Fast DRR generation for intensity-based 2D/3D image registration in radiotherapy

D. Sarrut [*,a,b] S. Clippe [b,a]

[a] *laboratoire LIRIS, Bât L, Université Lumière Lyon 2, 5 av. Pierre Mendès-France, 69676 Bron*

[b] *Radiotherapy departement, Centre de Lutte Contre le Cancer Léon Bérard, 28 rue Laennec 69008, Lyon*

Abstract

Conformal radiotherapy requires accurate patient positioning according to a reference given by an initial 3D CT image. Patient setup is controlled with the help of portal images (PI), acquired just before patient treatment. To date, the comparison with the reference by physicians is mostly visual. Several automatic methods have been proposed, generally based on segmentation procedures. However, PI are of very low contrast, leading to segmentation inaccuracies. In the present study, we propose an intensity-based method, with no segmentation, associating two portal images and a 3D CT scan to estimate patient positioning. The process is a 3D optimization of a similarity measure in the space of rigid transformations. To avoid time-consuming DRRs (Digitally Reconstructed Radiographs) at each iteration, we used 2D transformations and sets of DRRs pre-generated from specific angles. Moreover, we propose a method for computing intensity-based similarity measures obtained from several couples of images. We used correlation coefficient, mutual information, pattern intensity and correlation ratio. Preliminary experiments, performed with simulated and real PIs, show good results with the correlation ratio and correlation coefficient (lower than 0.5 *mm* median RMS for tests with simulated PI and 1.8 *mm* with real PI).

Key words: conformal radiotherapy, image registration, Digitally Reconstructed Radiographs, correlation ratio,

---

* Corresponding author.
Email addresses: `dsarrut@univ-lyon2.fr` (D. Sarrut), `clippe@lyon.fnclcc.fr` (S. Clippe).

Preprint submitted to Elsevier Science *3 June 2003*

1 Introduction

For more than one century radiation therapy has used X-rays to treat cancer. It has become one of the three main cancer treatment modalities, together with surgery and chemotherapy. To be efficient, radiation therapy must deliver a maximum dose of X-rays (now produced by linear accelerators) to the tumor while sparing surrounding normal tissue.

Before the beginning of treatment, physicians and physicists have to establish a Radiotherapy Treatment Planning (RTP). The RTP defines the number of beams, their size, their shape, their tilt and the beam energy. This is now done in 3D with the help of a computed tomography (CT) scan of the patient. However, delivering X-ray doses is a fractionated process.

For instance, at least 35 daily fractions are necessary to treat prostate cancer. But, installing the patient in exactly the same position every day is very difficult. This position is defined by the CT-scan used for establishing the RTP. Thus, the main difficulty is the day-to-day reproducibility of the patient setup. To check the positioning of the patient on the treatment couch, radiation therapists usually use only skin marks.

For many years, displacements have been related. Mean setup errors are between 5.5 mm and 8 mm [1, 2, 3] with a maximum, though rarely reported, of 18 mm [4] or 16 mm [5]. Even in recent series, using of immobilization devices, displacements remain important: 22 % are between 5 and 10 mm [6] and 57 % are over 4 mm [7]. If setup errors have often been measured, their consequences have rarely been evaluated. Three studies have reported a degradation of the therapeutic ratio caused by discrepancies between the planned and the delivered treatment positions [8, 9, 10].

The first solution proposed to reduce setup errors and their potential serious consequences is patient immobilization. Immobilization devices such as polyurethan foam cast or thermoplastic mask have been developed. Numerous studies have shown their usefulness in reducing setup error rates [2, 3, 11, 12]. However they do not eliminate all errors and several recent series failed to show any improvement with the use of immobilization devices [13, 14].

The second solution to improve patient setup is called *Portal Imaging*, which is generally used as a complement of immobilization systems. Films acquired on the beam exit site of the patient allow the verification of patient position by comparing them with a reference image. Recently developed Electronic Portal Imaging Devices (EPID) have several advantages [15, 16, 17]. First, images are obtained immediately, unlike films which need to be processed. EPID thus allow *on-line* (immediate) setup error detection and correction. Secondly, EPID provide digital imaging, with image processing abilities. Figure 1 shows a schematic view of one type of EPID.

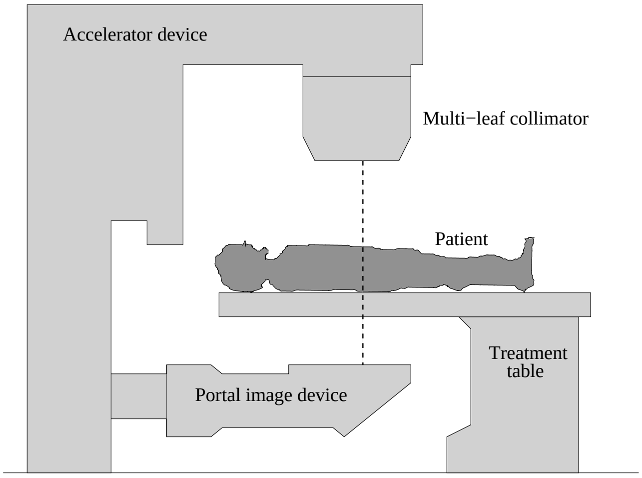

Fig. 1. Electronic Portal Imaging Device (EPID).

Control images are used to detect and quantify setup errors relative to the planned position defined by a reference image. To date this detection has only consisted in a visual inspection by the physician. But this is inaccurate, labor intensive and time-consuming. Hence, there is a real need for tools to help physicians in this tedious task. Moreover, conformal radiotherapy adds reduced margins around the target volume. Therefore, acute patient positioning becomes essential to make sure that no target is missed, and minimize the risk of local recurrence. After setup error estimation, the aim of radiotherapists is to correct patient position before each treatment session. To that end, remote-controlled treatment couches [18] and a tables with six degrees of freedom [19] can be used.

In this paper, we propose to develop a fully 3D, automatic method for detecting setup errors in conformal radiotherapy. The paper is organized as follows. Section 2 is an overview of image registration techniques used in the context of patient positioning in radiation therapy. The general principles and notations are described in section 3. Section 4 describes the fast digitally reconstructed radiographs (DRR) generation. In section 5, we describe similarity measures used to compare DRR and PI. Section 6 presents experimental results, analysis and discussion. Finally, Section 7 concludes the paper.

2 Background

For several years, image registration methods used to determine patient position have compared portal images, which represent the real position of the patient on the treatment couch, and a reference image which represents the expected position. In this paper we focus on rigid transformations, assuming that the patient's displacement is rigid. Methods are classified into two categories: *feature-based* methods using a segmentation step, and *intensity-based* with no segmentation step. Other criteria could be used, such as the different types of reference images: either simulator films, initial portal image validated by the physician, electronic portal images, or digitally reconstructed radiographs (DRRs). DRRs are 2D projection images [20]. They are computed with a specific volume-rendering from a CT-scan of the patient acquired before treatment which defines the reference position. Methods may either be 2D (three parameters to be determined: 2 translations and one rotation) or 3D (six parameters: 3 translations and 3 rotations). See also [21] for a review of setup verification in clinical practice.

2.1 Feature-based methods

Feature-based methods, using a segmentation step, are the most frequent methods used. Segmentation can be manual or automatic. In 1991, Bijhold *et al.* [22] presented one of the first methods of setup error measurement using portal images and a film as reference image. Image segmentation is done manually by delineating the bony outlines visible on both portal and reference images. Then, only the extracted, identical features are registered. Other methods use anatomical landmarks: three [23, 7] or five [24] homologous points determined in both images are matched. The main difficulty is the accurate definition of homologous points in the two images. The main drawback of these methods is that they generally only consider in-plane (2D) information, which is now known to be inaccurate in case of out-of-plane rotation or large translation [22, 25, 26, 27].

3D methods were then developed. The first one uses several portal images taken from different projections [28, 23]. In this case, the registration is still 2D; it is performed independently with each PI, so it has the same drawbacks related beforehand. Marker-based methods of registration have also been developed [29, 30]. Radio-opaque markers are implanted in the body of the patient to overcome the problems induced by tumor movements. However the markers have to be fixed in the tumor volume itself, which is an important restriction for implantation. The second impediment of this method is the uneasy detection of these markers on very low contrast portal images. A fully 3D method is proposed in [31, 32]. It is based on the registration of a 3D surface extracted from the CT scan with several image contours, segmented from the PI. Registration uses a least square optimization of the energy necessary to bring a set of projection lines (from camera to contour) tangent to the surface. Once structures have been extracted, the registration process is very quick. However it has never been experimented with real PI, which segmentation is much less accurate than with kilo-voltage radiographies.

Fully 3D techniques could also require DRRs. One difficulty in this context is the computational time. Two methods are proposed. DRRs could either be computed at each step of the optimization process (*on-line*), but this must be very quick because the patient is awaiting on the treatment couch, or before the registration step (*off-line*), when there is no time constraint, each DRR corresponding to a known position of the patient. In 1996, Gilhuijs *et al.* [33, 34] developed a 3D method with partial DRR (computed with several selected points in the scanner) in order to speed-up the process. Murphy [35] also use partial DRR. The method was further evaluated by Remeijer *et al.* [27] in 2000: it is quick but limited by a high failure rate of the segmentation step in the portal images. It consequently requires human manual correction.

To date, the segmentation step remains very difficult in PI because of their very low contrast, due to the high energy (5 - 20 Megavolt) of the X-rays used to acquire control images.

