Proceedings of the 29th Annual International
Conference of the IEEE EMBS
Cité Internationale, Lyon, France
August 23-26, 2007.

2D/3D Registration of Multiple Bones

Burton Ma, James Stewart, David Pichora, Randy Ellis, and Purang Abolmaesumi

Abstract— **The problem of 2D/3D registration is, given a 3D image of an object and one or more 2D images of the object in known poses, to recover the 3D pose of the object. We propose a solution for registering multiple bones in 2D radiographic images and 3D CT images using normalized correlation coefficient template matching. We performed tests using synthetic radiographs and CT volumes of two knees and one wrist. We obtained good registration results (less than 2° and 2 mm registration error) for all of the larger bones, but were unable to successfully register the small carpal bones of the wrist with high accuracy.**

I. INTRODUCTION

Registration of bony anatomy between 2D radiographic images and 3D CT images is a potentially useful tool for computer-aided surgery, orthopaedics research, and biomechanics research. The goal of 2D/3D registration of bony anatomy is to find the rigid transformation between the CT coordinate frame and the coordinate frames associated with the radiographic images.

Each image, be it radiographic or CT, has a coordinate frame associated with it. The CT volume has its own natural coordinate system. In computer-aided surgical applications, the radiographic images are typically provided by a C-arm fluoroscopic device that is tracked in 3D by a navigation system; each image has a pose relative to a tracked object (e.g., the patients anatomy). Knowledge of the relative poses of the 2D images and the imaging characteristics of the radiographic imager makes it possible to attempt a solution to the registration problem.

There are two common approaches to solving the 2D/3D registration problem when fiduciary markers are not used. The first approach extracts the shape of the bony anatomy from the CT scan, and uses shape features such as surface points or curves that are registered to the contours of the 2D images by fitting the shape features to the back-projected lines going from the 2D images to the x-ray source [2], [4]. The success of such techniques depends on the accuracy of the segmentation of the 2D and 3D images.

The second approach uses the CT volume to generate digitally reconstructed radiographs (DRRs). These intensity-based methods compare the DRRs to the 2D radiographs using a similarity metric. Registration is performed by finding the DRRs that are maximally similar to the radiographs [1], [3], [5], [7]–[9], [11].

We are interested in solving the problem of registering multiple bones or bony fragments that appear in the 2D and 3D images. This problem was studied in [1] where images of the knee were used to separately register the femur and tibia; however, they used a very good initialization of the registration estimate ($\pm 3^\circ$ and ± 3 pixels). Our approach is based on the well-studied technique of template matching using normalized cross-correlation (for example [12]). For each bone fragment, we generate a template from the DRR computed using the current registration estimate. The template is matched to the 2D image and the maximum value of the normalized cross-correlation coefficient is used as the similarity metric. We search for the template that produces the highest value of the similarity metric. Because template matching has quadratic complexity in the template and image sizes, we use a multi-scale approach to attain acceptable execution times. The search occurs over the widest range at the coarsest scale, and narrows at finer scales. We test our algorithm on knee and wrist images with initial registration estimate errors of 15° of rotation and 10 mm of translation.

II. METHODS

Our approach is based on our experience in planning three-dimensional corrections of bone deformities for computer-aided surgery. We typically use three orthogonal views of the anatomy for surgical planning. The rotational deformity is corrected by sequentially correcting the in-plane rotation (the rotation about the viewing direction) in each of the frontal, sagittal, and axial views. Usually a few iterations are required to correct the rotational deformity because the rotational components are not independent. The translational component of the deformity is easily corrected once the rotational component is corrected. This approach solves the 6-dimensional problem of finding the three rotational and three translational correction parameters by solving six one-dimensional optimization problems. When registering CT to x-ray images, the axial view is usually not available, and the axial component of the rotation must be estimated from the other views.

We assume that viewing directions of the 2D radiographic images and an estimate of the registration transformation for each bone fragment is available. We also assume that the CT scan has been processed to isolate the bone fragments. An explicit segmentation of the bones from soft tissue is not necessary; we only require that the bone fragments be available as individual volumes.

This work was supported by the Canadian Institutes of Health Research.

