

Manifold Learning for Multi-Modal Image Registration

Christian Wachinger

wachinge@in.tum.de

Nassir Navab

navab@in.tum.de

Computer Aided Medical Procedures

Technische Universität München

Munich, Germany

Abstract

The standard approach to multi-modal registration is to apply sophisticated similarity metrics such as mutual information. The disadvantage of these measures, in contrast to simple L1 or L2 norm, is the increased computational complexity and consequently the prolongation of the registration time. An alternative approach, which has so far not yet gained much attention in the literature, is to find image representations, so called structural representations, that allow for the direct application of L1 and L2 norm. Recently, entropy images [24] were proposed as a simple structural representation of images for multi-modal registration. In this article, we propose the application of manifold learning, more precisely Laplacian eigenmaps, to learn the structural representation. It has the theoretical advantage of presenting an optimal approximation to one of the criteria for a structural description. Laplacian eigenmaps search for similar patches in high-dimensional patch space and embed the manifold in a low-dimensional space under preservation of locality. This can be interpreted as the identification of internal similarities in images. In our experiments, we show that the internal similarity across images is comparable and notice very good registration results for the new structural representation.

1 Introduction

A challenge of registration is to align images that contain significant intensity variations. These variations can originate from a multitude of sources, such as illumination changes in optical images, field inhomogeneities in magnetic resonance (MR) images, and, simply, different imaging modalities. The common approach in intensity-based registration is to integrate similarity metrics that model those intensity variations, assuming a functional or statistical intensity relationship, instead of an identical one. Feature-based registration approaches, on the other hand, apply methods that are robust to intensity variations for keypoint extraction and description. An example would be the calculation of histograms of image gradients, as it is done for SIFT [11] and GLOH [14]. The registration of images from different modalities is, however, affected by more substantial intensity variations, making those approaches not directly applicable.

The approach that we present in this article is a mixture of feature- and intensity-based registration. The idea is to calculate dense descriptors that represent the structural information of image patches. The structural information of a patch is only dependent on the

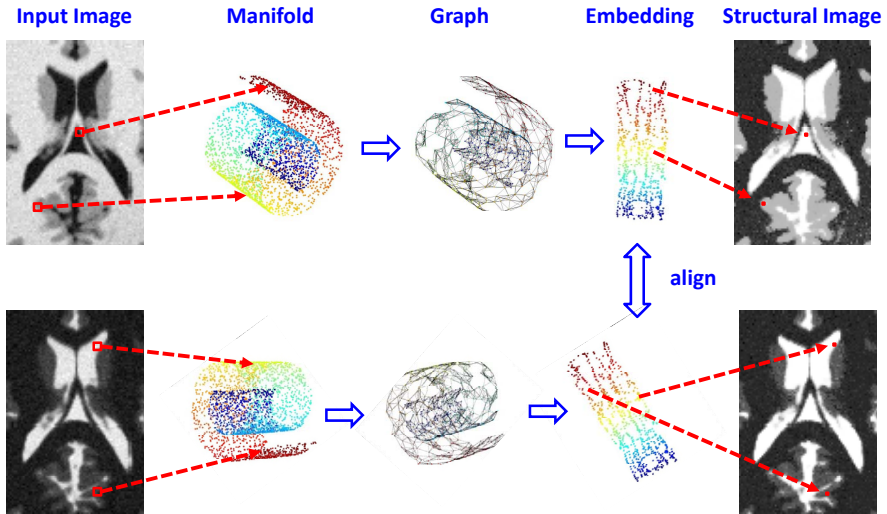


Figure 1: Structural representation with Laplacian eigenmaps. Patches of images lie on a manifold in high-dimensional patch space. The manifold is approximated by the neighborhood graph. The low-dimensional embedding is calculated with the graph Laplacian. Embeddings from different modalities have to be aligned to obtain the final representation.

structures in the patch, not the intensity values with which the structure is displayed. The input images are replaced by the dense set of descriptors, on which a regular intensity-based registration is performed. This guarantees a seamless integration into existing registration frameworks. The advantage of the structural representation is that the standard L1 or L2 norm can be used for the registration of the structural images. These are computationally less expensive and enable therefore a faster registration. This is even more important for groupwise registration, where the influence becomes quadratic, as shown in [26]. Moreover, a more efficient optimization scheme, the efficient second order minimization (ESM, [4, 25]), can be applied for multi-modal registration [26]. Finally, L1 and L2 norms are much better suited for parallelization than complex multi-modal similarity measures, which is important with respect to transferring the calculation to the GPU for further speed-up.

