

A Stochastic Iterative Closest Point Algorithm (stochastICP)

G.P. Penney, P.J. Edwards, A.P. King, J.M. Blackall, P.G. Batchelor, and D.J. Hawkes

Division of Radiological Sciences and Medical Engineering, The Guy's, King's and St. Thomas' Schools of Medicine and Dentistry, Guy's Hospital, London, SE1 9RT, UK.

Abstract. We present a modification to the iterative closest point algorithm which improves the algorithm's robustness and precision. At the start of each iteration, before point correspondence is calculated between the two feature sets, the algorithm randomly perturbs the point positions in one feature set. These perturbations allow the algorithm to move out of some local minima to find a minimum with a lower residual error. The size of this perturbation is reduced during the registration process. The algorithm has been tested using multiple starting positions to register three sets of data: a surface of a femur, a skull surface and a registration to hepatic vessels and a liver surface. Our results show that, if local minima are present, the stochastic ICP algorithm is more robust and is more precise than the standard ICP algorithm.

1 Introduction

The iterative closest point (ICP) algorithm [1] has been used for a wide variety of applications both in the field of medical imaging [2,3] and engineering. We are primarily concerned with using the ICP algorithm to register intraoperative ultrasound data to preoperative 3D modalities for use in image-guided surgery or interventions. This can enable surgeons or interventionists to navigate towards their surgical targets by viewing preoperative images which are spatially aligned with either the physical coordinate system of the patient or the intraoperative modality.

The ICP algorithm is usually an efficient method to minimise the root mean square (RMS) distance between two feature sets. However, it can be prone to finding local minima, especially in noisy feature sets. Previous attempts have been made to increase the robustness of the ICP algorithm. Masuda and Yokoy [4] improve robustness by using a random sample of points (rather than the whole dataset) which are altered at each iteration. In Luck *et al.* [5] robustness is improved by using both a simulated annealing and an ICP algorithm. Essentially the simulated annealing algorithm is used to produce "good" starting points for the ICP algorithm.

We propose a novel approach to increase the robustness of the ICP algorithm by adding random Gaussian noise to perturb the position of the features in one feature set. The magnitude of the noise is reduced during the registration process until finally it is set to zero when the algorithm reverts back to standard ICP.

The structure of our paper is as follows. We initially outline the standard ICP algorithm and detail our modifications. We then describe our experiments to register three sets of data: a femur, a skull and a liver dataset.

2 Method

2.1 Outline of ICP Algorithm

The iterative closest point algorithm is widely used and well known, and so only a very brief summary is given in this paper. Our description will concentrate on the case where the source dataset consists of n points $\mathbf{p}_j, j = 1, \dots, n$, whereas the target feature set can consist of points, lines or surfaces.

The algorithm registers two sets of features by repeating the following steps.

1. For each source point, calculate the closest point, or position on a line or surface in the target dataset.
2. Calculate the transformation matrix \mathbf{T} which minimises the RMS distance between the two sets of corresponding points.
3. Repeat the above steps until a suitable stopping criterion is met.

2.2 Modifications to the Algorithm

We have modified the algorithm in two significant but related ways. The first is to introduce an extra step to be carried out before point correspondence is calculated i.e. step 0 in the framework shown in section 2.1.

0. Add a random Gaussian perturbation to the position of the source points: $\mathbf{p}'_j = \mathbf{p}_j + \mathbf{s}_j, j = 1, \dots, n$, where \mathbf{s}_j represents a vector in a random direction, with a magnitude which has standard deviation σ .

Steps 1 and 2 are now carried out using the perturbed \mathbf{p}'_j rather than the original \mathbf{p}_j source points.

