Adaptive Video Compression using PCA Method

Mostafa Mofarreh-Bonab Department of Electrical and Computer Engineering Shahid Beheshti University, Tehran, Iran

ABSTRACT

In this paper, a new PCA based method for video compression is introduced. This method extracts the features of video frames and process them adaptively based on required accuracy. This idea improves the quality of compression effectively. In this paper, we focused on the fact that video is a composition of sequential and correlated frames, so we can apply the PCA to these high correlated frames. Most of other video compression methods use DCT transform to compression. DCT causes to large damage in the edges of frames which plays fundamental role in quality of video. Our method in this paper doesn't reduce the bandwidth of frequency response, so the edges of frames don't fade.

Keywords

Video Compression, Correlation, PCA, SVD, 2DPCA, Frame, Feature extraction, Database, PSNR, Bit Rate.

1. INTRODUCTION

PCA -also known as KL Transform for images- is a statistical approach which is used widely in pattern recognition and compression of various type of databases, especially image databases which their components have high correlation -i.e. frontal face databases. This method first used by Kirby and Sirovich for compressing human face image database [7], [8].In PCA method, features of images are extracted by means of images correlations and these information are mapped to an orthogonal space and the features that are correspond to less important components -smaller eigenvalues- are ignored. In this approach, compression is reached by very little loss of information [1]. In fact, the eigenvectors and their corresponding eigenvalues are extracted using the covariance matrix of images. Since the eigenvectors which are related to small eigenvalues have very low information, their elimination has almost no effect on the quality of reconstructed images and compression is done by eliminating those components. There are other Optimal PCA-based methods for compressing images which show better results than conventional PCA, i.e. 2DPCA, K2DPCA, KPCA, L1norm-Based2DPCA and so on[3], [11], [10], [9]. As mentioned in the following, a combination of improved PCA presented by M.Mofarreh et.al. in [1] and the introduced method in [2] for compressing one image show better results rather than other methods. Also we will show that other PCA based methods don't have proper performance for video sequences.

2. PCA method

Suppose there areMgrayscale N × Pimages. N × Pgrayscale images are equivalent to N × P matrixes that the values of the components of the matrixes are the light intensities of the corresponding pixel's location. Put N × P = Q. By reshaping the matrixes, the image can be expressed as 1 × Qvectors \underline{F}_i in equation 1. Applying PCA, images are transferred to another field. All images are put in **X** matrix that its elements are the intensity values of images. Mohamad Mofarreh-Bonab Electrical and Electronic Engineering school, University of Bonab, East Azerbaijan, Iran

$$\mathbf{X} = \begin{bmatrix} \underline{F}_1 \\ \vdots \\ \underline{F}_M \end{bmatrix}_{M \times 0}, \underline{F}_i = (\mathbf{x}_{i1}, \mathbf{x}_{i2}, \dots, \mathbf{x}_{iQ})_{1 \times Q} \qquad \text{Eq. 1}$$

The term \underline{F}_i indicates the ith image which is converted to a vector. In order to applying PCA, some definitions should be considered;

The mean vector, \underline{M}_x : that contains mean values of each image and expressed as:

$$\underline{\overline{M}}_{x} = \frac{1}{Q} \begin{bmatrix} \sum_{k=1}^{Q} x_{1k} \\ \sum_{k=1}^{Q} x_{2k} \\ \vdots \\ \sum_{k=1}^{Q} x_{Mk} \end{bmatrix}_{M \times 1} = \begin{bmatrix} m_{1} \\ m_{2} \\ \vdots \\ \vdots \\ m_{M} \end{bmatrix}$$
Eq. 2

 $\widetilde{\mathbf{M}}_{\mathbf{x}}$ matrix, that contains the values of $\underline{\mathbf{M}}_{\mathbf{x}}$ for *M* times and expressed as:

Covariance matrix C_x for *M* rows of *X* matrix is [1]:

$$\mathbf{C}_{\mathrm{x}} = \left[\mathbf{c}_{\mathrm{i},\mathrm{j}} \right]_{\mathrm{M} \times \mathrm{M}}$$

