AUTOMATED MULTI-MODALITY IMAGE REGISTRATION BASED ON INFORMATION THEORY

A. COLLIGNON, F. MAES, D. DELAERE, D. VANDERMEULEN, P. SUETENS AND G. MARCHAL

Laboratory for Medical Imaging Research, ESAT-Radiology
K. U. Leuven, Kardinaal Mercierlaan 94, 3001 Heverlee, and
U. Z. Gasthuisberg, Radiology, Herestraat 49, 3000 Leuven,
Belgium

Abstract. We propose an information theoretic approach to the rigid body registration of 3D multi-modality medical image data. The "mutual information" of grey-value pairs is proposed as a new matching criterion. The new voxel similarity based (VB) registration algorithm is applied to real world images of the human head (CT, MR, and PET volumes), and compared to stereotactic and 2 other voxel based registration solutions. Our results show that subvoxel accuracy can be obtained completely automatically and without any pre-segmentation.

1. Introduction

Rigid body registration is a fundamental task in three-dimensional (3-D) medical image processing that serves to align two or more 3-D scans that were taken at a different time, or generated on different medical imaging modalities. There are numerous applications. Medical diagnosis benefits from the complementarity of the information in images of different modalities. During radiation therapy planning dose calculation needs to be done based on the CT data while tumor outlining is often much better performed in the corresponding MR scan. The analysis of brain function requires the correlation of physiological data from PET data with anatomical data from the corresponding MR data, etc.

A general overview of image registration techniques has been compiled by Brown [1]. Two other articles survey image registration techniques and their

applications in medical imaging [9], [13]. From these surveys we learn that
the bulk of registration algorithms are either frame based, point landmark
based (using external and/or internal landmarks), surface based, or voxel
based. Point landmark and surface based registration algorithms are typ-
ically used for the alignment of anatomical scans, while voxel similarity
based registration algorithms seem to be more robust for alignment prob-
lems involving functional scans ([7] and [15]), although many people are
using surface based registration algorithms in the latter case also.

Voxel based (VB) registration algorithms optimize a function measuring
the similarity of all possible pairs of voxel grey-values (or derived features,
e.g. gradients). The main advantage of VB methods is that feature calcula-
tion is straightforward or even absent when only grey-values are used and
thus, in theory, accuracy is not limited by segmentation errors as in surface
based algorithms. Simple correlation of grey-values works only for different
modalities if there exists a *linear* mapping between grey-values [1], which is
unfortunately not always the case. Nonetheless, interest in VB registration
algorithms has recently revived by the success of Woods' AIR package in
clinical practice [16], which has been made more flexible by Hill et al [7].
Their basic assumption is that the grey-value mapping should be a func-
tion (e.g. to each MR grey-value correspondends at most one PET-value).
They accomplish this heuristically by demanding that a global measure of
the variance of the grey-value ratios, with the function's domain variable
as the ratios' denominator, be minimal. We recently compared these and
some other VB matching criteria [2]. All of them are heuristic in nature.
In the same work we first suggested an information theoretic approach to
the design of a non-heuristic multi-modality matching criterion and we pro-
posed to use the entropy of the joint probability distribution $p(g_1, g_2)$ of
the combined grey-values of all common voxel pairs in the two images to
be matched:

$$E(g_1; g_2) = \Leftrightarrow \sum_{g_1, g_2} p(g_1, g_2) \log (p(g_1, g_2)) \qquad (1)$$

In this article we use a basic theorem of statistical inference theory to de-
rive a new version of our entropy based VB matching criterion: the mutual
information or relative entropy of the joint probability distribution of the
grey-values of two (or more) scans. This criterion is more robust with re-
spect to the partial overlap problem and has some interesting mathematical
properties as well. Furthermore it is based on the original grey-values only
and requires no segmentation at all! Therefore, our new VB algorithm is
more robust than any of the surface based registration algorithms and than
any of the heuristic VB registration algorithms. We show results obtained

with real world data in section 3. Those results are compared with that of stereotactic and 2 other VB registration methods. First, however, we will present the theory and discuss the implementation issues in detail. In the discussion of the results we discuss robustness and accuracy of our approach, and compare it to other VB algorithms.

