

# How to achieve fast, accurate, and robust rigid registration between fluoroscopic X-ray and CT images

L. Joskowicz and D. Knaan

School of Engineering and Computer Science, The Hebrew Univ. of Jerusalem, Israel.  
E-mail: josko@cs.huji.ac.il, WWW site: <http://www.cs.huji.ac.il/~josko>

This paper presents algorithmic refinements to our method for rigid registration between preoperative CT and intraoperative X-rays and studies the influence of various set-ups and factors on the methods' performance. Experimental results on five different anatomical structures show that the new three-phase algorithm achieves a mean accuracy of 0.7mm (max=1.1mm) 95% of the time with an initial pose offset of up to 60mm in 80 secs average with three X-ray images and CT slice spacing of 2.4mm and handles image clutter well.

## 1. INTRODUCTION

Practical, accurate, and robust registration between a preoperative study and the intraoperative situation is a key technical challenge in Computer Aided Surgery. Current systems rely on implanted fiducials, which require additional surgery, or on points harvested on the surface anatomy by direct contact, which usually requires additional anatomical exposure. Anatomical image-based registration uses intraoperative images instead, and so is non-invasive, potentially faster, and less error-prone. However, it is technically much harder because it requires analyzing intraoperative images, which are usually low-resolution, noisy, have distortions and a reduced field of view, and may contain surgical tools and implants not present in the preoperative study.

Rigid registration between a preoperative CT and the intraoperative situation using fluoroscopic X-ray images has received much attention since these are the choice imaging modalities in many procedures. Geometry-based algorithms match selected geometric features from each data set by finding the transformation that minimizes the sum of distances between paired features [1–3]. Intensity-based algorithms match the intensities of one image data set with the intensity of the other by minimizing a measure of difference between them [4–9]. The main obstacles remain robustness, accuracy, and efficiency.

## 2. MATERIALS AND METHODS

We have developed a practical, fast, and robust approach to anatomical image-based rigid registration between preoperative CT images and intraoperative fluoroscopic X-ray images obtained with a tracked C-arm [9]. The approach is generic, customizable, and is targeted for use in orthopaedic surgery. In this paper, we present algorithmic refinements and study the influence of different set-ups and factors on the algorithm's performance.

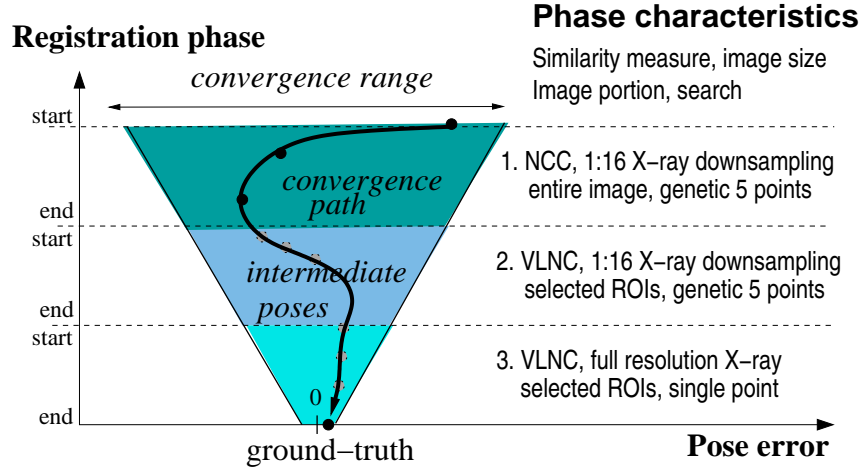


Figure 1. The registration funnel. The horizontal and vertical axes are the distance from the ground truth transformation and the three registration phases with their characteristics and convergence range (gray regions). Dots on a typical convergence path indicate anatomy pose estimates as they are refined during the registration process.

