (EPID). Discrepancies in the patient setup can be derived from the comparison of the prescribed and given treatment field through image matching. Consequently, the patient's position can be corrected in subsequent treatments.
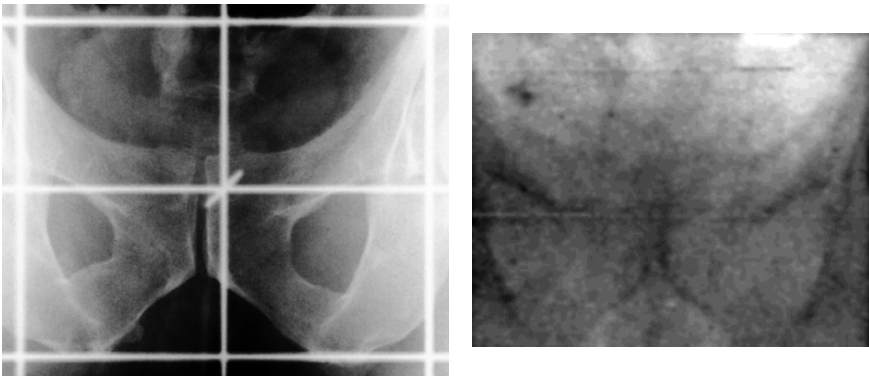
Fig. 1. Examples of a simulator (left) and portal image (right) of the pelvis. Both images were linearly stretched for clarity. The outer white 'rectangle' in the simulator image represents the treatment area.

In the 'Bernard Verbeeten Institution', the reference image consists of an simulator image because the DRR modality is not available. To do the image registration, a human operator manually delineates appropriate bone features in the simulator image. Subsequently, the template is positioned by hand at the corresponding location in the portal image through rotation and translation. Thus, the 'prescription image' (the simulator image) is related to the treatment area (depicted by the portal image) via the projected anatomy of the patient.

Although the sketched approach is generally accepted in clinical practice, it is well known to be sensitive to observer variability and time consuming due to the necessary user interaction. An ideal solution should meet the following requirements:

- fast
- automatic,
- accurate as well as precise,
- generic, i.e. not use explicit knowledge about the visualized anatomy.

In the literature, not many articles describe *X-ray - portal* image matching. In [3] [4] fiducial marks are extracted manually from the simulator image and a top-hat filter is used to recover them from the portal image. The registration is achieved through chamfer matching. A knowledge based approach was proposed in [5] applying a snake fitting algorithm to segment bone crests from images of the pelvis. At last [6] introduces a matched filter approach to emphasize bone features. Subsequently, the filtered images are registered through correlation. Although good results are given (a success rate of 95%), the experimental outcome is incomparable for the lack of a well described criterion of 'success'.

Clearly, none of the proposed solutions completely satisfies the requirements mentioned. In this paper a new technique will be described in which features are automatically extracted from both the simulator and the portal image without using *a priori* knowledge.

The paper is structured as follows. Section 2 gives a global description of our technique. Then, in Section 3 the effectiveness specifically regarding the accuracy and precision will be evaluated. For testing we will use images of the pelvic area made for prostate cancer treatment. Finally, in Section 4 we will summarize our findings and draw conclusions.

2   Approach

2.1   Outline

The image processing involved in the verification process comprises a number of stages as illustrated in Figure 2.
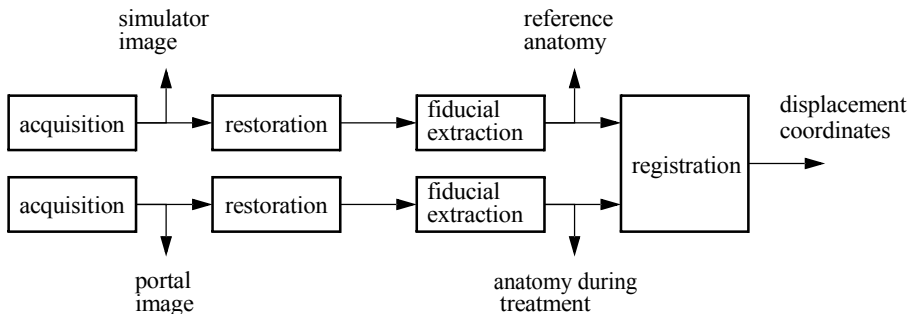
Fig. 2. Overview of the position verification system. The output consists of the displacement between two images.