2. Method

2.1. THEORY

In the sequel we restrict the geometric transformations to *rigid body transformations* T_α (α consists of 3 translations and 3 rotations). The rigid body assumption is well satisfied inside the skull in 3-D scans of the head, if abstraction is made of scanner calibration problems (that can lead to inaccurately known voxel sizes) and problems of geometrical distortions (e.g. in MRI), both of which can be dealt with by careful calibration and scan parameter selection respectively. The applicability of the mutual information matching criterion is not restricted to rigid body registration, however.
The selection of the new VB matching criterion (like in Woods [15], Hill [8], and Collignon [2]) is based on the observation that the two-dimensional grey-value histogram or scatter plot of the common voxel pairs of a pair of registered images will be dispersed by misregistration as a function of $\mid \alpha^* \Leftrightarrow \alpha \mid$, where α^* is the registration solution.
We will now present a new measure of scatter-plot dispersion that also has an interesting mathematical interpretation: relative entropy or mutual information of the scatter-plot. Intuitively, we expect the dispersion to be minimal when the interdependence of the grey-values of both images to be registered is maximal. Two signals are maximally interdependent when they are related by a one-to-one mapping, i.e. when g_1 is a function of g_2 and vice versa. This is seldom so for multi-modality data. As it happens, the mathematical definition of mutual information [3]:

$$S(X;Y) = \sum_{x,y} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) \tag{2}$$

also known as Shannon information of two random variables X and Y with marginal and joint probability distributions $p(x)$, $p(y)$ and $p(x,y)$ resp., satisfies the following properties (for a proof see [11]):

1. Symmetry: $S(X;Y) = S(Y;X)$
2. $S(X;Y) = 0 \Leftrightarrow (X$ and Y are independent)
3. Non-negative and bounded: $0 \leq S(X;Y) \leq \min(S(X;X), S(Y;Y))$
4. Invariance: (T is a one-to-one mapping) $\Rightarrow S(X;T(X)) = S(X;X)$

Interpretation of these properties confirms our intuition. Mutual information measures the degree of interdependence, which is bounded below by complete independence and bounded above by one-to-one mappings. We conclude that a mathematically sound matching criterion can be obtained by substitution of X and Y for the grey-values g_1 and g_2 of the images to be registered in (2). Moreover, it allows for much more general grey-value relationships than do the matching criteria proposed by Woods and Hill et al or than any of the other currently available heuristic ones (see [2] for an overview).

2.2. IMPLEMENTATION

The algorithm can be summarized by the following equation:

$$\alpha^* = \arg\ \min_{\alpha}\ \left(\Leftrightarrow\sum_s p(f_1(T_\alpha s), f_2(s))\log_2\left[\frac{p(f_1(T_\alpha s), f_2(s))}{p(f_1(T_\alpha s))\ p(f_2(s))}\right] \right) \quad (3)$$

where s is a sample of coordinates inside the volume of feature image f_2, $f_1(T_\alpha s)$ is the set of grey-values from feature image f_1 sampled at the rigidly transformed coordinates $T_\alpha s$, and $p(f_1, f_2)$ is obtained by normalizing the 2-D scatter-plot, $H(f_1, f_2)$, of the grey-values of that part of sample s that falls in the common volume of f_1 and f_2, into a joint probability distribution as follows:

$$p(f_1, f_2) = \frac{H(f_1, f_2)}{N} \quad (4)$$

where N is the number of coordinates in the overlapping part of sample s, and where the features $f_i = \frac{(g_i - g_i^{min})}{n_i}$ are rescaled versions of the original grey-values. n_i are the rescaling factors. Due to the rigid body assumption we have restricted T_α to a superposition of a 3-D rotation and a 3-D translation: $T_\alpha s = R_\alpha s + t_\alpha$.

