Image Distance Functions for Manifold Learning

Richard Souvenir, Robert Pless *

Washington University in St. Louis
Department of Computer Science and Engineering
One Brookings Drive, Campus Box 1045
St. Louis, Missouri, 63130 USA

Abstract

Many natural image sets are samples of a low-dimensional manifold in the space of all possible images. When the image data set is not a linear combination of a small number of basis images, linear dimensionality reduction techniques such as PCA and ICA fail and nonlinear dimensionality reduction techniques are required to automatically determine the intrinsic structure of the image set. Recent techniques such as ISOMAP and LLE provide a mapping between the images and a low-dimensional parameterization of the images. This paper specializes general manifold learning by considering a small set of image distance measures that correspond to key transformation groups observed in natural images. This results in more meaningful embeddings for a variety of applications.

Key words: Isomap, manifolds, nonparametric registration
PACS: 02.60.Ed, 87.80Pa

1 Introduction

Faster computing power and cheap large scale memory has led to a surge in research in the machine learning community on the topic of dimensionality reduction, which finds structure in a large set of points embedded in a very high-dimensional space. Many problems in computer vision can be cast in this framework, as each image can be considered to be a point in a space with one

* Corresponding author. Tel: +01 314 935 7546; Fax: +01 314 935 7302
Email addresses: rms2@cse.wustl.edu (Richard Souvenir),
pless@cse.wustl.edu (Robert Pless).

dimension for each pixel. When an image data set is generated by varying just a few parameters, such as a combination of pose, lighting, or camera viewpoints, then this set can be considered to be sampling a continuous manifold of the space of all possible images. Given a set of images, understanding this manifold and automatically parameterizing each image by its place on this manifold has emerged as an important tool in the model-free interpretation of image data.

Especially for the analysis of the variation in images of a single object, this approach was long foretold:

> In a very large part of morphology, our essential task lies in the comparison of related forms rather in the precise definition of each; and the deformation of a complicated figure may be a phenomenon easy of comprehension, though the figure itself may have to be left unanalyzed and undefined.
> —D'Arcy Thompson, [1]

Algorithms for inferring properties of image manifolds by comparing related images has been codified in a family of computational techniques exemplified by Isomap [2] and Locally Linear Embedding (LLE) [3]. These techniques extend a sparse set of local relationships between similar images to a global low-dimensional parameterization of all images. This work uses Isomap as an exemplar of this class of nonlinear dimensionality reduction tools, and the results will directly apply to other methods including Semidefinite Embedding [4], and could be extended to LLE, alignment of local representations (LLC) [5], and Hessian Eigenmaps [6].

The main contribution of this paper is to explore the application of Isomap to video imagery, and to guide the process of specializing Isomap for particular problem domains. Several earlier papers have visualized the parameterization of image sets and observe that it highlights perceptually relevant features. Here we emphasize that the parameterization produced by Isomap is a function of the input data set *and* the image distance metric.

A formal theory of the statistics of natural images and natural image variations — Pattern Theory — gives tools for defining relevant image distance metrics. We postulate that for natural image data sets, a small number of distance metrics are useful for many important applications. This paper proposes a set of distance measures that correspond to the most common causes of transformation in image sets and gives examples of how these significantly improve performance on a variety of application domains, including the de-noising of cardiac MR imagery.

2 Differential structure in dimensionality reduction

Dimensionality reduction is an important tool in image data analysis, because images are large, and when treated as a vector of pixel intensity values, lie in a very high-dimensional space. Here we give a brief introduction to Isomap as one tool for nonlinear dimensionality reduction. This is explicitly compared with linear dimensionality reduction as typified by principal components analysis (PCA). We argue that PCA is poorly suited to the analysis of many natural image sets, especially those which include motion. We then consider the structure of both Isomap and PCA embeddings.

2.1 Background of Isomap

Given an input set $\mathcal{I}$, which is a finite subset of $\mathcal{R}^D$, (where D is the number of pixels in an image), the dimensionality reduction techniques of Isomap and LLE produce a mapping function $f : \mathcal{I} \longrightarrow \mathcal{R}^d$. Very briefly, Isomap begins by computing the distance between all pairs of images (using the square root of the sum of the squared pixel errors, which is the L_2 norm distance if the images are considered points in $\mathcal{R}^D$). Then, a graph is defined with each image as a node and undirected edges connecting each image to its k-closest neighbors (usually choosing k between 5 and 10). A complete pair-wise distance matrix is calculated by solving for the all-pairs shortest paths in this sparse graph. Finally, this complete distance matrix is embedded into $\mathcal{R}^d$, by solving an Eigenvalue problem using a technique called Multidimensional Scaling (MDS) [7]. d is the dimension of the low-dimensional embedding and can be chosen as desired, but, ideally, is the number of degrees of freedom in the image set. LLE is a method with similar aims that creates a mapping that preserves linear relationships between nearby points. The original papers for Isomap [2] and LLE [3] have pointers to online, free implementations of the algorithm, a tradition which has been continued for the successors of these algorithms, including Hessian Eigenmaps [6], Laplacian Eigenmaps [8], and Semidefinite Embedding [4].