The second alteration to the algorithm is to change the stopping criterion (step 3) to be a set of criteria which define when the noise σ should be reduced. We would like to reduce σ when the algorithm has reached an optimum position. In most cases the algorithm tends to move towards the optimum position and then fluctuates about this position. In our stochastICP algorithm, due to the random perturbations, the residual error can both increase and decrease. Therefore, it is not obvious how the residual error value can be used to define the noise reduction criterion. Instead we have used the following criterion which is based on the rigid body parameters $\mathbf{E} = (\theta_x, \theta_y, \theta_z, X, Y, Z)$. These parameters can be calculated by decomposing the transformation matrix \mathbf{T} . The values of these parameters are logged after each iteration and if the current set of parameters \mathbf{E}_N are within a certain threshold t of a previous set of parameters \mathbf{E}_i then the algorithm reduces the size of σ ,

$$\exists i_0 \in \{1, \dots, N - 5\} : \max_{k=1, \dots, 6} |E_{i_0}^k - E_N^k| < t \tag{1}$$

where E^k represents the individual elements in the 6-tuple \mathbf{E} , i_0 represents a particular iteration and N equals the current number of iterations. The last five sets of parameters are not included in the search i.e. only $\mathbf{E}_i, i = 1, \dots, N - 5$ are used for the following reason. By excluding the last five iterations we aim to distinguish between the following

two cases. Firstly, where the algorithm is simply moving slowly through the search space; in this case the algorithm is still moving towards the optimum position and so we do not want to alter σ . Secondly, where the algorithm has revisited a position in the 6-dimensional parameter space; this should indicate that the algorithm is fluctuating about an optimum position and so σ should be reduced. The exclusion of the last five parameters was empirically chosen.

For the experiments described in this paper we begin with $\sigma = 16\text{mm}$. When the noise reduction criterion is met σ is reduced by a factor of $1/\sqrt{2}$ and the algorithm continues until $\sigma < 0.25\text{mm}$. At this point the noise is set to zero. The final stopping criterion for the algorithm is a standard stopping criterion i.e. where the change in the residual error falls below a threshold (10^{-4}mm for the experiments described in this paper).

The parameter t was set equal to $\sigma/5$ for the experiments using the femur and liver datasets and was set equal to $\sigma/50$ for the experiments using the skull dataset. These ratios were chosen empirically. Our reasons for choosing these values for t are as follows. The parameter t defines how precisely the algorithm must revisit a position for the noise reduction criterion to be met. Our choice for the size of t depends on the size of the variations in the values of the rigid body parameters \mathbf{E} between different iterations. If these parameters vary by a large amount, then t must also be fairly large, otherwise the algorithm may take an extremely large number of iterations to reach registration. However, if \mathbf{E} varies by only a small amount, then t should also be small to prevent the noise reduction criterion being met before the algorithm has reached an optimum position. The rigid body parameters \mathbf{E} vary due to two factors, the size of the noise σ and the speed with which the algorithm moves towards the optimum position. We account for the first of these factors by setting t to be a fraction of σ , therefore, the size of t is reduced as the noise decreases. The two different values for this ratio were necessary because the parameters \mathbf{E} varied by a much smaller amounts between successive iterations when registering the skull dataset, compared to the femur and liver datasets. This is probably due to the almost constant curvature of the upper part of the cranium.

2.3 Experiments

Experiments have been carried out using three sets of data. We are primarily concerned with the registration of intraoperative ultrasound data with preoperative 3D modalities for use in image guided surgery or interventions. Consequently each dataset used in this paper simulates ultrasound to CT or MR registration. However, our modifications do not detract from the generic nature of the ICP algorithm and so it should be useful for a wide range of registration problems.

The three target datasets were: a CT scan of a phantom femur (see figure 1), a MR scan of a volunteer's head and a MR scan of a volunteer's liver. In each case the source modality was freehand 3D ultrasound. The datasets are summarised in table 1. The 3D surfaces were extracted using a marching cubes algorithm [6] implemented in VTK [7] which was applied to binary images segmented from the 3D volumes. The binary volumes were produced by applying a threshold to segment the femur surface from the CT volume and by manual segmentation for the two MR volumes using the ANALYZE software package (Biomedical Imaging Resource, Mayo Foundation, Rochester, MN,

USA). The line features in the liver dataset were created from a number of manually picked points which define the centrelines of blood vessels.

