Automated 3D Registration of Truncated MR and CT Images of the Head

C.Studholme, D.L.G.Hill, D.J.Hawkes
Radiological Sciences
UMDS, Guys Hospital
London SE1 9RT
C.Studholme@umds.ac.uk

Abstract

This paper presents a fully automated technique for the rigid body registration of clinically acquired MR and CT images of the head. We describe our multi-resolution approach to the optimisation of voxel similarity measures and evaluate the performance of a number of measures when presented with clinical images with a small overlapping volume. Results are compared with those derived from manual registration by identification of corresponding point landmarks. We show that by limiting the measures to intra-cranial regions of the images, not containing deformable skin surface features, a greater accuracy may be provided for certain types of truncated image.

1 Introduction

It is common for patients to be imaged with more than one tomographic radiological imaging modality for the purposes of diagnosis or treatment planning. These three dimensional modalities provide complementary information about pathology, and associated normal anatomy. Variations in patient orientation, and differences in resolution and contrast of the modalities can make it difficult for a clinician to mentally fuse all the image information accurately. For this reason, there has been considerable interest in using image registration techniques to transfer all the acquired image information into a common coordinate frame. Current medical image combination techniques aim to present the registered images to the clinician in a way that makes the fusion process easier.

Most medical image registration algorithms have been developed for use in the head, where accurate registration is clinically important, and because it is normally possible to assume that the rigid bone of the skull restricts the transformation to the six degrees of freedom of a rigid body (three translations and three rotations).

In this paper, we concentrate on the registration of MR and CT images of the head, using rigid body transformations. Several algorithms have been proposed to solve this problem. These techniques make use of equivalent image features (point landmarks [1] and surface matching [2, 3, 4]), or voxel similarity measures [5, 6, 7]. The algorithms that match equivalent features almost invariably require some user interaction to identify these features, whereas the approaches based on purely voxel similarity can be fully automated. In addition features such as surfaces may not correspond to the same anatomical structure [4] and there may be significant inaccuracies in their location.

BMVC 1995 doi:10.5244/C.9.3

All the techniques cited above have been shown to work for whole head data. Many clinical images are, however, highly truncated, ie: only a small proportion of the head is imaged in one or both modalities. This is especially true of CT data, where the number of slices acquired is minimised in order to keep radiation dose to the patient low. To further reduce the radiation dose, it is common for some slices to be acquired with a larger slice spacing, and lower in-plane resolution. MR data can also be truncated, in order to keep imaging times to a minimum.

A clinically usable registration algorithm ideally needs to be fully automated and robust when used on typical clinically acquired data. We have implemented several voxel similarly measures and developed a simple and efficient multi-resolution approach to optimisation. We compare the way they perform when presented with clinical data in which one or both modalities is truncated.

2 Method

2.1 Voxel Similarity Measures

In this section we describe a number of similarity measures in terms of the 2D distribution or histogram of voxel intensities. Figure 3 provides an example MR and CT image from the same anatomical region of the same patient to illustrate the difference in intensities in each modality. Figure 1(a) shows a 2D distribution of intensities for manually registered MR and CT image volumes from a single patient. Bright areas in the images correspond to large numbers of voxels with those MR and CT intensities. These occur near the origin (both MR and CT have low values for air) and in the strong vertical line in the centre of the distribution where a large range of soft tissue features differentiated by MR correspond to a small range of CT intensities. The diagonal features of positive gradient at lower CT intensities correspond to air/soft tissue boundaries, and the diagonal features of negative gradient correspond to soft tissue/bone boundaries. Figure 1(b) shows a distribution due to a translational misregistration. In a normalised distribution $p(m, c)$ is the probability of an MR voxel with value m having a corresponding CT value c. The probability of a particular MR value occurring is simply $p(m) = \sum_c p(m, c)$ and of a CT value occurring is $p(c) = \sum_m p(m, c)$.

