Commun. Combin., Cryptogr. & Computer Sci., 1 (2021), 129–135

Communications in Combinatorics, Cryptography & Computer Science

Journal Homepage: www.vonneumann-publishing.com/cccs

Apply Tensorial methods for detecting and recovering facial shapes

Ali Reza Shojaeifard[a,*], Hamid Reza Yazdani[b], Mohsen Shahrezaee[c]

[a] Imam Hossein Comprehensive University, Department of Mathematics and Statistics Tehran, Iran.
[b] Imam Hossein Comprehensive University, Department of Mathematics and Statistics Tehran, Iran.
[c] Imam Hossein Comprehensive University, Faculty of Defense and Engineering, Tehran, Iran.

Abstract

This paper proposes a fast 3-D facial shape recovery algorithm from a single image with general, Unknown lighting. To derive the algorithm, we formulate a non-linear least-square problem with two-parameter vectors which are related to personal identity and light conditions. We then Combine the spherical harmonics for the surface normal of a human face with tensor algebra and show that in a certain condition, the dimensionality of the least-square problem can be further reduced to one-tenth of the regular subspace-based model by using tensor decomposition (N-mode SVD), which speeds up the computations. To enhance the shape recovery performance, we have incorporated prior information in updating the parameters. The proposed algorithm takes less than 0.4 s to reconstruct a face in the experiment and shows a significant performance improvement over other reported scheme.

Keywords: Facial shape recovery, Image processing, Statistical face model, Tensorial methods.

2020 MSC: 65F99, 15A69, 58C05.

1. Introduction

Tensors (i.e., multiway arrays) provide an effective and faithful representation of structural properties of the data, especially for multidimensional data or data ensembles affected by multiple factors [1]. For instance, a video sequence can be represented by a third-order tensor with the dimensionality of ($\mathtt{height} \times \mathtt{width} \times \mathtt{time}$); an image ensemble measured under multiple conditions can be represented by a higher-order tensor with the dimensionality of ($\mathtt{pixel} \times \mathtt{person} \times \mathtt{pose} \times \mathtt{illumination}$) [3].

The goal of this paper is to provide a practical method, which can be applied to a single picture taken by an ordinary camera and which achieves good accuracy in the recovery of facial shape in a short time. When taking a picture using an ordinary camera in a general environment, control of the pose of a face is easy but control of the light conditions is not. Hence, we assume that the face in an input image is in frontal pose under general light conditions, which are unknown [4]. We then formulate a non-linear least-square problem of two-parameter vectors by the use of the spherical harmonics for the surface normal of a human face to handle the general light conditions, based on the Lambertian assumption [5]. In order to speed up the calculation, we introduce tensor algebra and show how to reduce the dimensionality of the least-square problem [6].

In this paper, after introducing the tensor completion problem, we state the tensorial face shape recovery method for image recovery and implement it on some examples.

*Corresponding author
Email addresses: ashojaeifard@ihu.ac (Ali Reza Shojaeifard), hamidreza.yazdani@gmail.com (Hamid Reza Yazdani), mshahrezaee@mail.bmn.ir (Mohsen Shahrezaee)

Received: November, 1, 2021 Revised: November, 15, 2021 Accepted: November, 20, 2021

2. Preliminaries

Here, we use tensor algebra and explanation for standard operations such as inner product, F-norm, and singular value decomposition **(SVD)**. Therefore, we briefly state some preliminaries for tensor calculus and tensor completion. For more details and information, please read [1], and [2].

Definition 2.1. A tensor is a multidimensional array, with a dimensionality that is referred to as its order. X stands for a Nth-order tensor (i.e. an N-way array) which is identified as N-dimensional or N-mode tensor, too. Here, the word "order" is used for referring the dimensionality of a tensor (like Nth-order tensor), and the word "mode" is employed for describing operations on a particular dimension (like mode-n product)[2]. We denote the set of all n-dimensional tensors of order m by $T_{m,n}$. The tensor A is called symmetric, if all $a_{i_1,\ldots,i_n}$ are invariant under any permutation of indices. The set of all real n-dimensional symmetric tensors of order m is shown with $S_{m,n}$.

Definition 2.2. A fiber of a tensor is defined as a vector obtained by fixing all indices but one. Fibers are generalizations of matrix columns and rows. Mode-n fibers are obtained by fixing all indices but n^{th}.

