Template-based CTA X-ray Angio Rigid Registration of Coronary Arteries in Frequency Domain

Timur Aksoy[a], Stefanie Demirci[c], Muzaffer Degertekin[b], Nassir Navab[c], Gozde Unal[a]

[a]Sabanci University, Tuzla, Istanbul, Turkey
[b]Yeditepe University Hospital, Istanbul, Turkey
[c]Technical University of Munich, Germany

ABSTRACT

This study performs 3D to 2D rigid registration of segmented pre-operative CTA coronary arteries with a single segmented intra-operative X-ray Angio frame in both frequency and spatial domains for real-time Angiography interventions by C-arm fluoroscopy. Most of the work on rigid registration in literature required a close initialization of poses and/or positions because of the abundance of local minima and high complexity that searching algorithms face. This study avoids such setbacks by transforming the projections into translation-invariant Fourier domain for estimating the 3D pose. First, template DRRs as candidate poses of 3D vessels of segmented CTA are produced by rotating the camera (image intensifier) around the DICOM angle values with a wide range as in C-arm setup. We have compared the 3D poses of template DRRs with the real X-ray after equalizing the scales (due to disparities in focal length distances) in 3 domains, namely Fourier magnitude, Fourier phase and Fourier polar. The best pose candidate was chosen by one of the highest similarity measures returned by the methods in these domains. It has been noted in literature that these methods are robust against noise and occlusion which was also validated by our results. Translation of the volume was then recovered by distance-map based BFGS optimization well suited to convex structure of our objective function without local minima due to distance maps. Final results were evaluated in 2D projection space rather than with actual values in 3D due to lack of ground truth, ill-posedness of the problem which we intend to address in future.

Keywords: 2D-3D Rigid Registration, Data Integration for the Clinic/OR, Image-Guided Therapy, C-arm Angiography Intervention, X-ray to CTA Registration

1. INTRODUCTION

The aim of 2D-3D rigid registration of pre-operative CTA coronary arteries with intra-operative X-ray Angio frames is to present the physicians locations of corresponding CTA vessels during C-arm Angiography interventions. The availability of CTA vessel correspondence information during the X-ray acquisition may help the cardiology specialist in analysis and diagnosis of plaques, atherosclerosis, other diseases within the vessels.

2D-3D registration of medical structures is an ill-posed problem because there may be multiple poses in 3D space yielding the similar 2D projections. Estimation of camera and transformation matrices would require many point correspondences between volume and X-ray. Moreover the search space is very large; recovering 3D rotation and translation without correspondences is difficult due to abundance of local minima and high computational complexity.

Feature based intrinsic 2D/3D registration has been performed using either extrinsic markers or extracting anatomical features in intrinsic case. Intrinsic methods require accurate and robust feature extracting algorithms and distinct features for establishing correct correspondences. For large point sets methods similar to Iterative Closest Point (ICP) have been used.[1] Geometric primitives other than points such as curves, surfaces and edges are also corresponded by minimizing their distances with back or forward projections[2].[3] In case of coronary arteries, distinct regions such as bifurcations or prominent geometric structures are not in high quantity, thus enough number of features do not exist.

Intensity based 2D/3D registration uses solely the information of pixels and voxels. Digitally Reconstructed Radiograph (DRR) or Maximum Intensity Projections (MIP) are generated from volume and compared to X-rays by similarity measures such as mutual information, normalized cross correlation, gradient correlation and

sum of squared differences[45].[6] Some authors have compared coefficients from wavelet transform or orthogonal Zernike moment decompositions instead. Intensity based methods would fail in coronary artery registration due differences in modalities and lack of distinctive intensity variations on arteries. DRR based methods also require close initialization due to local minima and suffer from high computational complexity despite speed improvements in new approaches.[7]

Hybrid methods segment the anatomical structures before generating DRRs and uses intensity similarity methods for comparing them with X-rays. Usually the segmented vessel or a surface model such as tetrahedral mesh are projected with DRR to obtain intensities of desired regions or with MIP to obtain binary vessels[89] .

Reconstruction approach requires several X-rays for generating a 3D model to register to the preoperative volume. However as number of fluoroscopic images decrease so does the quality of the reconstructed model. Some authors have proposed a robust similarity technique for comparing poor quality reconstructed model to high quality CTA or MRA[10].[11]

This study performs a library and segmentation based registration where a desired number of rotational poses of CTA vessels are generated and saved after segmentation. 3D translation is then recovered through MIP and binary similarity measure as in hybrid methods since translation can be recovered fast and accurately with distance maps. Rotational poses of segmented vessels are compared to segmented X-ray vessels in Fourier domain in order to remove differences in translation.

