

Improved inter-modality image registration using normalized mutual information with coarse-binned histograms

Haewon Nam^{1,*},[†], Rosemary A. Renaut², Kewei Chen³, Hongbin Guo²
and Gerald E. Farin⁴

¹*Samsung Electronics Co., Ltd., Kiheung-gu nongso-dong san #24, Yongin 446-711, Korea*

²*School of Mathematics and Statistics, Arizona State University, AZ, U.S.A.*

³*Alzheimer Institute and Banner Good Samaritan Positron Emission Tomography Center, AZ, U.S.A.*

⁴*Department of Computer Science and Engineering, Arizona State University, AZ, U.S.A.*

SUMMARY

In this paper we extend the method of inter-modality image registration using the maximization of normalized mutual information (NMI) for the registration of [¹⁸F]-2-fluoro-deoxy-D-glucose (FDG)-positron emission tomography (PET) with T1-weighted magnetic resonance (MR) volumes. We investigate the impact on the NMI maximization with respect to using coarse-to-fine grained B-spline bases and to the number of bins required for the voxel intensity histograms of each volume. Our results demonstrate that the efficiency and accuracy of elastic, as well as rigid body, registration is improved both through the use of a reduced number of bins in the PET and MR histograms, and of a limited coarse-to-fine grain interpolation of the volume data. To determine the appropriate number of bins prior to registration, we consider the NMI between the two volumes, the mutual information content of the two volumes, as a function of the binning of each volume. Simulated data sets are used for validation and the registration improves that obtained with a standard approach based on the Statistical Parametric Mapping software. Copyright © 2008 John Wiley & Sons, Ltd.

Received 1 April 2008; Revised 17 July 2008; Accepted 23 July 2008

KEY WORDS: registration; positron emission tomography; magnetic resonance

1. INTRODUCTION

The goal of object registration, which is used in various fields of image recognition and analysis, is to find a spatial deformation that matches two volumetric images. Our interest, and the focus of

*Correspondence to: Haewon Nam, Samsung Electronics Co., Ltd., Kiheung-gu nongso-dong san #24, Yongin 446-711, Korea.

[†]E-mail: haewonn@yahoo.com

Contract/grant sponsor: Arizona Department of Health Services
Contract/grant sponsor: NIH; contract/grant number: EB 2553301

this paper, is the mapping of functional data from [^{18}F]-2-fluoro-deoxy-D-glucose (FDG) positron emission tomography (PET) volumes with volumetric T1-weighted magnetic resonance (MR) volumes. Thus, throughout, when we refer to PET and MR volumes, it is understood that these are FDG-PET and T1-weighted MR.

Medical volume registration has emerged in recent years as a particularly active research field due to its many clinical applications. For the non-expert, we recommend [1] by Modersitzki in which a complete introduction and general mathematical framework of image registration are presented. Here we, therefore, do not repeat details that can be found in that text, but provide a brief overview of distinctive features of registration methods, most relevant to our own work. A method of registration is classified by the nature of the transformation used to obtain the registration, the cost function utilized to find the best fit for the given map, and whether or not specific points/landmarks are used as constraints in the optimization. Although options for the transformation include rigid, affine and elastic, the latter of which permits nonlinear deformations, global rigid and affine transformations are most frequently used [1]. Linear transformation, however, cannot completely overlay two different volumes [2], and for more detailed matching it is crucial to adopt a local non-rigid/elastic registration or to use a landmarks/feature-based approach. Transformations directly based on intensity-based matching, which assumes linear correlation of the intensities between the volumes, are inappropriate for inter-modality registration. The landmark-based approach requires intervention for feature identification, which may introduce further errors and increase computational time [1, 3]. Consequently, methods based on the statistics of the inter-modality volumes, using maximization of mutual information (MI), which was independently introduced in [3, 4], have been adopted in many applications [1, 5]. Most significant to our goal of matching volumes from PET and MR modalities is the extension of MI for non-rigid deformation using normalized mutual information (NMI) [6] and its implementation using a hierarchical B-spline technique [7].

Our work extends the investigation in [8] on the influence of parameter choices for the accuracy and efficiency of NMI registration. This is especially pertinent when not only are the modalities different, with differing characteristics with respect to the signal-to-noise ratios (SNR), but also the volume resolutions are not identical. For example, for the data under consideration in this paper, the MR volume is obtained on a voxel grid of dimension $256 \times 256 \times 124$, where each voxel is of size $0.938 \times 0.938 \times 1.5 \text{ mm}^3$, whereas the PET volumes are on a voxel grid of dimension $128 \times 128 \times 31$, with voxel dimensions $1.88 \times 1.88 \times 3.375 \text{ mm}^3$. Therefore, in seeking to maximize the NMI between two volumes of this kind, it is important to recognize that the NMI is limited by the actual information content of each volume. While [8] suggested the use of coarse-grained binning for inter-modality NMI, with the same number of bins in each volume, neither full investigation was presented on how to determine the appropriate binning, nor that this might be different for each modality. Here, we show that different binning strategies should be used for each volume, dependent on their underlying statistics. The number of bins to choose can be determined prior to registration by calculating the NMI of the non-aligned given volumes as a function of the binning.