Isomap has several very important limitations. First, the Isomap algorithm defines a mapping from the original image set to $\mathcal{R}^d$. That is, Isomap computes a mapping $f : \mathcal{I} \longrightarrow \mathcal{R}^d$ and not, as might be more convenient, $f : \mathcal{R}^D \longrightarrow \mathcal{R}^d$. This means that once the Isomap embedding of an image set $\mathcal{I}$ is computed, for $I' \notin \mathcal{I}$, the value of $f(I')$ is not well defined. Additionally, the inverse mapping is also problematic. For a point $x \in \mathcal{R}^d$, if x is not in the set of points defined by $f(\mathcal{I})$, then $f^{-1}(x)$ is also not well defined. Although approaches have been proposed to compute these "out of sample" projections [9], this remains, both theoretically and practically, a challenge for Isomap and other dimensionality

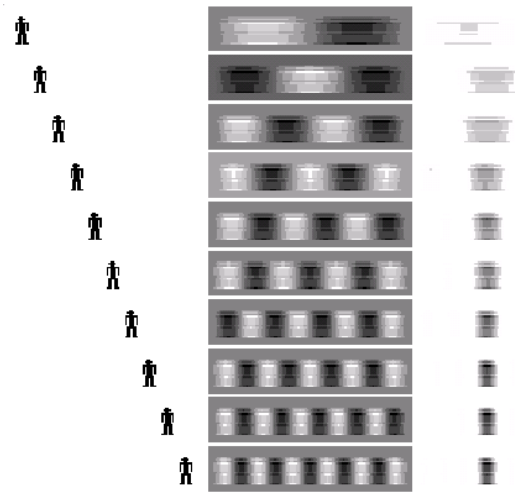

Fig. 1. (Left) Ten images from a 100 frame movie of a stick figure sliding left to right. (Middle) The first 10 principal images of this image set. (Right) The reconstruction of one frame of the original movie as the linear combination of progressively more principal images.

reduction techniques.

2.1.1 Comparison to PCA

It is instructive to view PCA in the same light. Given an input set of images $\mathcal{I}$ (still a finite subset of $\mathcal{R}^D$), Principal Component Analysis computes a function f which projects each image onto a set of basis images. The image set, $\mathcal{I}$, is used to derive a set of orthonormal basis images $\vec{b}_1, \vec{b}_2, \ldots \vec{b}_d$, and then the function f which maps an image x in $\mathcal{R}^D$ to a set of coefficients in $\mathcal{R}^d$ is:

$$f(x) = (x^\top \vec{b}_1, x^\top \vec{b}_2, \ldots, x^\top \vec{b}_d) = (c_1, c_2, \ldots, c_d)$$

Therefore, although the basis images are defined based upon an Eigen-analysis of the image data set $\mathcal{I}$, the function f is defined for all possible images of D pixels:

$$f_{PCA} : \mathcal{R}^D \longrightarrow \mathcal{R}^d$$

In addition to being more computationally efficient, the projection function f of PCA remains well defined for images that are not present in the original set $\mathcal{I}$. Also, the inverse function is defined as well, so that any point in the coefficient space can be mapped to a specific image by a linear combination of basis images:

$$f_{PCA}^{-1}(c_1, c_2, \ldots c_d) = c_1\vec{b}_1 + c_2\vec{b}_2 \ldots + c_d\vec{b}_d \tag{1}$$

Differential changes to the coefficients correspond to changes in weights of the linear basis functions. Consider an image x with corresponding coefficients $(c_1, c_2, \ldots c_d)$. The partial derivative of the inverse mapping function (Equation 1) describes how the image varies when changing the c_i coefficient:

$$\frac{\partial}{\partial c_i} f_{PCA}^{-1}(c_1, c_2, \ldots c_d) = \vec{b}_i$$

Equivalently, moving through the coefficient space can be interpreted as an operator: changing coefficient c_i by ϵ changes the image x by the addition of the b_i basis image: $x' = x + \epsilon\vec{b}_i$.

However, this is not usually the type of image change that underlies natural image variations. Natural changes to images, for example those due to variation in pose or shape deformations, are very poorly approximated by changes in linear basis functions. Figure 1 shows the PCA decomposition of an icon moving smoothly from left to right. Despite the fact that this image set has only one degree of freedom, it takes many principal components to reconstruct any of the original images effectively. This leads to the question: what local variations dominate the relationships between similar images in natural settings?

2.2 Differential Structure in Image Manifolds

Nonlinear dimensionality reduction, despite its drawbacks, has been successful at finding natural parameterizations, or "perceptual organizations" [2], of a variety of different image sets, including pose estimates in rigid body motions [10], visualization of biomedical image data sets [11], and a limited set of deformable/articulated hand poses [2].

Our exploration of Isomap starts by considering a data set of a woman running on a treadmill, captured by a standard camcorder. Figure 2 shows the output of the Isomap algorithm using the most common image distance metric, the sum of the squared pixel intensity differences.

To study the differential properties present in the Isomap embedding of a discrete set of natural images, we examine how the image set is parameterized and consider a sequence of nearby images along a straight line within the embedding. For the woman running on a treadmill, the cyclic nature of the running motion leads Isomap to embed the points in a circle. The "thickness" of the circle arises from the variation in the image appearance for images taken at the same part of the running cycle.

