

# Hierarchical Manifold Learning

Kanwal K. Bhatia<sup>1</sup>, Anil Rao<sup>1</sup>, Anthony N. Price<sup>2</sup>,  
Robin Wolz<sup>1</sup>, Jo Hajnal<sup>2</sup>, and Daniel Rueckert<sup>1</sup>

<sup>1</sup> Biomedical Image Analysis Group,  
Department of Computing, Imperial College London, London, UK  
<sup>2</sup> Division of Imaging Sciences and Biomedical Engineering,  
King's College London, London, UK\*

**Abstract.** We present a novel method of Hierarchical Manifold Learning which aims to automatically discover regional variations within images. This involves constructing manifolds in a hierarchy of image patches of increasing granularity, while ensuring consistency between hierarchy levels. We demonstrate its utility in two very different settings: (1) to learn the regional correlations in motion within a sequence of time-resolved images of the thoracic cavity; (2) to find discriminative regions of 3D brain images in the classification of neurodegenerative disease.

## 1 Introduction

In recent years, the use of manifold learning has become increasingly widespread in medical imaging, being used to uncover underlying structure both within a subject [14][17], and across populations [2][9][11][15]. Manifold learning techniques aim to discover the intrinsic dimensionality of data: a low-dimensional embedding which retains local structure of the data. Typical methods operate on images as a whole, with an entire image represented by a single point in vector space. A single medical image, however, may consist of several anatomical structures, which may vary to different extents between subjects across a population or within a time sequence of, for example, images of the thoracic cavity. An automated manifold learning technique to investigate images on a regional basis, without any prior information, forms the goal of our work.

Prior work in this area has suggested dividing an image into regular patches and finding the embedding for all patches simultaneously [4]. This was shown to offer advantages over whole image, as well as independent patch embedding, in learning the cardiac and respiratory cycles in 2D time sequences of the heart. One issue with this method is that the patch size needs to be chosen in some way, reflecting the size of the structures of the data in order to be useful. In addition, the embedding requires the solution of a matrix of size *number of patches* ×

---

\* We thank Marc Modat and M. Jorge Cardoso from the Centre for Medical Image Computing, University College London, for their advice and assistance. The work is partially funded under the 7th Framework Programme by the European Commission (<http://cordis.europa.eu/ist/>).

*number of images*. This makes it unsuitable for large sets of 3D data or for analysis at finer scales. In our work, we continue the idea of patch embeddings; however, we take a multiscale approach to remove the need to choose patch size. This results in a hierarchy of low-dimensional embeddings created using large to successively smaller patch sizes. To ensure consistency of the embeddings across anatomy, we enforce similarity of the embeddings between hierarchy levels. We use this *hierarchical manifold learning* (HML) for the analysis of regional variations within the thoracic cavity. Additionally, we apply it to a population of 3D brain images. Using this, we are able to automatically detect regions in the brain relevant to the classification of neurodegenerative disease.

## 2 Background

**Single Subject Analysis: Cardiac Image Analysis.** Being able to learn the cardiac cycle from images is of great potential benefit. Previous work has used the manifold structure of the cardiac cycle as an aid to left ventricle segmentation [17]. Knowledge of the physical motion of the heart is also becoming of increasing interest in patient selection for cardiac resynchronisation therapy, with *mechanical* dyssynchrony derived from image data shown to be a better predictor of outcome in certain patients [7]. Being able to automatically determine the phase differences between different regions of the heart directly from image data could therefore prove clinically useful.

Cardiac image analysis is, however, complicated by the presence of respiratory motion. This can also be analysed using manifold learning and previous work has employed this for ultrasound gating [14] and for lung CT volume reconstruction [8]. Two techniques for measuring respiration are through respiratory bellows and free-breathing navigator-gated strategies. Bellows monitor the physical movement of the chest wall during breathing, while navigators measure the displacement of the diaphragm [12]. Both methods acquire surrogates for the respiratory motion at single physical locations which then need to be correlated with the motion of the heart. This motivates the investigation of how the physical movements at various locations within the abdominal/thoracic cavity are related, and how this can be learned directly from images.

We apply hierarchical manifold learning in order to learn the different motions occurring in a sequence of real-time cardiac MRI, obtaining spatially-varying respiratory and cardiac correlations.

**Population Analysis: Feature Selection for Disease Classification.** Manifold learning has been used to build efficient representations of large sets of brain images [9] [11]. Reducing the dimension in this way has shown, for example, to be of benefit in the classification of Alzheimer’s disease [15]. Previous work, however, has solely focussed on whole images or specific segmented structures. We use HML to classify subjects from a population of subjects with Alzheimer’s disease and normal controls. In addition to this, we show how our method can be used to automatically detect the most discriminative regions for classification.