In the remainder of the paper we briefly review in Section 2 the underlying aspects of the NMI algorithm, which are relevant to our application, referring to the standard references for further details. Experiments and results are described in Section 3. An extensive validation study is presented for the simulated data, in which we compare with the rigid body coregistration algorithm implemented in the Statistical Parametric Mapping (SPM) toolbox [9], www.fil.ion.ucl.ac.uk/spm/. We conclude, contrary to the observation in [8], that for this specific inter-modal registration, greater efficiency and accuracy is obtained by limiting the number of the histogram bins for both volumes. This effectively tunes the choice of total number of bins to the information content of the

volumes, as is demonstrated by considering the NMI between aligned and non aligned volumes as a function of the number of the bins in each volume. The main contribution of this paper is thus to show that analysis of the information content within a volume guides the optimal choice of binning to be used in the NMI-based registration. Moreover, our implementation also improves the registration as compared with SPM through including a limited hierarchical coarse-to-fine grained elastic transformation.

2. METHODS

We suppose there are two volumes that provide two different representations of the same object, or one represents a standard/template against which the second is to be compared. The object registration task is to find a suitable spatial transformation that allows the mapping of the volumes to each other as closely as possible [1]. The volume to be transformed is called the test or floating volume, whereas the target is called the reference. Here we assume that the floating object, a T1-weighted MR volume, is to be mapped to a reference FDG-PET volume; we seek to map the anatomical structure indicated in the MR volume to the functional information observed in the PET volume. Figure 1 illustrates a slice of a typical FDG-PET volume as compared with a standard T1-weighted MR. Notice that the PET volume, which provides only functional information, does not show the skull. Practically, therefore, for registration of these modalities, scalp-peeling is first applied to the MR volume so as to remove the high-intensity values of the skull, [10], Figure 1(c). Additionally, we also consider the use of a clustered PET image, Figure 1(d) [11], as reference.

2.1. Normalized mutual information

To avoid dependence on the amount of volume overlap between two volumes, we maximize the NMI between the two images rather than the MI [6, 12, 13]. Because no assumptions are made

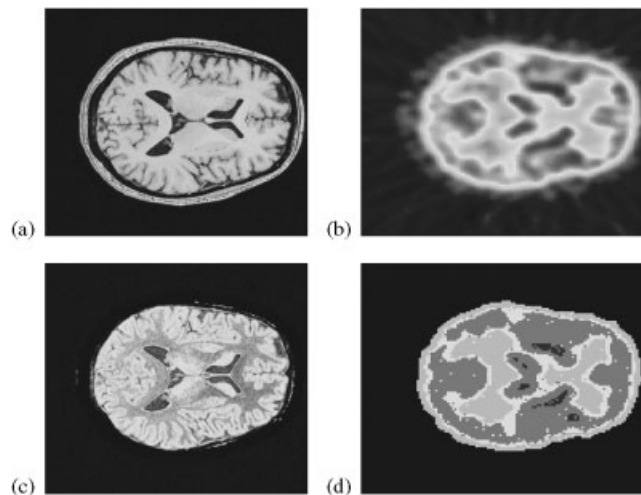


Figure 1. Comparison of PET/MR: (a) T1-MR; (b) PET; (c) scalp-peeled MR; and (d) clustered PET.

regarding the nature of the relation between the voxel intensities in both modalities, this criterion is very general and can be applied in a large variety of applications [14]. The entropy of the voxel intensity distribution of a volume is calculated via $H(A) = -\sum_{a \in A} P(a) \log P(a)$, where $P(a)$ is the probability distribution of the voxel intensities of volume A . If $H(A)$, $H(B)$ are used to denote the marginal entropies of volumes A and B , and $H(A, B)$ denotes their joint entropy, then the NMI, which expresses the amount of information that volume A contains about the second volume B , is given by $\text{NMI}(A, B) = (H(A) + H(B)) / (H(A, B))$. The mutual entropy of the two volumes is similarly defined in terms of the joint probability distribution of the voxel intensities, summed over all overlapping voxel pairs.

2.2. Histograms

To calculate the marginal and joint probability distribution functions, we use the histogram method [15], and rebin the volume intensities to n_R and n_T bins for the reference and test volumes, respectively. As compared with previous studies [6–8, 15] we allow $n_R \neq n_T$, and investigate using a coarse-grained binning closely related to the information content of the underlying volumes. Although the MR volumes typically exhibit good SNR, high resolution, and relatively clearly defined anatomical structure, one might not expect, from a volume with in a plane grid of 256×256 , 256 different intensity levels representing 256 different features. With the PET data, the SNR is poor and the resolution reduced, and the ability to detect a high number of significant features is further lowered. In particular, if we consider the histogram binning as providing a segmentation, then, when a segmentation is obtained with a large number of bins relative to the number of voxels, it is not meaningful to consider bins with very few voxels as providing reliable information [8].