2. METHODOLOGY

2D-3D registration is performed on segmented vessels. The most distinct element that determines the position and orientation of vessel projections are their overall shape rather than intensities or features. 3D segmentation was achieved by vessel tractography algorithm in (Cetin et al).[12] This method has successfully segmented the left and right tree coronary arteries. Then with connected components, the tree of interest (left side) was selected because the angio interventions are usually applied on the three main branches in that tree. This segmentation has extracted only the three or four main branches; since catheter interventions are performed only on vessels with large diameters, these results are sufficient.

2D segmentation was started with the Frangi filter preceded by smoothing Gaussian convolution.[13] However in the output image, artifacts due to undesired intensity variations, responses from non-vascular structures with thin and long shapes were present. We have applied in order median filter (to erase small artifacts like salt and pepper noise), histogram thresholding, morphological closing and opening (to eliminate thin vessels and small noises). Then the image was thresholded to obtain the binary map. The Catheter tube which resembles vessel was removed manually.

The registration is rigid with 6 degrees of freedom. The 6 variables to be found are the intensifier or camera's 3 rotation angles around the volume's center and the volume's translation on world coordinate x-y-z axes. x and y axes form the coronal plane of the patient. Intensifier's rotation axes are C-arm detector's primary (RAO to LAO), secondary (CRAN to CRAUD) angles and its optical axis. Thus a geometric setup resembling an actual C-arm has been constructed in digital environment to simulate X-ray images.

In 2D, the similarity spatial transform between the images p and r can be expressed as:

$$r(x,y) = p(\delta x + s(xcos(\theta) - ysin(\theta)), \delta y + s(xcos(\theta) + ysin(\theta))) \quad (1)$$

where $\delta x, \delta y$ are translation distances, θ is rotation angle, and s is the scaling factor.

2.1 Preprocessing

The preprocessing phase aims to determine the camera or intensifier's rotation angles accurately. As prior information for rotation angles, the DICOM header values in CTA and X-ray have been used since both CTA and C-arm devices were calibrated accurately. The values used in DICOM headers are source-patient distance, source-detector distance, intensifier's primary and secondary angles, CTA and X-ray's pixel spacings and the patient's lying position. Because DICOM header values are not entirely reliable due to patient and intensifier's motion during filming, they are used only as initial parameters before preprocessing.

This study estimates the rotation angles of the camera from a single X-ray frame where the vessel structure is most visible. For preregistration, binary DRR samples are created by rotating the camera around x-y-z axes in a specific angle range centered around the DICOM parameters and then saved as templates. The goal is to find the template or the candidate DRR which is most similar to the real X-ray image in terms of the rotation pose. But since 3D translation difference between DRR and real X-ray image is not known, the methods comparing DRRs and the X-ray have to be invariant to translation.

It is well known in literature that Fourier magnitude or power spectrum of a digital signal is invariant to translation and large amount of shifts in plane add only noise to the spectrum.[14] Discrete Fourier transform assumes that finite signals are periodic with their sizes as if the signal repeats itself. So DFT acts as if the opposite edges of the image are actually neighbors. This may lead to artifacts in the transform with the shape of a plus sign. If the vessel in an image is touching an edge as in case of clipping, Fourier transform will respond as if that edge continues at the other end and the vessel structure changes abruptly. Thus there will be high frequency artifacts due to this unintended abrupt change. To prevent that, a low-pass, blurring and circular Gaussian filter is applied to remove sudden cutoffs at edges and decrease the values of high frequency components that account for small details and noise. The relation between p and r images in Fourier domain is as follows:

$$F_r(w_x, w_y) = e^{j2\pi(w_x\delta x + w_y\delta y)/s} s^2 F_p((w_x cos(\theta) + w_y sin(\theta))/s, (w_x cos(\theta) - w_y sin(\theta))/s) \tag{2}$$

The first method returns the DRR candidate whose Fourier magnitude is most similar to the real X-ray's as the DRR with the closest rotation pose after equalizing the scales of DRRs and X-ray in 2D. In order to equalize the scales, the 0th moments i.e. the areas of binary vessels, then their ratios are computed. The scale difference is removed by inverse scaling of the real X-ray's Fourier Magnitude transform as seen in the equation. The reason of equalization of scales in Fourier rather than spatial domain is that the enlarged vessels may move out of the frame in space resulting in loss of the field of view. Because scale and translation differences are now eliminated, the dissimilarity between two spectra can only be due to the rotation pose.

