Published in final edited form as:

Phys Med Biol. 2010 October 7; 55(19): N465–N471. doi:10.1088/0031-9155/55/19/N01.

Efficient implementation of the rank correlation merit function for 2D/3D registration

M Figl, C Bloch, C Gendrin, C Weber, S A Pawiro, J Hummel, P Markelj, F Pernuš, H Bergmann, and W Birkfellner

Center for Medical Physics and Biomedical Engineering, Medical University Vienna, AKH 4 L, Währinger Gürtel 18-20, A-1090 Vienna, Austria and Laboratory of Imaging Technologies, Faculty of Electrical Engineering, University of Ljubljana, Trzaska 25, SI-1000 Ljubljana, Slovenia

Abstract

A growing number of clinical applications using 2D/3D registration have been presented recently. Usually, a digitally reconstructed radiograph is compared iteratively to an x-ray image of the known projection geometry until a match is achieved, thus providing six degrees of freedom of rigid motion which can be used for patient setup in image-guided radiation therapy or computer-assisted interventions. Recently, *stochastic rank correlation*, a merit function based on Spearman's rank correlation coefficient, was presented as a merit function especially suitable for 2D/3D registration. The advantage of this measure is its robustness against variations in image histogram content and its wide convergence range. The considerable computational expense of computing an ordered rank list is avoided here by comparing randomly chosen subsets of the DRR and reference x-ray. In this work, we show that it is possible to omit the sorting step and to compute the rank correlation coefficient of the full image content as fast as conventional merit functions. Our evaluation of a well-calibrated cadaver phantom also confirms that rank correlation-type merit functions give the most accurate results if large differences in the histogram content for the DRR and the x-ray image are present.

1. Introduction

2D/3D registration, the computation of six degrees of freedom (dof) of rigid body motion from an iterative comparison of projective images such as radiographs or portal images, has gained clinical acceptance as a non-invasive tool for patient setup and motion control in radiotherapy (Bednarz *et al* 2009; Cho *et al* 2008) and image-guided interventions in general. Real-time tumor motion control based on intra-fractional x-ray imaging (Cho *et al* 2008; Sharp *et al* 2004) is also a mid-term research objective in image-based 2D/3D registration; for intensity-based registration not requiring further invasive measures such as implantation of markers (Sharp *et al* 2004) or transponders (Sawant *et al* 2009), the robustness and convergence range of the merit function used for registration is a prerequisite since 2D/3D registration based on digitally rendered radiographs (DRRs) is indeed a complex registration task. On the other hand, it avoids the drawbacks of invasive methods, namely the trauma associated with such a procedure and further implications on follow-up imaging using magnetic resonance imaging (MRI) (Zhu *et al* 2009).

Numerous merit functions for 2D/3D registration were presented (Penney *et al* 1998); specific problems related to 2D/3D registration include the fact that mutual information—a

universal and well-known merit function for multi-modal 3D/3D registration—may prove unreliable in the case of 2D/3D registration (Penney *et al* 1998; Birkfellner *et al* 2009) due to the sparse population of the joint histogram used for the computation of Shannon's entropy. Furthermore, some merit functions known to give very exact results may show poor success rates due to a poor convergence behavior; most methods based on difference images and gradient images fall in this category. On the other hand, merit functions based on classical correlation measures show a good convergence behavior by avoiding strong local minima and featuring a wide convergence range. Among these are cross correlation, which by definition requires a strict linear relationship between image intensities in the DRR and the reference image (Birkfellner *et al* 2007) and local normalized cross correlation (Khamene *et al* 2006), the latter showing good performance but requiring a careful radiometric calibration of the imaging device. From our experience, one cannot easily favor one of these methods above the other since the problem of 2D/3D registration is not well defined, and therefore an arsenal of possible registration measures is required to solve a specific task.

In Birkfellner *et al* (2009), a merit function called *stochastic rank correlation* (SRC) based on Spearman's rank correlation coefficient was presented. Rank correlation requires a monotonic relationship between gray values and is, by definition, robust against outliers. It is therefore not suitable for multimodal image registration, but these properties nevertheless render this measure as an attractive alternative for 2D/3D registration.