The simulator and portal images are acquired with beams of very different photon energies (approximately 100 kV and 10 MV respectively). It can be seen in Figure 1 that portal images lack contrast and sharpness in the projected patient anatomy compared to low energy X-ray images. Moreover, they contain disturbing artefacts requiring restoration (for an elaborate description we refer to [7]). Also, corrections have to be included for the crosswires in the simulator image (Figure 1). It is unavoidable that by their presence image areas are occluded. The previously noticed difference in portal/simulator image structure has proven to be too large for correlation based methods [6][8]. For this reason we have opted to use so called fiducial 'structures' in the image that are assumed to be immobile with respect to the tumor (that often cannot be detected).

In the following sections we will investigate two types of fiducial features, describe techniques to extract them from the images and elaborate on the registration.

2.2   Definition of fiducial features

To allow proper verification of the patient's position fiducial features should satisfy the following prerequisites:

- detectable in both simulator and portal image,
- fixed with respect to the treatment target,
- in a reproducible position with respect to the radiation energy.

The most prominent aspects of portal images are the projections of bony structures, air cavities and body outlines. It has been proven that local extremum lines (ridges in the simulator image, troughs in the portal image) correspond to lines of maximum bone thickness [9]. In the remainder we will refer to such features as 'creases'. In the context of prostate treatments, it is generally assumed in clinical practice that the pelvic bone satisfies the second condition [10][11]

Let us consider the influence of photon energy on the position of two types of image features, namely *creases* and *edges* (i.e. transitions in image intensity). To this end consider a mono-energetic source and an image line detector of which the measured signal is linear with the number of incoming photons. Additionally, let us assume that the treatment site consists only of soft tissue and bone. Then, the measured signal on position x at the detector can be modelled by:

$$I(x) = I_0 \cdot e^{-\mu_b(\varepsilon)d_b(x) - \mu_s(\varepsilon)d_s(x) - \mu_a(\varepsilon)d_a(x)}, \tag{1}$$

where I_0 reflects the source intensity and the sensitivity of the detector, $\mu_y(\varepsilon)$ represents the attenuation coefficient of a material y as a function of energy, and $d_y(x)$ the pathlength through the material. The subscripts are b for bone, s for soft tissue and a for air.

The position of *creases* in the image (i.e. lines of maximal bone thickness) are found where the *first derivative* of I is zero. Furthermore, *edges* (due to bone/soft tissue transitions) are at locations where the *second derivative* has a zero crossing.

By definition:

$$C_{sd} = d_b(x) + d_s(x) + d_a(x) \tag{2}$$

is constant, representing the distance between source and detector.

To find an expression for both types of features, we will assume that the thickness of the patient is a linear function of the position x. Consequently:

$$d_a(x) = C + kx, \tag{3}$$

with k a constant.

Setting the first derivative of $I(x)$ (Equation 1) to 0 and solving for $d\,d_b(x)/dx$ we find (using Equation 2 and Equation 3):

$$\frac{d}{dx}d_b(x) = k\frac{(\mu_s(\varepsilon) - \mu_a(\varepsilon))}{(\mu_b(\varepsilon) - \mu_s(\varepsilon))}. \tag{4}$$

Similarly, setting the second derivative of $I(x)$ to zero results in:

$$\frac{d}{dx}d_b(x) = k\frac{(\mu_s(\epsilon) - \mu_a(\epsilon))}{(\mu_b(\epsilon) - \mu_s(\epsilon))} \pm \frac{1}{\sqrt{\mu_b(\epsilon) - \mu_s(\epsilon)}}\sqrt{\frac{\partial^2}{\partial x^2}d_b(x)} \; . \tag{5}$$

Note that Equation 4 recurs in Equation 5.