The second method for finding the correct pose is analysis of Fourier Phase relations of DRRs and real X-ray image to conclude whether there are any differences other than pure translation. Images p and r having only 2D translation motion $\delta x, \delta y$ between them satisfy the following relation:

$$Phase_r(v_x, v_y) = 2\pi(v_x\delta x + v_y\delta y) + Phase_p(v_x, v_y) \tag{3}$$

Since the power spectra of these images are the same, their cross correlation in Fourier domain R_{rp} would yield only the phase difference and can be expressed as:

$$R_{rp}(w_x, w_y) = e^{j2\pi(w_x\delta x + w_y\delta y)} \tag{4}$$

The inverse Fourier transform of R_{rp} is a delta dirac function translated by $(\delta x, \delta y)$.

Fourier phase correlation of two same shapes differing only by translation would yield a delta dirac function translated by same amount in the spatial domain.[15] If such output is obtained then it will be assumed that the shapes in the binary images differ only by translation. In order to analyze whether this output is a delta dirac function, 1 is subtracted from the largest real element and then the Frobenius norm is computed. Frobenius norm measures the distance of all elements of a matrix from zero. In this case, a lower Frobenius norm means that the output is nearly a null matrix and closer to the delta dirac function. Frobenius norms of scale equalized Fourier phase correlations were computed for all DRRs. The DRR candidate with the lowest Frobenius norm is selected as the best candidate for closest rotation pose. An advantage of this method is that phase correlation in frequency domain is robust to noise and occlusions.

For the third method, Fourier power spectrum has been mapped into the polar space and integrated along the radial axis (r). In similarity transform, rotation and scaling in spatial domain result in negative rotation and inverse scaling in Fourier domain respectively. In Fourier Polar domain, rotation would reflect as translation in

angle axis (θ), and scaling would reflect as inverse scaling multiplied by a constant amplitude coefficient.[16] This means rotation and scaling are decoupled into separate axes in this domain. The radius axis is integrated to obtain a 1D signal invariant to scale and translation differences. This signal, called rotation signature stores the rotation angle of a shape as a shift. The amplitude coefficients in the rotation signature due to scale difference don't need not be normalized since normalized correlation between two signals are not effected by constant scaling (variance) and mean values of the signals. The highest rotation signature correlation indicates the closest rotation pose.

The integration of Polar Fourier Transform signal along the radius axis (r) is expresses as:

$$\kappa(\theta) = \int M(rcos(\theta), rsin(\theta))dr \tag{5}$$

where $M(w_x, w_y) = |F(w_x, w_y)|$ is the Fourier Magnitude.

The Euler angles were recovered from the saved angle of the candidate DRR returned by these methods.

2.2 Similarity Measures for Preprocessing

Three similarity measures have been obtained from the three methods explained in previous section. All binary candidate DRRs have been compared with the segmented X-ray to find the closest rotation pose by the following individual and combined similarity measures:

1. Fourier power spectrum correlation
2. Integrated Fourier Polar signal correlation
3. Frobenius Norm of Inverse Fourier phase correlation
4. Weighted average of measures 1 and 2
5. First measure divided by the third
6. Fourth measure divided by the third

Among these measures, results showed that the fourth measure always returned the closest pose if the main vessels in X-ray images are not occluded or clipped to a great extent. If a large part of the CTA vessels is not visible in X-ray, then the second measure yielded the closest result. In experiments, the second measure was used only once due to occlusion.

Only a part of the whole vessel structure in segmented CTA may be visible in X-ray image due to vessels not covered by field of view, the contrast agent not penetrating to all ROI branches or rarely occlusion by other tissues. As the rotation pose estimation takes place in the frequency domain and depend on statistical correlation measures, they are robust against occlusion, clipping to a certain degree. When scale ratio cannot be computed accurately in such cases, since the first and third measures require scale difference, the second i.e. measure rotation signature measure should be used. In this study, among the frames in X-ray video sequence, the one with the highest visibility of the vessels is selected, therefore occlusion was not frequent.

