

Similarity Measures for Non-rigid Registration

Peter Rogelj and Stanislav Kovačič

Faculty of Electrical Engineering, University of Ljubljana

ABSTRACT

Non-rigid multimodal registration requires similarity measure with two important properties: locality and multimodality. Unfortunately all commonly used multimodal similarity measures are inherently global and cannot be directly used to estimate local image properties. We have derived a local similarity measure based on joint entropy, which can operate on extremely small image regions, e.g. individual voxels. Using such small image regions reflects in higher sensitivity to noise and partial volume voxels, consequently reducing registration speed and accuracy. To cope with these problems we enhance the similarity measure with image segmentation. Image registration and image segmentation are related tasks, as segmentation can be performed by registering an image to a pre-segmented reference image, while on the other hand registration yields better results when the images are pre-segmented. Because of these interdependences it was anticipated that simultaneous application of registration and segmentation should improve registration as well as segmentation results. Several experiments based on synthetic images were performed to test this assumption. The results obtained show that our method can improve the registration accuracy and reduce the required number of registration steps.

Keywords: similarity measure, registration, segmentation, entropy, joint distribution

1. INTRODUCTION

Non-rigid registration^{1,2} is a process for maximizing spatial image correspondence of two images within the constraints of an image transformation model, to bring the features of first image into alignment with those of the second image. Image correspondence is measured using similarity measures, which compare the data values at corresponding points in the images. A variety of similarity measures is suitable for images obtained using the same acquisition method³, where the most often used are methods based on difference of scalar intensity at corresponding image locations, attractive because of their simplicity, and correlation based methods, which generally produce better results. The situation is quite different when the images to be non-rigidly registered are acquired using different imaging procedures. There are two main issues to be addressed: first, the similarity measure must be capable of determining the correspondence between images of different measurement type, and second, the measure must be sensitive to local image differences. The problem is made difficult because the two properties desired of our similarity measure are contradictory.

In practice, image differences appear on a range of scales, from “global” to “local”. The aim of non-rigid registration is to correct those that comply with the image transformation model. Therefore, it is desired to be able to correct even the smallest image discrepancies, but to do so they must previously be detected. Therefore, similarity must be estimated from appropriately small image regions, which in extreme case become individual voxels. It is noted^{4,5} that registration based on individual voxels is ill-posed if displacements are calculated separately for each image voxel. For this purpose non-rigid registration must always incorporate a spatial model, e.g. elastic⁶ or viscous fluid model⁷, which defines the relations between individual region/voxel displacements.

1.1. Multimodal similarity measures

Multimodal similarity measures are capable of determining the correspondence between images of different measurement type, and therefore the correspondence between different intensity values for the same anatomical structure in different images. The relation between intensities is not linear, as is the case when registering images of the same modality. Usually it is so complex, that it is not possible to uniquely map intensities of one image to the intensities of the other one. In general the relation is not injective as the intensity values of the first image can correspond to

Corresponding author:

E-mail: peter.rogelj@fe.uni-lj.si

Address: Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, 1001 Ljubljana, Slovenia

more than one intensity value of the other image. Generally multimodal similarity measures model this complex and unknown relationship between image intensities statistically.

The most commonly used multimodal similarity measure is mutual information. It was first proposed and brought to the medical imaging field from information theory by Viola and Wells⁸. When used as a similarity measure it measures the statistical dependence between the image intensities. In principle, it reveals how much one image tells us about the other image, and takes on maximum value when the images are geometrically aligned. Given two images A and B mutual information $I(A, B)$ between them is defined as

$$I(A, B) = H(A) + H(B) - H(A, B), \quad (1)$$

where $H(A)$ and $H(B)$ denote marginal entropies of A and B , and $H(A, B)$ is their joint entropy. The entropies can be calculated using well-known Shannon definition:

$$H(\cdot) = - \sum p(\cdot) \log p(\cdot), \quad (2)$$

where $p(\cdot)$ stands for either marginal or joint probability distributions, estimated from image intensities. There are also other types of multimodal similarity measures like energy similarity measure⁹ and various types of generalized entropies¹⁰. Nevertheless, due to their statistical nature they are all inherently global in the sense that they all require relatively large image regions to achieve sufficiently high statistical significance of joint distribution. Various solutions to improve their locality and to overcome the problem of small sample size were proposed, e.g. using prior joint probability¹¹, reducing number of joint histogram bins, resorting to one dimensional statistics $H(A - B)$ ¹², etc. Nevertheless, among all methods that have been proposed only local multimodal similarity measures enable estimation of similarity from arbitrarily small image regions, e.g. individual voxels. We have derived a similarity measure of that type, which is based on joint entropy, and is described in the next section.

2. LOCAL MULTIMODAL SIMILARITY MEASURES

One of the frequently used multimodal similarity measures is the Shannon entropy $H(A, B)$ calculated from joint intensity distribution. With respect to spatial image correspondence it is also the most significant term of mutual information measure (Eq.(1)). Other two terms, i.e. the marginal entropies, are much less important, because when using partial volume interpolation¹³, they can only change due to different image covering at image borders. This is why both similarity measures under certain circumstances yield practically identical indication about image correspondence, see Fig. 1.