2.2.1. *Sample selection*
Feature values for image f_2 are known for a finite number of discrete voxel positions only. We need to decide how many and which voxel grey-values to include in our calculation of the mutual information matching criterion. Either the coordinates of all voxels of f_2, or a subset, or a superset thereof are used in the calculation of the mutual information for every α considered by the optimisation algorithm. In case a superset of the voxel coordinates is used, we need to select an interpolation method to find the grey-values at coordinates that do not coincide with the voxel coordinates of f_2. In order to keep the speed performance as high as possible we have simply used nearest neighbour interpolation to obtain samples from f_2. The aim of sub-sampling is to increase speed performance. It should be noted that some continuity

of grey-values is assumed if we expect the mutual information to be well-behaved in case of sub-sampling. The grey-value continuity assumption implies that $p(f_1(\alpha))$ will vary smoothly with α. This additional assumption is not a restriction as long as sub-sampling factors (defined as the fraction of samples taken per voxel size, in a certain direction) do not become too small. The aim of super-sampling is to increase accuracy. This can only work if the same continuity assumption is valid in both images, or equivalently if a meaningful interpolation method can be defined in both images.

2.2.2. *Interpolation*

Due to the continuity of the registration parameters α, regardless of whether the sample s is a sub-, or a super-set of all voxels in f_2, the transformed sample coordinates will always fall in between voxel coordinates of f_1, and thus interpolation is needed to obtain values from f_1. An interpolation method other than nearest neighbour interpolation is required to be able to guarantee subvoxel accuracy. The next best choice in speed is trilinear interpolation. The problem of trilinear or for that matter any other classical interpolation method is that they may introduce new grey-values, i.e. rescaled grey-values not present in the set of the original grey-values of f_1. Thus, any classical interpolation method other than nearest neighbour interpolation may introduce unpredictable changes into $p(f_1)$ as a function of changes in α. To avoid this problem we propose to use *trilinear partial volume distribution* which is dual to trilinear interpolation. It uses the same interpolation weights but instead of using them to average grey-values from f_1, they are used to distribute the voxel volume of the grey-value from f_2 *in the scatter-plot* over the grey-values of all nearest neighbours in f_1.

2.2.3. *Probability Estimation*

The evaluation of (2) requires the knowledge of grey-value probabilities which can be estimated from the grey-value scatter-plot. If quantisation of the scatter plot is too fine, application of parzen-windowing [5] can be used to reduce the effect of image noise and to increase the reliability of the probability estimates. However, parzen-windowing is rather expensive from a computational point of view. Clearly some compression of the data will benefit both the memory requirements and the quality of the probability distribution estimates. It is expected that small grey-value distortions of the observations do not induce large deviations in the registration solution. Based on these observations we propose to use simple rescalings of the images, f_1 and f_2, instead of the original grey-values, g_1 and g_2. For practical purposes it should be noted that rescaling is equivalent to binning of the grey-value histogram, which can be considered to be a crude approximation of parzen-windowing.

2.2.4. *Search Strategy*

In the search for the optimal value of α we have taken a pragmatic approach: 1) we assume we are able to select a meaningful initial value inside the search engine's attraction pool of the optimal value, 2) we then use a locally converging optimisation algorithm as the search engine. The implementation details are as follows: 1) Put the origins of the physical world coordinate systems in the image centres and assign zero values to those registration parameters that represent the relative translations of these coordinate systems as initial estimates. This is meaningful under the assumption that both images represent approximately the same volume of interest. 2) Apply rotations that are multiples of 90° so that the images are approximately identically oriented. This can be automated based on knowledge about the scanner and the scanning sequence. 3) We have used Powell's locally converging optimisation algorithm [10] for maximization of the mutual information matching criterion. The first 2 to 3 iterations may be performed using nearest neighbour interpolation for its speed, while later iterations may be performed using more complex interpolation for more accuracy. 4) For the same reasons we also suggest to use a subsample of the floating image during the first iterations, and a super-sample during the final iterations.