B. Ma is with the Human Mobility Research Centre at Kingston General Hospital, Kingston, Ontario, Canada. mab@cs.queensu.ca

D. Pichora is with the Department of Orthopaedics at Kingston General Hospital, Kingston, Ontario, Canada

P. Abolmaesumi, J. Stewart, and R. Ellis are with the School of Computing Science, Queen's University, Kingston, Ontario, Canada. {purang,jstewart,ellis}@cs.queensu.ca

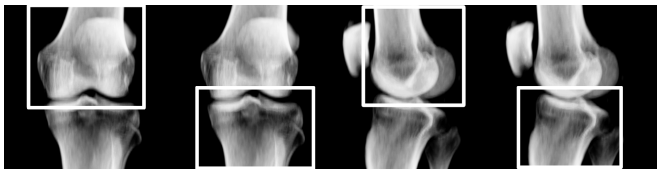

Fig. 1. Regions of interest identifying the femur and tibia. Note that the regions of the femur and tibia overlap, and that the regions all include bones other than the one of interest.

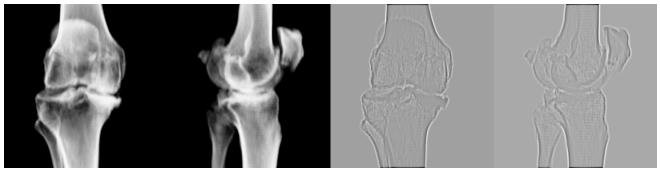

Fig. 2. Original x-ray images and their Laplacian of Gaussian images.

A. Algorithm Description

The first step of the algorithm is user identification of the bone fragments. Each bone fragment that is to be registered is identified with a rectangular region of interest. The regions will overlap if the bone fragments to be registered are adjacent. We do not use regions that include only the bone of interest; portions of adjacent bones are included in the regions. Regions of interest for the knee are shown in Fig. 1.

Our algorithm uses a multi-scale template matching approach for estimation; thus, we downsample the x-ray image dimensions by a factor of two and four in the second step. The search range for the registration parameters is widest at the lowest scale where the template matching can be performed quickly. The search narrows at each increasing scale as the parameter estimates improve. The multi-scale approach is necessary to achieve acceptable running times because the template matching complexity is quadratic in the image and template sizes.

Edge-like images are useful for orthopaedic registration problems because the soft-tissue component of the image tends to be attenuated. Rather than performing explicit edge-detection, we use Laplacian of Gaussian (LoG) images that can be efficiently computed using two one-dimensional convolutions (because the LoG filter is separable). The LoG filter used had standard deviation $\sigma = 2$ pixels and was truncated at 3σ. LoG image examples are shown in Fig. 2.

Image similarity is computed using normalized cross-correlation (NCC) for template matching. The template is generated by using 3D texture-mapping volume rendering to produce DRRs from the CT scan; we use the VTK volume renderer for DRR generation (`www.vtk.org`). The template is taken to be the portion of the DRR that coincides with the ROI defined on the x-ray image. The image is taken to be the zero-padded ROI on the x-ray image. The amount of padding controls the amount of translational misalignment that may be tolerated; larger amounts of padding allow for larger translational misalignment at the expense of greater computation time. The images are padded on each side by $1/2$, $1/4$, and $1/8$ of the template dimensions at the coarsest, middle, and original scales, respectively. The output of the template matching is an image of NCC coefficients of size $(I_w - T_w + 1) \times (I_h - T_h + 1)$, where I_w, I_h, T_w, T_h are the image width, image height, template width, and template height, respectively. The image similarity is the maximum value in the image of NCC coefficients. We use the OpenCV implementation of NCC template matching (`sourceforge.net/projects/opencvlibrary`).

Our registration algorithm attempts to maximize the NCC coefficient by searching for the best template. The search begins on the first x-ray image. The in-plane rotation (rotation about the direction of projection used to generate the DRR) is exhaustively searched at a resolution of $\delta\theta^\circ$ over a range of $\pm\Theta^\circ$ around the initial estimate of the registration transformation; that is to say, a sequence of search angles $\theta_j = -\Theta, -\Theta + \delta\theta, ..., \Theta$ is generated. The bone fragment is rotated about the direction of projection by angle θ_j and a DRR is generated. The template is extracted from the DRR and matched to the x-ray image. The angle producing the template that yields the largest similarity value is taken to be current estimate of the in-plane rotation, and the bone fragment is rotated by the current estimate. The process of estimating the in-plane rotation is repeated using the second x-ray image. For the purposes of this article, we used $\Theta = 20, 5, 2^\circ$ and $\delta\theta = 2, 1, 0.5^\circ$ at the coarsest, middle, and original scales, respectively.