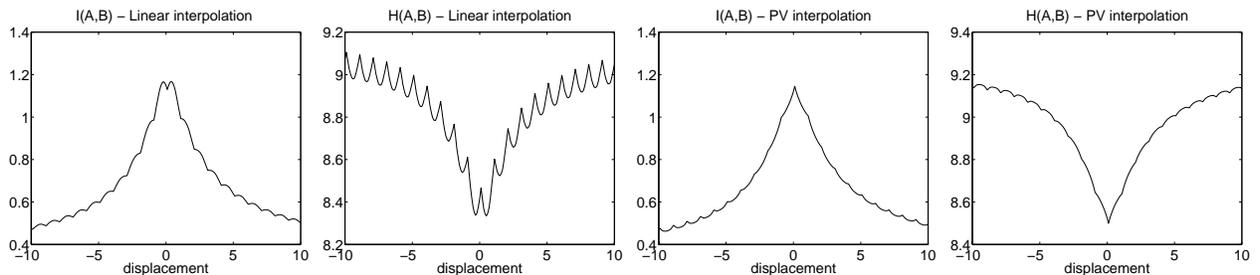


Figure 1. Similarity measures I (mutual information) and H (joint entropy) as a function of image translation for linear and partial volume (PV) interpolation.

We propose a local similarity measure derived from joint entropy that can be used for estimating similarity of arbitrarily sized image regions, including the regions with the size of only one image voxel. Let us rewrite Eq. (2) in the following form,

$$H = - \sum_{\mathbf{i}} p_{\mathbf{i}} \log p_{\mathbf{i}} = - \sum_{\mathbf{i}} \frac{N_{\mathbf{i}}}{N} \log p_{\mathbf{i}} \quad ; \quad \mathbf{i} = [i_A, i_B]^T, \quad (3)$$

where \mathbf{i} is an intensity vector consisting of the intensity values of the images A and B , and $p_{\mathbf{i}}$ is the joint probability of intensity pair \mathbf{i} . $N_{\mathbf{i}}$ is the number of occurrences of this intensity pair, and N is the total number of intensity

pairs in the image, which is usually equal to the number of image voxels. Let us assume that the images are divided into smaller non-overlapping regions R_r , each containing, say, $N_{\mathbf{i}}^r$ occurrences of intensity pair \mathbf{i} . The total number of occurrences $N_{\mathbf{i}}$ in the whole image can be obtained by summing $N_{\mathbf{i}}^r$ over all regions,

$$N_{\mathbf{i}} = \sum_r N_{\mathbf{i}}^r. \quad (4)$$

Substituting Eq. (4) in Eq. (3) we obtain

$$H = - \sum_r \sum_{\mathbf{i}} \frac{N_{\mathbf{i}}^r}{N} \log p_{\mathbf{i}}. \quad (5)$$

If the image is divided into individual voxels v , then

$$N_{\mathbf{i}}^r = N_{\mathbf{i}}^v = \begin{cases} 1 & ; \quad \mathbf{i} = [A(v), B(v)]^T \\ 0 & ; \quad \text{otherwise} \end{cases}, \quad (6)$$

and Eq. (5) can be simplified to

$$H = - \sum_r \frac{1}{N} \log p_{\mathbf{i}} = \sum_v - \frac{1}{N} \log p_v = \frac{1}{N} \sum_v h(v), \quad (7)$$

where p_v is the probability $p_{\mathbf{i}}$ for $\mathbf{i} = [A(v), B(v)]^T$, and $h(v) = -\log p_v$ is the ‘‘uncertainty’’ of this intensity pair. Thus, it is possible to calculate the entropy, which is a global similarity measure, from similarities of individual voxels $S^v = h(v)$ by summing up (averaging) over the whole image. Of course, summation can also be performed over arbitrary smaller image region R_r and the result is local similarity S^r of this region.

$$S^r = \frac{1}{N_r} \sum_{v \in R_r} h(v). \quad (8)$$

Notice that $h(v)$ is always estimated using the whole images. Thus, it also retains a global meaning, which is desirable feature, because it also shows us the quality of global alignment.

2.1. Generalization of entropy based similarity measure

Observing Eq. (7) it is easy to see its close relation to *log likelihood* similarity measure¹⁴. The only difference is in the way how the probability $p(\cdot)$ is derived. In our case, it is estimated from the image pair, while in case of likelihood similarity measure the probability $p(\cdot)$ is a prior information. It is also possible to use a combination of both¹¹. The use of prior information improves joint distribution by making it more similar to the expected joint distribution of correctly registered images. That prevents images to be misregistered because of low initial image correspondence and reduces the number of required registration steps.

Our registration process tends to move each voxel of image B , which is being non-rigidly deformed, in direction of largest improvement of local (voxel) similarity. As image B intensities are known and do not change, it is reasonable to replace joint intensity probability $p_{\mathbf{i}} = p(\mathbf{i}) = p(i_A, i_B)$ with conditional probability $p(i_A|i_B)$ ⁵. It is also possible to use other types of probability, e.g. probability of intensity difference $p(i_A - i_B)$ in order to improve statistical significance¹².