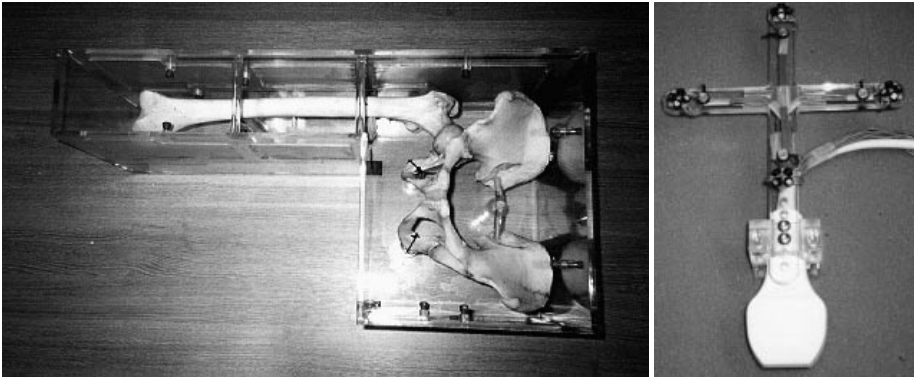


Fig. 1. Femur phantom (left) and a 10.5MHz ultrasound probe with infra-red LEDs attached (right).

The source points were manually picked from sets of freehand 3D ultrasound images. A 10.5MHz ultrasound probe (see figure 1), tracked using an Optotrak 3020 optical tracking system from Northern Digital Inc., was used to acquire source bone points for the femur and volunteer skull datasets. A 3.5MHz probe, tracked using a Polaris optical tracking system from Northern Digital Inc., was used to acquire the liver surface and midline of hepatic vessels for the liver dataset. The approximate 3D spatial accuracy of the Optotrak and Polaris are 0.2mm and 0.35mm respectively.

For the phantom femur dataset a “gold-standard” registration was calculated using fiducial markers attached to the box which contained the femur. For the volunteer head data a “gold-standard” registration was calculated by attaching a locking acrylic dental stent (LADS) [8] to the volunteer. Markers attached to the LADS were used to calculate a point based image to physical registration. No “gold-standard” was available for the liver dataset.

Target Modality	Target Feature	Source Feature
CT 320×320×177 voxels 1.094×1.094×3mm	femur surface 11484 facets	835 ultrasound bone points
MR 256×256×200 voxels 0.898×0.898×1.2mm	skull surface 12146 facets	375 ultrasound bone points
MR 256×256×26 voxels 1.328×1.328×10mm	liver surface (3044 facets) + hepatic vessels (175 line segments)	71 ultrasound liver surface 46 ultrasound vessel points

Table 1. Summary of the datasets showing voxel sizes, image dimensions, type of target feature and number of source points.

One hundred registrations were carried out using each dataset. The starting positions were the “gold-standard” registration position with random noise added to each of the six rigid body degrees of freedom. The size of this perturbation was an estimation of how accurately an approximate registration could be calculated during a procedure. The random noise used was 15mm or degrees for the femur and skull datasets and 30mm or degrees for the liver dataset. When registering using the liver dataset no “gold-standard” registration was available. Instead an initial registration was estimated by placing the probe on the inferior end of the volunteer’s sternum and acquiring an axial ultrasound slice. This ultrasound slice was approximately registered to the MR volume by making the following assumptions: that the x and y axes of the ultrasound image are parallel to the x and y axes in the MR volume and that the centre of the ultrasound image corresponds to the centre of the MR volume.

Registrations were carried out using our *stochastICP* algorithm and for comparison the standard ICP algorithm was also used. The stopping criterion for the standard ICP algorithm was identical to the final stopping criterion using the *stochastICP* algorithm i.e. when the change in the residual error falls below 10^{-4} mm.

3 Results

The registration results were analysed in the following ways. Firstly, failed registrations were removed. A registration was deemed to be a failure if its mean target registration error (TRE) [3] was greater than five times the minimum value of mean TRE over the set of 100 registrations. It was not possible to calculate a TRE for the liver dataset as there was no “gold-standard” registration, instead a registration was deemed to be a failure if the residual error was more than 25% larger than the minimum residual error for the set of 100 registrations.

A mean TRE was calculated by transforming a number of points, one for each voxel within a region of interest, firstly by the “gold-standard” matrix and secondly by the final registration matrices and calculating the RMS distance between the two positions. The regions of interest chosen for the three datasets were all the voxels within either the femur, the skull or the liver.

