Mode-kn Factor Analysis for Image Ensembles

Shuicheng Yan[1], Huan Wang[2], Jilin Tu[1], Xiaoou Tang[2] and Thomas S. Huang[1]

[1]Beckman Institute, University of Illinois at Urbana-Champaign, USA

{scyan,jltu,huang}@ifp.uiuc.edu

[2]Department of Information Engineering, Chinese University of Hong Kong, Hong Kong

{hwang5,xtang}@ie.cuhk.edu

Abstract

To obtain closed-form solution for estimating extra-factors, we present in this paper a statistical learning technique called *mode-kn Factor Analysis* for exploiting image ensembles, expressed as an n-th order tensor with the n-th dimension characterizing *intra-features* while other dimensions representing *extra-factors*. In the learning stage, for the k-th dimension of the tensor, the mode-kn patterns are constructed by concatenating the intra-features dimension and the k-th extra-factor dimension, and then a mode-kn factor analysis model is built based on the mode-kn patterns unfolded from the original tensor. In the inference stage, for a new image, the mode classification of the k-th dimension is performed within a probabilistic framework. The advantages of mode-kn factor analysis over conventional tensor analysis algorithms are two-fold: 1) a closed-form solution, instead of iterative sub-optimal solution as conventionally obtained, is achieved for estimating the extra-factor modes of new data; and 2) the classification capability is enhanced by interacting with the process of synthesizing data of all other modes in the k-th dimension. Experiments on the Pointing'04 and CMU PIE databases for pose and illumination estimation both validate the superiority of the proposed algorithm over conventional algorithms for tensor analysis.

I. INTRODUCTION

Computer vision research has witnessed an increasing interest in tensor-based learning techniques for image analysis [11][8]. These techniques can be roughly divided into two categories. The first category, e.g., TensorFaces [12], Multilinear Independent Component Analysis (ICA) [14] and Wang's work in [15], describes each image as a concatenated vector, and various images with different extra-factors (such as identity, pose and illumination) finally constitute a high order tensor for further analysis. The second category instead considers the original or processed image as a 2nd or high order tensor, and its integration with conventional learning techniques elicits a large set of tensor counterparts, such as Generalized Low Rank Approximations of Matrices [17], Two-dimensional Linear Discriminant Analysis [18], and Discriminant Analysis based on Tensor Representation [16].

The work presented in this paper is motivated from several observations from the first category of tensor-based learning algorithms: 1) most algorithms are enlightened from the algebraic point of view instead of statistical learning, and the learning as well as inference stages are based on certain image reconstruction criteria; 2) the mode matrices along the extra-factor dimensions often do not contain explicit semantics; and 3) in the inference stage, the estimation of the extra-factors often results in an optimization problem without a closed-form solution, and only an iterative sub-optimal solution is available.

In this paper, we present a statistical framework for modeling and analyzing the tensor-formed image ensemble, and consequently offer a closed-form solution for estimating the extra-factors of new data. All of these are achieved by the following three steps. First, we propose a novel tensor unfolding approach, called mode-kn unfolding of the n-th order tensor with the assumption that the n-th dimension of the data tensor characterizes the underlying features. The mode-kn unfolding operator unfolds a tensor into a matrix whose row number is the product of the sizes of the k-th and n-th dimensions, while the column number is the product of the sizes of other dimensions. A column vector of the unfolded matrix concatenates all the features in different modes along the k-th extra-factor dimension. Then, Factor Analysis [10] [3] [6] is performed on the unfolded matrix by considering each column vector as a new sample. Finally, for a new datum, a probabilistic inference method for handling incomplete data is applied for determining the mode at the k-th dimension, and the solution is obtained in a closed-form manner.

The rest of the paper is organized as follows. Section 2 gives a brief review of the related works on tensor analysis and the consequent motivations. We introduce the details of the mode-kn unfolding of a tensor and the mode-kn factor analysis in Section 3. Section 4 demonstrates a set of validating experiments and we conclude this paper in Section 5.

II. PRELIMINARIES AND MOTIVATIONS

A. Tensor Algebra

A *tensor* is a higher order generalization of a matrix. In this subsection, we first briefly introduce some concepts and operators on tensor algebra.