Finally, it is possible to replace the log function with any other function $f(p)$ that meets the requirement that the first derivative of function $p \cdot f(p)$ is strictly monotonically increasing or decreasing. If it is increasing/decreasing, higher/lower value means better correspondence. Linear function $f(p) = p$ is often used and similarity measure is in this case often called *energy similarity measure*, because its global version can be computed as the sum of squared histogram values⁹.

Based on observations mentioned above the local similarity measure derived from entropy, can be more generally written as

$$S^v = f(p); \quad S^r = \frac{1}{N_r} \sum S^v = \frac{1}{N_r} \sum f(p), \quad (9)$$

where S^v stands for the voxel similarity, S^r is the similarity of region r , and $f(p)$ can be any function that meets the requirement mentioned above. Probability p may be derived from image pair p_{image} , it may be given in advance p_{prior} , or in the most general case, it can be defined as a weighted sum of both:

$$p = (1 - \lambda) \cdot p_{image} + \lambda \cdot p_{prior} \quad ; \quad \lambda \in [0, 1] \quad (10)$$

where λ is a weighting parameter.

3. JOINT DISTRIBUTIONS

As seen in the previous section, local multimodal similarity measures are closely related to the joint intensity distribution. To better understand these measures it is necessary to know how the degree of misregistration reflects in joint intensity distribution¹⁵.

Let us imagine we have two simple images A and B representing the same object. Let each image consist of only two intensity values (i_{1A} and i_{2A} for image A , and i_{1B} and i_{2B} for image B , where i_{1A} corresponds to i_{1B} and i_{2A} corresponds to i_{2B}). Similarity between two voxels is according to the described similarity measures related to the joint probability of intensity pair, such that higher value means better correspondence. When images are correctly registered, the joint distribution consists of only two extrema that appear at intensity pairs (i_{1A}, i_{1B}) and (i_{2A}, i_{2B}) as the intensity regions perfectly overlap. When images do not overlap exactly there are some regions with intensity pairs (i_{1A}, i_{2B}) and/or (i_{2A}, i_{1B}) , as shown in Fig. 2. The joint probability of mismatched intensity pairs is related to the degree of misregistration where worse registration means higher probability. This also means that similarity function based on probability estimation for mismatched regions increases by degree of misregistration. When mismatched regions are bigger than correctly matched ones, the estimated similarity of mismatches is higher than similarity for correct match and registration in this case cannot be successful. Due to this fact it is advantageous to use prior information, from which approximate joint distribution of correctly registered images is known in advance, and therefore the phenomenon mentioned above shall not appear.

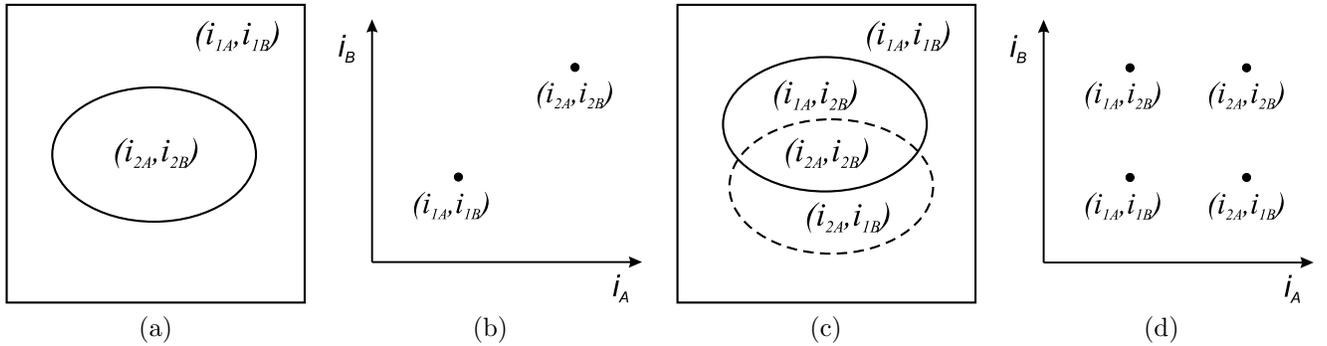


Figure 2. Joint distributions of registered images (a,b) and misregistered images (c,d).

In real images each tissue is not represented by a single intensity pair, but by a range of intensity pairs, forming an intensity class, see Fig. 3. Thus, in case of registered images the number of intensity classes equals the number of biological tissues K . When registering two images, their joint distribution consists of M joint distribution classes C , where each class represents a tissue type pair and all together form a set $\mathcal{C} = \{C_m\}; m = 1 \dots M$. Among them there are at most K single-tissue classes that correspond to the same tissue type in both images and form a subset $\mathcal{C}_S \subset \mathcal{C}$. All other classes are mixed-tissue classes and correspond to different tissue types in different images. During the registration the amplitude of mixed-tissue classes is expected to decrease until they finally disappear. Joint distribution then contains only single-tissue classes. If there are no intensity inhomogeneities each class can be modeled by two dimensional Gaussian function with mean value μ_m , amplitude A_m and covariance matrix Σ_m . Class parameters A_m , and Σ_m depend on the amount of image noise, where increase of noise reflects to increased class widths and decreased amplitudes. The similarity estimation is getting worse, as the probability of each intensity pair depends only on the number of its occurrences. Actually, image registration does not register biological features, but their intensities. Each tissue type is represented by a whole class of intensity pairs, which are treated independently. If their close relations within the same tissue type were taken into account, they could improve the registration.