2.3. Mapping and refinement

Mathematically speaking, let $R(\mathbf{x})$ and $T(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ be different modality reference and test volumes, respectively. We seek the optimal transformation $\mathbf{f}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, such that $T(\mathbf{f}(\mathbf{x}))$ occupies as closely as possible the same spatial volume as $R(\mathbf{x})$ [1], while also accepting that due to the different voxel intensity distributions of the volumes we cannot expect to make $T(\mathbf{f}(\mathbf{x})) \approx R(\mathbf{x})$. Moreover, there may be real local nonlinear differences in the volumes when $R(\mathbf{x})$ and $T(\mathbf{x})$ are from different human subjects, or acquired from the same person but before and after the onset of a disease, a treatment or simply after a long time interval. In our implementation, we compose a rigid body and nonlinear elastic deformation so as to allow for more accurate comparison of these small alterations. Specifically, transformation \mathbf{f} is obtained as the composition of a global transformation $\mathbf{f}_{\text{global}}$ with the elastic local map $\mathbf{f}_{\text{local}}$; $\mathbf{f}(\mathbf{x}) = \mathbf{f}_{\text{local}} \circ \mathbf{f}_{\text{global}}$.

Practically, we first find the global linear map followed by free form deformation (FFD) using a hierarchical cubic B-spline representation of the data. The movement of just one control point from the control grid of the hierarchical B-spline representation only affects the spline coefficients in a local neighborhood of size 4 in each dimension of the control point. The FFD thus accounts for modeling of local non-rigid deformations by local refinements of the control grid [6, 7, 16, 17]. In our work, the control space is halved as needed, and the values of the new control points are calculated from the old values. Mathematically, let $\phi_{j,k}^i$ be the (j, k) th control point at level i of control point mesh Φ^i . We will assume that the control point spacing is halved in every step. In this case, the position of control point $\phi_{2j,2k}^{i+1}$ coincides with that of $\phi_{j,k}^i$. All control points at the $i+1$ st level can be calculated directly from those of the i th level by a standard B-spline

subdivision algorithm [16, 18]. Each control point mesh, Φ^i , defines the spline-based FFD, and hence defines a local transformation $\mathbf{f}_{\text{local}}^i$ at each level of resolution. Their composition defines the local transformation

$$\mathbf{f}_{\text{local}}(\mathbf{x}) = \mathbf{f}_{\text{local}}^L \circ \dots \circ \mathbf{f}_{\text{local}}^0(\mathbf{x}) \quad (1)$$

To determine whether the grid needs to be locally refined through the FFD, we note that $\text{NMI}(A, B) \in [1, 2]$, and refine only when $\text{NMI}(A, B) < 2 - \varepsilon$. Tolerance ε has to be chosen such that refinements will occur inside the actual brain volume and not outside, where the locally calculated NMI is always close to 2. Moreover, for the data inside the brain volume, we can only match the volumes up to a limit, which is strictly less than 2 because of the noise level. Therefore, by choosing a suitable value for ε , we avoid unnecessary matching for data outside the brain, while assuring the best possible match within the brain. Notice that $\sum_i \text{NMI}(A_i, B_i) \neq \text{NMI}(A, B)$ where $\{A_i\}$ and $\{B_i\}$ are subdivisions of A and B , respectively. Indeed, NMI depends not only on the relation $H(A, B)$ between two volumes, but also on the volumes themselves $H(A)$, $H(B)$, and achieves 2 if and only if two volumes are perfectly aligned, $H(A) = H(B) = H(A, B)$. Finding an ‘optimal’ ε , which is not the purpose of this paper, is an important and interesting topic, and depends on the problem set at hand, the instrumentation and noise levels of each modality. Although this is worthy of consideration, it is a parameter that will need to be rechosen dependent on the modalities in question and the noise levels. We have found for our experiments, by trial and error, that $\varepsilon = 0.2$ is a suitable choice that prevents the algorithm subdividing virtually perfectly matched regions outside the brain volume. In order to also maintain the quality of the joint histogram calculated on a subregion, namely so that it is defined by sufficient data samples for our 2D simulations, we limit the number of levels of refinement to $L = 1$, as noted in [19].

2.4. Interpolation

It is well known that the mapping of a volume, whether by rigid or FFD transformation, moves a voxel in the original volume to a new point, which need not coincide exactly with any voxel location in the original grid [3, 4, 6, 15]. Therefore, interpolation is required to estimate the intensity values at the new grid points. We adopt the standard method of partial volume interpolation, which was determined to be most accurate for this application, as compared with the trilinear and nearest neighbor approaches [15].

2.5. Optimization

Many reports that use B-spline FFD as a local deformation include the thin-plate spline energy as a smoothing regularization term [6, 17]. This is particularly appropriate when dense control point meshes are utilized and it makes sense then to look for local smoothness. For our application, the two different volumes do not have the same inherent structure and we can neither expect to use very fine meshes nor expect perfect alignments. Moreover, introduction of the regularization introduces an additional regularization parameter and computational overhead. Here, therefore, we do not augment the NMI functional by a regularization term. In the optimization, we find the optimal \mathbf{f} given by

$$\mathbf{f} = \underset{\mathbf{f}}{\text{argmin}} \mathcal{C}(\mathbf{f}) \quad (2)$$

where $\mathcal{C}(\mathbf{f}) = -\text{NMI}(R(\mathbf{x}), T(\mathbf{f}(\mathbf{x})))$. First the global mapping that gives the minimum over all rigid transformations is obtained. It is important to perform this as an independent first stage so as

to make sure that the objects are oriented correctly relative to each other, with respect to translation and rotation. Local differences are then handled in the second stage in which we seek the local elastic modifications of this rigid map. We use the simple gradient descent method, as described in [6], only modified so as to incorporate the knot insertion for local refinement of the control grid in the regions where the test and target volumes have a large deviation as measured by the calculation of the NMI for that region. Although our 2D experiments use $L = 1$, we give the algorithm for the general case with arbitrary L .