Next, the rotation about the axis given by the cross-product of the two x-ray viewing directions is estimated; for x-rays in the frontal and sagittal planes this rotation would correspond to an axial rotation. The same process used to estimate the two in-plane rotations is used. The rotation is estimated on the first x-ray and then again on the second x-ray.

The process of estimating the rotational components is iterated three times at the coarsest scale, and only once at each finer scale. The in-plane translational components are estimated after the rotational components have been estimated at each scale. Starting with the first x-ray, a single template match is performed; the location of the maximum value in the NCC coefficients image is related to the in-plane translation component of the registration transformation. If the maximum value is in the center of the NCC coefficients image then the template and image are optimally aligned; otherwise, a translation towards the maximum value from the image center is required. The bone fragment is translated by small amount in the necessary direction, and the template matching process is iterated until the maximum value in the NCC coefficients image is centered or a fixed number of iterations is reached. The translation amount at each iteration is 1, 0.5, and 0.25 mm at the coarsest, middle, and original scales, respectively. The process is repeated for the second x-ray. We have found that it is beneficial to re-estimate the axial rotation after estimating the translation at the coarsest scale.

When registering multiple bones, all of the bones are considered in sequence at each step rather than performing all of the steps for each bone one at a time. There does not appear to be any computational reason to prefer one approach

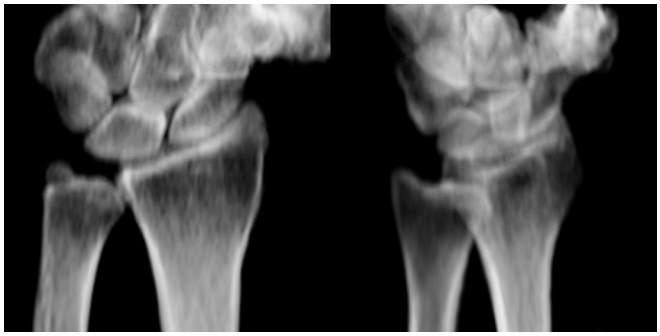

Fig. 3. X-ray images of the wrist. The largest bone is the radius. The two smaller bones articulating immediately above the radius are the lunate and scaphoid.

TABLE I

REGISTRATION ERRORS REPORTED AS *mean (standard deviation)*.

	Rotation degrees				Trans. mm	Failed %
	Frontal	Sagittal	Axial	Total		
Femur 1	-0.07	-0.07	-0.42	0.51	0.46	0.3
	(0.17)	(0.16)	(0.26)	(0.23)	(0.20)	
Tibia 1	-0.03	-0.10	-0.16	0.42	0.51	5.8
	(0.15)	(0.15)	(0.37)	(0.20)	(0.25)	
Femur 2	0.13	-0.13	0.10	0.38	0.40	2.1
	(0.19)	(0.16)	(0.25)	(0.17)	(0.18)	
Tibia 2	-0.24	0.04	0.14	0.44	0.49	0.0
	(0.15)	(0.24)	(0.27)	(0.19)	(0.14)	
Radius 1	-0.06	-0.04	-0.47	0.87	0.51	1.3
	(0.33)	(0.29)	(0.73)	(0.44)	(0.30)	

over the other.

B. Experimental Validation

CT scans of three patients were obtained from Kingston General Hospital. All patients had given informed consent for the use of their medical images for research purposes, and we had ethics approval from all institutions involved in this project. Two CT scans were of knees and one was of a wrist. All scans were performed using single slice CT machines with slice resolution of 512×512 pixels. Voxel sizes were approximately $0.6 \times 0.6 \times 2.0$mm for the knee scans, and $0.3 \times 0.3 \times 1.0$mm for the wrist scan.