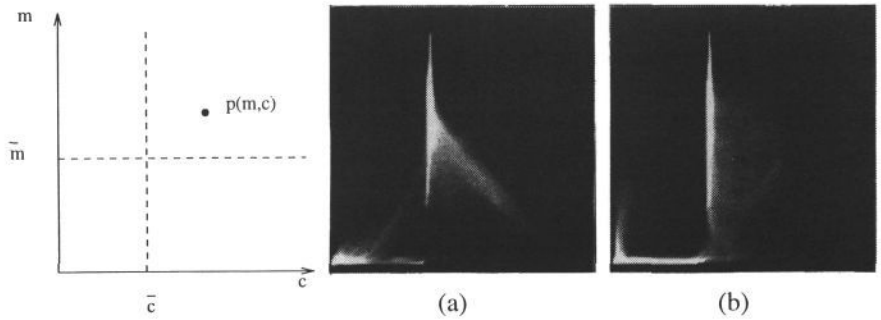

Figure 1: 2D histograms created from a CT (horizontal) and MR (vertical) image pair registered (a) and misregistered by translation along the cranio-caudal axis by 5 voxels (b).

2.1.1 CT Soft Tissue Correlation (measure (i))

These methods attempt to exploit the diagonal features present in the 2D histogram. Van den Elsen et al [7] proposed a number of modifications to direct cross correlation of MR and CT intensity. In these algorithms, ranges of CT intensities are remapped prior to calculation of the correlation coefficient. In this work we have used a triangular remapping over the region of CT intensities corresponding to the Hounsfield units for soft-tissue. Van den Elsen used a form of multi-resolution and multi-start exhaustive search to avoid the local optima. In our implementation we use the single start multi-resolution optimisation described in section 2.2. In addition we have used the correlation coefficient γ (which is normalised with respect to the number of overlapping voxels) as a measure between the MR and modified CT rather than direct correlation:

$$\gamma = \frac{\sum_{m,c}(m-\overline{m}).(c-\overline{c}).p(m,c)}{\{\sum_{m,c} p(m,c)(c-\overline{c})^2 . \sum_{m,c} p(m,c)(m-\overline{m})^2\}^{1/2}}$$

2.1.2 Corresponding MR and CT Variance (measures (ii) and (iii))

These are based on a measure used by Woods [9] for the registration of segmented MR images of the brain to PET images. Hill proposed the minimisation of Corresponding MR variance for MR-CT registration in [5] (measure (ii)). It makes the assumption that, for a given CT intensity, the range of intensities of corresponding MR voxels is minimised at registration. This can be expressed as a weighted sum of normalised standard deviations of corresponding MR values. This assumes that, for a given CT intensity, there is a single corresponding MR intensity at registration.

$$\sigma'_m = \sum_c [\sigma_m(c).p(c)]$$

The weighting ensures that the measure is influenced most strongly by MR variance for the most common CT values. For these tests we have compared both the minimisation of Corresponding MR variance and Corresponding CT variance (measure (iii)):

$$\sigma'_c = \sum_m [\sigma_c(m).p(m)]$$

2.1.3 Relative Entropy (measure (iv))

This measure, otherwise known as mutual information, is derived from an information theoretic approach to the dependence of one variable on another, when the precise functional form is unknown [8]. It has been independently applied to medical image data by Collignon[6] for tomographic image registration and by Viola [10] for projection image registration. It is based on the shared information or relative entropy between the overlapping regions in the two images, which at registration should be maxmimised. The mutual information between an MR image with intensity m and a CT image with intensity c defined for the 2D distribution is:

$$S(M;C) = \sum_{m,c} \left\{ p(m,c) \log \frac{p(m,c)}{p(m)p(c)} \right\}$$

2.2 Calculation and Optimisation of Measures

To solve the rigid body registration problem we must find the transformation T which maps voxels in the MR volume to voxels within the CT volume. The data is generally sampled at different rates within plane and between plane. In addition sampling is usually not the same in MR and CT acquisitions. To simplify the registration problem we first resample all the data to $1mm^3$ voxels using voxel averaging for resolution reduction by the nearest integer factor and tri-linear intensity interpolation. A multi-resolution pyramid of these images is then produced using voxel averaging.

