Real-time Registration of 3D Cerebral Vessels to X-ray Angiograms

Yasuyo Kita[1] Dale L. Wilson[2] J. Alison Noble[2]

[1] Electrotechnical Laboratory, Ibaraki 305-8568, Japan
(terashi@robots.ox.ac.uk Tel:+44-1865-282185 Fax:+44-1865-273908)
[2] University of Oxford, Oxford OX1 3PJ, United Kingdom

Abstract. A quick method to obtain the 3D transformation of a 3D free-form shape model from its 2D projection data is proposed. This method has been developed for the real-time registration of a 3D model of a cerebral vessel tree, obtained from pre-operative data (eg. MR Angiogram), to a X-ray image of the vessel (eg.Digital Subtraction Angiogram) taken during an operation. First, the skeleton of the vessel in a 2D image is automatically extracted in a model-based way using a 2D projection of a 3D model skeleton at the initial state (up to ±20 degree difference in rotation). Corresponding pairs of points on the 3D skeleton and points on the 2D skeleton are determined based on the 2D Euclidean distance between the projection of the model skeleton and the observed skeleton. In the process, an adaptive search region for each model point, which is determined according to the projected shape, effectively removes incorrect correspondences. Based on a good ratio of correct pairs, linearization of a rotation matrix can be used to rapidly calculate the 3D transformation of the model which produces the 2D observed projection. Experiments using real data show the practical usefulness of the method.

Key words: 3D-2D registration, ICP (Iterative closest point) algorithm, multi-modal fusion and augmented reality visualizations.

1 Introduction

This work is being developed to aid the endovascular treatment of intracranial aneurysms by coil embolisation. In current practice, the neuroradiologist guides a catheter through a vessel while viewing its 2D projection (X-ray angiogram). It is hard for a neuroradiologist to visualize the complex 3D shapes of the vessels from one 2D projection, even with the 3D shape information from pre-operative data (eg. MR A(ngiography)). To help the neuroradiologist's understanding, Wilson and Noble[1] developed a method for reconstructing a 3D model of cerebral vessels from slices of MRA data. Fig. 1a shows a result of the reconstructed 3D model. If this 3D structure is superimposed on a 2D intra-operative X-ray image and the location of the catheter is displayed on the 3D reconstruction, it may aid the neuroradiologists in accurately deciding how they should manipulate the catheter. For this purpose, real-time determination of the posture and position of the 3D model from its 2D projection is required.

The determination of the position and posture of a 3D model from its 2D view is a fundamental and important problem in Computer Vision research. In the case that the object has some prominent features (points, edges etc) that can be robustly extracted and matched between the 3D model and its 2D view, the approach based on feature-matching can be taken. However, usually both robust feature extraction and robust feature matching are not easy, especially in the case of a free-form object. The iterative closest point (ICP) algorithm[2], originally developed for 3D-3D rigid registration, has appropriate characteristics for free-form shape registration. The basic idea of the method is to use iterative transformations of the 3D model towards the correct position and posture using the corresponding pairs between the observed and the model points, which are matched on the basis of the closeness at each state. If the initial position and posture is not far from the correct position and posture, so that the corresponding pairs include a high ratio of correct pairs, the model can converge to the correct state. For registration of a 3D model registration to its 2D view, that is, for obtaining the best 3D transformation of a model which produces a given 2D view, the difficulties of the extension of this approach are mainly two-fold:
I) The difficulty of finding correct pairs between the projection of the 3D model and the observed 2D view using only the 2D distance, and;
II) Even after finding the pairs, it is not easy to feedback the 2D difference to the 3D transformation of the model.

Concerning (I), in [3], the tangent of the projection of the 3D model and the observed 2D curve was used to decrease the number of bad correspondences. Although the effectiveness of using such additional attributes (curvatures, grey level etc) in addition to the geometrical distance has been shown in 3D-3D registration of free-form objects[4], it is not so effective in the 3D-2D case for two reasons: a) the tangent on the 2D image is not an invariant feature; and b) the projection of the complex 3D model often causes complicated self-overlapping, where the robust calculation of geometric features can be difficult.