- Beam hardening in tissue could be taken into account when rendering the DRR. When using SRC, this does not have to be compensated since the change in pixel intensity is monotonic. Therefore, the computational load of DRR rendering is reduced.
- Rendering thresholds introduced for efficient DRR computation (Birkfellner *et al* 2005) do not affect registration performance.
- The dynamic range of modern amorphous silicon x-ray detectors used in kV-imaging units for image-guided radiotherapy has not to be taken into consideration.
- Variations in image intensity of subsequent DRRs, which are encountered when employing graphics boards (Spoerk *et al* 2007), can be neglected.
- Variable histogram content in subsequent radiographs due to a different dose input on the digital detector can be compensated.

The drawback of rank correlation is the fact that all gray values encountered in the reference radiograph and the DRR have to be sorted, which has a deleterious effect on the iterative registration process. SRC compensates this by computing a static random mask valid for both images which selects a small percentage of image content. From this subsample of the images, the rank correlation coefficient is computed. In Birkfellner *et al* (2009), it is shown that a rank correlation coefficient computed from 3 to 5% of image content can provide significantly better results compared to classical cross correlation on image pairs with equalized histograms while the computational effort is similar. In this note, we present a variation of the SRC algorithm which allows for using full image content without significantly corrupting registration speed.

2. Materials and methods

The basic concept of SRC is described in Birkfellner *et al* (2009); in order to avoid the time-consuming sorting of pixel gray values using the introsort algorithm while avoiding random subsampling of the image, we utilize the fact that the number of pixels in a medical image is usually much higher than the intensity depth of the image. In 2D/3D registration, it is for

instance common to select an appropriate window for the source computed tomography (CT) scan used for generating the DRR and to scale the volume down to 8 bit, resulting in 256 levels of gray. Given this circumstance, the sorting problem with a performance of order $N \log(N)$ on average can be replaced by a simplified bucket sort. N is the number of pixels used. Here, a histogram of the DRR is computed; since all pixels belonging to the same histogram class have the so-called *tied rank* since they are equivalent, an average rank computed from the boundaries of the histogram bin is assigned to each bin and further sorting of gray values is not necessary. The average rank $\bar{\rho}_i$ for each histogram bin i —that is, the mean of ranks belonging to the same histogram class in the ordered list of pixel gray values—is computed by

$$\bar{\rho}_i = \frac{1}{2} (\rho_{i_{\min}} + \rho_{i_{\max}}), \quad (1)$$

where $\rho_{i_{\min}}$ and $\rho_{i_{\max}}$ are the minimum and maximum rank for the pixels belonging to one histogram bin. $\rho_{i_{\min}}$ is determined by the number of pixels that are in bins $0, \dots, (i-)$, whereas $\rho_{i_{\max}}$ is determined by adding the number of pixels in the histogram bin i to $\rho_{i_{\min}}$. If the number of histogram bins is equal to the intensity depth of the images used for registration, the process of sorting a hash table of gray values and associated pixel coordinates can be replaced by a histogram computation of order N . The computation of Spearman's rank correlation coefficient $\mathcal{M}_{\text{RC}}$ (Birkfellner *et al* 2009)—which is used as the merit function for optimization—from the tied ranks is straightforward:

$$\mathcal{M}_{\text{RC}} = \frac{6 \sum_{i=0}^I \Delta \rho_i^2}{N(N^2 - 1)}, \quad (2)$$

with I being the number of histogram bins, equivalent to image intensity depth, $\Delta \rho_i^2$ the $(\bar{\rho}_i(I_{\text{DRR}}) - \bar{\rho}_i(I_{\text{x-ray}}))^2$, N the number of pixels used. Equation (2) is defined in such a manner that it yields a minimum for an optimal match of the DRR I_{DRR} and the radiograph $I_{\text{x-ray}}$.

The algorithm was implemented on a standard Dell Precision 490 PC with 4 GB of RAM and an nVidia Quadro 7500 graphics card. DRRs were generated using a GPU-based splat rendering routine (Birkfellner *et al* 2005) described in Spoerk *et al* (2007). A simplex optimization algorithm from the GSL (Free Software Foundation, Boston/MA) was used for minimization. The implementation of the RC-based merit function $\mathcal{M}_{\text{RC}}$ was compared to merit functions based on cross correlation (CC), histogram-based normalized mutual information (MI) (Penney *et al* 1998), correlation ratio (CR) (Skerl *et al* 2008) and SRC using 5% image content (Birkfellner *et al* 2009). The image content percentage for SRC was chosen in accordance to the results from Birkfellner *et al* (2009), where it was shown that a higher percentage does not improve registration outcome.