Given a transformation T mapping CT image c in volume V_c to MR image m in volume V_m , we must employ some form of interpolation to decide which CT value corresponds to a given MR value when evaluating a similarity measure M. We need to find the value of T such that for the measure M which is minimised at registration,

$$T_r = \min_T \{M(m(V_m), c(TV_c \cap V_m))\} \tag{1}$$

Choice of interpolation scheme and method of optimisation can both have a large effect on the computation time and robustness of the algorithm. Other approaches have used a variety of optimisation schemes including Powells Conjugate Gradient [6], multi-resolution exhaustive searches [7] and the Genetic Algorithm [5]. For this work we have combined multi-resolution evaluation and computationally efficient nearest neighbour interpolation. For a starting estimate T_0 at a particular data resolution r_l at level l in the pyramid, we evaluate a measure for a set of 13 transformations $\mathcal{T}(T_0, r_l)$. These are the current starting estimate and the starting estimate with increments and decrements of each of the 3 translations ($\pm\delta t$) and 3 rotations ($\pm\delta\theta$) by a step size related to the data resolution r_l to avoid aliasing effects. We can look for a better estimate of the registration transformation T_1 such that:

$$T_1 = \min_{T \in \mathcal{T}(T_0, r_l)} \{M(m(V_m), c(TV_c \cap V_m))\} \tag{2}$$

If $T_1 \neq T_0$ then we can repeat the search with $\mathcal{T}(T_1, r_l)$ until no further improvement can be made. Selection of step size to create $\mathcal{T}(T_0, r_l)$ based on r_l is important. We start the optimisation at low data resolution with r_l large and optimise removing large differences in registration. At a resolution r_l we set the translational step size δt to r_l and the rotational step size is determined by the transaxial field of view f and r_l such that $\delta\theta = \arctan(2r_l/f)$. Following optimisation at level l with sampling resolution r_l we continue to optimise at this resolution but with a step size based on the resolution at the next level down in the pyramid r_{l-1}. When this is complete we proceed to level $l-1$.

For all results presented here optimisations were initiated at a data sampling resolution of 16mm. In addition to the oct-tree resolutions of 1mm, 2mm, 4mm, 8mm and 16mm, intermediate resolutions of 1.5mm, 3mm, 6mm and 12mm were also used (constructed by tri-linear interpolation from the next highest resolution).

2.3 Comparison of Measures

The data used for the tests consisted of five pairs of clinically acquired MR and CT images of the brain. The image pair for patient A was acquired for the planning of epilepsy surgery and B to E were used to plan the removal of skull base tumours. The MR images were all T1 weighted Gadolinium enhanced 2D acquisitions except for patient A which was a 3D

gradient echo image. The CT datasets range in size from the largest (patient A) to the more common targeted volumes of patients B to E (see table 1). The CT image for patient A was synthesised by combining 11 slices acquired around the skull base and 17 slices acquired around the top of the head both with 512×512 pixel $3mm$ slices, with seven 256×256 $6mm$ slices (which we tri-linear interpolated to 512×512 $3mm$ slices) acquired through the centre of the skull.

The images were initially misregistered because of differences in patient orientation in the two scanners. All slices were nominally transaxial and were initially registered by manual identification of corresponding point landmarks providing an accuracy of between $1mm$ and $2mm$ intracranially. These estimates are shown in table 2, where the X axis is the patient right-left, the Y axis from front to back and the Z axis feet to head. The main rotational difference occurs in the X axis (patient nodding) which is due to differences in head rest construction and CT targeting parameters used to avoid radiation dose to the eyes.

The five image pairs were registered using all four similarity measures. For the starting estimate the centre of the two datasets were co-located with no rotations. The registration solutions obtained using the similarity measures were assessed by comparison with the manual solutions and by visual inspection. No ground truth registration was available and in practice is very hard to obtain with sufficient accuracy. Visual inspection adequately demonstrates mis-registration of a voxel or greater perpendicular to tissue boundaries.

Differences in head restraints and patient positioning can cause deformation of soft tissue structures such as the ears and sub-cutaneous fat, both of which appear with high contrast in MR and CT. When the field of view is large only a small proportion of voxels are affected. When only a small axial volume is sampled, particularly when this is at the base of the skull, the proportion of deformed voxels can be much higher. In an extension of the algorithm V_c in equation (2) is replaced by $V_c^{1/4}$ where this is the CT volume truncated transaxially to 125mm in X and Y to exclude most extracranial tissue as illustrated in figure 3.