The registration algorithm proceeds by generating Digitally Reconstructed Radiographs (DRRs) of the anatomy at its estimated pose for each C-arm viewpoint, computing the anatomy pose difference by comparing the DRRs with the real fluoroscopic X-ray images, and finding the new pose that minimizes the difference. The process is repeated with the new pose estimate until no further progress can be made. The algorithm incorporates five techniques: 1. Fast DRR generation from precomputed CT rays; 2. Two similarity measures to compare DRRs and X-ray images based on Normalized Cross-correlation (NCC and VLNC); 3. Multiresolution X-ray fluoroscopic images; 4. Dynamic Regions of Interest (ROIs) around bone contours, and; 5. Genetic search to avoid local minima.

We found it necessary to combine the above techniques into three phases with different characteristics to achieve the desired results (Fig 1). Each phase brings the initial estimate closer to the ground truth within the convergence range. In the first phase, we use the Normalized Cross Correlation (NCC) comparison measure because it is robust to noise, invariant to linear transformations, has a wide convergence range, but is inaccurate. We use 1:16 downsampled fluoroscopic X-ray with a single ROI (the entire image) and search the initial pose estimate and four additional poses nearby. For the second phase, we use the Variance Weighted Sum of Local Normalized Correlation (VLNC) because it enhances the influence of regions with potential useful information, is accurate, but has a narrower convergence range. We use 1:16 downsampled fluoroscopic X-ray images with dynamic 9x9 pixel ROIs and start the search with the pose estimate from the first phase and four others nearby. For the third step we use VLNC, full resolution fluoroscopic X-ray images, dynamic ROIs (10–15% of the image pixels) and the best pose estimate from the second phase because the computed transformation is very close to the final one. Since full-resolution fluoroscopic X-ray images are used in the last phase, the final accuracy is not affected by the multiresolution approach.

### 3. EXPERIMENTAL RESULTS

We conducted experiments on five anatomical structures in different scenarios: a living human pelvis CT study and simulated X-ray images, a dry spine, a dry femur, a cadaver lamb hip, and a cadaver human pelvis. Each scenario (ideal, realistic, and worst-case) is defined by the CT slice spacing (0.6, 2.4, and 4.2mm) and the presence of foreign objects in the fluoroscopic X-ray images (none, some, and some).

To obtain the ground truth registration transformation, we implanted in all but the first case several (4–7) 6mm aluminum spheres and CT scanned them at 0.6mm slice interval and 0.5mm slice thickness. We extracted the sphere centers to a mean accuracy of 0.1mm (max=0.3mm). In the operating room, we acquired sets of several fluoroscopic X-ray images (800x600 pixels, 8-bit gray-scale, pixel size =  $0.45mm^2$ ) at various C-arm orientations. We performed C-arm calibration to a mean accuracy of 0.3mm (max=0.6mm) [10]. We performed fiducial contact-based registration on the spheres and established the ground-truth transformation. We then applied the registration algorithm and computed the surface Target Registration Error (sTRE), which is the distance between the ground truth and the computed position of points on the bone surface.

For the spine, we obtained an average sTRE of 0.4mm (max=0.7mm) with 100% (95%) success for an initial sTRE pose offset of 0–10mm (0–15mm) with an average running time of 100 secs. For the femur we obtained an average sTRE of 0.9mm (max=1.8mm) with 100% (95%) success for an initial sTRE pose offset of 0–25mm (0–40mm) with an average running time of 45 secs. The spine convergence range is smaller than that of the femur because of the presence of adjacent vertebrae. For the cadaver lamb hip, we obtained an average sTRE of 1.1mm (max=1.8mm) with 100% (95%) success for an initial sTRE pose offset of 0–40mm (0–55mm), with an average running time of 85 secs. For the human pelvis we obtained with four fluoroscopic X-ray images, an average sTRE of 0.9mm (max=1.1mm) with 100% (95%) success for an initial sTRE pose offset of 0–45mm (0–55mm), with an average running time of 110 secs. The presence of foreign objects and surrounding bones had minor influence on the final results.