3. Results

3.1. AN ARTIFICIAL TEST USING A REAL MR IMAGE

We transformed a T1-weighted MR image (from the Heidelberg reference data set, see [4] for a detailed description) over (10.39, -3.27, 2.03) degrees and over (7.06, 1.21,18.26) mm using trilinear resampling. We then used this transformed version MR^T as a reference image to be registered with the original MR image. Eventhough large parts of some slices (slices 21 to 29 of 30) of the reference image have been set to zero in MR^T due to the transformation, the images could be registered. Accuracy results are summarized in table 1 for three different interpolation methods.

3.2. COMPARISON WITH STEREOTACTIC REGISTRATION: CT/MRI/GD-ENHANCED MRI

The data set used in the previous test also contains corresponding CT and Gd-enhanced MR (MRE) images. All the images were acquired under stereotactic conditions. In this section we summarize our registration results relative to the stereotactic registration solution. Note that since MR and MRE were registered by acquisition we used MR^T instead of MR for matching with the MRE image. This test is similar to, but much more re-

TABLE 1. This table gives the registration error for our algorithm when applied to an MR and its transformed version MR^T. Transformation parameters were (10.39, -3.27, 2.03) degrees and (7.06, 1.21,18.26) mm. Voxels are (1.33 x 1.33 x 4.0) mm^3 in both images. Subsample "3.0/1.0" means that only 1 in 3 voxels are sampled in X and Y direction of the image planes, and that all voxels are used in the perpendicular Z direction. The "error"-column gives us an arithmetic average registration error for all voxels in the central quarter of all slices of the MR volume. Binning parameter values were: $n_i = 16$. For MR^T: $g_{max} = 2550$ and $g_{min} = 0$. For MR: $g_{max} = 3359$ and $g_{min} = 2$.

f_1	f_2	Interpolation	sub-sampling factors	error (mm)
MR^T	MR	nearest neighb.	3.0/1.0	0.17
MR^T	MR	trilinear	3.0/1.0	0.0048
MR^T	MR	partial volume	1.0/1.0	0.17

alistic than the one in the previous section, because: 1) the noise will be different now, 2) because of the Gd-enhancement, and 3) due to the visibility of pathology (a tumor) in the MRE volume. A true multi-modality test is of course the registration of CT with MR. Table 2 gives the results.

TABLE 2. Same as table 1, but for MR^T combined with MRE. For MRE: $g_{max} = 3582$ and $g_{min} = 2$. For CT: $g_{max} = 4094$ and $g_{min} = 2$.

f_1	f_2	Interpolation	sub-sampling factors	error (mm)
MR^T	MRE	nearest neighb.	3.0/1.0	0.51
MR^T	MRE	trilinear	3.0/1.0	0.54
MR^T	MRE	partial volume	1.0/1.0	0.40
CT	MR	partial volume	1.1/0.3	2.50

3.3. COMPARISON WITH OTHER VB REGISTRATION ALGORITHMS USING 3-D MRI/CT

We have also compared the outcome of our algorithm with that of two other VB registration algorithms by van den Elsen ([12], [14]) on a CT/MR pair. In both cases her registration was performed automatically, without user interaction or use of fiducial markers. In [12] hierarchical correlation techniques were applied to "ridge"-feature images, obtained from the original data using geometrical operators in scale space. In [14] grey-value correlation is proposed between a grey-value mapping of the CT and the original

MR volumes. The difference between the solutions found are summarized in table 3. By visual inspection it was impossible to see the differences between our solution and the grey-value correlation based method of van den Elsen unless extremely large zooming factors were used. Fig. 1 visualizes our result.