In this article, we propose the application of manifold learning for finding an optimal structural representation of images. Laplacian eigenmaps present an optimal solution to one of the requirements for a structural representation for image registration. This is the preservation of locality, meaning that patches that are close in high-dimensional patch space are mapped to a close structural representation. The requirement that the structural representation of images is similar across modalities is guaranteed by their comparable internal similarities. This was previously exploited in a novel framework for multi-modal registration based on internal similarity by Penney *et al.* [16]. The two approaches are similar because small image patches are compared to find local similarities inside images. The consecutive utilization of this information is, however, entirely different. While in [16] a few internal similarity structures are identified and then transferred to the image from the other modality, we use all the patch information to build a neighborhood graph, approximating the manifold embedded in high dimensions. Subsequently, the graph Laplacian is calculated to find an optimal mapping to low-dimensional space. Since the embedding in low-dimensional space

is arbitrary, as long as it preserves the locality, we have to align embeddings from different modalities with an affine registration. This finally leads to the structural representation that is used in the intensity-based registration, as schematically illustrated in figure 1.

2 Related Work

There are two groups of methods for registering multi-modal images. The first one consists of methods that apply advanced similarity measures, such as mutual information (MI) [5, 30] or the correlation coefficient (CR) [19]. A nice overview of mutual information based techniques is provided in [18]. The second group of techniques, which is more related to our work, consists of methods that reduce a multi-modal to a mono-modal registration. These can again be differentiated into two classes. The first ones try to simulate one modality from the other. Examples are X-Ray to CT (Computed Tomography) registration with the creation of digitally reconstructed radiographs [15] and ultrasound to CT registration with the simulation of ultrasound images [14]. The second group consists of methods that transfer both images into a third, artificial modality. Examples are (i) the application of morphological tools [13], (ii) recoloring images depending on the variances of the image regions [1], (iii) the usage of edge- and ridge-information [12], (iv) cross-correlating gradient directions [2], and (v) the creation of shadow-invariant optical images [17]. In our case, we are interested in a general structural representation, so that the application specific approaches are not applicable. Further, (ii) and (iv) use cross-correlation for the comparison, indicating that the description is not truly identical. The morphological approach [13] mainly leads to a surface extraction, and although it has gray values instead of only binary values, much internal information is lost. Finally, edge, ridge, and gradient estimation is problematic for points where more than two regions are meeting, *e.g.* T-junctions, as discussed in [26].

Further related is the already mentioned work of Penney *et al.* [26], which builds upon the internal similarity in images. Shechtman *et al.* [22] also consider the internal similarity in images, however only locally, for constructing image descriptors. Another approach that we want to refer to is the DAISY descriptor [24], which also replaces the original images by a dense descriptor. Finally, Lee *et al.* [11] worked on learning similarity measures for multi-modal registration. For the supervised learning they use max-margin structured output learning. The approach seems related because they also consider learning, however, we are interested in finding structural representations, instead of learning the similarity measure, and our unsupervised learning with Laplacian eigenmaps does not need any training. The training is not simple because correctly aligned images from the target modalities must be available. It is even more challenging for magnetic resonance (MR) images, which vary significantly in their appearance for different echo and repetition times (TE/TR).

3 Structural Image Registration

Consider two images $I, J : \Omega \rightarrow \mathbf{I}$ defined on the image grid Ω with intensity values $\mathbf{I} = \{1, \dots, \alpha\}$. The registration is formulated as

$$\hat{T} = \arg \max_{T \in \mathbf{T}} \mathbf{S}(I, J(T)), \quad (1)$$

with the space of transformations \mathbf{T} and the similarity measure \mathbf{S} . For images with structures being depicted with the same intensity values, so $I(x) = J(\hat{T}(x))$ for $x \in \Omega$, the L1 or L2 norm