A RMS distance measure was also calculated to indicate the spread of the final registration positions (i.e. a measure of precision). This was calculated using the following method. Each point within the region of interest was taken in turn and transformed using each of the successful final registration matrices. For a given point the mean transformation position was determined and then the RMS distance to the mean position was calculated. The RMS distance quoted is the RMS value of these RMS distances calculated over the entire region of interest.

The registration results are given in table 2. This table shows the mean residual error, RMS distance, mean TRE and mean number of iterations for the successful registrations. It also shows the algorithm failure rate. Renderings, showing the target features and the final position of the source points for a typical successful registration, are shown in figure 2.

The results for all three datasets showed that the algorithms produced very similar residual error values. The *stochastICP* algorithm recorded lower RMS distance values

Feature	Algorithm	Mean Residual (mm)	RMS Distance (mm) (precision measure)	Mean TRE (mm) (accuracy measure)	Mean No. Iter.	Fail Rate (%)
Femur	ICP	0.72	0.15	1.18	81	36
	stochastICP	0.72	0.03	1.17	122	0
Skull	ICP	0.86	1.98	2.89	357	1
	stochastICP	0.86	1.57	2.65	479	1
Liver	ICP	6.14	2.60	-	102	24
	stochastICP	6.12	1.10	-	140	7

Table 2. Registration results, showing the mean residual, RMS distance, mean TRE and mean number of iterations for the successful registrations. Failure rate is also shown.

and mean TRE values for all three datasets, in particular when using the liver dataset where the RMS distance value was less than half that recorded using the standard ICP algorithm.

The largest differences between the two algorithms occurred in terms of the number of iterations required to reach registration and the failure rate. The stochastICP algorithm produced a much lower failure rate when using the femur and liver datasets, however, it required between 34% – 51% more iterations to reach registration.

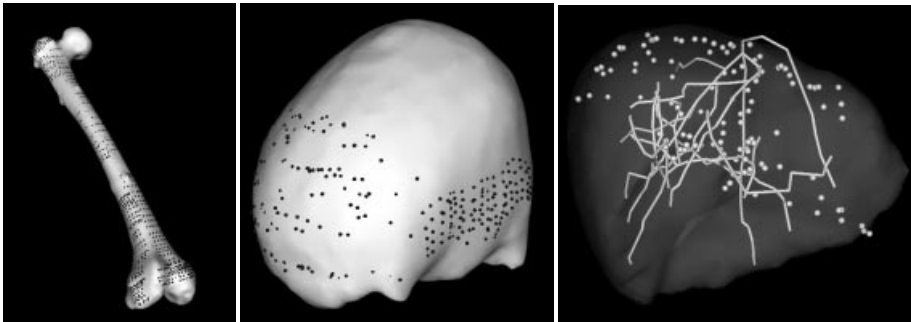


Fig. 2. Registration results from a typical successful registration position showing the source points overlaid onto the target surface and line features.

4 Discussion

The results using both the femur and the liver datasets show that the stochastICP algorithm is more robust than the standard ICP algorithm. The likely reason for this improvement is the presence of local minima in these datasets. It is clear that the standard ICP algorithm is not able to manoeuvre out of a local minimum, whereas the stochastICP algorithm, due to the random perturbations added to the point positions, can escape

from some local minima. In the case of the femur data the standard ICP algorithm resulted in solutions which were a long way from the “gold-standard” position. Figure 3 shows a rendering of the source points and target surface for one of the failed registrations using the standard ICP algorithm. This local minimum is caused by the points from the ultrasound images, which were mainly acquired from the anterior surface of the femur (as the posterior surface can be more difficult to image in a patient), crossing over to register to opposite sides of the femur surface. Because the femur is a long thin structure a very small rotational misalignment can place the feature sets in such a local minimum.

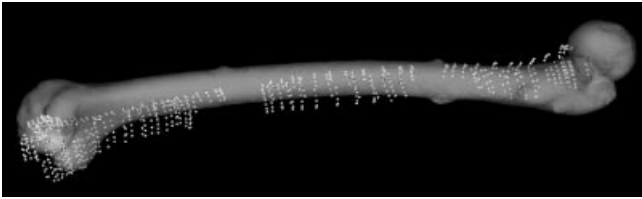


Fig. 3. A rendering of the source points and target surface femur dataset showing a misregistration position found using the standard ICP algorithm. The opacity of the femur surface has been reduced to show target points within the centre of the femur. The final registration position had a residual error of 5.54mm and a TRE of 18.74mm.