Because of insufficient image resolution and image blurring, some voxels belong to more than one tissue type. The intensity value i_v of such partial volume voxel v is a weighted sum of class intensities i_k , where weights t_{vk} are portions of tissues $k = 1 \dots K$ that exist at that location. Partial volume voxels are in joint distribution positioned on lines between the classes, see Fig. 3. Their contribution to registration is low because of their low probability and thus low estimated similarity. Matching of these voxels is not reliable, although edges are most information-rich parts of the images. Image processing operations, e.g. linear interpolation and linear image filtering, increase the number of such voxels and therefore worsen the registration. These voxels can be correctly matched only when knowing which classes exist in each voxel and what their portions are.

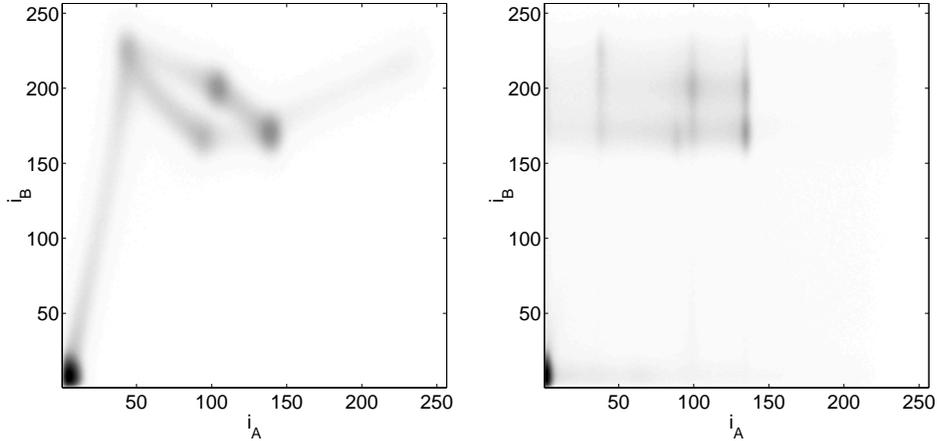


Figure 3. Expected joint distributions of real medical images without intensity inhomogeneities. (Left registered, Right misregistered).

4. SIMILARITY MEASURES BASED ON SEGMENTATION

In order to deal with the problems of image noise and partial volume voxels it is necessary to know which intensity pairs form each individual class and what portion of each partial volume voxel belongs to each class. This can be estimated by performing image segmentation¹⁶. When registering pre-segmented images similarity measure is perfectly defined by being 1 for the same tissue type on both images and 0 otherwise. Considering partial volume voxels, similarity measure of a voxel can be defined as a sum of corresponding voxel portions:

$$S^v = \sum_k \min(t_{kA}, t_{kB}) \quad ; \quad k = 1 \dots K, \quad (11)$$

where t_{kA} and t_{kB} are portions of tissue type k that exist at corresponding voxels in image A and B , respectively. Unfortunately, image segmentation is not trivial task, either. From the segmentation point of view, images can be segmented by registering them to a pre-segmented template. The image registration process is therefore dual to the segmentation. Both of them work better when using the others results. We anticipate that by performing both processes simultaneously it is possible to improve registration and segmentation results.

Based on observations mentioned above we propose a segmentation-based local similarity measure to improve registration results. The basic idea is to estimate the probability that certain intensity pair \mathbf{i} represents the correctly matched tissues. Note that joint distribution of correctly registered images consists of only single-tissue classes $C_k \in \mathcal{C}_S$, $k = 1 \dots K$, which correspond to image regions with correctly matched tissues. When images are not registered, there are also some mixed-tissue classes that represent different tissues at the same position in different images. Firstly, probabilities $p(C_m|\mathbf{i})$ of intensity pair \mathbf{i} belonging to certain class C_m , for all classes $m = 1 \dots M$, have to be estimated. Furthermore it is necessary to know which of the classes are single-tissue classes representing the correct match. Without prior information this is impossible to know for sure. Nevertheless, we can estimate the probability $p(\mathcal{C}_S|C_m)$ that C_m is a single-tissue class. The sum $\sum p(\mathcal{C}_S|C_m) \cdot p(C_m|\mathbf{i})$ over all classes then indicates the probability $p(\mathcal{C}_S|\mathbf{i})$ of \mathbf{i} belonging to one of the single-tissue classes. This is exactly the probability we are trying to estimate, since it indicates the similarity of intensities forming an intensity pair \mathbf{i} .