2.2 Intensity-based methods

A second class of methods uses the gray-level values of all the pixels in the images that must be compared. These methods assume that there is a statistical link between the pixel values of the images to be compared and that this link is maximum when the images are registered. For instance, Dong *et al.* [36] and Hristov *et al.* [37] used the linear correlation coefficient in a 2D method. They applied such measure to register portal images to modified DRRs (DRRs that have been filtered to resemble megavoltage images).

Recently, a method using mutual information (MI) was proposed by Maes *et al.* [38] and Wells *et al.* [39]. It obtained interesting results in the context of multimodality image registration between PET, MRI and CT-scan. Other intensity-based similarity measures have been studied, such as the correlation ratio [40]. Preliminary studies on image registration using MI in the context of patient setup in radiotherapy have been published [41, 42, 43].

3 Overview of the 2D/3D registration method

Considering published data and our preliminary works [41, 44], we decide to based our method on the following points:

- *No segmentation.* PI are noisy images with low contrast for which accurate, robust and automatic segmentation is uneasy [27]. Intensity-based methods seems to be more suitable to compare DRR and PI, because, in this case, the accuracy of the final estimation is not related to the precision of the PI segmentation. Such methods are based on a similarity measure (see section 5), denoted $\mathcal{S}$.
- *Use of several PI.* Some patient displacements cannot be accurately evaluated from a single PI (typically, a translation along the viewing direction leads to very small image scaling). Thus, several PI must be acquired from different projections. We denote n the number of PI used simultaneously, $\mathcal{I}^i$ the i^{th} PI and $\boldsymbol{P}^i$ the projection matrix associated with $\mathcal{I}^i$. Matrix $\boldsymbol{P}^i$ is given by a calibration procedure. In practice, the number of PI is limited by the acquisition time and by the amount of radiation received by the patient, and two PI ($n = 2$), acquired from orthogonal projections, are used. However, our approach can be used with any n.
- *DRR.* The 3D CT scan is denoted $\mathcal{V}$. We denote $\mathcal{D}^i_{\boldsymbol{U}}$ a DRR acquired according to the projection $\boldsymbol{P}^i$, for a given patient displacement denoted $\boldsymbol{U}$: $\mathcal{D}^i_{\boldsymbol{U}} = \boldsymbol{P}^i\boldsymbol{U}(\mathcal{V})$. $\boldsymbol{U}$ is a rigid transformation matrix with 6 parameters (3 translations, 3 rotations).
- *3D estimation.* It has been shown that out-of-plane rotations lead to inaccurate or false estimations [26]. Thus, 3D methods are required. Hence, our approach consists in performing a 3D optimization over the search space defined by the type of the transformation (here only rigid transformations are studied). We denoted $\overline{\mathcal{I}}$ the vector of n PI, and $\overline{\mathcal{D}_{\boldsymbol{U}}}$ the vector of the n corresponding DRR. The similarity criterion is defined as a global similarity measure $\mathcal{S}$ in each pair of PI-DRR.

$$\hat{\boldsymbol{U}} = \arg\max_{\boldsymbol{U}} \mathcal{S}(\,\overline{\mathcal{D}_{\boldsymbol{U}}}\,, \overline{\mathcal{I}}\,) \tag{1}$$

  The optimization of eq. (1) is done with the Powell-Brent procedure described in [45].
- *Combination of similarity.* We propose in section 5.3 a way to define similarity between *several* pairs of images.
- *DRR pre-computation.* At each iteration of eq. (1), there are n DRR generations and one evaluation of the similarity measure between n pairs of PI-DRR. However, DRR computation is a heavy task. Until now, most of the studies performing 3D estimation with DRR have focused on speeding up DRR generation [46, 33], but were detrimental to DRR quality (and thus to the quality of the similarity criterion). We propose here to replace the ex-

pensive DRR generation by another two-pass approach: pre-computed DRR and in-plane transformation. Such approach was also investigated in [47], but we propose here a mathematically justified decomposition.

4 Fast DRR generation

In order to overcome the expensive computational time of volume rendering, some authors [33] suggest that it is possible to generate a set of DRR computed according to a sampling of the search space before patient treatment. Such a step is only done once, before patient treatment (*offline*) and therefore does not have time constraints. However, the search space has six degrees-of-freedom, and even with as few as ten points along each axis, it leads to a database of 10^6 DRR, which is not tractable in practice. However, we see in the next section that pre-computing DRR can be very useful.

4.1 Principle

The procedure relies on a precomputed set of DRR. During the optimization (eq. (1)), the generation of a DRR according to a given projection $\boldsymbol{P}_1$ is done in two steps: first, an adequate DRR is chosen in the set and then it is deformed according to a 2D affine transformation $\boldsymbol{L}$. The next paragraphs shows how to choose the adequate out-of-plane DRR and how to compute $\boldsymbol{L}$.

Each image in the set of precomputed DRR is the projection of a rotation of the scanner, and this rotation is out-of-plane according to the optical axis of the initial projection, denoted $\boldsymbol{P}_0$ (the initial projection correspond to the projection of the portal image). Hence, such rotation has two degrees of freedom (see appendix F). The set of DRR is build by sampling the two parameters α, β. Out-of-plane rotations are bounded to a set of plausible rotations ($\pm 6°$) and sampled with a sampling step which allow sufficient accuracy (see experiments section 6).

4.2 Geometrical decomposition

Let $\boldsymbol{P}_0$ be the *initial projection* known by calibration process. Let $\boldsymbol{P}_1$ be the *objective projection*, corresponding to the DRR we want to generate. Parameterization is given in appendix A. Our goal is to find a projection $\boldsymbol{P}_h$ which is out-of-plane according to $\boldsymbol{P}_0$, and a matrix $\boldsymbol{L}$ such that $\boldsymbol{P}_1 \approx \boldsymbol{L}\boldsymbol{P}_h$. To do this, we consider two intermediate projections $\boldsymbol{P}_2$ and $\boldsymbol{P}_3$ such that we know how to build $\boldsymbol{P}_2$ from $\boldsymbol{P}_1$, $\boldsymbol{P}_3$ from $\boldsymbol{P}_2$ and $\boldsymbol{P}_h$ from $\boldsymbol{P}_3$. The figure 4.2 depicts the intermediate projections and their relationship in 2D (the optical center of projection $\boldsymbol{P}_i$ is denoted $\boldsymbol{c}_i$).

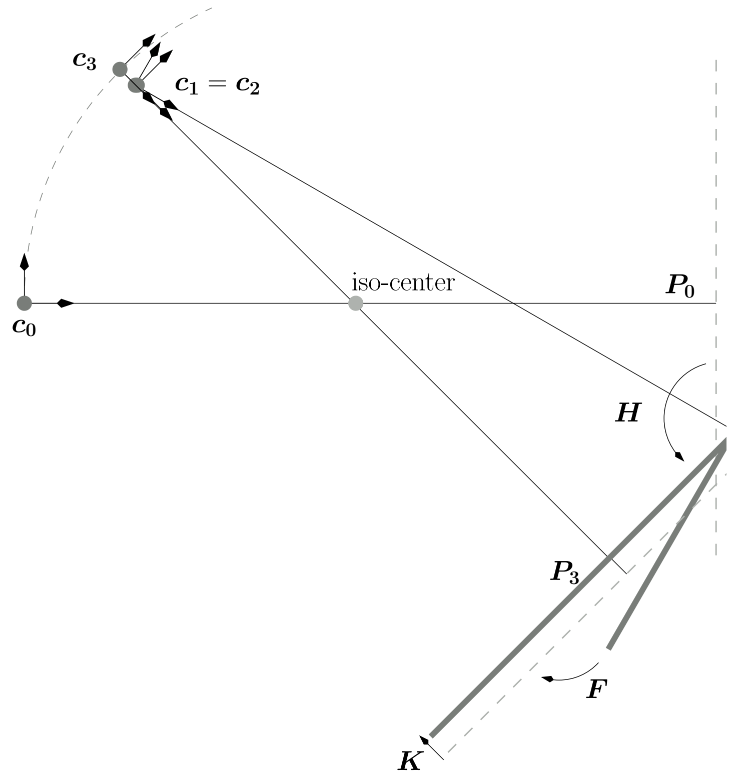

Fig. 2. Illustration in 2D of the projections $(\boldsymbol{P}_0, \boldsymbol{P}_1, \boldsymbol{P}_2, \boldsymbol{P}_3)$ and their optical centers $(\boldsymbol{c}_0, \boldsymbol{c}_1 = \boldsymbol{c}_2, \boldsymbol{c}_3)$. Some 2D transformations (the rectification $\boldsymbol{F}$ and the scaling matrix $\boldsymbol{K}$) are also depicted.

- We first consider the *corrective rotation* $\boldsymbol{R}$ and the projection $\boldsymbol{P}_2$ build as described in appendix B. The optical center $\boldsymbol{c}_2$ of $\boldsymbol{P}_2$ is the same that the optical center $\boldsymbol{c}_1$ of $\boldsymbol{P}_1$. Thus it exists a *rectification matrix* $\boldsymbol{F}$ such that $\boldsymbol{F}\boldsymbol{P}_2 = \boldsymbol{P}_1$ (see appendix C).

