PCA vs. Tensor-Based Dimension Reduction Methods: An Empirical Comparison on Active Shape Models of Organs

Jiun-Hung Chen and Linda G. Shapiro

Abstract— How to model shape variations plays an important role in active shape models that is widely used in model-based medical image segmentation, and principal component analysis is a common approach for this task. Recently, different tensor-based dimension reduction methods have been proposed and have achieved better performances than PCA in face recognition. However, how they perform in modeling 3D shape variations of organs *in terms of reconstruction errors* in medical image analysis is still unclear.

In this paper, we propose to use tensor-based dimension reduction methods to model shape variations. We empirically compare two-dimensional principal component analysis, the parallel factor model and the Tucker decomposition with PCA in terms of the reconstruction errors. From our experimental results on several different organs such as livers, spleens and kidneys, 2DPCA performs best among the four compared methods, and the performance differences between 2DPCA and the other methods are statistically significant.

I. INTRODUCTION

Modeling shape variations is a significant step in active shape models [1] that is widely used in model-based medical image segmentation. A standard method for this step is principal component analysis (PCA). Unlike PCA that uses vector-based representations, varied tensor-based dimension reduction methods [2][3][4] have been recently proposed and achieved better performances than PCA in face recognition. In contrast with conventionally using a vector representation to represent a shape, tensor-based dimension reduction methods can represent a shape by a two-dimensional matrix directly or can represent the whole training set of shapes as a tensor [5]. For example, two-dimensional principal component analysis (2DPCA) [2] constructs the image covariance matrix directly by using the original image matrices without transforming them into 1D vectors and uses its eigenvectors as principal components. The parallel factor model (Parafac) [6][5] and the Tucker decomposition [7][5] are two major tensor decomposition methods that decompose a tensor into components.

However, we have not seen any work that has used tensor-based dimension reduction methods in medical image analysis except [8] that compared 2DPCA [2] with PCA on a normal/abnormal left ventricle shape classification task. In addition, in contrast with previous papers that mainly focus on classification, our work requires accurate 3D reconstructions of 3D organs whose shape can vary significantly.

In this paper, we propose to model shape variations with tensor-based dimension reduction methods. We report on

our empirical comparison of four reconstruction methods, including PCA, 2DPCA, Parafac and the Tucker decomposition, on several different organs such as livers, spleens and kidneys. From our experimental comparisons, 2DPCA achieves the best performance among the four compared methods and there are statistically significant differences between the performance of 2DPCA and those of the other methods.

II. METHODS

Assume that we have a training set of N 3D shapes and each shape is represented by M 3D landmark points. Conventionally, we can represent each such shape by a vector of $3M \times 1$.

A. PCA

The total scatter matrix $\mathbf{S}$ is defined as

$$\mathbf{S} = \sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})^t(\mathbf{x}_i - \bar{\mathbf{x}}) \tag{1}$$

where $\mathbf{x}_i$ is the i-th training shape vector and $\bar{\mathbf{x}}$ is the mean shape vector as defined below.

$$\bar{\mathbf{x}} = \frac{\sum_{i=1}^{N}\mathbf{x}_i}{N} \tag{2}$$

PCA finds a projection axis $\mathbf{b}$ that maximizes $\mathbf{b^tSb}$. Intuitively, the total scatter of the projected samples is maximized after the projection of a sample onto $\mathbf{b}$. The optimal L projection axes $\mathbf{b}_l, l = 1, \ldots, L$ that maximize the above criterion are the eigenvectors of $\mathbf{S}$ corresponding to the largest L eigenvalues[1]. For a shape vector $\mathbf{x}$, we can use its reconstruction $\tilde{\mathbf{x}}$ defined below to approximate it.

$$\tilde{\mathbf{x}} = \bar{\mathbf{x}} + \sum_{l=1}^{L} c_l \mathbf{b}_l \tag{3}$$

where $c_l = (\mathbf{x} - \bar{\mathbf{x}})^t \mathbf{b}_l$.

