

A Computationally Efficient Approach for 2D-3D Image Registration

Md. Nazmul Haque, *Student Member, IEEE*, Mark R. Pickering, *Member, IEEE*, Moyuresh Biswas, *Member, IEEE*, Michael R. Frater, *Member, IEEE*, Jennie M. Scarvell and Paul N. Smith

Abstract— 2D-3D image registration has become an important tool in many clinical applications such as image-guided surgery and the kinematic analysis of bones in knee and ankle joints. A limitation of this approach is the need to recalculate the voxel values in the 3D volume for every iteration of the registration procedure. In this paper we propose a new 2D-3D image registration algorithm which uses the projected 2D data from the original 3D CT volume. For the majority of the iterations of the algorithm, only this 2D data is updated rather than the 3D volume. Experimental results show that similar registration accuracy to the approach which employs 3-D updates at every iteration can be achieved with our method if we employ 3-D updates only in the last few iterations. As a result of reducing the number of 3-D updates, the proposed approach reduces the time required to perform the registration by approximately a factor of five.

I. INTRODUCTION

2D-3D image registration is a process to estimate the 3D position of an object by comparing a projected image generated from the 3D volume with a captured 2D image of the object. This projected image is commonly referred to as a digitally reconstructed radiograph (DRR). In clinical applications, the 3D volume is usually captured using Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) and the 2D image is a single X-ray or a frame from a Fluoroscopy video.

Two primary components of the registration process are – the optimization procedure and the similarity measure [1, 2]. The registration accuracy is quantified with a similarity measure which must be multi-modal when the images to be registered are from dissimilar sensors, as is the case with the medical imaging application in this paper. Gradient-based optimization, though iterative in nature, is widely used in practice. These gradient-based approaches include, for example, Gauss-Newton and Levenberg-Marquardt algorithms [2].

Previous approaches to 2D-3D registration with regards to medical applications use different combinations of optimization and similarity measures. In [3], the authors

used Mutual Information and sparse histograms to register 2D fluoroscopy images to 3D CT data. Mutual Information has been widely used [4] and in [5] a newer measure known as Cross-Cumulative Residual Entropy (CCRE) was proposed.

Most 2D-3D image registration techniques iteratively update the position of the 3D object and generate a DRR for each new position. These DRR images are then compared with x-ray images to calculate the similarity measure [6, 7]. The most computationally expensive part of this procedure is the calculation of the updated 3D volume. To mitigate this problem of computational complexity some efficient methods for DRR generation have been proposed. In [8, 9] a number of DRRs are generated offline for a set of predefined out-of-plane rotations. These sample DRRs are used as a reference to estimate out-of-plane rotations during registration. These offline DRR generation methods can speed up the registration procedure but incur the penalty of reduced accuracy in out-of-plane measurement accuracy.

In this paper, our goal was to develop a fast, yet accurate method for 2D-3D registration. Our method is built upon our recently developed 2D-3D registration algorithm [1], which uses a gradient-based multi-modal approach to register 2D fluoroscopy images of normal bones in the knee to 3D CT data of the same bones. In the approach adopted in this paper, we substitute the 3D volume in the iterative stage with 2D projected data. The effect of the 3D transformation is then simulated in the 2D plane of the projection. This allows the algorithm to converge to the desired registered position without the need for the time-consuming recalculation of all the 3D pixel values.

II. 2D-3D REGISTRATION

A. Background

When registering 2D and 3D images it is necessary to project the 3D data onto a 2D plane so that it can be compared with an existing 2D image. In [1], for example, the voxel values of the 3D volume are summed along its depth to produce the 2D image. The registration algorithm of the approach in [1] proceeds as follows:

M. N. Haque, M. R. Pickering, M. Biswas and M. R. Frater are with the School of Engineering and Information Technology, University of New South Wales at the Australian Defence Force Academy, Canberra, Australia.

J. M. Scarvell and P. N. Smith are with the Trauma and Orthopaedic Research Unit in the Department of Surgery at The Canberra Hospital, Canberra, Australia.

