IMAGE MANIFOLD INTERPOLATION USING FREE-FORM DEFORMATIONS

Richard Souvenir, Qilong Zhang, and Robert Pless

Department of Computer Science and Engineering
Washington University in St. Louis
St. Louis, MO, 63130 USA

ABSTRACT

An important class of image data sets depict an object undergoing deformation. When there are only a few underlying causes of the deformation, these images have a natural low-dimensional structure which can be parameterized using manifold learning. This paper presents a method to solve for the deformation field as a function of the manifold coordinates – implicitly optimizing the deformation between all pairs of images simultaneously. Additionally, we provide a mechanism to create images for arbitrary coordinates of the manifold, addressing an important limitation of manifold learning algorithms for the case of images related through deformations. We give quantitative results in an artificial image morphing example and illustrate the method by finding the deformations relating all images of a cardiopulmonary MR image sequence.

Index Terms— Image processing, Interpolation, Biomedical image processing

1. INTRODUCTION

Typical biomedical image sets, such as the sample heart MR frames shown in Figure 1, suffer from poor resolution both spatially and temporally. Improvements in the imaging apparatus and protocols have helped to mitigate these problems, but current diagnostic biomedical images and videos are still of low quality. In this work, we describe a method for biomedical image deformation analysis which incorporates all of the video data simultaneously in order to facilitate inter-image interpolation and de-noising. This method exploits the low-dimensional manifold structure of a set of images.

Image manifolds have been shown to be useful representation tools for such varied input as MRI sequences, facial images, and animation data. However, a key limitation to manifold analysis of images in many applications is the lack of projection functions from the low-dimensional embedding space back to image space for samples not in the original training set. There has been some work on this problem, in the general case, where linear interpolation of nearby images is performed using generalized radial basis functions [1]. However, these function approximation methods do not generalize to approximating image manifolds where the major cause of image change is non-rigid deformation. In this work we seek to integrate explicit models of image deformation with manifold representations of image data sets.

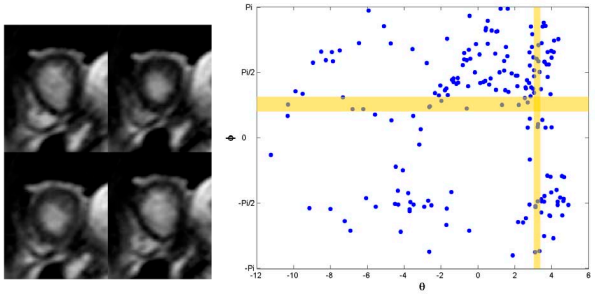

Fig. 1. Sample frames from a heart MR sequence and the 2D Isomap embedding from this sequence. Images with similar y-coordinates arise from similar parts of the heartbeat cycle; images with similar x-coordinates arise from similar parts of the breathing cycle. Our work characterizes image deformations in terms of these manifold coordinates.

In the next section, we review general methods for dimensionality reduction and some of their limitations. Section 3 describes models of image deformation and how to solve for them in the context of an image manifold. We then discuss interpolating images on this manifold and show results in in Section 5.

2. RELATED WORK

Classical dimensionality reduction techniques rely on Principal Component Analysis (PCA) [2] and Independent Component Analysis (ICA) [3]. These methods seek to represent data as linear combinations of a small number of basis vectors. However, natural image sets, specifically biomedical images, tend to vary due to variation in pose or shape deformations, which are very poorly approximated by changes in linear basis functions.

Such problems have led to a number of methods (typified by Isomap [4] and LLE [5]) which parameterize data sets drawn from low-dimensional nonlinear manifolds. These al-

gorithms work well for general manifold learning, but do not have the desirable quality of PCA for "back projection" analysis. These unsupervised algorithms return a mapping of input data points (or images) to low-dimensional coordinates rather than a function from the low-dimensional space to the original input (image) space. The next section introduces our method to solve this problem for a certain class of image manifolds by integrating manifold learning with models of local image deformation.