Thus, the Isomap embedding separates the image variations into two compo-

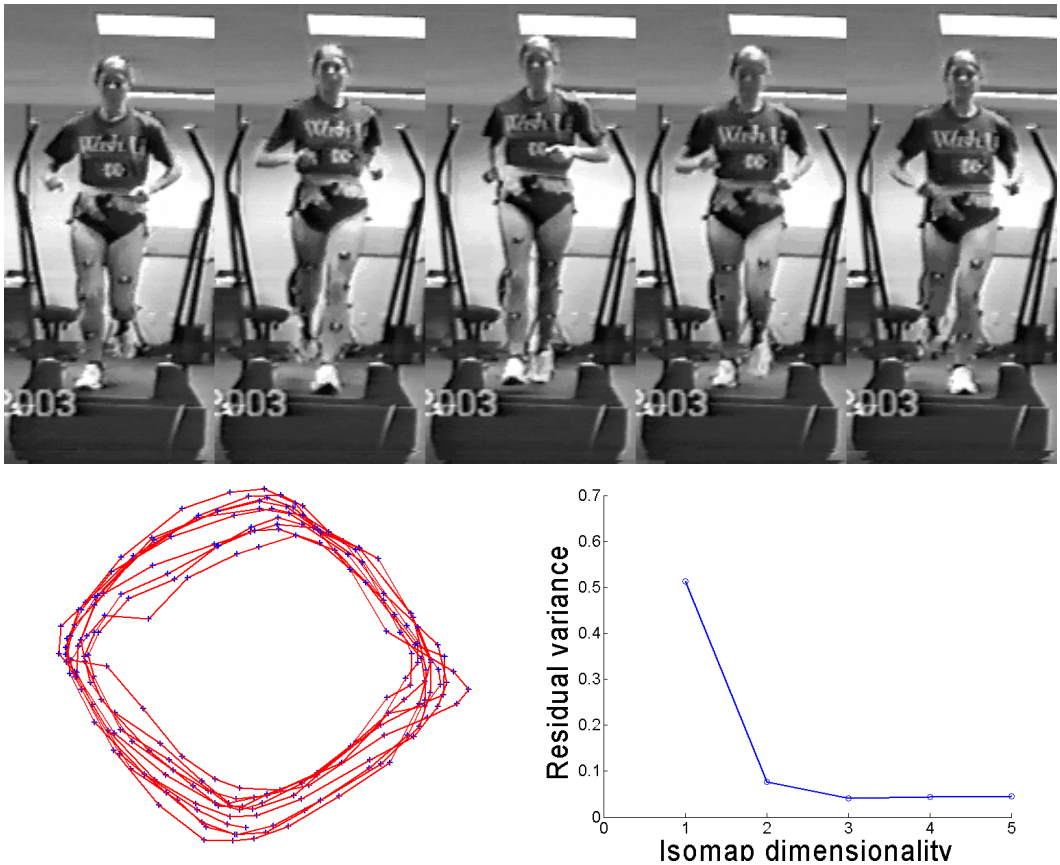

Fig. 2. Sample frames of a video data set of a woman running on a treadmill. (Bottom left) The two-dimensional Isomap embedding of this data set (using $k = 8$ neighbors). Each blue dot is the (nonlinear) projection of an original image and the red line connects the points from consecutive frames. (Bottom right) Plot of residual error shows that two dimensions capture almost all of the information in these local distance measurements.

nents. Tangential motion in the coefficient space (moving around the circle) corresponds to changes in the phase of the the running cycle. Radial motion in the coefficient space encodes the residual variation; in this sequence, the dominant change is the left to right position of the runner on the treadmill. Figure 3 shows an expanded view of the Isomap embedding, and the image set generated by moving radially through the coefficient space. This visualization of the data set is, in itself, a useful diagnostic tool. In this case, Isomap decomposes the data set into its main types of variation. However, this leads us to ask: what different types of image variation are likely to arise in natural images and can we improve Isomap in order to better understand the differential structure in the output mapping?

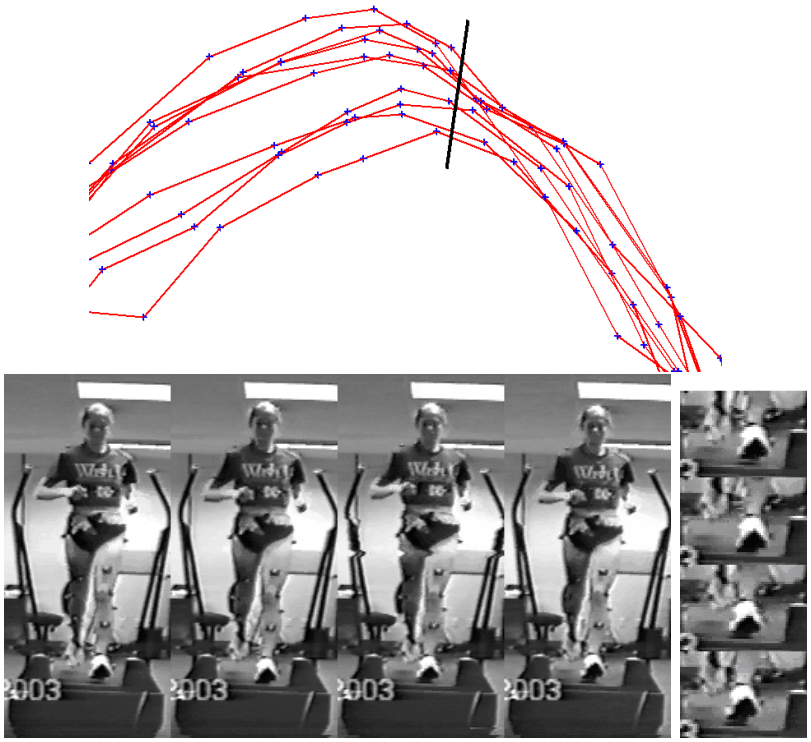

Fig. 3. An expanded view of the top of the trajectory shown in Figure 2. The radial variation show images taken at the same part of the running cycle. The dominant variation here is translation to the left, and can be seen most clearly in the enlarged view of the feet shown at the right.