TABLE 3. This table gives registration differences between our mutual information based registration algorithm and both VB algorithms proposed by van den Elsen et al. Voxels are (0.9765625 by 0.9765625 by 1.0) mm^3 for the MRI and (0.9375 by 0.9375 by 1.55) mm^3 for the CT volume. As before the average error Δ_{avg} was calculated over the central quarter of all planes of f_2, i.e. of the CT-volume. $\Delta_t(c)$ is the distance between the transformed coordinates of the volume centroid of f_2. Δ_θ is the angle between the unit eigenvectors corresponding to the unit eigenvalue of the rotation matrix. The solutions using nearest neighbour interpolation were found in less than 7 minutes and sub-sampling (1 voxels in 4, in all directions) while the solutions using partial volume distribution and no sub-, nor super-sampling were found in approximately 4 hours on an IBM RS/6000 workstation.

reference	interpolation	Δ_{avg} (mm)	$\Delta_t(c)$ (mm)	$\Delta_\theta(deg)$
[12]	nearest neighbour	2.72	2.2398	3.6776
	partial volume	2.02	1.6245	0.8957
[14]	nearest neighbour	1.62	0.8563	3.8177
	partial volume	1.26	0.0582	0.8674

3.4. TESTS WITH MRI/PET

The algorithm has been applied to MRI and PET data of Multiple Sclerosis patients for improved diagnosis, and to MRI and PET data of volunteers. In the latter case the registration serves to enhance the accuracy of a functional analysis of the human brain that has been performed earlier [6] using PET data alone. Due to the relatively low resolution of the PET images, even full sample application of our algorithm is fast (approximately 6 minutes for the data shown in fig. 1).

4. Discussion

There never exists a perfect one-to-one grey-value mapping between images to be registered, even in the registered position, due to noise and often also due to the nature of the data involved in multi-modality registration, e.g. skull and background have overlapping intensities in MR while they are completely different in CT. The algorithm proposed by Woods is not
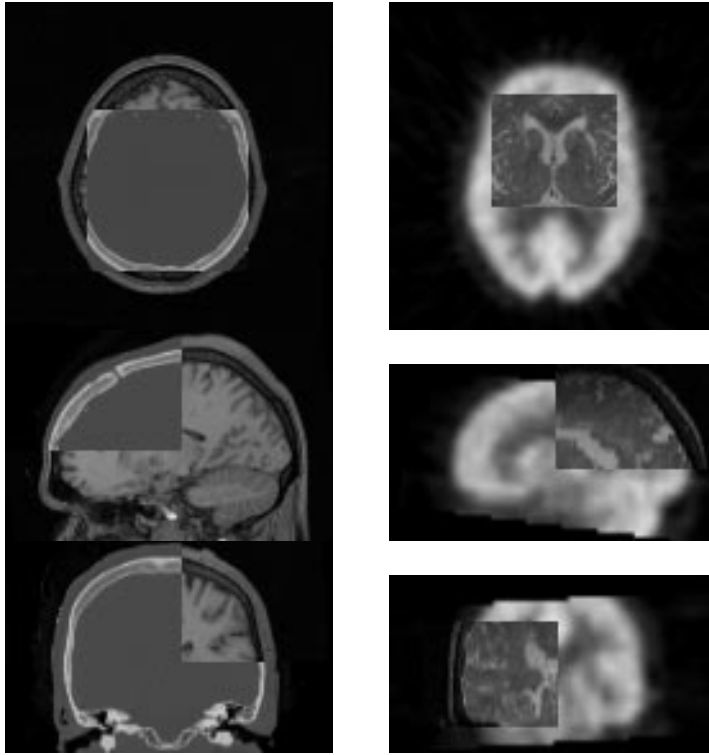
Figure 1.   Left: MRI/CT combination registered using mutual information with partial volume distribution, $n_i = 16$, sampling factors equal to 1. Right: PET/MRI combination of an MS patient. The exact registration solution is unknown. This result was found to be clinically acceptable.

