Published in final edited form as:

Med Phys. 2009 August ; 36(8): 3420–3428.

Stochastic rank correlation: A robust merit function for 2D/3D registration of image data obtained at different energies

Wolfgang Birkfellner^{a)}

Center for Biomedical Engineering and Physics, Medical University Vienna, Waehringer Guertel 18-20 AKH 4L, A-1090 Vienna, Austria

Markus Stock

Department of Radiotherapy, Division of Medical Radiation Physics, Medical University Vienna, Waehringer Guertel 18-20 AKH 4L, A-1090 Vienna, Austria

Michael Figl, Christelle Gendrin, Johann Hummel, and Shuo Dong

Center for Biomedical Engineering and Physics, Medical University Vienna, Waehringer Guertel 18-20 AKH 4L, A-1090 Vienna, Austria

Joachim Kettenbach

Department of Radiology, Medical University Vienna, Waehringer Guertel 18-20 AKH 4L, A-1090 Vienna, Austria

Dietmar Georg

Department of Radiotherapy, Division of Medical Radiation Physics, Medical University Vienna, Waehringer Guertel 18-20 AKH 4L, A-1090 Vienna, Austria

Helmar Bergmann

Center for Biomedical Engineering and Physics, Medical University Vienna, Waehringer Guertel 18-20 AKH 4L, A-1090 Vienna, Austria

Abstract

In this article, the authors evaluate a merit function for 2D/3D registration called stochastic rank correlation (SRC). SRC is characterized by the fact that differences in image intensity do not influence the registration result; it therefore combines the numerical advantages of cross correlation (CC)-type merit functions with the flexibility of mutual-information-type merit functions. The basic idea is that registration is achieved on a random subset of the image, which allows for an efficient computation of Spearman's rank correlation coefficient. This measure is, by nature, invariant to monotonic intensity transforms in the images under comparison, which renders it an ideal solution for intramodal images acquired at different energy levels as encountered in intrafractional kV imaging in image-guided radiotherapy. Initial evaluation was undertaken using a 2D/3D registration reference image dataset of a cadaver spine. Even with no radiometric calibration, SRC shows a significant improvement in robustness and stability compared to CC. Pattern intensity, another merit function that was evaluated for comparison, gave rather poor results due to its limited convergence range. The time required for SRC with 5% image content compares well to the other merit functions; increasing the image content does not significantly influence the algorithm accuracy. The authors conclude that SRC is a promising measure for 2D/3D registration in IGRT and image-guided therapy in general.

I. INTRODUCTION

Exact estimation of patient and target position is a crucial issue in modern fractionated radiation therapy since accurate dose delivery over the entire period of treatment is mandatory for therapeutic success. The current interest in image-guided radiation therapy (IGRT) has been stimulated by the advent of novel radiation therapy facilities with integrated kilovoltage (kV)-imaging units^{1,2} and the broad availability of semiconductor electronic image detectors.³⁻⁵ In the context of daily patient position verification, 2D/3D registration⁶⁻⁸ was always considered as the technology of choice since its introduction to medical imaging. Here, a 2D projection image (usually an x ray) is iteratively compared to a perspective volume rendering derived from volumetric 3D imaging. In the case of patient positioning in radiation therapy, typically megavoltage (MV) electronic portal images (EPIs) or, more recently, a kV x ray and digitally rendered radiographs (DRRs) derived from computed tomography (CT) volume data are used. Additional information from other imaging sources can also be integrated in such a volume.⁹ The latest generation of linear accelerators (LINACs) provides x-ray imaging equipment mounted together with the treatment system for intrafractional planar x-ray and cone beam CT imaging.¹

As far as 2D/3D registration in general is concerned, the main issues remain algorithm robustness and execution speed of the iterative registration process. A number of possible merit functions, rendering algorithms for the generation of DRRs and optimization strategies were proposed. These include the following.

- Various merit functions like cross correlation (CC),^{6,7} local normalized correlation (LNC),¹⁰ pattern intensity (PI),⁷ gradient-based methods, Zernike moment decompositions,¹¹ and possibly the most frequent due to its popularity for 3D/3D multimodal volume registration, mutual information (MI) and its variants.^{12-14,7,15-17}
- Fast DRR rendering algorithms, such as shear-warp algorithms,¹⁸ light field rendering,¹⁹ ray casting using subregions of the images,²⁰ estimation of histograms from stochastic ray casting,²¹ graphics processor unit (GPU) accelerated methods,^{22,23} and splat-rendering approaches.²⁴
- Decoupling²⁵ or pseudodecoupling²⁶ of degrees of freedom which accelerates the optimization process by utilizing special projection geometries.

The properties of merit functions need further discussion since this article deals with a novel type of similarity measure that has advantageous properties for the 2D/3D registration of images with substantial intensity discrepancies. In general, one can categorize merit functions as follows:

- Merit functions for intramodal registration: Images of similar histogram content are registered by establishing a relationship between pixel intensities. This can be done by computing a correlation measure between the corresponding pixel locations by, for instance, CC, or by minimizing the entropy in a difference image—PI being one of the best known measures of this kind.
- Merit functions for intermodal registration: Here, the assessment of pixel similarities is replaced by the likelihood for a pixel position being occupied. Different intensities or even color-coded information can be used as the base data for the histograms (considered a probability distribution in this case), and the comparison of histograms yields a merit function. MI, the measure that gives the degree of mutual statistical dependence of two random variables, is the merit function of choice in this case.

