2D-3D Intensity Based Registration of DSA and MRA – A Comparison of Similarity Measures

John H. Hipwell¹, Graeme P. Penney¹, Tim C. Cox²,
James V. Byrne³, and David J. Hawkes¹

¹ Division of Radiological Sciences, UMDS, Guy's & St Thomas' Hospitals
London SE1 9RT, UK
{John.Hipwell@kcl.ac.uk}

² National Hospital for Neurology and Neurosurgery, Department of Radiology
Queens Square, London, WC1N 3BG, UK

³ Department of Radiology, University of Oxford, The Radcliffe Infirmary, Oxford
Oxford, OX2 6HE, UK

Abstract. We have compared the performance of six similarity measures for registration of three-dimensional (3D) magnetic resonance angiography (MRA) to two-dimensional (2D) x-ray angiography images of the cerebral vasculature. The accuracy and robustness of each measure was investigated using a ground truth registration of a neuro-vascular phantom which was obtained using fiducial markers, and using “gold-standard” registrations of four clinical data sets calculated using manual alignment by a neuro-radiologist. Of the six similarity measures, pattern intensity, gradient difference and gradient correlation performed consistently accurately and robustly for all data sets. Using these similarity measures, and for starting positions within 8° rotation, 3mm in-plane translation and 50mm out-of-plane translation from the gold-standard/ground-truth positions, we obtained a success rate of greater than 80% for the clinical data sets, whilst none of the phantom registrations failed. The root-mean-square (rms) target reprojection error was less than 1.3mm for the clinical data sets. The rms target reprojection error for the phantom images was less than 1mm when using the most accurate similarity measures.

1 Introduction

Registration of interventional digital subtraction angiography (DSA) to pre-operative magnetic resonance angiograms (MRA) can greatly enhance visualisation during minimally invasive neuro-interventions and introduces potentially useful complementary information such as three-dimensional (3D) blood flow. Whilst there have been a number of papers describing 2D-3D registrations of MRA and x-ray images, these studies have tended to favour a feature based approach in which, for instance, 2D and 3D vascular skeletons are extracted and matched using a suitable distance metric [3,4,5].

In this paper we apply the intensity-based registration of Penney *et. al.* [6] to the registration of MRA and DSA of the cerebral vasculature. In order to determine the most appropriate similarity measure for this new application, we compare the performance of six measures when applied to the registration of both a physical phantom and routinely acquired clinical images.

2 The Registration Algorithm

Six rigid-body extrinsic parameters describe the position (X , Y , Z) and orientation (θ_x , θ_y , θ_z) of the 3D data set. These parameters are iteratively varied and digitally reconstructed radiographs (DRRs) generated and compared to the DSA image using a suitable similarity measure. DRRs of the vasculature are produced by casting rays through the segmented 3D data set, from the x-ray source to each pixel location in the DSA image. As each ray passes through a spherical volume of interest (VOI) selected by the user, the intensities of the intercepted voxels are integrated and projected onto the imaging plane to produce a DRR.

A gradient descent search strategy [6] is used to search the extrinsic parameter space to optimise the similarity measure. To reduce processing time and improve the robustness of the algorithm, a multi-resolution strategy has been adopted and a pair of concentric regions of interest (ROI) are specified. The smaller of the two ROI is used to obtain an initial approximate match between the images and the larger is used to refine this registration. The radii of these ROI are set to a quarter and a half of the projected VOI radius.

We have compared the performance of normalised cross correlation (CC) gradient correlation (Grad. CC) entropy of the difference image (Entropy), mutual information (Mut. Info.), pattern intensity (Pat. Int.) and gradient difference (Grad. Diff.) when used to quantify the similarity of the DRR and DSA images. Please refer to [6] for an overview of these measures.

Fig. 1. The neuro-vascular phantom. Left: surface rendering of a thresholded CT scan of the phantom with the positions of the eight fiducial markers clearly visible. Middle: the maximum opacity image from the DSA sequence showing the concentric ROI masks used. Right: DRR corresponding to a registration with 0.5mm target reprojection error.

3 Phantom Experiment

In order to assess the accuracy of the algorithm we have applied it to images acquired from a physical silicon neuro-vascular aneurysm phantom (middle cerebral artery bifurcation aneurysm, figure 1). This phantom was made by Professor D. Rufenacht and Dr. K. Tokunaga of the Division of Neuro-Radiology, Geneva University Hospital, Switzerland. It was mounted in a perspex box and filled with a 15% (by weight) aqueous solution of gelatin to give realistic x-ray attenuation and scatter.