That:

$$\begin{split} c_{i,j} &= \frac{1}{Q-1} \times \\ \sum_{k=1}^{Q} \left[\left(x_{ik} - \underline{\overline{M}}_{x}(i,1) \right) \times \left(x_{jk} - \underline{\overline{M}}_{x}(j,1) \right) \right] \qquad \text{Eq. 4} \end{split}$$

For C_x matrix, *M*eigenvectors \underline{v}_i , i = 1, 2, ..., M and *M* eigenvalues λ_i , i = 1, 2, ..., M can be found, which satisfy equation 5:

The modal matrix Λ will be obtained by putting all eigenvectors in a matrix, that its columns are the eigenvectors of C_x as shown below [2]:

$$\mathbf{\Lambda} = [\underline{\mathbf{v}}_1, \underline{\mathbf{v}}_2, \dots, \underline{\mathbf{v}}_M]$$
 Eq. 6

Now we can define *V* matrix:

$$\mathbf{V} = \mathbf{\Lambda}^{-1}$$
 Eq. 7

Λis a unitary matrix. So:

$$\mathbf{\Lambda}^{-1} = \mathbf{\Lambda}^{\mathrm{T}} \Longrightarrow \mathbf{V} = \mathbf{\Lambda}^{\mathrm{T}}$$
 Eq. 8

$$\Rightarrow \mathbf{V} = [\underline{v}_1, \underline{v}_2, \dots, \underline{v}_M]^T_{M \times M}$$
 Eq. 9

$$\mathbf{V} = \begin{bmatrix} v_1(1) & \cdots & v_M(1) \\ \vdots & \ddots & \vdots \\ v_1(M) & \cdots & v_M(M) \end{bmatrix}_{M \times M}$$
Eq.10

Applying PCA produces **Y** matrix:

$$\mathbf{Y} = \mathbf{V}.\left(\mathbf{X} - \widetilde{\mathbf{M}}_{\mathrm{x}}\right)$$
Eq. 11

with ,V and \tilde{M}_x matrixes and inversing mentioned process, Xmatrix can be calculated as equation 12:

$$\mathbf{Y} = \mathbf{V}. \left(\mathbf{X} - \widetilde{\mathbf{M}}_{\mathbf{x}} \right) \Rightarrow \mathbf{V}^{-1}. \, \mathbf{Y} = \mathbf{X} - \widetilde{\mathbf{M}}_{\mathbf{x}}$$
$$\Rightarrow \mathbf{X} = \mathbf{V}^{-1}. \, \mathbf{Y} + \widetilde{\mathbf{M}}_{\mathbf{x}} \qquad \text{Eq. 12}$$

According to equation 8 we can write:

$$\mathbf{V}^{-1} = (\mathbf{\Lambda}^{\mathrm{T}})^{-1} = (\mathbf{\Lambda}^{-1})^{-1} = \mathbf{\Lambda} = \mathbf{V}^{\mathrm{T}}$$
Eq. 13

Now, X matrix can be retrieved using equation 14:

For compressing **X**, some definitions are made as below[3]:

$$\begin{split} \mathbf{V}_{k} &\triangleq [v_{1}, v_{2}, \dots, v_{k}]^{T}{}_{M \times k}, \ k \in \{1, 2, \dots, M\} & \text{Eq. 15} \\ \mathbf{Y}_{k} &\triangleq \begin{bmatrix} Y(1, 1) & \cdots & Y(1, Q) \\ \vdots & \ddots & \vdots \\ Y(k, 1) & \cdots & Y(k, Q) \end{bmatrix}_{k \times Q} & \text{Eq.16} \end{split}$$

$$\widehat{\mathbf{X}} \triangleq \mathbf{V}_{k}^{\mathrm{T}} \cdot \mathbf{Y}_{k} + \widetilde{\mathbf{M}}_{x}$$
 Eq. 17

Considering the first k Eigen vectors from the M Eigen vectors (K<M), the $\hat{\mathbf{X}}$ matrix that is slightly different from **X** matrix, can be retrieved.