- Now, we consider the projection $\boldsymbol{P}_3$ (see appendix D), which has the same orientation than $\boldsymbol{P}_2$. The difference between $\boldsymbol{P}_2$ and $\boldsymbol{P}_3$ is the distance between the optical center and the isocenter. We build a scaling matrix $\boldsymbol{K}$ (see appendix E), such that $\boldsymbol{K}\boldsymbol{P}_3 \approx \boldsymbol{P}_2$. Note that if $\boldsymbol{c}_1$ is located on a sphere centered on the isocenter $\boldsymbol{s}$, with radius $\|\boldsymbol{d}\|$, we have $\boldsymbol{c}_1 = \boldsymbol{c}_2 = \boldsymbol{c}_3$, $\boldsymbol{K}$ is the identity matrix and $\boldsymbol{P}_3 = \boldsymbol{P}_2$.
- The last step consists in using the out-of-plane/in-plane rotation decomposition described in appendix F in order to write $\boldsymbol{P}_3$ according to an out-of-plane rotation of the initial $\boldsymbol{P}_0$. We thus obtain $\boldsymbol{P}_3 = \boldsymbol{C}'\boldsymbol{P}_h$, with $\boldsymbol{P}_h$ an out-of-plane projection and $\boldsymbol{C}'$ an in-plane rotation matrix.

Finally, we have:

- Intermediate projection $\boldsymbol{P}_2$ is build from $\boldsymbol{P}_1$ by an in-plane rectification matrix $\boldsymbol{F} : \boldsymbol{P}_1 = \boldsymbol{F}\boldsymbol{P}_2$
- Intermediate projection $\boldsymbol{P}_3$ is build from $\boldsymbol{P}_2$ by an in-plane scaling matrix $\boldsymbol{K} : \boldsymbol{P}_2 \approx \boldsymbol{K}\boldsymbol{P}_3$
- Out-of-plane projection $\boldsymbol{P}_h$ is build from $\boldsymbol{P}_3$ by the decomposition : $\boldsymbol{P}_3 = \boldsymbol{C}'\boldsymbol{P}_h$
- Thus, we can write the objective projection as an in-plane affine transformation $\boldsymbol{L}$ of an out-of-plane projection $\boldsymbol{P}_h$:

$$\boldsymbol{P}_1 \approx \boldsymbol{F}\boldsymbol{K}\boldsymbol{C}'\,\boldsymbol{P}_h = \boldsymbol{L}\boldsymbol{P}_h \tag{2}$$

The computation of $\boldsymbol{L}$ and $\boldsymbol{P}_h$ only requires some vectors and matrices manipulations and has thus a negligible computational cost according to the remaining of the procedure. Moreover, applying the affine transformation $\boldsymbol{L}$ on a 2D image is straightforward and very fast. Finally, DRR generation (computation of $\boldsymbol{L}$ and $\boldsymbol{P}_h$ and application of $\boldsymbol{L}$ on a previously computed DRR) takes less than 20 milliseconds on a common $1.7Ghz$ PC.

5 Similarity measures

In the previous sections, we have advocated the use of intensity-based similarity measures to compare DRR and PIs. Such measures compare the relative position of two images. One image is the *template* image, denoted I_t, and the other one is the *floating* image I_f. In this paper, an *intensity-based similarity measure* $\mathcal{S} : I_t \times I_f \mapsto \mathbb{R}$ is a criterion which quantifies some type of dependence between the intensity distributions of the images. It does not require any segmentation. The most widely known similarity measures are the correlation coefficient, entropy or mutual information [38, 39, 46]. In this section, we describe the bases of these measures and discuss the choices that we have made.

5.1 Joint histograms

Joint histograms [38] are the underlying common base of most similarity measures used in image registration [48] (even if some measures do not require it, they can all be computed from joint histograms). This is a 2D histogram computed given a transformation $\boldsymbol{U}$: we denote it $H_{\boldsymbol{U}}$. It is defined on the intensity distribution of the images (called $\mathcal{D}_t$ and $\mathcal{D}_f$): $H_{\boldsymbol{U}} : \mathcal{D}_t \times \mathcal{D}_f \rightarrow \mathbb{R}^+$. $H_{\boldsymbol{U}}(i,j) = n_{ij}$ is the number of points such as $I_f(\boldsymbol{x}) = j$ and $I_t(T(\boldsymbol{x})) = i$. Because of the discrete nature of images, $T(\boldsymbol{x})$ does not generally coincide with a point of I_f, and an interpolation procedure must be applied.

Probabilities p_{ij} must be estimated from quantities n_{ij}. Like most authors, we use a frequential estimation of the distributions: $p_{ij} = \frac{n_{ij}}{n}$ (where n is the number of points in the overlapping part of the images). Joint histogram is thus a contingency table and criterion $\mathcal{S}$ measures some type of dependence between the two intensity distributions. According to the type of the dependence (*e.g.* linear, functional), the considered type of intensities (numerical or categorical), and the different *diversity* measures used (variance, entropy), several measures can be defined [48].

5.2 Choice of the similarity measure

Four measures have been studied: *correlation coefficient* and *mutual information* have already been used in the PI-DRR comparison; *Pattern intensity* was proposed by Penney *et al.* [46] for the registration of a 3D CT with a 2D fluoroscopy image; finally *correlation ratio* was proposed by Roche *et al.* [49].

The correlation coefficient (denoted $\mathcal{S}_{CC}$) [36, 37], assumes that a linear relation exists between the images intensities and measures the strength of this relation. The mutual information ($\mathcal{S}_{MI}$) [38, 50, 39] computes a distance to the independence case, by way of the relative Shannon's entropy. $\mathcal{S}_{MI}$ is maximal when a functional dependence exists between the intensities. The pattern intensity ($\mathcal{S}_{PI}$) [46] measure the "smoothness" of the image difference $I_{diff} = I - sJ$ for each pixel in a small neighborhood. It requires three parameters: σ (see eq.(6)), r which define the size of the neighborhood and s which is a scaling factor used to build I_{diff}. The goal of this measure is to enhance bony structure correspondence, but it was designed for DRR/fluoroscopy registration and not for DRR/PI. The correlation ratio [40] (denoted by $\mathcal{S}_{CR}$) assumes a functional relation between the intensity distributions and measures its strength by the way of a proportional reduction of variance. Eq. (3), (4), (5) and (6) express the three measures (with the mean $m_I = \sum_i ip_i$, the variance $\sigma_I^2 = \sum_i (i - m_I)^2 p_i$, and the conditional variance $\sigma_{I|j}^2 = \frac{1}{p_j} \sum_i (i - m_{I|j})^2 p_{ij}$):

$$\mathcal{S}_{CC}(I, J) = \sum_i \sum_j \frac{(i - m_I)(j - m_J)}{\sigma_I \, \sigma_J} p_{ij} \tag{3}$$

$$\mathcal{S}_{MI}(I, J) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_i p_j} \tag{4}$$

$$\mathcal{S}_{CR}(I|J) = 1 - \frac{1}{\sigma_I^2} \sum_j p_j \sigma_{I|j}^2 \tag{5}$$

$$\mathcal{S}_{PI}(I, J) = \sum_{x,y} \sum_{(k-x)^2+(l-y)^2)<r} \frac{\sigma^2}{\sigma^2 + (I_{diff}(x, y) - I_{diff}(k, l))^2} \tag{6}$$

$\mathcal{S}_{CR}$ is not a symmetric criterion (whereas the others are), an image (I) must be chosen for estimating the other (J). Optimization (1) compares the same PI with several DRR and $\mathcal{S}_{CR}$ is normalized according to I, so we decided to define $I = DRR$ and $J = PI$.

5.3 Similarity of several pairs of images

The main optimization eq. (1) consists in optimizing a *combination* of similarity measures between several pairs of PI/DRR images (a pair for each PI acquisition). A simple procedure would consist in a linear combination of the similarity values $\sum_i^n \alpha_i \mathcal{S}(PI_i, \mathcal{D}_i)$. However, it is not clear how to determine the weights α_i. In this study, we advocate the following method:

- For measures explicitely based on joint histogram (mutual information): instead of computing a similarity value for each pair and aggregating the obtained value, we update the same *identical* joint histogram for all the PI/DRR pairs. Then, we compute a single similarity measure from this unique histogram. This is a general method which makes possible to use any histogram-based similarity measure (it was also used in [51]). Moreover, such a method forces the criterion (whatever its nature) to measure the *same* type of intensity dependence for each pair of images. On the opposite, if similarities were computed for each pairs independently and then aggregated, each one would measure different kinds of relations between PI and the DRR intensities.
- For the correlation ratio and the correlation coefficient: the computation of the joint histogram is not needed. However, the same idea is applied, intensities probabilities are computed with all the PI/DRR pairs.

- For the pattern intensity, the measures are added.

6 Experimental tests

6.1 Methods

Several tests were performed to progressively verify the validity of our assumptions. In all these experiments, error between two rigid transformations is evaluated with the *RMS error* (Root Mean Square). It is computed as the average distance (in millimeters) between a set of 3D points transformed by the two compared transformations. We used 1000 points spread inside a sphere of $150\,mm$ diameter centered at the target point (the tumor) in order to obtain a realistic error; a RMS error of $x\,mm$ means that the average distance between each point and its desired position is $x\,mm$.