In the case of the liver data there are some local minima a long distance away from the registration position which caused some registrations to fail. Again the stochasticICP algorithm was less affected by these local minima than the standard ICP algorithm. In both the liver and the skull datasets there are a number of local minima clustered around the global minimum. The lower RMS distance value when using the stochasticICP algorithm is believed to be because this algorithm was able to skip over or climb out of these local minima, whereas the standard ICP algorithm tended to become trapped in the first local minimum which was encountered.

The usual method of avoiding local minima using the standard ICP algorithm is to use a number of starting positions. Therefore, although the stochasticICP algorithm requires more iterations, its increased robustness should mean that fewer starting positions are required and so the stochasticICP algorithm may be ultimately more computationally efficient.

Our stochasticICP method could be thought of as combining a simulated annealing optimisation method [9] with the ICP algorithm, where the size of perturbation is analogous to the temperature parameter. This is a much more integrated combination of the two types of algorithm than the method proposed by Luck *et al.* [5], where a simulated annealing algorithm is used to provide good starting estimates for an ICP algorithm. Our method could also be thought of as combining a multi-scale or multi-resolution optimisation method (which have been found to be robust methods for voxel based registration [10]) with the ICP algorithm. In this case the size of the perturbation is analogous to the size of the blurring kernel used.

Future work on the stochastICP algorithm will include an investigation into the factors affecting the values of the additional parameters, in particular, σ and t . The values of these parameters are expected to be influenced by factors such as the shape of the object being registered and also the amount of noise in the source and target datasets. Work is also required to investigate the robustness of the algorithm to changes in these parameters. We also intend to extend our validation of the algorithm by greatly increasing the number of datasets used in our investigation and to compare the algorithm with other, more robust, ICP algorithms [4,5].

5 Conclusions

We have developed a stochastic ICP algorithm. From our results using three datasets it appears that, although our stochastICP algorithm requires more iterations, in the presence of local minima it is more robust (i.e. yields fewer failures) and more precise than the standard ICP algorithm.

6 Acknowledgements

The authors would like to thank the EPSRC (grant numbers GR/M53752, GR/L62221 (via the MedLINK programme), GR/N04867 and PhD studentship for J.M. Blackall). We would also like to thank Deputy for providing the femur phantom.

References

1. P.J. Besl and N.D. McKay. A method for registration of 3-D shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.
2. J. Feldmar, N. Ayache, and F. Betting. 3D-2D projective registration of free-form curves and surfaces. *Comput. Vision Image Understanding*, 65(3):403–424, 1997.
3. C.R. Maurer, Jr., G.B. Aboutanos, B.M. Dawant, Maciunas R.J., and J.M. Fitzpatrick. Registration of 3-D images using weighted geometrical features. *IEEE Trans. Med. Imaging*, 15(6):836–849, 1996.
4. T. Masuda and N. Yokoya. A robust method for registration and segmentation of multiple range images. *Comput. Vision Image Understanding*, 61(3):295–307, 1995.
5. J.P. Luck, W.A. Hoff, R.G. Underwood, and C.Q. Little. Registration of range data using a hybrid simulated annealing and iterative closest point algorithm. submitted to IEEE PAMI. available at <http://egweb.mines.edu/whoff/publications/2000/pami2000.pdf>
6. W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3-D surface reconstruction algorithm. *Computer Graphics*, 21(4):163–169, 1987.
7. W. Schroeder, K. Martin, B. Lorensen, L. Avila, R. Avila, and C. Law. *The Visualization Toolkit: An Object-Oriented Approach to 3-D Graphics*. Prentice-Hall, 1997.
8. M.R. Fenlon, A.S. Jusczyck, P.J. Edwards, and A.P. King. Acrylic resin dental stent for image guided surgery. *J. of Prosthetic Dentistry*, 83(4):482–485, 2000.
9. S. Kirkpatrick, C.D. Gelatt, Jr., and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598), 1983.
10. C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated 3D registration of MR and CT images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.