Medical Image Compression Using Block-to-Row Principal Component Analysis (BTRPCA)

Sin Ting Lim1, Nurulfajar Bin Abd Manap2

Corresponding Author, Faculty of Engineering and Technology, Multimedia University (MMU) Melaka, Malaysia; E-mail: stlim@mmu.edu.my
Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia; E-mail: nurulfajar@utem.edu.my

Abstracts: Principal Component Analysis (PCA) is capable of completely decorrelating input data in the transform domain. However, PCA is limited in image compression because there is a need to transmit or store the eigenvectors of the input data over a communication link and thereby affects the rate-distortion performance. In an effort to improve rate-distortion performance, this work proposed a block-to-row PCA (BTRPCA) algorithm that employs the eigenvectors from the model image of the same image modality coupled with a row vectorization approach. It is found from this work that the proposed method achieves PSNR improvements of up to 10 dB compared to its PCA counterparts at compression ratio of 64:1. At the same compression ratio, the proposed BTRPCA managed to achieve PSNR of 40 dB while the comparing algorithms scores well below 40 dB at the same configuration. This approach successfully improves the rate-distortion performance by reducing the overwhelming side information and computation overhead associated with PCA.

Keywords: Principal Component Analysis, Medical Image Compression, Model Image, MRI Brain Scans, Telemedicine.

1. INTRODUCTION

With an increasing advancement in telemedicine technology, the demand to store and transfer medical images between health professionals within or beyond healthcare institutional has grown by leaps and bounds. In teleradiology, high resolution medical images such as CT and MRI scans are shared among medical professionals in a universal server called the Picture Archive and Communication System (PACS). Teleradiology does not only facilitate the diagnostic interpretation on the images, it also allows the doctors to provide consultation for patients located at remote places via telecommunication link that further improve the quality of care of a healthcare system. As the volume of pixel data grows, the need for a larger storage space increase, so as the speed requirement of a communication link. Therefore, the compression and decompression scheme play an important role in telemedicine to achieve sufficient reduction of image file size and at the same time maintain maximum information carried by the image.

Lossy image compression can be accomplished by exploiting the redundancy present in the image termed as spatial redundancies where there exist significant correlation or redundancies among neighbouring pixels. Figure 1 shows the basic building blocks of a lossy image compression system that are used to compress and decompress images.

![Fig 1. Basic building blocks of a lossy image compression system.](image-url)
The original image is first fed to a data mapping step whose goal is to use a mapping model to de-correlate the data. JPEG or JPEG2000 are popular data mapping models used for de-correlating the image data in such a way that the new signal has the whole information represented in a few transformed coefficients. The mapping operation itself does not achieve data compression. The second step in data compression is quantization and this is the step where loss of information occurred in order to achieve compression. Quantization allows one to reduce the amount of information at certain frequencies which are less perceived by the human eye. The final step is an encoding process that allows the quantized values to be represented by the lowest possible bits. These three blocks make up the building blocks of any image compression encoding process. The decoding process is simply the inversion of the encoding process where original image is to be restored from the binary code stream.

Principal Component Analysis (PCA), being also a lossy compression method, begins to receive attention in medical image compression due to its high capability in data decorrelation. PCA is also termed as Karhunen-Loeve (KL) transform in image compression [1]. It is a statistical approach that transforms the image data to an orthogonal plane and discards principal components with smaller eigenvalues. However, since the transferred data must include both eigenvector matrix and mean matrix, the conventional PCA failed to achieve high compression ratio (herewith refer to Standard PCA for simple convention). This limitation is apparent in a telemedicine setting where the transferred data is to be transmitted over a link.

PCA applications in image compression have been documented in various papers [2-12]. Standard PCA [3,5,7,8], block-by-block PCA [2] and block-to-row bi-directional PCA (BTR-BD) [9] are parts of the block-based PCA algorithm identified in the literature. Image compression using non-block-based PCA such as the algorithms proposed in [13-14] requires the eigenvector matrix for the whole image to be transferred together with the compressed data itself to enable reconstruction. This is because the representation is so image specific that the eigenvectors must be included with the compressed data to enable reconstruction. The aim of this study was to propose a block-based PCA method that reduce the amount of transfer data required to represent the image in a telemedicine setting. This new method termed as block-to-row PCA (BTRPCA) differs from the Standard PCA technique in a way that a model image is pre-selected to take part in both compression and decompression process. The detailed description of the algorithms can be found in Section II of this paper. The resultant compression results and performance evaluation of our proposed algorithm are presented in Section III. Finally, Section IV concludes the proposed approach.

2. MATERIALS AND METHODS

A. Available block-based PCA algorithms

Three block-based PCA algorithms have been proposed in the literature, but all have its own limitation in certain aspects. In this section, these techniques will be briefly reviewed and their drawbacks will be discussed in Section III. The readers are to refer to the original quoted papers for a more detailed description of the algorithms.