Concerning (II), most proposed methods take similar approaches to a gradient descent method (eg. [5]) to find the best 3D transformation which minimizes the sum of 2D distances between corresponding pairs (or maximizes the similarity between the projection of the model and the observed data) over the six degrees of freedom. However, such approaches are time consuming and are not suitable for use in real-time applications. Fortunately, to address this problem, the Active Vision Research field has made advanced steps towards real-time object tracking from time-sequential 2D images. One solution that has been proposed is to linearize the 3D non-linear transformation[6]. The main difference between their application and ours is that, in their case, the feature correspondences are easier to find since special features (like corners) can be used. To use thier approach, we need to solve the problem (I) and robustly obtain a high ratio of correct pairs.

In this paper we propose a fast registration method which overcomes the two difficulties noted above as follows. Concerning (I), the model-based strategy plays an important role both in extracting the vessels from the X-ray image and
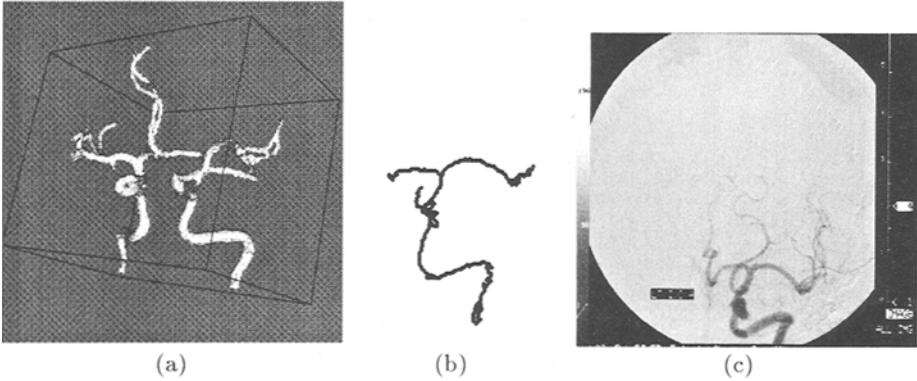
Fig. 1. The 3D model of the cerebral vessels and a digital subtraction angiography (DSA) image of the same vessels: (a) the 3D model of both carotid circulations; (b) a skeleton of the left internal carotid circulation; (c) a DSA image of the left internal carotid circulation.

in finding correct point matching pairs between the vessel and the 3D model. In particular, the search area for finding the corresponding observed points is effectively adjusted for each point in the 3D model depending on the shape of the model projection, so that most of the wrong pairs are excluded. Secondly, concerning (II), taking advantage of the high-ratio of correct pairs obtained by the first process, the 3D transformation of the model is calculated using separation of the translation effect and linearization of the rotation matrix. Although the correct position and posture of the model is not obtained at once, because of inaccurate matching pairs and linearization errors, the 3D model quickly converges to the correct state by iterating the point matching and model transformation processes.

2   Model-based 2D vessel extraction

2.1   Preprocessing

The input to our method is a skeleton of the 3D vessel model (eg. Fig.1b), obtained from the full 3D vessel reconstruction (Fig. 1a)[7], and a digital subtraction angiography (DSA) image of the vessels (eg. Fig. 1c). For full automation, the region of interest (which is almost a circle) is extracted from the X-ray image with simple image processing. The small black rectangle containing text is also removed from the region of interest.

2.2   Initial localization

Here we briefly explain our 3D coordinate frame, (X, Y, Z). The X-ray source of the X-ray machine is defined as the origin of the coordinate system. The image plane is on the $Z = f$ plane, where f is the distance between the source and the plane. The X and Y axes are defined as the same directions as I and J
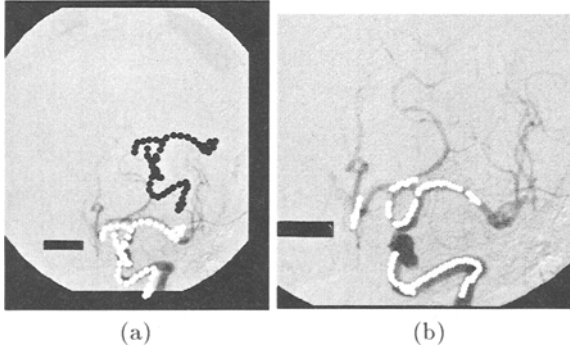
Fig. 2. Model-based extraction of 2D vessel skeleton: (a) initial translation (t_x, t_y) of 3D model (from black to white points); (b) resultant extraction (white lines).

of the image coordinates respectively. The 3D vessel model is placed between the source and the image plane and perspectively projected to the image. The model's initial position and posture is approximately known.