2.3 Optimization for Translation

The rotation pose and the scale differences have been recovered in the previous section. In C-arm perspective projection geometry, the scale difference is proportional to square root of the source to detector distance (focal length). It is assumed that the scale difference results from the erroneous measurement of this distance. Euclidean distance maps of binarized CTA volume and X-ray image are loaded into the program we have developed with QT,[17] VTK[18] and ITK.[19] Then the best DRR candidate's rotation angles and the corrected source to detector distance are entered. Other DICOM parameters are not altered.

Translation-only optimization is simpler since the minima occurs only when the shapes are aligned on top of each other. The objective function value needs to decrease as the overlap measure between the shapes increase.

In order to provide monotonicity and continuity for gradient-based optimization, the distance maps of binary vessels in 2D and 3D have been formed and registered for translation. In the distance maps, the values inside the vessels are 0 and background pixels take the value of the shortest millimeter distance to the vessel structure in single precision floating point format.

Gradient based BFGS optimization algorithm[20] is used for translation-only registration. First order derivatives are computed by central finite difference scheme instead of analytically. Since distance maps turn the function into a convex form, finite differencing does not degrade the accuracy.

During optimization, MinIP (Minimum Intensity Projection) instead of DRR in VTK library[18] has been used to project the 3D distance map. The objective function to be minimized is the total shortest distance of all 2D vessel pixels to the projected 3D vessel plus the total shortest distance of all projected 3D vessel pixels to the 2D vessel plus the overlap ratio of both vessels. The objective function is written as:

$$E(T) = \frac{\int \phi(I_1) I_2 d\mathbf{x}}{\int \phi(I_1) d\mathbf{x}} + \frac{\int \phi(I_2) I_1 d\mathbf{x}}{\int \phi(I_2) d\mathbf{x}} - \lambda \frac{2 \int \phi(I_2)\phi(I_1) d\mathbf{x}}{\int \phi(I_1) d\mathbf{x} + \int \phi(I_2) d\mathbf{x}} \tag{6}$$

where $I_1 = P(T(D))$, D is the 3D distance map undergoing rigid transform T, $P(\cdot)$ is the MIP operator, I_2 is the distance map of the real X-ray image, λ is a regularization constant and $\phi(.)$ is a one-sided delta dirac function expressed as:

$$\phi(y) = \begin{cases} 1 & if y = 0 \\ 0 & if y > 0 \end{cases}$$

3. RESULTS AND EXPERIMENTS

Registration was performed on both synthetic and real X-ray images. First candidate DRRs were produced for both sets. The range of angles and sampling rate effects the success of registration. In this study, in order to show that the true pose may be recovered even when the DICOM angles in C-arm are far from the actual ones, the rotation range was kept large and sampling rate was kept moderate.

In real X-ray image experiment, the camera has been rotated around the DICOM parameters along RAO-LAO direction (primary angle) and CRAN-CAUD (secondary angle) direction with 70 degree range centered around the initial position. Around the z-axis (optical axis), the camera was rotated half of that range i.e. 35 degrees since intensifier's rotation around its optical axis is not a common occurrence and could happen in small amounts due to errors and patient displacement. The sampling rate was 10 for primary and secondary angles and 5 for z-axis. In the discretized angle space in the given range, all combinations of angles were projected. Therefore $10 \cdot 10 \cdot 5 = 500$ candidate DRRs were produced. One such DRR set was formed for every real X-ray frame.

In synthetic set, the range was increased to 90 for primary and secondary angles, to 45 degrees around the optical axis in order to evaluate the accuracy of the preprocessing phase. The sampling rates were 15 for primary and secondary angles, 7 for optical axis. This corresponds to 1575 DRR candidates.

3.1 Validation on Synthetic Dataset

As a synthetic volume, a T-shape with 0.4, 0.4 and 0.5 x-y-z pixel spacings was created. Instead of the real X-rays, 5 DRR projections of the volume in different scales and angles have been produced. The volume was translated by the same amount in all projections. The actual angles are within the range of DRR candidates. Three of the results are in Figure 1.

In the synthetic shape, despite 2D alignments are nearly perfect, results in Table 1 indicate that actual rotation angles deviate from the experiment results. The reasons are explained in Section 4.