- Test 1 aims to evaluate the accuracy of the matrix decomposition described section 4 according to the objective projection matrix. We generate 1000 random rigid transformations $\boldsymbol{U}_i$, with translations lower than $20mm$ and rotation lower than $6°$ (uniformly distributed). Such transformations lead to RMS errors between $10mm$ and $30mm$. It corresponds to observed displacements (see section 2). For each position $\boldsymbol{U}_i$, we compare $\boldsymbol{P}_0\boldsymbol{U}_i$ and $\boldsymbol{L}\boldsymbol{P}_h$ with the RMS error (where $\boldsymbol{L}$ and $\boldsymbol{P}_h$ are computed as described section 4).
- Test 2 aims at quantify of the error due to the scaling matrix. In appendix E, we explain that we approximate the projection by using a scaling matrix. We generate 1000 other random rigid transformations $\boldsymbol{U}_i$, with translations lower than $100mm$. We measure the RMS error according to the scaling factor κ.
- Test 3 compares volume-rendering computed DRR with the corresponding approximated DRR. For test 3,5 and 6, we use a CT scan of a pelvis (52 $5\,mm$-thick slices, composed of 512×512 $0.93mm^2$ pixels). In order to evaluate the quality of the approximated DRR, we generate 1000 reference DRR (with 1000 random $\boldsymbol{U}_i$). Then, for each $\boldsymbol{U}_i$, we generate an approximated DRR by applying $\boldsymbol{L}$ to an out-of-plane DRR, corresponding to the exact $\boldsymbol{P}_h$. We do not consider sampling here (see test 4). We compare the reference DRR to the approximation with the mean intensity absolute difference (*MIAD*) computed by averaging the intensity difference between corresponding pixels. We also compute the *maximum* intensity difference between pixels. When applying the 2D affine transformation $\boldsymbol{L}$, we test nearest neighbor and bilinear interpolations.
- Test 4 quantifies the influence of the sampling step of the precomputed set of DRR (section 4). Test 1 has been repeated by changing out-of-plane rotations sampling, from $0°$ to $3°$ every $0.25°$ (for both out-of-plane angles).

- Test 5 evaluates the accuracy of the method without the multimodality registration difficulty (very low contrast in PI). For 100 random rigid transformations $\boldsymbol{U}_i$, we generate two DRR corresponding to the front and lateral view (with orthogonal projections $\boldsymbol{P}^1$ and $\boldsymbol{P}^2$), which play the role of the portal images. The procedure is then run in order to retrieve the correct $\boldsymbol{U}_i$. The starting point is vector null, the similarity measure is correlation ratio, the interpolation is nearest neighbor, the sampling step of the set of out-of-plane DRR is 1°, this set contains (for each view) 225 DRR (from $-7°$ to $7°$). For each run, RMS error is computed between the correct $\boldsymbol{U}_i$ and the estimation found.
- Test 6 evaluates the complete procedure with real PI and compares several similarity measures. The last test uses PI acquired on a patient corresponding to the previously described CT scan (see figure 3). PI are acquired with the iView GT (aSi) from Elekta. Irradiation field is $117 \times 130\,mm$ (490×540 pixels). We first manually register these images according to the scanner and denote the resulting position $\boldsymbol{U}_{ref}$. The manual registration is performed by a radiation oncologist with the help of interactively generated DRR (computed with the presented method) and manual delineation of anatomical structures projected in both portal images and DRR. We simulate 30 known patient displacements by performing the procedure with a random starting point corresponding to a rigid transformation $\boldsymbol{U}_i$. For each position estimation we compute the RMS error according to $\boldsymbol{U}_{ref}$. The four presented similarity measures are tested. Nearest neighbor interpolation is used (linear or partial volume interpolation increases computational time and does not clearly improve the results). Parameters for pattern intensity was $r = 3$ pixels and $\sigma = 10$ [46].

6.2 Results

- Test 1. The mean RMS error for this 1000 experiments is $0.04\,mm$.
- Test 2. The figure 4 depicts the RMS error according to κ (see appendix E). Each point shows the RMS error for a given $\boldsymbol{U}_i$. Note that the RMS error is quasi linear according to κ (the asymptotic standard error is lower than 0.006).
- Test 3. The quality of approximated DRR (with 256 grey levels) is summarized in table 1. Linear interpolation seems to gives better results than nearest neighbor interpolation. The mean error is very low: visually, the images are indistinguishable.
- Test 4. Figure 5 shows the RMS error according to the sampling step of the out-of-plane rotations. Error lower than $1mm$ can be achieved even with a sampling step lower than 1.5°.
- Test 5. 6% of the tests failed to achieve an correct position estimation that is to say led to a RMS greater than $3\,mm$. Such failures are due to local optimum in the optimization procedure. The remaining estimations (94%)

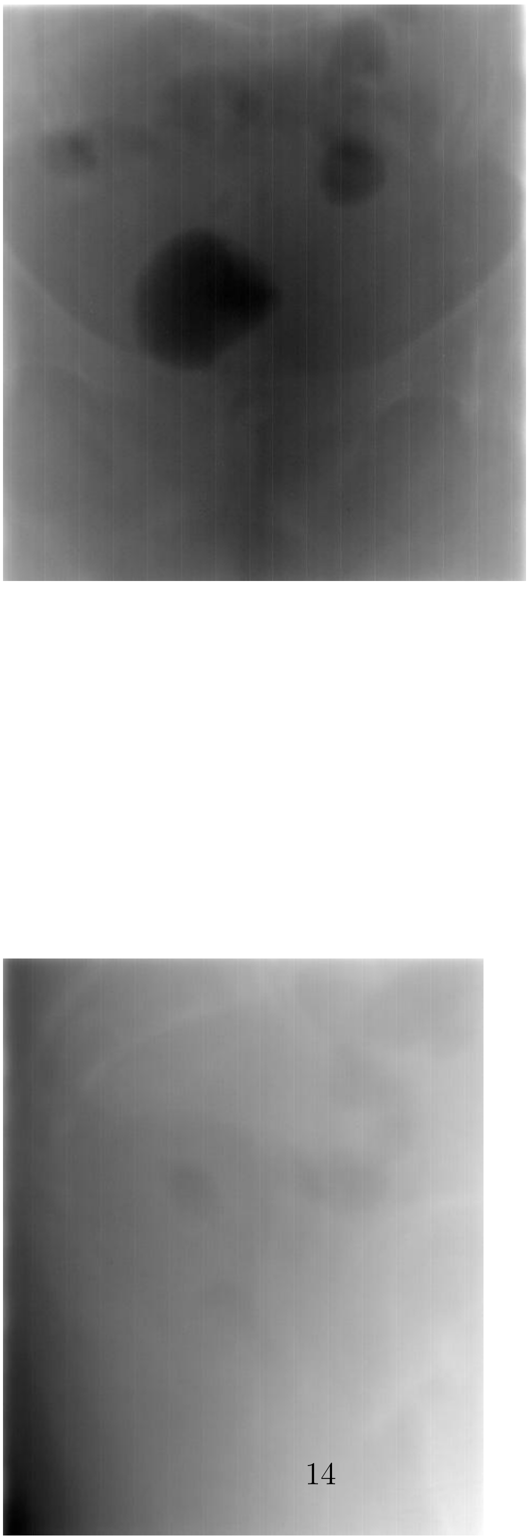

Fig. 3. Portal images (front and lateral view). Note the very poor contrast of the lateral view.

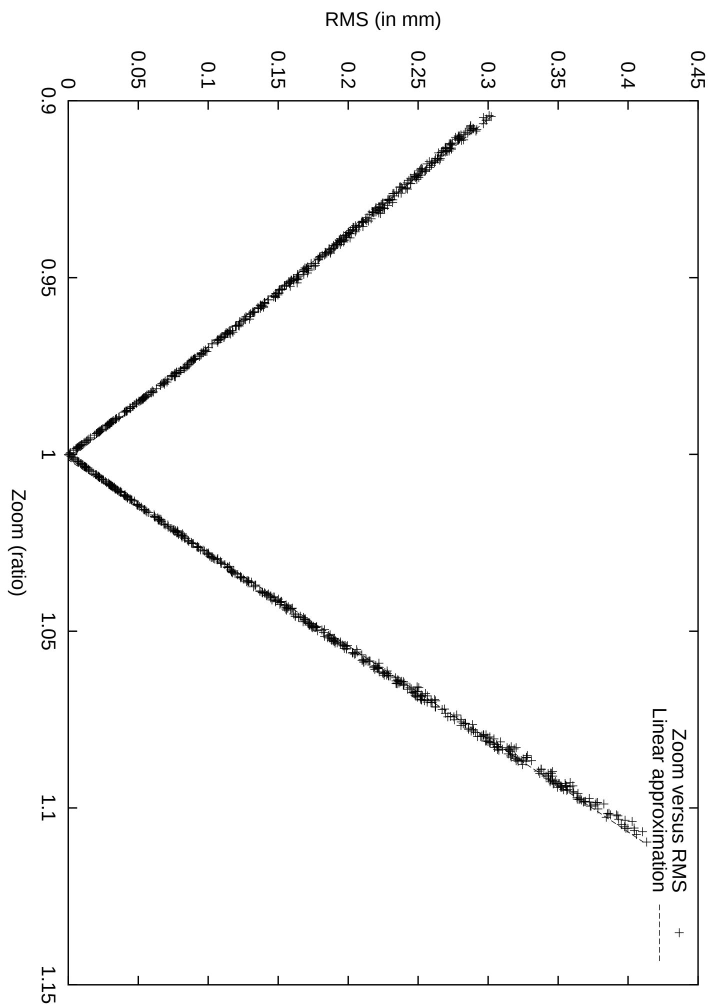

Fig. 4. Scaling ratio according to RMS error.

	MIAD	Max
Linear interpolation	0.44	17
Nearest neighbor	0.75	27

Table 1
Quality of approximated DRR. Results are mean of 1000 DRR comparisons (unity is grey level among 256).

have RMS error lower than $1.8mm$, 89% lower than $1mm$ and 56% lower than $0.5mm$ (median is $0.45mm$).

