

Incorporating a Measure of Local Scale in Voxel-Based 3-D Image Registration

László G. Nyúl, Jayaram K. Udupa*, *Senior Member, IEEE*, and Punam K. Saha, *Member, IEEE*

Abstract—We present a new class of approaches for rigid-body registration and their evaluation in studying multiple sclerosis (MS) via multiprotocol magnetic resonance imaging (MRI). Three pairs of rigid-body registration algorithms were implemented, using cross-correlation and mutual information (MI), operating on original gray-level images, and utilizing the intermediate images resulting from our new scale-based method. In the scale image, every voxel has the local “scale” value assigned to it, defined as the radius of the largest ball centered at the voxel with homogeneous intensities. Three-dimensional image data of the head were acquired from ten MS patients for each of six MRI protocols. Images in some of the protocols were acquired in registration. The registered pairs were used as ground truth. Accuracy and consistency of the six registration methods were measured within and between protocols for known amounts of misregistrations. Our analysis indicates that there is no “best” method. For medium misregistration, the method using MI, for small and large misregistration the method using normalized cross-correlation performs best. For high-resolution data the correlation method and for low-resolution data the MI method, both using the original gray-level images, are the most consistent. We have previously demonstrated the use of local scale information in fuzzy connectedness segmentation and image filtering. Scale may also have potential for image registration as suggested by this work.

Index Terms—Correlation, image analysis, image processing, image registration, image scale, magnetic resonance imaging (MRI), multiple sclerosis, mutual information.

I. INTRODUCTION

THE purpose of image registration is to represent information pertaining to the same object system of study in a common coordinate system. Registration becomes necessary primarily in the following situations: 1) Scenes are acquired for the same body region from different modalities, for example, computed tomography (CT), Magnetic resonance imaging (MRI), functional MRI (fMRI), and positron emission tomography (PET), for a patient’s head. The need for registration comes from the fact that fMRI and PET provide

mainly functional information and lack anatomic details, and *vice versa* for CT and MRI. Registration helps associating the functional activity information with specific anatomic regions. 2) When using different protocols within the same modality for the same body region, for example, T1-weighted images of the brain before and after administration of Gd-DTPA contrast agent. Here registration helps to identify the regions that enhance and thereby to study certain disease processes. 3) Scenes are acquired for the same body region using the same modality for different time instances. The time instances may be close to each other for studying the motion or displacement of an object in the body region, or they may be far apart for studying longitudinally the growth or change in an object. 4) In certain interventional procedures, information derived from acquired scenes is utilized to provide navigational aids for the devices used in the procedure. In these situations, it becomes necessary to register the device, the body region, and the scene. 5) Scenes acquired for a given body region are to be matched to a computerized atlas or model for the same body region. This is often helpful in studying statistically the variations in certain object measures over a subject population as well as in scene segmentation. No registration is needed between two scenes if both the following conditions are satisfied: a) The same objects of study are observable in both scenes and they do not change shape. b) The relationship among them and to the scene coordinate system does not change. For any objects, even static, in any body region, these conditions are difficult to fulfill, and therefore, registration becomes necessary in the situations described earlier.

Registration methods may be categorized in several ways. For surveys covering and classifying most of the image registration literature, see [1] and [2]. Broadly, registration methods may be divided into two groups: *scene-based* and *object-based*. In scene-based methods [3], the transformation needed to convert a given scene to match spatially with another given scene is estimated by matching the intensity pattern in the two scenes. In object-based methods [4], structure information derived from the two scenes is matched for this estimation. The registration transformation itself may be *rigid* [3] or *elastic* (also known as *deformable*) [5]. In the rigid case, the mismatch is assumed to be due to a global translation and rotation of one of the scenes/structures. Scaling, if required, is usually handled easily by an interpolation of one of the scenes prior to registration (or extraction of structures) to make its voxels equal in size to that of the other. In the elastic case, it is assumed that local deformations are needed in addition to a global translation and rotation, to achieve a match.

Manuscript received May 2, 2002; revised October 23, 2002. This work was supported in part by the National Institutes of Health (NIH) under Grants NS 37172 and Grant AR 46902, and in part by the Hungarian Ministry of Education under Grant FKFP 0908. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was C. Meyer. *Asterisk indicates corresponding author.*

L. G. Nyúl is with the Department of Applied Informatics, University of Szeged, H-6701 Szeged, Hungary.

*J. K. Udupa is with the Medical Image Processing Group, Department of Radiology, University of Pennsylvania, 423 Guardian Drive, 4th Floor Blockley Hall, Philadelphia, PA 19104-6021 USA (e-mail: jay@mipg.upenn.edu).

P. K. Saha is with the Medical Image Processing Group, Department of Radiology, University of Pennsylvania, Philadelphia, PA 19104-6021 USA.

Digital Object Identifier 10.1109/TMI.2002.808358

A. Registration as Applied to Studying Multiple Sclerosis (MS)

Much of the motivation for this investigation originated from our on-going research in MS image analysis. Therefore, we shall describe briefly in this section this basis as related to MS. The most commonly used metric to clinically measure disease progression in MS is the expanded disability status scale introduced by Kurtzke [6]. The clinical quantification of disease severity in this method is subjective and sometimes equivocal. The development of new treatments demands objective outcome measures for relatively short trials. Therefore, MRI has become one of the most important paraclinical tests for diagnosing MS and in monitoring disease progression in MS [7].

MRI has proven to be a very sensitive marker of the MS disease since macroscopic areas of damage or loss of myelin can be imaged with hyper or hypo intensity with respect to the surrounding tissues. A variety of protocols are being investigated for improving the conspicuity of the lesions in the images in the early inflammatory stage and in the advanced stage for both the microscopic and macroscopic processes. T1-weighted imaging without and with Gd-DTPA enhancement (T1E), dual spin-echo imaging providing the T2 and the proton density (PD) values for each image element, and magnetization transfer (MT) imaging are all utilized for characterizing and quantifying MS.

It is clear that the various MRI methods provide somewhat disparate pieces of the information relating to the MS puzzle. The use of quantitative MRI in predicting the onset of MS and in assessing disease severity in established MS is actively pursued at present [8]. In all such efforts, image registration to combine the complementary information from the images acquired under different protocols becomes vital. Many of the image segmentation methods for MS require one or more preprocessing operations, including registration to handle multichannel features [9]–[14]. Registration is also utilized in the evaluation of techniques that try to assign tissue-specific meaning to the numerical intensity values in commonly used protocols such as T1, T1E, T2, and PD [15]. New quantitative measures are also being developed from the images for each of several tissue regions including the whole brain parenchyma (BP), gray matter (GM), white matter (WM), lesions (LS), and cerebrospinal fluid (CSF), for patients with relapsing remitting, secondary progressive, and chronic progressive MS. These methods also depend on image registration. In summary, the study of MS can greatly benefit from image registration. However, efforts in this direction in actual clinical studies are rather sparse and focussed mainly on observing enhancement in postcontrast images upon subtraction from registered precontrast images.

B. Summary of Work

We describe the registration methods compared in this work in a general framework in Section II. A brief description of the new scheme using object scale is given here as well as the details of our implementation. The clinical data and experiments are described in Section III. Results and discussion are presented in Section IV, and concluding remarks in Section V.

II. METHODS

We represent an *image* or *scene* by a pair $\mathbf{C} = (C, f)$ where C (called the *scene domain*) is a three-dimensional (3-D) array of volume elements (voxels) covering a body region of the particular patient for whom image data \mathbf{C} are acquired, and f is a function (called *intensity* function) whose domain is C . All that is presented in this paper generalizes in a straightforward manner to vector-valued scenes, that is, when f is a vector-valued function, although we assume that f is scalar-valued.

