3D-2D IMAGE REGISTRATION BY NONLINEAR REGRESSION

Ana R. Gouveia[1], Coert Metz[2], Luís Freire[3], Stefan Klein[2]

[1]CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Covilhã, Portugal
[2]Biomedical Imaging Group Rotterdam, Erasmus MC, Rotterdam, the Netherlands
[3]Escola Superior de Tecnologia da Saúde de Lisboa, Instituto Politécnico de Lisboa, Lisboa, Portugal

ABSTRACT

We propose a 3D-2D image registration method that relates image features of 2D projection images to the transformation parameters of the 3D image by nonlinear regression. The method is compared with a conventional registration method based on iterative optimization. For evaluation, simulated X-ray images (DRRs) were generated from coronary artery tree models derived from 3D CTA scans. Registration of nine vessel trees was performed, and the alignment quality was measured by the mean target registration error (mTRE). The regression approach was shown to be slightly less accurate, but much more robust than the method based on an iterative optimization approach.

Index Terms **—** 2D/3D image registration, image guided interventions, regression

1. INTRODUCTION

Image guided interventions often require integration of preoperative image with intraoperative images A well-known example is the registration of preoperative 3D CT images and intraoperative 2D X-ray images. Most methods in the literature are based on simulated X-ray projection images – digitally reconstructed radiographs (DRRs) – computed from the preoperative CT scan. In these methods, the image registration is estimated by iteratively optimizing a similarity measure, measuring the difference between the DRR and the X-ray image [1]. Due to local maxima of the similarity measure, such iterative optimization procedures usually have a small capture range and therefore require initialization close to the searched pose [1,2].

In this paper, we present a robust 3D/2D registration method that addresses the capture range drawback. Our approach treats image registration as a nonlinear regression problem instead of using an iterative traditional approach. The nonlinear regression function is determined in a supervised learning stage and relates a DRR image to the 3D transformation parameters of the 3D object. To the best of our knowledge, no such approach to the 3D/2D registration problem has been previously presented in the literature. The method is quantitatively evaluated, and compared with a conventional iterative optimization based registration method [3].

2. METHODS

2.1. Registration by regression

Our proposed 3D/2D registration method consists of a nonlinear regression model that relates image features of the 2D projection image to the (translation and rotation) transformation parameters of the 3D image. A training phase takes place before the intervention (e.g. surgery), during its planning, when data are available and time constraints are not an issue. After the learning process, a nonlinear function is known and we are able to compute the transformation parameters for the 2D images acquired during the intervention.

The training set required for the learning process of our approach is a set of simulated 2D images (DRRs) obtained by manual transformations of the pre-interventional 3D image followed by projection. The features extracted from the DRR and its corresponding transformation parameters form an input-output pair in the training set. During the intervention, the image features of the 2D projection image are computed and fed as input to the regression function, which returns the estimated 3D translation and rotation parameters of the 3D image.

The formulation we propose for the 3D/2D registration is similar to the viewpoint determination problem described in [4], where the solution is determined by an interpolation function considering input-output examples from 2D perspective projection of a 3D scene.

2.2. Input features

In this work we consider two different sets of image features. The first set consists of the three 2D (geometric) moments of the images [5]. The moment of order zero, $\{M_{00}\}$, represents the total mass of a given image (or the area in the case of a binary image). The first order moments, $\{M_{01},M_{10}\}$, representing the centre of mass of the image, yield information about the object position; whereas the second order moments, $\{M_{02},M_{11},M_{20}\}$, also known as moments of inertia, contribute for its orientation.

The second set of features considered consists of the eigenvalues and eigenvectors computed from a Principal Component Analysis (PCA) on the pixels of the object of interest after a coarse segmentation of the image into object and background objects. The PCA was performed in two ways (which will be compared in the Experiments and Results section): a) on the 2D pixel position vectors (x,y) and b) on a combination of the pixel locations and their corresponding intensity values, i.e., a 3D vector with x, y and $I(x,y)$ as variables, where $I(x,y)$ is the intensity value of the point at position (x,y). In both cases, PCA was preceded by computing the z-score of the features, where we used an identical mean and standard deviation for x and y, to not loose pose information in this normalization procedure. The normalization is necessary since the intensities $I(x,y)$ have a different unit than the pixel positions x and y.