- Results of Test 6 are summarized in table 2. It shows mean and median RMS error, the number of RMS error greater than $3\,mm$ and the number

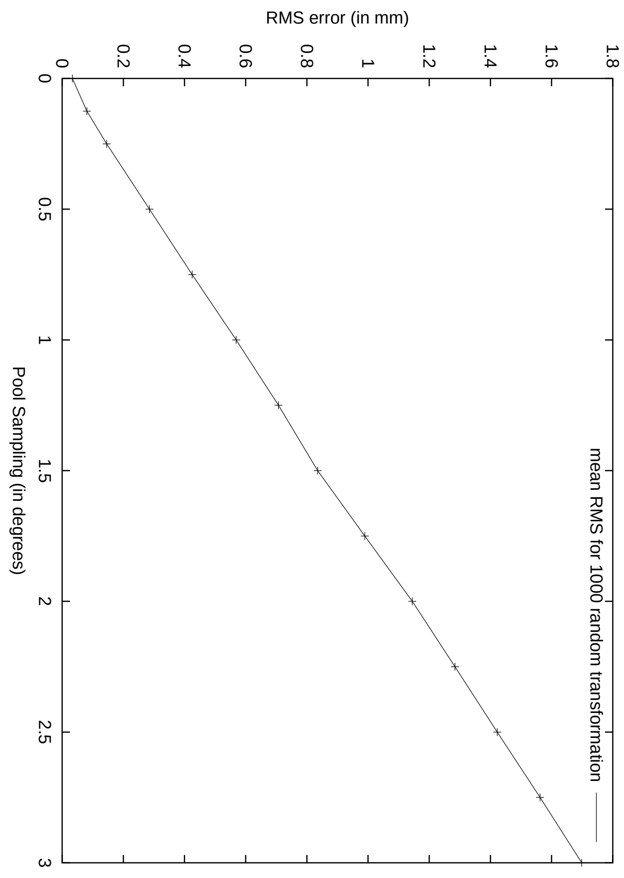

Fig. 5. Each point is the mean RMS error of 1000 random transformations according to different sampling (from 0° to 3° every 0.25°)

of registration which improve the RMS error of the starting position. We also indicate the mean number of iteration needed to converge. $\mathcal{S}_{MI}$ and $\mathcal{S}_{PI}$ are not suited for PI/DRR registration. $\mathcal{S}_{CC}$ and $\mathcal{S}_{CR}$ gives interesting results: RMS error close to $2\,mm$. The computational time depends on the time needed by one iteration and by the number of iteration. Time needed to compute $\mathcal{S}_{CC}$ is taken as reference ($T = 1$). Similarity evaluation is faster for $\mathcal{S}_{CC}$ and $\mathcal{S}_{CR}$ (no joint histogram is build) than for $\mathcal{S}_{MI}$ (1.8 times slower) and for $\mathcal{S}_{PI}$ (3.8 times slower). $\mathcal{S}_{CR}$ takes generally less iterations than $\mathcal{S}_{CC}$ to converge.

Measure	Mean	Median	#F	#I	Iter.	T
$\mathcal{S}_{CC}$	2.1	1.7	7	30	715	1
$\mathcal{S}_{MI}$	23	20	30	2	482	1.8
$\mathcal{S}_{CR}$	2.1	1.8	7	29	530	1
$\mathcal{S}_{PI}$	40	33	30	0	698	3.8

Table 2
Registration results of test 6, with different similarity measures for 30 tests. Column *Mean* is mean RMS, *Median* is median RMS, $\#F$ is the number of registration with RMS greater than $3\,mm$, $\#I$ is number of registration which improve the starting position. *Iter* is the mean number of iterations needed to converge and T represents the time needed for one iteration.

6.3 Comments

Tests 1, 2, 3 and 4 show that DRR generation by application of a 2D affine transformation on a precomputed DRR is accurate (RMS error is lower than $0.5mm$ for a sampling step of $1°$, less than 1/256 grey level difference) and fast (lower than 20 milliseconds). This allows us to perform optimization of eq. (1) with a sufficient number of iterations to be accurate within clinically acceptable time constraints (entire procedure takes about 30 seconds with $\mathcal{S}_{CR}$).

Test 5 shows that 2D/3D registration can be accurately performed and that all the 6 parameters of the rigid transformation are accurately retrieved (RMS lower than $1mm$ in 89% of the tests) with two 2D images.

Test 6 shows that patient pose estimation with portal images and CT scan is feasible: with correlation ratio or correlation coefficient, median RMS error is close to $1.7mm$.

7 Conclusion

This paper introduce an original method for patient pose estimation, using 2D portal images and a 3D CT scan. The method we proposed is based on the following points: it is fully 3D, it avoids segmentation, it uses several PI and pre-generated DRR, and it is an intensity-based registration procedure. The main contributions are (1) a geometrical transformation decomposition allowing the generation of any full resolution DRR without the usual time-consuming volume-rendering, (2) the use of the correlation ratio as similarity measure for PI/DRR images, and (3) the combination of several similarities between pairs of images. Experimental tests yield to good position estimations both for simulated and real PI (median RMS are lower 0.5 mm for simulated images and 1.7 mm for real PI).

One drawback of the method is the optimization procedure used for eq. (1): algorithm is sometimes trapped in local maxima. Another drawback could be the need for DRR (temporary) storage. However, with (compressed) DRR of 140KB, for 100 patients, we need about $225 \times 2 \times 100 \times 140 = 6.3$GB, which is easily tractable nowadays. Another difficulty could be the additional computational time required by the DRR pre-computation. However, DRR have to be generated only once and it can be done overnight.

This approach has many advantages. First, the geometrical transformation decomposition allows to use high quality DRR without increasing the computational time. The sets of DRR have to be computed only once *before* patient treatment. The quality of the DRR is also important, other methods usually require lower DRR quality in order to decrease the computational time. Poor quality PIs need to be compensated by high quality DRR so that the similarity measure indicates a higher value at the optimal point in the transformation space.

Our approach can be used with any number of PI (thanks to the unique joint histogram, updated for each pair of images) and with any type of portal image (by using different similarity measures). The overall procedure is fast, since it takes less than 30 seconds to complete on a common 1.7Ghz PC, without fine-tuning optimization.

The method is fully automatic, it requires no user intervention, but several initial parameters still need to be tuned. First, the choice of the similarity measure depends on the type of images. In our experiments, we used megavoltage portal images and observed that $\mathcal{S}_{CC}$ and $\mathcal{S}_{CR}$ lead to better results. In case of other types of control images, such as ultra-sounds [52], experiments conducted in different contexts [48] suggested that $\mathcal{S}_{MI}$ or χ^2 could be used. The discretization of the set of pre-computed projections also plays a role in the accuracy of the estimation. Experimental results show that 1° is a good tradeoff between precision and volume storage.

Further works are ongoing to improve the optimization procedure. The presented method is only valid for rigid body transformations, but we plan to include non-rigid deformations and organ displacements using the same principles (DRR generation and 2D transformations).

Acknowledgments

This work is supported by the *Région Rhône-Alpes* under the grant *AdéMo* ("Modélisation et suivi spatio-temporel pour le diagnostique et la thérapie"). The authors want also to thank all the physicians and physicists of the radiotherapy department of the *Léon Bérard* anti-cancer institute in Lyon for their fruitfull discussion and remarks. The authors also thanks Elekta company for their support.

A Notations

A.1 Projections

Perspective projections are parameterized by an *image plane* and an *optical center*. It is a affine transformation which projects a 3D point into the *image plane*, such that the image point is the intersection between the image plane and the line containing the initial 3D point and the optical center. Let $\boldsymbol{P}$ be a 3×4 perspective projection matrix, and $\boldsymbol{x} = \{x_1, x_2, x_3\} \in \mathbb{R}^3$. We denote $\dot{\boldsymbol{x}} = \{x_1, x_2, x_3, 1\}$ its homogeneous coordinates. The application of $\boldsymbol{P}$ leads to an image point $\dot{\boldsymbol{y}}$ such that $\dot{\boldsymbol{y}} = \boldsymbol{P}\dot{\boldsymbol{x}} = \{y_1, y_2, w\}$. The final 2D image point is given by $\boldsymbol{y} = \{\frac{y_1}{w}, \frac{y_2}{w}\} = \{u, v\}$. If $w = 0$, it means that $\boldsymbol{x}$ is in the focal plane and is projected to infinity.

In the following, we will use the notation $[\boldsymbol{A}|\boldsymbol{a}]$ with $\boldsymbol{A}$ a 3×3 matrix and $\boldsymbol{a}$ a column vector, where $[\boldsymbol{A}|\boldsymbol{a}][\boldsymbol{B}|\boldsymbol{b}] = [\boldsymbol{AB}|\boldsymbol{Ab} + \boldsymbol{a}]$. A perspective projection matrix $\boldsymbol{P}$ is decomposed into the following matrices:

$$\boldsymbol{P} = \boldsymbol{A}[\boldsymbol{G}|\boldsymbol{d}] \tag{A.1}$$

$$\boldsymbol{A} = \begin{bmatrix} f\kappa_u & 0 & u_0 \\ 0 & f\kappa_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{A.2}$$

$\boldsymbol{A}$ is the *intrinsic* parameters matrix, f is the *focal length*, (κ_u, κ_v) is the pixels size, (u_0, v_0) is the coordinate of the *principal point*. The 3×4 matrix $[\boldsymbol{G}|\boldsymbol{d}]$ is a rigid transformation matrix composed of a 3×3 rotation matrix $\boldsymbol{G}$ (orthogonal and $\det(\boldsymbol{G}) = 1$) and a translation vector $\boldsymbol{d}$. The optical center $\boldsymbol{c}$ of $\boldsymbol{P}$ is $\boldsymbol{c} = -\boldsymbol{G}^{-1}\boldsymbol{A}^{-1}\boldsymbol{d}$.

A.2 Rotations

Rotations are parameterized using the *rotation vector* paradigm. A rotation matrix $\boldsymbol{R}$ is characterized with a vector $\boldsymbol{r}$, such that $\boldsymbol{R}$ is the rotation of angle $\theta = \|\boldsymbol{r}\|$ around the unitary axis $\boldsymbol{n} = \frac{\boldsymbol{r}}{\|\boldsymbol{r}\|}$.

