Fast Tensor Singular Value Decomposition Using the Low-Resolution Features of Tensors

Cagri Ozdemir
Department of Computer Science & Engineering
South Dakota Mines
Rapid City, South Dakota
Email: cagri.ozdemir@mines.sdsmt.edu

Randy C. Hoover
Department of Computer Science & Engineering
South Dakota Mines
Rapid City, South Dakota
Email: randy.hoover@sdsmt.edu

Kyle Caudle
Department of Mathematics
South Dakota Mines
Rapid City, South Dakota
Email: kyle.caudle@sdsmt.edu

***Abstract*—The tensor singular value decomposition (t-SVD) based on an algebra of circulants is an effective multilinear subspace learning technique for dimensionality reduction and data classification. Unfortunately, the computational cost associated with computing the t-SVD can become prohibitively expensive, particularly when dealing with very large data sets. In this paper, we present a computationally efficient approach for estimating the t-SVD by capitalizing on the correlations of the data in the *temporal* dimension. The approach proceeds by extending our prior work on fast eigenspace decompositions by transforming the tensor data from the spatial domain to the spectral domain in order to obtain reduced order harmonic tensor. The t-SVD can then be applied in the transform domain thereby significantly reducing the computational burden. Experimental results which are presented on the extended Yale-B, COIL-100, and MNIST data sets show the proposed method provides considerable computational savings with the approximated subspaces that are nearly the same as the true subspaces as computed via the t-SVD.**

***Index Terms*—fast tensor singular value decomposition, tensor singular value decomposition, multilinear subspace learning**

I. Introduction

Subspace methods serve as important tools in many areas of machine learning, data analytics, and artificial intelligence. Arguably, the most popular linear subspace learning techniques are principal component analysis (PCA) [1], [2], linear discriminant analysis (LDA) [3], [4], and locally preserving projection (LPP) [5]–[9], these have been used extensively in a number of application such as face characterization and recognition [4]–[7], [10], object recognition and pose detection [11]–[16]. All these techniques take advantage of the fact that a set of highly correlated images can be represented by a low-dimensional subspace. While traditional PCA, LDA, and LPP methods have shown much success over the years, with the pervasive increase in 2-dimensional sensors, extensions to such techniques using multi-dimensional tensors (referred to as n-mode or n-way arrays) have been applied. Most notably are tensor decompositions based on the canonical polyadic (CP), parallel factors (referred to as CANDECOMP/PARAFAC), or the Tucker decomposition regarded as a High Order Singular Value Decomposition (HOSVD). In either approach, both factor an n-mode tensor into a collection of n factor matrices multiplied by a core tensor where the "tensor product" and core tensor differ depending on the approach [17]–[19].

In [20]–[28], a new approach to decomposing a third-order tensor has been proposed that's based upon Fourier theory and an algebra of circulants (referred to as the tensor SVD or the t-SVD for short as outlined in Section II-B). The t-SVD is unique in that it decomposes a third-order tensor into the product of two tensor-orthogonal third-order tensors and a front-face diagonal third-order tensor (referred to as f-diagonal) that are somewhat analogous to its matrix SVD counterpart. Once the left orthogonal tensor is computed (which contains of left-singular matrices as its lateral slices), online classification can be performed in real-time via a simple projection and nearest neighbor search. However, the offline calculation required to determine both the appropriate number of left-singular matrices, as well as the left-singular matrices themselves can be computationally expensive when dealing with large data sets. Furthermore, the discrete Fourier transform (DFT) used for computing t-SVD converts real-value pixel values to complex values of Fourier coefficients. Therefore, the complex arithmetic exacerbates the computational burden when computing the t-SVD using the DFT.

In this paper, we address the computational expense of computing the left orthogonal tensor consisting of desired left-singular matrices by exploiting the low-frequency information in datasets. Because most data of interest is highly correlated (e.g., video surveillance, face databases, object detection and pose estimation, etc.), most of the relevant signal energy is contained in relatively few low-frequency Fourier harmonics as outlined in [12]–[15], [29], [30]. In addition, we show that the discrete wavelet transform (DWT) can be used to: a) estimate the appropriate number of harmonics for information recovory and b) efficiently reduce the overall dimensionality of the training tensor thereby significantly reducing the computational cost associated with computing the t-SVD.

We present experimental results from some classic image data sets (namely the extended Yale-B, COIL-100, and

The current research was supported in part by the Department of the Navy, Naval Engineering Education Consortium under Grant No. (N00174-19-1-0014) and the National Science Foundation under Grant No. (2007367). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Naval Engineering Education Consortium or the National Science Foundation.

MNIST). Comparisons are made with the traditional t-SVD and it is shown that the proposed method is capable of computing the t-SVD along with significant computational savings.

The remainder of the paper is organized as follows: In Section II we provide some mathematical preliminaries and define the original t-SVD with the DFT as outlined in [22], [26], [27], [31]. In Section III we propose the fast t-SVD algorithm using the low-resolution information of data tensors. In Section IV, we compare the subspaces obtained from both the original t-SVD and the proposed fast t-SVD, and in Section V we provide some summary remarks and outline some avenues for future research.

II. MATHEMATICAL PRELIMINARIES

In this section, we provide a brief overview of the tensor definitions and mathematical operations for the original t-SVD based on the DFT, extended details may be found in [22], [26], [27], [31].