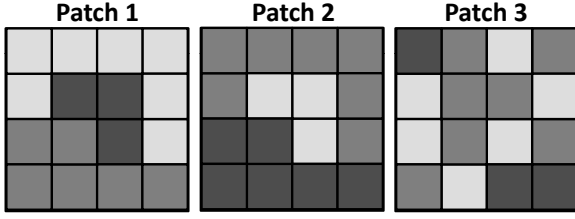


Figure 2: Three different patches. Patches 1 and 2 are structurally equivalent, Patch 3 is not.

are a good choice for \mathbf{S} . For more complex intensity relationships between the images, such as affine, functional, or statistical ones, typical choices for \mathbf{S} are the correlation coefficient, correlation ratio, and mutual information, respectively. These are, however, more computationally expensive. Our goal is therefore to find structural representations that replace I and J in the optimization of equation (1) and for which we can set \mathbf{S} to the L1 or L2 norm.

3.1 Original Structural Representation [26]

In order to achieve a structural representation of an image, we break the problem down to the simpler problem of finding a structural representation of an image patch. We denote patches that are defined on the local neighborhood \mathbf{N}_x around x as $P_x : \mathbf{N}_x \rightarrow \mathbf{I}$. Our objective is to find a function $f : P_x \mapsto D_x$ that assigns each patch a descriptor D_x so that the descriptor captures the structural information of the patch. Since we calculate a descriptor for each location x , we obtain a new image with the original intensities replaced by the descriptors.

We define two patches P_x, P_y to be structurally equivalent $P_x \sim P_y$, if there exists a bijective function $g : \mathbf{I} \rightarrow \mathbf{I}$ such that $\forall z \in \mathbf{N}_x : P_x(z) = g(P_y(z))$. For an illustration, the first two patches in figure 2 are structurally equivalent, in contrast to the third one. According to [26], two properties a function f has to fulfill are

(P1) Locality preservation:

$$\|P_x - P_y\| < \varepsilon \implies \|f(P_x) - f(P_y)\| < \varepsilon' \quad (2)$$

(P2) Structural equivalence:

$$P_x \sim P_y \iff f(P_x) = f(P_y) \quad (3)$$

with reasonable ε and ε' depending on the chosen norm. The motivation behind the first property is to ensure that similar patches are mapped to similar descriptors, which is important for the robustness to noise and the capture range of the registration. The second property states that descriptors are equivalent, if and only if, the patches are structurally equivalent. This ensures, on the one hand, the desired structural representation, and on the other hand, it avoids trivial solutions, such as mappings to constant values, and it improves the discrimination.

3.2 Revised Structural Representation

The structural representation in [46] is not precisely modeling the requirements for multi-modal registration. Subsequently, we first state two revised properties and later on explain their advantages. To this end we consider patches Q_x, Q_y to be part of image I , and R_x to be a patch of image J . Moreover, we introduce a function for each of the modalities, in the following denoted by f and f' .

(Q1) Locality preservation:

$$\|Q_x - Q_y\| < \varepsilon \implies \|f(Q_x) - f(Q_y)\| < \varepsilon' \quad (4)$$

(Q2) Structural equivalence:

$$Q_x \sim R_x \iff f(Q_x) = f'(R_x) \quad (5)$$

(Q1) in comparison to (P1) does no longer compare patches from both modalities, but restricts the comparison to patches within one image. This is better because the calculation of the norm $\|\cdot\|$ between images from different modalities is not well defined. In fact, due to the depiction of structures with different intensities, patches from multi-modal images may be similar although they do not depict the same structures. This could lead to further local optima and consequently mis-registrations.

We are as well more specific in the formulation of the structural equivalence (Q2). We require the structural equivalence only for patches of different images. The inclusion of patches from the same image, as it is done in (P2), is not meaningful, since no re-mapping of intensity values is required in the same image. It could in fact lead to ambiguities. The presented, more precise, modeling is no longer satisfiable by a global function f . We consequently have to define a local function for each modality, indicated with f and f' . In the next section, we explain the application of manifold learning for calculating the desired structural representation.

4 Manifold Learning

Manifold learning is an approach applied for non-linear dimensionality reduction and data representation. The task of dimensionality reduction is to find the underlying structure in a large set of points embedded in a high-dimensional space and to map these points to a low-dimensional space preserving the structure. Manifold learning has recently gained much attention to assist image processing tasks such as segmentation [48], registration [8, 47], tracking [9, 49], recognition [9, 47], and computational anatomy [8]. Common techniques for manifold learning are Isomap [43], local linear embedding [47], and Laplacian eigenmaps [9]. We focus on Laplacian eigenmaps because the technique is well founded on mathematical concepts (Laplace Beltrami operator) and computationally efficient.