The *mode-k product* operation of a high order tensor $\mathcal{Y} \in \mathbb{R}^{m_1 \times \ldots m_{k-1} \times m'_k \times m_{k+1} \times \ldots \times m_n}$ and a matrix $U^k \in \mathbb{R}^{m_k \times m'_k}$ produces another tensor $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times \ldots \times m_n}$, denoted as $\mathcal{Y} \times_k U^k$, that can be calculated as

$$\mathcal{X}_{i_1, i_2, \ldots, i_n} = \sum_{j_k=1}^{m'_k} \mathcal{Y}_{i_1, \ldots, i_{k-1}, j_k, i_{k+1}, \cdots, i_n} U^k_{i_k, j_k}. \tag{1}$$

For a tensor $\mathcal{X}$, a *mode-k pattern* is an m_k-dimensional vector formed by varying index i_k with all other indices fixed. We can arrange all mode-k patterns in order of indices into an $m_k \times (m_1 \times \cdots \times m_{k-1} \times m_{k+1} \times \cdots \times m_n)$ matrix, which is called *mode-k flattened matrix* of $\mathcal{X}$, denoted as X^k. The flattened matrices and tensor product

take on an important relation:

$$\mathcal{X} = \mathcal{Y} \times_k U^k \Leftrightarrow X^k = U^k Y^k. \tag{2}$$

The matrix operator Singular Value Decomposition (SVD) also has a higher-order generalization, *n-Mode SVD*, which expresses a tensor $\mathcal{X}$ as follows:

$$\mathcal{X} = \mathcal{D} \times_1 U^1 \times_2 U^2 \times_3 \ldots \times_n U^n \tag{3}$$

where $\mathcal{D} \in \mathbb{R}^{m_1' \times m_2' \times \ldots \times m_n'}$ is a *core tensor* coordinating the interaction of *mode matrices*: $U^1, U^2, \ldots, U^n$.

B. Related Works

Constituting a high order tensor with the image ensemble of different subjects at different expressions, viewpoints, and illuminations has been very popular in the past decade. In this paper, the n-th dimension of an n-th tensor is assumed to represent the underlying features for each image while other dimensions are for extra-factors such as expressions, viewpoints, and illuminations. The pioneering work of *TensorFaces* was contributed by Vasilescu et al. [13]. TensorFaces utilizes the mode matrices $U^1, U^2, \ldots, U^n$ for characterizing the variations of different factors. It has been successfully applied for face recognition with considerable improvement compared with the unsupervised algorithm Principal Component Analysis [2]. A further extension of TensorFaces to iteratively refine these mode matrices for better reconstruction performance was demonstrated in [12]. Vasilescu et al. [14] also extended the TensorFaces algorithm to Multilinear Independent Component Analysis (ICA) and utilized the ICA technique for computing the mode matrices along different tensor dimensions.

Wang et al. [15] proposed to analyze and synthesize human expressions by using the core tensor $\mathcal{D}$. For a new image datum x, the expression vector denoted as v_1 and identity vector denoted as v_2 are estimated by minimizing

$$\|x - \mathcal{D} \times_1 v_1^T \times_2 v_2^T \times_3 U^3\|, \tag{4}$$

where U^3 is assumed to be the mode matrix for the underlying image features. The vectors v_1 and v_2 are obtained by iteratively optimizing one vector with the other fixed. Vasilescu et al. [14] presented a more elegant method for such a purpose. First, the vectors v_1 and v_2 are combined into one vector via the Kronecker product, and then the combined vector can be obtained in closed-form manner; finally the the vectors v_1 and v_2 are derived by a rank-1 factorization approach. But the derived solution cannot be guaranteed to be globally optimal. The work in [8] and [11] demonstrated the application of a mode-k flattening based tensor decomposition method for segmentation and motion analysis, respectively.

C. Motivations

This work is motivated from the following observations on conventional tensor analysis techniques:

1) The computed mode matrices from the extra-factor dimensions lack explicit semantics and are of no use in the subsequent classification process. For example, for the expression mode matrix denoted as U^1, commonly it was believed that its column vectors characterize the expression variations, but one mode-1 pattern consists of the image gray levels of different expressions only on a certain point within the image plane, and hence cannot sufficiently characterize the expression variations. Moreover, for the subsequent classification of identity or expression, commonly no dimensionality reduction was conducted for identity and expression dimensions since the sizes in these dimensions are already small. As in Eqn. (4), if we set $v_1 = U^1 v_1'$, the estimation of expression will be independent of the mode matrix U^1 yet without sacrificing any accuracy.