Usually, in 2D/3D registration, intramodal images are registered, although efforts to register magnetic resonance (MR) volume images and DRRs exist²⁸ or are subject to research for slice-to-volume registration.^{16,17,29,30} With regard to the daily verification of patient position in IGRT, we are dealing with a quasi-intramodal matching problem; image intensity and histogram content are highly disparate due to the different energies and detector characteristics used for CT kV imaging from either the LINAC-integrated cone beam CT or a diagnostic CT image from another machine; both may provide the volume data for iterative DRR rendering in the 2D/3D registration process. The x-ray imaging from the LINAC-integrated imager in the treatment room provides the reference image. However, the physical process—the attenuation of ionizing radiation by tissue—is the same in both cases.

The robustness of the 2D/3D registration process is governed by the global shape of the merit function. A wide convergence radius (sometimes also referred to as “capture range” in literature) with a smooth concave shape is desirable for optimum convergence even in the presence of large initial misregistrations. Too many local or false global minima increase the risk of registration failure, whereas sharp, spike-shaped merit functions may give accurate results in the case of convergence but require the use of time-intensive global optimization algorithms.¹⁷

In a recent study on slice-to-volume registration of FluoroCT and CT imaging,¹⁷ we found CC to be an excellent merit function for 2D/3D matching if an exact linear relationship exists between pixels in the images under comparison. This was also confirmed by other recent studies^{10,16} but careful radiometric calibration of image histograms when rendering the DRR is necessary to achieve good results.¹⁰ The reason for this constraint of strict linearity is clear from the very definition of CC and is further addressed in Sec. II. MI, on the other hand, appears suitable for the problem of 2D/3D matching of imaging data obtained at different energies but was of questionable stability due to the sparse histogram population resulting in false global minima at the boundaries of the optimization range.^{7,17} In this article, we introduce stochastic rank correlation (SRC), a novel merit function for 2D/3D registration of images which exhibit a continuous monotonic relationship between pixel intensities. In particular, we have tested SRC on a reference dataset of a cadaver spine.²⁷

II. MATERIALS AND METHODS

II.A. Cross correlation, rank correlation, and stochastic rank correlation

The CC-merit function is, in fact, the Pearson product moment correlation coefficient that measures the level of linear dependence between two variables measured on the same subject. It is defined as

$$\mathcal{M}_{CC} = \frac{\sum_{ij} (I_{ijBase} - \bar{I}_{Base}) (I_{ijMatch} - \bar{I}_{Match})}{\sqrt{\sum_{ij} (I_{ijBase} - \bar{I}_{Base})^2 \sum_{ij} (I_{ijMatch} - \bar{I}_{Match})^2}}, \quad (1)$$

where I_{ij} denotes the pixel grayscale values of the base image and the match image, respectively. $\bar{I}$ is the average grayscale pixel value of the image. It is a well-known fact from basic statistics that the product moment correlation coefficient determines an exact linear relationship between the grayscale pixel values in I_{Base} and I_{Match} . Another property of $\mathcal{M}_{CC}$ is evident from Eq. (1). Variances are derived as deviations from the average value $\bar{I}$, which is a measure prone to error if outliers are present. In the case of medical image registration, high intensity artifacts jeopardize registration success. Careful radiometric calibration, for instance, by means of histogram equalization, can somewhat compensate for this problem.¹⁰

For variables measured on an ordinal level, statistics provides Spearman's rank correlation (RC) coefficient, a measure for assessing arbitrary monotonic functions. It does not require a linear relationship, and the variables need not be measured on a metric scale. For the purpose of medical image registration, a merit function based on RC can be defined as follows:

- Copy all pixel grayscale values γ , together with their coordinates $\vec{x}=(x,y)^T$ in a $3 \times N$ matrix $\mathbf{R}$, where N is the number of pixels in the images I_{Base} —the reference x ray and I_{Match} —the DRR. Two “raw score” matrices, $\mathbf{R}_{\text{Base}}$ and $\mathbf{R}_{\text{Match}}$, are the result.
- Rearrange $\mathbf{R}_{\text{Base}}$ and $\mathbf{R}_{\text{Match}}$ in such a manner that the values γ are ordered; the result is two “ranked score” lists, $\mathbf{R}'_{\text{Match}}$ and $\mathbf{R}'_{\text{Base}}$. The first row of matrix $\mathbf{R}'$ can be thought of as a “longitudinal histogram” whose number of classes equals the gray level depth of the image. The associated pixel coordinates $\vec{x}$ are in the second and third rows of $\mathbf{R}'$. The number of entries with identical values of γ would represent the height of the column in a conventional histogram. In practice, this ordering process uses a hash-table procedure so that the pixel coordinates for the gray values are preserved. In our implementation, we have used introsort, the sorting procedure implemented in the STL of the g++ compiler. The process is of complexity order $\mathcal{O}(N \log N)$ in the worst case. The rank ρ of the coordinate is the column index in this sorted array.
- Strictly speaking, pixels with identical grayscale values γ will have tied ranks—within a class in the longitudinal histogram, the rank is therefore arbitrary and a result of the sorting process; statistics demand that in such a case, the average rank $\bar{\rho}$ for all identical values of γ is computed, which is simply given as $\bar{\rho}(\gamma) = \frac{1}{2} * (\rho(\gamma)_{\min} + \rho(\gamma)_{\max})$. This process is of the order of $\mathcal{O}(N)$. We have found this step to be of little relevance in practice. Using the current index of the sorted coordinates as rank ρ should work as well.
- Discard the first row with grayscale values γ and replace it with the average rank $\bar{\rho}(\gamma)$. What remains is two submatrices $\mathbf{r}$ for the coordinate values $\vec{x}$ of I_{Base} and I_{Match} , which are ordered according to their associated (and now abandoned) value γ . In a simplified version of the algorithm where the computation of average ranks $\bar{\rho}(\gamma)$ is omitted, the rank ρ of the pixel coordinates is given by the index in the matrices $\mathbf{R}'$.
- For each pixel coordinate $\vec{x}_{\text{Base}_n}$, $n \in \{1, \dots, N\}$ in $\mathbf{r}_{\text{Base}}$, the rank ρ_{Match_m} , $m \in \{1, \dots, N\}$ of the corresponding pixel position in $\mathbf{r}_{\text{Match}}$ is sought, and the squared rank difference $\Delta \rho_n^2 = (\rho_{\text{Base}_n} - \rho_{\text{Match}_m})^2$ is recorded for each pixel. In the worst case, this is a process of complexity order $\mathcal{O}(N^2)$.
- Compute the merit function $\mathcal{M}_{\text{RC}}^*$ based on the rank correlation coefficient as