The results were validated on a novel porcine reference dataset developed by our group. A fresh cadaver head of a minipig was implanted with seven spherical polytetrafluoroethylene (PTFE) markers of 10 mm diameter; the markers were attached to the skull of the pig by cutting a thread using standard Schanz-screws from a orthopaedic fixateur externe set. A radiolucent screw holding the spheres was inserted. The skull was scanned using a Philips Brilliance 64 multislice CT with a resolution of $0.566 \times 0.566 \times 0.4 \text{ mm}^3$ and a slice thickness of 0.8 mm. The volume was interpolated to a cubic voxel size of 1 mm^3 . An x-ray of the phantom was taken using the kV-imaging unit of an Elekta Synergy linear accelerator (see figure 1) with 10 mm aluminum spheres attached to the radiolucent screws. The x-ray detector was a Perkin-Elmer XRD amorphous silicon detector driven in a 2×2 binning

mode with a resolution of 0.8mm^2 . Since we are mainly interested in high-performance 2D/3D registration for organ motion compensation in radiotherapy and since most clinical scans are not acquired with the high resolution of the original CT dataset, we settled for 1mm^3 voxels; this also increases the performance of the algorithm due to the smaller data size. The gold standard registration was determined by a direct linear transform. The markers were not removed from the images; in additional experiments where the markers were excluded by masking the 2D image, we did not encounter a significant change in the results. Details on the phantom can be found in Pawiro *et al* (2010). An intrinsic advantage of this phantom is the large amount of soft tissue connected to the cadaver, which gives a realistic clinical situation even for registration of organs other than the head.

For evaluation of the merit functions, known gold standard registration parameters were randomly changed in such a manner that the initial target registration error (iTRE) as computed from 400 randomly chosen positions on the cadaver specimen ranged from 0 to 15 mm. For the monitoring of intra-fraction tumor motion, this is a reasonable range if the patient is precisely aligned and if the tumor target volume is selected as the region-of-interest for registration. The protocol is in accordance with other standard evaluations found in the literature (van de Kraats *et al* 2005). One hundred and fifty registrations were carried out for each merit function. We have recorded both the resulting projected TRE (mean projection distance mPD) and the average Euclidean distance of the resulting 3D positions for the reference locations in 3D, which gives the mean target registration error (mTRE). The minimum rendering threshold applied to the CT dataset varied from -868 to -243 HU.

3. Results

Table 1 gives the results for all registration runs as mPD and mTRE, both as mean and standard deviation, and median. In general, it shows that RC, SRC computed from 5% image content and CR give the best results, which are usually very close. However, RC tends to outperform the CR in cases where the image histogram content shows strong differences due to the rendering threshold used for DRR generation. The time required for the evaluation of a single merit function is dependent on the rendering threshold used. This is a property of the volume-driven splat rendering method (Birkfellner *et al* 2005) used for DRR generation. However, the average times necessary are very similar (0.12 s for CC, RC, CR and MI, and 0.07 s for SRC 5%—the latter being an exception since the runtime heavily depends on the percentage of the image content used for evaluation). Our optimization setup—a repeated optimization by using a simplex algorithm—requires a very similar number of iterations for convergence; therefore, the runtime for all merit functions excluding SRC 5% was very similar. From the data on time requirements for SRC presented in Birkfellner *et al* (2009), it can be extrapolated that SRC with 12% image data would on average require the same amount of time as CC, RC, CR and MI.

4. Discussion

The proposed modification of the SRC algorithm renders rank correlation as an attractive alternative to common merit functions in 2D/3D registration, provided that the intensity depth of the images used is considerably smaller than the number of pixels used—an image of 256×256 pixels image size that uses the full range of 16 bit intensity depth would not benefit from the proposed algorithm in its current form. However, this scenario is rather academic since diagnostic CT scans used as source volumes for DRR rendering usually feature 12 bit resolution, and it is questionable whether full exploitation of 16 bit depth provides any advantage for the purpose of registration after appropriate windowing. Even the original image data from commercial flat panel detectors, which are provided in 16 bit resolution, only use a small part of the intensity range possible in our experience.