The right hand side in *Equation 4* is approximately constant (1.04 at 100 kV and 1.10 at 10 MV with k = 1). Consequently, there is practically no dependence on the energy level. The position of ridges is determined merely by the first derivative of d_b. The second term in *Equation 5*, however, may change dramatically as a function of energy as the factor before the square root varies from 2.5 to 7.1. Thus, the remaining influence of the attenuation coefficients may easily result in displacement of edge features. This comparison illustrates the superiority of creases over transitions as fiducial marks.

2.3   Fiducial feature extraction

In previous work we have tested several automatic crease extraction techniques [12]. Following this study we have opted to use a crease measure based on the structure tensor[13]. The structure tensor (F) is defined by the dyadic product of the gradient vector with itself:

$$\bar{F} = \overline{(\nabla f)(\nabla f)^t} = \begin{bmatrix} \overline{f_x f_x} & \overline{f_x f_y} \\ \overline{f_x f_y} & \overline{f_y f_y} \end{bmatrix} \; . \tag{6}$$

Here, f_x and f_y denote the partial derivatives of an image function $f(x,y)$. They are calculated through Gaussian derivatives. Also, a Gaussian is used to 'smoothen' the structure tensor (depicted by the overbar).

Let $\lambda_1 \geq \lambda_2$ be the eigenvalues of the structure tensor at a certain point on f, and v_1, v_2 the corresponding eigenvectors. A point is qualified as a ridge point if the first derivative in the direction of v_1 is zero.

In the description above, it can be seen that the width of the Gaussian derivative introduces a (first) scale parameter. This parameter is used only to suppress the effect of spatially uncorrelated noise on the derivatives. Therefore, it is chosen small, to correspond as much as possible to the inner scale of the image (i.e. the smallest features of interest). A second parameter is introduced by the Gaussian tensor smoothing. This parameter is tuned to prevent spurious creases.

It should be noticed that in our approach, changing the amount of tensor smoothing does not change the position of creases. It only influences the scale on which they are found. If a scale space of the input image were used instead, creases are found in different positions for different scales of the image [14].

In addition to the structure tensor approach, we have implemented the technique proposed by Gilhuijs [4] for comparison. Here, a top-hat transformation [15] is applied to the input image:

$$f_{out} = f_{in} - (f_{in} \bullet K) \quad , \tag{7}$$

where the operator • represents greyscale opening (i.e. morphological erosion followed by dilation). The kernel is rectangular of small size, typically 7x7 pixels. To identify the creases the transformed image (f_{out}) is merely thresholded. Both the exact kernel size and threshold level are to be determined experimentally.

In Section 3 we will evaluate both approaches.

2.4   Registration

To match the crease images, standard chamfer matching [16] was modified as follows. Instead of taking the distance transform to calculate the offset between two images, the 'reference-ridge image' is inverted after applying a Gaussian filter. Firstly, this allows the match to take into account extra information about the strength of individual points. Secondly, the influence of found crease points is limited locally, which is controlled by the width of the Gaussian. This is important in case of both over- and under-detection.

While performing the registration we will only consider translations and in-plane rotations. Thus, out-of-plane rotations are neglected. (Earlier, only poor results have been reported regarding the estimation of out-of-plane rotations based on two-dimensional reference images [9]).

From the preceding description it can be seen that our technique does not require explicit knowledge about the patient anatomy (conform requirement 3 in the introduction). Moreover, the user interaction is limited to a posteriori check of the result.

The remaining requirement (accuracy) will be explored in the next section.

3   Results

To evaluate our method we will focus here on the accuracy and the precision. It may be expected that they differ with the principal direction, i.e. transversal or longitudinal, due to differences in orientation information. For instance, in the images of the pelvis vertical structures dominate. Also, outliers will effect the outcome as a consequence of the occlusion of key bone structures by the crosswires. At last estimating in-plane rotation in addition to translation may contribute to a larger standard deviation due to the surplus in degrees of freedom.

To monitor the effect of such influences we conducted the following experiments. Over one year we randomly selected 57 patients visiting our institution for prostate cancer treatment. For each patient an anterior simulator image was obtained as well as several corresponding portal images (portal images were acquired in approximately 11 treatments per patient). Two experienced clinical physicists aligned each portal-simulator image pair manually. The averaged results of these two matches defined our gold standard. This ground truth was compared with the outcome by radiographers in clinical practice, the top-hat based technique and our method. The parameters of the latter two approaches were experimentally determined by minimizing their standard deviation.