$$\mathcal{M}_{\text{RC}}^* = 1 - \frac{6 \sum_{n=1}^N \Delta \rho_n^2}{N(N^2 - 1)}. \quad (2)$$

Note that $\mathcal{M}_{\text{RC}}^*$ is defined in such a manner that totally uncorrelated images would deliver a maximum value, and perfectly correlated images give a minimum so that a convex

optimization problem is given. For obtaining a concave minimization problem, one may define

$$\mathcal{M}_{\text{RC}} = \frac{6 \sum_{n=1}^N \Delta \rho_n^2}{N(N^2 - 1)}. \quad (3)$$

- Minimize iteratively for $\mathcal{M}_{\text{RC}}$ by producing new DRRs with different parameters.

The advantages of using RC rather than CC are evident.

- A linear relationship between pixel intensities is not necessary; therefore radiometric calibration can be avoided. In fact, pixel intensities are discarded at an early stage of merit function evaluation.
- Mean pixel intensity is not used for assessment of relative deviations in pixel intensities; therefore, RC is expected to be a more robust merit function if high intensity artifacts are present. In general, rank correlation is considered a measure robust against outliers.

The disadvantage of RC is also obvious. The processes of raw score matrix sorting and comparison of ranks is of considerable complexity and requires a large amount of computing time.

The basic idea of SRC is to obtain an estimate of the raw score lists, $\mathbf{R}_{\text{Base}}$ and $\mathbf{R}_{\text{Match}}$, by sampling the two matrices at randomly selected but identical pixel positions in I_{Base} and I_{Match} . In practice, one determines a mask of pixel positions prior to algorithm initialization for both images. This mask is computed from a uniform distribution using a random number generator. It is of utmost importance to emphasize that the random mask is constant during the registration process and identical for both images; hereby, the rank correlation coefficient is always computed from the image intensities at the same pixel location in $\mathbf{R}_{\text{Base}}$ and $\mathbf{R}_{\text{Match}}$. The effect of this operation is that the number of pixels N can be reduced significantly. The performance of such a merit function

$$\mathcal{M}_{\text{SRC}} = \frac{6 \sum_{n=1}^N \Delta \rho_n^2}{\mathcal{N}(\mathcal{N}^2 - 1)}, \quad \mathcal{N} \cdots \text{number of randomly selected pixels}, \quad \mathcal{N} \ll N \quad (4)$$

is assessed in this article on a reference 2D/3D DRR/x-ray dataset of a cadaver spine.²⁷ Figure 1 gives an overview of the 2D/3D registration process using stochastic rank correlation.

II.B. Algorithm evaluation

Evaluation of the algorithm took place with x-ray/DRR datasets of an available reference dataset for 2D/3D registration of a spine segment.²⁷ The phantom consists of a spine segment with five vertebral bodies. Image data are provided via <http://www.isi.uu.nl/Research/Databases/GS/>. For the evaluation, we used the CT data and the lateral fluoroscopic image from the example data. The CT was acquired using a 16-row multislice scanner (MX8000) at a tube voltage of 80 kV and the x ray was taken with a Integris BV5000 (both from Philips Medical Systems, Best, The Netherlands) at approximately 60 kV according to Ref. 27. Our main interest for applying SRC stems from image-guided radiotherapy with LINAC-mounted kV-imaging systems such as the Elekta Synergy system. Since the kV-imaging unit of such a system is rigidly attached to the LINAC, only one projection can be acquired at a time—therefore we confined ourselves to 2D/3D registration using only a single projection. The coordinate system used for registration in our

implementation is fixed to the imaging plane, as opposed to the local coordinate system used in Ref. 27. For this reason, the gold standard position data were transformed to the convention as used in our setup. Furthermore, the CT scan was interpolated from its original resolution of $0.30664 \times 0.30664 \times 0.492 \text{ mm}^3$ to a homogeneous resolution of 0.5 mm^3 using ANALYZE AVW 9.0. The reference x ray was also interpolated from its original resolution of 0.6271 mm^2 to a resolution of 0.5 mm^2 and cropped to a 256×256 matrix. The distance of the x-ray focus and the imaging plane is 802.37 mm. Finally, both the x-ray image and the CT were saved in 8 bit depth, which is also the original depth of the provided CT scan and the x-ray image. Figure 2 shows the reference x ray together with six DRRs rendered from the CT data. The splat-rendering algorithm used here²⁴ requires a minimum threshold for perspective volume rendering—six different intensity thresholds of 40, 50, 60, 70, 80, and 90 (out of 256) gray values were used for the evaluation of the algorithm with changing histogram content. Since the CT scan does not feature 12 bit depth, translating these thresholds to Hounsfield units (HU) is not possible; when transforming the 8 bit gray values to HU, one can estimate the minimum thresholds for DRR rendering to be in the range of -384 – 416 HU .