3. A MANIFOLD OF DEFORMATION FIELDS

We seek to build a parametric representation of the deformation fields which characterize the image variation on the manifold. In this section, we describe the free-form deformation (FFD) [6] model, then illustrate how to describe an image manifold using a parameterized model of image deformations. Finally we use this model to reconstruct images corresponding to arbitrary points on the manifold.

3.1. Deformation Model

Using the FFD model, the resulting deformations can be written as the 2D tensor product of standard one-dimensional cubic B-splines:

$$f(x,y) = \sum_{m=0}^{3}\sum_{n=0}^{3} B_m(\tilde{x})B_n(\tilde{y})\Phi_{m+i,n+j} \qquad (1)$$

where Φ denotes a $n_x \times n_y$ lattice of control points which parameterize the FFD, i, j denote the indices of the control points, and $\tilde{x}, \tilde{y}$ represent the relative positions of x, y in lattice coordinates (i.e. relative to the B-spline control point grid).

Commonly, when optimizing the FFD deformation to fit noisy image data, it is important to impose a smoothness term to prevent artifacts such as folding. A common smoothing term is:

$$\mathcal{S}[f] = \int \left[\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 f}{\partial xy}\right)^2 + \left(\frac{\partial^2 f}{\partial y^2}\right)^2 \right] dx\, dy \qquad (2)$$

Since global transformation can still be represented by a rigid or affine transformation of the control lattice of FFD, we simply use FFD to model the transformation f between two images.

3.2. Integrating Manifold Constraints

Manifold learning algorithms provide an automated tool to parameterize data sets, in our case, images, in terms of the variations introduced by a small number of causes. In this section, we propose a method to solve for deformation fields in all images of the data set simultaneously. This offers a powerful new constraint by penalizing variation in the FFD relative to nearby images in the manifold.

We illustrate the algorithm on an image set which varies due to two main causes. We use Isomap to automatically parameterize all images and obtain u and v, the 2D image manifold coordinates. The deformation fields on the manifold are also a function of u and v, and our goal is to describe these deformation fields explicitly by a set of FFD control points. Figure 1 shows example images and the Isomap embedding of a 100 frame cardiopulmonary image data set, for which the image changes are caused by the heart beating and breathing of the patient.

Given such an image data set $\mathcal{I} = \{\mathrm{I}_1, \mathrm{I}_2, \cdots, \mathrm{I}_N\}$ the transformation between any image I_i and the reference image I_{ref} is denoted by a transformation f_i. This transformation is the FFD transformation described in the last section and defined by the motion of $n_x \times n_y$ lattice of control points: $\Phi_i = \{\phi_{i,1},\ \phi_{i,2},\ \cdots,\ \phi_{i,n_x n_y}\}$. The set Φ contains all of the control points in all images, where $\phi_{i,j}$ is the j-th control point in the i-th image. However, image i is associated with manifold coordinates (u_i, v_i), so we can express these control points as a function of the manifold coordinates (u_i, v_i). We do this with another FFD to express the variation of the j-th control point as a function of the manifold coordinates (u, v). We parameterize this FFD with variables Θ_j, and there is one such FFD for each control point used to define the image warps.

Summarizing, we express the transformation of image i as a function $\mathcal{F}(u_i, v_i)$. In order to express a transformation, $\mathcal{F}$ describes the process of creating image warping FFD control points as $\{\Theta_1(u_i, v_i), \Theta_2(u_i, v_i), \ldots \Theta_{n_x n_y}(u_i, v_i)\}$. These control points define the warp for image i. The only free variables in this system are the parameters of the mapping Θ_j between manifold coordinates and the control point positions. Thus, to solve for the deformation fields, we minimize the following joint energy functional, over the set $\mathcal{I}$ of all N images, with respect to Θ (which affects the $\mathcal{F}$ term):