3 Pattern Theory, Image Variation and Distance Metrics

Pattern Theory builds upon the characterization of shapes defined by Grenander [12], and encodes variations in shapes of natural objects as the results of applying elements of a small group of transformations. The core of the research in Pattern Theory has been to develop tools to define probability distributions over these transformations. However, we notice that in many videos of particular objects, the set of observed transformations is quite limited. For example, a human runner cycles through a particular set of 3D shapes (related to one another through the action of diffeomorphic transformation) as she/he passes in front of a camera (varying the rigid transform relating the 3D object to the camera). When the observed deformations lie on such a low-dimensional manifold, developing distance measures within this manifold suffices to discover interesting structures.

Deformable template analysis [13,14] is one instantiation of pattern theory that applies to images. Small deformations in the neighborhood of a particular

image $I_a \in \mathcal{R}^D$ can be expressed in terms of three components, image motion, photometric changes, and noise. Following a presentation is adapted from [13], these variations can be expressed as:

$$I_{ah}(\vec{p}) = I_a(\vec{p} - hv(\vec{p})) + h\sigma^2 z(\vec{p}) + N(\vec{p}), \tag{2}$$

Here, the first term uses a displacement field v to define the spatial motion of image regions (pixel $\vec{p}$ in image I_a provides support for the pixel $\vec{p} + hv(\vec{p})$ in image $I_{ah}(\vec{p})$), the second term uses an additive term z defined at each pixel and an overall scaling factor σ^2 to specify variations in image appearance not accounted for by motion (lighting changes for example), and the third term describes imaging noise, which ought to be independent of the magnitude of the overall transformation h.

In developing our distance measures, we also distinguish between global motion patterns caused by changes in the camera orientation or translations of the object, and local motions caused by the non-rigid object deformations. Our goal for this section, then, is to propose distance measures that approximate geodesic distances along each group of transforms:

Transform Group	Distance Measure
Image Noise	Euclidean Distance
Rigid Motions / Projection Changes	Global Motion Estimates
Non-rigid Motions	Local Motion Estimates
Intensity Variation	Local Contrast Change Estimates

Insofar as possible, it is useful to have each distance measure be easy to compute, and invariant to the other transform groups. Within these constraints, the exact choice of distance measure is likely to be flexible. The key is that it is not necessary to have a strong model of the exact transform (i.e., it is not necessary to have an explicit model of the stride of the woman on the treadmill), rather it is only necessary to know the type of transform (e.g. diffeomorphic deformation).

3.1 Euclidean Distance Measures Noise

The most common implementation of Isomap to analyze images starts by computing the Euclidean distance (square root of the sum of the squares of the pixel intensities) between each pair of images. Define $||I_a - I_b||_2$ to be the Euclidean distance between of two images. Does this this distance measure have any concrete interpretation with respect to our deformation models?

If I_a and I_b are separate images of the same object (under the same deformation), then, from Equation 2, $v(\vec{p})$, and $z(\vec{p})$ are uniformly zero, and the Euclidean distance between I_a and I_b is:

$$\sum_{\vec{p}} ||I_a(\vec{p}) - I_b(\vec{p})||_2 = \sum_{\vec{p}} N(\vec{p})$$

If this noise is i.i.d, Gaussian and zero-mean, then the Euclidean distance $||I_a - I_b||^2$ is a negative log-likelihood that the two images are of the same object. That is, under this model of image formation, the distance measure commonly used in Isomap is most directly a measure of how unlikely it is that they are the same image, rather than a measure of how different the two images are. The following sections consider different definitions of image distances, so that the image embedding function may be based on local distances more closely tied to magnitude of the image deformations.

3.2 Rigid Motion

Some changes to the imaging geometry lead to globally consistent image transformations; rigid translations of an object lead to translations and scale changes of the image, and changing camera parameters (calibration and zoom) are well approximated by affine image warping. Measuring the magnitude of these changes between two images can be expressed as finding the image warp that makes those images the most similar.

For example, we can express the allowable warping of an image I_b as AI_b, for $A \in T$, where T represents a class of allowable transforms that define a global motion across the image (such as affine warps). Then, the distance measure can be written as the magnitude of the transform that minimizes the image difference:

$$||I_a - I_b||_T = || \arg\min_{A \in T} (I_a - AI_b)||$$

However, manifold learning techniques are most relevant to the understanding of non-rigid motions. To understand non-rigid motion in natural data sets, it is sometimes important to ignore the image distances caused by rigid motion. A rigid motion-invariant distance measure can be written:

$$||I_a - I_b||_{\text{invar(T)}} = \min_{A \in T} ||I_a - AI_b||$$

This rigid motion-invariant distance measure can used in combination with the non-rigid motion distance measures of the next section and an example of

this is shown in Section 4.1.

3.3 Non-Rigid Motion

For the case of unknown non-rigid transformations, the generic class of diffeomorphic deformations is a natural choice of transform groups. These deformations may not have a global structure, so we propose to measure the magnitude of the transformation by accumulating measures of local motion over the image.