2) The classification of identity, pose or other extra-factors is performed by minimizing the reconstruction error, which is intuitively illuminated from the algebraic point of view, and a more theoretical solution is desired for the whole classification process. Moreover, the estimation of the identity and expression vectors suffers from the local optimum issue in conventional tensor analysis techniques [15][14].

In this paper, we present a statistical framework for tensor analysis, and the main targets and the expected superiorities over conventional algorithms are: 1) modeling the extra-factors in a way easy for visualization and understanding; 2) inferring extra-factors with closed-form solutions; and 3) conducting the classification and synthesis of the modes at certain dimensions simultaneously.

III. Mode-kn Factor Analysis

In this section, we introduce our new solution to modeling the extra-factors of the tensor-formed data set. By following the terminology in Section 2, the data tensor is denoted as $\mathcal{X}$, and the n-th dimension of the tensor $\mathcal{X}$ represents the underlying features extracted from the images. The following subsections will introduce the learning and inference stages of our statistical framework for tensor analysis.

A. Learning Stage

The learning stage of our statistical framework includes two sub-steps: mode-kn flattening of an n-th order tensor and mode-kn factor analysis.

1) Mode-kn Pattern and Flattening: Mode-k flattening of a tensor is widely used for the subsequent n-mode SVD of a tensor. In this section, we present a novel flattening approach, called mode-kn flattening, to transform a tensor into a matrix. For the mode-kn flattening, the extra-factor dimension k and the feature dimension n of a tensor are combined to form a mode-kn pattern for subsequent analysis, and all the mode-kn patterns constitute

the so-called mode-kn flattened matrix $X^{kn} \in \mathbb{R}^{(m_k \times m_n) \times (m_1 \times \cdots \times m_{k-1} \times m_{k+1} \ldots \times m_{n-1})}$,

$$X_{ij}^{kn} = \mathcal{X}_{i_1, i_2, \cdots, i_n}, \tag{5}$$

where $i = (i_k - 1) \times m_n + i_n$ and $j = 1 + \sum_{l=1, l \neq k}^{n-1} (i_l - 1) \prod_{o=l+1, o \neq k}^{n-1} m_o$.

Each column vector of matrix X^{kn} is a single mode-kn pattern, which has explicit semantics. For example, if the 4-th order tensor $\mathcal{X}$ encodes the images of different subjects (dim-1) at different poses (dim-2) and illuminations (dim-3), a column vector of the mode-$(2,4)$ flattened matrix $X^{(2,4)}$ means the concatenated image features of different poses for a certain subject with certain illumination. The corresponding mode-$(2,4)$ pattern characterizes the correlation among the images of different poses. Owing to this explicit semantics and its sufficiently large row number of the flattened matrix, we can easily build a statistical model by taking the mode-kn pattern as a new object to be analyzed.

2) Mode-kn Factor Analysis: Factor analysis (FA) seeks to reveal the relationship between an observed vector variable and a latent variable of reduced dimensionality. It has been widely used for many applications involving high-dimensional observed data.

Given a $(m_k \times m_n)$-dimensional observation x, factor analysis tries to explain x by relating it to a q-dimensional latent variable z as:

$$x = \mu + Wz + e, \tag{6}$$

where W is a $(m_k \times m_n) \times q$ matrix, μ is the mean, and e is assumed to be white noise, namely $e \sim N(0, \delta^2 I)$, here, where I is a $q \times q$ identity matrix. The variable z is conventionally assumed to follow a Gaussian distribution $N(0, I)$, and then we have

$$x|z \quad \sim N(0, \delta^2 I), \tag{7}$$

$$x \quad \sim N(\mu, WW^T + \delta^2 I). \tag{8}$$

In this work, we take the mode-kn patterns as observed vectors, denoted as $\{X_1^{kn}, X_2^{kn}, \cdots, X_{\tilde{m}}^{kn}\}$, where $\tilde{m} = m_1 \times \cdots \times m_{k-1} \times m_{k+1} \times \cdots m_{n-1}$ and X_i^{kn} is the i-th column vector of the matrix X^{kn}, namely the i-th mode-kn pattern. Then the parameters of the above model can be inferred by maximizing the likelihood probability,

$$p(X_1^{kn}, \cdots, X_{\tilde{m}}^{kn} | \mu, W, \delta) = \prod_{i=1}^{\tilde{m}} p(X_i^{kn} | \mu, W, \delta). \tag{9}$$