$$\mathbf{E}[\mathcal{I}, \mathrm{I}_{ref}; \mathcal{F}] = \sum_{i=1}^{N} \mathcal{D}[\mathrm{I}_i^{\mathcal{F}(u_i,v_i)}, \mathrm{I}_{ref}] + \lambda \mathcal{S}[\mathcal{F}(u_i, v_i)] \qquad (3)$$

where $\mathcal{D}$ measures the error between image I warped according to the manifold coordinates and the reference image (using SSD, correlation or mutual information), λ is a weighting parameter and the second term corresponds to the regularizer defined as:

$$\begin{aligned}\mathcal{S}[\mathcal{F}] = \sum_{i=1}^{N} \int \Bigg[&\left(\frac{\partial^2 \mathcal{F}(u_i,v_i)}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 \mathcal{F}(u_i,v_i)}{\partial xy}\right)^2 \\ &+ \left(\frac{\partial^2 \mathcal{F}(u_i,v_i)}{\partial y^2}\right)^2 \Bigg] dxdy + \gamma \sum_{j=1}^{n_x n_y} \int \Bigg[\left(\frac{\partial^2 \mathcal{F}_j}{\partial u^2}\right)^2 \\ &+2\left(\frac{\partial^2 \mathcal{F}_j}{\partial uv}\right)^2 + \left(\frac{\partial^2 \mathcal{F}_j}{\partial v^2}\right)^2 \Bigg] dudv \end{aligned} \qquad (4)$$

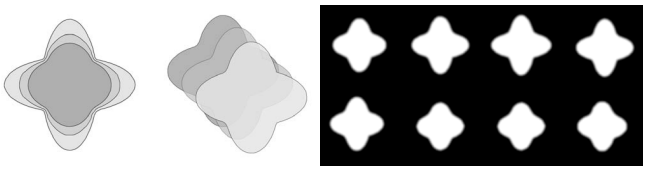

Fig. 2. An artificial data set constructed by composing two deformations (illustrated on the left), a non-rigid variation and a rigid translation. Eight sample frames are shown on the right

where γ is a weighting parameter. The regularizer (the corollary to Equation 2 in the single image case) ensures the smoothness of the deformation fields. The first term constrains the transformation between images to be smooth, while the second term penalizes the large local variations of control points Φ on manifold space.

The optimal deformation fields are found using gradient decent minimization of Equation 3 with respect to Θ. To avoid the high computation cost associated with the transformation complexity required to capture the deformation fields, one can use a multi-resolution approach [7] in which the resolution of the control points mesh Φ increases along with the image resolution, in a coarse-to-fine manner.

4. INTERPOLATION ON AN IMAGE MANIFOLD

Once manifold deformation fields $\mathcal{F}$ are obtained, we can calculate the transformation of any image on the manifold with respect to the reference image. Given two images I_p and I_q, as well as their associated transformation functions f_p and f_q respectively, we can approximate the transformation $f_{p,q}$ between these two images by computing the thin-plate spline [8] on image point position correspondences $\{f_p(x,y) \leftrightarrow f_q(x,y)\}$.

Now, given a query point p on the image manifold, we want to approximate its associated image I_p. The approximation $\hat{\mathrm{I}}_p$ can be obtained, in a straightforward manner, by transforming the image I_i closest on the manifold to position p, that is $\hat{\mathrm{I}}_p = \mathrm{I}_i^{f_{i,p}}$. To reduce the bias introduced by transforming a single image, one could consider $\hat{\mathrm{I}}_p$ as a weighted sum of images transformed from neighbors by using natural neighbor interpolation [9]

$$\hat{\mathrm{I}}_p = \sum \omega_i(p)\, \mathrm{I}_i^{f_{i,p}} \tag{5}$$

where $\omega_i(p)$ is the natural neighbor coordinate of image I_p with respect to the image I_i.

