Isomap and Nonparametric Models of Image Deformation

Richard Souvenir and Robert Pless
Department of Computer Science and Engineering
Washington University in St. Louis
St. Louis, MO, 63117

Abstract

Isomap is an exemplar of a set of data driven non-linear dimensionality reduction techniques that have shown promise for the analysis of images and video. These methods parameterize each image as coordinates on a low-dimensional manifold, but, unlike PCA, the low dimensional parameters do not have an explicit meaning, and are not natural projection operators between the high and low-dimensional spaces. For the important special case of image sets of an unknown object undergoing an unknown deformation, we show that Isomap gives a valuable pre-processing step to find an ordering of the images in terms of their deformation. Using the continuity of deformation implied in the Isomap ordering allows more accurate solutions for a thin-plate spline deformation from a specific image to all others. This defines a mapping between the Isomap coordinates and a specific deformation, which is extensible to give projection functions between the image space and the Isomap space. Applications of this technique are shown for cardiac MRI images undergoing chest cavity deformation due to patient breathing.

1. Introduction

Understanding non-rigid deformations remains a challenging problem in the analysis of images. In this paper, we consider a special case of this problem: given many images of an object undergoing an unknown (but small) set of deformations, characterize the deformations and give each image parameters which describe how they have been deformed.

This problem is quite common in the medical imaging community, and one important application is cardiac MR imagery. Complete low-resolution images can be captured in modern MRI machines in about 60 ms. MR imagery taken of the same plane varies primarily due to three causes, deformation of the heart during the heartbeat, deformation of the heart due to breathing, and (potentially) contrast agents permeating slowly through the tissues.

Common practice for diagnostic cardiac MRI is to isolate the effects of heartbeat by triggering the MR image capture at the same part of the heartbeat cycle, and to isolate the effects of breathing by asking the patient to hold their breath. This leaves images which vary only because of the permeating contrast agent. In this paper we aim to alleviate the necessity that the patient hold their breath, by learning a model of the image deformation caused by breathing. This would allow cardiac MRI to be performed on unconscious patients unable to hold their breath.

A gated-cardiac image set consist of images taken at the same part of the heartbeat cycle. These images, for any specific patient, fall near (but because of image noise, not exactly on) a 2-D manifold within the space of images (which has as many dimensions as the image has pixels). This manifold has natural parameters, the breathing cycle and the permeation of the contrast agent. Our goal can be phrased as learning a mapping between the image set and these natural parameters, and then learning a model of the image deformation caused by breathing.

Because these image sets can be parameterized by a small number of parameters, reasoning about such image sets is an ideal application of dimensionality reduction. Since the manifold is not a linear subspace of the image space (the images are not linear combinations of a small number of basis images), linear dimensionality techniques such as PCA and ICA fail. However, there exists a class of techniques that attempt to discover the intrinsic low-dimensional parameterization of a point set. Isomap [12] is one exemplar of this class of non-linear dimensionality reduction (NLDR) algorithms. Isomap defines a mapping:

$$f : \mathcal{I} \to \mathbb{R}^k$$

where k is usually a small number such as 2 or 3, and is ideally the number of free parameters that were varied in creating the image data set. However, the meaning of these parameters is not defined in the Isomap procedure. In this paper we intend to approach this deficiency in Isomap for the specific case of image deformations. That is, for the case of images that come from a one-parameter set of deformations of a single object, we seek to extend to the function f to become:

$$f : \mathcal{I} \to \mathbb{R}^k \to \mathcal{D},$$

1

where the specific form of $\mathcal{D}$ depends on a choice of how to represent an image deformation.

The contributions of this paper are the demonstration of Isomap as a pre-processing tool and the intelligent selection of distance functions that define Isomap so that it can parse an image set and automatically parametrize each image by its magnitude of variation due to different causes. After the images are so parameterized, surprisingly simple and naive approaches may suffice to capture complex non-rigid deformations. Furthermore, the mapping of the Isomap parameterization onto a specific image deformation defines natural interpolation and projection models, lacking in classical Isomap.

2 Isomap

Isomap was introduced as a general tool for visualizing sociological data sets that may be derived from many different sources. Applied to images, the Isomap algorithm proceeds as follows:

Isomap: Given input images $\{X_1, X_2, \ldots, X_n\}$

1. Compute $d(i, j) = ||X_i - X_j||_2$

2. Create graph $G(V, E)$ with a node for each image, and an edge between pairs of images whose distance is small, and set the edge weight to that distance.

3. Compute all pairs shortest path distance on G. Let $d(i, j)$ be the length of the shortest path between node i and node j. Define distance matrix D such that $D(i, j) = d(i, j)^2$.