2.3. Nonlinear Regression Model

As a regression model we use Neural Networks (NN), which have a great ability of modelling complex (nonlinear) functions and are considered universal approximators [6]. The central idea of this "nonlinear statistical model" is to extract linear combinations of the inputs as derived features, and then model the output as a nonlinear function of these features [7].

By applying a set of labeled training input-output examples (i.e., a set of inputs and corresponding desired responses), a neural network learns by the modification of its processing units or synaptic weights. The synaptic weights are optimized such that the difference between the desired response and the actual response of the network is minimized.

3. EXPERIMENTS AND RESULTS

3.1. Imaging Data

In this work we focused on the registration of 3D preoperative coronary CTA and 2D intraoperative X-ray angiography. To evaluate the new registration method in a controlled setting, we do not use real X-ray images in our experiments, but simulated projection images of the coronary vessel tree, with known ground truth transformation. To this end, we made coronary segmentations at end-diastole of ten patients [3] to obtain binary vessel tree models. From these 3D models, DRRs were generated using the computation procedure described in [8]. The projection geometry for the computation of the DRRs as well as the initial orientation of the preoperative data were derived from an interventional X-ray image, thereby mimicking a clinically relevant view. The field of view and voxel size of CTA were 256x256 voxels and 0.7x0.7x0.9 mm^3, respectively, and the size and voxel size of the DRR images were 512x512 voxels and 0.22x0.22 mm^2, respectively.

For each patient, 11000 DRRs were generated, 10000 to train the regression model and 1000 to test the performance of our method. The transformations were drawn from a uniform distribution, with a wide yet relevant range, i.e. between -10 and 10 degrees for rotations and between -10 and 10 mm for translations.

3.2. Implementation details of features extraction and regression model

For the PCA features, in order to select the points in the object, we segmented the DRR using a threshold value manually chosen. It should be noted that this step was trivial, since we used simulated projections of the vessel tree, which do not suffer from vessel-like structures in the background. With real X-ray images, this is a larger challenge, not yet addressed in this paper.

The NN was designed following the general recommendations in the literature [6,9]. Topology used is feed-forward, specifically a Multi-Layer Perceptron (MLP) with one hidden layer. The training algorithm is backpropagation (the generalization of the Widrow-Hoff learning rule) with conjugate gradient as the optimization technique. The weights are randomly initialized within small ranges ([-1,1]).

The activation functions used are the sigmoid (hyperbolic tangent) and linear functions for the hidden and output layers, respectively. For the hidden layer, we tried several numbers of units in initial trial-and-error experiments on the first patient (see Section 3.5) to find the best option. The output layer has one output unit per registration parameter, following general recommendations, which means our problem has one MLP per transformation parameter, i.e. our registration tool comprises six independent MLPs. All units are fully connected with each unit of the next layer.

For a proper use by the NN, the input and output vector elements are rescaled [6], to a range between 0 and 1 and between -1 and 1, respectively.
The Neural Network was implemented in MatLab 7.11.0.584.

3.3. Conventional 3D/2D registration method

The registration method described in the present work was compared to the method based on iterative optimization proposed in [3], which was designed for the application considered. This method uses a nonlinear conjugate gradient optimizer and a similarity metric based on a distance transform of a projection of the 3D coronary segmentation onto the X-ray image, and a fuzzy segmentation of vessel structures in the 2D image. As with our method, this fuzzy segmentation step became trivial, because we only used DRRs of the binary vessel tree, instead of real X-rays.

3.4. Evaluation methodology

The evaluation of the registration approaches was performed by the computation of the mean target registration error (mTRE) before and after registration. This enables us to assess registration success or failure and the consequent capture ranges for both registration approaches. The mean target registration error (mTRE) is computed as the mean distance to the ground truth (pose of the centerline tree at the known transformation parameters) of all center points of the 3D vessel centerline trees:

$$mTRE(T) = \frac{1}{N}\sum_{n=1}^{N}\left\|T(p_n) - T_{gold}(p_n)\right\| \qquad \text{(Eq. 3.1)}$$

where T is the resulting transformation of one of the registration methods to assess, T_{gold} is the known transformation and p_n are points on the centerline of the 3D vessel tree.