One implementation of this is to define a distance measure that uses the response of a collection of Gabor filters to estimate local motions. Complex Gabor filters are applied to the same positions in both images, and the phase difference of the complex response is summed over all locations. Given two images I_a, I_b and $G_{(\omega,\{V|H\},\sigma)}$ which is defined to be the 2D complex Gabor filter with frequency ω, oriented either vertically or horizontally, with σ as the variance of the modulating Gaussian, the distance can be expressed as:

$$\begin{aligned} ||I_a - I_b||_M = &\sum_{x,y} \Psi(G_{(\omega,V,\sigma)} \otimes I_a, G_{(\omega,V,\sigma)} \otimes I_b) \\ &+\Psi(G_{(\omega,H,\sigma)} \otimes I_a, G_{(\omega,H,\sigma)} \otimes I_b) \end{aligned}$$

where Ψ returns the absolute value of the phase difference of the pair of complex Gabor responses.

This distance function is dependent upon the choices of ω, and σ. The wavelength of the Gabor filter should be at least twice as large as the image motion caused by small deformations, and σ can be chosen as approximately the wavelength. In practice, this metric is surprisingly robust to the choice of σ.

Because it is based on the phase of the local image structure, this image distance measure is robust to small changes in the local contrast. Furthermore, because the Gabor filters are computed over small regions of the image, the effect of pixel noise is minimized.

Other distance measures are appropriate when the shape of the object is defined by its silhouette and the object can be cleanly segmented from the background. If the segmentation is robust, a distance metric is invariant to *any* changes in illumination or contrast as long as it relies only on binary valued data. For a pair of images I_a, I_b with point sets falling inside the silhouette P_a, P_b, we can employ the symmetric Hausdorff distance, $h(P_a, P_b)$. Extending this to become an affine invariant distance measure (as described

in Section 3.2) requires an additional minimization step:

$$||I_a - I_b||_A = \min_{A \in T} h(P_a, AP_b),$$

where AP_b is the point set of the second image after deformation by an affine transform A.

3.4 Intensity Variation

For image sets derived from an object undergoing intensity changes (e.g., contrast changes, lighting, shading, and fog), we exploit a different function of the Gabor filter bank responses. Given two images I_a, I_b and $G_{(\omega,\{V|H\},\sigma)}$ which is defined as in Section 3.3, the image distance can be expressed as:

$$\begin{aligned} ||I_a - I_b||_C = & \sum_{x,y} \Big| |G_{(\omega,V,\sigma)} \otimes I_a| - |G_{(\omega,V,\sigma)} \otimes I_b| \Big| \\ & + \Big| |G_{(\omega,H,\sigma)} \otimes I_a| - |G_{(\omega,H,\sigma)} \otimes I_b| \Big| \end{aligned}$$

where $|\cdot|$ returns the magnitude of a complex value.

Small motions of an image region may change the phase a Gabor filter response, but do not significantly affect the magnitude of the filter response, so this distance measure also has the desirable property of being largely invariant to small motions.

4 Applications

In this section, we illustrate the use of the proposed distance metrics from the previous section on two example application domains: a bird flying against a blue sky towards the camera and a cardiac MRI data set.

4.1 Rigid and Deformable Motion

We consider a data set of a flying bird captured against a clear sky. This data set exhibits two important properties. First, the clear sky background allows very simple and robust segmentation of the bird. Second, there exists an obvious dominant motion – the wings flapping. The wing flapping is a non-rigid deformation that is complicated to parameterize without an explicit bird

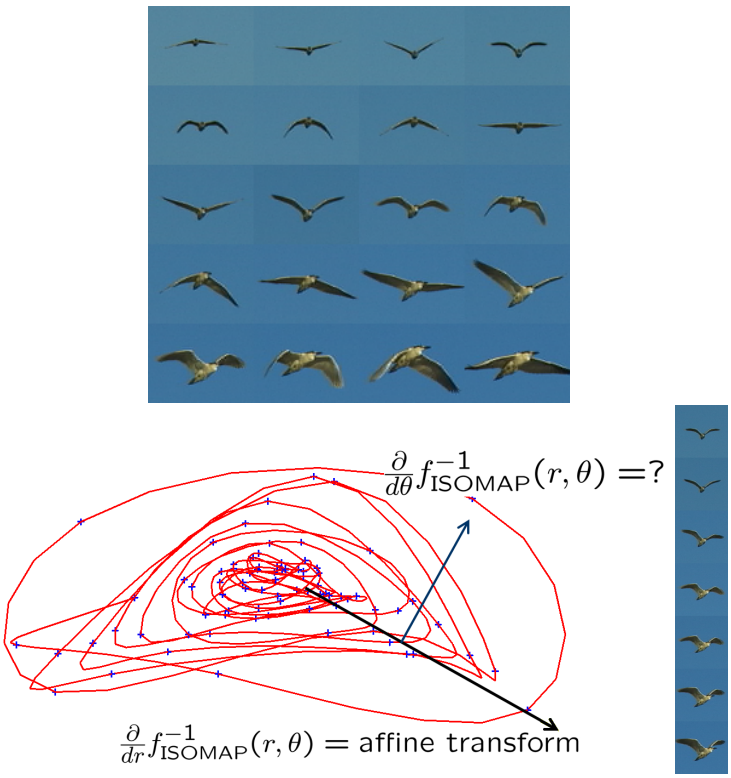

Fig. 4. (Top) Sample images of a video sequence of a bird flying across the sky. (Bottom left) The Isomap embedding of this set of images. Moving radially in the embedding corresponds, locally, to an affine transformation of the image that depends only on the relative position of the bird to the camera. The transform required to move tangentially in the Isomap space varies by location and requires a motion model of the bird. (Bottom right) Images closest to the dark radial arrow.

dynamics model. Furthermore, the bird is flying past the camera, so the rigid transformation relating the bird and camera position is continuously changing. Therefore, the variability in this data set is a combination of rigid and non-rigid motions. These properties of the input data set suggest that using the Hausdorff distance measure discussed at the end of Section 3.3 may elucidate relevant structures within the Isomap embedding.