A. Mathematical Preliminaries

Multidimensional data can be defined as multidimensional arrays of numbers, commonly referred to as *tensors* [19]. The dimensions of a tensor are called *ways* or *modes*. The number of modes determines the *order* of a tensor. If, for example, $\mathcal{A} \in \mathbb{R}^{\ell\times m\times n}$ then we say $\mathcal{A}$ is a third-order tensor. Thus, matrices and vectors are second-order and first-order tensors, respectively. Fundamental to the results presented in this section is based on recent developments based upon Fourier theory and and algebra of circulants as outlined in [**?**], [20]–[27]. In addition, recently defined multiplication operation on third-order tensors has been shown in [22], [26]–[28], [31].

First, we introduce the basic notations and definitions outlined in [20]–[22]. It will be convenient to break a tensor $\mathcal{A}$ in $\mathbb{R}^{\ell\times m\times n}$ up into various slices and tubal elements, and to have an indexing on those. The i^{th} lateral slice will be denoted $\mathcal{A}_i$ whereas the j^{th} frontal slice will be denoted $\mathcal{A}^{(j)}$. In terms of MATLAB indexing notation, this means $\mathcal{A}_i \equiv \mathcal{A}(:,i,:)$ while $\mathcal{A}^{(j)} \equiv \mathcal{A}(:,:,j)$.

We use the notation $\mathbf{a}_{ik}$ to denote the i,k^{th} tube in $\mathcal{A}$; that is $\mathbf{a}_{ik} = \mathcal{A}(i,k,:)$. The j^{th} entry in that tube is $\mathbf{a}_{ik}^{(j)}$. Indeed, these tubes have special meaning for us in the present work, as they will play a role similar to scalars in $\mathbb{R}$. Thus, we make the following definition:

Definition 1. An element $\mathbf{c} \in \mathbb{R}^{1\times 1\times n}$ is called a **tubal-scalar** of length n.

In order to discuss multiplication between two tensors we must first introduce the concept of converting $\mathcal{A} \in \mathbb{R}^{\ell\times m\times n}$ into a block circulant matrix. If $\mathcal{A} \in \mathbb{R}^{\ell\times m\times n}$ with $\ell \times m$ frontal slices then $\mathbf{circ}(\mathcal{A})$ is a block circulant matrix of size $\ell n \times mn$.

$$\mathbf{circ}(\mathcal{A}) = \begin{bmatrix} A^{(1)} & A^{(n)} & A^{(n-1)} & \dots & A^{(2)} \\ A^{(2)} & A^{(1)} & A^{(n)} & \dots & A^{(3)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ A^{(n)} & A^{(n-1)} & \ddots & A^{(2)} & A^{(1)} \end{bmatrix}.$$

We anchor the **MatVec** operator to the frontal slices of the tensor. **MatVec**$(\mathcal{A})$ takes an $\ell \times m \times n$ tensor and returns a block $\ell n \times m$ matrix

$$\mathbf{MatVec}(\mathcal{A}) = \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(n)} \end{bmatrix}.$$

The operation that takes **MatVec**$(\mathcal{A})$ back to tensor form is the **fold** operator:

$$\mathbf{fold}\big(\mathbf{MatVec}(\mathcal{A})\big) = \mathcal{A}.$$

Definition 2. Let $\mathcal{A} \in \mathbb{R}^{\ell\times p\times n}$ and $\mathcal{B} \in \mathbb{R}^{p\times m\times n}$ be two third order tensors. Then the **t-product** $\mathcal{A} * \mathcal{B} \in \mathbb{R}^{\ell\times m\times n}$ is defined as

$$\mathcal{A} * \mathcal{B} = \mathbf{fold}\big(\mathbf{circ}(\mathcal{A}) \cdot \mathbf{MatVec}(\mathcal{B})\big).$$

Definition 3. The **identity tensor** $\mathcal{I} \in \mathbb{R}^{m\times m\times n}$ is the tensor whose first frontal slice is the $m\times m$ identity matrix, and other frontal slices are all zeros.

Definition 4. If $\mathcal{A}$ is $\ell \times m \times n$, then the **tensor transpose** $\mathcal{A}^T$ is the $m \times \ell \times n$ tensor obtained by transposing each of the frontal slices and then reversing the order of transposed frontal slices through n.

Definition 5. A tensor $\mathcal{A} \in \mathbb{R}^{n\times n\times n}$ has an **tensor inverse** $\mathcal{B}$ provided

$$\mathcal{A} * \mathcal{B} = \mathcal{I} \text{ and } \mathcal{B} * \mathcal{A} = \mathcal{I},$$

where $\mathcal{I} \in \mathbb{R}^{n\times n\times n}$. The tensor inverse is computed as

$$\mathcal{A}^{-1} = \mathbf{fold}\big(\mathbf{circ}(\mathcal{A})^{-1}\big).$$

Definition 6. The tensor norm used through this paper is the **Frobenious norm** which for the tensor $\mathcal{A} \in \mathbb{R}^{\ell\times m\times n}$ is given by

$$\|\mathcal{A}\|_F = \sqrt{\sum_{i=1}^{\ell}\sum_{j=1}^{m}\sum_{k=1}^{n} a_{ijk}^2},$$

where a_{ijk} is the i,j,k^{th} element of $\mathcal{A}$.