Using the known registration parameters, we evaluated the capture range and the accuracy achievable by performing 145 registrations with varying starting positions. Following a simple power analysis, the sample number should be sufficient for giving a significant result. The correct parameters were distorted by a random number. The range of displacements was chosen to be $\pm 5^\circ$ and $\pm 10 \text{ mm}$. These displacements should also reflect the typical range of motion encountered in radiotherapy, for instance, in lung irradiation.⁵ After registration, the position of the maximum voxel coordinate $\vec{x} = (186, 186, 256)^T$ using the found registration parameters was computed, and the result was compared to the position of that voxel as given by the gold standard parameters. The absolute difference was recorded as the target registration error (TRE). We did not select the results in terms of successful and unsuccessful registrations; only those results where the optimization did not converge after a fixed set of iterations were removed. This problem, however, only occurred for one of the merit functions under consideration.

The merit functions under comparison were CC, CC after histogram equalization of the DRR and the x ray, and SRC using 3%, 5%, and 10%. Furthermore, SRC was also computed for 50% and 66% of image content—these results are, however, to be considered out of competition since the time required for computing the merit function becomes unacceptable. The results were recorded, and a t-test for independent samples was carried out for the TRE using SPSS 17.0. The resulting *p* values are given in Sec. III.

PI was also included. PI is known to give excellent results for 2D/3D registration,⁷ but suffers from a very limited convergence range, which was also confirmed by our own experience.^{17,23,31,32} The merit function operates on a difference image I_{diff} and is given as

$$\mathcal{M}_{\text{PI}} = \sum_{i,j} \sum_{d^2 \leq r^2} \frac{\sigma^2}{\sigma^2 + (I_{\text{diff}}(i,j) - I_{\text{diff}}(v,w))^2}, \quad (5)$$

$$d^2 = (i - v)^2 + (j - w)^2.$$

Since PI operates on a difference image, the threshold for DRR rendering also appears critical since the disorder of the difference image is evaluated as a registration measure. For our method of evaluation, we used PI together with histogram equalization. It turned out that the algorithm converged only for approximately 50% of all trials, with very varying accuracies when using internal parameters of $\sigma = 10$ and $r = 3$ or $r = 5$. A detailed description of the algorithm can be found in Ref. 7.

MI, the other merit function that is of known usefulness when dealing with multimodal 3D/3D registration problems, is a natural choice for images with different histogram content. However, it is also known for its problematic behavior in the case of 2D/3D registration due to the sparse population of the joint histogram.^{7,17} MI is highly dependent on the chosen implementation.⁵ In our case, a simple approach where the probability distribution was derived from a histogram with 32 bins was also of limited success. Therefore, MI was finally excluded from the experiments, which are strictly aiming at quasi-intramodal 2D/3D registration. When registering x-ray data to volume data from another modality but CT, the case is of course different.

II.C. Implementation

The algorithm was implemented using GCC 4.1.2 on Fedora 8 on a Dell Precision 490 personal computer. User interfaces were programed using the Qt 3.3 toolkit (Trolltech, Norway). Basic image processing functionality was added using the `ANALYZEAVW` 9.0 library (BIR, Mayo Clinic, Rochester/MN).³³ In principle, the software setup was adopted from the setup used in Ref. 17. DRRs were generated using the wobbled splatting algorithm as described in Ref. 24. The Downhill-Simplex algorithm used for optimization in the iterative registration process used was taken from the GSL package (Free Software Foundation, Boston, MA). Registration takes place in such manner that after convergence of the algorithm, the result is used as a starting point for another trial. After three subsequent registrations, the result was considered to be final.

III. RESULTS

In a first trial, the stability and accuracy of the implementation was tested. For this purpose, a DRR was rendered at a threshold of 70, and the result was fed into the program as the x ray. The known initial position used for rendering the DRR was distorted by random positions of 5° and 10 mm in each coordinate. Fifteen registrations were carried out. Such an evaluation on a phantom image gives only little information on the actual performance of an algorithm, but it gives a figure of the quality of the implementation including the optimization algorithm. For this ideal configuration, no significant differences in the maximum TRE were found. All merit functions converged properly. PI suffers from its small range of convergence, which affects the overall algorithm accuracy. The detailed results can be found in Table I.

The crucial evaluation is the registration of the spine phantom dataset; the results can be found in Table II. Here, SRC using 3%–10% of image content always outperformed CC and CC_{HEq}. For CC, results were really bad, especially when using another threshold than 70, which gave the most homogeneous results for CC, CC_{HEq}, and SRC. Maximum TRE for CC_{HEq} ranged from 14.3±6.8 mm when using DRR threshold 70 to 18.9±8.6 mm at threshold 90. The results for SRC (excluding the registration efforts when using 50% and 66% image content, which did not show any improvement compared to SRC using 5% or 10% image content) ranged from 12.6±5.7 mm (SRC 5%, DRR threshold 70) to 14.7±9.0 mm (SRC 3%, DRR threshold 40). For all trials, the best results were provided by SRC. It is the very nature of the volume-driven splat-rendering routine to perform better for higher thresholds,²⁴ which results in faster registration if a higher threshold is chosen. Detailed results for the time requirements are given in Table II; however, SRC for 3% and 5% image contents required usually 20–120 s, which is similar to the performance of CC and CC_{HEq}. SRC with 50% or 66% of image content took up to 10 min and is therefore only of academic interest.

PI showed only limited stability and was terminated in almost 50% of all trials. Beyond that, a tendency to converge to local minima was evident, which spoiled the registration results.

Maximum TRE for those cases where convergence was achieved ranged from 15.0 ± 7.3 mm (PI with $r=3$, $\sigma=10$ at DRR threshold 60 to 17.0 ± 8.3 mm for the same parameters at DRR threshold 80. The results are given in Table III.