Isomap is performed on this data set using the symmetric Hausdorff distance and $k = 8$ neighbors. This gives the embedding shown in Figure 4. There is a circular motion in the trajectory caused by the cyclic nature of the data. However, there is also a larger scale consistent radial motion, caused by image differences that arise from the approach of the bird toward the camera. Thus the Isomap embedding automatically de-couples the cyclical, non-rigid component of the bird motion from the rigid component of the bird approaching the camera. To highlight this effect, the right side of Figure 4 shows the images closest to a radial line in the Isomap embedding. The images nearest this line are approximately related by a rigid transformation.

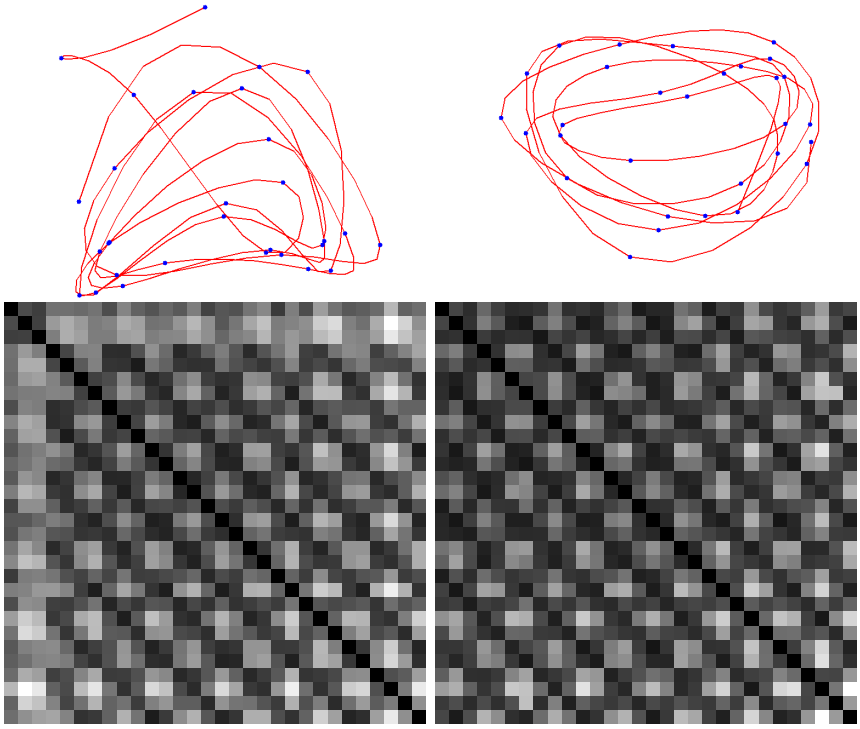

Fig. 5. (Top left) Isomap embedding of the bird image sequence where each image is a dot and the line connects the images in order. (Bottom left) the complete distance matrix defined by the Euclidean distance metric. (Right) The plots for the same sequence using the affine invariant distance measure described in Section 3.2.

In order to emphasize the deformable motion of the bird, we desire the distance function to ignore, as much as possible, variation caused by anything other than the deformable motion. Small rigid transformations of an object lead to locally affine distortions of the image, so here we consider the affine-invariant Hausdorff distance measure described in Section 3.3.

Figure 5 shows the result of using this affine-invariant distance measure. This defines a more clearly cyclic mapping of the bird images, emphasizing the variation in the non-rigid deformation of the bird shape by minimizing the image distance due to rigid variation. Furthermore, the solution for the best-fitting affine matrix A between two images offers an image warping operator for interpolating between images. This could form the basis for better "out-of-sample" inverse projections, and could be used to create more realistic image interpolation.

4.2 Cardiac MRI: Non-rigid Motion and Contrast Changes

MRI data is typified by large data sets which are often noisy. An image of a particular subject may vary for a number of reasons, including noise inherent

in the sensor itself, motion of the subject during data capture, and time-varying effects of contrast agents that are used to highlight particular types of tissue. These variations are difficult to parameterize in a very general setting, but for a particular subject, the images are likely to lie on a low dimensional manifold.

The direct application of Isomap (using sum of the squared pixel intensity differences and $k = 8$ neighbors) to a particular MR image set is shown in Figure 6. This image set contains real-time cardiac MR images, captured during a 60 ms window during the systolic part of consecutive heartbeats. The data set includes 180 such images from the same patient. The variation in these images has three causes. First, between images there is variation in the position of the heart (and liver, which is visible at the bottom of the images) due to the compression of the chest cavity during breathing. Second, a contrast agent is slowly permeating through the tissues. Third, the MRI images are noisy.