4.1 Laplacian Eigenmaps

Laplacian eigenmaps [9] build upon the construction of a neighborhood graph that approximates the manifold, on which the data points are lying on. Subsequently, the graph Laplacian is applied to calculate a low-dimensional representation of the data that preserves locality.

Considering k points $\mathbf{a}_1, \dots, \mathbf{a}_k$ in \mathbb{R}^N lying on a manifold \mathcal{M} , we want to find a set of corresponding points $\mathbf{b}_1, \dots, \mathbf{b}_k$ in the low-dimensional space \mathbb{R}^n ($n \ll N$). We assume a twice differentiable function $m : \mathcal{M} \rightarrow \mathbb{R}^n, \mathbf{a}_i \mapsto \mathbf{b}_i$. It is shown in [9] that the following holds for the relationship between the distances on the manifold and embedding space

$$\|m(\mathbf{a}_i) - m(\mathbf{a}_j)\| \leq \text{dist}_{\mathcal{M}}(\mathbf{a}_i, \mathbf{a}_j) \cdot \|\nabla m(\mathbf{a}_i)\| + o(\text{dist}_{\mathcal{M}}(\mathbf{a}_i, \mathbf{a}_j)), \quad (6)$$

with $\mathbf{a}_i, \mathbf{a}_j \in \mathcal{M}$. We see that $\|\nabla m(\mathbf{a}_i)\|$ indicates how close nearby points are mapped. Consequently, a map that best preserves locality on average is found with the following minimization

$$\arg \min_{\|m\|_{L^2(\mathcal{M})}=1} \int_{\mathcal{M}} \|\nabla m(\mathbf{a}_i)\|^2 d\mathbf{a}_i. \quad (7)$$

It is this optimization problem, for which Laplacian eigenmaps provide an optimal solution, by calculating eigenfunctions of the Laplace Beltrami operator.

This optimally locality preserving embedding is exactly what was required for the structural representation in (Q1). We only have to identify the points $\mathbf{a}_i, \mathbf{a}_j$ with the patches Q_x, Q_y . The mapping m is therefore a suitable candidate for the function that provides the structural representation f . For the second property (Q2), we consider manifolds \mathcal{M} and \mathcal{M}' for two different modalities with patches $Q_x \in \mathcal{M}$ and $R_x \in \mathcal{M}'$. Since the intensity, with which objects are depicted in the images, varies with the modality, the two manifolds are not directly comparable. Applying, however, the assumption that the internal similarity of both modalities is equivalent, as in [16], we conclude that the structure or shape of both manifolds is similar. Since Laplacian eigenmaps preserve locality when embedding the manifold in a low-dimensional space, this structure is preserved in low dimensions. We could then directly use the coordinates of \mathbf{b}_i as descriptor for the corresponding location D_x . This is, however, not possible because the embedding of the structure in low-dimensional space is arbitrary, as long as it preserves the locality. The embeddings of both manifolds \mathcal{M} and \mathcal{M}' are therefore only similar when correcting for rotation, translation, and scale. Consequently, an affine registration of the point sets \mathbf{b}_i and \mathbf{b}'_i has to be performed. The coordinates of the registered embeddings finally provide the structural descriptors.

We conclude that m fulfills the revised properties (Q1) and (Q2) for a structural representation. It has therefore theoretical advantages in comparison to entropy images, since they do not fulfill the revised properties, and further, the preservation of locality is optimal for m .

4.2 Application of Laplacian Eigenmaps

In this section, we describe the application of Laplacian eigenmaps to our problem in more details. We consider one dimension of the ambient space for each image pixel of the patches. We use patches of size 15×15 , so that $N = 225$. This size proved to be a good compromise between too small patches that do not contain enough structural information, and too large patches that contradict with the required locality and further lead to a higher computational burden.