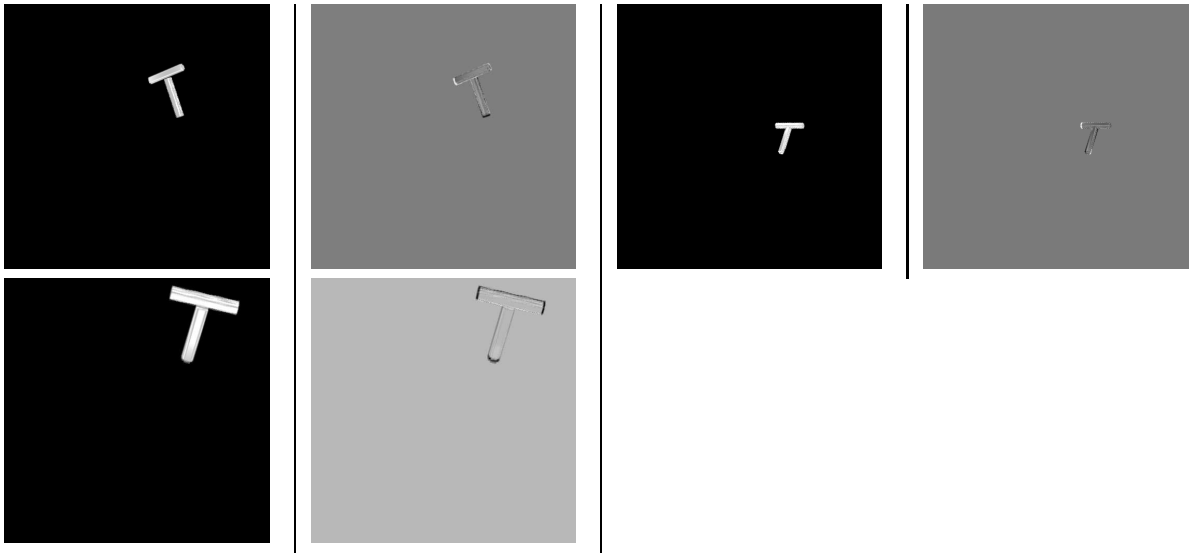

Figure 1. T-shape Registration Results: DRR projections and Difference Images

T-shape Image Number	Actual Rotation Values (x,y,z)	Actual Translation Values (x,y,z)	Rotation Values Found	Translation Values Found	Objective Function Value
1	-6, 33, 18	12, 14, -23	-15, 15, 19	23.67, 12.32, -2.29	-26.90
2	-32, 33, -19	12, 14, -23	-33, 27, -16	21.20, 3.38, -8.10	-26.70
3	39, -5, -16	12, 14, -23	39, -5, -16	11.89, 22.05, -15.88	-25.98

Table 1. Numerical Results of T-shape Registration

3.2 Validation on Clinical Dataset

Clinical experiments were implemented on 7 real X-ray images of 3 different patients. The methodology is same as of the synthetic dataset. A subset of the results are in Figure 2.

Unlike the synthetic shape, real X-ray images do not have ground truth values of the actual spatial transforms. Ground truth for real X-rays can be obtained by putting markers in patient's anatomy and precise measurements. As such data was not available, the evaluation was done only through visual inspection and 2D pose dissimilarities.

4. DISCUSSION AND CONCLUSIONS

In synthetic dataset, it is observed that preprocessing phase recovered the closest poses of all 5 projections and distance map based BFGS algorithm overlaid the shapes on top of each other. As a measure of accuracy, experimental results can be compared with actual results. However this measure is not suitable for this work due to ambiguity in Euler angles and reduction in dimension meaning more than one angle set can yield the same projection. Depth ambiguity is another source of error. For instance T-shape projection when it is tilted forward with a certain angle could be very close to the pose where it is tilted backwards with the same angle. In order to avoid depth ambiguity, more than one simultaneous projections as in bi-plane or multi-frame Angios are necessary.

In clinical dataset, the results are successful but one of the reasons for imperfect alignment is that some of the vessels in C-arm frames are out of view so the scale differences cannot be computed precisely. Another reason is that vessels undergo non-rigid or non-linear deformation due to breathing and heart beat motion. Because patient's deformation in C-arm frames may be different than in CTA volume, the rigid transform cannot completely align vessels as in the synthetic shape. A way to remedy the heart beat motion disparities is equalization of heart phases in CTA and X-ray. Some C-arm frames also have corresponding heart phase information attached to them. Using this information, the frame with the same heart phase of CTA image can be extracted from the X-ray sequence. Future work includes nonlinear registration studies following our rigid registration technique presented in this paper.