Several parameter estimation methods have been proposed for the factor analysis model [1], [5], and commonly they can guarantee a numerically correct and stable solution.
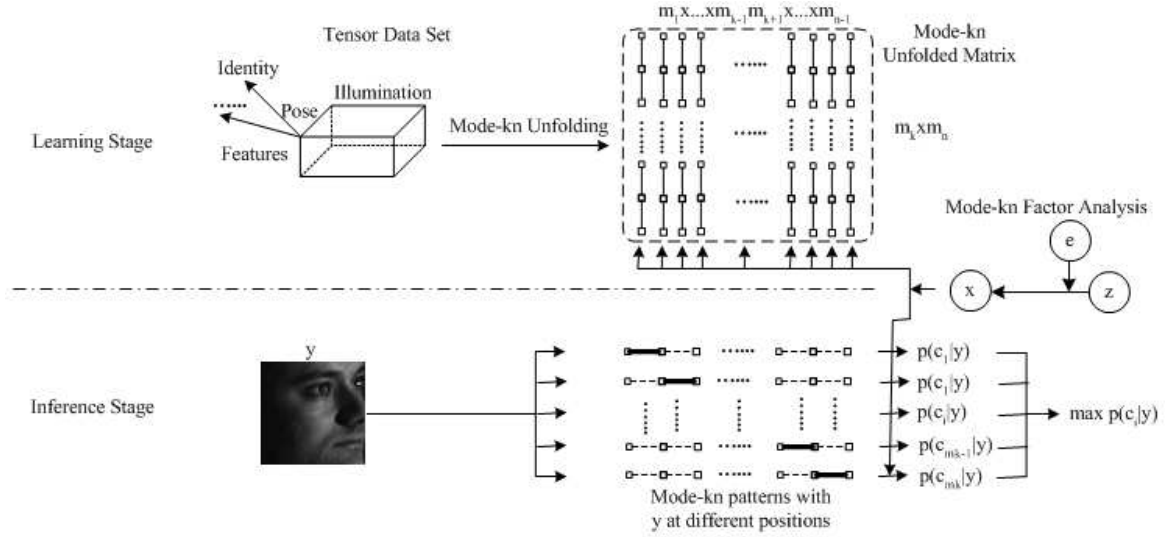
Fig. 1. Schematic diagram showing the learning and inference stages of the mode-kn factor analysis algorithm for tensor analysis.

B. Inference Stage

In the inference stage, the basic task is to predict the identity, pose or illumination mode/label for any new image datum $y \in \mathbb{R}^{m_n}$. In this work, we perform the estimation of extra-factors one by one instead of all of them together, but all can be done within a unified framework.

In this work, we utilize a simplified criterion for measuring the confidence that the image y belongs to the c-th mode of the k-th extra-factor, namely the maximal joint probability of the latent variable and a certain mode-kn pattern with y as the c-th segment. It is formally defined as

$$\hat{p}(c|y) = \max_{z, x^{kn}(y,c)} p(z, x^{kn}(y,c)), \tag{10}$$

where $x^{kn}(y, c)$ means a mode-kn pattern with y as the c-th segment. Note that a mode-kn pattern has m_k segments of equal lengths, and we can directly derive the probability for y to be the c-th segment of a mode-kn pattern from Eqn. (8), but it will result in a complex computation problem. Hence we propose the above simplified statistical confidence measure $\hat{p}(c|y)$.

The term $p(z, x^{kn}(y, c))$ in Eqn. (10) can be further rewritten as

$$p(z, x^{kn}(y,c)) = p(z)p(x^{kn}(y,c)|z). \tag{11}$$

Then, we can have

$$\max_{z, x^{kn}(y,c)} p(x^{kn}(y,c)|z) \tag{12}$$

$$= \max_z \exp\{-||y - (\mu + Wz)_c||^2/\delta^2\}, \tag{13}$$

where $(\mu + Wz)_c = \mu_c + W_c z$ means the c-th segment of the vector $\mu + Wz$. Similarly, μ_c is the c-th segment of the vector μ and W_c is the c-th segment of the matrix W along the row direction. It means that the elements of $x^{kn}(y, c)$ except y are with the same values as Wz.