The $\hat{\mathbf{X}}$ matrix is compressed matrix which obtained from \mathbf{X} . So the compression ratio - the ratio of the required memory to save $\hat{\mathbf{X}}$ to the required memory to save \mathbf{X} - can be calculated as equation 18 shows:

mem1is the required memory to save \mathbf{V}_k^{T} , \mathbf{Y}_k , $\mathbf{\tilde{M}}_x$.

mem2is the required memory to save X.

$$\Rightarrow$$
 Memoryratio = $\frac{kM + kQ + M}{MQ}$ Eq. 19

3. IMPROVED PCA:

M. Mofarreh et.al. in [1] show that, it is possible to address the most erroneous pixels in the reconstructed images and improve the quality of images exactly at those pixels. According to [1], the \mathbf{Y}_k matrix is converted to $\widehat{\mathbf{Y}}_k$ as Eq. 20, which nonzero columns in the bottom of the matrix are applied to compensate the effect of error:

$$\widehat{\mathbf{Y}}_{k} = \begin{bmatrix} Y(1,1) & \cdots & Y(1,\alpha) & \cdots & Y(1,N^{2}-1) & Y(1,N^{2}) \\ Y(2,1) & \cdots & Y(2,\alpha) & \cdots & Y(2,N^{2}-1) & Y(2,N^{2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y(k,1) & \cdots & Y(k,\alpha) & \cdots & Y(k,N^{2}-1) & Y(k,N^{2}) \\ 0 & 0 & Y(k+1,\alpha) & \cdots & 0 & 0 \\ 0 & 0 & Y(k+2,\alpha) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & Y(M,\alpha) & \cdots & 0 & 0 \end{bmatrix}_{M \times N^{2}}$$

International Journal of Computer Applications (0975 – 8887) Volume 44– No.21, April 2012

Eq. 20

4. 2DPCA Method:

For applying PCA method, all image matrixes should be converted to vectors previously, so the resulting image vectors of images construct a high dimensional image vector space which its complexity doesn't let accurate enough calculations of covariance matrixes. 2DPCA technique uses 2D matrixes rather than 1D vectors in conventional PCA and this idea leads to smaller matrix spaces rather than high dimensional vectors. In fact, in 2DPCA technique, an image covariance matrix which its size is much smaller than covariance matrix of PCA is generated directly from image matrixes. So 2DPCA has two fundamental advantages over conventional PCA: easily construction of accurate covariance matrixes and smaller required time to calculate corresponding eigenvectors - because of the smaller size of covariance matrixes-.

Suppose **A** is an N × P image and \underline{X} is a P-dimensional unitary column vector. Projecting **A** onto \underline{X} by the linear transformation $\underline{Y} = \mathbf{A} \cdot \underline{X}$ leads to produce an N-dimensional projected vector \underline{Y} that is called the projected feature vector of image **A**.the goal of 2DPCA is to find the proper projection vector \underline{X} . from this point of view, following criterion is adopted:

That $\mathbf{S}_{\mathbf{x}}$ is the covariance matrix of the projected feature vectors of the training samples and $tr(\mathbf{S}_{\mathbf{x}})$ is the trace of $\mathbf{S}_{\mathbf{x}}$. Maximizing the criterion in 21 leads to find a projection direction \underline{X} , so that the total scatter of the resulting projected samples is maximized.

The covariance matrix \mathbf{S}_{x} is defined as:

$$\begin{aligned} \mathbf{S}_{\mathrm{x}} &= \mathrm{E}(\underline{\mathrm{Y}} - \mathrm{E}\underline{\mathrm{Y}})(\underline{\mathrm{Y}} - \mathrm{E}\underline{\mathrm{Y}})^{\mathrm{T}} \\ &= \mathrm{E}(\mathbf{A}\underline{\mathrm{X}} - \mathrm{E}(\mathbf{A}\underline{\mathrm{X}}))(\mathbf{A}\underline{\mathrm{X}} - \mathrm{E}(\mathbf{A}\underline{\mathrm{X}}))^{\mathrm{T}} \\ &= \mathrm{E}\left((\mathbf{A} - \mathrm{E}\mathbf{A})\underline{\mathrm{X}}\right)\left((\mathbf{A} - \mathrm{E}\mathbf{A})\underline{\mathrm{X}}\right)^{\mathrm{T}} \qquad \mathrm{Eq. 22} \\ \mathrm{So,} \quad \mathrm{tr}(\mathbf{S}_{\mathrm{x}}) &= \underline{\mathrm{X}}^{\mathrm{T}}\left[\mathrm{E}\left((\mathbf{A} - \mathrm{E}\mathbf{A})\right)^{\mathrm{T}}(\mathbf{A} - \mathrm{E}\mathbf{A})\right]\underline{\mathrm{X}} \qquad \mathrm{Eq. 23} \end{aligned}$$

Let
$$\mathbf{G}_{t} = \mathbf{E}((\mathbf{A} - \mathbf{E}\mathbf{A}))^{T}(\mathbf{A} - \mathbf{E}\mathbf{A})$$
.

 G_{t} is called the image covariance matrix and is a P × Pnonnegative definite matrix and can retrieved directly using the training image samples.

Suppose there are M images. Jth image is expressed by A_j , j = 1, 2, ..., M which is a N × P matrix. Let show the mean image of all images with \overline{A} , so the G_t matrix can be expressed as:

$$\mathbf{G}_{t} = \frac{1}{M} \sum_{j=1}^{M} (\mathbf{A}_{j} - \overline{\mathbf{A}})^{T} (\mathbf{A}_{j} - \overline{\mathbf{A}})$$
Eq. 24

And criterion 21 can be written as:

$$\mathbf{I}(\underline{\mathbf{X}}) = \underline{\mathbf{X}}^{\mathrm{T}} \mathbf{G}_{\mathrm{t}} \underline{\mathbf{X}}$$
 Eq. 25

This criterion is called the generalized total scatter criterion. The unitary vector <u>X</u>which maximizes the criterion is called optimal projection axis. This axis is a vector that maximize the J(<u>X</u>). Inother words, this is the eigenvector of \mathbf{G}_t which corresponds to the largest eigenvalues. Generally, only one vector can't satisfy all the requirements, so a set of projection orthonormal vectors $\underline{X}_1, \underline{X}_2, ..., \underline{X}_d$ are used to maximizing the J(<u>X</u>). In fact, optimal projection vectors are such orthonormal

eigenvectors of $\mathbf{G}_{\mathbf{t}}$ that have largest corresponding eigenvalues.

Optimal projection vectors are used to extracting features of images. For a sample image **A** projected feature vectors are:

$$\underline{Y}_{k} = A\underline{X}_{k}, \quad k = 1, 2, ..., d$$
 Eq. 26

So, a set of projected feature vectors for image A are obtained that called principal components (vectors) of the image. Note that the principal components in 2DPCA are vectors, in oppose with conventional PCA which their principal components were scalars.

Obtained principal vectors are used to build a $N \times d$ matrix that is called feature matrix or feature image as below:

$$\mathbf{B} = [\underline{Y}_1, \dots, \underline{Y}_d]$$
 Eq. 27

-

5. 2DPCA based Image Reconstruction

Principal components and eigenvectors (eigenfaces) are used to reconstruct images in PCA. Equivalently, in 2DPCA we retrieve images as following approach:

Suppose d orthonormal eigenvectors which correspond to d largest values of eigenvectors of covariance matrix \mathbf{G}_t are $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_d$. After projecting image onto these d axes, resulted principal component vectors are:

$$\underline{Y}_{k} = \mathbf{A}\underline{X}_{k}, \quad k = 1, 2, ..., d \quad \text{Eq. 28}$$
Now, let $\mathbf{V} = [\underline{Y}, ..., \underline{Y}_{d}]$ and $\mathbf{U} = [\underline{X}_{1}, ..., \underline{X}_{d}]$, so:
 $\mathbf{V} = \mathbf{A}.\mathbf{U} \quad \text{Eq. 29}$