B. Computation of the t-SVD

As outlined In [20]–[22] the authors show that, for $\mathcal{A} \in \mathbb{R}^{\ell\times m\times n}$, there exists orthogonal tensors $\mathcal{U}$ and $\mathcal{V}$, a *front-face diagonal* (f-diagonal) tensor $\mathcal{S}$ such that

$$\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T = \sum_{i=1}^{min(\ell,m)} \mathcal{U}_i * \mathbf{s}_i * \mathcal{V}_i^T, \tag{1}$$

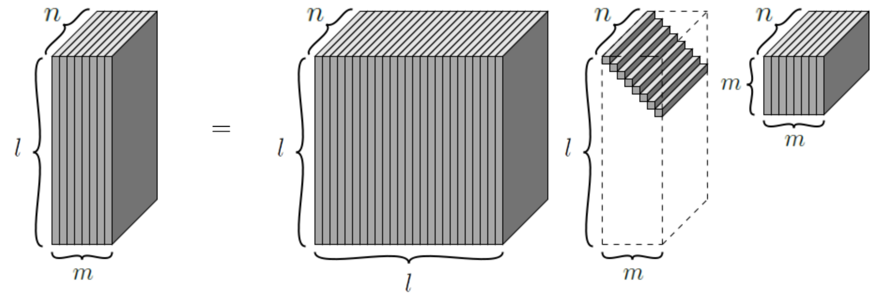

Fig. 1. Graphical depiction of the **t-SVD**.

where $\mathcal{U} \in \mathbb{R}^{\ell \times \ell \times n}$, such that $\mathcal{U} * \mathcal{U}^T = \mathcal{I}_{\ell \times \ell \times n}$ is the tensor of left-singular matrices, $\mathcal{V} \in \mathbb{R}^{m \times m \times n}$, such that $\mathcal{V} * \mathcal{V}^T = \mathcal{I}_{m \times m \times n}$ is the tensor of right-singular matrices, $\mathcal{S} \in \mathbb{R}^{\ell \times m \times n}$ is an *f-diagonal* tensor where $\mathbf{s}_i = \mathcal{S}(i, i, :)$ are the singular tuples. In the context of the current work, the lateral slices of $\mathcal{U}$ (the left-singular matrices) are analogous to the left-singular vectors of the matrix SVD and serve as the *tensor* principal components. A graphical depiction of the **t-SVD** is shown in Fig. 1 where the tensor on the left-hand-side is the image data tensor $\mathcal{A}$ and each lateral slice $\mathcal{A}^{(i)}$ is an $l \times n$ image in our data set. We note that, l is the row-space of our image set and referred to as mode-1, n is the column-space of our image set and referred to as mode-3, whereas m is the "temporal-space[1]" of our image set and referred to as mode-2.

Computation of the **t-SVD** comes from the constructive proof outlined in [21], [22], [26], which will be restated here for completeness. It is well known in matrix theory that a circulant matrix can be diagonalized via left and right multiplication by a discrete Fourier transform (DFT) matrix. Similarly, a block circulant matrix can be block diagonalized via left and right multiplication by a DFT matrix. For example, consider the tensor $\mathcal{A} \in \mathbb{R}^{n \times n \times n}$, then

$$(F_n \otimes I_n)\mathbf{circ}(\mathcal{A})(F_n^* \otimes I_n) = \mathbf{diag}(D_1, D_2, ..., D_n), \quad (2)$$

where each of the D_i are $n \times n$, I is an $n \times n$ identity matrix, F_n is the $n \times n$ DFT matrix, F_n^* is its conjugate transpose, and $\otimes$ is the Kronecker product. To construct the **t-SVD** defined in (1), the matrix singular value decomposition is performed on each of the D_i, i.e., $D_i = U_i S_i V_i^T$ shown in Figure 1. Taking the first block column of each block circulant matrix and applying the **fold** operator results in the decomposition $\mathcal{U} * \mathcal{S} * \mathcal{V}^T$. Note that for simplicity, as well as computational efficiency, this entire process can by performed using the fast Fourier transform (FFT) in place of the DFT matrix as illustrated in [20]–[22].

III. Fast tensor singular value decomposition

Previous works based on the t-SVD outlined in [26], [27], and the authors used as high-resolution as is available for the tensor data sets in order to maintain as much information as possible, however this may result in an impractical computational burden when dealing with large data sets. If, for example, we have $m \times n$ images with ℓ samples, then we concatenate all samples into the tensor $\mathcal{A} \in \mathbb{R}^{m \times \ell \times n}$. To perform the original t-SVD, we first apply the FFT along mode-3 direction, then we compute the matrix SVD for each frontal slices as illustrated in [20]–[22]. The time complexity of performing the FFT along the each tube of tensor $\mathcal{A}$ is $O(m\, \ell\, n\, log(n))$, whereas the time complexity of calculating the matrix SVD for the each frontal slices of tensor $\mathcal{A}$ is $O(2n\; min(m\; \ell^2,\; \ell\; m^2))$. Thus, as ℓ is the number of samples and generally significantly large, it determines the computational cost of the t-SVD.