A. General Scheme for Registration

In this paper, we confine ourselves to the evaluation of rigid-body scene-based registration methods. Here, we are looking for a coordinate transformation that maps coordinates of scene \mathbf{C}_{i2} into coordinates in scene \mathbf{C}_{i1} such that some similarity measure between \mathbf{C}_{i1} and the transformed \mathbf{C}_{i2} is maximal. Formally, in scene-based registration, the input is a pair of scenes $\mathbf{C}_{i1} = (C_{i1}, f_{i1})$ and $\mathbf{C}_{i2} = (C_{i2}, f_{i2})$ of the same body region and the output is a pair of scenes $\mathbf{C}_{o1} = (C_{o1}, f_{o1})$ and $\mathbf{C}_{o2} = (C_{o2}, f_{o2})$. Assuming that \mathbf{C}_{i2} (the *test scene*) is to be matched with \mathbf{C}_{i1} (the *reference scene*), the output scenes are such that $\mathbf{C}_{o1} = \mathbf{C}_{i1}$ and $\mathbf{C}_{o2} = T(\mathbf{C}_{i2}) = (C_{i1}, f_{o2})$, where T consists of a geometric transformation T_G to register \mathbf{C}_{i2} with \mathbf{C}_{i1} and an interpolation operation T_I to estimate scene intensities at the new locations for each voxel c in \mathbf{C}_{i2} . An appropriate scene-based or object-based interpolation method is utilized for this purpose. A majority of the existing scene-based registration methods are of the rigid type [3], [16]–[19]. They can all be characterized more-or-less by the following process.

- 1) (Optionally) interpolate \mathbf{C}_{i1} and \mathbf{C}_{i2} to make voxels cubic and equal in size in both scenes.
- 2) Choose a *criterion function* G for determining the degree of match/mismatch between (the transformed) \mathbf{C}_{i2} and \mathbf{C}_{i1} and select a *stopping criterion* S .
- 3) Initialize T by setting it to an initial translation and rotation of \mathbf{C}_{i2} to roughly align with \mathbf{C}_{i1} .
- 4) Optimize $G(T)$ as a function of T [and of $S(T)$] by incrementally modifying T and re-evaluating $S(T)$.

The essential differences in the different methods are in how the mismatch criterion function G and the stopping criterion S are selected and how $G(T)$ is optimized. Examples of criterion functions used include correlation coefficients of intensities or of intensity features [17], [20]–[22], variance of intensity ratios [22], [23], histogram dispersion [24], and mutual information (MI) [25], [26]. The search technique in Step 4) may use, for example, Powell’s method [27]–[29], the downhill simplex method [30], stochastic search [25], genetic methods [24], or (semi-)exhaustive search.

B. Scale as a New Feature

In scene-based methods, the match/mismatch criterion function G may be evaluated directly from the input scenes or from some feature scenes extracted from the original scenes. Several features (such as edges, ridges, troughs) have been proposed for scene-based image registration in the past [17]. We propose a

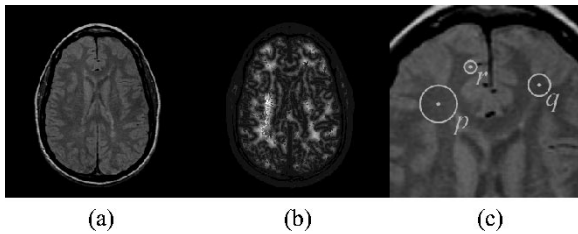


Fig. 1. Illustration of local structure size (scale). An MRI slice of a MS patient's head (a), the corresponding scale image (b), and part of the original slice enlarged (c) to show details. In (b), scale values are not shown for the background part of the slice.

new feature—the (local) object-scale—that has been recently used for aiding image segmentation [31] and filtering [32].

“Scale” is a fundamental, well-established concept in image processing [33]. Our idea of “scale” [31], which is somewhat different from a similar concept used in computer vision [33], is to determine the size of local structures under a prespecified region-homogeneity criterion. For example, in Fig. 1, local size of the structure to which voxel p belongs is bigger than that to which q or r belongs. There are other approaches to determine local scale measures. Pizer *et al.* [34] defined optimal scale at each spatial location as the scale for which “medialness” measure reaches a local maximum along the scale direction in the scale-space representation. In their medialness model, scale represents the distance at which a pair of opposing boundaries are expected. In their interpretation, no continuity of homogeneity is used, whereas, in our approach, the concept is rather simple, object size is determined by looking for discontinuity of homogeneity in images, and the computation is simple [31].

We define *scale* in an image \mathbf{C} at any voxel $c \in C$ as the radius $r(c)$ of the largest ball centered at c which lies entirely within the same object region (defined under a prespecified region-homogeneity criterion). This adaptive scale concept has evolved in connection with fuzzy connectedness principles [35] and represents the local size of the object at the voxel. In [31], a simple and effective algorithm has been described which estimates $r(c)$ at every location $c \in C$ in any image \mathbf{C} without explicit object segmentation but based on continuity of region homogeneity. It has been shown in [31] that this information of scale leads to a significantly improved object segmentation. For completeness, the algorithm for scale computation is reproduced below.

Let $\nu = (\nu_1, \nu_2, \nu_3)$ be the size of the voxels in the given scene $\mathbf{C} = (C, f)$, where ν_1, ν_2 , and ν_3 are the length of the voxel in the three principal directions in some measurement unit (millimeters in our application). A ball $B_{k,\nu}(c)$ of radius $k \geq 0$ and with center at $c = (c_1, c_2, c_3) \in C$ in \mathbf{C} is defined by

$$B_{k,\nu}(c) = \left\{ e \in C \mid \sqrt{\sum_{i=1}^3 \frac{\nu_i^2 (c_i - e_i)^2}{\min_j [\nu_j^2]}} \leq k \right\}. \quad (1)$$

Note that $B_{k,\nu}(c)$ forms a digital ball when the spatial resolution is isotropic and a digital ellipsoid when it is not. The equation is formulated so that, in both cases, $B_{k,\nu}(c)$ represents a (quantized) ball in the scene domain. We will refer to $B_{k,\nu}(c)$ as a ball irrespective of whether or not ν is isotropic.

For a ball $B_{k,\nu}(c)$ of any radius $k > 0$ and centered at c , we define a *fraction of object*, $FO_{k,\nu}(c)$, that indicates the fraction of the ball boundary occupied by a region which is homogeneous with c , by

$$FO_{k,\nu}(c) = \frac{\sum_{e \in B_{k,\nu}(c) - B_{k-1,\nu}(c)} W_\psi(|f(c) - f(e)|)}{|B_{k,\nu}(c) - B_{k-1,\nu}(c)|} \quad (2)$$

where $|B_{k,\nu}(c) - B_{k-1,\nu}(c)|$ denotes the number of voxels in $B_{k,\nu}(c) - B_{k-1,\nu}(c)$ and W_ψ is a homogeneity function. In this paper, we use a zero-mean unnormalized Gaussian function for W_ψ . The algorithm for object scale estimation is as follows.

Algorithm OSE

Input: \mathbf{C} , $c \in C$, W_ψ , a fixed threshold t_s .