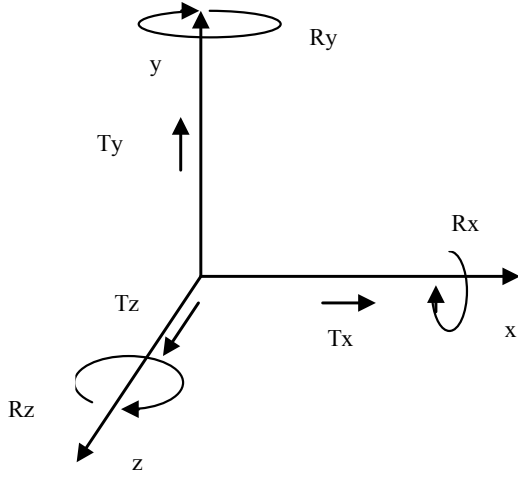


Figure 1. The orientation of the 6 parameters that are used to describe 3D rigid body motion

1. A geometric transform is applied on the 3D CT volume to find the warped position of the bone.
2. The warped volume is projected onto a 2D plane to produce the 2D DRR.
3. Both the DRR and 2D fluoroscopy frame are then filtered with a Laplacian-of-Gaussian (LoG) filter.
4. The similarity between the DRR and the fluoroscopy frame is then measured.
5. An update vector for the geometric transform is then found.
6. The algorithm is repeated with the updated transform parameters.

B. Proposed Fast Method

In this section, the new registration algorithm is described for the application of registering 3D CT data to 2D single plane fluoroscopy frames. Our approach to fast 2D-3D registration consists of simulating in the 2D plane the effect of the 3D transformation followed by projection onto the plane. This enables us to reconstruct a 2D DRR only once and use it throughout the iterative stage, as opposed to the algorithm in [1] where a new DRR is constructed at each iteration.

1) Simulating 3D translations in 2D

The rigid body motion of any 3D object can be modeled with 6 parameters, as shown in Figure 1. They are translations in x , y and z directions (T_x , T_y and T_z respectively) and rotations about the x , y and z axes (R_x , R_y and R_z respectively). Based on this model, the geometric transformation matrix for translations is as follows:

$$(x' y' z' 1) = (x y z 1) * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T_x & T_y & T_z & 1 \end{bmatrix} \quad (1)$$

where (x, y, z) are the original coordinates of the voxel and (x', y', z') are the transformed coordinates. Since the DRR is produced by summing the CT data in the z direction, translations in the x and y directions of the 3D CT volume will correspond to identical translations in the x and y directions for the 2D DRR.

However, translation in the z direction is different. The X-ray beams in the fluoroscopy unit diverge from a common point as they pass through the object and arrive at the image intensifier. Consequently, as an object moves closer to the X-ray source, it will produce a larger image at the image intensifier. This means that relative translation in the z direction will result in a change in scale of the object in the fluoroscopy frame. Hence, the out-of plane translation can be estimated by measuring the small amount of zoom that must be applied to the DRR to register it with the Fluoroscopy frame. This change in scale is usually difficult to measure precisely as a large translation in the z direction will result in only a very a small scale change in the fluoroscopy frame (see [10] for more details).

2) Simulating 3D rotations in 2D

Rigid body rotations of θ around the z , x and y axes are given in (2), (3) and (4) respectively.

$$\begin{aligned} x' &= x * \cos \theta - y * \sin \theta \\ y' &= x * \sin \theta + y * \cos \theta \\ z' &= z \end{aligned} \quad (2)$$

$$\begin{aligned} x' &= x \\ y' &= y * \cos \theta - z * \sin \theta \\ z' &= y * \sin \theta + z * \cos \theta \end{aligned} \quad (3)$$

$$\begin{aligned} x' &= z * \sin \theta + x * \cos \theta \\ y' &= y \\ z' &= z * \cos \theta - x * \sin \theta \end{aligned} \quad (4)$$

It is clear from (2) that rotation around the z axis can be applied directly in the 2D space. For rotations around the x and y axes, the object in the 2D DRR will stretch or shrink depending on its depth in the z direction. To accurately simulate these rotations in 2D, we need an estimate of the z coordinates of the edges of the object. We adopt the following estimate from [1].

$$\hat{z}_i = \frac{\sum_z z |V_e|}{\sum_z |V_e|} \quad (5)$$

In (5), \hat{z}_i are estimates of the object edges in the z -direction and V_e is a Laplacian-of-Gaussian filtered version of the 3D CT data. In our approach, the \hat{z}_i values are estimated only once before the first iteration. It is assumed that, for the small rotations considered in our work, the position of object edges in the z direction remain nearly constant and hence we can use the initially estimated \hat{z}_i values during the iterative update.