Our experiments show that the algorithm takes on average 70–90 secs and requires 3,000 DRRs: 5 secs for initialization, 10 secs and 400 DRRs for the first phase, 30 secs and 2,400 DRRs for the second, and 25 secs and 200 DRRs for the third.

To justify our choice of methods and further test the influence of different parameters, we conducted the following technical experiments.

**1. Number of fluoroscopic X-ray images:** To determine how many images should be used, we conducted experiments on the femur data in an ideal scenario and on the cadaver pelvis in a realistic scenario. Two or more images from different C-arm viewpoints with pose differences of at least  $25^\circ$  between them were taken. Fig. 2 shows the results for the femur. We note a slight decrease (0.1mm) in accuracy with two fluoroscopic X-ray, and no significant increase with more than three images. For the cadaver pelvis, we observed more significant improvement between three and four images (0.8mm instead of 1.2mm mean sTRE, same convergence range) at the expense of 30% more running time.

**2. Two-stage comparison measure:** To demonstrate the need for a two-stage comparison measure, we conducted the following experiments on the lamb hip in an ideal scenario. We performed the registration with NCC only, VLNC only, and NCC followed

Number of X-ray images	final sTRE		Convergence range		Average running time (secs)
	mean	max	100%	95%	
2	0.8	1.1	1-20	1-30	18
3	0.7	1.0	1-25	1-40	32
4	0.7	0.9	1-25	1-40	37
5	0.7	0.9	1-25	1-40	47

Figure 2. Relation between the number of fluoroscopic X-ray images and the accuracy, convergence range, and running time of the registration algorithm for the femur data set.

by VLNC. The NCC converges for initial pose estimates that are far from the final pose (100% success for sTRE offsets of 1-40mm, and 95% for 1-55mm), although its accuracy is poor (mean sTRE=8mm). VLNC is more accurate, with an average sTRE of 1.1mm, but its convergence is only 95% in the 1-20mm range. The combination of both measures is superior: NCC widens the convergence range and provides pose estimates in the range for VLNC to convergence 100% with maximum sTRE of 0.3mm. The convergence range is well within the range of practical coarse initial pose estimation.

**3. Genetic algorithm:** To demonstrate the need for a genetic search algorithm, we conducted experiments on the spine in an ideal scenario. The convergence range is relatively small (0-15mm) due to the small size of the vertebra and the presence of adjacent vertebrae. Fig. 3(a) shows the results with/without genetic search. We note that simple search achieves a cumulative success rate of 100% (95%) for the 0-2mm (0-5mm) range in average running time of 60 seconds, while the range is expanded to 0-10mm (0-15mm) with average running time of 95 secs when genetic search is added. The final accuracy of both is very similar since in the final phase of search a single pose estimate is used.

**4. Full resolution:** To quantify the influence of downsampling on the registration accuracy, we conducted experiments on the spine in an ideal scenario. Fig. 3(b) shows that an average final sTRE of less than 1mm (max=2mm) is already achievable with 1:16 down-sampled fluoroscopic X-ray images in 50 secs on average. With full resolution, the final average sTRE decreases to 0.3mm (max=0.5mm) at the expense of an additional 45 secs of running time. Note that the final accuracy is independent of the initial pose estimate. Full resolution can thus be traded off for running time at the expense of accuracy.

**5. Robustness:** To demonstrate the robustness of the algorithm to soft tissue, surrounding anatomy, and foreign objects in the fluoroscopic X-ray images, we conducted experiments on the cadaver lamb hip (Fig. 4). Two sets of three fluoroscopic X-ray images with initial sTRE pose offset of 10mm were taken. Soft tissue and part of the surrounding pelvis appear naturally; dark contours were added manually to simulate the presence of metallic surgical instruments. The registration succeeded for the first set to a final sTRE of 1.5mm, while for the second it failed (final sTRE=9.9mm). The difference between them was the presence of noise right on the bone contours, which contain the most useful information. The added image clutter is way over what is normally seen.

