# Manifold Learning for Multi-Modal Image Registration

Christian Wachinger wachinge@in.tum.de Nassir Navab navab@in.tum.de Computer Aided Medical Procedures (CAMP) Technische Universität München (TUM) Munich, Germany

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 [3] 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, see figure 1 for an overview of the procedure. It has the theoretical advantage to present an optimal approximation to one of the criteria for a perfect structural description. Laplacian eigenmaps search for similar patches in highdimensional 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 Structural Image Registration

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

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

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 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.

### 2 Structural Representation

The structural representation in [3] 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:

$$||Q_x - Q_y|| < \varepsilon \implies ||f(Q_x) - f(Q_y)|| < \varepsilon'$$
 (2)

#### (Q2) Structural equivalence:

$$Q_x \sim R_x \iff f(Q_x) = f'(R_x)$$
 (3)

(Q1) in comparison to [3] 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 ||.|| 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

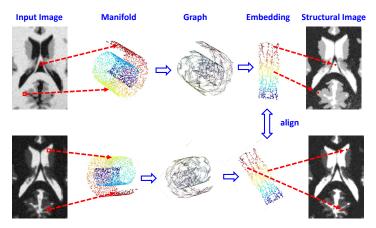


Figure 1: Structural representation with Laplacian eigenmaps. Patches of images lie on a manifold in high-dimensional patch space. The manifold is approximated by the neighborhood graph. The low-dimensional embedding is calculated with the graph Laplacian. Embeddings from different modalities have to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

different images. The inclusion of patches from the same image, as it is done in [3], 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'.

## 3 Laplacian Eigenmaps

Laplacian eigenmaps [1] 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.

It is exactly this preservation of locality that makes Laplacian eigenmaps interesting for a structural representation, since it optimally fulfills property (Q1). In the following we explain, why also (Q2) is fulfilled. 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 [2], we conclude that the structure or shape of both manifolds is similar. Since Laplacian eigenmaps preserve locality when embedding the manifold in a lowdimensional space, this structure is preserved in low dimensions. We could then directly use the low-dim coordinates 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 has to be performed. The coordinates of the registered embeddings finally provide the structural descriptors, see figure 1.

- Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.*, 15(6), 2003. ISSN 0899-7667.
- [2] 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.
- [3] Christian Wachinger and Nassir Navab. Structural image representation for image registration. In MMBIA, 2010.