4.1. Implementation of segmentation based local similarity measure

Let us assume that each class C_m ; $m = 1..M$ can be modeled by a two-dimensional Gaussian function $p(\mathbf{i}, C_m)$ with mean value μ_m and covariance matrix Σ_m . Joint distribution $p(\mathbf{i})$ can be approximated by the following sum:

$$p(\mathbf{i}) = \sum_l p(\mathbf{i}, C_l) = \sum_l p(\mathbf{i}|C_l) \cdot p(C_l). \quad (12)$$

It is expected that classes are far enough from each other to achieve dominance of class C_m in its neighborhood \mathcal{O}_m , such that contributions of all other classes $p(\mathbf{i}, C_l)$; $l = 1..M, l \neq m$ can be neglected. Therefore, number of classes M , their mean values μ_m and amplitudes A_m can be estimated by extensive search for maxima in joint intensity distribution. When maximum is found, its position is used as a class mean value μ , while value itself is used as a class amplitude A . Probability of intensity pairs in \mathcal{O}_m can be approximated by

$$p(\mathbf{i})|_{\mathbf{i} \in \mathcal{O}_m} \approx p(\mathbf{i}, C_m) = A_m \exp\left(-\frac{1}{2}(\mathbf{i} - \mu_m)^T \Sigma_m^{-1} (\mathbf{i} - \mu_m)\right); \quad m = 1 \dots M. \quad (13)$$

By taking a log of (13) we get

$$2 \ln \left(\frac{A_m}{p(\mathbf{i})} \right) = (\mathbf{i} - \mu_m)^T \Sigma_m^{-1} (\mathbf{i} - \mu_m) \quad ; \quad \mathbf{i} \in \mathcal{O}_m \quad (14)$$

$$\Sigma_m^{-1} = \mathbf{U} = \begin{bmatrix} u_{11} & u_{12} \\ u_{12} & u_{22} \end{bmatrix} \quad (15)$$

$$2 \ln \left(\frac{A_m}{p(\mathbf{i})} \right) = u_{11}(i_A - \mu_{mA})^2 + 2u_{12}(i_A - \mu_{mA})(i_B - \mu_{mB}) + u_{22}(i_B - \mu_{mB})^2 \quad ; \quad \mathbf{i} \in \mathcal{O}_m, \quad (16)$$

which can be solved for \mathbf{U} using least squares method for all intensity pairs \mathbf{i} in the neighborhood \mathcal{O}_m . The covariance matrices $\Sigma_m = \mathbf{U}^{-1}$ can be then used to estimate the class a priori probabilities $p(C_m)$.

$$p(C_m) = \int p(\mathbf{i}, C_m) d\mathbf{i} = \int A_m \exp\left(-\frac{1}{2}(\mathbf{i} - \mu_m)^T \Sigma_m^{-1} (\mathbf{i} - \mu_m)\right) d\mathbf{i} \quad ; \quad m = 1..M \quad (17)$$

$$p(C_m) = 2\pi A_m |\Sigma_m| \quad ; \quad m = 1..M \quad (18)$$

Theoretically the sum of all a priori probabilities C_m , $m = 1..M$, should be 1. In reality this is seldom the case even if all class parameters are estimated absolutely correct, due to the fact that some intensity pairs with low joint probability do not belong to any of the estimated classes. The majority of such intensity pairs represent partial volume voxels, which are correctly matched only when images are registered with subvoxel accuracy. Currently we are not dealing with partial volume voxels explicitly. Therefore, to prevent incorrect matching we introduce an additional class C_0 with small probability

$$p(\mathbf{i}, C_0) = \varepsilon \quad (19)$$

for each intensity pair. We set ε to the probability of a single image voxel in joint intensity distribution.

A posterior probability of class C_m , showing the chance of intensity pair to belong to a certain class C_m is according to Bayes rule the following:

$$p(C_m|\mathbf{i}) = \frac{p(\mathbf{i}, C_m)}{\sum_{l=0}^M p(\mathbf{i}, C_l)}. \quad (20)$$

Furthermore, to derive the similarity measure it is necessary to estimate the probability $p(\mathcal{C}_S|C_m)$ that a certain class C_m ; $m = 1..M$ is a single-tissue class. Let us assume that each tissue type has unique intensity representation μ . Therefore, among all maxima positioned at the same intensity of image A (or image B) only one can belong to the set of single-tissue classes. All the other maxima represent mixed-tissue classes, see Fig. 2. Let subset $\mathcal{C}_{\mu A} \subset \mathcal{C}$

consists of all classes with the same mean value μ_A . Probability $p_A(\mathcal{C}_S|C_m)$ of class C_m being a single-tissue class according to intensity of image A can be estimated as

$$p_A(\mathcal{C}_S|C_m) = \frac{p(C_m)}{\sum_{C_l \in \mathcal{C}_{\mu_A}} p(C_l)}. \quad (21)$$

Such estimation of probability that class is a single-tissue class is not necessarily sufficient, as all classes with higher probability may belong to the same subset \mathcal{C}_{μ_B} , which consist of all classes with the same value μ_B . Such situations can be prevented by using probability $p_B(\mathcal{C}_S|C_m)$ that class C_m is a single-tissue class according to intensity of image B ,

$$p_B(\mathcal{C}_S|C_m) = \frac{p(C_m)}{\sum_{C_l \in \mathcal{C}_{\mu_B}} p(C_l)}. \quad (22)$$