3.5. Optimization and parameter settings for regression based registration

Prior to the evaluation of the registration method proposed, some experiments were performed to optimize the method. To this end, the image set of one of the patients was used, whereas the sets of the nine remaining patients were used for the evaluation in Section 3.6. For every training and testing process, two datasets were randomly drawn, one with a maximum of 10000 images used to train and validate (using 70% and 30% of the data, respectively). The other dataset (#1000), composed by unseen data, was used for evaluating the NN's generalization ability [10].

Based on these tuning experiments, we set the number of hidden units to twice the number of input units. The stopping epoch (i.e., a stopping criteria for training of each MLP) was tuned separately for each MLP. The effect of using different features (Section 2.1), individually and combined, was investigated on this training set as well. Table 3.1 summarises the results. Individually, Set 1 was the best contributor to the registration, being especially valuable for finding the translation parameters. Though Set 2 gave the worst results individually, when combined with Set 1 it helped to decrease the registration errors. The configuration Set1+Set3 was found to be the optimal combination of feature sets. In this configuration, the input layer of the neural network has 18 units: 6

TABLE 3.1. Registration results for the feature sets considered: moments set (1), the eigenvalues and eigenvectors of the points in the object of interest (2) and the eigenvalues and eigenvectors set for the same points and corresponding intensity values. Parameters errors are given for rotations and translations (R_A, R_B and R_G – rotations in degrees around x-, y- and z-axis, respectively, and Tx, Ty and Tz – translations in x, y and z directions in mm, respectively), mean TRE errors (mTRE) for all registrations performed (of the entire testing set) and just considering the successful registrations (i.e. mTRE<2mm) are presented.

	Set 1		Set 2		Set 3		Set 1 + Set 2		Set 1 + Set 3	
	Moments		Eigen [*X Y*]		Eigen [*X Y I*(*x*,*y*)]					
	mean	std	mean	std	mean	std	mean	std	mean	std
R_A (º)	0,29	0,29	3,01	2,31	0,23	0,36	0,13	0,15	0,05	0,05
R_B (º)	0,30	0,26	2,84	2,14	0,36	0,43	0,12	0,12	0,07	0,08
R_G (º)	0,41	0,39	1,79	1,47	0,25	0,22	0,14	0,18	0,07	0,09
T_X (mm)	0,15	0,12	4,86	2,94	2,68	2,45	0,04	0,06	0,02	0,03
T_Y (mm)	1,36	1,54	5,04	2,87	5,02	2,88	0,65	1,08	0,33	0,52
T_z (mm)	0,12	0,11	4,70	3,00	0,73	0,82	0,07	0,08	0,03	0,04
mTRE (mm)	1,45	1,55	9,76	3,02	6,31	2,85	0,69	1,09	0,35	0,55
mTRE (mm) (mTRE<2mm)	0,87	0,50	1,70	0,37	1,34	0,46	0,48	0,37	0,30	0,29

moments, 3 eigenvalues and 3x3 components of the eigenvectors. This combined feature set was used in all following experiments.

From Table 3.1 it can also be observed that a problematic transformation in the 3D/2D registration is the translation along the source-detector axis. Since the projection is in the y-axis direction, the corresponding translation parameter has less influence on the appearance of the projected model than the other parameters [11]. Reduced performance in this direction has also been reported for the conventional registration approach described in Section 3.3 [3]. Additional tuning experiments were performed to understand how large should be the training set. Four training set dimensions were considered and the results presented in Fig. 3.1. Registration quality improved slightly with increasing training size. Since the training is a preoperative task, our selection criteria was the accuracy and robustness of the registration and we used the largest training set.

3.6. Test results

The results of the 3D/2D experiments for all patients and for both methods are summarized in Fig. 3.2 and Fig. 3.3.

Considering this registration results, the regression method was less accurate for the majority of datasets compared to the iterative optimization based registration method, with a median error larger for 7 of the 9 datasets (Fig. 3.2). However, the number of outliers for both approaches clearly showed a higher resistance to outliers for the regression approach (Fig. 3.2). That robustness of the regression method is also stressed in Fig. 3.3 where the mTRE values before registration against the mTRE values after registration are shown.