In updating the transformation parameters, we follow a three level update sequence. At the first level (for the first 10 iterations), only T_x , T_y and R_z are updated. Then up until iteration 25 (second level), we update all parameters except T_z . After iteration 25 (third level), all parameters including T_z are updated.

Understandably only updating the 2D DRR in the iterative optimization removes the expensive task of computing warped coordinates of the 3D volume and is expected to improve time complexity. One issue with this approach could be the registration accuracy. We discovered that if 3D updates are incorporated in the loop for only a few iterations, the resulting algorithm performs almost as accurately as the full 3D registration procedure.

We experimented with different ways to incorporate the 3D updates into the iterative optimization e.g. one 3D update after five 2D updates or one 3D update after 10 2D updates. Eventually we found that performing a 3D update at the beginning of each level and five 3D updates at the end of the procedure worked best in practice. Hence our algorithm uses the 2D DRR image for the first 40 iterations (except for the start of each level) and updates it with the simulated 2D transformation parameters. For the last 5 iterations, the 3D CT volume is updated and the transform parameters are initialized with the final 2D values. The similarity measure used in the procedure was the Sum of Conditional Variance (SCV) measure from [1].

Summarizing the above, the proposed algorithm proceeds as follows:

Pre-compute:

1. Construct the 2D DRR image of the original CT volume.
2. Estimate the value of \hat{z}_i .

Iterate the following steps:

1. For the start of each level and for the last 5 iterations apply a 3D transform on the CT volume and construct a DRR image from the updated

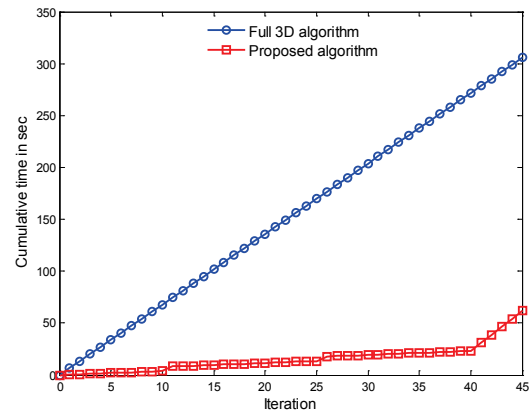


Figure 2. Computational complexity for the proposed algorithm and the full 3D algorithm from [1].

volume. Otherwise apply a 2D transform on the DRR image.

2. Apply (LoG) filter to both the DRR and Fluoroscopy frames to highlight the edges of the object.
3. Update the transformation parameters using the SCV similarity measure [1].
4. Repeat the algorithm with the updated parameters.

III. EXPERIMENTAL RESULTS

In our experiments, we compared the performance of the proposed algorithm with the multi-modal gradient based registration technique of [1]. This is because our approach is based on [1] and we aimed to improve the time computational efficiency with our fast algorithm. The algorithms were used to register segmented 3D CT data of a female tibia and femur to synthetic 2D fluoroscopy images. We have used low resolution versions of the original data for this experiment (1 pixel = 1 mm). The exact 3D position of the CT that was used to produce the synthetic fluoroscopy image is known and is used as the reference standard to evaluate the registration performance.

Test volumes were generated from the original 3D CT data using randomly chosen values between -5 and +5 for the 6 motion parameters. For a robust comparison, the registration was performed for 100 such random starting positions. The algorithms under investigation were then used to register the transformed CT images to the fluoroscopy image. Although the initial starting positions were chosen randomly, they were the same for both algorithms. For each starting position, the algorithms are allowed to make 45 iterations (which on average is enough for convergence [1]). If, at the end of 45 iterations, the values of all transformation parameters become less than 1, the algorithm is considered to have successfully converged.