Distance functions that measure variation due to breathing motion and are invariant to the contrast changes or vice-versa — instead of the Euclidean image distance which varies due to both effects — give distance measures that are more isometric to the underlying manifold parameters. We use the pair of functions based on Gabor filter responses proposed in Sections 3.3 and 3.4. Embedding the images in one dimension using Isomap with these distance functions gives each image two coordinates. These coordinates are plotted at the right of Figure 7. The y-axis corresponds to the embedding based upon the Gabor filter phase difference (which measures local motions, but is largely invariant to contrast changes), and correlates to the different deformations of the chest cavity. The x-axis variation is based upon the Gabor filter magnitude change (which measures local contrast changes, but is largely invariant to small motions). Because the contrast change is due to a contrast agent permeating through the tissue, this is related to the original ordering of the data.

The images whose projections onto the y-axis are similar are taken at essentially the same part of the breathing cycle. A video sequence that plays the original images in the order they appear when projected onto the y-axis shows a very slow deformation, because the frames are ordered by what part of the breathing cycle they capture. Taking a window of 10 consecutive frames within this movie (all of which have similar y-axis projections) gives 10 images of the heart at the same part of the breathing cycle. These can be averaged together to de-noise the image without introduce motion blur. One frame of this sliding window average is shown at the bottom of figure 7.

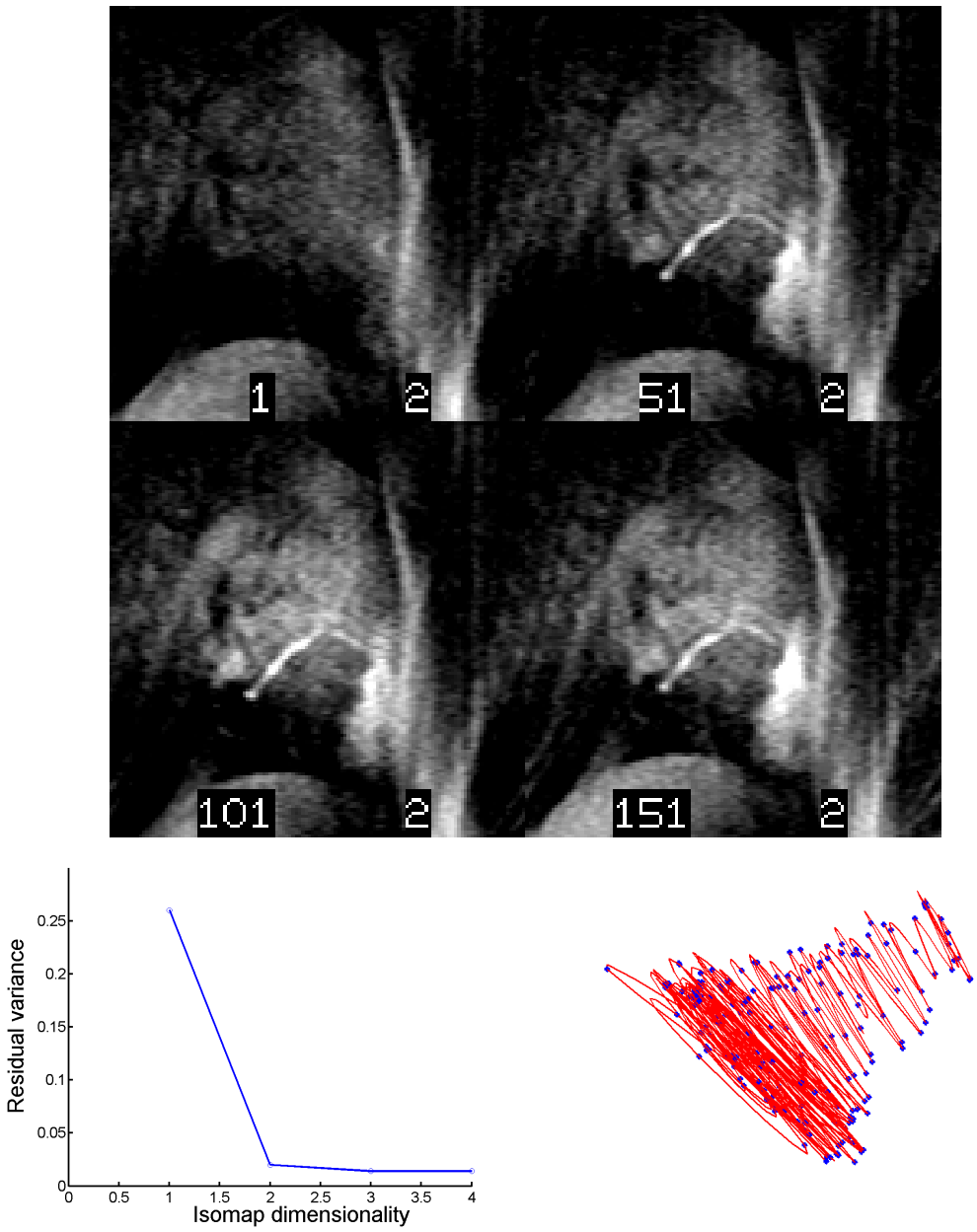

Fig. 6. Four samples of a sequence of MR images, and the associated Isomap embedding (using $k = 8$ neighbors.) The plot of Isomap dimension versus residual error indicates that 2 dimensions suffice to capture most of the distance information. The red line connects the images in order, the temporal sequence corresponds to variation due to a contrast agent permeating through the tissue.