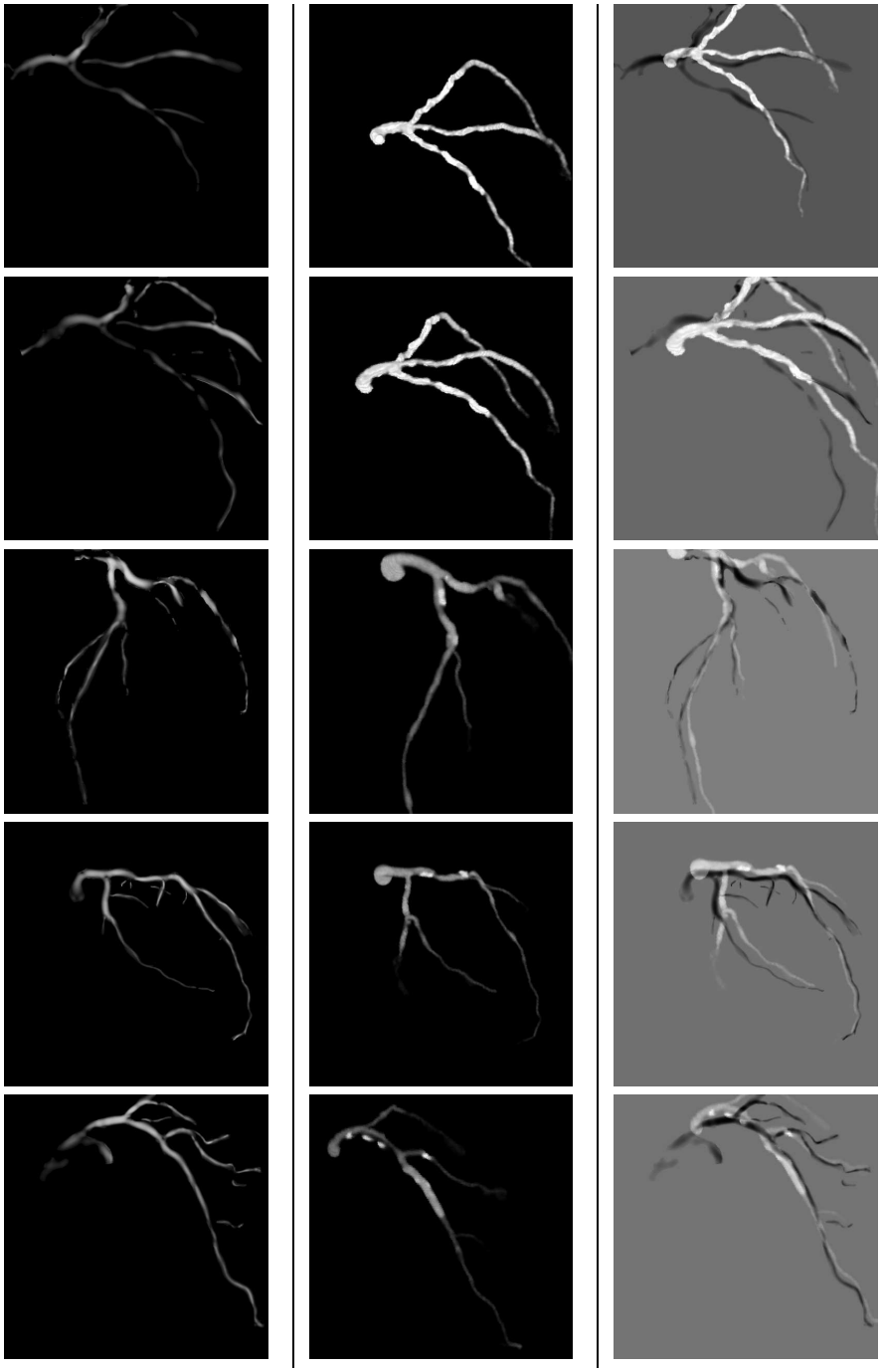

Figure 2. Real X-ray Registration Results: Real X-ray Image, Best DRR candidate and Difference Images

5. ACKNOWLEDGMENT

This project was funded by the Scientific and Technological Research Council of Turkey (TUBITAK) -BMBF Germany Intense Co-operation Grant, project number: 108E162 and Turkish Academy of Sciences Young Scientist Award (TUBA-GEBIP).