MR and CT Image Volume Parameters				
	MR		CT	
Patient	Num Voxels	Voxel size (mm)	Num Voxels	Voxel size (mm)
A	256x256x124	0.938x0.938x1.5	512x512x42*	0.488x0.488x3.0
B	256x256x36	0.918x0.918x3.0	512x512x28	0.488x0.488x3.0
C	256x256x30	0.820x0.820x3.0	512x512x25	0.488x0.488x3.0
D	256x256x30	0.898x0.898x3.0	512x512x19	0.488x0.488x3.0
E	256x256x7	0.859x0.859x3.0	512x512x18	0.488x0.488x1.5

Table 1: MR and CT image volume parameters. *after interpolation as described in the text.

3 Results

3.1 Whole Image Registration

Patient A containing an almost complete skull image in both MR and CT provided consistent estimates from all of the measures, given the $1mm^3$ data resolution and optimisation

Manual MR-CT Registration Estimates								
Patient	Translation (mm)			Rotation (deg.)			R.M.S.	Num.
	Tx	Ty	Tz	Rx	Ry	Rz	Error	Points
A	8.4	33.7	2.1	24.1	-6.4	-0.9	2.55	12
B	3.56	24.0	-0.2	17.8	-0.6	-5.14	1.37	13
C	0.3	12.0	-9.5	1.3	-0.7	2.3	1.37	13
D	-0.4	13.9	4.9	0.9	-4.4	4.9	2.37	11
E	5.8	-0.1	-3.3	11.4	1.2	0.89	1.64	8

Table 2: Transformation parameters estimated by manual point landmark identification.

used. An expert visually inspected the results. The automatic solutions were indistinguishable one from another, but all were found to be more accurate than the manual solution. Images for patient B produced comparable results with all but the corresponding MR variance measure. Comparison of the solution found with measures (i),(ii) and (iv) and manual estimates by visual inspection again showed that the automated algorithms produced a more accurate result.

The solutions found for patient C using measures (ii), (iii) and (iv) were visually indistinguishable, and appeared more accurate than the solution found using manual registration as illustrated in figure 2. The result using soft tissue correlation was appreciably less accurate. Applying further optimisation by reducing the step size for $1mm^3$ resolution data resulted in an estimate closer to that of the other three measures. This may indicate that a more complex registration parameter space is produced by this measure.

Results for patient D show three similar registrations from measures (i),(iii) and (iv), with a complete failure for the MR variance measure. Visual inspection of the results obtained using the successful measures shows a considerable improvement over that produced by manual registration. The interleaved multi-slice 2D spin echo MR sequence used for this patient contained motion artefact which probably contributed to a poor point based estimate.

The image set for patient E was the most severely truncated, with only a limited axial field of view targeted around a small acoustic neuroma in both the MR and CT images. Optimisation of measures (i) and (iii) failed to improve on the starting estimate. Both the relative entropy and corresponding MR variance produced results which were near to but noticeably worse than the solution found using manual registration. The images contained a considerable proportion of deformed external skin features as illustrated in figure 3 (right).

3.2 Intra-Cranial Image Registration

To remove the effects of the majority of skin surface deformations on the intra-cranial rigid body registration, the images from patient E were re-registered using a restricted axial volume as shown in figure 3. All the similarity measures except the relative entropy measure produced poor results as shown in table 3. Final estimates for measures (i),(ii), and (iii) are pathological as optimisation was terminated when the overlap between datasets fell below a pre-set threshold of 40%. This seems to indicate that these measures were more dependent on the air/skin boundary for their registration estimate. Visibly the entropy result showed considerable improvement over the full volume registration parameters as can been seen in figure 4.