We construct a graph with a node for each point Q_x and with edges connecting neighboring nodes. The neighborhood can be defined with an δ -neighborhood around each point, so Q_y is in the δ -neighborhood of Q_x if $\|Q_x - Q_y\|^2 < \delta$, with $\|\cdot\|$ the Euclidean norm. Although this is geometrically motivated, a disadvantage is the selection of the parameter δ . In our implementation, we search instead for the $l = 500$ nearest neighbors and add edges

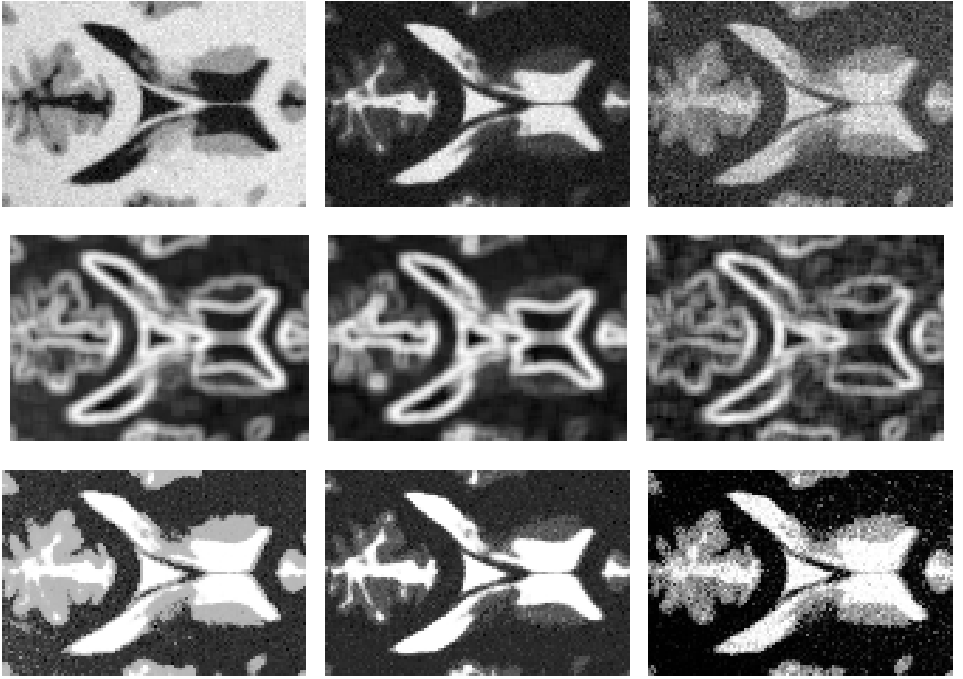


Figure 3: T1 (left), T2 (middle), and PD (right) images. First line: original images. Second line: entropy images. Third line: Laplacian images.

between them in the adjacency graph. Further, heat kernel-based weights are assigned to the edges with $w_{xy} = e^{-\|Q_x - Q_y\|^2 / (2 \cdot \sigma^2)}$ and σ^2 the variance.

We select $n = 1$ as dimension for the low-dimensional space. The reasons for not increasing this value are, first, that we obtain good results, and second, that for $n > 1$ we would have to store a vector in each pixel position instead of a scalar. This increases the computational complexity and memory consumption of the registration, and moreover, makes the visualization more challenging. Additionally, the low dimensionality facilitates the affine alignment to a correction of scale and shift. In fact, a robust normalization accounting for outliers and flipping, makes the point-based registration in this case even superfluous.

5 Experiments

We perform experiments on T1-, T2, and PD-weighted MR images from the BrainWeb database¹. We work with BrainWeb images containing 3% noise and 20% intensity non-uniformity, in order to achieve realistic results. The ground truth alignment for the images is given. In figure 3 we show the original images, the entropy images, and the images from the Laplacian eigenmaps, in the following shortly referred to as Laplacian images. We can clearly observe the different nature of the entropy and Laplacian images. Entropy images resemble gradient images, pronouncing boundaries and changes in the images. The Laplacian

¹<http://www.bic.mni.mcgill.ca/brainweb/>

Table 1: Registration study. Errors for translation in mm and rotation in degree.