### 3 Methods

#### 3.1 Manifold Learning

Manifold learning algorithms are based on the premise that data are often of artificially high dimension. Indeed, they can be well-represented by a manifold of much lower dimensionality embedded in the high dimensional space. The aim of manifold learning algorithms (see [6] for a summary), is to discover the *embedding coordinates* approximating this manifold, thereby reducing the dimensionality of the data. In order to do this effectively, it is necessary to maintain the local structure of the data in the new embedding. This structure can be represented by constructing a graph  $G(V, E)$ , where the vertices  $V$  correspond to data samples and edges  $E$  represent neighbourhood (defined using k-nearest neighbour or  $\epsilon$ -ball distance in the space of the original data) similarities between the data points. These similarities can be encapsulated in a similarity matrix  $\mathbf{W}$  given by:

$$W_{ij} = \begin{cases} e^{-\frac{\|u_i - u_j\|^2}{2}} & \text{if } i, j \text{ neighbours;} \\ 0 & \text{otherwise.} \end{cases}$$

where  $\|u_i - u_j\|^2$  represents the Euclidean ( $\ell_2$ ) distance between data points  $i$  and  $j$ . We can then define the graph Laplacian operator as  $\mathbf{L} = \mathbf{D} - \mathbf{W}$  where  $\mathbf{D}$  is a diagonal matrix with non-zero elements,  $\mathbf{D}_{ii} = \sum_j W_{ij}$ , representing the degree of the  $i$ th vertex.

**Laplacian Eigenmaps.** One commonly-used manifold learning technique is Laplacian Eigenmaps (LE) [5]. This preserves structure in the data by ensuring that data points which are "close" in the high-dimensional space remain "close" in the low-dimensional embedding. This is done by minimising the following cost function:

$$C(x) = \sum_{ij} (x_i - x_j)^2 W_{ij} \quad (1)$$

which minimises the weighted Euclidean distance between the embedding coordinates  $x_i$  and  $x_j$  of data points  $i$  and  $j$ , respectively, in the low-dimensional embedding. The measure of closeness between points  $i$  and  $j$  is defined by the weight,  $W_{ij}$ , which indicates their similarity. For medical imaging applications, these points can represent whole images or image patches. Some common similarity metrics include functions of the Euclidean norm distance (equivalent to sums-of-squared differences between the images) and the correlation coefficient [14] as well as those based on Gabor filter responses [13]. One advantage of LE over other manifold learning techniques is its capacity to additionally handle non-metric similarity measures such as Normalised Mutual Information [15]. The  $n$ -dimensional solution to (1) is given by the eigenvectors  $\mathbf{x}$  of the generalised eigenvalue problem  $\mathbf{L}\mathbf{x} = \lambda\mathbf{D}\mathbf{x}$ , corresponding to the  $n$  smallest non-zero eigenvalues ( $\lambda$ ).

### 3.2 Hierarchical Manifold Learning (HML)

Previous work on localised manifold learning [4] has shown that naively partitioning images into regular patches is insufficient: anatomical structures rarely fall conveniently within a regular grid. To overcome this, the authors add a weighting term between adjacent patches to keep their low-dimensional embeddings similar, thus utilising neighbourhood information in the construction of the joint embeddings. One drawback of this approach is its inability to scale to very large numbers of images or patches. Furthermore, it lacks a systematic method of selecting patch size. Different regions may prove to be more or less relevant at different scales and choosing a single scale may miss trends evident at other levels. Hierarchical manifold learning aims to address these issues. In our work we present our framework and show how it can be applied to cardiac images at even very fine scales (2x2mm). In addition, we apply this to the classification of subjects with Alzheimer’s disease and show how HML can be used to automatically determine the best regions for classification.