Consequently, the optimization problem is simplified as

$$\min_z ||y - (\mu_c + W_c z)||^2/\delta^2 + ||z||^2, \tag{14}$$

and the optimal z can be derived as

$$z = (\delta^2 I + W_c^T W_c)^{-1} W_c^T (y - \mu_c). \tag{15}$$

Finally, the classification of the mode for a given extra-factor is performed by

$$c^* = \arg\max_c \hat{p}(c|y). \tag{16}$$

The whole statistical framework for tensor analysis is demonstrated in Figure 1. The learning stage and inference stage are further detailed in the diagram.
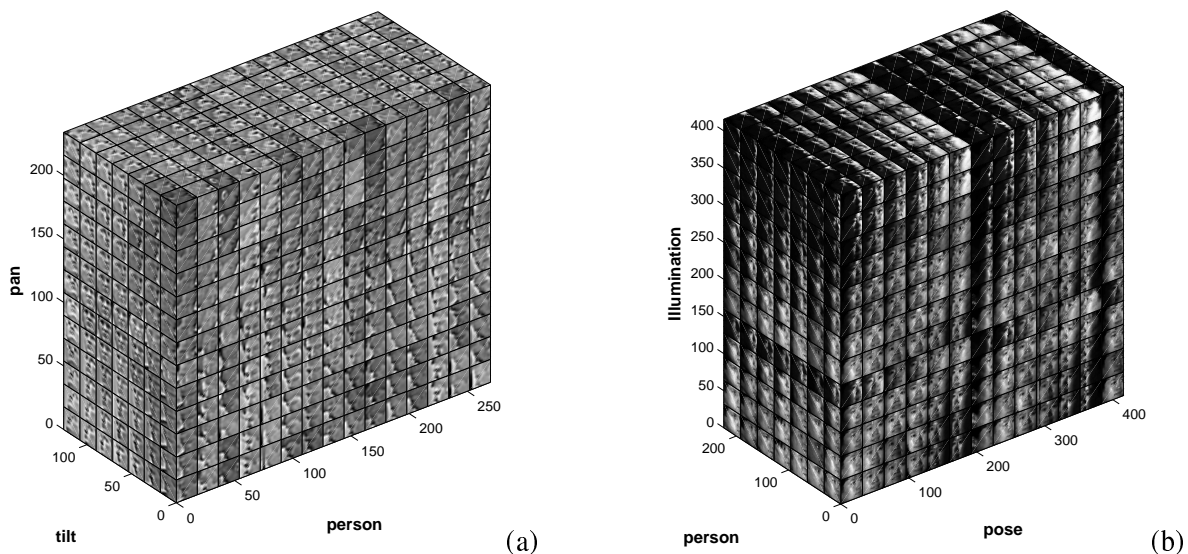

Fig. 2. Data sets shown in a 4D structure: a) Pointing'04 and b) CMU PIE. Note that to improve display, some persons and illuminations are not shown for CMU PIE database.

C. Discussion

In this subsection, we would like to highlight some aspects of our proposed mode-kn factor analysis algorithm.

1) Mode-kn pattern vs. Mode-k Pattern: Most previous algorithms [11][8] with tensor analysis directly analyze the mode-k pattern. When the k-th dimension of the tensor data is for an extra-factor, the mode-k pattern lacks explicit semantics and the size is typically small, e.g., the number of poses or illuminations is commonly at most

21 in our experiments. Consequently, the statistical analysis of these patterns is of limited importance, and no dimensionality reduction is necessary for these dimensions in real applications, as has been indicated with the algorithms [15][14].

A mode-kn pattern characterizes the combined features for all modes of a certain extra-factor, and the statistical modeling of these patterns can reveal the correlations among different modes of this extra-factor. In addition, the size of the mode-kn pattern is large, and the number of patterns is also reasonably large, which greatly facilitate the subsequent statistical modeling and learning process.

2) Classification by Interacting with Synthesis: As demonstrated in Section 3.2, the classification of the mode of a certain extra-factor for a new image datum is defined as a search for the optimal latent variable z and the image data in other modes, such that the joint probability of the latent variable and the complete mode-kn pattern is maximized. In the classification process, the derived latent variable has encoded all the information for synthesizing the images for other modes of the given extra-factor, namely

$$x^{kn} = \mu + Wz, \tag{17}$$

where z is the derived latent variable in Eqn. (13).