IV. DISCUSSION AND CONCLUSIONS

In this article we present a merit function suitable for 2D/3D registration of intramodal data sets acquired with x-ray-based equipment. The specific strength of SRC is the independence of pixel intensity differences in the image histograms as long as they can be modeled as monotonic transforms. Thus, radiometric calibration is unnecessary. However, SRC also has limitations by design. As it is the case with most intensity-based merit functions, depth of the images is crucial; if only a few actual values of possible gray values γ are available, the number of tied ranks will be considerable, resulting in an inconclusive minimum for the merit function. It is also necessary to use images with lots of content—an x ray showing a small structure surrounded by air will be definitely result in only a few random pixels with different ranks. Furthermore, noise may affect the performance of SRC, as it is the case with most other intensity-based merit functions. In this article, we have not examined the influence of noise on SRC. However, the evaluation was carried out with image data acquired on standard imaging equipment, and the splat-rendering routine used²⁴ produces images stricken with Gaussian noise by definition, which is reduced by applying a spatial low pass filter prior to registration of the DRR to the x ray. We therefore assume that the noise level encountered here represents a real-life scenario.

Our results demonstrate that, in the given setup, SRC outperforms both classical CC and CC with histogram equalization of the images under comparison; in the case of SRC, any kind of fine tuning in histogram space is not necessary by definition. Therefore, SRC is ideally suited for registration tasks involving image data obtained at different energies, and the experimental data on the spine phantom support this conclusion. However, a patient study using actual IGRT equipment is a pending task. Furthermore, the robustness of SRC in the presence of high intensity artifacts is an open question. It is in the very nature of the rank correlation coefficient that it is more robust against outliers; the extent of this theoretical advantage is to be examined.

The use of stochastic sampled image information for merit function design is not new; in Ref. 21, a similar approach was used to derive a MI-based merit function while reducing time requirements for computing the DRR. We have not explored this approach since wobbled splatting,²⁴ as a volume-driven method, does not necessarily benefit from rendering just a few pixels. On the other hand, it is more efficient compared to raycasting, and it offers huge potential for GPU-based implementation.²² A direct comparison of a partial raycasting approach as proposed in Ref. 21 and splat rendering is, however, also an open task.

One possibly astonishing fact is the range of error encountered in the experiments and the fact that PI does not perform as good as one might expect from the literature. The reason for this behavior lies in the fact that the initial misalignment in the experimental series is larger than the convergence range typically used in most evaluations. The very limited convergence range of PI, however, is also known and documented.^{23,31,17,37} In Ref. 27, a convergence range of only a few millimeters is considered to give safe results. With tumor motion tracking being our objective, it makes little sense to distinguish between failed and successful registrations—only those results where the optimization algorithm did not converge were excluded; the range of motions possible used in this evaluation is realistic, and beyond that, the TRE computed gives a worst case scenario. PI, a merit function that gives a very sharp minimum, here is a victim of this setup. It gives very good results for

small displacements, but outside the convergence range, it does not even provide a gradient that might finally point in the appropriate direction. Future research work, for instance, on improved optimization schemes as proposed in Ref. 31, is, however, necessary to improve convergence behavior and algorithm accuracy. Since it is not possible to compute analytical derivatives of a rank correlation type of merit function, options for alternative optimization schemes are somewhat narrowed down.

Another problem we encountered is the little availability of reference datasets and reference algorithm implementations. The spine dataset²⁷ is a good initial effort in this direction but suffers from limitations such as its reduced intensity range and the small size of the specimen; a pending task for the community is therefore the development of more representative phantoms, which are to be provided together with gold standard data and reference implementations of the registration algorithms. Given the methodological nature of this article, an exhaustive evaluation with more merit functions is therefore also a pending task.

The other question is the intended use of inter- and intrafractional imaging in combination with 2D/3D registration in IGRT. The most obvious application—exact positioning of the patient prior to treatment—is widely discussed and well known both in image-guided therapy^{6-8,20,21,25,16,17,28} and radiation oncology.^{2-5,19,10,26} For this purpose, several x-ray images from various positions or even a cone beam CT can be made, the latter providing the known advantages of a full 3D volume. However, numerous tumor locations require intrafractional position monitoring due to organ motion during irradiation. It appears as if the advent of on-board imaging units has further stimulated this field of research, which spans approaches from motion prediction^{34,35} to permanent x-ray monitoring of implanted fiducial markers³⁶ and registration.³⁷⁻⁴¹ 2D/3D registration using only one projection⁴² as in our experimental series appears to be a promising method of choice to achieve this goal of minimizing dose delivered outside the target volume despite the additional irradiation from the low-energy x-ray system. With this perspective in mind, we have designed our experiments; here, the sensitivity and reliability of motion detection is of greater interest than the absolute accuracy (which can be achieved by performing a 3D/3D registration of the cone beam CT and the diagnostic CT prior to irradiation anyhow). The future will show whether image-based motion correction at a high rate can compete with motion sensing using miniature active⁴³ and passive⁴⁴ tracking solutions. To achieve this goal, a massive increase in registration speed, which may be achievable using GPU-based parallel approaches is necessary.

Acknowledgments

ANALYZE^{AVW} and ^{AVW} were provided courtesy of Dr. R. A. Robb, BIR, Mayo Clinic, Rochester/MN. The authors are grateful to Mary Mc Allister, MD, Johns Hopkins University Hospital, Baltimore, for manuscript assistance. The spine dataset and ground truth for the 2D-3D registration used in this work was provided by the Image Sciences Institute, University Medical Center Utrecht, The Netherlands. This work was supported by the Eurasia-Pacific Uninet Foundation and the Austrian Science Foundation FWF Grant Nos. P19931 and L503.