4. Use MDS (defined below) to embed D into the desired dimension.

MDS (Multi-Dimensional Scaling) Given $n \times n$ matrix D, such that D(i,j) is the desired squared distance from point i to point j.

1. Define $\tau = -HDH/2$, (H is the centering matrix, $H_{ij} = \delta_{ij} - \frac{1}{n}$.

2. The top k eigenvectors of matrix τ are the coordinates of the k dimensional embedding which whose pairwise point distances best match the given distance matrix.

The threshold in step 2 of the Isomap algorithm — the "small enough" distance — is often set so that each point has an edge to a small number (5-10) of its closest neighbors. Even more important for the case of images is the choice of distance function. Considering images as vectors in a high-dimensional space, the L_2-norm distance squared is equivalent to the sum of the squared pixel intensity differences. It has been noted recently that better image distance measures often significantly improve the performance

of Isomap on image sets [9]. The next section gives a framework for addressing what is the best distance for a given application.

3 Paired Image Distance Measures

Recently work has shown the value of choosing appropriate distance measures when using Isomap to parameterize image data [9]. The axes used to define the coordinates in a 2-D parameterization produced by Isomap are difficult to automatically interpret, so instead we seek to find a pair of distance measures matched to the causes of the image deformation such that the Isomap parameterizations using each distance measure correlates with only one of the causes of the deformation.

More formally, consider an image set $\mathcal{I}$ that is parameterized by two parameters (b, c) (which might stand for breathing and contrast), and suppose that $I(b, c)$ defines the noise-free image that would be generated for a particular part of the breathing cycle and contrast permeation. We would like to define a distance measures d_1, d_2 such that Isomap applied to $\mathcal{I}$ using d_1 gives a 1-D parameterization such that image $I(b, c)$ is mapped to the real number b, and Isomap applied to $\mathcal{I}$ using d_2 gives a 1-D parameterization such that image $I(b, c)$ is mapped to the real number c. In addition to being as insensitive to noise as possible, this requires:

- $d_1(I(b, c), I(b + \delta, c)) = \delta$

- $d_1(I(b, c), I(b, c + \epsilon)) = 0$

- $d_2(I(b, c), I(b + \delta, c)) = 0$

- $d_2(I(b, c), I(b, c + \epsilon)) = \epsilon$

That is, we require that each distance measure accurately measure small variations in the deformation parameter it is assigned to capture, and furthermore the distance measure should be invariant to the image changes caused by small changes in the other mode of deformation. Because only nearby points are used in the Isomap procedure, it is not required that the distance measures are globally invariant to the other deformation. Depending upon the application, weaker conditions may suffice, such as requiring the image distance to be monotonic with respect to (rather than equal to) an intrinsic parameterization of the deformation.

For the example case of a gated cardiac MRI data, there are two causes of motion, that manifest themselves locally as image motion and contrast change. The next section introduces a pair of image measures ideally suited to parsing the effects of these two causes.

3.1 Gabor Filter Bank Response Distances

The best distance measures relate directly to transformation groups that define the image change [**?**]. For the case of unknown deformations, the generic class of diffeomorphic deformations is a natural choice of transformation groups, and measuring local motions is an effective distance measure. Since Isomap only requires that local distance be computed, simple distance measure is possible.

The following defines a distance measure that uses the response of a collection of Gabor filters to estimate motions between images. 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.

Local Image Deformation Distance

1. Given images I_1, I_2

2. Define $G_{(\omega,V,H,\sigma)}$ to be the 2D complex Gabor filter with frequency ω, oriented either vertically or horizontally, with σ as the variance of the modulating Gaussian.

3.

$$
\begin{aligned}
D_{(\omega,\sigma)} \quad = \quad & \sum_{x,y} \Psi(G_{(\omega,V,\sigma)} \otimes I_1, G_{(\omega,V,\sigma)} \otimes I_2) \\
& + \Psi(G_{(\omega,H,\sigma)} \otimes I_1, G_{(\omega,H,\sigma)} \otimes I_2),
\end{aligned}
$$

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

This distance function is dependent upon 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 approximately the wavelength. The results shown in Section 5 are 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 region of the image, the effect of pixel noise is minimized.

Fortuitously, the complex Gabor filter responses can also be used to compute the second half of this pair of distance measures.

Local Image Contrast Distance

1. Given images I_1, I_2

2. Define $G_{(\omega,V,H,\sigma)}$ to be the 2D complex Gabor filter with frequency ω, oriented either vertically or horizontally, with σ as the variance of the modulating Gaussian.