B. Tensor-Based Dimension Reduction Methods

In contrast with conventionally using a vector representation to represent a shape, tensor-based dimension reduction methods represent a shape by a two-dimensional matrix representation. In other words, let $\mathbf{X}$ be a $3 \times M$ matrix to represent a shape. In the following, we give a very brief introduction on tensors. For more details about tensors, please refer to [5]. A tensor is a generalization of vectors

Jiun-Hung Chen and Linda G. Shapiro are with Computer Science and Engineering, University of Washington, Seattle, WA 98195. {jhchen,shapiro}@cs.washington.edu

---

[1]To be consistent, we will use L in the following discussions to denote the number of components used in reconstructing a shape.

and matrices. The order (or mode) of a tensor is the number of dimensions. We use a third-order tensor $\mathbb{X} \in R^{3 \times M \times N}$ to represent the whole training set of N shapes where the first mode represents the x,y,z dimension of a point, the second mode represents the order of points and the third mode represents different patients. Although we focus on using third-order tensors in this paper, it is easy to extend our concepts to higher-order tensors. For example, if the above training set changed at regular intervals of time, then a fourth-order tensor in which the fourth order represents time can be used.

1) 2DPCA: 2DPCA [2] projects a shape matrix $\mathbf{X}$, which is a $3 \times M$ matrix onto a vector, $\mathbf{b}$, which is a $M \times 1$ vector, by the linear transformation.

$$\mathbf{c} = \mathbf{X}\mathbf{b} \tag{4}$$

The image scatter matrix $\mathbf{G}$ is defined as

$$\mathbf{G} = \sum_{i=1}^{N}(\mathbf{X}_i - \bar{\mathbf{X}})^t(\mathbf{X}_i - \bar{\mathbf{X}}) \tag{5}$$

where $\mathbf{X}_i$ is the shape matrix that represents the i-th training shape and

$$\bar{\mathbf{X}} = \frac{\sum_{i=1}^{N}\mathbf{X}_i}{N} \tag{6}$$

Similar to PCA, the goal of 2DPCA is to find a projection axis that maximizes $\mathbf{b}^t\mathbf{G}\mathbf{b}$. The optimal L projection axes $\mathbf{b}_l, l = 1, \ldots, L$ that maximize the above criterion are the eigenvectors of $\mathbf{G}$ corresponding to the largest L eigenvalues. For a shape matrix $\mathbf{X}$, we can use its reconstruction $\tilde{\mathbf{X}}$ defined below to approximate it.

$$\tilde{\mathbf{X}} = \bar{\mathbf{X}} + \sum_{l=1}^{L}\mathbf{c}_l\mathbf{b}_l{}^t \tag{7}$$

where $\mathbf{c}_l = (\mathbf{X} - \bar{\mathbf{X}})\mathbf{b}_l$.

2) Parallel Factor Model: Parafac [6][5] factorizes a tensor into a weighted sum of component rank-one tensors [5]. In other words, given a tensor $\mathbb{X} \in R^{I \times J \times K}$, Parafac decomposes it as

$$\mathbb{X} \approx \sum_{l=1}^{L}\lambda_l\mathbf{a}_l \circ \mathbf{b}_l \circ \mathbf{c}_l \tag{8}$$

where $\mathbf{a}_l \in R^{I \times 1}$, $\mathbf{b}_l \in R^{J \times 1}$, $\mathbf{c}_l \in R^{K \times 1}$, $\lambda_l \in R^1$ for $l = 1, \ldots, L$ and $\circ$ represents the vector outer product [5]. The alternating least squares (ALS) method [6][5] is commonly used to find the Parafac decomposition.

After the decomposition is computed, for a test shape, different methods [9][10] can be used to find the associated coefficient vectors and to compute the reconstruction that approximates it. In this paper, we follow the linear projection method in [9]. Given a shape matrix $\mathbf{X}$, we calculate its reconstruction $\tilde{\mathbf{X}} = \sum_{l=1}^{L} c_l\lambda_l\mathbf{a}_l \circ \mathbf{b}_l$ to approximate it by solving the following equation.

$$\min_{\tilde{\mathbf{X}}} ||\mathbf{X} - \tilde{\mathbf{X}}|| \tag{9}$$

where $||\mathbf{X}||$ is the Frobenius norm of $\mathbf{X}$.