References

1. Dawson LA, Jaffray DA. Advances in image-guided radiation therapy. *J. Clin. Oncol.* 2007; 25(8): 938–946. [PubMed: 17350942]
2. Elshaikh M, Ljungman M, Haken R, Ten, Lichter AS. Advances in radiation oncology. *Annu. Rev. Med.* 2006; 57:19–31. [PubMed: 16409134]
3. Jans HS, Syme AM, Rathee S, Fallone BG. 3D interfractional patient position verification using 2D-3D registration of orthogonal images. *Med. Phys.* 2006; 33(5):1420–1439. [PubMed: 16752578]

4. Meyer J, Richter A, Baier K, Wilbert J, Guckenberger M, Flentje M. Tracking moving objects with megavoltage portal imaging: A feasibility study. *Med. Phys.* 2006; 33(5):1275–1280. [PubMed: 16752562]
5. Künzler T, Grezdo J, Bogner J, Birkfellner W, Georg D. Registration of DRRs and portal images for verification of stereotactic body radiotherapy: A feasibility study in lung cancer treatment. *Phys. Med. Biol.* 2007; 52(8):2157–2170. [PubMed: 17404461]
6. Lemieux L, Jagoe R, Fish DR, Kitchen ND, Thomas DGT. A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs. *Med. Phys.* 1994; 21(11): 1749–1760. [PubMed: 7891637]
7. Penney GP, Weese J, Little JA, Desmedt P, Hill DLG, Hawkes DJ. A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Trans. Med. Imaging.* 1998; 17(4): 586–595. [PubMed: 9845314]
8. Turgeon GA, Lehmann G, Guiraudon G, Drangova M, Holdsworth D, Peters T. 2D-3D registration of coronary angiograms for cardiac procedure planning and guidance. *Med. Phys.* 2005; 32(12): 3737–3749. [PubMed: 16475773]
9. Balter JM, Kessler ML. Imaging and alignment for image-guided radiation therapy. *J. Clin. Oncol.* 2007; 25(8):931–937. [PubMed: 17350941]
10. Khamene A, Bloch P, Wein W, Svatos M, Sauer F. Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. *Med. Image Anal.* 2006; 10:96–112. [PubMed: 16150629]
11. Dong S, Kettenbach J, Hinterleitner I, Bergmann H, Birkfellner W. The Zernike expansion—an example of a merit function for 2D/3D registration based on orthogonal functions. *MICCAI.* 2008; 11(2):964–971. [PubMed: 18982698]
12. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging.* 1997; 16(2):187–198. [PubMed: 9101328]
13. Wells WM III, Viola P, Atsumi H, Nakajima S, Kikinis R. Multi-modal volume registration by maximization of mutual information. *Med. Image Anal.* 1996; 1(1):35–51. [PubMed: 9873920]
14. Studholme C, Hill DL, Hawkes DJ. Automated 3D registration of MR and CT images of the head. *Med. Image Anal.* 1996; 1(2):163–175. [PubMed: 9873927]
15. Pluim JP, Maintz JB, Viergever MA. Mutual-information-based registration of medical images: A survey. *IEEE Trans. Med. Imaging.* 2003; 22(8):986–1004. [PubMed: 12906253]
16. Fei B, Duerk JL, Boll DT, Lewin JS, Wilson DL. Slice-to-volume registration and its potential application to interventional MRI-guided radio-frequency thermal ablation of prostate cancer. *IEEE Trans. Med. Imaging.* 2003; 22(4):515–525. [PubMed: 12774897]
17. Birkfellner W, Figl M, Kettenbach J, Hummel J, Homolka P, Scherthaner R, Nau T, Bergmann H. Rigid 2D/3D slice-to-volume registration and its application on fluoroscopic CT images. *Med. Phys.* 2007; 34(1):246–255. [PubMed: 17278510]
18. Lacroute, P.; Levoy, M. Fast volume rendering using a shear-warp factorization of the viewing transformation. *Proceedings of SIGGRAPH '94*; 1994; New York: ACM; p. 451–458.
19. Russakoff DB, Rohlfing T, Mori K, Rueckert D, Ho A, Adler JR Jr. Maurer CR Jr. Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration. *IEEE Trans. Med. Imaging.* 2005; 24(11):1441–1454. [PubMed: 16279081]
20. Knaan D, Joskowicz L. Effective intensity-based 2D/3D rigid registration between fluoroscopic x-ray and CT. *MICCAI.* 2003; 6(1):351–358.
21. Zöllei, L.; Grimson, E.; Norbash, A.; Wells, W. 2D-3D rigid registration of x-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators. *Proceedings of IEEE CVPR*; 2001; New York: IEEE Computer Society;
22. Spoerk J, Bergmann H, Wanschitz F, Dong S, Birkfellner W. Fast DRR splat rendering using common consumer graphics hardware. *Med. Phys.* 2007; 34(11):4302–4308. [PubMed: 18072495]
23. Kubias A, Deinzer F, Feldmann T, Paulus D, Schreiber B, Brunner T. 2D/3D image registration on the GPU. *Pattern Recognition and Image Analysis.* 2008; 18(3):381–389.