Definition 2.3. *Mode-n matricization (unfolding)* of tensor X, denoted as $X_{(n)}$, is obtained by arranging all mode-n fibers as columns of a matrix. The precise order in which fibers are stacked as columns is not important as long as it is consistent. Figure 1 shows the fibers of 3-tensor. *The folding* is the inverse operation of matricization/unfolding.

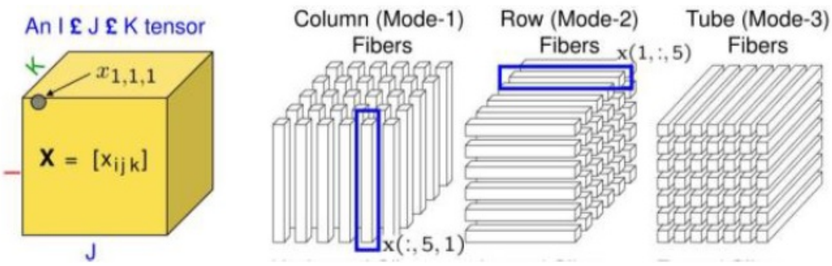


Figure 1: Fibers of a tensor from rank 3.

Definition 2.4. Mode-n product of tensor X and matrix A is denoted by $X \times_n A$, and defined by

$$Y = X \times_n A \Longleftrightarrow Y_{(n)} = AX_{(n)}. \tag{2.1}$$

This product is commutative (when applied in distinct modes), i.e.

$$(X \times_n A) \times_m B = (X \times_m B) \times_n A. \tag{2.2}$$

for $m \neq n$.

Definition 2.5. The inner product of two tensors X and Y of same size is defined as $< X, Y >$. Unless otherwise specified, we treat it as dot product defined as follows [3]:

$$< X, Y >= \sum_{i_1 \cdots i_m = 1}^{n} x_{i_1 \cdots i_m} y_{i_1 \cdots i_m}. \tag{2.3}$$

The F-norm of a tensor X (Generalized from matrix Frobenius norm), is defined as [2]:

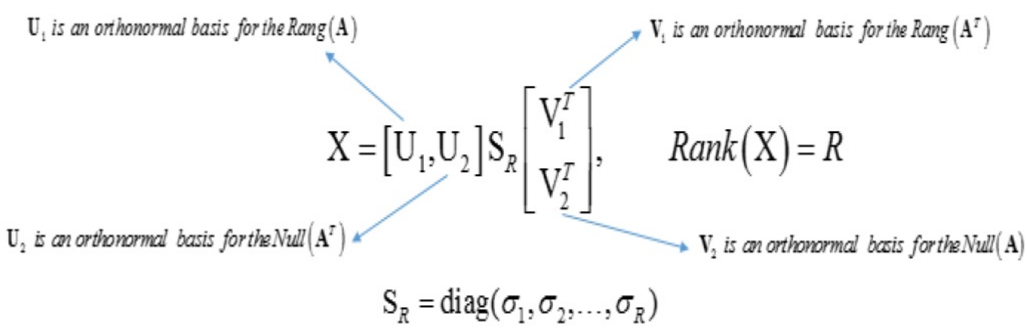
$$||X||_F := \sqrt{< X, X >} = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} X^2_{i_1, i_2, \cdots, i_N}}. \tag{2.4}$$

Definition 2.6. Suppose X is a symmetric tensor of $S_{m,n}$, r is a positive integer number and $u^{(k)} \in \mathbb{R}^n$ for $k \in \{1, \cdots, r\}$ exist such that

$$X = \sum_{k=1}^{r} (u^{(k)})^m. \tag{2.5}$$

Therefore, X is called a *completely positive tensor (CP)*, and $\mathrm{Eq.}(3)$ is a *CP-decomposition* of X (For example, see Figure 2). In the CP-decomposition of $\mathrm{Eq.}(3)$, the minimum of r is called *CP-rank* of X [5].

Definition 2.7. In the general, The *Singular Value Decomposition (SVD)* is a factorization of a real or complex matrix that generalizes the eigen decomposition, which only exists for square normal matrices to any $m \times n$ matrix via an extension or the polar decomposition. In the tensor calculus, similar concepts proposed as follows:



Where U_i (for $i = 1, 2$) are orthonormal and can be extended to orthonormal basis, S_R is a diagonal and positive definite of dimension R, with R as the number of non-zero eigenvalues of tensor X^*X and $V = [V_1, V_2]$ is a unitary tensor of $\mathrm{rank}(X)$ [1]. The representation of SVD shown in figure 2.