(a) Livers   (b) Left kidneys
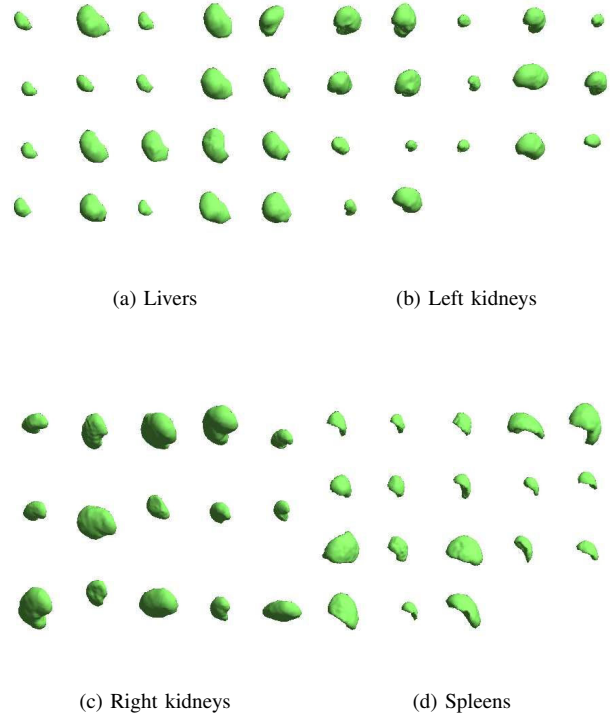
(c) Right kidneys   (d) Spleens

Fig. 1: The 3D triangular meshes of different organs we use in the experiments.

3) Tucker Decomposition: In contrast with Parafac, which decomposes a tensor into rank-one tensors, the Tucker decomposition is a form of higher-order principal component analysis that decomposes a tensor into a core tensor multiplied by a matrix along each mode [5]. Given a tensor $\mathbb{X} \in R^{I \times J \times K}$, the Tucker decomposition is given by

$$\mathbb{X} \approx \mathbb{G} \times_1 \mathbf{A} \times_2 \mathbf{B} \times_3 \mathbf{C} \tag{10}$$

$$= \sum_{p=1}^{P}\sum_{q=1}^{Q}\sum_{r=1}^{R}g_{pqr}\mathbf{a}_p \circ \mathbf{b}_q \circ \mathbf{c}_r \tag{11}$$

where $\mathbb{G} \in R^{P \times Q \times R}$ is called the core tensor, $\mathbf{A} \in R^{I \times P}$, $\mathbf{B} \in R^{J \times Q}$, $\mathbf{C} \in R^{K \times R}$, $\mathbf{a}_p \in R^{I \times 1}$ is the p-th column in $\mathbf{A}$, $\mathbf{b}_q \in R^{J \times 1}$ is the q-th column in $\mathbf{B}$, $\mathbf{c}_r \in R^{K \times 1}$ is the r-th column in $\mathbf{C}$ and $\times_n$ is the n-mode matrix product operator for multiplying a tensor by a matrix in mode n [5]. ALS can be used to find the Tucker decomposition.

Let $\mathbf{V}_{(3)}$ be the matrix formed by mode-n matricizing [5] the tensor $\mathbb{G} \times_1 \mathbf{A} \times_2 \mathbf{B}$ with respect to the third mode. Based on the above linear projection idea [9], given a shape vector $\mathbf{x}$, we calculate its reconstruction $\tilde{\mathbf{x}} = \sum_{l=1}^{L} c_l\mathbf{v}_l$ where $\mathbf{v}_l$ is the l-th column of $\mathbf{V}_{(3)}$ to approximate it by solving the following equation.

$$\min_{\tilde{\mathbf{x}}} ||\mathbf{x} - \tilde{\mathbf{x}}|| \tag{12}$$

III. EXPERIMENTAL RESULTS AND DISCUSSIONS

We have 3D mesh models of 20 livers, 17 left kidneys, 15 right kidneys, and 18 spleens as shown in Figure 1. All these 3D triangular meshes are constructed from CT scans of different patients and the 3D point correspondence problems among different 3D mesh models of the organs are solved[2]. All the mesh models of the same organ have the same number of vertices (2563) and the same number of faces (5120), and all vertices are used as landmarks to represent shapes. The tensors are constructed in the similar manner described in [9][10]. The task is to measure the difference between an unknown shape model and its reconstruction. We use leave-one-out cross validation to determine how accurately an algorithm will be able to predict data that it was not trained on. In order to reflect different definitions of shape similarities, two different commonly used metrics, Euclidean distances (i.e, the sum of the distances between all pairs of corresponding landmarks) and Hausdorff distance[3], are used to measure the shape difference between two shapes. We use the tensor library [12] for tensor decomposition. In addition, we further investigate how the numbers of components, L, affect the reconstruction errors.