The methodology is based on the idea of manifold alignment [10]. Given a first level embedding  $\bar{\mathbf{x}}$  of the full image, we can then recursively subdivide the image into equal smaller parts (for example, into four in a 2D image, or eight in 3D). As each subdivision is contained within the current level, we would expect embeddings at successive levels to be similar in some way. We obtain the new embedding  $\mathbf{x}$  of one of these sub-parts by amending the LE cost function such that the new embedding is also close to the embedding obtained at the previous level, that is, we align the new embedding to its ”parent” embedding  $\bar{\mathbf{x}}$ .

$$C(x) = \sum_i \mu(x_i - \bar{x}_i)^2 + (1 - \mu) \sum_{ij} (x_i - x_j)^2 W_{ij} \quad (2)$$

where  $\mu$  is a weighting parameter to determine the influence of each term. Low values for  $\mu$  reduce the strength of the inter-manifold alignment and the embedding moves towards the embedding of individual patches. Higher values for  $\mu$ , in contrast, lead to closer aligned embeddings. Since the right-most term in Eq. (2) is equivalent to  $\mathbf{x}^T \mathbf{L} \mathbf{x}$ , it can be shown that the analytical solution to (2) is given by the linear solution:

$$\mathbf{x} = (\mu \mathbf{I} + (1 - \mu) \mathbf{L})^{-1} \mu \mathbf{I} \bar{\mathbf{x}} \quad (3)$$

Here,  $\mathbf{L}$  is the graph laplacian for the particular patch under consideration, computed in the same way as in the standard LE formulation (and so has dimension equal to the number of images). This avoids the need to solve eigenvalue problems of large, augmented laplacian matrices as in [4]. Additionally we can keep recursively subdividing to very fine image scales to produce sensitive spatially-varying embeddings. By keeping each embedding close to its previous level, we avoid issues with smaller patches being dominated by noise.

This method requires calculating pairwise similarities between all images for each patch at each level - but at each level this has the same number of operations as a single level manifold embedding of the whole image. The linear equation [2]

is solved for every patch with the same complexity. The algorithm can therefore be applied at fine scales of large sets of 3D images.

## 4 Results

**Application to Cardiac Image Analysis.** Healthy volunteers were imaged using real time MR, using a balanced steady state free precession (SSFP) sequence with spatial resolution of  $2\times 2\times 10\text{mm}$  and temporal resolution of 117ms per frame (FA/TE/TR = 20/1.2/2.4ms). Short axis, 2-chamber and 4-chamber cardiac views were acquired for 200 dynamics (the last 180 being used for analysis). Traces from respiratory bellows were recorded alongside sequence event markers in order to align images to physiological motion.

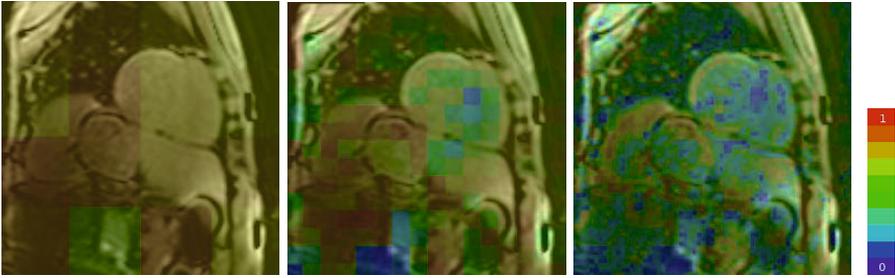
We apply HML to a time-varying sequence of 2D cardiac images. At each level, and at each patch, we construct a graph where each node represents part of a 2D frame at one point in time, in a similar way to the conventional method of manifold learning. Since we expect similar contrasts for all frames, we weight each pair of frames using the distance defined in Section 3.1, retaining the 24 nearest neighbours at each data point.

Reducing to two dimensions, we find that one dimension shows stronger correlation with the bellows trace and the other with the area of the left ventricle (a proxy for the cardiac cycle). Figure 1 shows the correlation coefficient,  $\rho = \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$ , between the first dimension of the embedding coordinates for each patch and the respiratory bellows trace, for three different levels of the hierarchy. It can be seen that there is a strong correlation between the trace and, in particular, the motion of the liver. Figure 2 shows the correlation between the embedding coordinates for each patch and the area of the left ventricle in each frame. As expected, the best correlations can be found in the heart and also in areas with blood vessels. Additionally, as the embeddings are, by design, in alignment, we can correlate the embedding at any location with any other location, using their coordinates directly.