An important fact for mode-kn factor analysis to possess great classification capability is that it performs the classification by interacting with synthesis. As shown in Eqn. (10), for the final classification, it synthesizes the images in all other modes of a given extra-factor, and ensures that the combined mode-kn pattern will have high probability along with the latent variable. Instead, most previous algorithms for tensor analysis make the prediction only by reconstructing the input image via the learned core matrix, but do not consider whether the synthesized images in other modes will have reasonable probabilities. This classification-by-synthesis approach directly leads to the superiority of the mode-kn factor analysis algorithm over conventional algorithms in potential classification capability.

IV. Experiments

In this section, we present two sets of experiments for evaluating the effectiveness of our proposed mode-kn factor analysis algorithm for classification of the extra-factors including pose and illumination.

A. Data Set Preparation

Two databases are used for the experiments. One is the Pointing'04 [7] data set. The Pointing'04 head pose database consists of 15 sets of images. Each set contains two series of 93 images of the same person with different poses. The nose tips are manually marked for this database, and we crop the faces with the assumption that the face sizes are similar. This database is used for pose estimation in two directions: pan and tilt. Another database

is the CMU PIE [9] database, where the facial images are cropped and resized to 32 by 32. Due to the data incompleteness for some subjects, we use a subset consisting of 59 subjects with 13 poses and 21 illuminations.

For the Pointing'04 database, we extract the first layer of Laplacian Pyramid [4] features to represent the faces. $91 = 7$ (tilt) $\times$ 13 (pan) poses are used for the experiments, and the images from the first series are used for training and from other series for testing. The gray-level features are used for the CMU PIE database, with the first 25 persons for training and the others for testing. The 4D structures of the extracted data tensor are displayed in Figure 2.

B. Visualization of Mode Matrices

In this subsection, we visualize the learned mode matrices from mode-kn factor analysis. The CMU PIE database is used for visualization. The mean vector μ and the mode matrix W along the pose (illumination) dimension are displayed in Figure 3 and Figure 4. The counterpart of mode matrices from TensorFaces are shown in Figure 5. We can see that the column vectors of W in mode-kn factor analysis characterize the joint variation of different modes for certain extra-factors, and another interesting observation is that different components of the pose pattern characterize different illuminations, while different components of the illumination pattern characterize different poses. For the mode matrices from Tensorfaces, we cannot observe meaningful characteristics, and even for the feature dimension, we cannot observe the explicit relationship with pose and illumination.

Fig. 3. The mean faces in μ (first row) and the first four column vectors in the matrix W (second to fifth rows) for the mode-kn factor analysis along the pose dimension. The results are obtained from the training data of CMU PIE database.

Fig. 4. The mean faces in μ (first row) and the first four column vectors in the matrix W (second to fifth rows) for the mode-kn factor analysis along the illumination dimension. The results are obtained from the same data as in Figure 3.
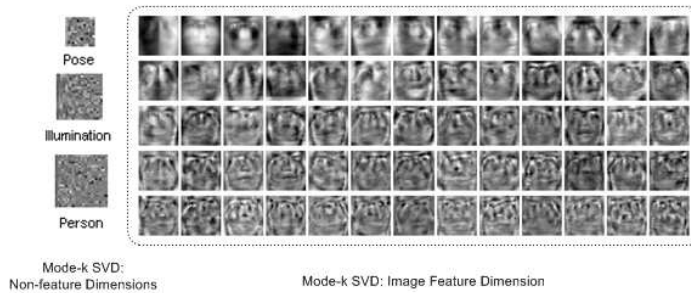
Fig. 5. The mode matrices of different tensor dimensions from Tensorfaces. Note that the column vectors of the mode matrix for the feature dimension of the tensor are reshaped into matrices to facilitate display, and the results are obtained from the training data of the CMU PIE database.

TABLE I

PAN POSE ESTIMATION ACCURACIES (%) ON THE POINTING'04 DATABASE. NOTE THAT THE SYMBOL *TF* MEANS TENSORFACES AND *MICA* MEANS MULTILINEAR ICA.