In Table 1: the mean distances and corresponding standard deviations are tabulated between the ground truth and each of the three alternatives. To this end 549 images were included without in-plane rotation according to the gold standard. Consequently, they were aligned using translation only. The field dimensions typically were $(100 \text{ mm})^2$

digitized at a resolution of 75 dpi. The execution time for one match was on the order of a few seconds.

Table 1: mean distances and corresponding standard deviations between matches of experts (defining the gold standard) and three alternative 'systems'.

matching by translation only		radio-graphers	tophat	structure tensor
# of images		549	549	549
transversal:	mean (mm)	-0.08	0.11	0.27
	st.dev (mm)	1.28	1.87	1.61
longitudinal:	mean (mm)	0.52	0.49	0.55
	st.dev (mm)	1.33	2.04	1.87

All three matching 'systems' appear about equally *accurate* and not significantly different from the gold standard. The systematic difference in transversal and longitudinal *precision* (standard deviation) is caused by the shape of the anatomy: vertical structures dominate. Also, the precision of both the tophat as well as the orientation tensor based technique is notably worse than the human observers. Further analysis showed that only a few outliers are at the basis of this outcome. As was noticed previously, the cross hairs in the imaging device may block the central part of the portal image. In that case the absence of key bone structures such as the symphysis contributes to a larger standard deviation.

Most of the outliers will be detected immediately by a human observer. For this reason it seems obvious to exclude the matches that exceed a minimum distance from the gold standard. Table 2: shows the result when matches are ignored that deviate more than 7.5 mm. In this case, apparently no large differences in accuracy are found with the gold

Table 2: As in Table 1:, but leaving out those matches that are more than 7.5 mm from the gold standard for each system separately.

matching by translation only		radio-graphers	tophat	structure tensor
discarded outliers (n=549)		0.2%	2%	2.2%
transversal:	mean (mm)	-0.08	0.05	0.19
	st.dev (mm)	1.28	1.42	0.90
longitudinal:	mean (mm)	0.5	0.45	0.57
	st.dev (mm)	1.27	1.43	0.98

standard. The precision of the structure tensor technique may not be dramatically smaller than both the radiographers or the tophat approach, but it is significantly better by the Fisher test.

To study the influence of rotation, 78 images from 8 patients were analyzed in which small in-plane revolutions were indicated by the gold standard (all smaller than 3°). Table 3: shows the outcome. It may come as a surprise that the radiographers matched the

Table 3: as in Table 2:, now also matching by rotation by radiographers and the structure tensor approach

matching by translation and rotation		radio-graphers	structure tensor
# of images		78	76
transversal:	mean (mm)	-0.28	0.11
	st.dev (mm)	1.49	1.11
longitudinal:	mean (mm)	0.42	0.42
	st.dev (mm)	1.45	1.05
rotation:	mean ($^{\circ}$)	-	0.36
	st.dev ($^{\circ}$)	-	0.89

images using translation only in *all* cases. As a logical consequence, the registration turned out to be less precise (with standard of deviations approximately 1.4 mm). In our view this only emphasizes the difficulty of retrieving in-plane rotations by humans. Contrasting to this, the rotation-angle appeared well estimable by the automatic techniques (with a precision on the order of 0.9°). The effect of the additional parameter on the transversal/longitudinal accuracy and precision was negligible.

4   Discussion/conclusion

Comparing our results using the tophat technique with those reported in the literature, we noticed remarkably worse performance in our outcome. In [4] the precision of both the radiographers as well as the tophat technique is approximately twice as good (with standard deviations of about 0.5 mm). This difference might be attributed to the quality of the portal images. It is well known that our fluoroscopic portal imaging systems yield worse signal to noise ratio than liquid filled ionisation chambers employed by [4] (see also [3]). Another remarkable difference is that in [4] the reference template is drawn manually.