| Datasets | Similarity | Rotation | Trans X | Trans Y | RMS |
|----------|------------|----------|---------|---------|--------------|
| T1-T2 | L2 | 4.879 | 9.019 | 6.471 | 7.000 |
| | MI | 2.325 | 3.768 | 5.226 | 3.954 |
| | Entropy | 2.084 | 4.539 | 5.231 | 4.180 |
| | Laplacian | 2.584 | 2.061 | 2.1680 | 2.271 |
| T1-PD | L2 | 2.760 | 6.422 | 5.755 | 5.227 |
| | MI | 2.304 | 4.138 | 4.907 | 3.937 |
| | Entropy | 2.283 | 4.782 | 4.750 | 4.108 |
| | Laplacian | 1.750 | 3.007 | 1.929 | 2.297 |
| T2-PD | L2 | 1.784 | 2.947 | 2.916 | 3.942 |
| | MI | 2.161 | 4.628 | 3.812 | 3.680 |
| | Entropy | 1.723 | 4.296 | 3.780 | 3.450 |
| | Laplacian | 1.171 | 2.350 | 1.984 | 1.900 |

images, however, look like the original images, but with a different coloring. We can further observe that the assumption of comparable internal similarities in the images is justified, because the appearance of the Laplacian images across the modalities is very similar.

In order to quantify the promising visual appearance for image registration, we show surface plots of the similarity measures for translation in figure 4. We compare the usage of the L2-norm on the original images, MI on the original images, and L2-norm on entropy and Laplacian images for all combinations of multi-modal alignment. The maxima indicate the best alignment. MI shows a very sharp peak at the correct position, but seems to have a limited capture range. Entropy images also indicate the correct position, but the cost functions contain several local maxima. We observe the cost function with the largest capture range for the Laplacian images. In figure 5, we show similarity plots for rotation. The results are comparable to those from the surface plots. We note the limited capture range for MI and the local maxima for entropy images. The Laplacian images lead to a wide and smooth peak, as desired.

Although similarity plots give a good intuition about the performance of different similarity measures, it is only a registration study that shows the final quality. We perform a registration study for all multi-modal image combinations. The random starting position deviates up to ± 15 mm in translation and $\pm 10^\circ$ in rotation from the correct pose. We show the average absolute error for translation and rotation, together with the overall root mean squared error (RMS), for 100 registration runs for each configuration in table 1. We weight 1 mm equal to 1° to quantify translational and angular displacement from the ground truth in one single value. We see that the positive impression of the Laplacian images from the similarity plots is confirmed by the registration results. We obtain a significantly lower error in comparison to MI and entropy images. The performance of entropy images and MI is comparable. Our experiments therefore confirm the theoretical advantages of Laplacian images in comparison to entropy images in practice.