Our evaluation reveals that the theoretic advantage of using rank correlation-type measures—the independence of histogram content in the DRR and the reference x-ray as long as a monotonic relationship of corresponding gray values is provided—is also visible in practice. This is especially true for large histogram discrepancy, where the CR is inferior to RC and SRC according to our results. The performance of MI—a measure known for its outstanding performance in multimodal 3D/3D registration—is somewhat disappointing. This is, however, a well-known fact from the literature (Penney *et al* 1998; Birkfellner *et al* 2009) as pointed out in the introduction.

Furthermore, one may ask whether the computation of RC results in an advantage over SRC, namely since the latter is even more performant by means of computational load in the present evaluation. The data on algorithm stability and performance do, however, indicate that RC gives more accurate results with a lower variance compared to SRC. We conclude that both methods—RC and SRC—have the potential to solve a given clinical task in 2D/3D registration. As pointed out in the introduction, a wide range of methods is desirable in this case since there is no golden rule.

The limitation of rank correlation lies in the fact that it is, by definition, a merit function for intramodal registration; a registration of x-ray image data and DRRs rendered from MRI is not feasible. The advantages of rank correlation—namely the robustness against outliers and the invariance to variations in image intensity and detector characteristics—do however outweigh this drawback for the specific problem of x-ray and CT data. A future task for further research is a throughout evaluation of the applicability of rank correlation to 2D/3D registration using electronic portal images. It is especially interesting to see whether a robust merit function such as RC can also cope with pairs of simultaneously acquired kV and MV images, which may be retrieved from a number of modern image-guided radiation therapy units.

Acknowledgments

We wish to thank D Georg, M Stock (Department of Radiotherapy, MUW), F Kainberger, I Nöbauer (Department of Diagnostic Radiology, MUW) and H Bergmeister (Center for Biomedical Research, MUW) for their help in the acquisition of the porcine reference image data. This work was supported by the Austrian Science Foundation FWF under projects P19931 and L503. JH was supported by FWF project L625.

References

- Bednarz G, Machtay M, Werner-Wasik M, Downes B, Bogner J, Hyslop T, Galvin J, Evans J, Curran W Jr, Andrews D. Report on a randomized trial comparing two forms of immobilization of the head for fractionated stereotactic radiotherapy. *Med. Phys.* 2009; 36:12–7. [PubMed: 19235368]
- 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; 34:246–55. [PubMed: 17278510]
- 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:N73–84. [PubMed: 15843725]
- Birkfellner W, Stock M, Figl M, Gendrin C, Hummel J, Dong S, Kettenbach J, Georg D, Bergmann H. Stochastic rank correlation: a robust merit function for 2D/3D registration of image data obtained at different energies. *Med. Phys.* 2009; 36:3420–8. [PubMed: 19746775]
- Cho B, Suh Y, Dieterich S, Keall PJ. A monoscopic method for real-time tumour tracking using combined occasional x-ray imaging and continuous respiratory monitoring. *Phys. Med. Biol.* 2008; 53:2837–55. [PubMed: 18460750]
- 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]