Automated MR-CT Registration Estimates							
Patient	Measure	Translation (mm)			Rotation (deg.)		
		Tx	Ty	Tz	Rx	Ry	Rz
A	(i)	9.0	33.3	0.3	24.0	-6.0	-0.3
	(ii)	8.6	32.5	0.8	23.6	-6.8	-0.8
	(iii)	9.0	33.0	1.0	24.0	-6.3	-0.3
	(iv)	9.0	32.0	0.3	23.0	-6.3	-0.3
B	(i)	3.3	24.0	1.0	18.3	-1.3	-6.0
	(ii)	3.8	24.6	2.7	14.9	-1.7	-5.3
	(iii)	4.0	24.0	1.7	16.3	-1.3	-5.3
	(iv)	4.0	24.0	0.0	17.3	-1.3	-5.3
C	(i)	0.0	11.6	-9.0	1.7	0.0	3.0
	(ii)	0.3	12.4	-8.0	-3.0	0.1	2.7
	(iii)	0.0	12.7	-8.3	-2.7	0.0	2.3
	(iv)	0.0	12.0	-8.7	-2.7	0.0	2.7
D	(i)	0.0	13.3	10.0	-0.3	-4.7	5.3
	(ii)*	2.0	10.0	26.1	0.5	-5.4	5.26
	(iii)	0.0	13.3	8.3	-3.3	-5.3	4.7
	(iv)	0.0	13.3	8.0	-2.7	4.7	4.7
E	(i)*	7.3	-6.5	-16.5	-12.3	-1.7	0.5
	(ii)	4.8	4.1	0.3	9.7	0.0	0.0
	(iii)*	5.3	2.2	-5.5	14.7	0.0	0.5
	(iv)	5.3	3.0	-1.0	11.0	0.0	0.3
E (trunc)	(i)*	26.0	-5.0	8.0	-2.9	-2.0	19.0
	(ii)*	-17.3	-3.0	0.0	0.0	2.6	0.0
	(iii)*	10.4	7.6	-11.7	-9.8	16.3	1.7
	(iv)	6.1	0.75	-2.8	13.4	0.3	1.2

Table 3: Transformation parameters estimated by optimisation of (i) CT soft tissue, (ii) Corresponding MR Variance,(iii) Corresponding CT Variance,(iv) Relative Entropy. *Indicates significant failure.

4 Conclusion

We have evaluated the performance of a variety of voxel similarity measures for the registration of MR and CT images of the head. In all cases we have used a simple and efficient multi-resolution gradient descent optimisation algorithm, with step size related to image resolution to avoid aliasing effects. We have confirmed that all these methods perform well when the images being registered have a large field of view, covering most of the head. Many clinically acquired images, however, have a limited field of view, and it is also common for a patient's CT scans to be composed of slices with different resolutions. We have found that the measures we are evaluating can be distinguished in terms of their ability to register the truncated datasets. The relative entropy algorithm was the only measure that successfully registered all the test datasets.

Registration using voxel similarity measures has the advantage that all overlapping (or corresponding) image data is used in the computation of the optimisation measure. They are, therefore, less prone to data truncation, image noise, image artefact and small tissue deformations than surface based methods. The disadvantage of using all overlapping im-

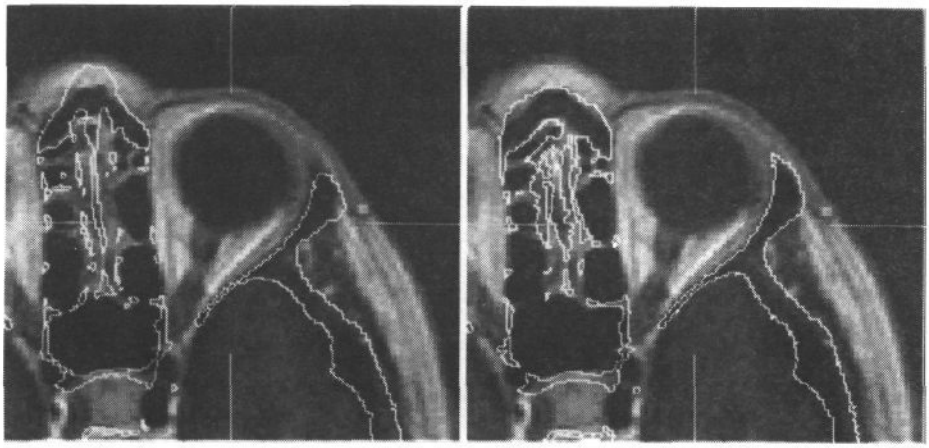

Figure 2: Images showing transaxial slices through the eye in the MR volume for patient C with CT bone threshold boundary in white for the manual registration estimate (left) and the estimate from minimisation of corresponding MR variance (right).

age data for registration is that regions of the patient that deform significantly (eg: the ear) can contribute to the registration solution, leading to inaccurate registration of rigid structures of interest. We have demonstrated that this problem can be overcome simply by excluding regions of the image that are likely to deform (eg: skin surface, mandible) from the calculation of the measure. This process does not require accurate segmentation. A further benefit of excluding some structures from the registration is that fewer voxels are used to calculate the measures, resulting in a reduced computational load.