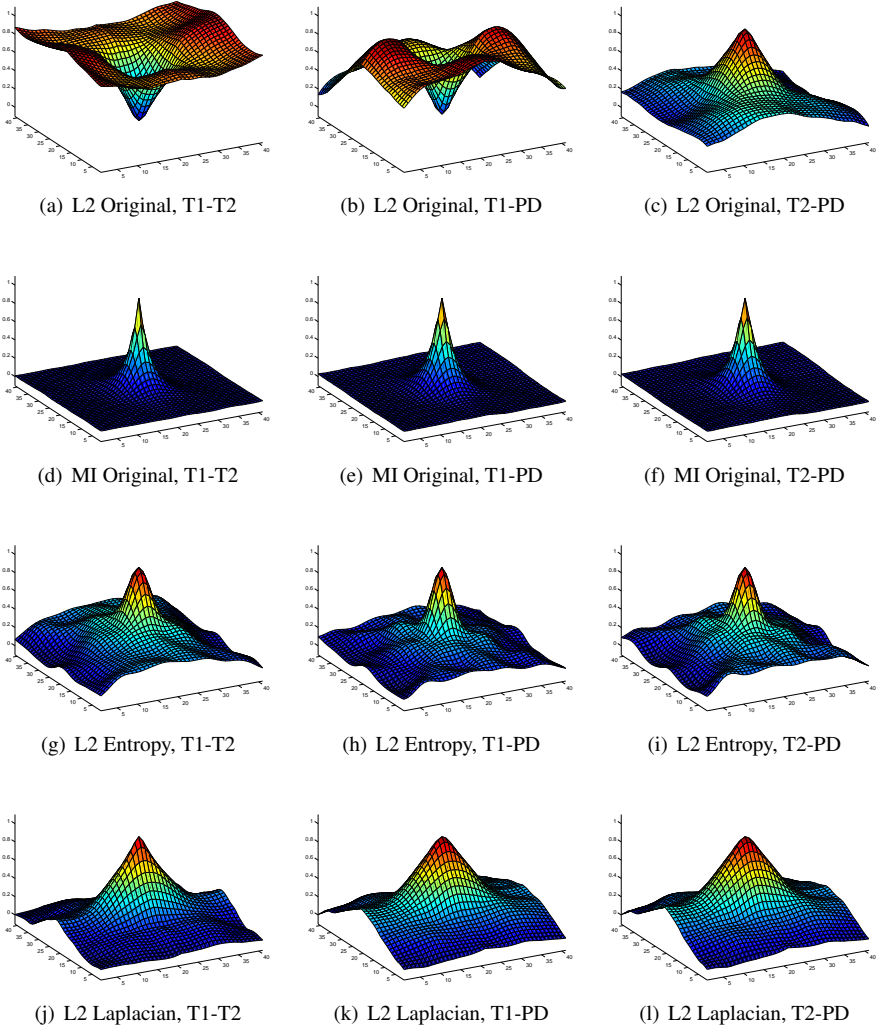


Figure 4: Plot of similarity measures with respect to translation in x and y direction. Maxima indicate best alignment.

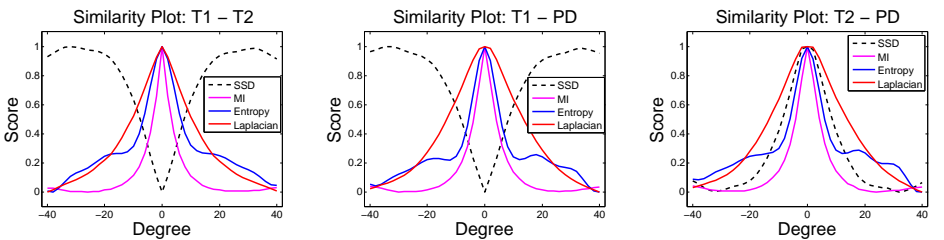


Figure 5: Plot of similarity measures with respect to rotation.

6 Conclusion

In this article, we presented a more accurate formulation of the requirements on a structural representation for multi-modal image registration. We subsequently showed that Laplacian eigenmaps fulfill these revised requirements. In fact, the first property is fulfilled optimally, since the embedding of Laplacian eigenmaps is designed to preserve locality. For the second property, we assumed comparable internal similarity across imaging modalities, as it was previously noted in the literature. Our results show that images from different modalities have indeed comparable internal similarity, and that Laplacian eigenmaps are well suited for providing a structural image representation. We obtained very good results for the Laplacian images in our experiments, confirming its theoretical advantages.

Acknowledgment: This work was partly funded by the European Commission.

References

- [1] A. Andronache, M. von Siebenthal, G. Székely, and P. Cattin. Non-rigid registration of multi-modal images using both mutual information and cross-correlation. *Medical Image Analysis*, 12(1):3–15, 2008.
- [2] O. Arandjelovic and R. Cipolla. A manifold approach to face recognition from low quality video across illumination and pose using implicit super-resolution. *ICCV*, 2007.
- [3] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.*, 15(6), 2003. ISSN 0899-7667.
- [4] S. Benhimane and E. Malis. Real-time image-based tracking of planes using efficient second-order minimization. In *IEEE/RSJ*, pages 943–948, 2004.
- [5] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. 3D multi-modality medical image registration using feature space clustering. In *Computer Vision, Virtual Reality and Robotics in Medicine*, 1995.
- [6] Samuel Gerber, Tolga Tasdizen, Sarang Joshi, and Ross Whitaker. On the manifold structure of the space of brain images. In *MICCAI*, volume 5761, 2009. ISBN 978-3-642-04267-6.
- [7] E. Haber and J. Modersitzki. Intensity gradient based registration and fusion of multi-modal images. *Methods of information in medicine*, 46(3):292–299, 2007.
- [8] Jihun Hamm, Christos Davatzikos, and Ragini Verma. Efficient large deformation registration via geodesics on a learned manifold of images. In *MICCAI*, volume 5761, pages 680–687, 2009. ISBN 978-3-642-04267-6.
- [9] C.S. Lee and A. Elgammal. Modeling view and posture manifolds for tracking. In *ICCV*, 2007.
- [10] Daewon Lee, M. Hofmann, F. Steinke, Y. Altun, N.D. Cahill, and B. Scholkopf. Learning similarity measure for multi-modal 3d image registration. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 186–193, 20-25 2009. doi: 10.1109/CVPR.2009.5206840.