In the registration process we use n points which have been sampled at regular intervals from the 3D model skeleton. The 3D coordinates of these points are $\mathbf{X}_i = (X_i, Y_i, Z_i)^\top (i = 1, ...n)$. The 3D transformation of the model is represented by $\mathbf{R}$(the 3×3 rotation matrix) and $\mathbf{T} = (t_x, t_y, t_z)^\top$(the translation vector). The 2D projections of the 3D model points after the transformation of $\mathbf{R}$ and $\mathbf{T}$ have the 3D coordinates $(x_i, y_i, f)^\top$, where $x_i = f X'_i / Z'_i$, $y_i = f Y'_i / Z'_i$ and $\mathbf{X}'_i = \mathbf{R}\mathbf{X}_i + \mathbf{T}$.

When real X-ray images are acquired, the patient's head is immobilized, and the X-ray source and the image plane are rotated together around the head. Here, inversely, we rotate the model (head) to give the same effect as the X-ray system rotation. Since the rotation angle of the system is known from the graduations, the position and posture of the 3D model can be estimated approximately. This includes about $\pm$ 20 degrees error in rotation and about $(\pm 100, \pm 100, \pm 200)$(mm) in translation, since the position and posture of the head is not calibrated and is changed a little during the acquisition of the X-ray images. It is this calibration, or determination of the change in the patient position and orientation between MR and X-ray, that we wish to find.

In Fig. 2a black points represent the projection of the 3D model skeleton at its initial state. Using simple template matching between the projected shape of the 3D model skeleton at its initial state and the X-ray image, t_x, t_y is roughly estimated so that the projection optimally overlaps the dark regions (possible vessel regions) . In the case of Fig. 2a, the model is translated by (-14.8, 39.8, 0.0)(mm); the white points show the projection of the model after the translation.
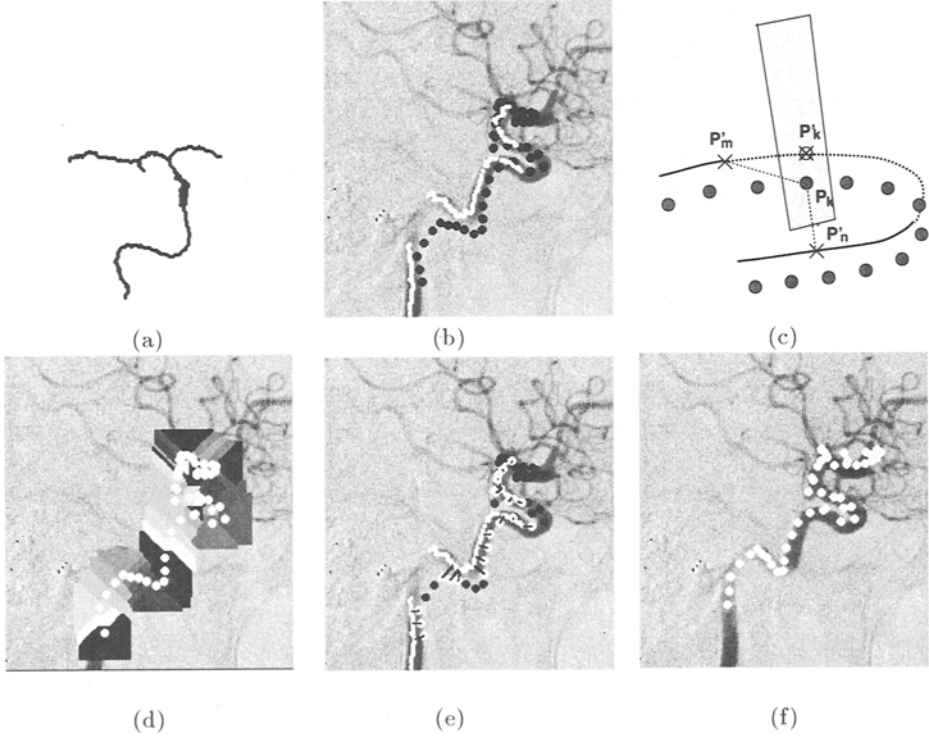
Fig. 3. Point matching search: (a) skeleton of the 3D model of the right internal carotid circulation; (b) initial position (black points) and extracted 2D skeleton (white lines); (c) anisotropy of the appropriate search regions(see text); (d) territory-based search regions; (e) result of point matching (only the large white dots made corresponding pairs); (f) the final registration result.