To summarize we have shown that the relative entropy of the joint intensity feature space is an effective measure for voxel based registration in this application. Careful attention to sampling and a simple multi-resolution search provides a robust and efficient optimisation algorithm. The algorithm is robust to the truncation that results from poor sampling, and exclusion of deformable regions such as the skin surface.

Acknowledgements

This work was funded by UK EPSRC Grant GR/H48170. We are grateful for the support and encouragement of our clinical colleagues in this work, in particular Prof. Michael Gleeson (ENT Surgeon), Mr. Anthony Strong (Neurosurgeon), Dr. Tim Cox (Neuroradiologist), and for the technical assistance of the Radiographic staff of Guy's, St. Thomas', The Maudsley and St. George's Hospitals in London. We are grateful to William Wells and Andre Collignon for useful discussions, and copies of work in progress.

References

[1] Hill DLG, Hawkes DJ, Crossman JE, Gleeson MJ, Cox TCS, Bracey EECML, Strong AJ, Graves P. "Registration of MR and CT images for skull base surgery using point-like anatomical features", British Journal of Radiology,**64**,1991,pp 1030-1035.

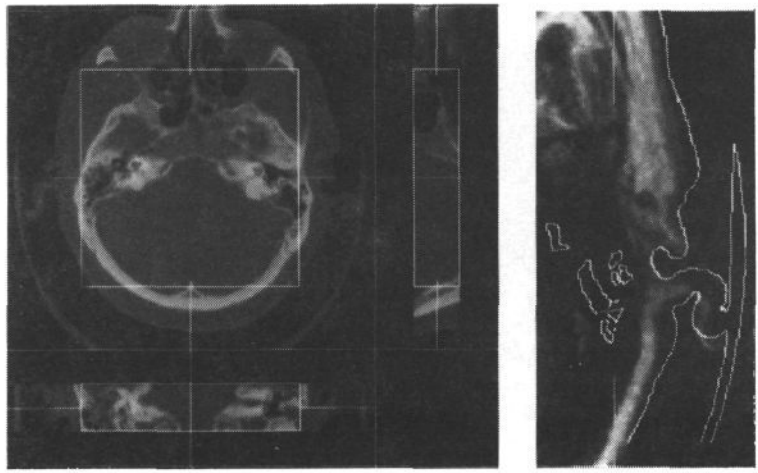

Figure 3: Images showing orthogonal slices through the full CT volume for patient E (left), resulting registration using relative entropy showing deformed CT skin surface threshold on grey level MR (right). White box outline (left) shows truncated CT volume used to limit registration measures to intra-cranial portions of the image.

[2] Pelizzari CA, Chen GTY, Spelbring DR, Weichselbaum RR, Chen C-T. "Accurate three dimensional registration of CT, PET and/or MR images of the brain", Journal of Computer Assisted Tomography, **13**, 1989,pp 20-26.

[3] Jiang H, Robb RA, Holton KS. "New approach to 3-D registration of multimodality medical images by surface matching", *in Proceedings of Visualisation in Biomedical Computing*, Chapel Hill, N.C.,U.S.A., 1992,pp 196-213.

[4] Hill DLG, Hawkes DJ. "Medical image registration using knowledge of adjacency of anatomical structures", Image and Vision Computing,**12**,1994,pp 173-178.

[5] Hill DLG, Hawkes DJ, Harrison NA, Ruff CF. "A Strategy for Automated Multimodality Image Registration Incorporating Anatomical Knowledge and Imager Characteristics", *in Proceedings of Information Processing in Medical Imaging*, Flagstaff,U.S.A., 1993,pp 182-196.

[6] Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marchal G. "Automated Multi-Modality Image Registration based on Information Theory", *in Proceedings of Information Processing in Medical Imaging*, Brest,France,1995,pp 263-274.

[7] Van den Elsen PA, Pol EJD, Sumanawaeera TS, Hemler PF, Napel S, Adler JR. "Grey Value Correlation techniques used for automatic matching of CT and MR brain and Spine images", *in Proceedings of Visualisation in Biomedical Computing*,Rochester Mn,U.S.A. ,1994, pp 227-237.