The final estimate of probability that class C_m is a single-tissue class is then

$$p(\mathcal{C}_S|C_m) = p_A(\mathcal{C}_S|C_m) \cdot p_B(\mathcal{C}_S|C_m). \quad (23)$$

The sum of $p(\mathcal{C}_S|C_m)$ over all m indicates the number of single-tissue classes. When there are ω single-tissue classes with the same μ_A or μ_B our presumption is incorrect. The estimated probabilities $p_A(\mathcal{C}_S|C_m)$ or $p_B(\mathcal{C}_S|C_m)$ are ω times too small. Nevertheless, as the relations between the probabilities are still correct, registration should still be successful but it may require some more registration steps.

Finally, local similarity measure based on image segmentation can be estimated as

$$S^v = p(\mathcal{C}_S|\mathbf{i}) = \sum_m p(C_m|\mathbf{i}) \cdot p(\mathcal{C}_S|C_m). \quad (24)$$

5. COMPARISON RESULTS

A reasonable approach to compare global similarity measures is to apply them on various transformations and to evaluate their properties, such as number of local extrema, their smoothness, position of global maximum, capture range, etc. However, comparison of local similarity measures turns to be more problematic, as any transformation of a single voxel region makes a drastic change in region overlap. Furthermore, the results do not depend only on similarity measure used but also on spatial model, e.g. elasticity. We argue that voxel based similarities cannot be tested in isolation without considering spatial model, and it appears that the only meaningful way to compare them is their treatment in the context of complete registration systems.

To show the relative performance of our local similarity measures we used two spatially aligned three-dimensional medical images. The first one, image A , was used as a reference, while the second one, image B , was transformed with a known transformation T_0 and then registered back to image A , see Fig. 4. The registration result was evaluated using the RMS voxel displacement error e , calculated from differences between voxel displacements achieved by the registration T_R and displacements derived by inverse of known transformation T_0^{-1} . Voxels representing a background were not included and were removed by mask Ω , as every matching of the background can be treated as correct. N_Ω denotes number of voxels accepted by the mask Ω and \mathbf{x} are spatial coordinates.

$$e = \sqrt{\frac{1}{N_\Omega} \sum_{\mathbf{x} \in \Omega} (T_{0\mathbf{x}}^{-1} - T_{R\mathbf{x}})^2} \quad ; \quad \mathbf{x} = [x_1, x_2, x_3]^T \quad (25)$$

The similarity measures were tested using BrainWeb simulated images¹⁷ to satisfy the requirement of initial spatial image alignment. The comparison results are shown in Table 1. Column 1 contains the similarity measures used. MRI-T1 images were non-rigidly registered to reference MRI-PD images using the same multiresolution elastic registration approach but based on different local multimodal similarity measures for individual voxels. Tests were performed for normal images, images with 9% of noise and images with 40% of intensity inhomogeneities (shading), all without any prior information, and for normal images using prior information, derived from the initial images. There was a relatively small number of iterations used, (2 for original resolution, 7 for half resolution, and 10 for other

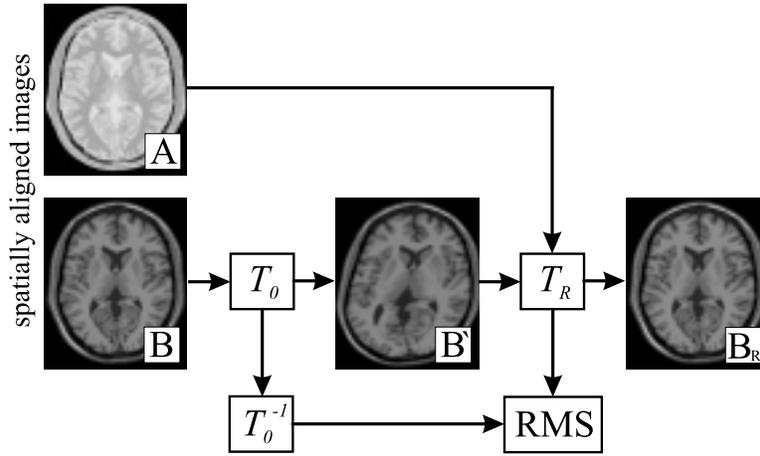


Figure 4. Evaluation scheme.

Table 1. Comparison results of local multimodal similarity measures.

S^v	normal	noise	shading	prior inf.
$p(\mathbf{i})$	6.7449	6.9147	7.4476	7.9025
$p(i_A i_B)$	2.6263	2.6614	3.7351	2.1834
$p(i_A - i_B)$	6.5745	6.6526	6.8558	7.0569
$\log(p(\mathbf{i}))$	2.5631	2.5754	2.8047	1.1952
$\log(p(i_A i_B))$	2.8207	2.8305	2.9997	1.3988
$\log(p(i_A - i_B))$	3.8948	4.0508	4.2280	3.2910
$p(\mathcal{C}_S \mathbf{i})$	1.2600	1.2537	2.1172	0.9445

three resolution levels). The initial transformation T_0 was composed of six Gaussian functions such that maximal initial displacement was 16.8 voxels and initial RMS error e was 6.9.