X-ray Image Number	DICOM Rotation Parameters (x,y,z)	Rotation Parameters Found	Translation Parameters Found	Objective Function Value
1	-14.7, 45.0, 180	77, 1, 200	12.15, 2.80, 28.40	-4.11
2	44.6, -14.9, 0	59, -19, 12	-9.17, -6.14, -11.62	-5.87
3	-35.7, 41.6, 180	43, 25, 168	3.82, 27.55, -1.40	-6.90
4	-11.7, -29.0, 180	-19, -21, 176	10.80, 9.04, 3.38	-5.10
5	35.1, 1.7, 180	35, 17, 200	-4.96, 11.97, 6.99	-6.96

Table 2. Numerical Results for X-ray Registration

REFERENCES

[1] Besl, P. J. and McKay, N., "A method for registration of 3-d shapes," *IEEE Trans. Pattern Analysis, and Machine Intelligence* **14**, 239 – 256 (1992).

[2] Alperin, N. L. and Pelizzari, D. N., "Retrospective registration of x-ray angiograms with mr images by using vessels as intrinsic landmarks," *Magnetic Resonence Imaging* **4**, 139 – 144 (1994).

[3] Zikic., D., Groher, M., Khamene, A., and Navab, N., "Deformable registration of 3d vessel structures to a single projection image," in [*Medical Imaging 2008: Image Processing, SPIE*], **691** (2008).

[4] Penney, G. P., Weese, J., Little, J. A., Desmedt, P., Hill, D. L. G., and Hawkes, D. J., "A comparison of similarity measures for use in 2d 3d medical image registration," *IEEE Trans. on Medical Imaging* **17**(4), 586–595 (1998).

[5] Zheng, G., "Effective incorporation of spatial information in a mutual information based 3d-2d registration of a ct volume to x-ray images," **5242**, 922–929 (2008).

[6] Maintz, J. B. A. and Viergever, M. A., "A survey of medical image registration," *Medical Image Analysis* **2**(1), 1–36 (1998).

[7] Markelj, P., Tomazevic, D., Likar, B., and Pernus, F., "A review of 3d-2d registration methods for image-guided interventions," *Medical Image Analysis* **16**, 642–661 (2012).

[8] Turgeon, G. A., Glen Lehmann, G. G., Drangova, M., Holdsworth, D., and Peters, T., "2d-3d registration of coronary angiograms for cardiac procedure planning and guidance," *Medical Physics* **32**(12), 3737 – 3749 (2005).

[9] Vermandel, M., Betrouni, N., Gauvrit, J., Pasquier, D., Vasseur, C., and Rousseau, J., "Intrinsic 2d-3d registration based on a hybrid approach: use in the radiosurgical imaging process," *Cell Molecular Biology* **52**, 44 – 53 (2006).

[10] Tomazevic, D., Likar, B., and Pernus, F., "3d-2d image registration: The impact of xray views and their number," **4791**, 450–457 (2007).

[11] Prummer, M., Hornegger, J., Pfister, M., and Dorfler, A., "Multi-modal 2d3d nonrigid registration," in [*Medical Imaging 2006: Image Processing, SPIE*], **6144** (2006).

[12] Cetin, S., Unal, G., Demir, A., Yezzi, A. J., and Degertekin, M., "Vessel tractography using intensity based tensor model," in [*Proceedings of 3rd MICCAI Workshop on Computation and Visualization for (Intra)Vascular Imaging*], 81–88 (2011).

[13] Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A., "Multiscale vessel enhancement filtering," in [*MICCAI*], 130–137 (1998).

[14] McGuire, M., "An image registration technique for recovering rotation, scale and translation parameters," (1998).

[15] Sarvaiya, J. N., Patnaik, S., and Kothari, K., "Image registration using log polar transform and phase correlation to recover higher scale," *Journal of Pattern Recognition Research* **7**, 90 – 105 (2012).

[16] Wolberg, G. and Zokai, S., "Robust Image Registration Using Log Polar Transform," in [*Proc. IEEE Int. Conf. on Image Processing*], (2000).

[17] Nokia, "Qt development framework 4.8.0," (2012).

[18] Kitware, "Visualization toolkit 5.8.0," (2011).

[19] Kitware, "Insight segmentation and registration toolkit 4.1.0," (2012).

[20] Naoaki Okazaki, J. N., "liblbfgs: a library of l-bfgs," (2010).