Output: $r(c)$.

begin

 set $k = 1$;

while $FO_{k,\nu}(c) \geq t_s$ **do**

 set k to $k + 1$;

endwhile;

 set $r(c)$ to $k - 1$;

 output $r(c)$;

end

The algorithm iteratively increases the ball radius k by one, starting from one, and checks for $FO_{k,\nu}(c)$, the fraction of the object containing c that is contained in the ball. The first time when this fraction falls below t_s , we consider that the ball contains an object region different from that to which c belongs. Following the recommendation in [31], we use $t_s = 0.85$.

To determine the standard deviation (SD) σ_ψ of W_ψ , we first compute the local in-plane intensity gradients over the entire scene domain between all pairs of neighboring voxels c and d where c and d are in the same slice. (We exclude local out-of-plane intensity gradients, since the out-of-plane resolution is usually coarser). Here, local intensity differences are taken in both x and y direction over the entire scene and pooled together. The upper 10% of the gradients are then discarded in order to account for inter object boundaries; the mean M_h and the SD σ_h are then computed over the lower 90% gradients. Finally, σ_ψ is taken to be $M_h + 3\sigma_h$. In this fashion, the only parameter required in the method is determined automatically. A pair of neighboring voxels will be then assigned a value of intensity similarity depending on their intensity difference and the noise in the entire scene that is measured by σ_ψ .

Several hundred studies have been processed automatically and successfully using this method of scale estimation in our ongoing projects for different protocol and body region combinations in several image segmentation applications. The scale image for the MR slice of an MS patient's head, shown in Fig. 1(a), is shown in Fig. 1(b). As can be seen from the image, interior regions have higher scale (brighter intensities) while detailed structured regions as well as boundary vicinities have lower scale (darker intensities). One set of our new methods used scale values directly as registration feature. In these cases, the registration method tries to match potential boundary regions to boundary regions, and homogeneous “flat” regions to

flat regions. Incorporating boundary information in this manner into the registration method results in improved matching accuracy in certain cases. Our approach has the advantage of not requiring edge detection or explicit segmentation (which is a challenging task on its own). It uses information relating to the approximate whereabouts of the boundaries (regions of small scale) in a fuzzy manner.

C. Transformation

In our application (registering MRI head datasets for the purpose of computing MS disease-specific measures), the rigid-body assumption is satisfied well, since, in all cases, the images utilized in this study were acquired in the same session on the same scanner. Therefore, scanner calibration problems, and geometric distortions, as well as patient-related changes (due to disease evolution) can be neglected.

The geometric part T_G of a 3-D rigid-body transformation T can be described by six parameters: three global translation components (t_x, t_y, t_z) and three rotation angles (ϕ_x, ϕ_y, ϕ_z) around the three coordinate axes. In our application, the origin of the coordinate system was placed at the center of the scene domain and the axes were defined by the natural directions of the digital scene. The transformation T_G is given by applying the three rotations and the translation component in the following order:

$$T_G = D \circ R_x \circ R_y \circ R_z \quad (3)$$

where T_G represents the global (composite) transformation, D represents translation, and $R_x, R_y,$ and R_z represent rotation around the $x, y,$ and z axes, respectively, and \circ denotes composition. For T_I , we used trilinear interpolation.

D. Matching Criteria

There are several measures proposed in the literature to define the degree of match/mismatch of two scenes. We utilized three different matching criteria to guide the search for optimum registration: cross-correlation, normalized cross-correlation, and mutual information.

Correlation techniques [3], [16], [36] perform well in mono-modal registration wherein there is a linear relationship between the measurements for the same spatial elements in the two acquisitions. However, they are generally not the best candidates for matching images from different modalities because this special relationship is usually missing. In our application, however, the correlation method was expected to work well since, in all cases, we register MR images and we expect some relationship among the intensity values in the various MRI protocols utilized in our application.

When using object-scale feature for computing the mismatch function, that is, when registering the two scale images, the task can be considered as a mono-modal registration problem, since the scale values represent the same property. However, the different protocols show different contrast between tissues and the scale images may not show exactly the same objects always. For example, in PD images, GM and WM are portrayed well, and CSF is somewhat less prominent, while in T2 images, CSF is the most visible object, and the interface between WM and GM

is usually blurred. As long as there are sufficient similar structures in the images, the matching algorithm performs well, even with some dissimilarities present in the images.

For any reference scene \mathbf{C}_1 and test scene \mathbf{C}_2 , the correlation measure $CR(\mathbf{C}_1, \mathbf{C}_2)$ is defined as

$$CR(\mathbf{C}_1, \mathbf{C}_2) = \sum_{c \in \mathbf{C}_1} f_1(c) \cdot f_2(c). \quad (4)$$

We confine the computation of $CR(\mathbf{C}_1, \mathbf{C}_2)$ to those voxels c which were within the head volume, thus excluding the large area of the more-or-less homogeneous background. The hypothesis for registration here is that the criterion function $G_{CR}(T) = CR(\mathbf{C}_1, T(\mathbf{C}_2))$ is maximized when the images \mathbf{C}_1 and $T(\mathbf{C}_2)$ are in spatial registration.

We define normalized correlation as

$$CR_n(\mathbf{C}_1, \mathbf{C}_2) = \frac{\sum_{c \in \mathbf{C}_1} f_1(c) \cdot f_2(c)}{\sqrt{\sum_{c \in \mathbf{C}_1} f_1^2(c) \cdot \sum_{c \in \mathbf{C}_1} f_2^2(c)}}. \quad (5)$$

The criterion function for the normalized case is $G_{CR_n}(T) = CR_n(\mathbf{C}_1, T(\mathbf{C}_2))$.

Given that both images to be registered show information about the same underlying anatomy, there is MI between the two images. MI is a basic concept from information theory, measuring the statistical dependence between two random variables or the amount of information that one variable contains about the other.

We treat the intensity values $a = f_1(c)$ and $b = f_2(c)$ in the two scenes that are to be registered as random variables A and B , respectively, over the scene domain. Estimations for the joint and marginal distributions of A and B can be obtained from the joint and marginal histograms of the two scenes. MI is defined in terms of entropy in the following way:

$$MI(A, B) = h(A) + h(B) - h(A, B) \quad (6)$$

where $h(A)$ is the entropy of the random variable A , and is defined as $h(A) = -\sum_a p(a) \ln p(a)$, $h(B)$ is defined analogously, and the joint entropy of two random variables A and B is $h(A, B) = -\sum_{a,b} p(a, b) \ln p(a, b)$. The criterion function in this case is simply defined as $G_{MI}(T) = MI(A, B')$, where A and B' are random variables representing intensities in \mathbf{C}_1 and $T(\mathbf{C}_2)$, respectively. Probability $p(a)$ for a given intensity value a in \mathbf{C} in the above entropy formulation is estimated as the number of voxels in \mathbf{C} having value a divided by the size of \mathbf{C} , that is

$$p(a) = \frac{1}{|\mathbf{C}|} \cdot |\{c \in \mathbf{C} | f(c) = a\}|. \quad (7)$$

MI as a match/mismatch measure for image registration has been investigated by several groups [19], [25], [26]. The published results show that MI gives one of the best registration results in multimodal image matching. This technique does not require any *a priori* model of the relationship between the modalities. It assumes only that one volume provides the most information about the other (i.e., $G_{MI}(T)$ is maximized) when they are correctly registered.