3.

$$
\begin{aligned}
D_{(\omega,\sigma)} \quad = \quad & \sum_{x,y} \big| |G_{(\omega,V,\sigma)} \otimes I_1| - |G_{(\omega,V,\sigma)} \otimes I_2| \big| \\
& + \big| |G_{(\omega,H,\sigma)} \otimes I_1| - |G_{(\omega,H,\sigma)} \otimes I_2| \big|,
\end{aligned}
$$

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 affect the magnitude of the filter response, so this second distance measure also has the desired invariant properties. The following section discuss methods to use the Isomap parameterization of the image set, and in particular give methods for the analysis of a data set including image which have undergone an unknown spatial deformation. The advantage given by Isomap is that the magnitude of this deformation is known, and the images can be re-ordered by their deformation. (The supplementary material shows a video of a cardiac MRI image set remapped to order the images by their deformation).

These results derive from a gated cardiac MRI study, with 180 images taken from an unknown part of the patient's breathing cycle. Using this distance function in Isomap, as described in Section 2, effectively gives a 1-D parameterization of the image set, with the free parameter corresponding monotonically with the breathing cycle.

4 Extracting Deformation Groups

For the class of image sets generated by multiple examples of an object undergoing a non-rigid transform, we address the principal shortcoming of Isomap and other non-linear dimensionality reduction algorithms, namely the inability to extract meaning for the low-dimensional coordinates and perform an inverse projection from a point not in the original set to a new point on the image manifold. By using an appropriate distance measure, as described earlier, the images are sorted relative to their major deformation. In order to solve for the parameters of this deformation, our method takes the following steps:

- Select an appropriate distance measure (Sec. 3)

- Use Isomap to find an ordering for the images

- Find point correspondences between images

- Extend point correspondences into image warps

4.1 Point Tracking

The main benefit to sorting the points relative to the deformation instead of using unsorted images is that point tracking is simplified. In general, point tracking is easier if the putative corresponding points are closer together. Naive methods such as iterative closest point matching [1] can effectively track points through hundreds of frames. In this work we use a simple feature tracker [8] which makes an initial guess of point correspondences and uses RANSAC [6] to improve the solution.

4.2 Thin Plate Splines

A thin-plate spline [2] is a two-dimensional interpolation function whose name refers to a physical analogy involving the bending of a thin sheet of metal [5]. Given an arbitrary set of points in $\mathbb{R}^2$ and some function $f(x, y)$ evaluated at those points, the thin plate minimizes what is known as the "bending energy" function:

$$\int\int_{\mathbb{R}^2} (f_{xx}^2 + 2f_{xy}^2 + f_{yy}^2) dx dy$$

Thin-plate splines have been used frequently in image analysis. This construct has been used with velocity encoded MR images [11], to calculate cardiac strain from MR images [7], and analyzing bone structure on radiographs [3].

The pervasiveness of this construct in the image analysis domain indicates that thin-plate splines provide a natural way to move from point correspondences to entire image warps. Let $P_t(i)$ be the coordinate of the i-th tracked point in frame t. The thin plate spline warping function is the function f that minimizes the bending energy above and simultaneously maps all points $P_1(i)$ in the first frame exactly onto their corresponding points $P_t(i)$ in the t-th frame. That is,

$$\forall_i f(p_{1,i}) = p_{t,i}$$

, and for all image points (x, y) that were not tracked in the first image, the function f maps them in such a was that the overall mapping minimizing the distortion measured by the bending energy. Figure 1 shows a representation of a thin-plate spline capturing the image deformation for two cardiac MR images. Using the image distance measure described above, the images represented in Figures 1a and b had a high inter-image distance. The thin-plate spline overlayed on these images represents the deformation of one image to the other. Figure 1c shows the result of transforming the image in 2b to that in 2a.