Algorithm 1

Tolerance ζ for the minimization of the gradient, tolerance $0 < \varepsilon < 0.5$ for the threshold on the NMI measurements, and the initial line search parameter μ_0 for the gradient descent method are given.

- (1) **Compute** global rigid body transformation $\mathbf{f}_{\text{global}} = \text{argmin}_{\mathbf{f}} \mathcal{C}(\mathbf{f})$.
 - (2) **Initialize** the control points Φ^1 . Set $l = 1$. $\mu = \mu_0$.
 - (3) **For** the control points related to the regions where NMI of the test and target images is less than the threshold $2 - \varepsilon$, **repeat**
 - (a) Compute the gradient vector $\nabla \mathcal{C} = \partial \mathcal{C}(\mathbf{f}(\Phi^l)) / \partial \Phi^l$ of the cost function (2) with respect to the non-rigid transformation parameters Φ .
 - (b) **While** $\mu \|\nabla \mathcal{C}\| > \zeta$, **do**
 - (i) **Until** $\mathcal{C}(\mathbf{f}(\Phi_{\text{new}}^l)) < \mathcal{C}(\mathbf{f}(\Phi^l))$

$$\Phi_{\text{new}}^l = \Phi^l - \mu \nabla \mathcal{C}, \quad \mu = \mu/2$$
 - (ii) $\Phi^l = \Phi_{\text{new}}^l$, $\mu = \mu_0$.
 - (iii) Recalculate the gradient vector $\nabla \mathcal{C}$.
 - (c) **Increase** the control point resolution by knot insertion. Set $l = l + 1$.
 - (d) **Increase** the image resolution.
- until** $l > L$, the finest level L is reached.

3. EXPERIMENTS

3.1. Data setup

Quantifying registration accuracy of real volumes is difficult, especially for inter-modality registration. We therefore validate the coarse-binned hierarchical scheme for registering PET/MR using the approach presented in [10] for validating and comparing registration tools for mapping of PET/MR modalities. There, real MR volumes are obtained and then used to generate a set of simulated PET volumes, which are randomly transformed by a rigid map relative to the original MR.

3.1.1. Real MR data. A T1-weighted pulse sequence (radio-frequency-spoiled gradient recall acquisition in the steady state, repetition time = 33 ms, echo time = 5 ms, flip angle = 30°, one excitation, field-of-view = 24 cm, imaging matrix = 256 × 192, slice thickness = 1.5 mm, scan time = 13 : 36 min) was used to acquire 124 contiguous horizontal MRI slices on a 256 × 256 grid with voxel dimensions 0.938 × 0.938 × 1.5 mm³ for a normal subject. The volume was visually confirmed to be free of artifacts, 22 non-brain slices were removed, and then resampled using trilinear interpolation to 256 × 256 × 153 to give voxel boxes of size 0.938 × 0.938 × 1 mm³.

3.1.2. *Simulated PET data sets.* The procedure, outlined below, for generating simulated PET volumes is similar to that in [10].

- (1) The MR volume is scalp-edited and segmented using the SPM2 software into gray matter, white matter and CSF components. The intensities of the segmented MR volume are then rescaled to give means of CSF from white to gray regions satisfying 1:3.3:10.
- (2) 10 different random 2D rigid body within plane transformations comprising of a rotation θ and a translation (t_x, t_y) are applied to the MR volume, where $\theta \in [-10^\circ, 0]$ and $(t_x, t_y) \in [0, 24] \text{ mm} \times [-24, 0] \text{ mm}$. In each case, the random transformation, θ, t_x, t_y , is recorded for use in the validation experiments and the relative misalignment is measured in two norm for the vector $(\theta/10, t_x/24, t_y/24)^T$.
- (3) Each of the 10 volumes is then smoothed by a three-dimensional Gaussian filter with a resolution of 7 mm FWHM in each direction.
- (4) Gaussian white noise of 10, 30, and 50% of the mean intensity of the brain voxels is added, yielding a total of 30 volumes.
- (5) Each volume is then smoothed to PET resolution of about 8 mm FWHM using a Gaussian filter.

In total, 30 simulated volumes are obtained. Figure 2 provides an example of a simulated PET at each noise level as compared with the intensity-mapped MR from Step 1. For our experiments, each transformed PET volume is also clustered to five different clusters, using the clustering tool described in [11], generating an additional 30 data sets.