A.3 Initial and objective projections

The *initial projection* is denoted $\boldsymbol{P}_0$. We parameterize projections according to the isocenter $\boldsymbol{s}$ because this is the reference point used by most of radiotherapy devices: the center of the tumor is assumed to be the isocenter. The optical center is denoted $\boldsymbol{c}_0$ ($\boldsymbol{Id}$ is the identity matrix). We define $\boldsymbol{P}_0$ such that :

$$\boldsymbol{P}_0 = \boldsymbol{A}[\boldsymbol{G}|\boldsymbol{d}][\boldsymbol{Id}| - \boldsymbol{s}] = \boldsymbol{A}[\boldsymbol{G}|\boldsymbol{d} - \boldsymbol{Gs}] \tag{A.3}$$
$$\boldsymbol{c}_0 = -\boldsymbol{G}^{-1}\boldsymbol{d} + \boldsymbol{s} \tag{A.4}$$

The *objective projection* corresponds to the initial projection of a displaced scene. The scene displacement is a rigid transformation matrix $\boldsymbol{U} = [\boldsymbol{T}|\boldsymbol{t}]$ composed of a 3D rotation $\boldsymbol{T}$ and a translation $\boldsymbol{t}$. The rotation $\boldsymbol{T}$ is parameterized as a rotation around the isocenter. The objective projection is denoted $\boldsymbol{P}_1$, the optical center $\boldsymbol{c}_1$. Thus, we define $\boldsymbol{P}_1 = \boldsymbol{P}_0\boldsymbol{U}$:

$$\boldsymbol{P}_1 = \boldsymbol{P}_0\,[\boldsymbol{T}| - \boldsymbol{Ts} + \boldsymbol{s} + \boldsymbol{t}] \tag{A.5}$$
$$\boldsymbol{P}_1 = \boldsymbol{A}\,[\boldsymbol{G}|\boldsymbol{d} - \boldsymbol{Gs}]\,[\boldsymbol{T}| - \boldsymbol{Ts} + \boldsymbol{s} + \boldsymbol{t}] \tag{A.6}$$
$$\boldsymbol{P}_1 = \boldsymbol{A}\,[\boldsymbol{GT}|\boldsymbol{d} - \boldsymbol{Gs} - \boldsymbol{GTs} + \boldsymbol{Gs} + \boldsymbol{Gt}] \tag{A.7}$$
$$\boldsymbol{P}_1 = \boldsymbol{A}\,[\boldsymbol{GT}|\boldsymbol{d} - \boldsymbol{GTs} + \boldsymbol{Gt}] \tag{A.8}$$
$$\boldsymbol{c}_1 = -\boldsymbol{T}^{-1}\boldsymbol{G}^{-1}\boldsymbol{d} + \boldsymbol{T}^{-1}\boldsymbol{s} - \boldsymbol{T}^{-1}\boldsymbol{t} \tag{A.9}$$

B Corrective rotation

Consider the rotation $\boldsymbol{R}$ of center $\boldsymbol{c}_1$ which bring the optical axis of $\boldsymbol{P}_1$ into the line $(\boldsymbol{c}_1, \boldsymbol{s})$. We can compute this rotation by using the spherical coordinates (in $\boldsymbol{P}_1$ frame) of $\boldsymbol{b} = [\boldsymbol{GT}|\boldsymbol{d} + \boldsymbol{Gt}]\boldsymbol{s} = \boldsymbol{d} + \boldsymbol{Gt}$. It is then possible to define a projection $\boldsymbol{P}_2 = \boldsymbol{A}\,[\boldsymbol{R}|\boldsymbol{0}]\,[\boldsymbol{GT}|\boldsymbol{d} - \boldsymbol{GTs} + \boldsymbol{Gt}]$. We show that $\boldsymbol{c}_2 = \boldsymbol{c}_1$ ($\boldsymbol{0}$ is the nul vector):

$$\boldsymbol{P}_2 = \boldsymbol{A}\,[\boldsymbol{R}|\boldsymbol{0}]\,[\boldsymbol{GT}|\boldsymbol{d} - \boldsymbol{GTs} + \boldsymbol{Gt}] \tag{B.1}$$
$$\boldsymbol{P}_2 = \boldsymbol{A}\,[\boldsymbol{RGT}|\boldsymbol{Rd} - \boldsymbol{RGTs} + \boldsymbol{RGt}] \tag{B.2}$$
$$\boldsymbol{c}_2 = -\boldsymbol{T}^{-1}\boldsymbol{G}^{-1}\boldsymbol{R}^{-1}\boldsymbol{Rd} + \boldsymbol{T}^{-1}\boldsymbol{G}^{-1}\boldsymbol{R}^{-1}\boldsymbol{RGs} - \boldsymbol{T}^{-1}\boldsymbol{G}^{-1}\boldsymbol{R}^{-1}\boldsymbol{RGt} \tag{B.3}$$
$$\boldsymbol{c}_2 = -\boldsymbol{T}^{-1}\boldsymbol{G}^{-1}\boldsymbol{d} + \boldsymbol{T}^{-1}\boldsymbol{s} - \boldsymbol{T}^{-1}\boldsymbol{t} \tag{B.4}$$
$$\boldsymbol{c}_2 = \boldsymbol{c}_1 \tag{B.5}$$

C Rectification matrix

It is known that if two projections $\boldsymbol{P}_a = [\boldsymbol{A}|\boldsymbol{a}]$ and $\boldsymbol{P}_b = [\boldsymbol{B}|\boldsymbol{b}]$ have the same optical center and the same intrinsic parameters, it exists a 3×3 rectification matrix $\boldsymbol{F} = \boldsymbol{B}\boldsymbol{A}^{-1}$, such that $\boldsymbol{F}$ map the image plane of $\boldsymbol{P}_a$ onto the image plane of $\boldsymbol{P}_b$, see [53]. Thus, we can compute the rectification matrix $\boldsymbol{F}$ between $\boldsymbol{P}_2 = [\boldsymbol{ARGT}|\boldsymbol{ARd} - \boldsymbol{ARGs} + \boldsymbol{ARGt}]$ and $\boldsymbol{P}_1 = [\boldsymbol{AGT}|\boldsymbol{Ad} - \boldsymbol{AGs} + \boldsymbol{AGt}]$:

$$\boldsymbol{F} = (\boldsymbol{AGT})^{-1}(\boldsymbol{ARGT})^{-1} \tag{C.1}$$

Thus, we have: $\boldsymbol{P}_1 = \boldsymbol{F}\boldsymbol{P}_2$.

D Projection $\boldsymbol{P}_3$

Now, we consider the projection $\boldsymbol{P}_3$ defined such that:

- $\boldsymbol{P}_3$ has the same orientation as $\boldsymbol{P}_2$ (rotation is $\boldsymbol{RGT}$),
- the distance between the optical center $\boldsymbol{c}_3$ and the isocenter $\boldsymbol{s}$ is equal to the distance between $\boldsymbol{c}_0$ and $\boldsymbol{s}$.

Thus, we can define κ such that:

$$(\boldsymbol{c}_3 - \boldsymbol{s}) = \kappa(\boldsymbol{c}_2 - \boldsymbol{s}) \qquad \kappa = \frac{\|\boldsymbol{c}_0 - \boldsymbol{s}\|}{\|\boldsymbol{c}_2 - \boldsymbol{s}\|} \tag{D.1}$$

By definition $\|\boldsymbol{c}_0 - \boldsymbol{s}\| = \|\boldsymbol{c}_3 - \boldsymbol{s}\|$ thus there exists a rotation $\boldsymbol{Q}$ of center $\boldsymbol{s}$ such that $\boldsymbol{c}_3 = \boldsymbol{Q}(\boldsymbol{c}_0 - \boldsymbol{s}) + \boldsymbol{s}$. Hence, projection $\boldsymbol{P}_3$ is $\boldsymbol{P}_3 = \boldsymbol{A}[\boldsymbol{RGT}| - \boldsymbol{RGT}(\boldsymbol{c}_3)]$. Thus, we have :

$$P_3 = A[RGT| - RGT(c_3)] \tag{D.2}$$
$$P_3 = A[RGT| - RGT(Q(c_0 - s) + s)] \tag{D.3}$$
$$P_3 = A[RGT| - RGT(Q(-G^{-1}d + s - s) + s)] \tag{D.4}$$
$$P_3 = A[RGT|RGTQG^{-1}d - RGTs] \tag{D.5}$$

E Scaling matrix

Let $\boldsymbol{P}_a = \boldsymbol{A}[\boldsymbol{M}|\boldsymbol{a}]$ and $\boldsymbol{P}_b = \boldsymbol{A}[\boldsymbol{M}|\boldsymbol{b}]$ be two projections with the same orientation matrix $\boldsymbol{M}$, the same intrinsic parameters $\boldsymbol{A}$ and different optical centers ($\boldsymbol{c}_a = -\boldsymbol{M}^{-1}\boldsymbol{a}$ and $\boldsymbol{c}_b = -\boldsymbol{M}^{-1}\boldsymbol{b}$). It is not possible to find a matrix which map the image plane of $\boldsymbol{P}_a$ onto the image plane of $\boldsymbol{P}_b$ because of the perspective projection: in-plane movements depend on the distance between the scene and the optical center. Hence, because the distances ($\|\boldsymbol{s} - \boldsymbol{c}_a\|$ and $\|\boldsymbol{s} - \boldsymbol{c}_b\|$) are larger than the distance between the two optical centers ($\|\boldsymbol{c}_a - \boldsymbol{c}_b\|$), we assume that the apparent motion look like a scaling (when the scene move closer to the optical center, it looks bigger and conversely). Thus, we build a 3×3 scaling matrix $\boldsymbol{K}$ of center the principal point $\boldsymbol{l} = (u_0, v_0)$ and where the scaling factor κ is the ratio of the distances between optical centers and the isocenter $\boldsymbol{s}$:

$$\boldsymbol{K} = -\boldsymbol{S}_\kappa \boldsymbol{l} + \boldsymbol{l} \qquad \boldsymbol{S}_\kappa = \begin{bmatrix} \kappa & 0 \\ 0 & \kappa \end{bmatrix} \qquad \kappa = \frac{\|\boldsymbol{c}_b - \boldsymbol{s}\|}{\|\boldsymbol{c}_a - \boldsymbol{s}\|} \tag{E.1}$$

Hence, $\boldsymbol{K}\boldsymbol{P}_b \approx \boldsymbol{P}_a$. Numerical experimentations in section 6 show that this approximation is sufficient for our purpose.