We generated one pair of synthetic x-ray images using raycast volume rendering for each CT scan. Each pair of synthetic x-ray images comprised an AP image and an approximate lateral image; the imaging directions were not constrained to be exactly perpendicular. The wrist lateral x-ray was not a true lateral image because the carpal bones are not visible in such images. All x-ray images were 512×512 pixels. We intentionally did not match the transfer function used to generate the simulated x-rays with the transfer function used to generate the DRRs in the registration algorithm (i.e. the x-rays and DRRs did not have the same luminosity and contrast).

For the knee experiments we attempted to simultaneously register the femur and tibia. Each bone was independently displaced by a rotation of 15° about a random axis passing through the knee center followed by a translation of 10mm in a random direction. We registered the bones using our algorithm and computed the rotational and translational errors of the estimated registration transformation. Each experiment was run 1000 times.

For the wrist experiment we attempted to register the radius, scaphoid, and lunate. The scaphoid and lunate are especially challenging to register because they overlap each other as well as the surrounding carpal and arm bones in the x-ray images. We used the fact that the radiocarpal joint contributes primarily to flexion/extension and abduction/adduction motion of the wrist; thus, an accurate registration of the radius greatly constrains the registration parameter space of the scaphoid and lunate. We displaced the radius by a rotation of 15° about a random axis passing through the wrist center followed by a translation of 10 mm in a random direction. The scaphoid and lunate were displaced relative to the displaced radius by a random flexion/extension and abduction/adduction rotation of 25° followed by a random translation of 5 mm. We first registered the radius, and then initialized the registration estimate of the scaphoid and lunate using the radius registration result. Because the range of axial rotation of the scaphoid and lunate is known to be small compared to the flexion/extension and abduction/adduction rotation, we did not estimate the axial rotation until the finest resolution images were used. This experiment was repeated 500 times.

III. RESULTS

We computed rotational registration errors as they would be measured by a clinician (ie. as angular errors projected into the frontal, sagittal, and axial planes). We also computed the total rotational error by converting the rotation to its screw representation. The translational errors were computed as the difference between the estimated translation and the known actual translation. Registrations with rotational errors greater than x° or translational errors greater than $y$mm were considered failures. Mean and standard deviations of the registration parameter errors were computed after stripping the failures from the results. These results are tabulated in Table I, except for the scaphoid and lunate.

The results for the lunate and scaphoid were poor compared to the other bones we tested, although the lunate results were reasonably good given the poor fidelity of the input radiographic images. The total rotational and translational errors for the scaphoid and lunate are shown in Fig. 4.

Registration times were approximately 50 seconds for three bone fragments using commodity PC hardware (CPU: AMD X2 4400, GPU: NVIDIA 8800 GTS) without multi-threading. The most time consuming aspect of our algorithm appears to be template matching on the finest scale images; it would be relatively easy to thread this part of the algorithm to take advantage of multiple processors.

IV. DISCUSSION AND FUTURE WORKS

We have presented preliminary results of a template matching algorithm for registering radiographic and CT images. The high success rate of our algorithm using a small

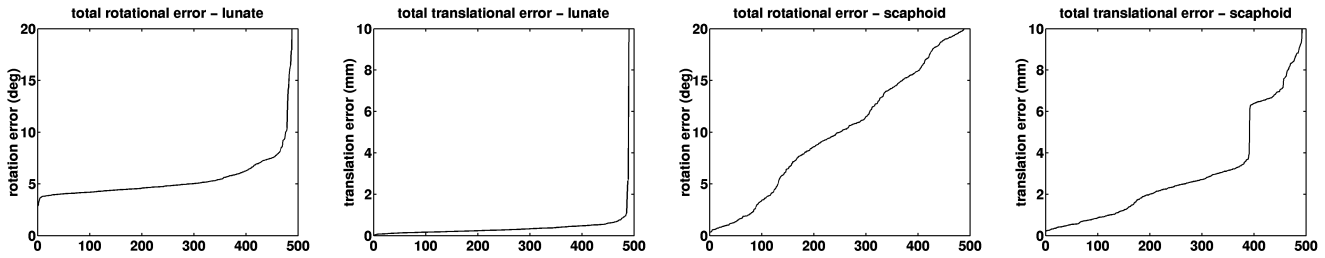

Fig. 4. Total rotational and translational errors for the lunate and scaphoid bones sorted in ascending order.

number of synthetic radiographic images makes us cautiously optimistic that our methods may be successful under more realistic circumstances.