Since $\underline{X}_1, ..., \underline{X}_d$ are orthonormal vectors, so we can reconstruct the image **A** easily. We call this reconstructed image $\widetilde{\mathbf{A}}$ which is calculated as below:

$$\widetilde{\mathbf{A}} = \mathbf{V}\mathbf{U}^{\mathrm{T}} = \sum_{k=1}^{\mathrm{d}} \underline{Y}_{k} \underline{X}_{k}^{\mathrm{T}}$$
Eq. 30

Suppose:

$$\widetilde{\mathbf{A}}_{k} = \underline{Y}_{k} \underline{X}_{k}^{\mathrm{T}}, \quad k = 1, 2, ..., d$$
 Eq. 31

Note that the size of \widetilde{A}_k is the same size of A and called a subimage of A. A combination of d subimages can reconstruct the image A. If the number of principal components are d = P, no information will be lost and $\widetilde{A} = A$. But if the number of principal components are less than P, i.e. d < P, some information have been lost and \widetilde{A} will be an estimation of A.

6. Improved PCA and 2DPCA for Video

In this section, the results of applying three mentioned methods: PCA, Improved PCA and 2DPCA for compression of 14 sample frames are presented. The frames are chosen from throwing an apple in a white background. Figure 1 shows the original frames. The frames are arranged from the top down and left to right.

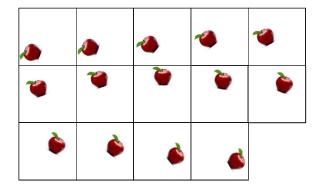


Fig 1: Original frames

Based on our background on PCA, 2DPCA and improved PCA in face database compression, we expect that 2DPCA will provide the best performance in video compression too but according to simulations, the improved PCA shows better results than two other methods. Figure.2 shows the first three frames of applying mentioned triple methods with same bitrates to the original frames. As it seen, the PCA method added fade effects to the reconstructed frames and applying 2DPCA causes vertical lines and black bands to the frames. So in this application (video compression), the best performance is provided by improved PCA method.

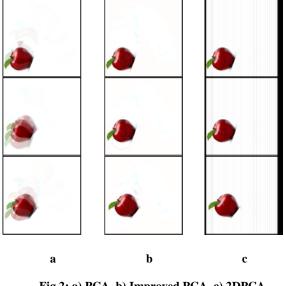


Fig 2: a) PCA, b) Improved PCA, c) 2DPCA

(Note to the vertical lines and black bands on the right side of 2DPCA results)

7. Mentioned Method

A piece of video can be supposed as a sequence of frames which have high correlation with each other. We expect that the correlation between neighboring frames be higher than the correlation between far ones. The mentioned approach in this paper is based on these correlations. Let us divide a $N \times M$ frame video to N subvideos that each subvideo has M frames. By choosing M little enough, the frames of subvideo may have similar frames with high probabilities, because of little distances between them. The main idea is that, we can put these frames as PCA inputs. According to high correlation between frames, we expect high performance using this technique. An important thing in video files is that, there are many similar sections in several frames and little parts of frames have sharp variations. As an example figure 3 represents 3 frames of a video. As it seen, some sections of these frames are constant in 3 frames.

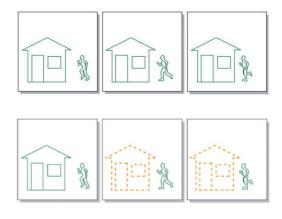


Fig 3: 3 sample frames to show constant and moving parts of the frames

Suppose 3 frames are 3 images. We apply PCA after converting them to vectors. As it investigated in the previous sections, PCA has some weaknesses in compression of images that belong to a moving object, so for improving the quality of erroneous pixels we should to increase the value of k, which in turn causes to unnecessary increase in other constant parts of the frame and this reduces the overall performance of method. For increasing the quality of moving parts of frames, we divided the frames in some bands, so we can apply the PCA to these bands separately. By this idea, constant and moving parts of frames have been separated from each other, so we can save constant parts of frames using little number of eigenfaces and save moving parts using more eigenfaces. This concept done by getting feedback from conventional PCA. Figure 4 shows the designed Encoder.