The frequency resolution through samples (mode-2 direction) of a tensor data can be obtained by applying the FFT along the tubes in mode-2 direction. Since we deal with highly correlated images, the frequencies with large "*power-spectra*" (magnitude) occur in low frequencies. Therefore, in spectral domain, we can easily obtain low-frequencies with large magnitude. In Fig. 2, after applying the FFT in mode-2 direction, we compute the magnitude of each lateral slices by computing tensor frobenious norm defined in **definition 6** and show that first lateral slice of the each data tensor has the largest power spectra. Therefore, we propose to use a set of pre-determined number low frequencies for the offline computation instead using all information in the tensor data. If, for instance, consider the tensor $\mathcal{A} \in \mathbb{R}^{m \times \ell \times n}$, then applying the original t-SVD results in the decomposition $\mathcal{A} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ where the left orthogonal tensor $\mathcal{U}$ comprising eigenmatrices is a $m \times m \times n$ tensor. In our work, after applying the FFT in mode-2 direction, first p lateral slices can be selected ($p << \ell$) and then applying inverse FFT in mode-2 direction along first p lateral slices turns the tensor data $\mathcal{A} \in \mathbb{R}^{m \times \ell \times n}$ into the tensor $\bar{\mathcal{A}} \in \mathbb{R}^{m \times p \times n}$. We propose to compute the t-SVD on this reduced tensor data ($\bar{\mathcal{A}} = \bar{\mathcal{U}} * \bar{\mathcal{S}} * \bar{\mathcal{V}}^T$) and obtain the left orthogonal tensor $\bar{\mathcal{U}}$ which is a $m \times m \times n$ tensor. We call this proposed method as **the fast tSVD-dft**. Because the eigenmatrices of tensor $\bar{\mathcal{A}}$ can be effectively used to approximate the high-resolution eigenmatrices of tensor $\mathcal{A}$, the tensor $\bar{\mathcal{U}}$ can be used for online computation instead of the tensor $\mathcal{U}$.

Although the low frequencies capture most of the information, a set of low-frequencies (lateral slices in spectral domain) must be determined to compute the fast t-SVD.

[1]Temporal in this context need not refer to time but rather samples in our training/testing set.

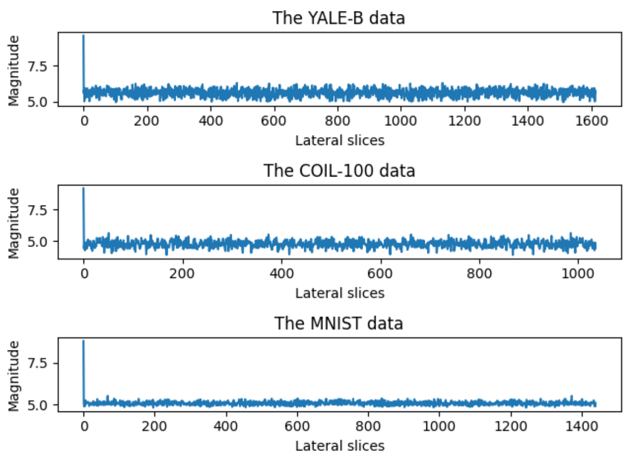

Fig. 2. Magnitude of the Fourier coefficients of lateral slices

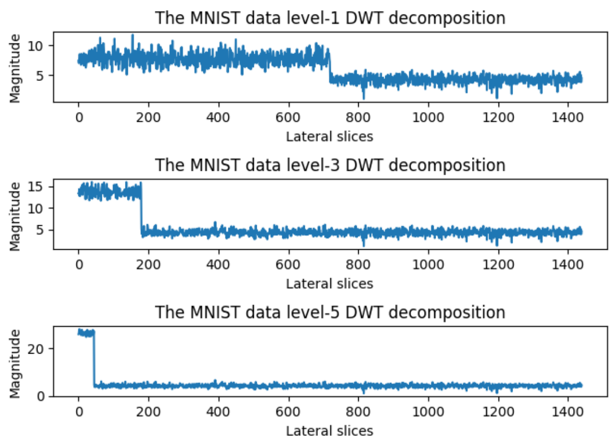

Fig. 3. Magnitude of the Wavelet coefficients of lateral slices

However, it seems problematic to determine a threshold based on the power spectra density plots in Fig. 2. We address this problem by applying the DWT instead the FFT. In the wavelet transform, a signal in time domain is decomposed by passing it through high-pass filter (resulting in detail coefficients) and low-pass filter (resulting in approximation coefficients) to produce low-pass and high-pass wavelet coefficients referred to as "*level-1 decomposition*". Low-pass version can be further decomposed by passing it to again a set of low-pass and high pass filters referred to as "*level-2 decomposition*". This process can be further continued to a pre-defined level, extended details may be found in [32]–[35]. There are different types of wavelets (e.g. Haar wavelet [32] and Daubechies wavelet [36]). In our work, we use the Haar wavelets due to its low computation cost and simplicity to apply as compared to other wavelets.

Therefore, we can directly use approximation coefficients to compute the t-SVD referred to as **the fast tSVD-dwt**. Fig. 3 illustrates the magnitude of each lateral slices after applying the DWT throughout mode-2 direction on the MNIST data tensor with level-1, level-3, and level-5 decomposition. In this case, a number of lateral slices corresponding to low-frequencies in the spectral domain is determined based upon the number of level decomposition used in the DWT. For example, we can use first 45 slices instead of 1440 to compute the t-SVD for the MNIST data, when the level-5 discrete wavelet transform is applied shown in Fig. 3.