The results in Table 1 show that segmentation based measure $S^v = p(\mathcal{C}_S|\mathbf{i})$ performs best among all similarity measures. In all the cases it yields the best result. Relatively good results are also achieved when using the original entropy based measure $S^v = \log(p(\mathbf{i}))$ and its derivatives $S^v = p(i_A|i_B)$ and $S^v = \log(p(i_A|i_B))$. Measures based on probability of intensity difference $p(i_A - i_B)$ and measure $S^v = p(\mathbf{i})$ are shown not to be appropriate for estimation of local properties.

It is shown that all similarity measures have low sensitivity to noise. A sensitivity to intensity inhomogeneities (shading) is much higher and is the highest for segmentation based measure $S^v = p(\mathcal{C}_S|\mathbf{i})$, as expected. The reason is in Gaussian modeling of intensity classes. As classes in the presence of inhomogeneities do not comply to Gaussian distribution, covariance matrices Σ_m are estimated incorrectly. This leads to bad estimation of probabilities $p(C_m|\mathbf{i})$ and thus incorrect estimation of similarity.

As expected prior information in general improves the registration. The biggest improvement is shown for original entropy based measure $S^v = \log(p(\mathbf{i}))$ and its derivative $S^v = \log(p(i_A|i_B))$. Segmentation based similarity measure show relatively small improvement. The reason for this is that prior information derived from correctly registered images does not contain any mixed-tissue classes. Therefore, discrimination between single-tissue and mixed-tissue classes cannot add any improvement.

Comparison of estimated similarities derived from the same joint intensity distribution for two similarity measures that show good performance is presented in Fig. 5. Similarities derived by similarity measure $S^v = \log(p(\mathbf{i}))$ differ from joint distribution $p(\mathbf{i})$ only in the logarithmic function, which makes maxima wider and their amplitudes less different. The segmentation based measure $p(\mathcal{C}_S|\mathbf{i})$ shows additional improvement by reducing the similarities for mixed-tissue intensity pairs based on additional knowledge gained by image segmentation.

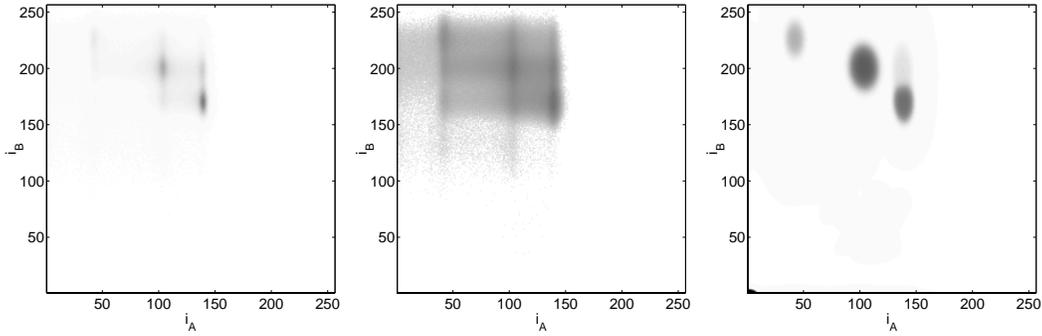


Figure 5. Joint distribution $p(\mathbf{i})$ (left) and estimated similarities: $S^V = \log(p(\mathbf{i}))$ (center), and $S^V = p(C_S|\mathbf{i})$ (right).

The by-product of segmentation based similarity measure is segmented image as follows. The whole image consists of single-tissue and mixed-tissue classes. All of them are present when images are misregistered, see Fig 6 (left). Registration reduces mixed-tissue classes and when images are correctly registered only the single-tissue classes remain, see Fig 6 (right). Each class then represent a single tissue type.

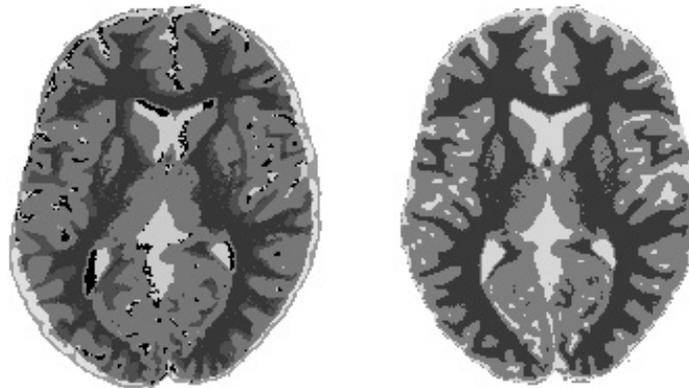


Figure 6. Result of the segmentation for misregistered (left) and registered images (right). Each class is represented by different gray value.

6. CONCLUSION

This paper describes various similarity measures for local multimodal image registration. They are all based on probability distribution determined from the whole image content or given in advance as prior information. As the same probability distribution is used for all image regions/voxels, these measures are all sensitive to intensity inhomogeneities.