Algorithm	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	Total
TF+SP	75.4	72.4	64.8	56.2	53.3	57.1	56.2	42.8	52.4	51.4	67.6	58.1	61.0	69.1
TF+FV	83.8	85.7	79.1	67.6	65.7	63.8	74.3	66.7	63.8	67.6	**79.1**	68.6	**78.1**	72.6
MICA+SP	78.1	74.3	67.6	58.1	70.5	62.0	69.5	57.1	61.0	61.0	60.0	62.9	73.3	65.8
MICA+FV	**84.8**	85.7	**82.9**	70.5	68.6	68.6	72.4	70.5	**73.3**	71.4	78.1	**70.5**	77.1	75.0
Mode-kn FA	83.8	**86.7**	81.0	**74.3**	**79.1**	**82.0**	**78.1**	**77.1**	71.4	**71.4**	79.1	66.7	74.3	*77.3*

C. Pose Estimation on Pointing'04

In this subsection, we evaluate the effectiveness of mode-kn factor analysis for extra-factor estimation on the Pointing'04 database. Comparisons are performed with two conventional techniques for tensor analysis: TensorFaces (TF) and Multilinear ICA (MICA). For these two baseline algorithms, two approaches are used for final classification. One is the subspace projection method presented in [13], referred to as *SP* in the results table, and the other one, referred to as *FV* [14], computes the extra-factor vectors as in Eqn.(4) and then conducts classification directly on these vectors with the Nearest Neighbor approach. For both baseline algorithms and mode-kn factor analysis, the column number of the mode matrix related to intra-features is tested with several values, namely [10% 20% 30% 40% 50% 60% 70% 80% 90%] of the feature number, and the best results are reported. The results listed in Tables 1-2 show that our proposed algorithm outperforms all of the other four methods for pose estimation in both

TABLE II

TILT POSE ESTIMATION ACCURACIES (%) ON THE POINTING'04 DATABASE.

Algorithm	T1	T2	T3	T4	T5	T6	T7	Total
TF-SP	78.0	63.6	50.8	50.0	54.4	50.8	67.7	60.6
TF-FV	**91.8**	76.9	64.1	73.3	80.5	85.1	**99.0**	81.5
MICA-SP	85.1	59.5	44.6	53.9	50.3	62.6	85.6	63.1
MICA-FV	87.7	72.8	66.7	73.3	77.4	**89.2**	**99.0**	80.9
Mode-kn FA	87.7	**81.0**	**72.3**	**77.4**	**81.0**	86.2	98.0	*83.4*

pan and tilt directions. Another observation is that the pose estimation at the near-frontal directions are the most difficult for the other four methods, while mode-kn factor analysis works much better than them in these directions.

D. Pose and Illumination Estimation on PIE

In this subsection, we evaluate the estimation accuracy of pose and illumination on the CMU PIE databases. The results listed in Tables 3-4 again show the superiority of our proposed mode-kn factor analysis over the other four methods. Especially for pose estimation, the accuracy is improved from 74.3% of the best result from the other four algorithms to 80.2%. Also, we observe that the FV-version method can achieve better performance than the SP-version method for both TensorFaces and Multilinear ICA.

Note that we focus on the classification of the extra-factors with mode-kn factor analysis in this work, and its applications for automatic image synthesis with different illuminations and poses are left for future publication due to the page limitation.

TABLE III

POSE ESTIMATION ACCURACIES (%) ON THE CMU PIE DATABASE. NOTE THAT P1-P13 MEAN DIFFERENT POSES IN BOTH TILT AND PAN DIRECTIONS.

Algorithm	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	Total
TF-SP	58.4	93.6	57.7	62.9	81.4	39.2	83.1	42.7	25.5	88.5	79.4	79.3	87.3	67.6
TF-FV	68.8	79.4	72.3	68.6	77.0	61.6	79.6	59.8	27.5	86.1	73.4	87.4	74.0	70.4
MICA-SP	59.0	92.4	60.1	62.0	80.1	40.6	76.8	38.7	23.4	89.9	76.2	71.6	**88.2**	66.1
MICA-FV	**77.7**	88.0	72.5	**82.1**	**82.2**	59.9	79.7	56.4	21.6	90.1	**86.6**	86.4	82.1	74.3
Mode-kn FA	73.4	**93.8**	**75.1**	78.7	80.4	**78.7**	**89.2**	**87.5**	**38.1**	**90.6**	83.5	**99.2**	73.9	*80.2*

TABLE IV

ILLUMINATION ESTIMATION ACCURACIES (%) ON THE CMU PIE DATABASE. NOTE THAT I1-I12 MEAN DIFFERENT ILLUMINATIONS WHILE THE ILLUMINATIONS FROM 13 TO 21 ARE COMBINED AS *Else* IN THE TABLE TO FACILITATE DISPLAY.