- [11] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, 2004. ISSN 0920-5691. doi: <http://dx.doi.org/10.1023/B:VISI.0000029664.99615.94>.
- [12] 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. *Medical image analysis*, 1(2):151–161, 1996.
- [13] J.B.A. Maintz, P.A. van der Elsen, and M.A. Viergever. Registration of 3D Medical Images Using Simple Morphological Tools. In *International Conference on Information Processing in Medical Imaging*, pages 204–217, 1997.
- [14] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27(10):1615–1630, 2005. URL <http://lear.inrialpes.fr/pubs/2005/MS05>.
- [15] G.P. Penney, J. Weese, J.A. Little, P. Desmedt, D.L.G. Hill, and D.J. Hawkes. A comparison of similarity measures for use in 2-d-3-d medical image registration. *Medical Imaging, IEEE Transactions on*, 17(4):586–595, Aug. 1998.
- [16] G.P. Penney, L.D. Griffin, A.P. King, and D.J. Hawkes. A novel framework for multi-modal intensity-based similarity measures based on internal similarity. *SPIE*, 6914, 2008.
- [17] Daniel Pizarro and Adrien Bartoli. Shadow resistant direct image registration. In *SCIA*, pages 928–937, 2007.
- [18] Josien P. W. Pluim, J. B. Antoine Maintz, and Max A. Viergever. Mutual information based registration of medical images: A survey. *IEEE Trans. Med. Imaging*, 22(8):986–1004, 2003.
- [19] Alexis Roche, Grégoire Malandain, Xavier Pennec, and Nicholas Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proc. of First Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI’98)*, volume 1496 of *LNCS*, pages 1115–1124, Cambridge, USA, October 1998. Springer Verlag.
- [20] Gustavo K. Rohde, Wei Wang, Tao Peng, and Robert F. Murphy. Deformation-based nonlinear dimension reduction: Applications to nuclear morphometry. In *ISBI*, 2008.
- [21] Sam T. Roweis and Lawrence K. Saul. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*, 290(5500):2323–2326, 2000.
- [22] E. Shechtman and M. Irani. Matching Local Self-Similarities across Images and Videos. *IEEE Conference on Computer Vision and Pattern Recognition*, 2007.
- [23] J.B. Tenenbaum, V. Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319, 2000.
- [24] Engin Tola, Vincent Lepetit, and Pascal Fua. A fast local descriptor for dense matching. In *Conference on Computer Vision and Pattern Recognition*, Alaska, USA, 2008.

- [25] Christian Wachinger and Nassir Navab. Similarity Metrics and Efficient Optimization for Simultaneous Registration. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.
- [26] Christian Wachinger and Nassir Navab. Structural image representation for image registration. In *CVPR Workshops, IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, San Francisco, California, June 2010.
- [27] Christian Wachinger, Diana Mateus, Andreas Keil, and Nassir Navab. Manifold Learning for Patient Position Detection in MRI. In *IEEE International Symposium on Biomedical Imaging*, Rotterdam, The Netherlands, April 2010.
- [28] Christian Wachinger, Mehmet Yigitsoy, and Nassir Navab. Manifold learning for image-based breathing gating with application to 4d ultrasound. In *MICCAI*, Beijing, China, September 2010.
- [29] W. Wein, S. Brunke, A. Khamene, M.R. Callstrom, and N. Navab. Automatic ct-ultrasound registration for diagnostic imaging and image-guided intervention. *Medical Image Analysis*, 12:577–585, October 2008.
- [30] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1996. URL <http://citeseer.ist.psu.edu/354937.html>.
- [31] Qilong Zhang, Richard Souvenir, and Robert Pless. On manifold structure of cardiac mri data: Application to segmentation. *CVPR*, 1:1092–1098, 2006. ISSN 1063-6919.