The obtained volumes are then downsampled to size $128 \times 128 \times 16$, with voxel sizes $1.88 \times 1.88 \times 3.375 \text{ mm}^3$, with the subvolume of 16 slices chosen to include significant features of the

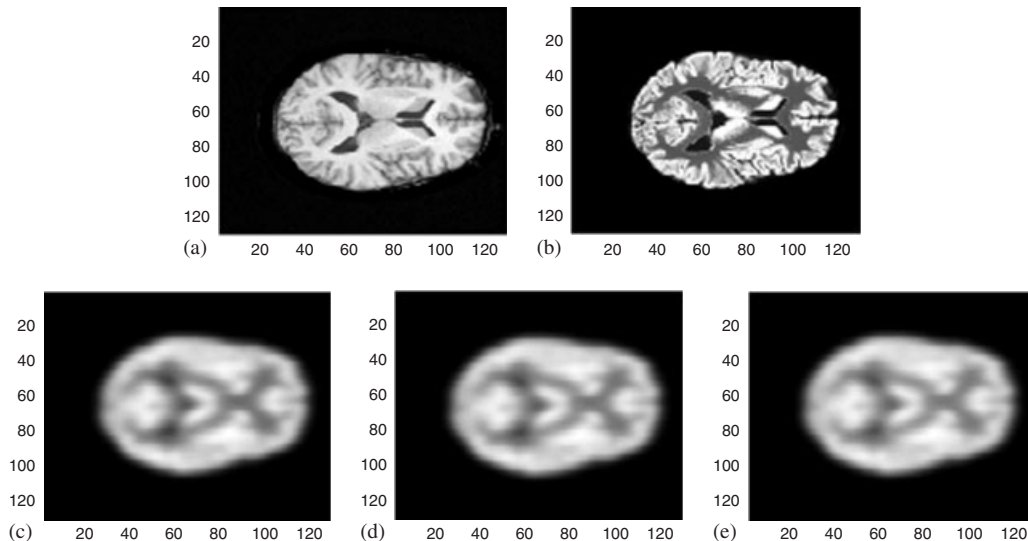


Figure 2. Simulated PET from original MR, at slice number 8, three different noise levels. (b) shows the segmented MR as described in Section 3.1.2 (1): (a) original MR; (b) segmented MR; (c) 10% noise; (d) 30% noise; and (e) 50% noise.

brain volume, sampled from volume slices 51 to 104, and representing a total of 35% of the volume.

3.2. Investigating dependence of NMI on the histogram binning and clustering

Experiments were carried out to evaluate the impact of changing the number of bins in the volume histograms, or the number of clusters in the volume, on the actual calculation of the NMI for aligned and misaligned volumes, prior to assessing the performance of the registration algorithms. In the interests of keeping the presentation clear and concise, in the following we present illustrative results for 2D; the 3D results are similar.

In Figures 3(a), (b) we show how the NMI between the MR and aligned simulated PET, for a representative slice with noise level 30%, varies with respect to changing the number of bins. Figure 3(a) illustrates the NMI for an aligned slice (solid line) and misaligned slice (crossed line) with misalignment 0.7794, using $n_T = 32$ and varying n_R , $1 \leq n_R \leq 100$. Figure 3(b) is similar to 3(a) but illustrates the impact on the NMI of varying n_T while $n_R = 32$ is fixed. Notice that the NMI is insignificant when the number of bins of either slice is reduced much below 10, but also as the number of bins increases significantly above 10, the NMI decreases slightly. For this specific example in Figure 3(a) the maxima occur at $n_R = 10$, aligned and misaligned, and in Figure 3(b) for $n_T = 16, 12$, for aligned and misaligned cases, respectively. This is representative of the average maxima obtained for the 30×16 simulated data sets, see Table I. As expected the NMI for the misaligned slices is less than for the aligned slices, increasing about 10% with alignment. This information guides the algorithm for registering slices. Assuming that a rough initial alignment is made, the actual NMI can only be improved through alignment by about 10% and it is thus unreasonable to impose a stopping criterion which forces the global NMI to increase significantly more than 10% above its starting value. Moreover, the calculation of the NMI should be based on relatively coarse-binned histograms.

In Figure 3(c) and Table II, we present a similar study that illustrates the change in NMI with respect to clustering of the PET slice for fixed $n_T = 32$, from which we conclude that the optimal number of clusters for maximizing NMI does not depend on the noise level.

The NMI increases with the alignment of the slices and is uniformly higher using the optimal choice $(n_R, n_T) = (10, 16)$, as compared with using 32 bins in each case, Figure 4. The x -axis

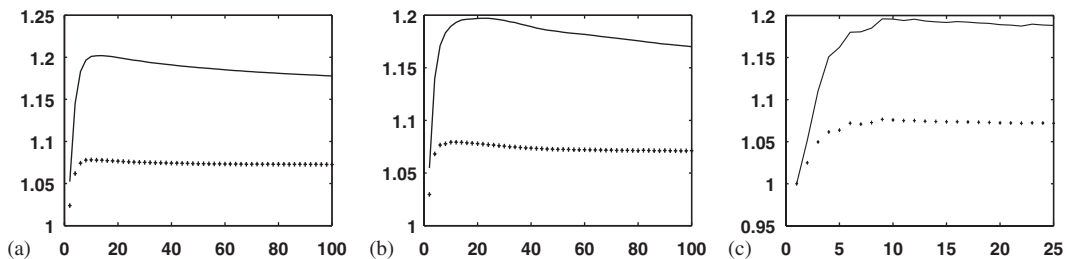


Figure 3. Illustrating the NMI with respect to the number of bins for perfectly aligned data (solid line), and not aligned (crossed line), for noise level 30%. Plot (a, varying n_R , $n_T = 32$) illustrates how NMI changes with respect to the number of PET bins, and Plot (b, varying n_T , $n_R = 32$) with respect to the number of MR bins. Plot (c, varying PET clusters) illustrates the NMI with respect to the number of clusters over the PET slice for perfectly aligned data, solid line, and not aligned, crossed line, for noise level 30%, when the number of bins for MR is 32.