IV. Experimental Results

In this section, we measure error between the true subspace (the original t-SVD) and the approximated subspace (the fast t-SVD) by: 1) applying the energy recovery ratio defined for tensors [27], 2) determining the distance between the subspaces, and 3) obtaining the residue between the subspaces [29]. We also determine the time required to compute these subspaces. For our experimental validation, we choose three common data sets used in the literature, namely: (a) Extended Yale-B data set that contains images of human faces under varying facial expressions and illumination directions [37], [38]. There are 38 subjects and 59 images per subject; (b) The COIL-100 data set that contains images of 100 different objects being rotated about a single degree of freedom to obtain 72 poses per object [39]. For computational efficiency, we use first 20 objects with 72 poses of each object to produce total data set of 1440 images; and (c) the classic MNIST data set of handwritten digits [40] where we randomly select 200 samples of each digit to produce a total data set of size 2000 images. We note that, for computational efficiency, each image in extended Yale-B, and COIL-100 data sets was resized to 32×32, whereas the MNIST data set was resized to 20×20.

A. *Information Recovery*

It has been shown in Section III that because the image samples in the tensor data are highly correlated, most of the information in the entire tensor data $\mathcal{A} \in \mathbb{R}^{m \times \ell \times n}$ is obtained by low-frequencies in spectral domain. Therefore, we construct the reduced tensor data $\bar{\mathcal{A}} \in \mathbb{R}^{m \times p \times n}$ by only selecting the first p lateral slices in the spectral domain and then backing them into the spatial domain. Dimension reduction is performed by projecting the tensor data $\mathcal{A}$ onto the left orthogonal tensor $\bar{\mathcal{U}} \in \mathbb{R}^{m \times m \times n}$ obtained from the reduced tensor data $\bar{\mathcal{A}}$ via t-SVD (We call this as the fast t-SVD). To determine how much information about the tensor data is captured by the left orthogonal tensor, we rely on energy recovery ratio defined for tensors [27] as

$$\rho(\mathcal{A}, \bar{\mathcal{U}}_k) = \frac{\sum_{i=1}^{k} \|\bar{\mathcal{S}}_i\|_F^2}{\|\mathcal{A}\|}, \tag{3}$$

where $\bar{\mathcal{S}}$ is the singular-tuples of reduced tensor data $\bar{\mathcal{A}}$, "$\|.\|_F$" donates the Frobenious norm, and "k" is the number of eigenmatrices. We may also determine the energy recovery ratio of left singular tensor $\mathcal{U}$ by substituting $\mathcal{S}$ for $\bar{\mathcal{S}}$.

In Fig. 5, we show energy recovery ratio of the original t-SVD, fast tSVD-dft, and fast tSVD-dwt on the MNIST data set. Both the DFT and DWT are applied along statistical mode-2 in order to obtain low-frequencies as shown in Section III.

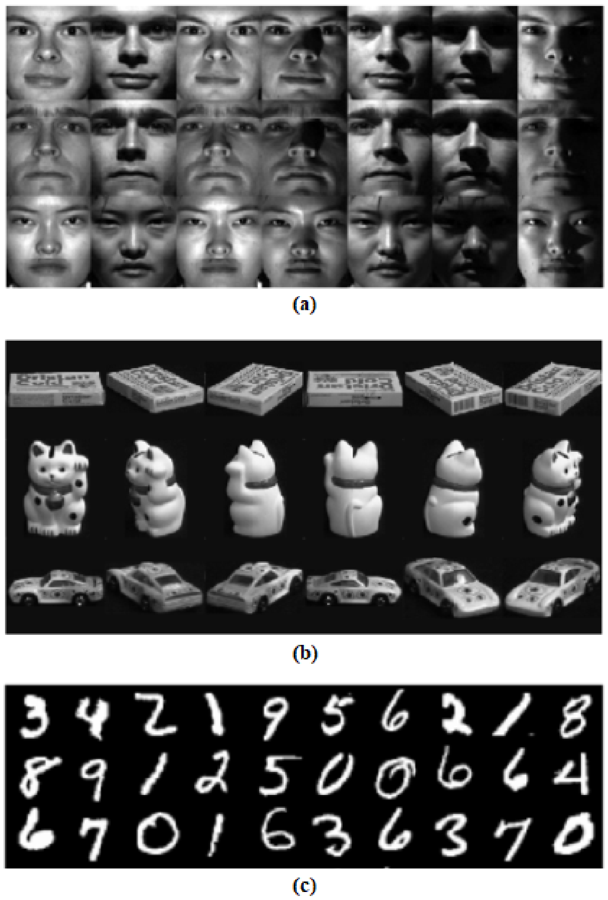

Fig. 4. A subset of images from each data set used in this research. (a) the extended Yale-B face database, (b) the COIL-100 data set, and (c) the MNIST data set.

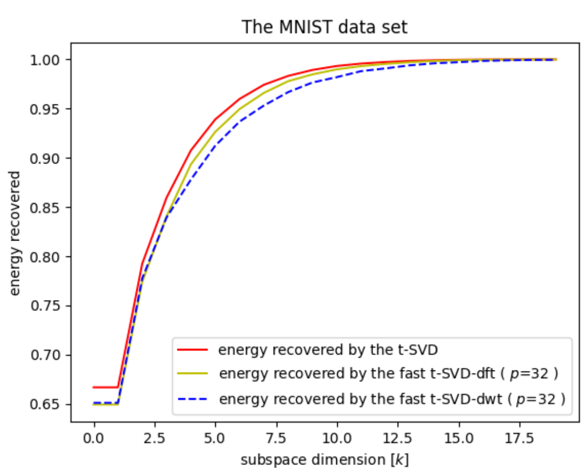

Fig. 5. Energy recovery ratio of different subspaces obtained via the t-SVD, fast t-SVD-dft, and fast t-SVD-dwt. p denotes the selected lateral slices in spectral domain.