Measures derived from joint entropy treat each intensity pair individually and do not consider relations between intensities within the same tissue type. This relationship is established by the segmentation based method that we have presented. Our method, however, does not solve the problem of partial volume voxels efficiently, as the segmentation method used lacks capability to detect such partial volume voxels. Consequently, images cannot be registered with sub-voxel accuracy.

Segmentation is usually based on presumption that pure single-tissue clusters conform to Gaussian distribution. This is not true when images are subjected to intensity inhomogeneities. Registration is in this case not necessarily successful. To alleviate this problem images should be preliminary corrected for intensity inhomogeneities¹⁸. It is

easier to perform such correction when images are already segmented. This indicates that all three mentioned tasks, registration, segmentation, and correction for intensity inhomogeneities, should be performed simultaneously.

Our plan for the future is to improve the segmentation based method to take into account partial volume voxels and thus enable subvoxel accuracy. Furthermore, we are planning to incorporate the correction for intensity inhomogeneities to improve the registration and segmentation of real images.

REFERENCES

1. A. W. Toga, ed., *Brain Warping*, Academic Press, San Diego, 1999.
2. H. Lester and S. R. Arridge, "A survey of hierarchical non-linear medical image registration," *Pattern Recognition* **32**(1), pp. 129–149, 1999.
3. M. Holden, D. L. G. Hill, E. R. E. Dent, J. M. Jarosz, T. C. S. Cox, T. Rohlfing, G. J., and D. J. Hawkes, "Voxel similarity measures for 3-d serial mr brain image registration," *IEEE Transactions on Medical Imaging* **19**, pp. 94–102, February 2000.
4. T. Gaens, F. Maes, D. Vandermeulen, and S. P., "Non-rigid multimodal image registration using mutual information," in *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, W. Wells, A. Colchester, and S. Delp, eds., no. 1496 in Lecture Notes in Computer Science, pp. 1099–1106, Springer-Verlag, (MIT, Cambridge, MA, USA), October 1998.
5. J. B. A. Maintz, H. W. Meijering, and M. A. Viergever, "General multimodal elastic registration based on mutual information," *Medical Imaging* **3338**, pp. 144–154, 1998.
6. R. Bajcsy and S. Kovačič, "Multiresolution elastic matching," *Computer Vision, Graphics and Image Processing* **46**, pp. 1–21, April 1989.
7. M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," *Springer Lecture Notes in Computer Science* **1131**, pp. 267–276, 1996.
8. P. Viola and W. Wells III, "Alignment by maximization of mutual information," in *Proceedings of the 5th International Conference on Computer Vision*, pp. 16–23, 1995.
9. T. M. Buzug, J. Weese, C. Fassnacht, and C. Lorenz, "Elastic matching based on motion vector fields obtained with a histogram based similarity measure for dsa-image correction," *Computer Assisted Radiology and Surgery*, pp. 139–144, 1997.
10. I. J. Taneja, "On generalized information measures and their applications," *Advances in Electronics and Electron Physics* **67**, pp. 327–413, 1989.
11. B. Likar and F. Pernus, "A hierarchical approach to elastic registration based on mutual information," in *Proceedings of the Workshop on Biomedical Image Registration - WBIR'99*, F. Pernus, S. Kovačič, H. S. Stiehl, and M. A. Viergever, eds., pp. 24–45, August 1999.
12. P. Rogelj and S. Kovacic, "Local similarity measures for multimodal image matching," in *Proceedings of the first International Workshop on Image and Signal Processing and Analysis – IWISPA 2000*, S. Loncaric, ed., pp. 81–86, University Computing Center, University of Zagreb, 2000.
13. J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information based image registration," *Computer Vision and Image Understanding* **77**(2), pp. 211–232, 2000.
14. M. Leventon and W. Crimson, "Multi-modal volume registration using joint intensity distributions," in *Proceedings of the 1st International Conference on Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, W. Wells, A. Colchester, and S. Delp, eds., no. 1496 in Lecture Notes in Computer Science, pp. 1057–1066, Springer-Verlag, (MIT, Cambridge, MA, USA), October 1998.
15. C. Studholme, H. D.L.G., and D. Hawkes, "Using voxel similarity as a measure of medical image registration proc.," in *Proceedings of the British machine vision conference – BMV'94*, E. Hancock, ed., pp. 235–244, 1994.
16. D. Laidlaw, K. Fleischer, and A. Barr, "Partial-volume bayesian classification of material mixtures in mr volume data using voxel histograms," *IEEE Transactions on Medical Imaging* **17**(1), pp. 74–86, 1998.
17. R. Kwan, A. Evans, and G. Pike, "An extensible mri simulator for post-processing evaluation," in *Visualization in Biomedical Computing (VBC'96)*, Lecture Notes in Computer Science, pp. 135–140, Springer-Verlag, 1996.
18. B. Likar, J. B. A. Maintz, M. Viergever, and F. Pernus, "Retrospective shading correction based on entropy minimisation," *Journal of Microscopy* **197**(3), pp. 285–295, 2000.