Table I. Average number (rounded) of bins at which NMI is maximized, compare Figure 3, using 32 bins in the other case. The average norm of misalignment is 0.8849.

Noise level (%)	Aligned		Misaligned	
	PET	MR	PET	MR
10	11	16	10	13
30	10	16	10	13
50	11	16	10	13

Table II. Average number (rounded) of clusters of the PET slice at which NMI is maximized, compare Figure 3(c).

Noise level (%)	Aligned	Misaligned
10	10	9
30	13	12
50	12	12

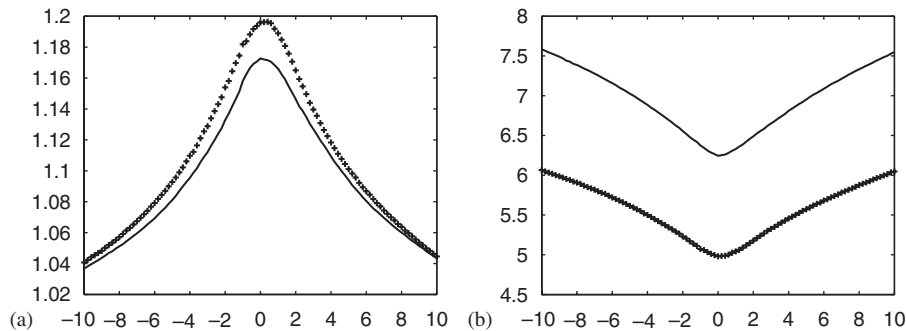


Figure 4. 2D case with 30% noise. Illustrating (a, plot of NMI against rotation) increase in NMI, and (b, plot of joint entropy $H(A, B)$ against rotation) decrease in joint entropy with alignment for optimal (10, 16) bins (+ crosses), suboptimal binning (32, 32) bins (solid line).

measures the displacement angle. The crossed line shows the NMI for the optimal choice and the solid line for a suboptimal choice. Figure 4(b) demonstrates that the joint entropy $H(A, B)$ is indeed less for the optimal binning of the slice histograms.

Finally, we compare visually the impact of histogram equalization and clustering, see, Figure 5, which shows the result of applying clustering and histogram equalization to 5 and 10 levels in each case, to a slice containing 50% noise. Both techniques serve to increase the SNR, but yield different rough segmentations. However, clustering introduces an additional preprocessing step prior to registration, while histogram binning is intrinsic to the NMI algorithm. Thus, unless clustering serves to improve the registration accuracy for reduced cost, there would be no particular benefit in its use. We will see later that clustering offers no such advantage for improving the NMI registration algorithm, in Section 3.3.

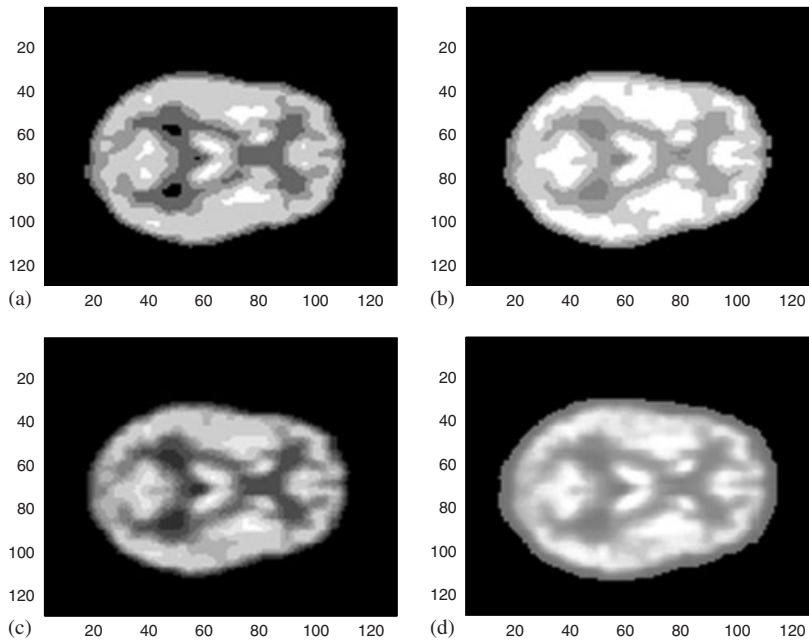


Figure 5. Illustrating impact of clustering and histogram equalization for a simulated PET slice with 50% noise, Figure 2(f): (a) clustered: 5 clusters; (b) equalized: 5 bins; (c) clustered: 10 clusters; and (d) equalized: 10 bins.