![Fig 2. An image is partitioned into blocks.](image)

Figure 2 illustrates the concept of how an MRI brain image of size $n$ can be partitioned into large number of small, non-overlapping blocks of size $n$ in block-based PCA algorithms. The number of blocks in an image, $b$ is defined as:
\[ b = \frac{MN}{n^2} \quad (1) \]

1) Standard Pca

The Standard PCA is the simplest and most widely used PCA algorithms in image compression. Each \( n \times n \) block of the pixels is rearranged in columns and the resulting matrix \( r(x, y) \) is organized as below:

\[
R = r(x, y) = \begin{bmatrix}
r(0,0) & r(0,1) & \ldots & r(0,b-1) \\
r(1,0) & r(1,1) & \ldots & r(1,b-1) \\
\vdots & \vdots & \ddots & \vdots \\
r(n^2-1,0) & r(n^2-1,1) & \ldots & r(n^2-1,b-1)
\end{bmatrix}_{r \times b}
\quad (2)
\]

The algorithm started by subtracting each element in \( r(x, y) \) with the mean of each data row \( \bar{m} \),

\[
R - \bar{m} = \bar{R} \quad (3)
\]

The crucial step in any PCA algorithm is to obtain the covariance matrix of \( \bar{R} \) and from the covariance matrix, the eigenvectors will be found. As shown in Equation (4), the feature matrix \( V \) contains all principal components arranged according to the descending order of the eigenvalues. It is sufficed to say that the first column of the \( V \) is the eigenvectors having the highest eigenvalues. \( k \) is the number of principal components selected in the feature matrix.

\[
V = V(x, y) = \begin{bmatrix}
V(0,0) & V(0,1) & \ldots & V(0,k-1) \\
V(1,0) & V(1,1) & \ldots & V(1,k-1) \\
\vdots & \vdots & \ddots & \vdots \\
V(n^2-1,0) & V(n^2-1,1) & \ldots & V(n^2-1,k-1)
\end{bmatrix}_{V \times k}
\quad (4)
\]

PCA deals with transforming the original data so that they are expressed in term of principal components. In this way, the PCA coefficients, \( Y \) can achieve dimensionality reduction when fewer principal components are taken into account in Equation (5). Eigenvectors are now used as the basic vectors for transformation. The PCA coefficients and the feature matrix are the compressed data that are ready to be quantized and losslessly encoded.

\[
Y = [ V^T \times \bar{R}]_{y \times k}\quad (5)
\]

2) Block-by-block PCA

Block-by-block PCA was proposed by Taur and Tao in 1996 with the aim to compress ROI and NROI arbitrarily by classifying the image blocks on mammograms. In this approach, the partitioned blocks have different number of principal components, depending on the classes of the tissue type. The background of the image is coded as zero and as for the tissue part the number of principal components used is chosen as the minimum number that can satisfy a predetermined error requirement. The compressed data contains the index that is used to tag the classes, the number of principal components needed to achieve the error requirements, feature matrix and the PCA transformed coefficients. Block-by-block PCA was one of the early studies that suggested classification of blocks for region-based compression. However, bit streams of the compressed data will increase substantially with the increase of the number of classes.
3) Block-To-Row Bi-Directional PCA (BTR-BD)

BTR-BD PCA was proposed with the aim of improving the rate-distortion performance. In this approach, two covariance matrixes $C_r$ and $C_c$ were computed from the row and column directions respectively from the mean subtracted matrix $\overline{R}$:

$$C_r = \overline{R}^T \overline{R}$$
$$C_c = \overline{R} \overline{R}^T$$  \hspace{1cm} (6)

Once the feature matrices $V$ and $U$ containing the number of principal components $k_u$ and $k_v$ are formed, the compressed data $Y$ is obtained as:

$$Y = [U^T \times \overline{R} \times V]_{(k_u \times k_v)}$$  \hspace{1cm} (7)

The compressed data for BTR-BD consists of the PCA coefficients and feature matrices obtained from row and column directions. The overhead information is expected to increase and it may not have improved the rate-distortion performance when quantization actually applied.

4) Proposed block-to-row PCA (BTRPCA)

To overcome the disadvantages posed by each method, a novel BTRPCA is proposed in this study. The mainstay of BTRPCA lies on the use of a model image for both compression and decompression process as illustrated in Figure 3 and Figure 4. PCA coefficients of the input image are derived using the universal basic vectors set from a model image. The selection criteria of the model image are that the image must be from the same class or same modality with the input image and the model image is free of feature abnormalities and noises.