Fig. 2. Bottom: DRRs of the gold standard registrations of the clinical data sets (left to right, patients 1 to 4). Top: The maximum opacity images of the DSA sequence (inverted) to which the DRRs are registered showing the concentric ROI masks used.

DSA images were acquired using an Advantx DX (GE Medical Systems) x-ray set (two views). The PAL composite video output from the x-ray set was digitised via a Pulsar frame capture card (Matrox Imaging) and saved via a PC workstation as 512×512 pixel matrix, 8 bit grey-level images. The x-ray tube voltage was 85 kV and the phantom was placed in the isocentre of the x-ray system. Two views were acquired with the C-arm orientated at approximately $\pm 45^\circ$ to the vertical. A distortion-correction phantom and software were used to correct for pincushion distortion in the fluoroscopy images.

Phase-contrast MR angiography of the phantom was performed using a GE Medical Systems Signa Echospeed 1.5T. The acquired image contained $256 \times 256 \times 124$ voxels, each with dimensions $0.86 \times 0.86 \times 1.0$ mm. Blood vessels were segmented as described in [1] to produce a binary image (the 3D phantom model).

The intrinsic perspective parameters of the “ground truth” registration were calculated from images of a 60mm acrylic calibration cube in which 14 radio-opaque ball-bearings are embedded at each of the vertices and at the centers of each face. The extrinsic rigid-body parameters were calculated using eight fiducial markers attached to the perspex box containing the phantom. The fiducial markers consisted of a post to which two different types of acrylic imaging caps could be attached. The MR imaging cap contained a void which was filled with contrast fluid (0.5 mM Gadolinium). The x-ray imaging cap contained a divot which contained a 3mm diameter steel ball bearing. The caps have been accurately manufactured so that the centre of the ball coincides with the centre-of-gravity of the contrast fluid.

4 Clinical Validation

We obtained clinical MRA and DSA images from three patients undergoing treatment for cerebral aneurysms and one patient with an arteriovenous malformation (AVM).

Table 1. Displacements of the test registration starting positions from the ground truth (phantom) or gold standard (clinical data) registrations, in terms of the extrinsic parameters X , θ_x , θ_y and θ_z . Also given are the mean target reprojection errors for these starting positions.

Start Position	δX	$\delta\theta_x$	$\delta\theta_y$	$\delta\theta_z$	No. of Reg'ns	Mean Reproj. Error (mm)
1	± 25 mm	$\pm 4^\circ$	$\pm 4^\circ$	$\pm 4^\circ$	16	2.4
2	± 50 mm	$\pm 8^\circ$	$\pm 8^\circ$	$\pm 8^\circ$	16	4.7
3	± 75 mm	$\pm 12^\circ$	$\pm 12^\circ$	$\pm 12^\circ$	16	6.9
4	± 100 mm	$\pm 16^\circ$	$\pm 16^\circ$	$\pm 16^\circ$	16	9.1

Digital subtraction angiograms were obtained for all patients (two views per patient) using a GE Medical Systems Advantx DX x-ray set. These images were acquired using a Matrox Meteor II frame grabber, captured at half second intervals and saved as 512×512 pixel matrix, 8 bit grey-level images. A distortion-correction phantom and software were used to correct for pincushion distortion of the acquired images.

Phase-contrast MR angiography was performed using a Siemens Magnetom Vision 1.5T. The acquired images contained $256 \times 256 \times 64$ voxels with a resolution of $0.8 \times 0.8 \times 0.5$ mm. Blood vessels were segmented from three of these data sets as described in [1] and from the fourth as described in [2] to produce binary images (the clinical 3D models).

No perspective calibration cube images were available for these clinical data sets, so the four intrinsic parameters were estimated from the known focal length and image resolution of the x-ray set. This estimation is not expected to introduce significant errors into the registration, however any errors that are present will be included in the estimated target reprojection error calculation (section 5).

The extrinsic parameters of the gold standard registrations were generated via manual manipulation of a surface rendering of the 3D model using a interactive, graphical tool. To assess the reproducibility of the gold standard registrations two of the data sets were chosen and eight additional manual registrations were carried out by two observers. The first observer was a consultant neuro-radiologist (JVB) and the second a research fellow in medical imaging science (JHH). The mean reprojection error (calculated over the points described in section 5) between the gold standard and these manual registrations was calculated to be 1.7mm (standard deviation 0.4mm).

5 Registration Accuracy and Robustness Experiments

From the phantom ground truth registration, and the gold standard registrations for the clinical data, a total of 64 starting positions were generated by altering the positions of the 3D data sets using the perturbations given in table 1. The in-plane translation (δY or δZ) is assumed eliminated using a trivial manual alignment procedure, however, we simulate errors in this alignment by introducing a random perturbation (3-mm standard deviation) of the in-plane (Y and Z) position.