3.3. Evaluating the registration algorithms

Experiments were carried out to validate the observations on the impact of clustering and binning on the performance of the registration algorithms, as well as on the performance of the NMI with respect to the levels of refinement used in the rigid and elastic registrations. It might be supposed that the elastic registration is in this case immaterial, because the simulation presents only a random rigid transformation. Yet the additional processes described in Section 3.1.2, by which noise, smoothing and downsampling are applied to the volumes, have the impact of adding small nonlinear changes to the volumes. It is thus of interest to find whether the elastic step allows an improvement beyond the rigid transformation, and hence accounts for these small nonlinear changes, with the understanding that elastic registration would truly be of use for matching volumes for which larger nonlinearities might be expected. Again, we only present results of our 2D experiments.

3.3.1. Registration experiments. For each generated random transformation, we record the obtained transformation in order to be able to assess the ability of the NMI algorithm to find the specific transformation. Then, given the exact known random transformation of the simulated PET with respect to the original MR, the robustness can be measured as a relative error of the parameters of the obtained rigid NMI transformation. When the elastic deformation is also included, however, we cannot use the same measure of accuracy. Hence, in all experiments we report instead the L_2 relative error in the matching of intensity values over the entire overlapped image, where the

values at the warped voxel locations are interpolated back to the nearest underlying grid values. All results are reported as averages over 160 cases, with standard deviation in parentheses. For the experiments we used initial line search parameter $\mu_0=0.1$, threshold on the NMI $\varepsilon=0.2$, threshold on the gradient $\zeta=10^{-8}$, and central finite difference approximations of second order for calculating the gradient vector.

In Table III the cases labeled A, B, C, D are experiments for which in A there are 5 clusters, for B 5 bins, C 10 bins, and D the standard of 32 bins. Case C is the optimal choice as determined by graphing NMI relative to number of bins against $n_T=32$. The rigid body algorithm uses the fine level grid, 128×128 . Results indicated as FFD are those obtained in which we seek to improve the rigid registration directly with a fine level FFD grid. Results indicated as FFD2 use one level of refinement for the FFD, namely from grids using 64 points to 128. Further coarsening the grid reduces the statistical content of the histograms. Experiments in which we first registered linearly at the coarse level, 64×64 , before refining to the fine grid did not yield improvement and are not reported. For comparison, the reported costs are normalized with respect to the standard case, D.

3.3.2. *Registration observations.* The different types of registration are contrasted in Table III with respect to the relative error, which is defined as below.

$$\text{Error} = \frac{\|T(\bar{\mathbf{f}}(\mathbf{x})) - T(\mathbf{f}^*(\mathbf{x}))\|}{\|T(\mathbf{f}^*(\mathbf{x}))\|}$$

where $\bar{\mathbf{f}}$ is the estimated transformation and \mathbf{f}^* is the *exact*-known solution, which perfectly aligns the reference and the target images. $\|\cdot\|$ represents the Euclidean length, i.e. $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n x_i^2}$ for

Table III. Average relative error over 10 simulated PET cases, where MR uses 32 bins, except for the optimal case where we use 16 bins for the MR. Standard deviation in parentheses.

Case	Binning (n_R, n_T)	Rigid		FFD		FFD2	
		Error	Cost	Error	Cost	Error	Cost
<i>10% noise level</i>							
A	(5, 32)	0.0354 (0.021)	0.9487	0.0342 (0.020)	1.9350	0.0323 (0.020)	1.5955
B	(5, 32)	0.0219 (0.014)	0.4898	0.0233 (0.012)	1.2630	0.0203 (0.014)	0.9610
C	(10, 16)	0.0209 (0.010)	0.7120	0.0232 (0.008)	1.2451	0.0207 (0.010)	1.1054
D	(32, 32)	0.0586 (0.046)	1.000	0.0511 (0.033)	3.3822	0.0557 (0.043)	2.4253
<i>30% noise level</i>							
A	(5, 32)	0.0332 (0.023)	0.9845	0.0297 (0.022)	1.5784	0.0288 (0.018)	1.4448
B	(5, 32)	0.0268 (0.026)	0.8359	0.0256 (0.023)	1.4615	0.0242 (0.024)	1.2230
C	(10, 16)	0.0260 (0.015)	0.8086	0.0246 (0.012)	1.2075	0.0221 (0.013)	1.0935
D	(32, 32)	0.0326 (0.023)	1.000	0.0331 (0.019)	2.4175	0.0315 (0.018)	2.1609
<i>50% noise level</i>							
A	(5, 32)	0.0235 (0.013)	0.5785	0.0241 (0.009)	1.3101	0.0211 (0.012)	0.9525
B	(5, 32)	0.0284 (0.016)	0.5194	0.0275 (0.016)	1.2831	0.0245 (0.016)	1.0236
C	(10, 16)	0.0215 (0.017)	0.5123	0.0202 (0.016)	0.9950	0.0207 (0.017)	0.8768
D	(32, 32)	0.0359 (0.026)	1.000	0.0329 (0.022)	2.8323	0.0328 (0.020)	2.4380

a vector and $\|T\| = \sqrt{\sum_{i=1}^m \sum_{j=1}^n T_{i,j}^2}$ for a two-dimensional image. It can be immediately noted from Table III that the presumed optimally binned case, case C, yields the minimum average relative error and certainly the lowest standard deviation over virtually all noise levels and algorithmic choices. Generally, the standard deviation is relatively large with respect to the average error, because NMI is a statistical measurement, and the 2D results tend to be unstable for small data sets. Note also that the cost for just the rigid stage of the registration decreases significantly when using a limited number of clusters or bins, as compared with case D. In particular, for high noise, the optimal binning reduces cost to 51% of the rigid standard case while generating much reduced error.