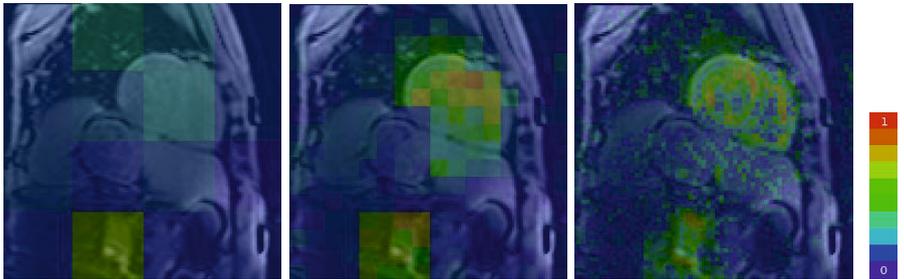
The weighting parameter  $\mu$  dictates how strongly each embedding should be aligned to its parent embedding. A higher value keeps the embeddings closer, but may lead to a lower discernibility between patches. This can be seen in Figure 3, which shows the correlation coefficients between each patch and the bellows trace, where a high value of  $\mu$  gives a more homogenous embedding.

Table 1 compares the computational time required for HML and the simultaneous embedding of [4], for varying patch sizes on a 2.4GHz processor. HML allows the embeddings of much smaller patch sizes to be efficiently computed.

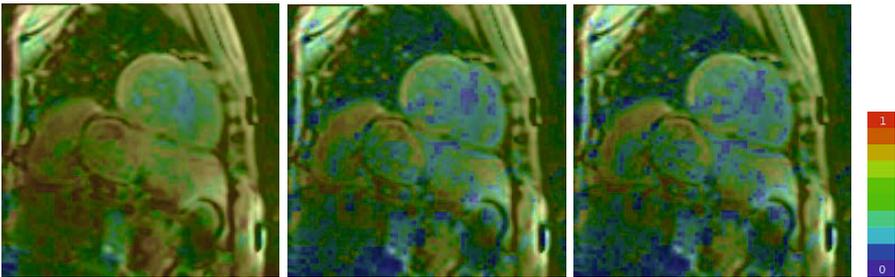
**Application to Feature Selection for Disease Classification.** We have applied HML to the Alzheimer’s Disease Neuroimaging Initiative (ADNI) [1] dataset of 429 subjects of size  $160 \times 192 \times 160\text{ mm}$ . This consists of 231 normal control subjects and 198 subjects with Alzheimer’s disease. Skull-stripped images have been intensity-normalised and aligned to MNI space using affine registration. We apply HML to this dataset using  $\mu = 0.1$  to allow for greater



**Fig. 1.** Correlation with respiratory bellows trace at various resolution levels ( $\mu = 0.5$ )



**Fig. 2.** Correlation with area of the left ventricle at various resolution levels ( $\mu = 0.5$ )



**Fig. 3.** Correlation with respiratory bellows trace using different weighting parameters  $\mu$ . From left to right:  $\mu = 0.9$ ,  $\mu = 0.5$ ,  $\mu = 0.1$ .

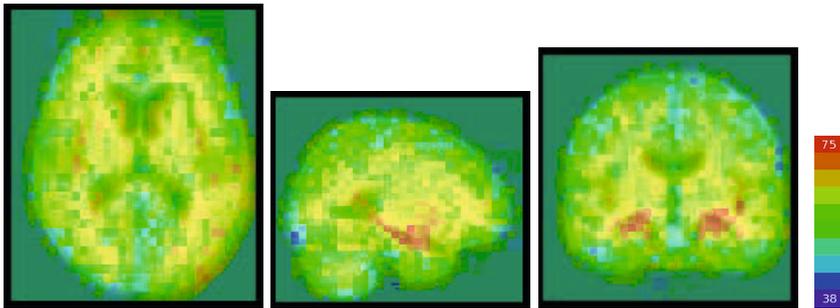
discernability between patches and reduce the data to 20 dimensions. Similarity between patches is computed as specified in 3.1. The neighbourhood is selected to be as many neighbours as needed to keep the graph fully connected. At each level, we classify the patches using linear support vector machines on the embedding coordinates of each patch individually. Ten-fold cross validation is used, with the same folds used for each patch. This allows us to assess the discriminative power of each patch. Figure 4 shows the classification accuracy with the smallest patch size of  $5 \times 6 \times 5 \text{mm}$ . The most discriminative patches can be seen to

**Table 1.** Time in seconds to run each embedding for various patch sizes

Patch size (mm)	Hierarchical	Simultaneous [4]
64	1.23	0.22
32	1.80	16.1
16	3.46	4872

be concentrated around the hippocampus, in keeping with established results. HML can therefore also be viewed as a method for feature selection.

Table 2 shows the classification accuracy of the most discriminative patch with patch size (HML). The best value is comparable to classification rates obtained through manifold learning using segmented hippocampal volumes [16]. For comparison, we also show the classification rates of the same patches obtained when constructing a separate manifold for each patch independently (IND), using standard Laplacian Eigenmaps (that is, without any hierarchical alignment, while keeping all other parameters the same). It is likely that the neighbourhood information, implicit in HML embedding, helps to improve classification accuracy.