Considering $G_{CR}(T)$, $G_{CR_n}(T)$, and $G_{MI}(T)$, and letting C_1 and C_2 also be scale images of the given reference and test scenes [for $G_{CR}(T)$ and $G_{MI}(T)$], we consider the following six strategies for registration:

- 1) correlation of original scenes (c,o);
- 2) correlation of scale scenes (c,s);
- 3) normalized correlation of original scenes (cn,o);
- 4) normalized correlation of scale scenes (cn,s);
- 5) MI in original scenes (m,o);
- 6) MI in scale scenes (m,s).

Our way of incorporating scale information into the registration measure is one of several possible. A different scale-space approach is used by another group [16], [17], [37] to combine geometrical information with the correlation measure for the purpose of improving registration. They use gradient operator derived from Gaussian filter of a certain scale parameter to determine the ridgeness of a voxel that is subsequently used in their registration measure. In their approach, the scale parameter is selected from a set of values after inspecting their effect on the ridgeness measure. Once selecting the “optimal” value, the same scaling parameter is used for the entire scene (in fact, for all scenes of the same kind). It is shown in [29] that incorporating gradient information as a weighting factor into the MI measure improves registration in terms of avoiding local optima while keeping the accuracy achieved by using the original MI measures. Here, the gradient is computed by using a Gaussian kernel of a fixed scale, too. In our approach, scale values are tied to the voxels and their surroundings, and therefore, are spatially variant [31].

E. Initialization

In almost all practical cases, the functions we try to optimize have several local optima. Assuming that reaching the global optimum is desirable, good initialization close to the global optimum is, therefore, very crucial for an iterative optimization process to avoid local optima as much as possible. In most cases, the two scenes to be registered cover almost the same body region and have similar orientation by acquisition. However, for different acquisitions, the field of view (FOV) may be different and repositioning errors always occur. Usually, one of the images can be automatically corrected for scale differences by interpolation. Identifying the same position (volume of interest) and orientation, however, usually requires manual help from the operator.

We initialized the transformation T by aligning the centroids and the principal axes of the two scenes. Sometimes this principal axes method itself provides the final registration [38], but often it is only the initialization step to a more sophisticated registration method. The background (outside the head) was segmented by thresholding, and the foreground was used to compute the centroid and the principal axes. Since we knew that the orientations of the original scenes are similar, the eigenvectors were selected such that the least amount of rotation is involved in the initial match. There may be other problems with this initialization. For example, when the eigenvalues are very similar, it is possible that the wrong axes are matched and large rotations are introduced. We prevented this by a constraint on the rotation

angles, and if the program detected such a problem, it swapped the principal axes.

F. Search

One of our previous implementations for multimodal image registration used exhaustive search to minimize the cross-correlation function [36]. Translations were varied in one voxel steps and rotations in a quarter-degree steps. Since this approach involves evaluation of the correlation function for a large number of positions, it was computationally expensive. To reduce computation but still have the necessary capture range and precision, we used a multilevel approach. Multilevel methods are widely used in the literature for registration [36], [39], [40]. They have proved effective in reducing computation time as well as in helping to avoid local minima. In this paper, we used a three-level Gaussian pyramid, each level being twice as coarse in resolution as the previous, level 0 being the original resolution. However, since our data are highly anisotropic (in z direction), we applied filtered subsampling within slices only. Thus, at the coarsest level (level 2), the data resolution was approximately isotropic.

At each hierarchy level, an optimal solution was determined which was used as the starting point for the next level. At the lowest resolution, search was initialized by the principal axes method. Powell’s multidimensional direction set method [41] was used to optimize the criterion function by using Brent’s one-dimensional optimization algorithm for line minimizations. At each hierarchy level, the direction matrix is initialized with translation steps corresponding to the in-plane size of one voxel at that resolution, and the rotation step size is chosen such that it corresponds to one translation step at 60 mm from the center of rotation. Following the suggestions of [26], we used the parameter ordering $(t_x, t_y, \phi_z, \phi_x, \phi_y, t_z)$, so that images get aligned within the plane first, which may facilitate the optimization of the out-of-plane parameters. The fractional tolerance for Brent’s line minimization method was set to $2 \cdot 10^{-4}$, that for Powell’s method to 10^{-3} , and the maximum number of iterations in one step was 25 and 10 for Brent’s and Powell’s method. The translation components were expressed in terms of ν_1 ($= \nu_2 = 0.86$ mm in all our data) and the rotations in degrees. Double precision floating point arithmetic was used throughout the optimization process.

III. EVALUATION

A. Data

MRI head datasets of 10 MS patients (two male and eight female, age between 36 and 50 years, mean 43, SD 4.5) were used in the experiments described below. The patients and datasets were randomly selected from our database which contains about 100 patients with about 4000 3-D scenes comprising fast-spin-echo dual-echo (FSE) T2 and PD studies, T1 and T1E studies, and MT studies. For each patient, six datasets comprising FSE T2 and PD, T1 and T1E, and MT without and with saturation pulse (referred to as MT1 and MT2 studies below) were selected. The datasets of each patient were acquired in a single session. FSE PD and T2 scenes are

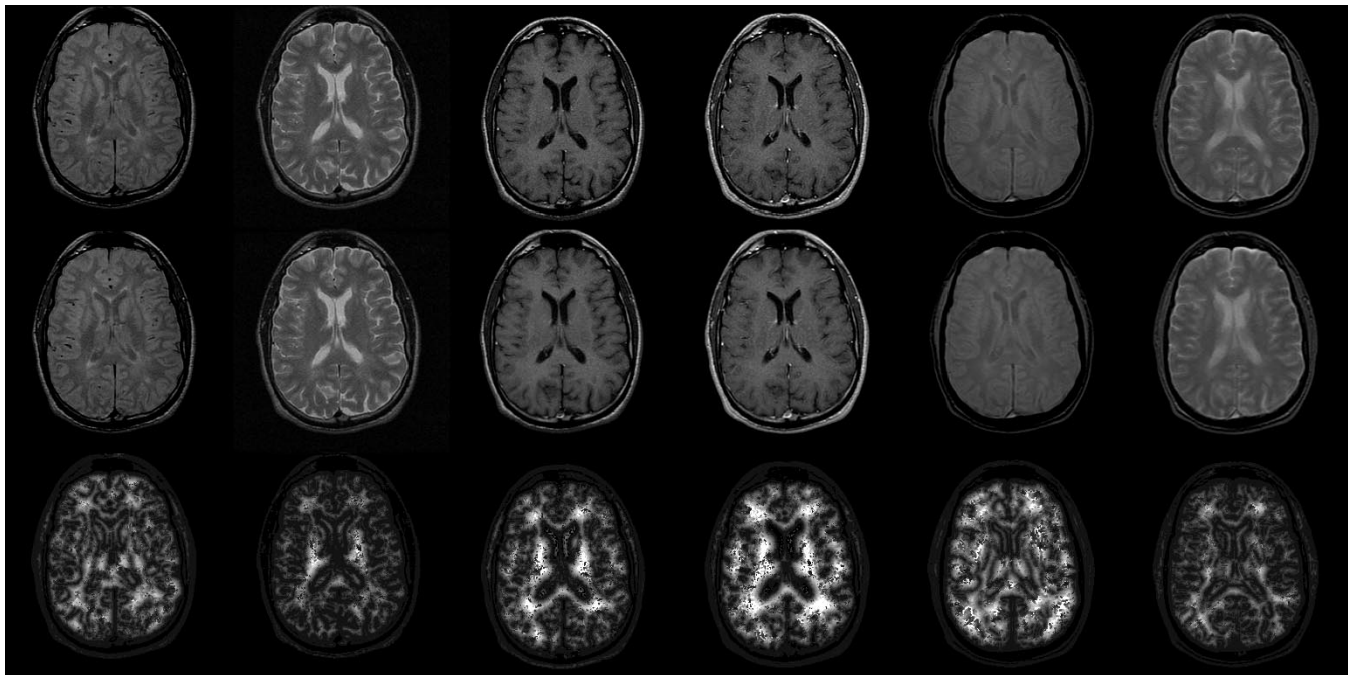


Fig. 2. Original corresponding slices of each of the six datasets (FSE PD, FSE T2, T1, T1E, MT1, MT2) of one patient before registration (first row) and after registration (second row). Slices of the scale images corresponding to those in the first row are shown in the third row. (Scale values are not shown for the background part of the slice). Note the similarity of the scale images across all protocols.

in registration by acquisition and so are the two MT scenes. The selected datasets were visually inspected and found to be of good quality having no severe artifacts. Although one dataset contained visible movement artifact, we kept it for the experiment. Unfortunately, artifacts occur in real situations and we wanted to evaluate the registration methods in nonidealistic real situations.