Similarly to Fig 5, TABLE I illustrates the energy recovery ratio of all three subspaces on the extended Yale-B and COIL-100 data sets. Although, each of the three data sets is reduced to $p = 32$ in the fast t-SVD-dft and fast tSVD-dwt, the subspaces obtained from the both methods capture almost the same energy as that of the original t-SVD.

TABLE I
ENERGY RECOVER RATIO ON THE EXTENDED YALE-B AND COIL-100 DATA SETS.

Data Set	Subspace dimension $[k]$	Energy rec. t-SVD	Energy rec. t-SVD-dft	Energy rec. t-SVD-dwt
Ext. Yale-B	2	0.768	0.764	0.763
	5	0.971	0.967	0.963
	7	0.984	0.979	0.975
	10	0.992	0.989	0.986
	15	0.998	0.996	0.992
COIL-100	2	0.906	0.906	0.906
	5	0.985	0.984	0.983
	7	0.993	0.992	0.992
	10	0.997	0.997	0.996
	15	0.999	0.999	0.998

B. Distance Between the Subspaces

True and approximated tensor subspaces can be compared in terms of the distance between these subspaces. As the left singular tensor is orthogonal, tensor product of the left orthogonal tensor and its tensor always provides the identity tensor defined in **definition 5** and can be written as

$$\mathcal{U}^T * \mathcal{U} = \mathcal{I}. \tag{4}$$

The multiplication of the original tensor subspace $\mathcal{U}$ by the transpose of the approximated subspace $\bar{\mathcal{U}}$ is close to identity tensor. Therefore, the distance between the subspace can be calculated as follows

$$\delta = \|(\mathcal{U}^T * \bar{\mathcal{U}}) - \mathcal{I}\|_F, \tag{5}$$

where $\|.\|_F$ is the tensor Frobenious norm defined in **definition 6**, and δ measures how close these two tensor subspaces.

C. Residue Between Subspaces

The possibility that the tensor subspace $\mathcal{U} \in \mathbb{R}^{m\times m\times n}$ can be rotated in to the tensor subspace $\bar{\mathcal{U}} \in \mathbb{R}^{m\times m\times n}$ is explored by solving the problem

$$\Delta = \min_{\mathcal{Q}} \|\mathcal{U} - \bar{\mathcal{U}} * \mathcal{Q}\|_F, \tag{6}$$

where $\mathcal{Q} \in \mathbb{R}^{m\times m\times n}$ is an orthogonal tensor, and Δ is the residue. As $\mathcal{U}$ and $\bar{\mathcal{U}}$ are orthogonal tensors, the residue can be calculated as

$$\mathcal{T} = \mathcal{U}^T * \bar{\mathcal{U}}, \tag{7}$$

$$\text{t-SVD}(\mathcal{T}) = \mathcal{U}_t * \mathcal{S}_t * \mathcal{V}_t^T, \tag{8}$$

$$\mathcal{Q}_{min} = \mathcal{U}_t * \mathcal{V}_t^T, \tag{9}$$

$$\Delta^2 = \|\bar{\mathcal{U}} - \mathcal{U} * \mathcal{Q}_{min}\|_F^2. \tag{10}$$

The smaller residue Δ, the closer $\mathcal{U}$ and $\bar{\mathcal{U}}$ are to representing the same subspace. Extended details may be found in [30].

In Fig. 6, we show the comparison of the approximated subspaces derived from low-resolution harmonics of the data