robust enough to handle this type of deviations from the one-to-one mapping assumption. Therefore, his algorithm requires prior segmentation of the brain from the MR image before application of his matching criterion to that part of the image. After careful inspection of the dispersion behaviour of the grey-value scatter plot Hill et al propose to select only specific bins of the scatter plot to obtain meaningful dispersion measurements. Their approach can be automated but in a data dependent manner only. Our first attempt in [2] to increase robustness in a data independent way was based on an intuitive information theoretic reasoning only: the scatter-plot entropy as a measure of multi-modality information content or signal complexity has all the properties of a data independent measure of scatter-plot dispersion. The main problem with the entropy matching criterion, however, is its sensitivity to the problem of partial overlap, and, even worse, also to the problem that the information content of the separate overlapping parts of the images is not constant, causing it to have multiple local

minima all over α-space. In this paper we have solved the latter problem also using relative entropy.

The results from sections 3.1 and 3.2 are all subvoxel accurate if the largest voxel size (mostly equals the slice distance) is taken as a reference, and if the stereotactic registration solutions are taken to be correct. The latter assumption was clearly not true for the pair of CT/MR volumes in section 3.2. In this case visual inspection showed that our mutual information based registration solution is better than the stereotactic one, which was particularly noticable in the caudal-cranial direction. This can be explained by the fact that the stereotactic localiser was fixed to the head by means of a mask and not by screws. Taking into consideration the nature of the data in these experiments we may also conclude that the mutual information is inherently robust against noise and local differences (e.g. pathological). Comparison of the results for the different interpolation methods tells us that nearest neighbour interpolation offers a very rapid solution with an acceptable, near subvoxel accuracy, even when it is used with a subsample of the floating volume, f_2. All experiments with nearest neighbour interpolation and sub-sampling were performed in less than 7 minutes. The trilinear partial volume distribution method was especially designed to overcome some edge effects, and effects of image grid interference, that are both inevitable when calculating the mutual information of a finite sample. Therefore, it does not go well with sub-sampling, which explains why no sub-sampling was used in tables 1 and 2. When the same tests were performed using the sampling factors that were used with nearest neighbour and trilinear interpolation the search algorithm got stuck in a local minimum far from the registration solution. Our implementation of the trilinear partial volume distribution has approximately the same complexity as that of trilinear interpolation which was about twice as slow as nearest neighbour interpolation. This factor two together with the differences in the in-plane sampling factor explains why the partial volume distribution based experiments in tables 1 and 2 took 90 minutes. The results from section 3.3 show us that it is possible to register images of different modalities without using any prior knowledge about the grey-value mapping between parts of the image, and yet obtain essentially the same registration results. E.g. both methods proposed by van den Elsen require the application of feature extraction algorithms to make simple correlation work. This approach makes her algorithms data dependent. The results obtained in section 3.4 have not yet been compared to that of other algorithms. However, our algorithm is more user-friendly than Woods' AIR software [16] because it does not require the prior segmentation of the brain surface from the MR images. Our algorithm does not require the selection of a part of the histogram, as does Hill's more robust version of Woods'

algorithm.

In all our experiments so far, using the search procedure described in section 2.2.4 we have observed translations up to 2.5 cm and rotations up to 11° that were all recovered fully automatically. Thus, for our practical purposes the mutual information criterion offers a well behaved optimisation function in a reasonably large region around the initial estimate of the registration parameter values. In theory we need to prove that it is strictly quasi-convex to guarantee its unimodality. Further experiments will be needed in the future to obtain an estimate of the range of quasi-convexity around the global optimum.

5.   Conclusion

The new voxel similarity based (VB) registration algorithm has been applied successfully to real world images of the human head (CT, MR, and PET volumes) which satisfy the rigid body assumption well. Comparison of the registration outcome with that of stereotactic and 2 other VB registration solutions indicates that mutual information is a subvoxel accurate matching criterion and that it is robust in multi-modality situations.

The experimental results indicate that the mutual information matching criterion is a well-behaved multi-modality 3-D registration criterion, and that we have found satisfactory solutions for all implementation problems. Most importantly, we have found that due to the finiteness of the sample of grey-value pairs involved in the calculation of the mutual information matching criterion the use of partial volume distribution of grey-values is required to obtain the most accurate results.