TABLE 1
MEAN AND STANDARD DEVIATION (SD) OF MOTION PARAMETER ERRORS FOR THE PROPOSED AND FULL 3D [1] ALGORITHMS.
(TRANSLATION PARAMETER UNITS ARE MILLIMETERS AND ROTATION PARAMETER UNITS ARE DEGREES)

Motion Parameters	Tibia		Femur	
	3D algorithm Mean/SD	Proposed algorithm Mean/SD	3D algorithm Mean/SD	Proposed algorithm Mean/SD
Tx	0.05/0.01	-0.08/0.00	0.07/0.00	0.07/0.00
Ty	-0.13/0.01	0.00/0.00	0.03/0.00	0.03/0.00
Tz	-0.41/0.06	-0.54/0.04	0.34/0.03	0.41/0.04
Rx	0.12/0.006	-0.00/0.01	0.13/0.01	0.00/0.02
Ry	-0.41/0.019	-0.36/0.01	-0.21/0.01	-0.13/0.03
Rz	-0.04/0.00	-0.01/0.00	0.04/0.00	0.04/0.00

For the tibia, the proposed algorithm was successful for 100% of the initial disparity cases and the success rate of the original “gradient-based with SCV” approach was 98%. For the Femur both the algorithms had a success rate of 96%. Figure 2 shows the cumulative time required for both these algorithms against the number of iterations. From Figure 2 it can be seen that the use of the 2D DRR in the iterative stage renders our algorithm about 5 times faster than the one in [1]. A closer look at Figure 2 reveals that the cumulative time taken by the proposed algorithm is approximately 60 seconds, of which a significant portion was consumed by the last five 3D iterations. However these 3D iterations are required to provide similar accuracy to the full 3D algorithm.

Table 1 shows a comparison of the mean and standard deviation of the proposed and full 3D approaches for the tibia and femur. From Table 1 we can see that the proposed algorithm provides similar accuracy to the full 3D “gradient-based with SCV” algorithm. For the T_z parameter, the proposed algorithm provides slightly less accuracy than the full 3D algorithm.

IV. CONCLUSIONS

In this paper we have described a new computationally efficient 2D-3D image registration algorithm based on gradient-descent optimization. The proposed method operates only on a 2D projection image in the iterative optimization instead of the 3D volume as used in other approaches in the literature. The algorithm directly calculates the DRR for the updated out-of-plane rotations and does not require the offline generation of sample 2D projections.

For registration accuracy, we have also shown a suitable way of incorporating 3D volumes in the procedure with a minimal effect on the time efficiency. The registration error of the resulting algorithm is not significantly greater than the errors reported for the best recently proposed techniques in the literature.

V. REFERENCES

- [1] M. Pickering, A. Muhit, J. Scarvell, and P. Smith, "A new multi-modal similarity measure for fast gradient-based 2D-3D image registration," in *Proc. IEEE Int. Conf. on Engineering in Medicine and Biology (EMBC)*, Minneapolis, USA, 2009, pp. 5821-5824.
- [2] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Medical Image Analysis*, vol. 3, pp. 373-386, 1999.
- [3] L. Zollei, E. Grimson, A. Norbash, and W. Wells, "2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators," in *Proc. Computer Vision and Pattern Recognition, CVPR*, vol. 2, pp. 696-703, 2001.
- [4] J. Pluim, J. Maintz, and M. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE transactions on medical imaging*, vol. 22, pp. 986-1004, 2003.
- [5] F. Wang and B. Vemuri, "Non-rigid multi-modal image registration using cross-cumulative residual entropy," *International Journal of Computer Vision*, vol. 74, pp. 201-215, 2007.
- [6] D. Russakoff, T. Rohlfing, K. Mori, D. Rueckert, A. Ho, J. Adler Jr, and C. Maurer Jr, "Fast generation of digitally reconstructed radiographs using attenuation fields with application to 2D-3D image registration," *IEEE transactions on medical imaging*, vol. 24, pp. 1441-1454, 2005.
- [7] J. Weese, R. Goecke, G. O. Penney, P. Desmedt, T. M. Buzug, and H. Schumann, "Fast voxel-based 2D/3D registration using a volume rendering method based on sharp-warp factorization," *Proc. SPIE 3661*, pp. 802-810, 1999.
- [8] H. Jans, A. Syme, S. Rathee, and B. Fallone, "3D interfractional patient position verification using 2D-3D registration of orthogonal images," *Medical physics*, vol. 33, pp. 1420-1439, 2006.
- [9] A. Lujan, J. Balter, and R. Ten Haken, "Determination of rotations in three dimensions using two-dimensional portal image registration," *Medical physics*, vol. 25, pp. 703-708, 1998.
- [10] J. Scarvell, M. Pickering, and P. Smith, "New registration algorithm for determining 3D knee kinematics using CT and single-plane fluoroscopy with improved out-of-plane translation accuracy," *Journal of Orthopaedic Research*, vol. 28, pp. 334-340, 2009.