The experimental results are shown in Figures 2 and 3. From these figures, we can see that 2DPCA performs best among the four methods, the Tucker decomposition performs slightly better than PCA and the Parafac decomposition is the worst. In fact, the performance differences between 2DPCA and other methods are statistically significant[4]. In addition, as the numbers of components increase, the reconstruction errors of all methods and the differences among different methods decrease. While 2DPCA and PCA had similar computation time in our experiments, the Parafac method and the Tucker decomposition were an order of magnitude slower.

We attribute the poor performance of the Parafac method to the limited expressiveness of rank-one tensors and vector outer products. Our results showing that the performance of 2DPCA is better PCA are in line with those reported for face recognition [2]. As explained in [2], because the dimension of the image scatter matrix is much smaller than that of the total scatter matrix, and we deal with small sample size problems where the number of training examples is much smaller than the dimensions of a shape, 2DPCA can capture

---

[2]We constructs the shape of an organ from manual segmentation of CT scans of a patient by using marching cubes in ITK-SNAP and then uses [11] to find correspondences among shape models.

[3]Given a pair of two 3D point sets, A and B, the Hausdorff distance between A and B is given by:

$$H(A, B) = \max\{\min_{p \in A} \min_{q \in B} d(p, q), \min_{q \in B} \min_{p \in A} d(p, q)\} \quad (13)$$

where $d(p, q)$ is the Euclidean distance between two 3D points, p and q. To compare a pair of two shapes, the Hausdorff distance between the two vertex sets of this given pair of shapes is computed.

[4]Given a method and a number of eigenvectors, a group of the reconstruction errors collected from each leave-one-out test was created. For a pair of two such groups, a standard two-sample t-test procedure was performed for comparing it. If the p-values is less than the reference probability, the result is statistically significant.
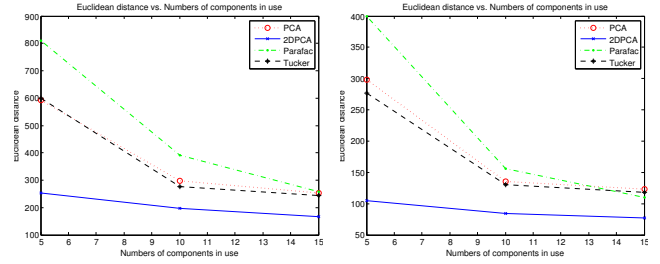
(a) Livers (Euclidean distance)　(b) Left kidneys (Euclidean distance)
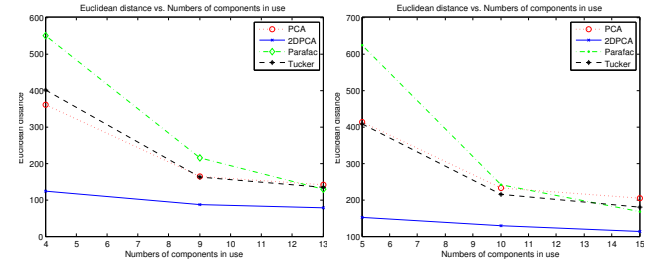
(c) Right kidneys (Euclidean distance)　(d) Spleens (Euclidean distance)

Fig. 2: Reconstruction errors in terms of Euclidean distances for different organs.

more accurate covariance information than PCA. Another possible reason is that the spatial dependencies between x, y and z are better preserved in 2DPCA.

Although the Tucker decomposition is a form of higher-order principal component analysis [5], its marginal improvement over PCA may be ascribed to the use of the linear projection method to compute the reconstruction, as the potential power of the tensor may be lost in this step. In contrast with the better performance gain from the Tucker decomposition to PCA reported in [3][9][10], a possible reason for the marginal improvement over PCA may be that the number of training examples used in this paper is much smaller than those used in these papers.

IV. CONCLUSIONS AND FUTURE WORK

In this paper, we propose to use different tensor-based dimension reduction methods to model shape variations. From our empirical comparisons of the reconstruction errors, 2DPCA is the best among the four compared methods and the performance differences between 2DPCA and the other methods are statistically significant. While we focus on modeling shape variations, it is easy to apply our ideas to different applications such as medical image classification and medical image retrieval and to different data formats from varied imaging modalities such as PET, MRI, ultrasound and CT scans.