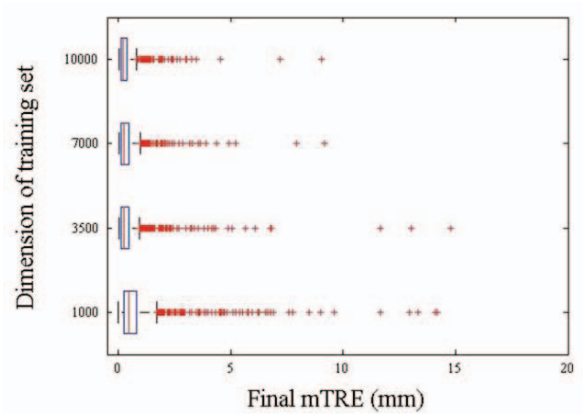


Figure 3.1 Registration results for different dimensions of the NN training set. Mean TRE errors (mTRE) for registrations performed are presented.

4. DISCUSSION

The results presented in this work show that the proposed regression-based 3D-2D registration approach remains robust in case of large misalignments. It could be used either stand-alone, or as an initialization step for a conventional 3D-2D registration to improve the final accuracy.

Our method is based on image features but not in the common meaning of being a point-to-point, curve-to-curve or surface-to-curve registrations. Instead, the correspondences are established more like library-based registrations where a library 2D template is generated for a predefined set of transformation parameters. Each library template represents the expected 2D appearance of a 3D geometric feature for a particular transformation. The template that is most similar to the 2D geometric feature is then aligned with the 2D data to define the remaining unknown transformation parameters. In 3D/2D registration, templates like 2D silhouettes [12], Fourier descriptors of a 2D silhouette [13] and shock graph representation of the projected 3D shape [14] are found in the literature, with different searching and/or matching schemes.

In the literature some other work can be found with a registration approach using a neural network (e.g. [15], [16], [17], [18], [19]). However, the problem formulation is different (i.e. the function to be determined does not relate the same input-outputs that we do), the applications are different and most of them do not compare the results with a conventional method. The most similar approach [19] presents a method designed for 3D rigid-body registration of fMRI time series, which relies on a limited number of Fourier coefficients of the images to be aligned.

5. CONCLUSION

We have proposed a novel approach for 3D/2D image registration, based on nonlinear regression. Promising results were demonstrated for the registration of coronary vessel trees. Compared with a conventional approach, the number of misregistrations was reduced substantially. Future work includes experiments using real X-ray images and the implementation and evaluation of different regression (other than NN).

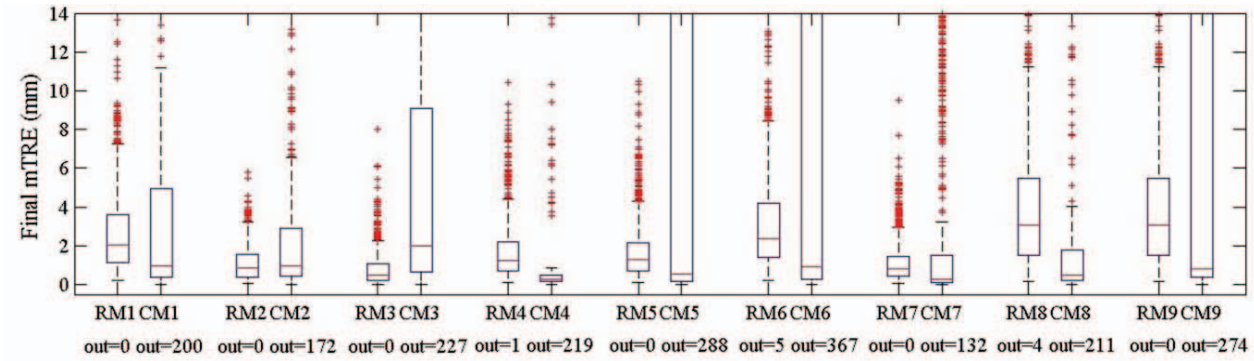


Figure 3.2 Comparison of the registration method by regression (RM) to the conventional registration method (CM), for each patient (labeled from 1 to 9). The number of points with mTRE > 14mm that fall outside the plot is indicated by "out".