[8] Carlson AB. *Communication Systems*, McGraw Hill,Singapore, 3rd Ed.,1986, pp 574-587.

[9] Woods RP, Mazziotta JC, Cherry SR. "MRI-PET registration with automated algorithm", Journal of Computer Assisted Tomography, **17**,1993, pp 536-346.

[10] Viola P, Wells W. "Alignment by Maximization of Mutual Information", *in Proceedings of the 5'th international conference on computer vision*, Boston, U.S.A.,1995, pp 16-23.

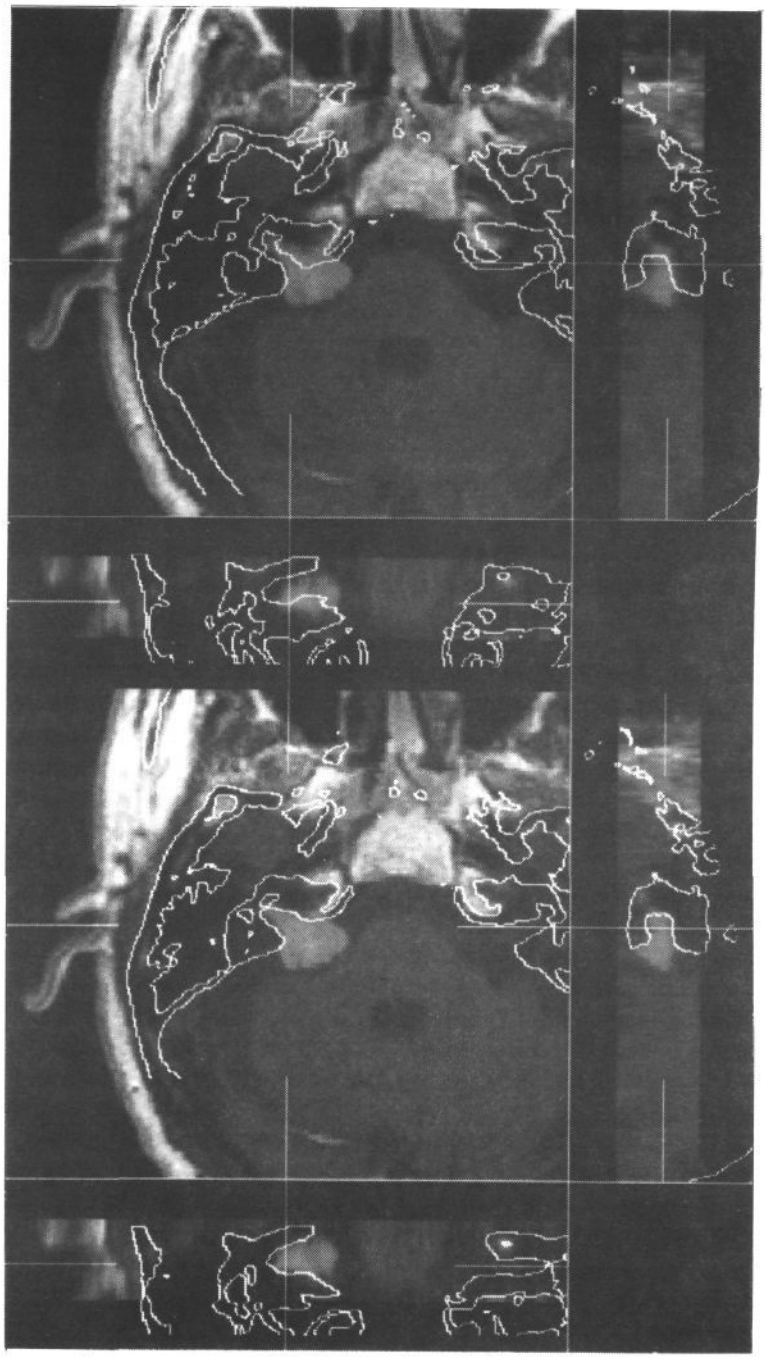

Figure 4: Orthogonal slices through a point inside the internal auditory canal in the MR grey level image for patient E. Registration results for relative entropy using full volume (top) and intra-cranialy truncated (bottom) CT are shown by overlay of CT bone threshold boundary in white.