6.   Acknowledgments

References

1. L.G. Brown, "A Survey of Image Registration Techniques", *ACM Computing Surveys*, Vol. 24, No. 4, pp. 325-376, December 1992.
2. A. Collignon, D. Vandermeulen, P. Suetens, G. Marchal, "3D Multi-Modality Medical Image Registration Using Feature Space Clustering", accepted for presentation at the *CVRMed'95 conference* on april 3-5, 1995, and publication in the *Lecture Notes in Computer Science* series, Springer-Verlag.
3. T.M. Cover, J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.
4. AIM Programme, Project A2003, COVIRA Computer Vision in Radiology, Deliverable (2) D1/1.1 Set of raw images and clinical description, 26 June 1992, distributed by the Commission of the European Communities DG XIII, Brussels.
5. R.O. Duda, P.E. Hart, *Pattern Classification and Scene Analysis*, Stanford Research Institute, Menlo Park, CA, USA, A Wiley-Interscience Publication, 1973.
6. P. Dupont, G.A. Orban, R. Vogels, G. Bormans, J. Nuyts, C. Schiepers, M. De Roo, L. Mortelmans, "Different Perceptual Tasks Performed with the Same Visual Stimulus Attribute Activate Different Regions of the Human Brain: A Positron Emission Tomography Study", *Proc. Natl. Acad. Sci., USA*, Vol. 90, pp. 10927-10931, December 1993, Neurobiology.
7. D.L.G. Hill, D.J. Hawkes, N.A. Harrison, C.F. Ruff, "A Strategy for Automated Multimodality Image Registration Incorporating Anatomical Knowledge and Imager Characteristics", *13th Int. Conf., Information Processing in Medical Imaging, IPMI'93*, Flagstaff, Arizona, USA, June 1993, Lecture Notes in Computer Science, Vol. 687,pp. 182-196, Springer-Verlag, Berlin Heidelberg.
8. D.L.G. Hill, C. Studholme, D.J. Hawkes, "Voxel Similarity Measures for Automated Image Registration", *Proc. SPIE Vol. 2359, Visualization in Biomedical Computing*, SPIE Press, Bellingham, WA, Rochester, MN, October 4-7, 1994.
9. C.R. Maurer, J.M. Fitzpatrick, "A Review of Medical Image Registration", *Interactive Image-Guided Neurosurgery*, R.J. Maciunas (Ed), Park Ridge, IL, American Association of Neurological Surgeons, 1993, pp. 17-44.
10. W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in C, The Art of Scientific Computing*, Cambridge University Press, 1988.
11. I. Vajda, *Theory of Statistical Inference and Information*, Kluwer Academic Publisher, Dordrecht, 1989.
12. P.A. Van den Elsen, *Multimodality matching of brain images*, Ph.D. Thesis, Utrecht University Thesis, June 1993.
13. P.A. van den Elsen, E-J.D. Pol, M.A. Viergever, "Medical Image Matching - A Review with Classification", *IEEE Eng. in Medicine and Biol.*, pp. 26-38, March 1993.
14. P.A. Van den Elsen, E.J.D. Pol, T.S. Sumanaweera, P.F. Hemler, S. Napel, J. Adler, "Grey value correlation techniques used for automatic matching of CT and MR brain and spine images", *Proc. SPIE Vol. 2359, Visualization in Biomedical Computing*, pp. 227-237, SPIE Press, Bellingham, WA, Rochester, MN, October 4-7, 1994.
15. R.P. Woods, J.C. Mazziotta, S.R. Cherry, "MRI-PET Registration with Automated Algorithm", *JCAT*, 17(4):536-546, July/August, 1993.
16. R.P. Woods, *Automated Image Registration (AIR): User's Manual and Program Documentation*, Division of Brain Mapping Neuropsychiatric Institute, UCLA School of Medicine, 8/7/1993.