5 Results and Summary

We applied our method to the analysis of MRI data. The image set we used is a "held breath" MRI of a heart. In
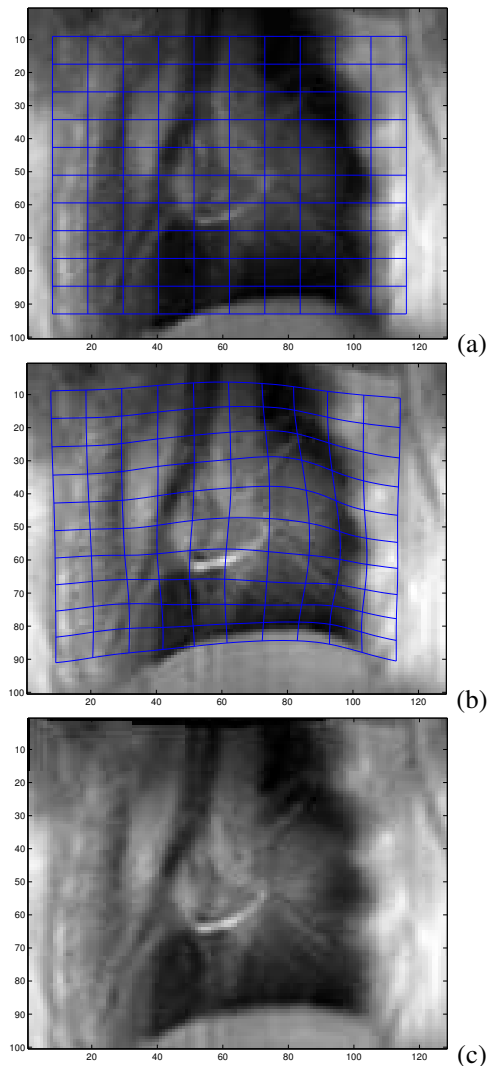

Figure 1: Using thin-plate splines to model deformations.

this experimental design, the patient is asked to hold their breath, and the MRI pulses are triggered at the same point in consecutive heart beats until enough pulses are captured to reconstruct an image. Each image is created in this way, and the data set includes 180 such images from the same patient. The variation in these images has three causes. First, the patient does not always hold their breath in exactly the same position, so between images there is variation in the position of the heart and liver (visible at the bottom of the images in Figure 1). Second, the contrast agent is slowly permeating through the tissues in view. Third, the MR images themselves have noise.

MRI data is typified by large data sets which are often noisy. In general, we can blur consecutive images in a video to remove image noise, however, this generally introduces

motion blur, which is a detriment to accurate point tracking. However, by reordering the images, the motion in between images is generally small enough that motion blur is not introduced. Figure 2 shows two consecutive images (after reordering using Isomap) and the result of blurring. In the original ordering, these are not proximal frames. In fact, due to the effects of the contrast agent used in the procedure, there is a difference in the average intensity of each image, which can clearly be seen in Figure 2b where the vasculature is more pronounced.

In a test to demonstrate the effectiveness of this method, we applied our method to the entire set of 180 MR images. Using the distance measure described in Section 3 and our procedure, as outlined in Section 4, we solved for the image deformation of each image relative to the first image in dataset. Figure 3 shows the results of capturing the unknown deformation with and without using the Isomap sorting step. In this experiment, point correspondences and image warps (to a reference frame) were applied to the set using the original ordering and the ordering obtained by Isomap sorting. *Mutual Information* was used to calculate how similar the warp frames were to the reference frame. Mutual Information [4, 13, 14] is a widely-used metric in medical image registration. [10]

Conclusions. This paper has illustrated the use of Isomap as a pre-processor for the analysis of deformable image sets where the deformation is unknown. The primary contribution was the introduction of pairs of image distances that made it possible to separate the effects of different causes of image variation. This allows the images to be parameterized by their unknown deformation, which facilitates several parts of the process of modeling the deformation, including image de-noising to make it easier to find and define feature points, and image re-ordering, so that feature points remain nearby in neighboring image. The general technique of combining non-linear dimensionality reduction with representations of image deformation provide the beginnings of a framework for the model-free understanding of non-rigid motions.

References

[1] P J Besl and N D McKay. A method for registration of 3-d shapes. *PAMI*, 14(2):239–256, February 1992.

[2] F L Bookstein. Principal warps: Thin plate splines and the decomposition of deformations. *Pattern Analysis and Machine Intelligence*, 11, June 1989.