In order to assess the performance of the registrations, sets of between 12 and 18 target points were chosen by two consultant neuro-radiologists on the 3D phantom and clinical models. These points coincided with features such as bifurcations and points of

Fig. 3. Robustness results for the phantom (left) and clinical data sets (right), comparing the performance of the six similarity measures when registering the segmented MRA to the maximum opacity DSA images.

high vessel curvature. For each registration the mean target reprojection error of these points was calculated with respect to the corresponding ground truth or gold standard registration. If this mean error was greater than 4mm then the registration was deemed to have failed.

Two images were generated from each image sequence. The first was a single, approximately mid-sequence frame exhibiting good opacity of all arterial blood vessels. The second was a maximum opacity image in which the intensity of each pixel was set equal to the maximum opacity achieved during the DSA sequence.

6 Results

The results for registering the phantom MRA to the maximum opacity DSA images are summarised in figure 3 (left) and table 2 (top). 100% success rates have been obtained for the two closest starting positions using pattern intensity, gradient correlation and gradient difference. These measures fail more often than entropy, however, as the starting position is moved further from the ground truth. Mutual information performs less well than these measures but correlation is the least successful. Pattern intensity and gradient difference are the most accurate of the similarity measures, both achieving target reprojection errors of less than 1mm for all successful registrations.

The results for registering the clinical MRA data sets to the maximum opacity DSA images for all four patients are summarised in figure 3 (right) and table 2 (bottom). There

Table 2. Target reprojection error results for the phantom (top) and clinical data sets (bottom), comparing the performance of the six similarity measures when registering the segmented MRA to the maximum opacity DSA image.

Phantom Reprojection Errors in mm (SD).						
Start Posn.	CC	Entropy	Pat. Int.	Grad. CC	Grad. Diff.	Mut. Info.
1	1.59 (0.01)	1.08 (0.02)	0.89 (0.05)	0.90 (0.03)	0.88 (0.02)	1.25 (0.07)
2	1.66 (0.13)	1.14 (0.01)	0.88 (0.05)	1.25 (0.36)	0.90 (0.05)	1.38 (0.06)
3	2.10 (0.25)	1.10 (0.06)	0.94 (0.05)	1.10 (0.03)	0.99 (0.08)	1.45 (0.35)
4	2.13 (0.39)	1.20 (0.07)	0.94 (0.00)	1.42 (0.13)	0.91 (0.05)	1.11 (0.08)
Clinical Reprojection Errors in mm (SD).						
Start Posn.	CC	Entropy	Pat. Int.	Grad. CC	Grad. Diff.	Mut. Info.
1	1.26 (0.01)	1.69 (0.02)	1.12 (0.04)	1.18 (0.04)	1.28 (0.02)	1.15 (0.05)
2	1.38 (0.00)	1.99 (0.07)	1.15 (0.04)	1.08 (0.04)	1.26 (0.04)	1.20 (0.05)
3	1.67 (0.03)	2.17 (0.20)	1.24 (0.04)	1.26 (0.03)	1.39 (0.02)	1.36 (0.25)
4	1.64 (0.00)	2.02 (0.06)	1.35 (0.08)	1.34 (0.10)	1.58 (0.01)	1.74 (0.03)

are a number of differences between these results and those obtained for the phantom. The first is that entropy performs markedly worse than all the other measures for this clinical data, whereas it was at least as good if not better than the majority of the other measures for the phantom data. Of the other measures gradient difference performs consistently well for the clinical data sets, followed by pattern intensity, which out-performs gradient difference for start position 2, and gradient correlation.

For these clinical data sets pattern intensity and gradient correlation achieve the lowest target reprojection errors of 1.1 to 1.4mm. The target reprojection errors of gradient difference and mutual information are less than 1.3mm for start positions 1 and 2 but rise more steeply for more distant start positions. Entropy produces the least accurate registrations.

The results for registering to the single DSA images were similar to those obtained for registration to the maximum opacity images. In nearly all cases, however, registering to the maximum opacity DSA images resulted in a higher success rate (by up to 10%) compared to registering to the single mid-sequence frame. The maximum opacity registrations were also more accurate (by up to 0.5 mm in some cases). This result is not unexpected as these images will have higher contrast and signal-to-noise ratio than the single frames.