All datasets were acquired on the same GE Signa 1.5T scanner and consisted of contiguous axial slices covering the entire brain with slice dimensions 256×256 , pixel size of 0.86 mm^2 , $\text{NEX} = 1$, and $\text{FOV} = 22 \text{ cm}$. The MT images were acquired at 5-mm slice thickness while the other four were acquired at 3 mm. For all patients, for the MT image acquisition $\text{TR} = 106 \text{ ms}$ and $\text{TE} = 5 \text{ ms}$, for the T1-weighted studies (with and without enhancement) $\text{TR} = 600 \text{ ms}$ and $\text{TE} = 27 \text{ ms}$. For the FSE acquisitions, for all patients $\text{TR} = 2500 \text{ ms}$, with $\text{TE} = 16\text{--}18$ and $80\text{--}96 \text{ ms}$. For each patient, the four 3-mm datasets consisted of the same number of slices (45–58), all 5-mm datasets included 28 slices. Slices of images of the six protocols from one patient are displayed in Fig. 2.

B. Experiments

The scenes were misregistered by known amounts of translation and rotation (3–3 each). The pairs that were in registration at acquisition were used as ground truth, and the accuracy and consistency of the six registration methods were measured within and between protocols for small, medium, and large misregistrations, thus resulting in a data set of 540 3-D scenes ($9 \times 6 \times 10$). (The extents of these misregistrations are described later.)

To assess *intramodality accuracy*, we used the following procedure: Take a scene \mathbf{C} from each protocol in the set

$\{PD, T2, T1, T1E, MT1, MT2\}$, and reslice \mathbf{C} to create an artificially misregistered version, $\mathbf{C}_k = T_k(\mathbf{C})$, by applying a known transformation T_k . For each method $m \in \{(c, o), (cn, o), (m, o), (c, s), (cn, s), (m, s)\}$, find transformation T_{k_m} which registers \mathbf{C}_k with \mathbf{C} , and compare T_k and T_{k_m} .

Using the pairs (PD, T2) and (MT1, MT2) which were in registration at acquisition, we can assess *intermodality accuracy* in a similar way. The only difference is that, the resliced scene is registered to the other scene from the pair (e.g., rotate and reslice the T2 scene and then register to the PD scene).

Intermodality consistency can also be assessed by utilizing the registered pairs. Register a scene \mathbf{C} of protocol \mathcal{P} to both scenes from a pair (e.g., register a T1 scene to a T2 and to the corresponding PD scene) and compare the two resulting transformations. This can be repeated for $\mathcal{P} \in \{T1, T1E, MT1, MT2\}$ by using the PD-T2 pair as reference, and for $\mathcal{P} \in \{PD, T2, T1, T1E\}$ by using the MT1–MT2 pair as reference. From the three groups together, we have a total of 54 experiments, each involving ten data sets.

C. Figures of Merit

In each experiment, we compare two transformations. In the *accuracy* tests, these are the known misregistration and the transformation found by the method in case, and they are expected to be the inverse of each other. In the *consistency* tests, the two registrations to scenes of the registered pair are compared and they are expected to be equal. We measure the error between two transformations by the mean of the root-mean-square error (RMSE) for the eight corner voxels of the box approximately bounding the head, i.e., the volume of interest in our application.

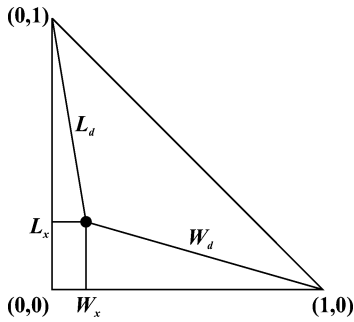


Fig. 3. Mapping the number of wins, losses, and nonsignificant differences into a single “goodness” value for comparison. See text for details.

We combined three different rotations (0, medium, and large angle) and three translations (0, medium, and large displacement) to introduce nine different misregistrations. All nonzero translations and rotations were such that all three directions/axes were involved: rotations were about 2.5° and 10° , respectively, and translations were about five pixels (4.3 mm) and 25 pixels (21.5 mm), respectively. To reduce the number of comparisons and possibly improve statistics, we divided these nine transformations into three groups of misregistrations. After examining the results for each transform, we decided to combine those involving the same amount of rotation, since the methods performed with similar errors within these groups.

We devised a ranking system for the methods as follows. Each registration task is performed on datasets of ten patients. For each method and misregistration (and patient data), for each comparison (i.e., accuracy or consistency test, see above), we computed the mean of the RMSE for the eight corner voxels of the box approximately bounding the head. We can use the paired t -test to compare the errors of two methods performing the same task (same misregistration, same protocol, ten different datasets). Based on the mean RMSE, one of the two methods is either significantly better than the other (i.e., smaller error), or there is no significant difference between the two in performing that particular task. For each misregistration group, for each reasonable pair of methods, for each task (protocol, accuracy/consistency test), we count the number of occurrences of significant wins, losses, and nonsignificant differences of the methods in question. Based on a combination of these counts, we rank the methods within a particular task. Let us consider a particular triplet of wins, losses, and nonsignificant differences (w, l, n) . The total number of comparisons for this particular case is the sum $X = w + l + n$. We normalize the number of wins and losses to get values in the range $[0, 1]$: $W_x = w/X$, $L_x = l/X$. A particular configuration of wins and losses can be identified by a point with coordinates W_x and L_x in the normalized win-loss plane (Fig. 3). The closer this point is to the $(1, 0)$ point (all wins) and the farther it is from the $(0, 1)$ point (all losses), the better the configuration is. We use the following formula to compute the “goodness” value (γ) of a particular win-loss configuration:

$$\gamma = \frac{L_d}{W_d} = \frac{\sqrt{(1 - L_x)^2 + W_x^2}}{\sqrt{(1 - W_x)^2 + L_x^2}}. \quad (8)$$

IV. RESULTS AND DISCUSSION

First, criterion functions are plotted in intramodality and intermodality registration situations to allow visual comparison of their robustness against translation and rotation errors. Then methods are compared numerically, illustrated by tables, using the composite “goodness” measure described in Section III.