However, including FFD in the registration may reduce the overall error of the registration, this does not come without additional cost. Although we have not yet sought to optimize our implementation, we can observe that the elastic registration increases the computational cost. On the other hand, using a multilevel FFD, namely from coarse-to-fine (FFD2) reduces cost, in some cases by as much as 30%.

The results presented support the premise for inter-mode registration of slices using maximization of NMI, one should base the NMI calculations on coarse-binned histograms. This offers the additional advantage of speeding the calculation while maintaining accuracy.

4. CONCLUSIONS

We have shown the power of utilizing NMI for registration of PET and MR slice and volume data, in which we speed the calculation through utilizing coarse-binned histograms of voxel intensities. The success of this approach validates our observation that the information content between PET and MR volumes is intrinsically limited both by the resolution capabilities of the instrumentation, and the actual information content that one modality provides about the other. Our results were validated through careful examination on simulated data. To determine the histogram binning to use for a given inter-modality registration, we suggest that the NMI of the given volumes be calculated as a function of the number of bins of each volume, and that the argument yielding the maximum be used for the object registration, possibly to then be refined at alignment if finer binning yields improved NMI.

Future work will include the extension of the application of the registration algorithm for application to real data sets and full 3D evaluation.

ACKNOWLEDGEMENTS

This work was partially supported by the Arizona Alzheimer's Disease Center which is funded by the Arizona Department of Health Services, and by NIH grant EB 2553301.

REFERENCES

1. Modersitzki J. *Numerical Methods for Image Registration* (1st edn). Oxford University Press: New York, U.S.A., 2004.
2. Zitova B, Flusser J. Image registration methods: a survey. *Image and Vision Computing* 2003; **21**:977–1000.
3. Collignon A, Maes F, Delaere D, Vandermeulen D, Seutens P, Marchal G. Automated multimodality image registration using information theory. In *Proceedings of the 14th International Conference on Information*

IMPROVED INTER-MODALITY IMAGE REGISTRATION

- Processing in Medical Imaging: Computational Imaging and Vision 3*, Bizails Y, Barillot Ca, Di Paola R (eds). Kluwer Academic Publishers: Boston, June 1995; 263–274.
4. Viola P. Alignment by maximization of mutual information. *Ph.D. Dissertation*, Massachusetts Institute of Technology, Cambridge, MA, 1995.
 5. Pluim JPW, Maintz JBA, Viergever MA. Mutual-information-based registration of medical images: a survey. *IEEE Transactions on Medical Imaging* 2003; **22**(8):986–1004.
 6. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging* 1999; **18**(8):712–721.
 7. Likar B, Pernuš B. A hierarchical approach to elastic registration based on mutual information. *Image and Vision Computing* 2001; **19**:33–44.
 8. Zhu Y-M, Cochoff SM. Influence of implementation parameters on registration of MR and SPECT brain images by maximization of mutual information. *Journal of Nuclear Imaging* 2002; **43**(2):160–166.
 9. Statistical Parametric Mapping. Wellcome Department of Cognitive Neurology. Available from: <http://www.fil.ion.ucl.ac.uk/spm/> (12 January 2003).
 10. Kiebel SJ, Ashburner J, Poline J-B, Friston KJ. MRI and PET coregistration—a cross validation of statistical parametric mapping and automated image registration. *NeuroImage* 1997; **5**:271–279.
 11. Guo H, Renaut RA, Chen K, Reiman E. Clustering huge data sets for parametric PET imaging. *Biosystems* 2003; **71**:81–92.
 12. Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognition* 1999; **32**:71–86.
 13. Studholme C, Hill DLG, Hawkes DJ. Automated 3D registration of MR and PET brain images by multi-resolution optimization of voxel similarity measures. *Medical Physics* 1997; **24**(1):25–35.
 14. Collignon A. Multi-modality medical image registration by maximization of mutual information. *Ph.D. Thesis*, 1998.
 15. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging* 1997; **16**(2):187–198.
 16. Lee S, Wolberg G, Shin SY. Scattered data interpolation with multilevel B-splines. *IEEE Transactions on Visualization and Computer Graphics* 1997; **3**(3):228–244.
 17. Xie Z, Farin GE. Image registration using hierarchical B-splines. *IEEE Transactions on Visualization and Computer Graphics* 2004; **10**(1):85–94.
 18. Forsey DR, Bartels RH. Hierarchical B-spline refinement. *ACM Transactions on Computational Graph* 1988; **22**(4):205–212.
 19. Xu X, Dony RD. Evaluation of hierarchical elastic medical image registration method. *Proceedings of the 2004 Canadian Conference on Electrical and Computer Engineering*, Ontario, Canada, vol. 3, 2–5 May 2004; 1289–1292.