Our results do not include the important effects of calibration errors in the x-ray imaging device. Previous work [10] has shown that calibration errors can have a large effect on registration accuracy. Also, the image characteristics of our synthetic radiographs are certainly different than those obtained from a real imaging device. We are planning on validating our results using real fluoroscopic images of cadaveric or animal specimens.

Our results for the lunate and scaphoid are disappointing but not surprising given the fact that the bones are small and have significant overlap with surrounding bones in the radiographic images. The lunate results were somewhat inaccurate in rotation but consistent for most of the trials. The scaphoid results were inaccurate and inconsistent over all of the trials. Unfortunately, the scaphoid is arguably the most interesting bone in the wrist for surgical purposes as it is one of the most frequently fractured bones in the body. We hope to continue research towards computer-aided percutaneous pinning of the scaphoid using fluoroscopic registration. Clearly, we will need to refine and extend our approach to address this challenging problem.

Our main justification for using LoG images is that their computation is very efficient. By computing the LoG images at different resolutions we are essentially performing Marr-Hildreth edge detection [6]. The disadvantage of this approach is that other edge detectors are known to have superior performance to the Marr-Hildreth operator. Furthermore, by only retaining the edge information we are losing intensity information that might prove useful when using real radiographic images.

Our method for generating templates is computationally efficient but not optimal. We generated templates by considering the area of the DRR that corresponded to the region of interest on the radiograph. If the registration estimate used to generate the DRR is poor then it is possible that the template will not include all of the bone of interest. A better method would be to isolate the template from the DRR without regard to the region of interest. This requires iterating over the entire DRR which becomes inefficient on the finest scale images.

One reason for performing this work is that we are interested in performing 2D/3D registration for fracture treatment. Future work includes attempting to register and track bone fragments in radiographic images.

REFERENCES

[1] R. Dalvi, R. Abugharbieh, M. Pickering, J. Scarvell, and P. Smith, "Registration of 2D to 3D joint images using phase-based mutual information," *in SPIE Medical Imaging*, 2007, pp 651209 1–9.

[2] J. Feldmar, N. Ayache, and F. Betting, 3D-2D projective registration of free-form curves and surfaces. *Computer Vision and Image Understanding*, vol. 65, no. 3, 1997, pp 403-424.

[3] D. Knaan and L. Joskowicz. "Effective intensity-based 2D/3D rigid registration be tween uoroscopic X-ray and CT", *in Medical Image Computing and Computer Assisted Intervention*, 2003, pp I:351-358.

[4] 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, vol. 17, no. 4, 1995, pp 378-390.

[5] H. Livyatan, Z. Yaniv, and L. Joskowicz, Gradient-based 2D/3D rigid registration of uoroscopic x-ray to CT, *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, 2003, pp 1395-1406.

[6] D. Marr and E. Hildreth, Theory of Edge Detection, *Proceedings of the Royal Society of London B*, vol. 207, 1980, pp 305–328.

[7] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, A comparision of similarity measures for use in 2-D-3-D medical image registration. *IEEE Transactions on Medical Imaging*, vol. 17, no. 4, 1998, pp 586-595.

[8] D. B. Russakoff, T. Rohlng, A. Ho, D. H. Kim, R. Shahidi, J. R. Adler Jr., and C. R. Maurer Jr, "Evaluation of intensity-based 2D-3D spine image registration using clinical gold-standard data", *in Biomedical Image Registration: Second International Workshop*, 2003, pp 151-160.

[9] T. Tang, "A hybrid 2D/3D atlas registration approach for medical images", Doctoral thesis, Queen's University, Kingston, Ontario, Canada, 2005.

[10] T. Tang, "Calibration and point-based registration of uoroscopic images", Masters thesis, Queens University, Kingston, Ontario, Canada, 1999.

[11] L. Zöllei, "2D-3D rigid body registration of x-ray uoroscopy and CT images", Masters thesis, Massachusetts Institute of Technology, 2001.

[12] K. Briechle and U. D. Hanebeck, "Template matching using fast normalized cross correlation", *in SPIE AeroSense Symposium*, 2001, pp 95–102.