Algorithm	I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	Else	Total
TF-SP	25.8	24.4	24.0	8.4	9.3	11.8	14.3	6.3	16.3	12.2	14.3	15.8	16.9	16.0
TF-FV	**93.2**	16.3	0.2	11.5	16.7	29.2	51.1	33.5	19.5	12.4	5.9	10.0	28.3	26.4
MICA-SP	23.5	26.9	**59.7**	17.2	15.2	39.4	35.1	19.0	37.8	20.0	21.7	27.1	34.6	31.1
MICA-FV	85.5	32.1	24.9	24.7	29.0	45.7	47.3	45.7	**39.4**	28.3	**27.4**	**37.3**	**44.5**	41.3
Mode-kn FA	92.1	**51.0**	10.4	**35.7**	**58.8**	**47.3**	**70.4**	**73.8**	26.0	**55.2**	5.2	29.4	36.3	*42.0*

V. CONCLUSION

In this paper, we have proposed a statistical framework for analyzing tensor-formed image ensemble. Three main contributions are offered for tensor analysis: 1) mode-kn unfolding of the tensor and the corresponding mode-kn factor analysis were proposed for tensor analysis; 2) a closed-form solution was provided for inferring the extra-factors of new data; and 3) the estimation of the mode for a given extra-factor and the synthesis for other modes

were performed in an interacting manner, instead of independently as done conventionally. An interesting direction for future work is to utilize a Mixture Factor Analysis model [6], instead of a factor analysis model, for expressing the possibly nonlinear generative model of the mode-kn patterns.

REFERENCES

[1] T. Anderson, and H. Rubin. Statistical inference in factor analysis. *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, pp. 111–150, 1956.

[2] P. Belhumeur, J. Hespanda, and D. Kiregeman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711–720, 1997.

[3] C. Bishop, and M. Tipping. A hierarchical latent variable model for data visualization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 281–293, 1998.

[4] P. Burt, and P. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, vol. 31, pp. 532–540, 1983.

[5] A. Dempster, N. Laird, D. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society, Series B*, vol. 39, pp. 1–38, 1977.

[6] K. Fukunnaga. Introduction to Statistical Pattern Recognition, second edition. *Academic Press*, 1991.

[7] N. Gourier, D. Hall, and J. Crowley. Estimating face orientation from robust detection of salient facial features. *Proceedings of Pointing, ICPR, International Workshop on Visual Observation of Deictic Gestures, Cambridge, UK*, 2004.

[8] V. Govindu. A tensor decomposition for geometric grouping and segmentation. *proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 1150–1157, 2005.

[9] T. Sim, S. Baker, and M. Bsat. The CMU pose, illumination, and expression database. *Proceeding of the European Conference on Computer Vision*, vol. 25, pp. 1615–1618, 2003.

[10] C. Spearman. General intelligence, objectively determined and measured. *American Journal of Psychology*, vol. 15, pp. 201–293, 1904.

[11] M. Vasilescu. Human motion signatures: analysis, synthesis, recognition. *Proceedings of the 16th International Conference on Pattern Recognition*, vol. 3, pp. 456–460, 2002.

[12] M. Vasilescu, and D. Terzopoulos. Multilinear subspace analysis for image ensembles. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 93–99, 2003.

[13] M. Vasilescu, and D. Terzopoulos. Multilinear analysis of image ensembles: Tensorfaces. *Proceedings of the European Conference on Computer Vision*, vol. 2, pp. 447–460, 2002.

[14] M. Vasilescu, and D. Terzopoulos. Multilinear independent components analysis. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol . 1, pp. 547–553, 2005.

[15] H. Wang, and N. Ahuja. Facial expression decomposition. *IEEE International Conference on Computer Vision*, vol. 2, pp. 958–965, 2003.

[16] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H. Zhang. Discriminant analysis with tensor representation. *Proceddings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 526–532, 2005.

[17] J. Ye. Generalized Low Rank Approximations of Matrices, *Machine Learning Journal*, vol. 61, pp. 167–191, 2005.

[18] J. Ye, R. Janardan, and Q. Li. Two-Dimensional Linear Discriminant Analysis, *Neural Information Processing Systems*, pp. 1569–1576, 2004.