2.3 Skeleton extraction

After the initial localization, the projection of the model is classified into horizontal and vertical segments on the image. In the vicinity of each segment, a histogram of the grey levels of the image is calculated to decide an adaptive threshold of vessel brightness. Although the estimated thickness of the vessel from the 3D model can be also used to establish an adaptive threshold, we have not yet implemented this. Using the adaptive information of approximate vessel direction and brightness, the corresponding opposite edges of vessels are tracked more robustly than without a model. In Fig. 2b, the skeleton is well extracted despite the weak image contrast.

3 Territory-based correspondence search

Here our sub-goal is to find as many correct corresponding points on the 2D vessel skeleton for each point on the 3D model while avoiding wrong pairs. Fig.

3 shows an example of a X-ray image of the right internal carotid circulation: Fig. 3a shows the 3D model skeleton; Fig. 3b shows the projection of the model (black points) after initial localization (Section **2.2**) plus the skeleton of the 2D vessel extracted from the image by the method described at Section **2.3** (white lines). As shown in Fig. 3b, some parts of the skeleton of the 2D vessel are not extracted because of insufficient contrast enhancement, a sudden turn of the vessel (acute curve), or self-overlap. Fig. 3c is a schematic diagram at such a location: the solid line and the dotted line show the extracted and unextracted parts of 2D vessel skeleton respectively; the large dots represent the projection of the 3D model points. If we try to find the corresponding 2D points for all of the 3D model points using only the closeness on the 2D image, the model points whose corresponding 2D vessel has not been extracted (eg. P_k) will definitely have the wrong correspondences (eg. P'_m or P'_n) causing serious errors.

Removing such model points (matching outliers), is not an easy task since it is not known which parts are successfully extracted before the correct registration has been computed. For example, in [3], human interaction is required for this removal process. Note, however, that in many cases wrongly selected 2D vessel points, like P'_m or P'_n for P_k, are the correct corresponding points to other model points. If we can set an appropriate search region (eg. the shaded region in Fig. 3c) by considering the projected shape, model points whose corresponding 2D vessel part has not been extracted can be rejected as a point that has no correspondence.

We propose a novel method to adaptively calculate such an anisotropic search region according to the projected shape of the model. The search region for each model point is determined by segmenting the image into territories of the model points as follows. On the image, each projected point tries to extend its territory in the vicinity of the point at the same speed as the other points and to some width (the largest search width). Each pixel in the image belongs only to the search region of the point that first reaches the pixel. As a result, search regions become akin to Voronoi regions. Concretely, we implement this process by region growing using mask-processing. Fig. 3d shows the resultant search regions. We can see that points around more complicated projected shapes tend to have smaller search areas. As a result, the chance that model points find a wrong corresponding point (which actually corresponds to different parts of the model) are greatly decreased; while some points find the proper match at a far position as seen in Fig. 3e. This search mechanism is essential when we deal with complex 3D free-form models which cause complicated self-overlapping in the 2D projections.

4   3D-2D matching pairs → 3D transformation

Now we have m corresponding pairs of the observed position $\mathbf{x}_i^o, (x_i^o, y_i^o, f)^\top$ and a 3D model point $\mathbf{X}_i, (X_i, Y_i, Z_i)^\top$ whose projection is $\mathbf{x}_i, (x_i, y_i, f)^\top$. In many methods, the nonlinear minimization, $min(\sum_{i=1}^m ((x_i^o - x_i)^2 + (y_i^o - y_i)^2))$ is solved iteratively to determine $\mathbf{R}$ and $\mathbf{T}$ at the same time. To speed up this

calculation, we take the approach presented in [6]. We explain this method very briefly.