7 Discussion

We have found that gradient correlation, pattern intensity and gradient difference perform best of the six similarity measures compared. This is in agreement with the findings of Penney *et. al.* [6] for the comparison of these measures when used to register computed tomography (CT) to fluoroscopy images of a spine phantom. This is despite the large differences in modality and anatomy between these two applications, and the fact that none of these similarity measures have been specifically developed for application to MRA-DSA registration.

Fig. 4. Registration of patient 4, view 2. Left: Histogram of final values of the gradient difference similarity measure. Right: Comparison of registrations (with sub-regions enlarged). Central column: maximum opacity DSA image to which the MRA is registered. Left column: the gold-standard registration. Right column: the “best” gradient difference registration, *i.e.* that producing the highest value of the gradient difference similarity measure.

We have found that the success rate of these registrations falls off rapidly once the start position exceeds position 2, that is 8° rotation, 3mm in-plane translation and 50mm out-of-plane translation from the gold-standard/ground-truth registration. This is representative of intensity-based registration algorithms which have a “capture range” within which a certain fraction of corresponding features must be approximately aligned. However, we have found that manual alignment to within these tolerances can be rapidly and easily achieved using an interactive tool.

The registration success rates varied significantly between the four clinical data sets. For patients 1, 2 and 3, for instance, 90% of the registrations obtained using the gradient difference similarity measure were successful. For the second view of patient 4, however, only 48% succeeded. The histogram of similarity measure values for these registrations of patient 4 (figure 4) reveals a small cluster of 9 of the 128 registrations that all have consistently high similarity measures and very similar extrinsic parameter values. This mean registration position differs from the gold-standard position by 10° , 6° and 10° rotations about the x , y and z axes respectively. Visual comparison of these two registration positions, however, (figure 4) suggests that the position found by the algorithm is much more accurate than the gold-standard position.

The algorithm currently takes approximately 10 minutes running on a 1.2 GHz AMD processor PC with 1 GByte of RAM, however, considerable speed-up could be achieved using techniques such as shear-warp factorisation. The timing variation with different similarity measures was found to be negligible.

8 Conclusions

We have applied an intensity based 2D-3D registration algorithm to the multi-modality alignment of MRA and DSA images. Of the six similarity measures compared, gradient

difference, pattern intensity and gradient correlation performed consistently accurately and robustly for all data sets. Using these similarity measures, and for starting positions within 8° rotation, 3mm in-plane translation and 50mm out-of-plane translation from the gold-standard/ground-truth positions, we obtained a success rate of greater than 80% (less than 4mm target reprojection error) for the clinical data sets. Whilst none of the phantom registrations failed. The root-mean-square (rms) target reprojection error of the clinical registrations was less than 1.3mm (less than the 1.7mm reprojection error estimated for the gold standard registrations) and for the phantom images less than 1mm when using the most accurate similarity measures.

Acknowledgments

We would like to thank Kawaldeep Rhode, Robert McLaughlin, Albert Chung and Paul Summers for their assistance in acquiring the images used in this paper and also Kawaldeep Rhode for distortion correcting the DSA images used. This research is funded by EPSRC grant GR/M55015.

References

1. A.C.S. Chung and J.A. Noble. Statistical 3D vessel segmentation using a Rician distribution. In *Proc. MICCAI*, pages 83–89, 1999.
2. A.C.S. Chung, J.A. Noble and P. Summers. Fusing speed and phase information for vascular segmentation in phase contrast MR angiograms. In *Proc. MICCAI*, pages 166–175, 2000.
3. J. Feldmar, G. Malandain, N. Ayache, S. Fernandezvidal, E. Maurincomme and Y. Troussset. Matching 3D MR angiography data and 2D x-ray angiograms. In *Proc. CVRMed/MRCAS*, pages 129–138. Berlin, Germany: Springer-Verlag, 1997.
4. Y. Kita, D.L. Wilson and J.A. Noble. Real-time registration of 3D cerebral vessels to x-ray angiograms. In *Proc. MICCAI*, pages 1125–1133, 1997.
5. A. Liu, E. Bullitt and S.M. Pizer. 3D/2D Registration via skeletal near projective invariance in tubular objects. In *Proc. MICCAI*, pages 952–963, 1998.
6. G.P. Penney, J. Weese, J.L. Little, Desmedt P., D.L.G. Hill, and D.J. Hawkes. A comparison of similarity measures for use in 2D-3D medical image registration. *IEEE Transactions on Medical Imaging*, 17((4):586–595, 1998.
7. G.P. Penney, P.G. Batchelor, D.L.G. Hill, D.J. Hawkes. and J. Weese. Validation of a two- to three-dimensional registration algorithm for aligning preoperative CT images and intraoperative fluoroscopy images. *Medical Physics*, 28(6):1024–1032, 2001.