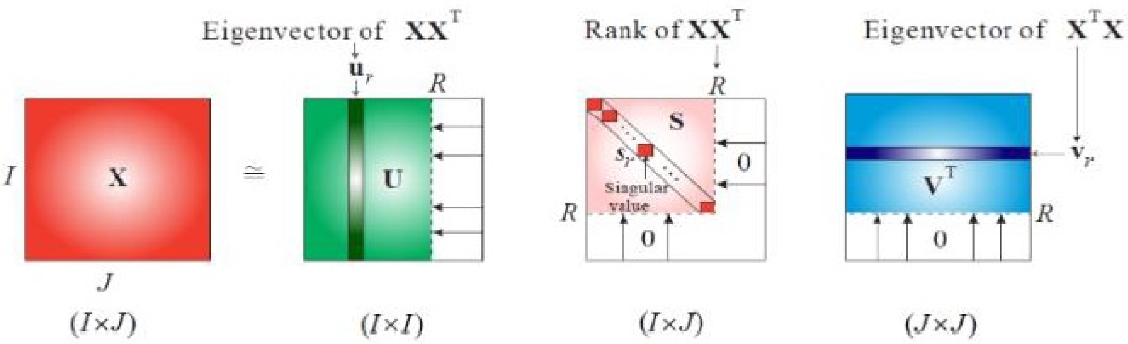


Figure 2: The representation of SVD.

Definition 2.8. Given a low-rank (either CP rank or other ranks) tensor T with missing entries, the goal of completing it can be formulated as the following optimization problem:

$$\begin{aligned} &\mathrm{Min}_X \mathrm{rank}_*(X), \\ &\mathrm{Subject} \quad \mathrm{to} \quad X_\Omega = T_\Omega. \end{aligned}$$

Where $\mathrm{rank}_*(X)$ denotes a specific type of tensor rank based on the rank assumption of given tensor T, X represents the completed low rank tensor of T and Ω is an index set of observations. For this paper, the specific rank is completed positive (CP) rank [3].

Definition 2.9. The generalized of SVD or N-mode SVD (also named Higher-order SVD) is defined as [3]:

$$D = G \times_1 U_1 \times_2 \cdots \times_N U_N. \tag{2.6}$$

Where G is the core tensor and U_k is derived from SVD of

$$D_{(k)} = U_k \sum_k V_k^T. \tag{2.7}$$

and G is defined as

$$D = G \times_1 U_1^T \times_2 \cdots \times_N U_N^T. \tag{2.8}$$

HoSVD of 3-tensor shown in figure 3.

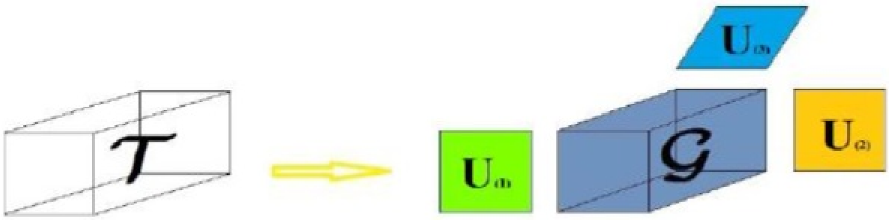


Figure 3: HoSVD decomposition of 3-tensor.

3. Algorithm Implementation

An image of a human face depends on various parameters, such as its 3D structure, head pose, light and exposure, surface reflection property, etc. This picture can be approximated as linear equation i.e.,

$$I(x, y) \approx f(x, y)^T S. \tag{3.1}$$

Where $I(x, y)$ is the brightness of the pixel $(x, y), s \in \mathbb{R}^{n_l}$ is the light condition vector and $f(x, y)$ is a n_l-dimensional vector which is related to the surface characteristics and is either the scaled normal ($n_l = 3$) or the spherical harmonic representation ($n_l = 4$ or 9) at the pixel (x, y). Figure 4 shows the main process of the algorithm.

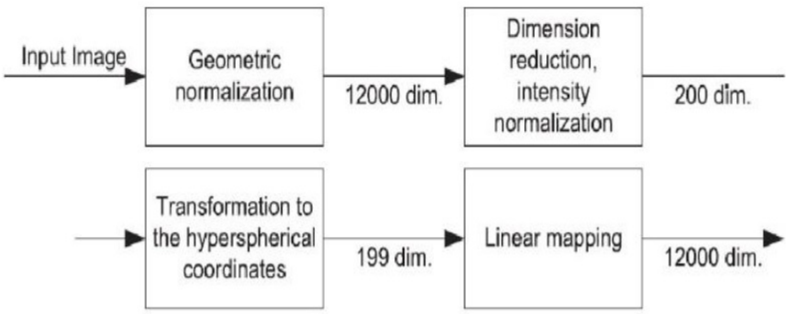


Figure 4: The schematic view of algorithm.