The minimization criteria is initially only based on rotation by separating out the translation effect using the following geometry. If we pick up two pairs from the m corresponding pairs, $\mathbf{X^o}_i - \mathbf{X^o}_j = \mathbf{R}(\mathbf{X}_i - \mathbf{X}_j)$ should be perpendicular to the vector $(\mathbf{x^o}_i \times \mathbf{x^o}_j)$. Therefore the optimal rotation is obtained by computing

$$\min_{\mathbf{R}} \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} ((\mathbf{x}_i \times \mathbf{x}_j) \cdot \mathbf{R}(\mathbf{X}_i - \mathbf{X}_j))^2.$$

The rotation matrix can be represented by quaternions $\mathbf{q} = (q_0, q_1, q_2, q_3)^\top$ and linearized with some small error as follows.

$$\mathbf{R} = \begin{bmatrix} (q_0^2 + q_1^2 - q_2^2 - q_3^2) & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_1 q_2 + q_0 q_3) & (q_0^2 - q_1^2 + q_2^2 - q_3^2) & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & (q_0^2 - q_1^2 - q_2^2 + q_3^2) \end{bmatrix}$$

$$\approx \begin{bmatrix} q_0 & -2q_3 & 2q_2 \\ 2q_3 & q_0 & -2q_1 \\ -2q_2 & 2q_1 & q_0 \end{bmatrix}$$

Using this linearization, the equation of minimization on $\mathbf{R}$ becomes

$$\min_{\mathbf{q}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} ((\mathbf{x}_i \times \mathbf{x}_j) \cdot (q_0(\mathbf{X}_i - \mathbf{X}_j) + 2\mathbf{q}' \times (\mathbf{X}_i - \mathbf{X}_j)))^2,$$

where $\mathbf{q}' = (q_1, q_2, q_3)^\top$. The $\mathbf{q}$ which minimizes this function is obtained by solving the simultaneous equations of the partial differentials on $\mathbf{q}$.

Once the rotation is known, the translation can be calculated by solving $2m$ simultaneous equations on (t_x, t_y, t_z):

$$\frac{f(X_i^{o'} + t_x)}{Z_i^{o'} + t_z} = x_i' \quad \text{and} \quad \frac{f(Y_i^{o'} + t_y)}{Z_i^{o'} + t_z} = y_i',$$

where $\mathbf{X^{o'}} = \mathbf{RX}$ using the calculated $\mathbf{R}$.

We tested this method using synthetic data. Although we omit the detail of the results, the experiments indicated that a high ratio of correct pairs (at least 50 %) is imperative for the success of this method.

5 Results

Fig. 3f shows a registration result on a real example. Although a lower part of the internal carotid artery is faded owing to insufficient contrast enhancement and many of the other arteries are self-overlapping, the 3D model is properly transformed so that the projection overlaps well with the 2D observed vessel. The final transformation is a 16.6 degree rotation around the axis (0.39, 0.67 , -0.64) and a (6.6, 6.5, -5.6)(mm) translation. Computational time for each step
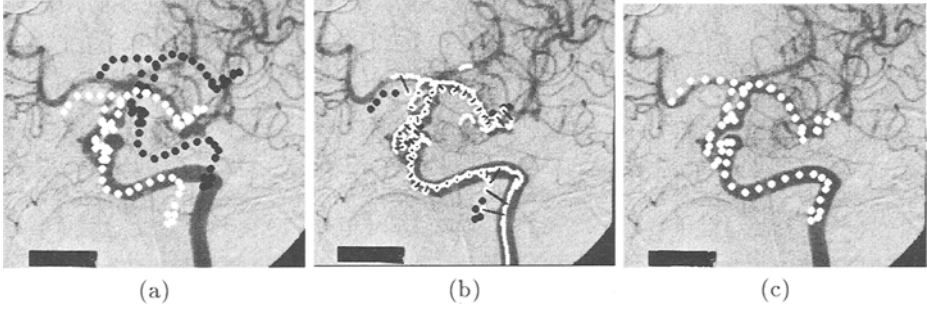
Fig. 4. Result: (a) initial localization (white points); (b) initial point correspondence; (c) final registration result.

was (1)preprocessing (0.2 sec), (2)initial localization (0.08 sec), (3)2D skeleton extraction (0.5 sec), (4)point correspondence (0.13sec) and (5)3D transformation (0.01 sec). Steps (4) and (5) were iterated 30 times for convergence and hence the total time was 5.5 sec (SUN IPX/ULTRA1).