5. EXPERIMENTAL RESULTS

We examine our approach on artificially generated data, for which we control the deformation parameters. We also test this method on a noisy cardiopulmonary MRI sequence. All of the experiments were initialized with the same set of parameters. The computation of deformation fields is performed in a multi-resolution fashion by successive refinement of the FFD $\mathcal{F}(u_i, v_i)$ using control points resolutions of 4×4, 7×7 and finally 13×13. The resolution of the FFD $\mathcal{F}_j$ controlling image control points is kept as 5×5.

We generated an artificial 100-frame data set by defining a shape and deforming it using a non-rigid deformation and a rigid translation. Thus, this data set has a 2D manifold structure, indexed by the magnitude of each deformation. The two deformations and eight sample frames of this data set are depicted in Figure 2. After learning the image manifold coordinates, the images on the convex hull on the manifold were chosen as the training image set. The goal was to reconstruct the remaining images from the data set.

To assess the quality of the interpolation results, we calculated the mean and variance of the SSD between the ground truth and the reconstructed images:

$$SSD_i = \frac{1}{n}\sqrt{\sum (\mathrm{I}_{ref} - \mathrm{I}^{f_i})^2} \tag{6}$$

where n is the number of pixels on the image. Additionally, we compute the normalized mutual information (NMI) [10].

The large image changes make this a challenging data set for standard image interpolation techniques. The top row of Figure 3 gives results for direct natural neighbor interpolation on the image manifold. We show the results of using our method in the second row. The SSD and NMI measures shown in Tables 1 show the increased performance of our method versus the direct method.

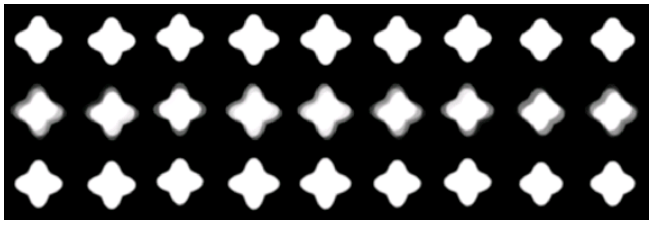

Fig. 3. Results on the artificial "star" data set. (1st row) Ground truth results. (2nd row) Results by directly interpolating between natural neighbors. (3rd row) Results using our method.

Method	SSD	NMI
Direct	0.90 ± 0.42	1.80 ± 0.048
Our Method	0.37 ± 0.12	1.91 ± 0.013

Table 1. Comparison for artificial "star" data set experiment between natural neighbor interpolation and our method. Our method improves both the SSD error measure, and the NMI similarity measure. Each cell shows the mean and variance over all of the images.

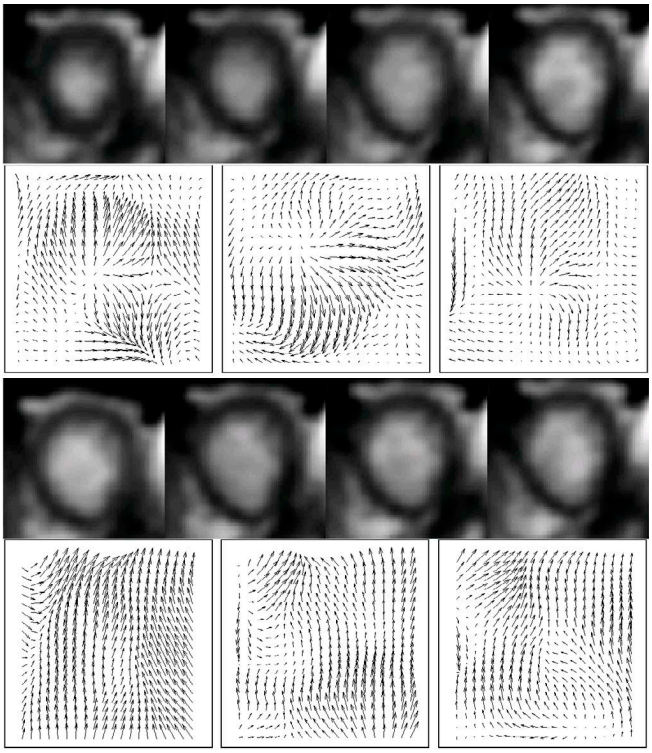

Fig. 4. Two examples of a series of images, and the deformation fields computed using the manifold constraints. The top row shows the expansion during a heart beat; the bottom row shows the dominant translation due to breathing.