From the preceding sections we conclude that a registration system for portal/simulator image matching should be generic, accurate, automatic and fast. Local extremum lines

can be detected without using a priori knowledge; their position is invariant with the photon energy. The accuracy of an automatic system based on the structure tensor is not significantly different from experienced physicists as well as radiographers (the latter doing the work in clinical practice). The precision is significantly better than that of human observers when outliers are discarded. Moreover, in-plane rotations are taken into account *without* sacrificing accuracy or precision. These results are obtained when user control is admitted to detect outliers only. The execution time of our system is less than 10 seconds which is sufficient for clinical application.

References

[1]   P. Rubin, "Clinical Oncology - A multidisciplinary approach for physicians and students", 7th edition, Saunders & Co, Philadelphia, 1993.

[2]   R. Bansal, L. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, J. Duncan, "A minimax entropy registration framework for patient setup verification in radiotherapy", Computer Aided Surgery, vol. 4, pp. 287-304, 1999.

[3]   K.G.A. Gilhuijs, J. Bijhold, M. van Herk, H. Meertens, "Automatic on-line inspection of patient set-up in radiation therapy using digital portal images", Medical Physics, vol. 20, pp. 667-677, 1993.

[4]   K.G.A. Gilhuijs, A. Touw, M. van Herk, R.E. Vijlbrief, "Optimization of automatic portal image analysis", Medical Physics, vol. 22, pp. 1089-1099, 1995.

[5]   F. Yin, K. Nie, L. Chen, C.W.Chen, "Automated extraction of pelvic features in portal and simulator images", SPIE Conference on Medical Imaging and Image Processing, Newport Beach, California, Vol. 2710, pp. 1020-1023, 1996.

[6]   F. Kreuder, B.Schreiber, C. Kausch, O. Doessel, "A structure based method for online matching of portal images for an optimal patient set-up in radiotherapy", Philips Journal of Research, vol. 51, no. 2, pp. 317-337, 1998.

[7]   J. Stoeckel, "Simulator portal image matching - a patient position verification system for radiotherapy", graduation report, Pattern Recognition Group, Technical University Delft, 1999.

[8]   J. Moseley, P. Munro, "A semiautomatic method for registration of portal images", Medical Physics, vol. 21, no. 4, pp. 551-557, 1994.

[9]   K.G.A. Gilhuijs, P.J.H. van de Ven, M. van Herk, "Automatic three-dimensional inspection of patient setup in radiation therapy using portal images, simulator images, and computed tomography data", Medical Physics, vol. 10, no. 6, pp. 389-399, 1996.

[10]   J.A. Antolak, I.I. Rosen, C.H. Childress, G.K. Zagars, A. Pollack, "Prostate target volume variations during a course of radiotherapy", International Journal of Radiation Oncology, Biology and Physics", vol. 42, no. 3, pp. 661-672, 1998.

[11]   M.J. Zelefsky, D. Crean, G.S. Mageras, O. Lyass, L. Happerset, "Quantification and predictors of prostate position variability in 50 patients evaluated with multiple CT scans during conformal radiotherapy", Radiotherapy and Oncology, vol. 50, no. 2, pp. 225-234, 1999.

[12]   J. Stoeckel, F.M. Vos, P.H. Vos, A.M. Vossepoel, "An evaluation of ridge extraction methods for portal imaging", submitted for publication to ICPR conference 2000.

[13]   M. Kass, W.Witkin, "Analyzing oriented patterns", Computer Vision, Graphics and Image Processing, vol. 37, pp. 362-385, 1987.

[14] D.S. Fritsch, E.L. Chaney, A. Boxwala, M.J. McAuliffe, S. Raghavan, A. Thall, J.R.D. Earnheart, "Core-based portal image registration for automatic radiotherapy treatment verification", International Journal for Radiation Oncology, Biology and Physics, vol. 33, no. 5, pp. 1287-1300, 1995

[15] J. Serra, "Image analysis and mathematical morphology", 2nd edition, Academic Publishers, San Diego, Ca, 1988.

[16] G. Borgefors, "Hierarchical chamfer matching: a parametric edge matching algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 10, no. 6, pp. 849-865, 1988.