A. Visual Comparison of Registration Functions

If criterion functions are plotted as a function of misregistration, visual inspection of these plots may point out differences between different functions and show their possible strengths or weaknesses if used in an optimization process. Plots of these kind for the six registration methods described in this paper are shown in Fig. 4 for intramodality and in Fig. 5 for intermodality registration. Criterion functions are shown as a function of translation in x direction (in one voxel step) and rotation around the z axis (in 1° steps). The functions were evaluated for the whole 3-D datasets, however, transformations did not involve out-of-plane components. For the intramodality example, a PD volume of a patient was used and for the intermodality example the same PD and the corresponding T2 volume was used. It is assumed that no transformation is required for perfect registration, since we register identical scenes and scenes that are perfectly registered at acquisition. Similar plots with similar trends were obtained—although not shown here—for other protocols (T2, T1, T1E, MT1 and MT2 intramodality, and MT1 and MT2 intermodality registration) and other datasets.

Figs. 4 and 5 are similar to plots in [29] and [42], thus allowing easy comparison of the behavior of our proposed registration measures with those proposed by other groups in the field. Horizontal axes correspond to the applied misregistration (in voxels or degrees, respectively). Vertical axes correspond to the criterion function values. Actual values along the vertical axes are not shown, however, since they differ in orders of magnitude among different functions and protocols and are not directly comparable, and the actual ranges are not relevant to a decent optimizer anyway.

Interpolation artifacts are studied in [29], [42], in the context of image registration using MI. The “typical” pattern due to interpolation [42] does not appear in our translation plots, since translation was applied in voxel steps and no interpolation actually took place in this case. However, linear interpolation was used in the rotation examples.

The extremely sharp peaks for $(m, 0)$ in Fig. 4 are due to the lack of interpolation errors and due to the fact that identical images are being matched. This MI peak spreads a little in the case of scale images (m, s) . Scale images in our implementation have only 12 distinct values and assume a strong smoothing effect. The correlation measures have similar shapes and have sharper peaks when applied to the scale images. In the rotation experiment (second row), normalized correlation on scale images (cn, s) seems to eliminate noise (perhaps due to interpolation) along the curve that shows on the standard correlation function on scale images (c, s) . It would be worth studying as to what effect the gray-scale resolution utilized for scale has on the performance of the three scale-based methods. It may improve their performance but will clearly be computationally more expensive.

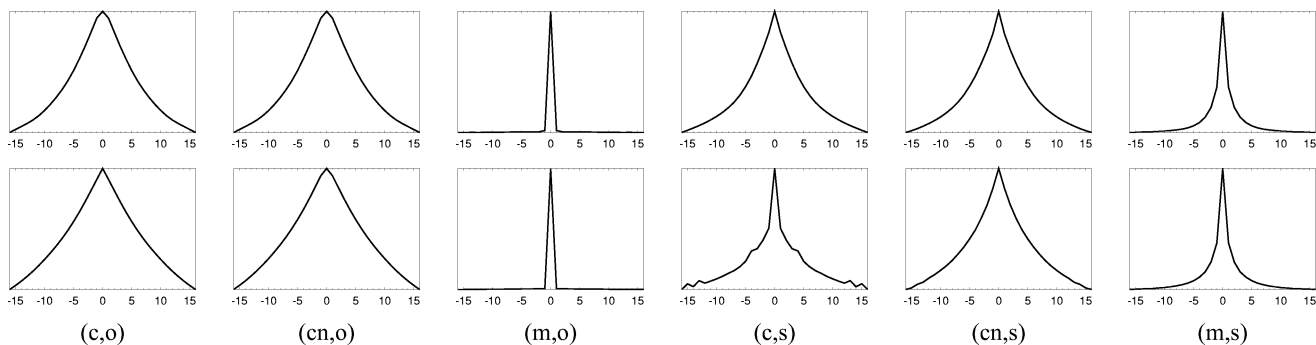


Fig. 4. Criterion functions for the six methods for PD intramodality matching, as a function of translation in x direction (first row) and as a function of rotation around the z axis (second row). See text for explanation.

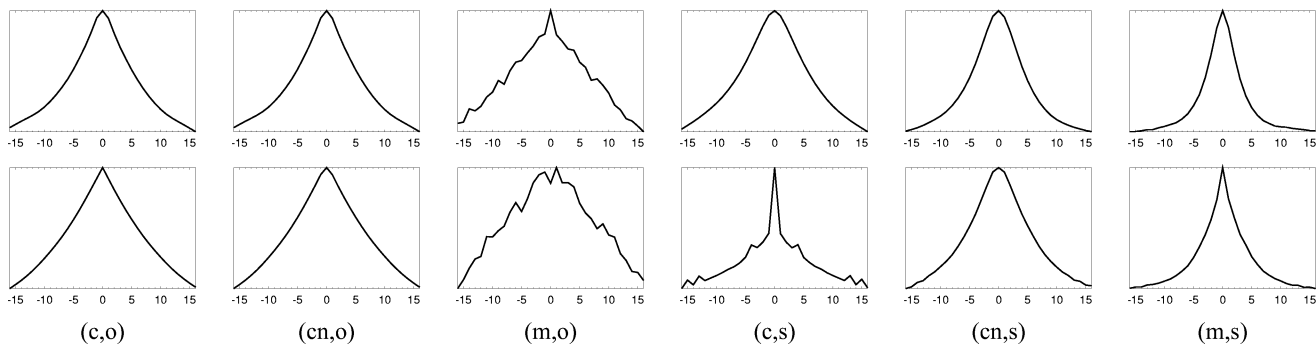


Fig. 5. Criterion functions for the six methods for PD-T2 intermodality matching, as a function of translation in x direction (first row) and as a function of rotation around the z axis (second row). See text for explanation.

Intermodality registration plots (see Fig. 5) differ mainly from those in Fig. 4 in that peaks are more spread out. Here, the two images contain different intensity patterns and perhaps different noise structure that make both correlation and entropy-based functions smoother. MI in the interprotocol example seems to improve if applied to scale images (m,s) instead of the original values (m,o) where local minimum is at the optimal location (as also noted in [29]).

In summary, all criterion functions studied herein seem to behave well in case of translation and rotation mismatch in both intraprotocol and interprotocol matching. Their numerical comparison follows.

B. Numerical Comparison of the Performance of Methods

In the majority of the 5400 registrations performed, the RMSE of the resulting transformation was within the subvoxel range. In a few cases, the RMSE was larger than the voxel size and there were a few situations in which the method did not find a reasonable registration at all. When applying any method to a large number of real images there is a chance that the method will fail. The success of a registration method cannot be guaranteed, and when it fails, special care is required. Most of these failure cases corresponded to registering scale images with large misregistration (rotation). Since no ground-truth registrations were available for our datasets, the results obtained are potentially valid only for the complex combination of input data, optimizer and criterion functions used.

For the 3-mm datasets, the RMSE was around 0.5 mm for misregistrations with rotation and around 1 mm for the zero-

rotation cases, most of the error coming from translation in the (low-resolution) z direction. The RMSE was slightly higher (around 1 and 3 mm, respectively) for the 5-mm datasets.

Our analysis is based on all possible sensible comparisons (135 in all). The “goodness” values γ are shown in Table I (intramodality), Table II (intermodality), and Table III (consistency). Values for the 3-mm datasets are shown in the top half of each table while those for the 5-mm datasets are in the bottom. Rows correspond to the methods, columns correspond to the misregistration groups (small, medium, and large), and the combined values over all groups (total).

Our overall observation is that there is no “best” method that stands out in all cases. In total, there is significant difference in performance for intraprotocol registration while for interprotocol registration the best methods have more similar performance. The most competitive among the methods seem to be (cn,o) and (m,o). We note that there are situations wherein (cn,o) is better than (m,o) and *vice versa*, and (m,s) is better than (m,o) and *vice versa*.