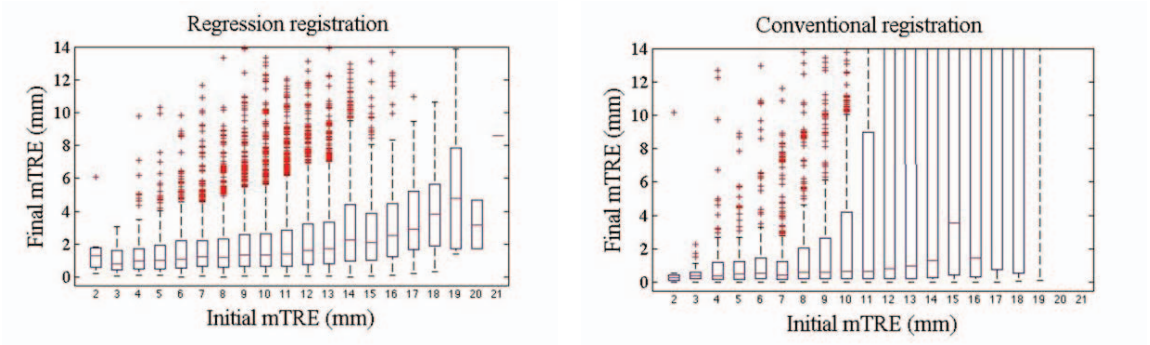


Figure 3.3 Comparison of the registration method by regression to the conventional registration method, considering the results for all patients. The plot shows the mTRE values before registration against the mTRE values after.

5. REFERENCES

[1] Markelj, P., et al. A review of 3D/2D registration methods for image-guided interventions. Med. Image Anal. (2010).

[2] van de Kraats, E.B., et al. Standardized evaluation methodology for 2-D–3-D registration. IEEE Trans. Med. Imag. 24 (9): 1177–1189, 2005.

[3] Metz, C.T., et al. Alignment of 4D Coronary CTA with Monoplane X-Ray Angiography, MICCAI Workshop Augmented Environments for Computer Assisted Interventions,2011. URL: http://www.bigr.nl/publication/694.

[4] Liang, B., and Chung, R. Viewpoint determination of image by interpolation over sparse samples. Journal Image and Vision Computing, 26 (7): 941-954, 2008.

[5] Liao, S.X., Image analysis by moments, PhD Thesis, 1996.

[6] S. Haykin, Neural Networks, A Comprehensive Foundation, 2nd Ed: Prentice-Hall, 1999.

[7] Hastie, T., et al. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Ed: Springer, 2009.

[8] Metz et al. GPU accelerated alignment of 3-D CTA with 2-D X-ray data for improved guidance in coronary interventions. Proceedings of IEEE International Symposium on Biomedical Imaging: from Nano to Macro, 2009.

[9] Sarle, W. Neural network frequently asked questions, ftp://ftp.sas.com/pub/neural/FAQ.html. 2005

[10] Flexer, A. Statistical evaluation of neural networks experiments: Minimum requirements and current practice. In Proceedings of the 13th European Meeting on Cybernetics and Systems Research, vol 2, pp. 1005–1008, Austrian Society for Cybernetic Studies, Vienna (1996).

[11] Edwards P.J., et al. Medical Image Registration – Chapter12: Guiding Therapeutic Procedures. CRC Press, 2001.

[12] Hermans, J.et al. Robust initialization for 2D/3D registration of knee implant models to single-plane fluoroscopy. In Medical Imaging 2007:Image Processing, vol. 6512. SPIE, San Diego, CA, USA, p. 651208.

[13] Banks, S.A. and Hodge, W.A. Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. IEEE Trans. Biomed. Eng. 43 (6), 638–649. 1996

[14] Cyr, C. et al. 2D–3D registration based on shapematching. In: IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2000). IEEE Computer Society, Hilton Head Island, SC, USA, pp.198–203. 2000.

[15] Elhanany et al "Robust Image Registration based on Feedforward NN, In 2000 IEEE SMC, 2000.

[16] Yan et al. A Neural Network approach for 3D surface modeling and registration. In IEEE IWBCS, 2004.

[17] Zhang et. al. Rapid surface registration of 3D volumes using a neural network approach. Image Vision Comput. , vol. 26 pp. 201-210, 2008.

[18] Abche et al. Image Registration based on Neural Network and Fourier Transform. In Proceedings of the 28th IEEE EMBS, New York City, USA, 2006.

[19] Freire, L.C., et al. FMRI 3D registration based on Fourier space subsets using neural networks. In IEEE EMBC, 2010.