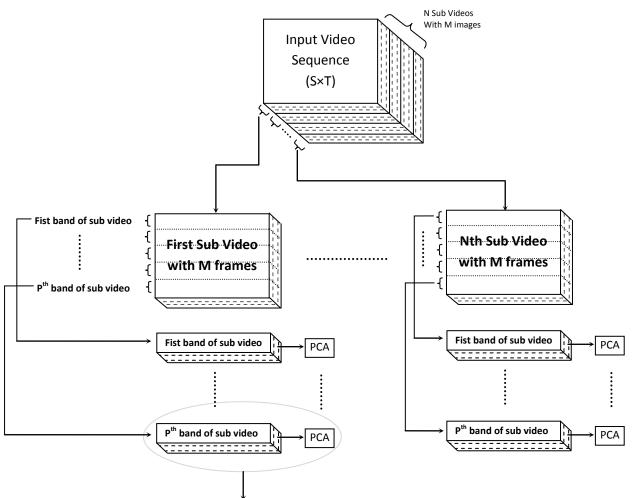


Fig 4: Block Diagram of designed Encoder for mentioned method

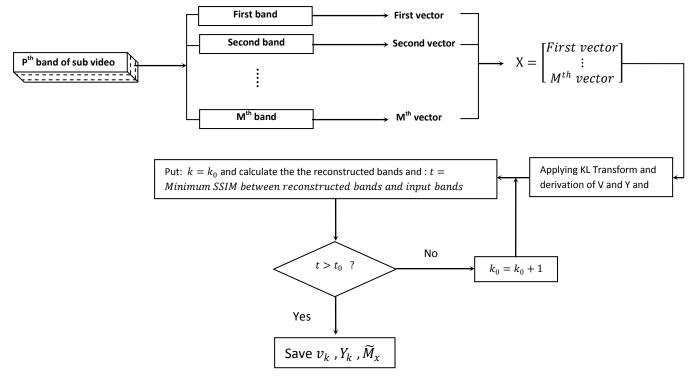


Fig 5: Diagram of designed Encoder for mentioned method, continued

Mentioned method is applied to video as follows:

- 1. Dividing input video to N subvideos that each subvideo has M frames
- 2. Dividing all subvideos to P bands with equal width
- 3. Converting *N* × *P* obtained frame groups to vectors and constructing *X* matrix. In fact, there are *N* × *P*, *X* matrixes.
- 4. Applying PCA to all $N \times P$ groups, according to required SSIM value. For each group, required number of eigenfaces should be saved.
- 5. Getting feedback of reconstructed frames and correcting most erroneous pixels according to SSIM.
- 6. Saving all \boldsymbol{Y}_k and \boldsymbol{V}_k matrixes.

One of the advantages of this method is the simplicity of its decoding procedure. The decoding procedure is as figure 6:

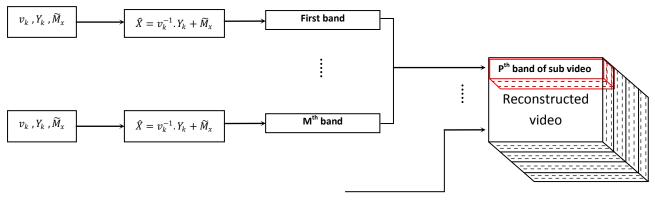


Fig 6: Designed Decoder for mentioned method

As figure 6 shows, decoding of coded information is done in a simple way. The required calculations for decoding a $R \times S$ video with $N \times M$ frames that each frame is splitted to P bands are as equation 32:

$$\widehat{\mathbf{X}} = \mathbf{v}_{M \times M}^{T} \cdot \widehat{\mathbf{Y}}_{k \times (\frac{R}{p} \times S)} + \widetilde{\mathbf{M}}_{x M \times (\frac{R}{p} \times S)}$$
Eq. 32

According to equation 32, it's clear that only adders and multipliers are needed for this decoder.