- Pawiro SA, Markelj P, Gendrin C, Figl M, Stock M, Bloch C, Weber C, Unger E, Noebauer I, Kainberger F, Bergmeister H, Georg D, Bergmann H, Birkfellner W. A new gold-standard dataset for 2D/3D image registration evaluation. *Proc. SPIE*. 2010; 7625:76251V.
- Penney GP, Weese J, Little JA, Desmedt P, Hill DL, Hawkes DJ. A comparison of similarity measures for use in 2-D–3-D medical image registration. *IEEE Trans. Med. Imaging*. 1998; 17:586–95. [PubMed: 9845314]
- Sawant A, Smith RL, Venkat RB, Santanam L, Cho B, Poulsen P, Cattell H, Newell LJ, Parikh P, Keall PJ. Toward submillimeter accuracy in the management of intrafraction motion: the integration of real-time internal position monitoring and multileaf collimator target tracking. *Int. J. Radiat. Oncol. Biol. Phys.* 2009; 74:575–82. [PubMed: 19327907]
- Sharp GC, Jiang SB, Shimizu S, Shirato H. Tracking errors in a prototype real-time tumour tracking system. *Phys. Med. Biol.* 2004; 49:5347–56. [PubMed: 15656282]
- Skerl D, Likar B, Pernus F. A protocol for evaluation of similarity measures for non-rigid registration. *Med. Image Anal.* 2008; 12:42–54. [PubMed: 17683976]
- Spoerk J, Bergmann H, Wanschitz F, Dong S, Birkfellner W. Fast DRR splat rendering using common consumer graphics hardware. *Med. Phys.* 2007; 34:4302–8. [PubMed: 18072495]
- 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; 24:1177–89. [PubMed: 16156355]
- Zhu X, Bourland JD, Yuan Y, Zhuang T, O'Daniel J, Thongphiew D, Wu QJ, Das SK, Yoo S, Yin FF. Tradeoffs of integrating real-time tracking into IGRT for prostate cancer treatment. *Phys. Med. Biol.* 2009; 54:N393–401. [PubMed: 19661570]

Figure 1.

The reference x-ray of the porcine phantom, taken from an antero-posterior view. The seven fiducial markers for the determination of the gold-standard reference registration are clearly visible. The markers were not removed from the image data for the evaluation. From our experience, the markers do not have a significant influence on registration results.

Table 1

Mean, standard deviation and median for 150 registrations using different rendering thresholds for DRR generation. Both the mTRE—based on the evaluation of 400 known positions inside the CT volume—and the mean error in the projected positions (mean projected distance mPD) were recorded. The upper rows give the mean and standard deviation, whereas the median is given in the lower row. RC generally provides the lowest results with the lowest scatter as indicated by the standard deviation, and appears to be the measure most robust against differences in image intensities which were introduced by changing the rendering threshold. The best values for mTRE and mPD are printed in boldface in each column.

	Threshold (HU)	CC	RC	SRC 5%	CR	MI
mPD (mm)	-868	2.96 ± 1.25 2.74	3.08 ± 2.62 2.66	3.82 ± 3.31 3.11	7.68 ± 4.65 6.45	3.88 ± 1.09 3.73
	-711	4.60 ± 2.38 4.01	2.50 ± 0.76 2.38	2.67 ± 1.52 2.16	2.00 ± 1.90 1.49	3.27 ± 2.03 2.86
	-555	4.79 ± 3.26 3.91	2.68 ± 0.43 2.62	3.12 ± 2.23 2.52	2.71 ± 2.57 1.52	3.36 ± 1.74 2.96
	-400	6.02 ± 4.55 4.38	3.02 ± 1.34 2.87	3.86 ± 2.89 3.22	2.54 ± 2.18 1.62	4.21 ± 2.35 3.51
	-243	14.05 ± 2.19 13.48	6.65 ± 8.11 4.3	6.84 ± 6.19 5.17	8.07 ± 4.90 6.67	15.38 ± 5.88 12.91
mTRE (mm)	-868	4.16 ± 2.84 2.93	4.04 ± 3.07 2.95	4.48 ± 3.45 3.20	6.55 ± 2.90 5.80	5.64 ± 2.90 4.72
	-711	5.09 ± 3.07 3.92	3.27 ± 2.52 2.66	3.77 ± 2.62 2.68	3.56 ± 3.11 2.18	4.29 ± 2.71 3.44
	-555	5.79 ± 4.27 3.85	3.69 ± 2.35 2.80	4.01 ± 2.97 2.87	3.80 ± 2.92 2.76	4.33 ± 2.98 3.39
	-400	6.80 ± 5.28 4.27	4.03 ± 2.51 2.90	4.34 ± 3.28 3.33	3.86 ± 2.95 2.93	5.13 ± 3.91 3.73
	-243	13.77 ± 1.80 13.15	6.94 ± 7.70 4.22	7.02 ± 5.93 5.03	8.10 ± 4.39 6.56	15.27 ± 5.63 12.76