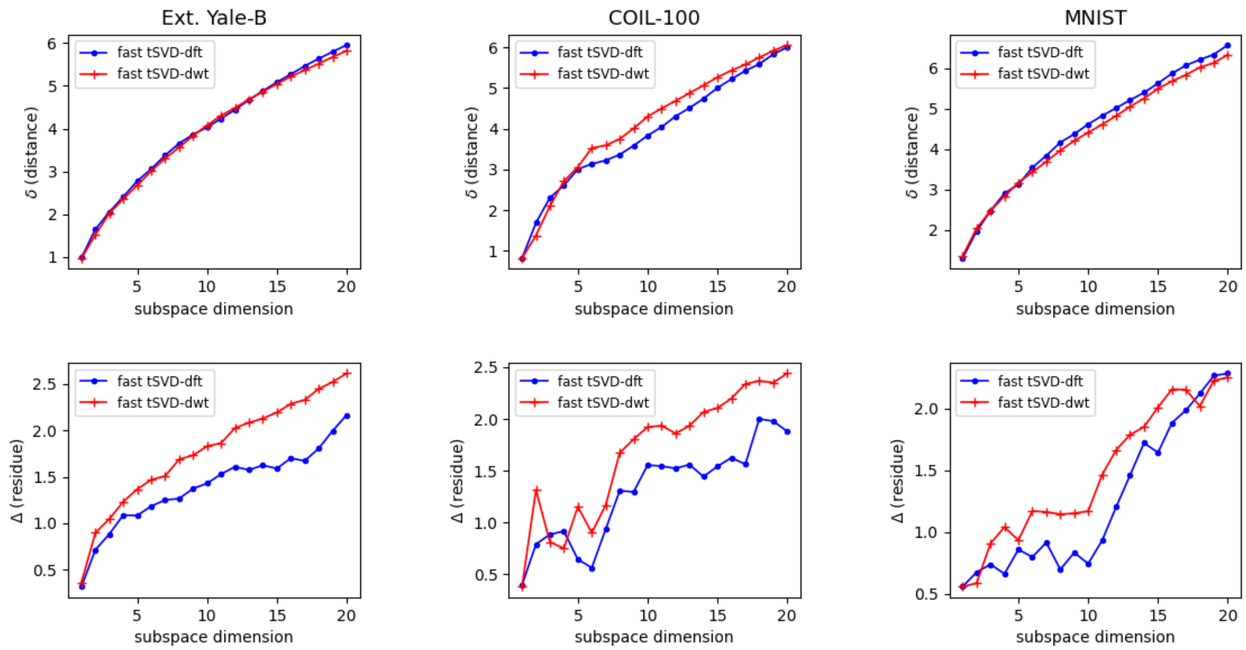

Fig. 6. Comparison of the two approximated subspaces and the true subspaces of all three data sets using distance and residue comparison measures.

sets and the true subspaces. The first row shows the comparison plots on all three data sets for distance (δ) comparison measure; whereas the second row illustrates the plots for residue (Δ) comparison measure. In particular, the distance and residue between the subspaces of the fast t-SVD and the true subspace show that the both approximated subspaces are nearly the same as the original subspace, as we observe very small distance and residue between the subspaces.

D. Computational Time

In this section, we apply the t-SVD on the original training sets and the reduced training sets in order to obtain the left orthogonal tensors. We reduced the training sets using low-frequency Fourier and wavelet harmonics. After obtaining the reduced training sets, t-SVD provides the approximated left orthogonal tensors referred to as the fast tSVD-dft and tSVD-dwt shown in Section III. The computation time of these three approach on all data sets is determined and illustrated in Table III.

For each of the three data sets, 20 different runs were evaluated where a random training/testing split was performed to ensure complete coverage of the training/testing space. Information regarding image size, training set and testing set, and the number of classes is outlined in Table II.

TABLE II
SPECIFICATIONS ON THE DATA SETS USED FOR EXPERIMENTAL VALIDATION.

Data Set	**Image Size**	**Train Set**	**Test Set**	**# of Classes (C)**
Ext. Yale-B	32×32	1614	628	38
COIL-100	32×32	1036	404	20
MNIST	20×20	1440	560	10

TABLE III
COMPUTATION TIME FOR EACH OF THE THREE METHODS ON ALL THREE DATA SETS.

	t-SVD	**fast tSVD-dft**	**fast tSVD-dwt**
Ext. Yale-B	13.524 ± 0.134	0.155 ± 0.002	0.041 ± 0.001
COIL-100	5.388 ± 0.123	0.054 ± 0.001	0.031 ± 0.001
MNIST	4.674 ± 0.038	0.020 ± 0.001	0.016 ± 0.002

TABLE III illustrates the computation time of the three methods on all data sets in seconds. As all methods are computed for 20 different cycles, we show the mean time $\pm$ the standard deviation in time across 20 runs. We use randomly chosen training sets for each data set shown in TABLE II, however we selected first 32 lateral slices from all training sets in the spectral domain with the FFT and the level-5 DWT for each data set (we selected the same number of lateral slices for computing the energy recovery in Fig. 5). We note that the time given for the fast tSVD-dft and fast tSVD-dwt comprises reconstructing the reduced training sets by using first 32 lateral slices in the spectral domain. As can be seen from the table and subspace comparison measures results, the fast tSVD-dft and tSVD-dwt provide significant computational savings, while obtaining the nearly same subspace as that of the original t-SVD shown via all three subspace comparison measures that are the energy recovery ratio, distance, and residue between the subspaces.

V. CONCLUSIONS AND FUTURE WORK

This paper presents a computationally efficient algorithm for the tensor singular value decomposition taking advantage of highly correlated image samples and using the low-frequency features that obtain most of the information of tensor data sets. The DFT and DWT are used for obtaining low-frequency harmonic tensors, and the t-SVD is performed in the transform domain to significanlty reduce the computational burden. Different subspace comparison measures are defined to compare the "estimated" subspace with the subspace as computed via the direct t-SVD. Experimental results on the extended Yale-B, COIL-100, and MNIST data sets show that our proposed method provides considerable computational savings for the off-line computation, while the subspaces obtained from our method reach the same quality as the subspaces of the original t-SVD. Future work will focus on methods that determine the optimal number of frequency harmonics for both the fast tSVD-dft and fast tSVD-dwt as well as improving the computational efficiency of online classification.

REFERENCES

[1] I. T. Jolliffe, "Principal components in regression analysis," in *Principal component analysis*. Springer, 1986, pp. 129–155.

[2] K. Fukunaga, *Introduction to statistical pattern recognition*. Elsevier, 2013.

[3] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of eugenics*, vol. 7, no. 2, pp. 179–188, Sept. 1936.

[4] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711–720, July 1997.

[5] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Transactions on Pattern analysis and Machine intelligence*, vol. 12, no. 1, pp. 103–108, Jan. 1990.

[6] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, Mar. 1991.

[7] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition," in *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1994, pp. 84–91.

[8] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang, "Face recognition using laplacianfaces," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 3, pp. 328–340, Jan. 2005.

[9] D. Cai, X. He, J. Han, and H.-J. Zhang, "Orthogonal laplacianfaces for face recognition," *IEEE transactions on image processing*, vol. 15, no. 11, pp. 3608–3614, Oct. 2006.