F In-plane/out-of-plane rotation decomposition

Let $\boldsymbol{M}$ be a rotation matrix. We want to decompose $\boldsymbol{M}$ into two rotation matrices $\boldsymbol{M} = \boldsymbol{C}\boldsymbol{H}$ such that $\boldsymbol{H}$ is a rotation around an axe which is included in the plane Oxy and $\boldsymbol{C}$ is a rotation around Oz axe. Such decomposition can be done by expressing the rotation with quaternions. The rotation $\boldsymbol{M}$ corresponds to the rotation vector $\boldsymbol{r}_0 = \theta_0 \boldsymbol{n}_0$ ($\boldsymbol{n}_0$ is an unit vector) and the following quaternion $\boldsymbol{q}_0 = \{q_0^0, q_0^1, q_0^2, q_0^3\}$:

$$\boldsymbol{M} = \boldsymbol{q}_0 = \begin{cases} \theta \\ \boldsymbol{r}_0 = \theta \boldsymbol{n}_0 \end{cases} \rightarrow \begin{cases} q_0^0 = \cos(\theta/2) \\ q_0^1 = \sin(\theta/2)\, n_0^1 \\ q_0^2 = \sin(\theta/2)\, n_0^2 \\ q_0^3 = \sin(\theta/2)\, n_0^3 \end{cases} \tag{F.1}$$

where v^i denote the i^{th} component of a vector $\boldsymbol{v}$. We denote $q_0^0, q_0^1, q_0^2, q_0^3$ by w, x, y, z.

Our goal is to write $\boldsymbol{q}_0 = \boldsymbol{q}_2\boldsymbol{q}_1$ or $\boldsymbol{M} = \boldsymbol{C}\boldsymbol{H}$, where $\boldsymbol{C}$ is a rotation matrix around the Oz axis and $\boldsymbol{H}$ is out-of-plane according to Oz. We have:

$$\boldsymbol{q}_1 = \begin{cases} \phi \\ \boldsymbol{r}_1 = \phi \boldsymbol{n}_1 = \begin{pmatrix} \alpha \\ \beta \\ 0 \end{pmatrix} \end{cases} \rightarrow \begin{cases} q_1^0 = \cos(\phi/2) \\ q_1^1 = \sin(\phi/2)\, \alpha \\ q_1^2 = \sin(\phi/2)\, \beta \\ q_1^3 = 0 \end{cases} \tag{F.2}$$

We denote q_1^0, q_1^1, q_1^2 by a, b, c. Then, we define $\boldsymbol{q}_2$:

$$\boldsymbol{q}_2 = \begin{cases} \gamma \\ \boldsymbol{r}_2 = \gamma \boldsymbol{n}_2 = \begin{pmatrix} 0 \\ 0 \\ \gamma \end{pmatrix} \end{cases} \rightarrow \begin{cases} q_2^0 = \quad \cos(\gamma/2) \\ q_2^1 = 0 \\ q_2^2 = 0 \\ q_2^3 = \quad \sin(\gamma/2)\, \gamma \end{cases} \tag{F.3}$$

We denote q_2^0, q_2^3 by d, e. So, the equality $\boldsymbol{q}_0 = \boldsymbol{q}_2\boldsymbol{q}_1$ may be written:

$$\begin{cases} q_0^0 = q_2^0 q_1^0 \\ q_0^1 = q_2^0 q_1^1 - q_2^3 q_1^2 \\ q_0^2 = q_2^0 q_1^2 + q_2^3 q_1^0 \\ q_0^3 = q_2^3 q_1^0 \end{cases} \text{or} \begin{cases} w = da \\ x = db - ec \\ y = dc + eb \\ z = ea \end{cases} \tag{F.4}$$

In (F.4), $\boldsymbol{q}_0$ is given, the parameters we want to compute in $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are α, β and γ. By considering the first and the last equalities, we have:

$$a = \frac{w}{\cos(\gamma/2)} \quad \text{so} \quad z = \sin(\gamma/2)\frac{w}{\cos(\gamma/2)} \tag{F.5}$$

$$\text{we get} \quad \gamma = 2\arctan(z/w) \tag{F.6}$$

Then, we can derive the other unknowns:

$$b = \frac{(xw + yz)w}{(w^2 + z^2)d} \qquad c = -\frac{(xz - yw)w}{(w^2 + z^2)d} \tag{F.7}$$

$$\phi = 2\arccos(w/d) \qquad \alpha = b\frac{\phi}{\sin(\phi/2)} \qquad \beta = c\frac{\phi}{\sin(\phi/2)} \tag{F.8}$$

Then, by using equations (F.1)(F.2)(F.3) and α, β, γ given by equations (F.6)(F.8), we compute the quaternions $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$, and thus the rotation matrices $\boldsymbol{H}$ and $\boldsymbol{C}$ which decompose $\boldsymbol{M}$ in a out-of-plane and a in-plane part.

Now, we apply this decomposition to the projection $\boldsymbol{P}_3$ previously defined. We consider $\boldsymbol{RGT}$ the rotation part of $\boldsymbol{P}_3$, and write $\boldsymbol{RGTG}^{-1} = \boldsymbol{CH}$. Thus, we have $\boldsymbol{RGT} = \boldsymbol{CHG}$ and $\boldsymbol{P}_3$ can be written $\boldsymbol{P}_3 = \boldsymbol{A}[\boldsymbol{CHG}|\boldsymbol{Cd} - \boldsymbol{CHGs}]$, with $\boldsymbol{H}$ an out-of-plane rotation and $\boldsymbol{C}$ an in-plane rotation. By construction, $\boldsymbol{C}$ is a rotation of angle γ around the optical axis of $\boldsymbol{P}_0$. Thus, it is possible to write $\boldsymbol{P}_3 = \boldsymbol{C}'\boldsymbol{P}_h$, with $\boldsymbol{P}_h = \boldsymbol{A}[\boldsymbol{HG}|\boldsymbol{d} - \boldsymbol{HGs}]$ and $\boldsymbol{C}'$ being the in-plane rotation matrix of center the principal point $\boldsymbol{l} = (u_0, v_0)$ and with the same angle γ than $\boldsymbol{C}$.

References

[1] V. Rudat, M. Flentje, D. Oetzel, M. Menke, W. Schlegel, and M. Wannenmacher. Influence of the positioning error on 3D conformal dose distributions during fractionated radiotherapy. *Radiother Oncol*, 33(1):56–63, 1994.

[2] E.M. Soffen, G.E. Hanks, C.C Hwang, and J.C. Chu. Conformal static field therapy for low volume low grade prostate cancer with rigid immobilization. *Int. J. Radiat. Oncol. Biol. Phys.*, 20(1):141–6, 1991.

[3] S.A. Rosenthal, M. Roach, B.J. Goldsmith, and et al. Immobilization improves the reproductibility of patient positioning during six-field conformal radiation therapy for prostate carcinoma. *Int. J. Radiat. Oncol. Biol. Phys.*, 27(4):921–6, 1993.

[4] PB. Greer, T.M. Mortensen, and C.C. Jose. Comparison of two methods for anterior-posterior isocenter localization in pelvic radiotherapy using

electronic portal imaging. *Int. J. Radiat. Oncol. Biol. Phys.*, 41(5):1193–9, 1998.

[5] W. De Neve, F. Van den Heuvel, M. De Beukeleer, and et al. et al. Routine clinical on-line portal imaging followed by immediate field adjustment using a tele-controlled patient couch. *Radiother Oncol*, 24(1):45–54, 1992.

[6] U. Mock, K. Dieckmann, U. Wolff, T.H. Knocke, and R. Potter. Portal imaging based definition of the planning target volume during pelvic irradiation for gynecological malignancies. *Int. J. Radiat. Oncol. Biol. Phys.*, 45(1):227–232, 1999.

[7] J.C. Stroom, M.J. Olofsen-van Acht, S. Quint, and et al. et al. On-line set-up corrections during radiotherapy of patients with gynecologic tumors. *Int. J. Radiat. Oncol. Biol. Phys.*, 46(2):499–506, 2000.

[8] J. Kinzie, G. Hanks, C. McLean, and S. Kramer. Patterns of care study: hodgkin's disease relapse rates and adequacy of portals. *Cancer*, 52:2223–2226, 1983.

[9] J.E. White, T. Chen, J. McCracken, and et al. et al. The influence of radiation therapy quality control on survival, response and sites of relapse in oat cell carcinoma of the lung. *Cancer*, 50:1084–90, 1982.