24. Birkfellner W, Seemann R, Figl M, Hummel J, Ede C, Homolka P, Yang X, Niederer P, Bergmann H. Wobbled splatting—a fast perspective volume rendering method for simulation of x-ray images from CT. *Phys. Med. Biol.* 2005; 50(9):N73–N84. [PubMed: 15843725]
25. Birkfellner W, Wirth J, Burgstaller W, Baumann B, Staedele H, Hammer B, Gellrich NC, Jacob AL, Regazzoni P, Messmer P. A faster method for 2D/3D medical image registration. *Phys. Med. Biol.* 2003; 48(16):2665–2679. [PubMed: 12974581]
26. Andrews DW, Bednarz G, Evans JJ, Downes B. A review of 3 current radiosurgery systems. *Surg. Neurol.* 2006; 66(6):559–564. [PubMed: 17145309]
27. van de Kraats EB, Penney GP, Tomazevic D, van Walsum T, Niessen WJ. Standardized evaluation methodology for 2-D-3-D registration. *IEEE Trans. Med. Imaging.* 2005; 24(9):1177–1189. [PubMed: 16156355]
28. Tomazevic D, Likar B, Pernus F. 3-D/2-D registration by integrating 2-D information in 3-D. *IEEE Trans. Med. Imaging.* 2006; 25(1):17–27. [PubMed: 16398411]
29. Frühwald L, Kettenbach J, Figl M, Hummel J, Bergmann H, Birkfellner W. A comparative study on manual and automatic slice-to-volume registration of CT images. *Eur. Radiol.* (in press).
30. Archip N, Statli P, Morrison P, Jolesz F, Warfield SK, Silverman S. Non-rigid registration of pre-procedural MR images with intra-procedural unenhanced CT-images for improved targeting of tumors during liver radiofrequency ablations. *MICCAI.* 2007; 10(2):969–977. [PubMed: 18044662]
31. Dey J, Napel S. Targeted 2D/3D registration using ray normalization and a hybrid optimizer. *Med. Phys.* 2006; 33(12):4730–4738. [PubMed: 17278825]
32. Russakoff, DB.; Rohlfing, T.; Ho, A.; Kim, DH.; Shahidi, R.; Adler, JR., Jr.; Maurer, CR, Jr.. Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data. *Proceedings of the second International Workshop on Biomedical Image Registration; 2003; Philadelphia, PA: Springer; p. 151-160.* Paper No. LNCS 2717
33. Robb RA, Hanson DP, Karwoski RA, Larson AG, Workman EL, Stacy MC. Analyze: A comprehensive, operator-interactive software package for multidimensional medical image display and analysis. *Comput. Med. Imaging Graph.* 1989; 13(6):433–454. [PubMed: 2688869]
34. Tewatia D, Zhang T, Tome W, Paliwal B, Metha M. Clinical implementation of target tracking by breathing synchronized delivery. *Med. Phys.* 2006; 33(11):4330–4336. [PubMed: 17153412]
35. Sharp GC, Jiang SB, Shimizu S, Shirato H. Prediction of respiratory tumour motion for real-time image-guided radiotherapy. *Phys. Med. Biol.* 2004; 49(3):425–440. [PubMed: 15012011]
36. Tang X, Sharp GC, Jiang SB. Fluoroscopic tracking of multiple implanted fiducial markers using multiple object tracking. *Phys. Med. Biol.* 2007; 52:4081–4098. [PubMed: 17664596]
37. Rohlfing T, Denzler J, Grassl C, Russakoff DB, Maurer CR Jr. Markerless real-time 3-D target region tracking by motion backprojection from projection images. *IEEE Trans. Med. Imaging.* 2005; 24(11):1455–1468. [PubMed: 16279082]
38. Zeng R, Fessler JA, Balter JM. Estimating 3-D respiratory motion from orbiting views by tomographic image registration. *IEEE Trans. Med. Imaging.* 2007; 26(2):153–163. [PubMed: 17304730]
39. Lehmann GC, Holdsworth DW, Drangova M. Angle-independent measure of motion for image-based gating in 3D coronary angiography. *Med. Phys.* 2006; 33(5):1311–1320. [PubMed: 16752566]
40. Shimizu S, Shirato H, Ogura S, Akita-Dosaka H, Kitamura K, Nishioka T, Kagei K, Nishimura M, Miyasaka K. Detection of lung tumor movement in real-time tumor-tracking radiotherapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2001; 51(2):304–310. [PubMed: 11567803]
41. Shirato H, Shimizu S, Kitamura K, Onimaru R. Organ motion in image-guided radiotherapy: Lessons from real-time tumor tracking radiotherapy. *Int. J. Clin. Oncol.* 2007; 12(1):8–16. [PubMed: 17380435]
42. Suh Y, Dieterich S, Keall PJ. Geometric uncertainty of 2D projection imaging in monitoring 3D tumor motion. *Phys. Med. Biol.* 2007; 52:3439–3454. [PubMed: 17664553]
43. Hummel J, Figl M, Kollmann C, Bergmann H, Birkfellner W. Evaluation of a miniature electromagnetic position tracker. *Med. Phys.* 2002; 29(10):2205–2212. [PubMed: 12408292]

44. Kupelian P, Willoughby T, Mahadevan A, Djemil T, Weinstein G, Jani S, Enke C, Solberg T, Flores N, Liu D, Beyer D, Levine L. Multi-institutional clinical experience with the Calypso System in localization and continuous, real-time monitoring of the prostate gland during external radiotherapy. *Int. J. Radiat. Oncol., Biol., Phys.* 2007; 67(4):1088–1098. [PubMed: 17187940]

FIG. 1.

The principle of 2D/3D registration of two images of different histogram content using a stochastic rank correlation merit function. Two significant properties of this process are apparent—only a small subsample of pixels is used for generation of raw score lists $\mathbf{R}$, and rather than pixel intensity itself, the position (or rank) of corresponding pixels in the ordered submatrix $\mathbf{r}$ is used for computing the approximated rank correlation coefficient.

FIG. 2.

The original x ray (right) from the reference dataset and the corresponding registered DRRs rendered with different thresholds (40–90). The images are given in 8 bit depth; the thresholds therefore refer to the minimum intensity used for rendering. For all thresholds, a series of 145 registration runs with random starting positions was carried out, and the difference in the result compared to the known gold standard registration parameters was recorded.