Method (m,o) using MI and the original intensities performs best on intramodality registration for cases with medium misregistration, and for high-resolution data with small, and low-resolution data with large misregistration. However, method (cn,o) using normalized cross-correlation and the original intensities is better for cases with large misregistration and for low-resolution data with small misregistration. In total, method (m,o) performs better for the 3-mm data and method (cn,o) for the 5-mm data. Methods using scale images did not get a high score; however, in the large misregistration groups, method (m,s) ranks third. Similarly, in intermodality

TABLE I
COMPARISON OF METHODS FOR INTRAPROTOCOL REGISTRATION.
THE GOODNESS VALUES γ ARE LISTED

thickness	method	small	medium	large	total
3 mm data	(c,o)	2.188	0.192	0.046	0.344
	(cn,o)	1.991	0.117	21.625	1.550
	(m,o)	3.222	421.000	6.250	10.346
	(c,s)	0.353	1.230	0.457	0.580
	(cn,s)	0.192	1.230	0.675	0.568
	(m,s)	0.707	1.000	0.812	0.831
5 mm data	(c,o)	1.000	0.310	0.053	0.322
	(cn,o)	4.281	0.831	inf*	5.200
	(m,o)	1.625	2.500	2.643	2.188
	(c,s)	0.233	0.737	0.068	0.279
	(cn,s)	0.737	1.000	0.500	0.744
	(m,s)	0.850	1.991	2.000	1.457

* The method significantly won in all cases, therefore, the denominator in the formula for γ is 0, thus, γ is infinity.

TABLE II
COMPARISON OF METHODS FOR INTERPROTOCOL REGISTRATION.
THE GOODNESS VALUES γ ARE LISTED

thickness	method	small	medium	large	total
3 mm data	(c,o)	1.355	1.176	1.000	1.168
	(cn,o)	0.400	0.615	4.281	0.940
	(m,o)	1.153	0.737	0.867	0.907
	(c,s)	1.413	1.153	0.737	1.053
	(cn,s)	1.000	1.153	0.615	0.907
	(m,s)	1.000	1.355	0.867	1.050
5 mm data	(c,o)	1.355	1.355	0.831	1.176
	(cn,o)	1.000	1.991	1.737	1.510
	(m,o)	1.625	1.991	2.500	1.991
	(c,s)	0.575	1.355	0.737	0.850
	(cn,s)	0.502	0.076	0.502	0.310
	(m,s)	1.355	0.707	0.867	0.949

TABLE III
COMPARISON OF METHODS FOR CONSISTENCY.
THE GOODNESS VALUES γ ARE LISTED

thickness	method	small	medium	large	total
3 mm data	(c,o)	2.828	1.153	1.286	1.555
	(cn,o)	0.502	0.928	1.991	0.972
	(m,o)	1.071	0.933	0.737	0.909
	(c,s)	1.679	1.071	0.737	1.080
	(cn,s)	0.557	0.801	0.801	0.717
	(m,s)	0.801	1.153	1.000	0.976
5 mm data	(c,o)	1.679	2.828	1.625	1.950
	(cn,o)	1.176	2.479	1.383	1.548
	(m,o)	2.828	2.828	1.355	2.142
	(c,s)	0.737	0.502	0.801	0.675
	(cn,s)	0.270	0.169	0.502	0.296
	(m,s)	0.850	0.535	0.801	0.722

registration, method (m,o) and (cn,o) beat the others in most groups, however, the scores are not that polarized. The MI method on scale images similarly ranks among the top three in four groups.

As far as consistency is concerned, the correlation method using original intensities ranks first in all groups for the 3-mm data and the MI method is more consistent for the low-resolution data. Methods using scale images seem to be considerably less

consistent than the others on low-resolution data, according to our measure.

Our current implementation requires that both reference and test scenes have identical level of discretization and size. It was not affecting the results, since the original datasets were “compatible” and there was no need for extra interpolation. Another implementation limitation is that the multiresolution pyramid is computed only in $x-y$ directions, such that, at the lowest resolution level, the reformatted data are close to isotropy. We chose this implementation to compensate for the highly anisotropic data.

V. CONCLUDING REMARKS

Three pairs of rigid-body registration algorithms were implemented, using cross-correlation, normalized cross-correlation, and MI, operating on original gray-level images and on the intermediate images (called scale images) resulting from our new scale-based method. The scale images, are derived directly from the original images.

Three-dimensional data of the head were acquired from 10 MS patients using six MRI protocols (T1, T2, PD, T1 with Gd, MT on, MT off) and both intraprotocol and interprotocol registration results were analyzed among the methods. Our analysis indicates that there is no “best” method, although there are significant differences in the performance of the methods, for the particular application of combining information from multiprotocol MR imagery for studying MS.

Our results showed that object scale information (like a set of intrinsic markers) can be used in image registration to provide good results. We have previously demonstrated the use of local scale information in fuzzy connectedness segmentation and image filtering [31], [32]. Scale may also have potential for image registration as suggested by this paper.

ACKNOWLEDGMENT

The authors would like to thank Dr R. I. Grossman for the MRI data sets utilized in this research.

REFERENCES

- [1] P. A. van den Elsen, E.-J. D. Pol, and M. A. Viergever, “Medical image matching—A review with classification,” *IEEE Eng. Med. Biol. Mag.*, vol. 12, pp. 26–39, Jan./Feb. 1993.
- [2] J. B. A. Maintz and M. A. Viergever, “A survey of medical image registration,” *Med. Image Anal.*, vol. 2, no. 1, pp. 1–36, 1998.
- [3] R. P. Woods, S. T. Grafton, C. J. Holmes, S. R. Cherry, and J. C. Mazziotta, “Automated image registration: I. General methods and intrasubject, intramodality validation,” *J. Comput. Assist. Tomogr.*, vol. 22, no. 1, pp. 139–152, 1998.
- [4] C. A. Pelizzari, G. T. Chen, D. R. Spelbring, R. R. Weichselbaum, and C. T. Chen, “Accurate three-dimensional registration of CT, PET, and/or MR images of the brain,” *J. Comput. Assist. Tomogr.*, vol. 13, no. 1, pp. 20–26, 1989.
- [5] R. Bajcsy and S. Kovacic, “Multiresolution elastic matching,” *Comput. Vis. Graph.*, vol. 46, no. 1, pp. 1–21, 1989.
- [6] J. F. Kurtzke, “Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS),” *Neurology*, vol. 33, no. 11, pp. 1444–1452, 1983.
- [7] S. A. Lukes, L. E. Crooks, M. J. Aminoff, L. Kaufman, H. S. Panitch, C. Mills, and D. Norman, “Nuclear magnetic resonance imaging in multiple sclerosis,” *Ann. Neurol.*, vol. 13, no. 6, pp. 592–601, 1983.
- [8] J. A. Frank, Ed., *Advances in Multiple Sclerosis*. Philadelphia, PA: Saunders, 2000, vol. 10(4), Neuroimaging Clinics of North America.

[9] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Suetens, "Automated segmentation of multiple sclerosis lesions by model outlier detection," *IEEE Trans. Med. Imag.*, vol. 20, pp. 677–688, Aug. 2001.

[10] F. B. Mohamed, S. Vinitzki, S. H. Faro, C. F. Gonzalez, J. Mack, and T. Iwanaga, "Optimization of tissue segmentation of brain MR images based on multispectral 3D feature maps," *Magn. Reson. Imag.*, vol. 17, no. 3, pp. 304–409, 1999.