Figure 5 shows a 3D-face reconstruction. The coordinate plane part of the face is cut in 3D and represented by the tensor representation methods (with various tensor analysis).

The overall procedure for the proposed method has two steps: modeling and Reconstruction. In the first step, we apply the affine transform to the harmonic images of each training sample, so that the centers

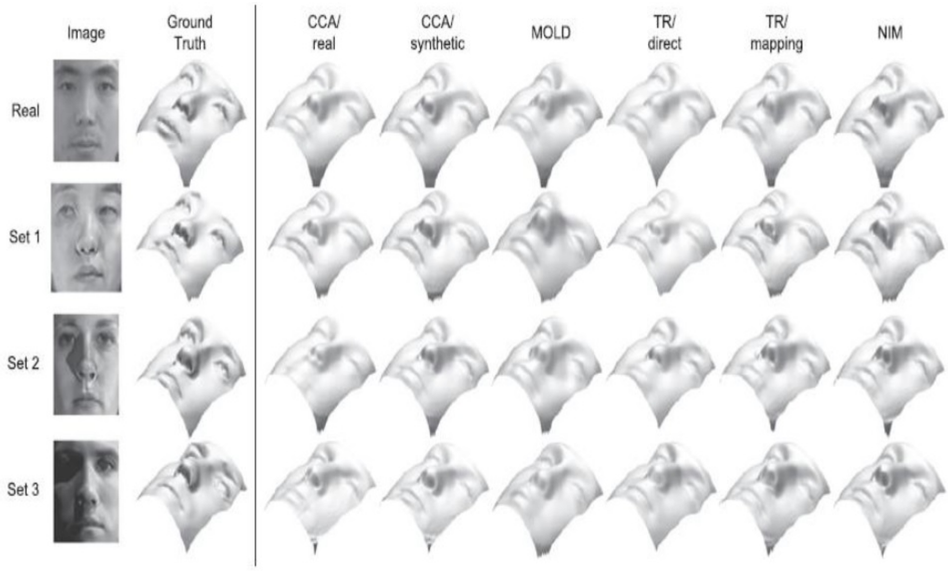


Figure 5: The samples of 3D face reconstructions by tensorial representation methods.

of the eyes and mouth of all samples are located at the same positions. After that, calculate the mean tensor of the flat face (F) and apply N-mode SVD to it (Q). Now, the train image is stored by the HoSVD version of the mean tensor. In the second step, apply the affine transform to a test image, the resultant image is denoted as I′. Now, calculate in the reduced dimensional space and denote its tensorized version as L. After that, if the norm distance between Q and L is minimized, the best result is obtained, otherwise, search the best Q for this purpose. In fact, the main problem is minimizing the problem of norm distance between L, Q.

For our experiments, we run **MATLAB R2020a** on the laptop system *(Asus A53sv)* with configuration as shown in table 1.

Table 1: The configuration of experiment system.

Laptop	**Asus A53sv**
CPU	Ci7-2670QM (6MB Smart Cache)
No of Cores	4 Physical/8 Thread
Frequency	2.2-3.1 GHz
RAM	16 GB (DDR3/1600 MHz)
H.D.D.	750 GB (7200 rpm)
GPU	GeForce Gt 630m (2GB/96 Cuda Cores)
Performance (FP32)	367 Gflops
O.S.	Win 10 Pro 64 bit

After importing the original picture and running the algorithm, the results in figure 6 are obtained.

For another example, we apply the algorithm on the other images by the different poses of head, background, and exposure. The result is shown in figure 7.

4. Conclusions

Practical experiments show that the proposed algorithm takes only a few hundredths to a few tenths of a second to reconstruct a face, and this improves performance dramatically. Studies show that the efficiency, accuracy, and speed of this method in recovering the shape of the face in different exposure