Table 1

Registration results for an ideal simulated DRR; these data give an idea of the accuracy to be expected from the reference spine dataset, the projection geometry used, and ideal circumstances given the implementation of used for the experiments. As expected, CC and SRC perform similar under such ideal conditions. A range of possible displacements from the known projection parameters of 5° and 10 mm was used. The TRE was computed for the maximum voxel coordinate available in the volume.

Result	Merit function					
	CC	SRC 3%	SRC 5%	SRC 10%	SRC 50%	PI r=3, $\sigma=10$
TRE (mm)	6.9±4.5	7.5±1.7	8.0±4.4	6.8±3.3	7.9±4.1	12.5±6.8
					6.8±3.7	12.3±6.8

TABLE II

Results of 145 registrations of an x ray of the Utrecht cadaver spine phantom to a DRR computed from the volume scan of the spine phantom. Different minimum thresholds for DRR rendering were chosen (column 2) for the 8 bit images under comparison. Each of the known registration parameters was displaced by a random number in the range of 5° and 10 mm. Merit functions used were CC, cross correlation applied to the images after histogram equalization (CC_{HEq}), and SRC derived from 3%, 5%, 10%, 50% and 66% image contents. A sample TRE was computed for the maximum voxel coordinates in the CT-volume available. The best results in each line are printed in boldface. Due to the considerable computing expenses, SRC with 50% and 66% image contents was not considered to be eligible for an optimum result. However, registration accuracy apparently does not increase considerably when using such large amounts of image data. SRC was compared to CC_{HEq} by applying an unpaired t-test to TRE data *p* values below 0.05 were considered to be significant and marked with a single asterisk. *p* values below 0.01 are considered highly significant and were marked with two asterisks.

Min. DRR threshold	Result	CC	CC _{HEq}	Merit function				
				SRC 3%	SRC 5%	SRC 10%	SRC 50%	SRC 66%
40	TRE (mm)	19.6±9.9	17.3±10.1	14.7±9.0	12.7±6.6	14.6±6.4	14.2±6.3	13.8±5.7
	<i>p</i> vs (CC _{HEq})	–	1	0.02*	0.000**	0.009**	–	–
	Time (s)	126±12	110±8	122±11	119±9	128±10	401±34	635±71
50	TRE (mm)	23.2±21.2	18.0±9.3	14.5±8.1	12.9±6.2	14.3±5.9	14.5±7.8	13.3±6.4
	<i>p</i> vs (CC _{HEq})	–	1	0.01*	0.000**	0.000**	–	–
	Time (s)	119±12	106±8	115±10	117±10	122±8	404±48	622±55
60	TRE (mm)	16.3±11.3	16.0±7.5	13.7±7.7	13.3±6.7	14.3±6.7	15.2±9.0	14.1±6.8
	<i>p</i> vs (CC _{HEq})	–	1	0.011*	0.001**	0.053	–	–
	Time (s)	103±11	88±8	93±9	98±7	101±7	413±48	609±52
70	TRE (mm)	14.3±6.2	14.3±6.8	13.0±6.7	12.6±5.7	14.0±6.5	13.9±6.5	14.0±6.0
	<i>p</i> vs (CC _{HEq})	–	1	0.106	0.02*	0.67	–	–
	Time (s)	58±7	46±4	52±5	49±4	55±2	336±40	563±66
80	TRE (mm)	29.3±24	16.2±8.9	14.5±7.5	12.7±6.2	13.7±5.7	12.7±5.1	13.1±4.8
	<i>p</i> vs (CC _{HEq})	–	1	0.08	0.000**	0.006**	–	–
	Time (s)	31±3	26±2	20±2	28±2	34±2	322±35	519±45
90	TRE (mm)	30.3±25.1	18.9±8.6	14.5±7.5	13.9±6.4	14.3±6.6	12.3±5.7	13.6±6.3
	<i>p</i> vs (CC _{HEq})	–	1	0.000**	0.000**	0.000**	–	–
	Time (s)	20±2	18±1	20±2	20±2	26±2	303±20	520±50

TABLE III

Results of 200 registrations using PI of an x ray of the Utrecht cadaver spine phantom to a DRR computed from the volume scan of the spine phantom. Again, the correct registration parameters were randomly distorted by 5° and 10 mm in each coordinate. The algorithm converged only in approximately 50% of all cases. Only those results were listed here. We conclude that for the given range of possible displacements, convergence behavior of PI clearly hampers the results. For this reason, it was also not compared to the other methods.

Min DRR threshold	Result	Merit function	
		$PI_{HEQ} (r=3, \sigma=10)$	$PI_{HEQ} (r=5, \sigma=10)$
40	TRE (mm)	15.5 ± 6.8	15.4 ± 6.1
	Success rate	103/200	112/200
	Time (s)	65 ± 10	75 ± 11
50	TRE (mm)	16.9 ± 7.2	16.2 ± 7.4
	Success rate	90/200	102/200
	Time (s)	63 ± 9	75 ± 10
60	TRE (mm)	15.0 ± 7.3	15.6 ± 6.4
	Success rate	98/200	101/200
	Time (s)	52 ± 7	62 ± 8
70	TRE (mm)	15.2 ± 6.9	15.7 ± 7.5
	Success rate	108/200	113/200
	Time (s)	32 ± 4	42 ± 6
80	TRE (mm)	17.0 ± 8.3	16.6 ± 7.9
	Success rate	111/200	109/200
	Time (s)	22 ± 4	40 ± 8
90	TRE (mm)	16.8 ± 7.4	15.9 ± 7.5
	Success rate	115/200	114/200
	Time (s)	17 ± 2	32 ± 3