[10] C. Carrie, S. Hoffstetter, F. Gomez, and et al. Impact of targeting deviations on outcome in medulloblastoma: study of the french society of pediatric oncology (SFOP). *Int. J. Radiat. Oncol. Biol. Phys.*, 45(2):435–9, September 1999.

[11] J.E. Marks and A.G. Haus. The effect of immobilisation on localisation error in the radiotherapy of head and neck cancer. *Clin Radiol*, 27(2):175–7, 1976.

[12] J.E. Marks, J.M. Bedwinek, F. Lee, J.A. Purdy, and C.A. Perez. Dose-response analysis for nasopharyngeal carcinoma: an historical perspective. *Cancer*, 50(6):1042–50, 1982.

[13] C.M. Nutting, V.S. Khoo, V. Walker, H. McNair, C. Beardmore, A. Norman, and D.P. Dearnaley. A randomized study of the use of a customized immobilization system in the treatment of prostate cancer with conformal radiotherapy. *Radiother. Oncol.*, 54(1):1–9, 2000.

[14] P.Y. Song, M. Washington, F. Vaida, and et al. et al. A comparison of four patient immobilization devices in the treatment of prostate cancer patients with three dimensional conformal radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 34(1):213–219, 1996.

[15] M. van Herk and H. Meertens. A matrix ionisation chamber imaging device for on-line patient setup verification during radiotherapy. *Radiother Oncol*, 11(4):369–78, 1988.

[16] A.L. Boyer, L. Antonuk, A. Fenster, and et al. et al. A review of electronic portal imaging devices (EPIDs). *Med Phys*, 19(1):1–16, 1992.

[17] P. Munro. Portal imaging technology: past, present, and future. *Sem. Radiat. Oncol.*, 5(2):115–133, 1995.

[18] A. Bel, O. Petrascu, I. Van de Vondel, L. Coppens, N. Linthout, D. Verellen, and G. Storme. A computerized remote table control for

fast on-line patient repositioning: implementation and clinical feasibility. *Med Phys*, 27(2):354–8, 2000.

[19] D.C. Hornick, D.W Litzenberg, K.L Lam, J.M. Balter, J. Hetrick, and R.K. Ten Haken. A tilt and roll device for automated correction of rotational setup errors. *Med Phys*, 25(9):1739–40, 1998.

[20] G.W. Sherouse, K. Novins, and E.L. Chaney. Computation of digitally reconstructed radiographs for use in radiotherapy treatment design. *Int. J. Radiat. Oncol. Biol. Phys.*, 18(3):651–8, 1990.

[21] C.W. Hurkmans, P. Remeijer, J.V. Lebesque, and B.J. Mijnheer. Set-up verification using portal imaging; review of current clinical practive. *Radiotherapy and Oncology*, 58:105–120, 2001.

[22] 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):1665–79, 1991.

[23] J. Van de Steene, F. Van den Heuvel, A. Bel, and et al. et al. Electronic portal imaging with on-line correction of setup error in thoracic irradiation: clinical evaluation. *Int. J. Radiat. Oncol. Biol. Phys.*, 40(4):967–976, 1998.

[24] M.A. Hunt, T.E. Schultheiss, G.E. Desobry, M. Hakki, and G.E. Hanks. An evaluation of setup uncertainties for patients treated to pelvic sites. *Int J Radiat Oncol Biol Phy*, 32(1):227–33, 1995.

[25] J.M. Michalski, J.W. Wong, W.R. Bosch, and et al. et al. An evaluation of two methods of anatomical alignment of radiotherapy portal images. *Int. J. Radiat. Oncol. Biol. Phys.*, 27(5):1199–206, 1993.

[26] J. Hanley, G.S. Mageras, J. Sun, and G.J. Kutcher. The effects of out-of-plane rotations on two dimensional portal image registration in conformal radiotherapy of the prostate. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1331–1343, 1995.

[27] P. Remeijer, E. Geerlof, L. Ploeger, K. Gilhuijs, M. van Herk, and J.V. Lebesque. 3-D portal image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.*, 46(5):1281–90, 2000.

[28] J. Hanley, M.A. Lumley, G.S. Mageras, and et al. et al. Measurement of patient positioning errors in three-dimensional conformal radiotherapy of the prostate. *Int. J. Radiat. Oncol. Biol. Phys.*, 37(2):435–44, 1997.

[29] K.L. Lam, R.K. Ten Haken, D.L. McShan, and A.F. Thornton Jr. Automated determination of patient setup errors in radiation therapy using spherical radio-opaque markers. *Med Phys*, 20(4):1145–52, 1993.

[30] K.P. Gall, L.J. Verhey, and M. Wagner. Computer-assisted positioning of radiotherapy patients using implanted radiopaque fiducials. *Med. Phys.*, 20(4):1153–9, 1993.

[31] S. Lavallée and R. Szeliski. Recovering the Position and Orientation of Free-Form Objects from Image Contours Using 3D Distance Maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(4):378–390, April 1995.

[32] L. Brunie, S. Lavallee, J. Troccaz, P. Cinquin, and M. Bolla. Pre- and intra-irradiation multimodal image registration: principles and first experiments. *Radiother Oncol*, 29(3):244–52, 1993.

[33] K.G. Gilhuijs, P.J. 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):389–99, 1996.

[34] K.G. Gilhuijs, K. Drukker, A. Touw, P.J. van de Ven, and M. van Herk. Interactive three dimensional inspection of patient setup in radiation therapy using digital portal images and computed tomography data. *Int. J. Radiat. Oncol. Biol. Phys.*, 34(4):873–85, 1996.

[35] M.J. Murphy. An automatic six-degree-of-freedom image registration algorithm for image-guided frameless stereotaxic radiosurgery. *Medical Physics*, 24(6):857–866, 1997.

[36] L. Dong and A.L. Boyer. An image correlation procedure for digitally reconstructed radiographs and electronic portal images. *Int. J. Radiat. Oncol. Biol. Phys.*, 33(5):1053–60, 1995.

[37] D.H. Hristov and B.G. Fallone. A grey-level image alignment algorithm for registration of portal images and digitally reconstructed radiographs. *Med Phys*, 23(1):75–84, 1996.

[38] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality Image Registration by Maximization of Mutual Information. *IEEE Transactions on Medical Imaging*, 16(2):187–198, 1997.

[39] W.M. Wells, P.A. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-Modal Volume Registration by Maximization of Mutual Information. *Medical Image Analysis*, 1(1):35–51, 1996.

[40] A. Roche, G. Malandain, X. Pennec, and N. Ayache. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration. In A. Colschester W.M. Wells and S. Delp, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI'98*, volume 1496, pages 1115–1124. Springer Verlag, Lecture Notes in Computer Science, 1998.

[41] D. Sarrut and S. Clippe. Patient positioning in radiotherapy by registration of 2D portal to 3D CT images by a contend-based research with similarity measures. In *Computer Assisted Radiology and Surgery (CARS)*, pages 707–712, San Fransisco, USA, 2000. Elsevier Science.

[42] D. Plattard, G. Champleboux, P. Vassal, J. Troccaz, and M. Bolla. EPID for patient positioning in radiotherapy: calibration and image matching in the entroPID system. In H.U. Lemke, editor, *Cars'99*, pages 265–9. Elsevier science, 1999.

[43] R. Bansal, L. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, and J. Duncan. A minimax entropy registration framework for patient setup verification in radiotherapy. *Computer Aided Surgery*, 4(6):287–304, 1999.

[44] D. Sarrut and S. Clippe. Geometrical transformation approximation for 2D/3D intensity-based registration of portal images and CT scan. In W.J. Niessen and M. A. Viergever, editors, *Medical Image Computing*

and Computer-Assisted Intervention MICCAI'2001, volume 2208, pages 532–540, Utrecht (Netherlands), 2001. Springer Verlag, Lecture Notes in Computer Science.

[45] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing.* Cambridge University Press, second edition, 1992.

[46] G.P. Penney, J. Weese, J.A. Little, P. Desmedt, D.L. Hill, and D.J. Hawkes. A comparison of similarity measures for use in 2D-3D medical image registration. *IEEE Transactions on Medical Imaging*, 17(4):586–595, 1998.

[47] A.E. Lujan, J.M. Balter, and R.K. Ten Haken. Determination of rotations in three dimensions using two-dimensional portal image registration. *Med Phys*, 25(5):703–8, 1998.

[48] D. Sarrut. *Recalage multimodal et plate-forme d'imagerie médicale à accès distant.* PhD thesis, Université Lumière Lyon 2, 2000.

[49] A. Roche, G. Malandain, and N. Ayache. Unifying Maximum Likelihood Approaches in Medical Image Registration. *Special Issue on 3D Imaging of the International Journal of Computer Vision of Imaging Systems and Technology*, 2000.

[50] P.A. Viola and W.M. Wells. Alignement by maximization of Mutual Information. *International Journal of Computer Vision*, 24(2):137–154, 1997.

[51] M.J. Clarkson, D. Rueckert, D.L.G. Hill, and D.J. Hawkes. Registration of multiple video images to pre-operative ct for image guided surgery. In K.M. Hanson, editor, *Medical Imaging: Image Processing*, volume 3661, pages 14–23, San Diego, California 1999, 1999. SPIE.

[52] T.R. Willoughby, M.S. Weinhous, and P.A. Kupelian. Daily transabdominal ultrasound localization of the prostate gland with the BAT system: 2705 alignments from 100 patients. In *42nd american society for therapeutic radiology and oncology*, volume 48 (3) of *Int. J. Radiat. Oncol. Biol. Phys.*, page p308, oct 2000. supplement.

[53] A. Fusiello, E. Trucco, and A. Verri. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12(1):16–22, 2000.