Fig. 4 shows another result using a X-ray image of the vessels of the 3D model which is shown in Fig. 1b. Although the projected shape of the model at initial state is quite different from the observed vessel, the 3D transformation of the model is correctly obtained. The transformation is a 14.6 degree rotation around the axis (-0.77, -0.61, -0.18) and a translation of (-9.7, 4.8, -127.9)(mm). The large translation in the Z direction agrees with the fact that the height of the bed where the human subject lay was altered by approximately 10~15 cm during the acquisition of the images. Even though, at the middle of the internal carotid artery in the image, there is very complicated self-overlapping at which the topology of the initial projected shape is different from the actual one, territory-based search restriction helps to retain the high ratio of correct corresponding pairs. Computational time for each step was (1)preprocessing (0.2 sec), (2)initial localization (0.11 sec), (3)2D skeleton extraction (1.5 sec), (4)point correspondence (0.31 sec) and (5)3D transformation (0.01 sec). A total of 18 iterations were required for convergence and the total time was 5.7 sec (SUN IPX/ULTRA 1). We have applied the method to a total of 9 X-ray images for the two 3D models. All data have similar results except one which failed because of poor 2D skeleton extraction.

6   Conclusion

We have proposed a method for the real-time determination of the pose and position of a complicated 3D free-form shape with respect to its 2D projection. The method is fully automated. The robustness and speed of the method are supported by two characteristics:
1. Robust model-based 3D-2D point matching using territory-based search restriction.

2. Linear solution to obtain the 3D transformation from 3D-2D point correspondences.

Although the second part is not original, the first part enables the second part to work effectively even on complication data. The results using real data show that the proposed method is quite promising.

Our future work will focus on:

1. Assessment of the method by comparing it with 3D ground truth data. Unfortunately, we do not have the true value of the 3D transformation of the model in the experiments shown here. We will apply the method to real data where fiducial markers have been used to compute the transformation.

2. Improvement of the 2D skeleton extraction process. Since the model gradually becomes closer to the correct state, incremental extraction is more desirable.

Other aspects we will consider include:

1. Comprehensible display of the 3D model with the 2D view.

2. Fusion of two (or more) 2D views. The information about the distance in Z direction is fundamentally poor if we use just one view. Fusion of two different views is desirable to increase the accuracy of the localization of the model. It is especially important in the case where the 3D model has a simple linear shape.

References

1. D.L. Wilson and J.A. Noble: "Segmentation of cerebral vessels and aneurysms from MR angiography data", In *Proc. of Information Processing in Medical Imaging*, pp.423–428, 1997.
2. P. J. Besl and N. D.Mckay: "A method for registration of 3D shapes", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.14, No.2, pp. 239–256, 1992.
3. J. Feldmar, G. Malandain, N. Ayache, S. Fernandez-Vidal, E. Maurincomme and Y. Trousset: "Matching 3D MR Angiography Data and 2D X-ray Angiograms", In *Proc. of CVRMed-MRCAS'97*, pp.129–138, 1997.
4. Y. Kita: "Force-based registration method using attribute values", In *Proc. of 14th International Conference on Pattern Recognition*, pp.34–39, 1996.
5. G.P. Penney, J.A. Little, D.L.G. Hill and D.J. Hawkes: "2D-3D registration for use in image guided intervensions", In *Proc. of Medical Image Understanding and Analysis '97*, pp.49–52, 1997.
6. J.J. Heuring and D.W. Murray: "Visual head tracking and slaving for visual telepresence", In *Proc. of IEEE Int Conf. on Robotics and Automation*, pp.2908–2914, 1996.
7. D.L. Wilson, J.A. Noble and C. Pudney: "From MR Angiography to X-Ray Angiography", In *Proc. of Medical Image Understanding ans Analysis'97*, pp.161–164, 1997.