5 Conclusion

Several concluding thoughts are in order. First, techniques such as Isomap and LLE are important tools in processing large video and image collections. These general statistical tools need to be specialized in order to take advantage of properties of natural images and deformations because image data

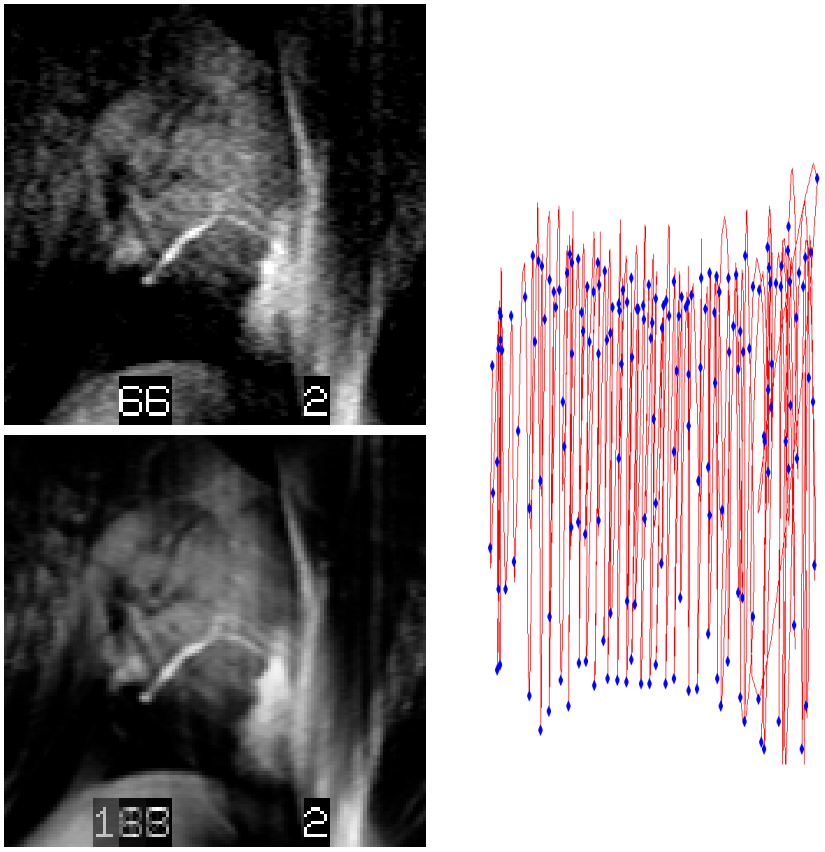

Fig. 7. (Right) Isomap embedding using a contrast invariant distance measure based on local Gabor phase. Note how the embedding aligns itself with two concrete degrees of freedom. (Top left) A sample image from the cardiac MR image set. (Bottom right) The result of averaging 10 images with similar y-component values. Since the y component encodes motion of the object in the image, averaging images with similar y-components does not result in spatial blurring, but does minimize pixel noise in individual images.

sets are (even locally) almost never linear combinations of other images. Finally, a small set of image transformation primitives gives powerful tools for registration of many different kinds of data sets.

References

[1] D. Thompson, On Growth and Form, Cambridge Universtiy Press, 1917.

[2] J. B. Tenenbaum, V. de Silva, J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (5500) (2000) 2319–2323.

[3] S. T. Roweis, L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (5500) (2000) 2323–2326.

[4] K. Q. Weinberger, L. K. Saul, Unsupervised learning of image manifolds by semidefinite programming, in: Computer Vision and Pattern Recognition, 2004.

[5] Y. W. Teh, S. Roweis, Automatic alignment of local representations, in: Advances in Neural Information Processing Systems, Vol. 15, 2003.

[6] D. L. Donoho, C. Grimes, Hessian eigenmaps: Locally linear embedding techniques for high-dimensional data, Proceedings of the National Academy of Science 100 (2003) 5591–5596.

[7] I. Borg, P. Groenen, Modern Multidimensional Scaling: Theory and Applications, Springer-Verlag, 1997.

[8] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, in: T. G. Dietterich, S. Becker, Z. Ghahramani (Eds.), Advances in Neural Information Processing Systems 14, MIT Press, Cambridge, MA, 2002, pp. 585–591.

[9] Y. Bengio, J. Paiement, P. Vincent, O. Delalleau, N. Le Roux, M. Ouimet, Out-of-sample extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering, in: S. Thrun, L. Saul, B. Schölkopf (Eds.), Advances in Neural Information Processing Systems 16, MIT Press, Cambridge, MA, 2004.

[10] R. Pless, I. Simon, Using thousands of images of an object, in: CVPRIP, 2002.

[11] I. S. Lim, P. H. Ciechomski, S. Sarni, D. Thalmann, Planar arrangement of high-dimensional biomedical data sets by isomap coordinates, in: Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems (CBMS 2003), New York, 2003.

[12] U. Grenander, Elements of Pattern Theory, Johns Hopkins University Press, Baltimore, 1996.

[13] A. Trouve, L. Younes, Local analysis of a shape manifold, Tech. Rep. 2002-03, Laboratoire d'Analyse, Geometrie et Applications, CNRS, Universite Paris (2002).

[14] A. Trouve, Diffeomorphism groups and pattern matching in image analysis, International Journal of Computer Vision 28 (1998) 213–221.