We are currently applying learned models from these tensor-based dimension reduction methods to model-based
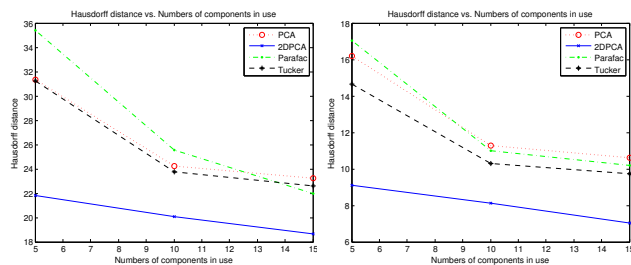
(a) Livers (Hausdorff distance) (b) Left kidneys (Hausdorff distance)
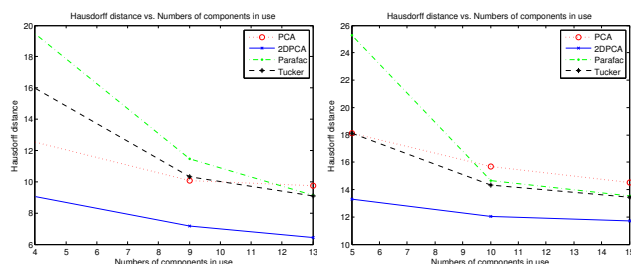
(c) Right kidneys (Hausdorff distance) (d) Spleens (Hausdorff distance)

Fig. 3: Reconstruction errors in terms of Hausdorff distances for different organs.

organ segmentations. We will also investigate different ways [13][10][4][5] to further improve 2DPCA and tensor decompositions. For example, the grid-sampling strategy [13] is proposed to further improve the performance of 2DPCA. In addition, in contrast to using the linear projection method, a multilinear method that can simultaneously infer the coefficient vectors in different modes [10] can be used to find better reconstructions.

V. ACKNOWLEDGMENT

REFERENCES

[1] G. J. E. T. F. Cootes and C. J. Taylor, "Active appearance models," *IEEE Transactions on PAMI*, vol. 23, pp. 681–685, Jan. 2001.

[2] J. Yang, D. Zhang, A. F. Frangi, and J. yu Yang, "Two-dimensional pca: A new approach to appearance-based face representation and recognition," *IEEE Transactions on PAMI*, vol. 26, no. 1, pp. 131–137, 2004.

[3] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear analysis of image ensembles: Tensorfaces," in *ECCV '02*. London, UK: Springer-Verlag, 2002, pp. 447–460.

[4] H. Wang and N. Ahuja, "A tensor approximation approach to dimensionality reduction," *Int. J. Comput. Vision*, vol. 76, no. 3, pp. 217–229, 2008.

[5] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, to appear (accepted June 2008).

[6] R. A. Harshman, "Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multi-modal factor analysis," *UCLA Working Papers in Phonetics*, vol. 16, no. 1, p. 84, 1970.

[7] L. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, September 1966. [Online]. Available: http://ideas.repec.org/a/spr/psycho/v31y1966i3p279-311.html

[8] H. Huang, F. Makedon, and R. McColl, "High dimensional statistical shape model for medical image analysis," *ISBI'08*, pp. 1541–1544, May 2008.

[9] M. Vasilescu and D. Terzopoulos, "Multilinear image analysis for facial recognition," *ICPR'02*, vol. 2, pp. 511–514 vol.2, 2002.

[10] ——, "Multilinear independent components analysis," *CVPR'05*, vol. 1, pp. 547–553 vol. 1, June 2005.

[11] T. Heimann, I. Oguz, I. Wolf, M. Styner, and H. M. H, "Implementing the automatic generation of 3d statistical shape models with itk," in *Open Science Workshop at MICCAII*, 2006.

[12] B. W. Bader and T. G. Kolda, "Algorithm 862: Matlab tensor classes for fast algorithm prototyping," *ACM Trans. Math. Softw.*, vol. 32, no. 4, pp. 635–653, December 2006. [Online]. Available: http://dx.doi.org/10.1145/1186785.1186794

[13] S. Shan, B. Cao, Y. Su, L. Qing, X. Chen, and W. Gao, "Unified principal component analysis with generalized covariance matrix for face recognition," *CVPR'08*, pp. 1–7, June 2008.