**6. Computational issues:** We explored two computational issues: the location of the anatomy reference coordinate frame location and the choice of the optimization method.

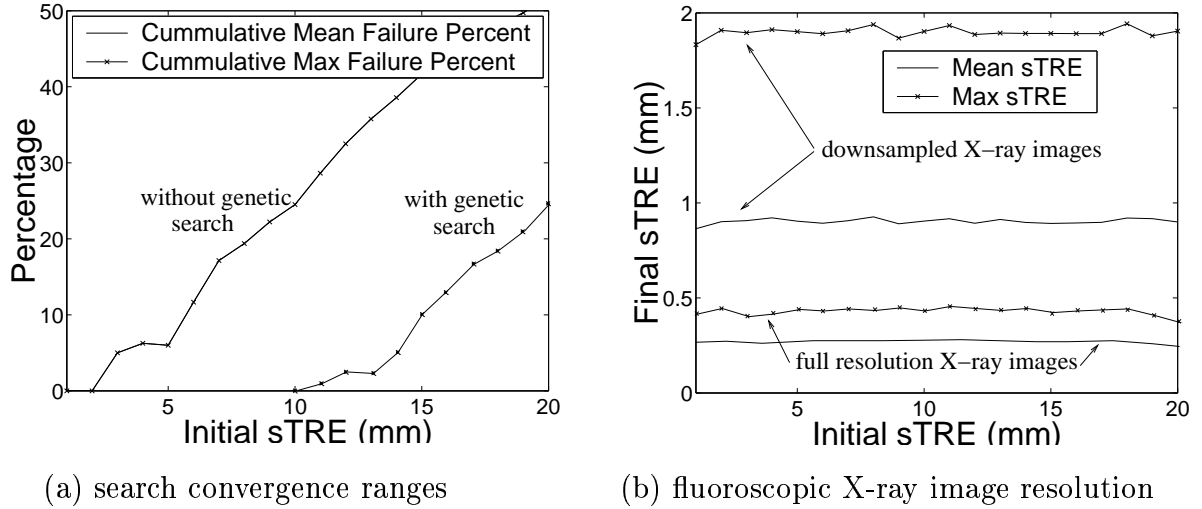


Figure 3. (a) Convergence ranges to a final registration of  $sTRE \leq 2\text{mm}$  (for the spine data set) with and without a genetic search, and; (b) final registration sTRE with 1:16 downsampled and full resolution fluoroscopic X-ray images.

We found that rotating the anatomy around the CT center or around the CT-volume frame origin has little influence on the convergence range and the final sTRE but that the running time decreased by an average of 20%, from 120 to 95 secs. This is most likely due to the looser coupling between the rotation and translation parameters, which causes the optimization to converge faster. We compared the Powell and the Downhill Simplex optimization methods for the clinical pelvis in a realistic scenario and found the nearly identical. Both resulted in the same final mean sTRE error (0.3mm). The success rate with Downhill Simplex was 100% (95%) on the 0–20mm (0–30mm) range with an average running time of 80 secs, compared to 100% (95%) success on the 0–15mm (0–20mm) range with an average running time of 115 secs with Powell’s method.

#### 4. CONCLUSION

The results of our experiments demonstrate that the proposed three-phase method is generic, practical, robust, and accurate. The new three-phase algorithm achieves a mean accuracy of 0.7mm (max=1.1mm) 95% of the time with an initial pose offset of up to 60mm in 80 secs average with three X-ray images and CT slice spacing of 2.4mm and handles image clutter well. These results are superior from those reported recently [8,11].

#### ACKNOWLEDGMENT

This research was supported in part by a grant from the Israel Ministry of Industry and Trade for the IZMEL Consortium on Image-Guided Therapy.