**Fig. 4.** Classification accuracy (%) of 5mm patches of Alzheimer's data**Table 2.** Comparison of patch classification rates at various resolution levels using embeddings obtained from hierarchical manifold learning (HML) and independently, without linkage to coarser levels (IND)

Patch side length (mm)	160	80	40	20	10	5
Classification rate of most discriminative patch (HML) (%)	64	68	70	74	76	75
Classification rate of most discriminative patch (IND) (%)	64	66	68	66	68	69

## 5 Discussion

We have presented Hierarchical Manifold Learning to explore regional variations within image data. We have shown how it can be used to determine local correlations within the cardiac and respiratory cycles in image sequences of the

thoracic cavity. The algorithm does not require pre-selecting regions of interest or patch size, and can be used to obtain embeddings even at very fine scales.

We have demonstrated the scalability of the algorithm by applying it to a population of 3D image datasets, in order to discover discriminative regions in the classification of Alzheimer's disease. In future work, we aim to investigate how to best combine these individual patch embeddings in more sophisticated ensemble classification schemes.

## References

1. <http://www.loni.ucla.edu/ADNI>
2. Aljabar, P., et al.: A combined manifold learning analysis of shape and appearance to characterize neonatal brain development. *IEEE TMI* 30(12), 2072–2086 (2011)
3. Helm, R.H., Lardo, A.C.: Cardiac magnetic resonance assessment of mechanical dyssynchrony. *Curr. Opin. Cardiol.* 23, 440–446 (2008)
4. Bhatia, K.K., Price, A.N., Hajnal, J.V., Rueckert, D.: Localised manifold learning for cardiac image analysis. In: Haynor, D.R., Ourselin, S. (eds.) *Proc. SPIE 2012* (2012)
5. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation* 15(6), 1373–1396 (2003)
6. Cayton, L.: *Algorithms for Manifold Learning*. Tech. report, UCSD (2005)
7. Delgado, V., Bax, J.J.: Assessment of systolic dyssynchrony for cardiac resynchronization therapy is clinically useful. *Circulation* 123, 640–655 (2011)
8. Georg, M., Souvenir, R., et al.: Manifold learning for 4D CT reconstruction of the lung. In: *IEEE Computer Society Workshop MMBIA* (2008)
9. Gerber, S., Tasdizen, T., Thomas Fletcher, P., Joshi, S., Whitaker, R., A.D.N.I.: Manifold modeling for brain population analysis. *Med. Im. An.* 14(5), 643–653 (2010)
10. Ham, J., Lee, D.D., Saul, L.K.: Semisupervised alignment of manifolds. *AI and Statistics*, 120–127 (2005)
11. Hamm, J., Davatzikos, C., Verma, R.: Efficient Large Deformation Registration via Geodesics on a Learned Manifold of Images. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I. LNCS*, vol. 5761, pp. 680–687. Springer, Heidelberg (2009)
12. Savill, F., et al.: Assessment of input signal position for cardiac respiratory motion models during different breathing patterns. In: *Proc. ISBI*, pp. 1698–1701 (2011)
13. Souvenir, R., Pless, R.: Image distance functions for manifold learning. *Image and Vision Computing* 25(3), 365–373 (2007)
14. Wachinger, C., Yigitsoy, M., Navab, N.: Manifold Learning for Image-Based Breathing Gating with Application to 4D Ultrasound. In: Jiang, T., Navab, N., Plum, J.P.W., Viergever, M.A. (eds.) *MICCAI 2010, Part II. LNCS*, vol. 6362, pp. 26–33. Springer, Heidelberg (2010)
15. Wolz, R., Aljabar, P., Hajnal, J.V., Rueckert, D.: Manifold Learning for Biomarker Discovery in MR Imaging. In: Wang, F., Yan, P., Suzuki, K., Shen, D. (eds.) *MLMI 2010. LNCS*, vol. 6357, pp. 116–123. Springer, Heidelberg (2010)
16. Wolz, R., et al.: Automatically determined hippocampal atrophy rates in ADNI. *Alzheimer's and Dementia* 6(4), S284 (2010)
17. Zhang, Q., Souvenir, R., Pless, R.: On manifold structure of cardiac MRI data: Application to segmentation. In: *Proc. IEEE CVPR*, vol. 1, pp. 1092–1098 (2006)