[3] H P Chang, P H Liu, H F Change, and C H Chang. Thin-plate spine (tps) graphical analysis of the mandible on cephalometric tadiographs. *Dentomaxillofacial Radiology*, 31:137–141, 2002.
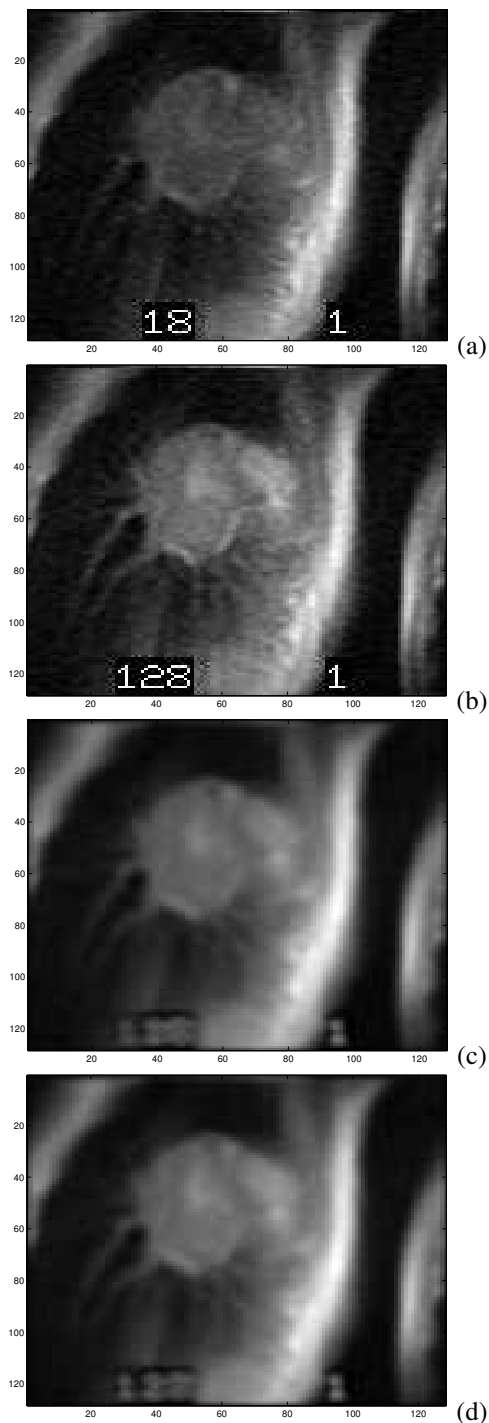
Figure 2: Reducing MR image noise by blurring with neighboring (in Isomap order) images. Images (a) and (b) are (originally) temporally distant and differ due to the effect of a contrast agent. Images (c) and (d) are the result of applying a blur using nearby (in Isomap order) images. Minimal motion blur is introduced using this method.
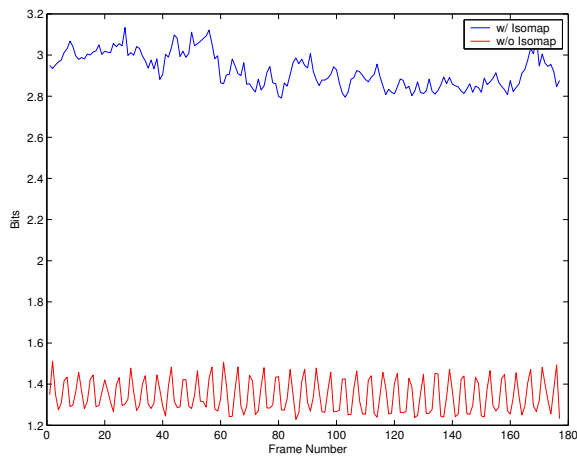
Figure 3: Mutual information (in bits) of frames in a video sequence warped to a reference image with and without the use of the Isomap sorting step.

[4] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multimodality image registration based on information theory. *Information Processing in Medical Imaging*, pages 263–274, 1995.

[5] Eric W. Weisstein et al. Thin plate spline. http://mathworld.wolfram.com/ThinPlateSpline.html.

[6] M A Fischler and R C Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. of the ACM*, 1981.

[7] Mark Griffin. Measuring cardiac strain using thin plate splines. In *Proc. Computational Techniques and Applications*, Brisbane, Australia, July 2001.

[8] Y Ma, S Soatto, J Kosecka, and S Sastry. *An Invitation to 3D Vision*, chapter 11. Spinger-Verlag, 2003.

[9] R Pless. Differential structure in non-linear image embedding functions. In *Articulated and Nonrigid Motion*, 2004.

[10] Josien P. W. Pluim, J.B. Antoine Maintz, and Max A. Viergever. Mutual information based registration of medical images: a survey. *IEEE Transactions on Medical Imaging*, 22(8):986–1004, August 2003.

[11] G.I. Sanchez-Ortiz, D. Rueckert, and P. Burger. Motion and deformation analysis of the heart using thin-plate splines and density and velocity encoded mr images. In *16th Statistical Workshop on Image Fusion and Shape Variability*, pages 71–78, Leeds, UK, July 1996.

[12] J. Tenebaum, V. Silva, and J. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, December 2000.

[13] P.A. Viola. *Alignment by maximization of mutual information*. PhD thesis, Massachusetts Institute of Technology, Boston, MA, 1995.

[14] W.M. Wells, P. Viola, and R. Kinkinis. Multi-modal volume registration by masimization of mutual information. *Medical Robotics and Computer Assisted Surgery*, pages 55–62, 1995.