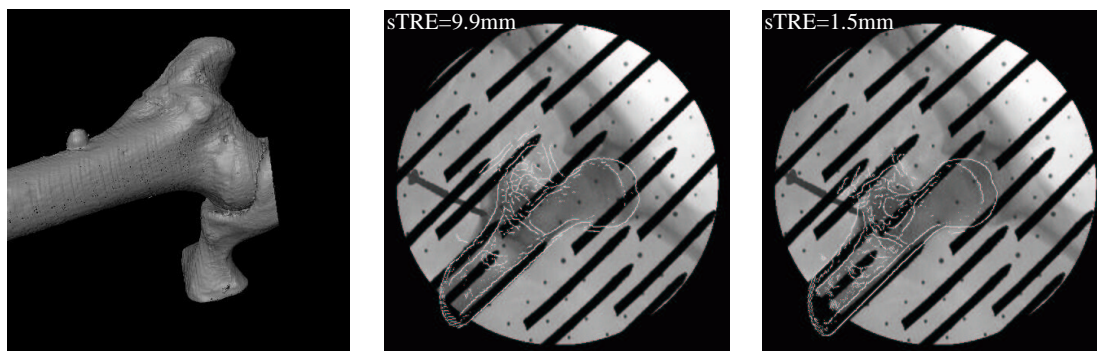


Figure 4. Robustness to image noise on the cadaver lamb hip experiment. The left image shows the hip model from the CT. The center and right images show fluoroscopic X-ray images with contours at initial and final pose superimposed on them (white contours)

## REFERENCES

1. Besl, P. and McKay, N. A method for registration of 3D shapes, *IEEE Trans. on Pattern Analysis and Machine Intelligence* 14(2), 1992.
2. Hamadeh, A., Lavallée, S. and Cinquin, P., 'Automated 3-dimensional computed tomographic and fluoroscopic image registration, *Computer Aided Surgery* 3(1), 1998.
3. Guéziec, A., Kazanzides, P., Williamson, B., and Taylor, R.H., Anatomy based registration of CT-scan and intraoperative X-ray images for guiding a surgical robot, *IEEE Transactions on Medical Imaging* 17(5), 1998.
4. Lemieux, L., Jagoe, R., Fish, R. *et al.*, A patient-to-computed-tomography image registration method based on DRRs, *Medical Physics* 21(11), 1994.
5. Murphy, M.J., An automatic six-degree-of-freedom registration algorithm for image-guided frameless stereotactic radiosurgery, *Medical Physics* 24(6), 1997.
6. Roth, M., Brack, C., Burgkart, R., and Czopf, A., Multi-view contourless registration of bone structures using a single calibrated x-ray fluoroscope, *Proc. of the Computer-Assisted Radiology and Surgery Conf.*, 1999.
7. LaRose, D.A, Bayouth, J., and Kanade, T., Transgraph: interactive intensity-based 2D/3D registration of X-ray and CT data, in *SPIE Image Processing*, 2000.
8. Tomazevic, D., Likar, B., Slivnik, T., and Pernus, F., 3-D/2-D registration of CT and MR to X-ray images. *IEEE Trans. on Medical Imaging*, Vol 22(11), 2003.
9. Knaan, D., and Joskowicz, L., Effective intensity-based 2D/3D rigid registration between fluoroscopic X-ray and CT, MICCAI'2003, *Lecture Notes in Computer Science 2879*. Ellis, R. and Peters, T. editors, Springer, 2003.
10. Livyatan, H., Yaniv, Z., and Joskowicz, L. Robust automatic C-arm calibration for fluoroscopy-based navigation: a practical approach, *Proc 5th Int. Conf. on Medical Image Computing and Computer-Aided Intervention*, Tokyo, Japan, 2002.
11. Livyatan, H., Yaniv, Z., and Joskowicz, L. Gradient-based 2D/3D rigid registration of fluoroscopic X-ray to CT. *IEEE Trans. on Medical Imaging*, Vol 22(11), 2003.