We also tested our proposed algorithm on the heart MR sequence depicted in Figure 1. Figure 4 shows examples of two sets of deformation fields. The top half shows three deformation fields relating four images along a vertical strip of the cardiopulmonary image manifold which roughly corresponds to variations in the heartbeat cycle, but not the breathing cycle. The bottom half shows images and deformation fields corresponding to a images that are translated due to breathing. Finally, Figure 5 shows the results of interpolating an image for the manifold coordinates corresponding to a given image, illustrating its de-noising properties.

6. CONCLUSIONS

This paper integrates tools for manifold learning with standard models of image deformation. For image sets such as cardiopulmonary MR data, using this manifold structure regularizes the solution to the deformation fields relating all images. Better deformation fields offer diagnostic value in measuring heart volume and dynamics and support image de-noising. Furthermore, for the case of image manifolds which vary due to image deformation, this offers the ability to reconstruct images for arbitrary manifold positions, lifting an important limitation of nonlinear manifold learning algorithms.

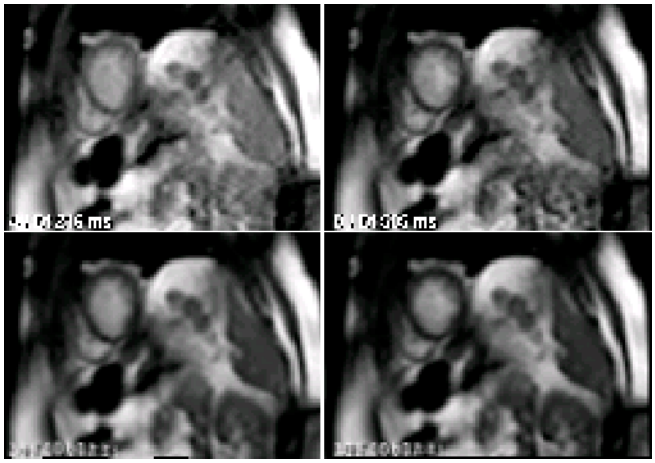

Fig. 5. An example of de-noising using learned deformation fields on the image manifold. On the top row are the original images and the bottom row shows the de-noising possible by correctly warping the nearby images.

7. REFERENCES

[1] A. Elgammal and Chan-Su Lee, "Separating style and content on a nonlinear manifold," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, June 2004, vol. 1, pp. 478–485.

[2] I T Jolliffe, *Principal Component Analysis*, Springer-Verlag, 1986.

[3] Aapo Hyvärinen, Juha Karhunen, and Erkki Oja, *Independent Component Analysis*, John Wiley and Sons, 2001.

[4] Joshua Tenebaum, Vin de Silva, and John Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, December 2000.

[5] Sam Roweis and Lawrence K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, pp. 2323–2326, December 2000.

[6] T. Sederberg and S. Parry, "Free-form deformation of solid geometric models," in *Proceedings of SIGGRAPH '86*, August 1996, pp. 151–160.

[7] M. Corvi and G. Nicchiotti, "Multiresolution image registration," in *Proc. International Conference on Image Processing (Vol. 3)-Volume 3*, Washington, DC, USA, 1995, p. 3224, IEEE Computer Society.

[8] F.L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. on Pattern Anaylsis and Machine Intelligence*, vol. 11, no. 6, pp. 567–585, June 1989.

[9] R. Sibson, "A brief description of natural neighbour interpolation," in *Interpreting Multivariate Data*, Vic Barnet, Ed., pp. 21–36. John Wiley and Sons, Chichester, 1981.

[10] C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3d medical image alignment," *Pattern Recognition*, vol. 32, pp. 71–85, 1999.