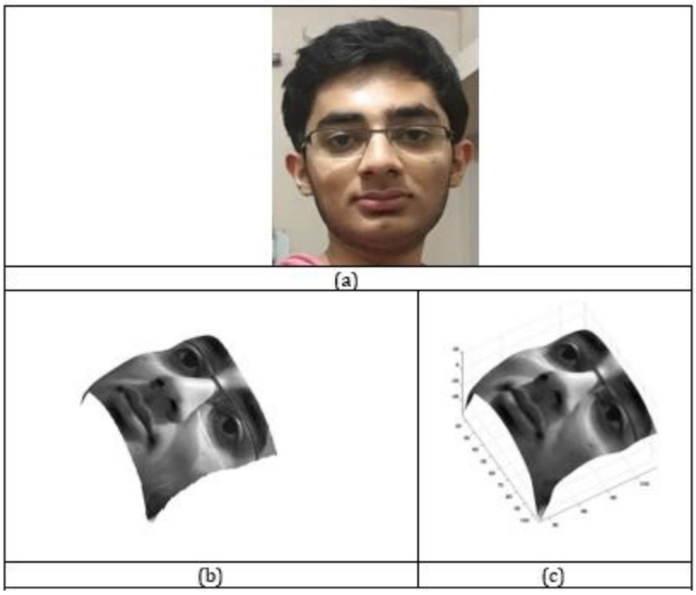


Figure 6: The implementation of Facial Shape Recovery by Tensorial methods. (a): original photo, (b): RR-tensor representation output and (c): SIFR-tensor representation output.

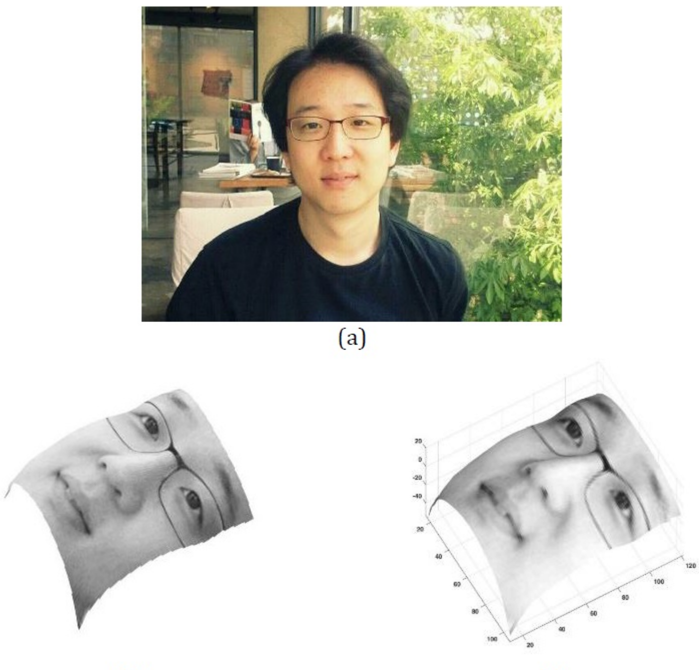


Figure 7: The implementation of Facial Shape Recovery by Tensorial methods. (a): original photo, (b): RR-tensor representation output and (c): SIFR-tensor representation output.

conditions is very high. The image space is created by selecting a suitable nonlinear function and approximating a set of multiplicative transformed samples. The final image is the result of minimizing the sum of all errors at multiple vertices. We use various tensor analyses such as CP and Tucker to reduce unwanted changes and improve efficiency. The experimental results show that the proposed algorithm has high reconstruction accuracy even in the presence of shadows in different exposure conditions and the processing speed is high enough for instantaneous (real-time) applications. Rapid reconstruction of face shapes from different angles and with different expressions can be a future research field.

References

[1] L. Qi, *The Spectral Theory of Tensors*, arXiv (2012): 1201.3424v1. 1, 2, 2.7

[2] A. Rovi, *Analysis of 222 Tensors, MAI mathematics: Master thesis* (2010), Linkopings University. 2, 2.1, 2.5

[3] Q. Song, H. Ge, J. Caverlee and X. Hu, *Tensor Completion Algorithms in Big Data Analytics*, ACM Trans- actions on Knowledge Discovery from Data (TKDD), **6 (13)** (2019), 1–48. 1, 2.5, 2.8, 2.9

[4] M. Lee, C-H. Choi, *Fast facial shape recovery from a single image with general, unknown lighting by using tensor representation*, Pattern Recognition, Elsevier, **44** (2011),1487–1496. 1

[5] M. Lee, C-H. Choi, *A robust real-time algorithm for facial shape recovery from a single image containing cast shadow under general, unknown lighting*, Pattern Recognition, Elsevier, **46** (2013), 38–44. 1, 2.6

[6] M. Lee, C-H. Choi, *Real-time facial shape recovery from a single image under general, unknown lighting by rank relaxation*, Computer Vision and Image Understanding, Elsevier, **120** (2014), 59–69. 1