[10] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 1, pp. 34–58, Aug. 2002.

[11] H. Murase and S. K. Nayar, "Illumination planning for object recognition using parametric eigenspaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 12, pp. 1219–1227, Dec. 1994.

[12] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Pose detection of 3-D objects using S^2 correlated images and discrete spherical harmonic transforms," in *2008 IEEE International Conference on Robotics and Automation*. IEEE, May 2008, pp. 993–998.

[13] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Pose detection of 3-D objects using images sampled on $SO(3)$, spherical harmonics, and wigner-d matrices," in *2008 IEEE International Conference on Automation Science and Engineering*. IEEE, 2008, pp. 47–52.

[14] R. C. Hoover, *Pose estimation of spherically correlated images using eigenspace decomposition in conjunction with spectral theory*. Ph.D dissertation, Colorado State University, 2009.

[15] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Eigendecomposition of images correlated on S^1, S^2, and $SO(3)$ using spectral theory," *IEEE transactions on image processing*, vol. 18, no. 11, pp. 2562–2571, June 2009.

[16] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts, "Fast eigenspace decomposition of images of objects with variation in illumination and pose," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 41, no. 2, pp. 318–329, July 2010.

[17] R. A. Harshman *et al.*, "Foundations of the parafac procedure: Models and conditions for an" explanatory" multimodal factor analysis," *University of California at Los Angeles Los Angeles, CA*, 1970.

[18] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, Sept. 1966.

[19] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, Aug.2009.

[20] M. E. Kilmer, C. D. Martin, and L. Perrone, "A third-order generalization of the matrix svd as a product of third-order tensors," *Tufts University, Department of Computer Science, Tech. Rep. TR-2008-4*, Oct. 2008.

[21] K. Braman, "Third-order tensors as linear operators on a space of matrices," *Linear Algebra and its Applications*, vol. 433, no. 7, pp. 1241–1253, Dec. 2010.

[22] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM Journal on Matrix Analysis and Applications*, vol. 34, no. 1, pp. 148–172, Feb. 2013.

[23] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra and its Applications*, vol. 435, no. 3, pp. 641–658, Aug. 2011.

[24] C. D. Martin, R. Shafer, and B. LaRue, "An order-p tensor factorization with applications in imaging," *SIAM Journal on Scientific Computing*, vol. 35, no. 1, pp. A474–A490, Feb. 2013.

[25] D. F. Gleich, C. Greif, and J. M. Varah, "The power and arnoldi methods in an algebra of circulants," *Numerical Linear Algebra with Applications*, vol. 20, no. 5, pp. 809–831, Oct. 2013.

[26] N. Hao, M. E. Kilmer, K. Braman, and R. C. Hoover, "Facial recognition using tensor-tensor decompositions," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 437–463, Feb. 2013.

[27] R. C. Hoover, K. S. Braman, and N. Hao, "Pose estimation from a single image using tensor decomposition and an algebra of circulants," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, Sept. 2011, pp. 2928–2934.

[28] C. Ozdemir, R. C. Hoover, and K. Caudle, "2dtpca: A new framework for multilinear principal component analysis," in *2021 IEEE International Conference on Image Processing (ICIP)*, 2021, pp. 344–348.

[29] C.-Y. Chang, A. A. Maciejewski, and V. Balakrishnan, "Fast eigenspace decomposition of correlated images," *IEEE Transactions on Image Processing*, vol. 9, no. 11, pp. 1937–1949, Nov. 2000.

[30] K. Saitwal, A. A. Maciejewski, and R. G. Roberts, "A comparison of eigendecomposition for sets of correlated images at different resolutions," in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, vol. 1. IEEE, Oct. 2003, pp. 1011–1017.

[31] R. C. Hoover, K. Caudle, and K. Braman, "Multilinear discriminant analysis through tensor-tensor eigendecomposition," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, Dec.2018, pp. 578–584.

[32] A. Haar, "Zur theorie der orthogonalen funktionensysteme," *Mathematische Annalen*, vol. 69, no. 3, pp. 331–371, 1910.

[33] P. Porwik and A. Lisowska, "The haar-wavelet transform in digital image processing: its status and achievements," *Machine graphics and vision*, vol. 13, no. 1/2, pp. 79–98, Nov. 2004.

[34] A. Jensen and A. la Cour-Harbo, *Ripples in mathematics: the discrete wavelet transform*. Springer Science & Business Media, June 2001.

[35] G. Strang and T. Nguyen, *Wavelets and filter banks*. SIAM, 1996.

[36] I. Daubechies, "Orthonormal bases of compactly supported wavelets ii. variations on a theme," *SIAM Journal on Mathematical Analysis*, vol. 24, no. 2, pp. 499–519, Mar. 1993.

[37] D. J. Kriegman, P. N. Belhumeur, and A. S. Georghiades, "Representations for recognition under variable illumination," in *Shape, Contour and Grouping in Computer Vision*. Springer, 1999, pp. 95–131.

[38] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible illumination conditions?" *International journal of computer vision*, vol. 28, no. 3, pp. 245–260, July 1998.

[39] S. A. Nene, S. K. Nayar, and H. Murase, "Columbia object image library (COIL-100)," Technical Report CUCS-006-96, Tech. Rep. CUCS-006-96, Feb. 1996.

[40] (2018) MNIST database of handwritten digits. [Online]. Available: http://yann.lecun.com/exdb/mnist/