[11] S. K. Warfield, M. Kaus, F. A. Jolesz, and R. Kikinis, "Adaptive, template moderated, spatially varying statistical classification," *Med. Image Anal.*, vol. 4, no. 1, pp. 43–55, 2000.

[12] G. Gerig, D. Welti, C. R. Guttman, A. C. Colchester, and G. Szekely, "Exploring the discrimination power of the time domain for segmentation and characterization of active lesions in serial MR data," *Med. Image Anal.*, vol. 4, no. 1, pp. 31–42, 2000.

[13] R. Kikinis, C. R. Guttman, D. Metcalf, W. M. Wells, G. J. Ettinger, H. L. Weiner, and F. A. Jolesz, "Quantitative follow-up of patients with multiple sclerosis using MRI: Technical aspects," *J. Magn. Reson. Imag.*, vol. 9, no. 4, pp. 519–530, 1999.

[14] S. Vinitzki, C. F. Gonzalez, R. Knobler, D. Andrews, T. Iwanaga, and M. Curtis, "Fast tissue segmentation based on a 4D feature map in characterization of intracranial lesions," *J. Magn. Reson. Imag.*, vol. 9, no. 6, pp. 768–776, 1999.

[15] Y. Ge, J. K. Udupa, L. G. Nyúl, L. Wei, and R. I. Grossman, "Numerical tissue characterization in MS via standardization of the MR image intensity scale," *J. Magn. Reson. Imag.*, vol. 12, no. 5, pp. 715–721, 2000.

[16] P. A. van den Elsen, J. B. A. Maintz, E.-J. D. Pol, and M. A. Viergever, "Automatic registration of CT and MR brain images using correlation of geometrical features," *IEEE Trans. Med. Imag.*, vol. 14, pp. 384–396, June 1995.

[17] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Comparison of edge-based and ridge-based registration of CT and MR brain images," *Med. Image Anal.*, vol. 1, no. 2, pp. 151–161, 1996.

[18] G. P. Penney, J. Weese, J. A. Little, P. Demedt, D. L. G. Hill, and D. J. Hawkes, "A comparison of similarity measures for use in 2-D-3-D medical image registration," *IEEE Trans. Med. Imag.*, vol. 17, pp. 586–595, Aug. 1998.

[19] J. West, J. M. Fitzpatrick, M. Y. Wang, B. M. Dawant, C. R. Maurer Jr, R. M. Kessler, R. J. Maciunas, C. Barillot, D. Lemoine, A. Collignon, F. Maes, P. Suetens, D. Vandermeulen, P. A. van den Elsen, S. Napel, T. S. Sumanaweera, B. Harkness, P. F. Hemler, D. L. G. Hill, D. J. Hawks, C. Studholme, J. B. A. Maintz, M. A. Viergever, G. Malandain, R. X. Pennec, M. E. Noz, G. Q. Maguire Jr, M. Pollack, C. A. Pelizzari, R. A. Robb, D. Hanson, and R. P. Woods, "Comparison and evaluation of retrospective intermodality registration techniques," *J. Comput. Assist. Tomogr.*, vol. 21, no. 4, pp. 554–566, 1997.

[20] D. L. Collins, T. M. Peters, and A. C. Evans, "Automated 3D nonlinear image deformation procedure for determination of gross morphometric variability in human brain," *Proc. SPIE*, vol. 2359, pp. 180–190, 1994.

[21] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3-D registration of MR and CT images of the head," *Med. Image Anal.*, vol. 1, no. 2, pp. 163–175, 1996.

[22] —, "Multiresolution voxel similarity measures for MR-PET registration," in *Proc. 14th Int. Conf. Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds., 1995, pp. 287–298.

[23] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, "MRI-PET registration with automated algorithm," *J. Comput. Assist. Tomogr.*, vol. 17, no. 4, pp. 536–546, 1993.

[24] D. L. Hill, C. Studholme, and D. J. Hawkes, "Voxel similarity measures for automated image registration," *Proc. SPIE*, vol. 2359, pp. 205–216, 1994.

[25] W. M. Wells III, P. Viola, H. Atsumi, S. Nakajuma, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Med. Image Anal.*, vol. 1, no. 1, pp. 35–51, 1996.

[26] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, pp. 187–198, Apr. 1997.

[27] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modal image registration based on information theory," in *Proc. 14th Int. Conf. Information Processing in Medical Imaging*, Y. Bizais, C. Barillot, and R. Di Paola, Eds., 1995, pp. 263–274.

[28] L. Lemieux, R. Jagoe, D. R. Fish, N. D. Kitchen, and D. G. Thomas, "A patient-to-computed-tomography image registration method based on digitally reconstructed radiographs," *Med. Phys.*, vol. 21, no. 11, pp. 1749–1760, 1994.

[29] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.*, vol. 19, pp. 809–814, Aug. 2000.

[30] C. R. Meyer, G. S. Leichtman, J. A. Brunberg, R. L. Wahl, and L. E. Quint, "Simultaneous usage of homologous points, lines and planes for optimal, 3-D, linear registration of multimodality imaging data," *IEEE Trans. Med. Imag.*, vol. 14, pp. 1–11, Mar. 1995.

[31] P. K. Saha, J. K. Udupa, and D. Odhner, "Scale-based fuzzy connected image segmentation: Theory, algorithms, and validation," *Comput. Vis. Image Understanding*, vol. 77, no. 2, pp. 145–174, 1999.

[32] P. K. Saha and J. K. Udupa, "Scale-based diffusive image filtering preserving boundary sharpness and fine structures," *IEEE Trans. Med. Imag.*, vol. 20, pp. 1140–1155, Nov. 2001.

[33] T. Lindeberg, *Scale-Space Theory in Computer Vision*. Boston, MA: Kluwer Academic Publishers, 1993.

[34] S. M. Pizer, D. Eberly, D. S. Fritsch, and B. S. Morse, "Zoom-invariant vision of figural shape: The mathematics of cores," *Comput. Vis. Image Understanding*, vol. 69, no. 1, pp. 55–71, 1998.

[35] J. K. Udupa, L. Wei, S. Samarasekera, Y. Miki, M. A. van Buchem, and R. I. Grossman, "Multiple sclerosis lesion quantification using fuzzy-connectedness principles," *IEEE Trans. Med. Imag.*, vol. 16, pp. 598–609, Oct. 1997.

[36] K. Palágyi and J. K. Udupa, "Medical image registration based on fuzzy objects," in *Proc. Summer Workshop Computational Modeling, Imaging and Visualization in Biosciences, COMBIO'96*, vol. KFKI-1996-06/M-H, Budapest, Hungary, 1996, pp. 44–48.

[37] J. B. A. Maintz, P. A. van den Elsen, and M. A. Viergever, "Evaluation of ridge seeking operators for multimodality medical image matching," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 353–365, Apr. 1996.

[38] K. D. Toennies, J. K. Udupa, G. T. Herman, I. L. Wornom III, and S. R. Buchman, "Registration of 3-D objects and surfaces," *IEEE Comput. Graph. Applicat.*, vol. 10, pp. 52–62, May 1990.

[39] F. Maes, D. Vandermeulen, and P. Suetens, "Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information," *Med. Image Anal.*, vol. 3, no. 4, pp. 373–386, 1999.

[40] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures," *Med. Phys.*, vol. 24, no. 1, pp. 25–35, 1997.

[41] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. New York: Cambridge Univ. Press, 1992.

[42] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Comput. Vis. Image Understanding*, vol. 77, no. 2, pp. 211–232, 1999.