In contrast with other famous video compression methods like DVC Intra, H.264 Intra and MJPEG-DPCM, mentioned method in this paper provides better performance and less complexity, as represented in fig 7.

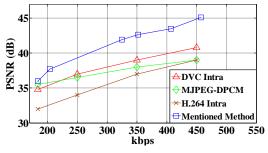


Fig 7: Curves of some low complexity encoders for video Compression applied to Akiyo frames

In order to illustrate the adaptively of mentioned method, note to two following frames which are first and 10th frames of 300 video sequence available frames of Akiyo at media.xiph.org/video/derf online at figure 8. These frames are divided into 8 bands and as it clear in figure 8, the bands 1, 2, 3, 5, 6, 7 and 8 are almost constant and its only the 4th band that changes. For compression with a distinct SSIM value for 50 frames, using a few numbers of eigenfaces for constant parts satisfies the quality requirements and it's necessary for variable parts to use more eigenfaces. The number of needed eigenfaces can be calculated by mentioned algorithm in previous sections.



Fig 8: first and 10th frames of Akiyo video sequence which is divided to 8 bands

8. ACKNOWLEDGMENTS

The authors would like to thank the reviewers for their beneficial comments and suggestions.

9. CONCLUSION

In this paper, PCA method is analyzed for video compression. Some other methods based on PCA like 2DPCA are also investigated and we showed that although they are high performance in face database compression but in video compression they didn't show such good performance. Mentioned method in this paper has several advantages like more simplicity in encoder and especially in decoder, because of using just adders and multipliers in their structures -in oppose to the other methods which use more complicated functions like inverse FFT and motion estimation functions-.

10. REFERENCES

- [1] Mostafa Mofarreh-Bonab and Mohamad Mofarreh-Bonab, "Face Database Compression By Hotelling Transform Using A New Method," 2nd World Conference on Information Technology, 23-27 Nov. 2011, Antalya, Turkey.
- [2] Mohamad Mofarreh-Bonab and Mostafa Mofarreh-Bonab, "A New Technique For Image Compression Using PCA," International Journal of Computer Science and Communication Networks, IJCSCN, vol. 2, pp. 111-116, 2012.
- [3] J. Yang, D. Zhang, A. F. Frangi and J.Y. Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, No. 1, pp. 131-137, 2004
- [4] Z. Gu, W. Lin, B. Lee, C. Lau, "Low-Complexity Video Coding Based on Two-Dimensional Singular Value Decomposition," IEEE Trans. Image Processing, vol. 21, No. 2, pp. 674-687, Feb 2012.
- [5] A.M. Aznaveh, F. Torkamani Azar, A. Mansouri, "Face Database Compression by Hotelling Transform Using Segmentation, Signal Processing and Its Applications, 2007. ISSPA 2007.
- [6] L. I. Smith, "A Tutorial on Principal Component Analysis [online], Available: csnet.otago.ac.nz/cosc453/student_tutorials/principal_co mponents.pdf, 2002
- [7] L. Sirovich and M. Kirby, "Low-Dimensional Procedure for Characterization of Human Faces," J. Optical Soc. Am., vol. 4, pp. 519-524, 1987.
- [8] M. Kirby and L. Sirovich, "Application of the KL Procedure for the Characterization of Human Faces," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 103-108, Jan. 1990.
- [9] X. Li, Y. Pang, Y. Yuan, "L1-Norm-Based 2DPCA," IEEE Trans. Systems Man and Cybernetics, vol. 40, No. 4, 2009.
- [10] Y. Li, X. Lei, B. Bai and Y. Zhang, "Information Compression and Speckle Reduction for Multi Frequency Polarimetric SAR Images Based on Kernel PCA," Journal of Systems Engineering and Electronics, vol. 19, pp. 493-498, 2008.
- [11] C. Yu, H. Qing and L. Zhang, "K2DPCA plus 2DPCA: An Efficient Approach for Appearance Based Object Recognition", 3rd International Conference on Bioinformatics and Biomedical Engineering, ICBBE, 2009.