Specifically, the input image and the model image are partitioned into fixed blocks of $n \times n$ that do not overlap each other. For example, the first block of the input image consists of intensity value which is denoted as:

![Fig 3. Proposed BTRPCA encoder](image-url)
The first block of the model image consists of intensity value is then denoted as:

$$H_{nn} = \begin{bmatrix} h(0,0) & h(0,1) & \cdots & h(0,n-1) \\ h(1,0) & h(1,1) & \cdots & h(1,n-1) \\ \vdots & \vdots & \ddots & \vdots \\ h(n-1,0) & h(n-1,1) & \cdots & h(n-1,n-1) \end{bmatrix}_{(n^2 \times b)}$$  \hspace{1cm} (9)

Each block in the input image and model image will then be vectorised into a $n^2 \times 1$ row vector of $x$ and $h$ vector respectively to obtain a transformed matrix $D_x$ and $D_m$ for the input image and model image respectively.

$$D_x = [x_1, x_2, x_3, \cdots, x_b]_{(n^2 \times b)}$$  \hspace{1cm} (10)

$$D_m = [h_1, h_2, h_3, \cdots, h_b]_{(n^2 \times b)}$$  \hspace{1cm} (11)

Where $b$ indicates the number of blocks and $x_i$ and $h_i$ contain all elements within a block.

$$x_i = \begin{bmatrix} f(0,0) \\ f(1,0) \\ \vdots \\ f(n^2-1,0) \end{bmatrix}_{(n^2 \times 1)}$$  \hspace{1cm} (12)

$$h_i = \begin{bmatrix} h(0,0) \\ h(1,0) \\ \vdots \\ h(n^2-1,0) \end{bmatrix}_{(n^2 \times 1)}$$

The mean vector for both $D_x$ and $D_m$ are calculated with mean $m$ and $m_m$ respectively and the matrices are adjusted by the mean to obtain $\overline{D_x}$ and $\overline{D_m}$. The PCA coefficients will be derived based on the feature matrix of the model image. Hence the feature matrix for the model image $V_m$ containing all the selected principal components is formed as follows:

$$V_m = \overline{D_m}$$
\[ \mathbf{V}_m = [\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_k]^{(n^2 \times k)} \]  

(13)

Where \( \lambda \) is the eigenvector from the row directions and \( k \) is the number of chosen principal components. The PCA coefficients \( \mathbf{Y} \) in BTRPCA approach is therefore derived:

\[ \mathbf{Y} = [\mathbf{V}_m^T \times \mathbf{D}_x]_{(k \times b)} \]  

(14)

The same model image will be used as the receiver end to reconstruct the compressed data as in Equation (15).

\[ \mathbf{X}_{PCA} = [\mathbf{V}_m^T \times \mathbf{Y}] + m \]  

(15)

The quantization step is the crucial part of any transform coding process. In BTRPCA, the PCA coefficients and mean vector are quantized by a uniform midtread quantizer similarly adopted by the JPEG [15] where the quantized value is calculated as

\[ \hat{y} = w \cdot \text{round} \left( \frac{y}{w} \right) \]  

(16)

\( w \) being the step size of the quantization level, \( y \) the individual element in PCA coefficient matrix \( \mathbf{Y} \) and function \( \text{round}(x) \) returns each element in the matrix to the nearest integer. In this study, the \( w \) is set according to the compression performance. For example, when \( w \) increases the number of coefficients with zero values increases resulting in reduction of number of bits. Run-Length Encoding (RLE) is further applied to encode the zero-valued coefficients [16].

To reconstruct the image, the process is exactly the inverse of the compression process as shown in Figure 4. The bit streams of the PCA coefficients are de-quantized and the PCA coefficients are decompressed by the corresponding inverse BTRPCA transformation using associated feature matrix and mean directly obtained from the model image. Akin to the private-key encryption network in cryptography, the encoder and decoder use the same common key - the model image to compress and recover the compressed data closest to the original image. With that, there is no need to store different set of eigenvectors in the compressed stream and overhead information can be greatly reduced.

### 1.1 Image Database

Two test images from two public datasets were selected for the experiments: MRI brain image [17] and retinal fundus image [18]. The size of the MRI brain image and retinal fundus image are and respectively. A brain lesion is clearly visible on the MRI brain image as shown in Figure 5 while the optic disk, fovea and blood vessels are clearly visible on the colour fundus image shown in Figure 6. The model images are selected so that they are free of any abnormalities and noise. Block size of \( n = 8 \) is used in this experiment.
Fig. 1. Image dataset for MRI image (a) Model image (b) Test image.

Fig 2. Image dataset for fundus image (a) Model image (b) Test image.

3. RESULTS AND DISCUSSION

The performance of the proposed BTRPCA is compared against Standard PCA algorithm, block-to-block PCA algorithm and BTR-BD PCA algorithm. To allow fair compression, all the algorithms under investigation are quantized and encoded using the same approach as in BTRPCA. The colour input image will first be separated into three channels (R, G, B-channels) and each channel is fed into their respective algorithms. Each compressed gray image will be combined to form the reconstructed colour image and each reconstructed image is evaluated based on the CR, PSNR and computation time.

<table>
<thead>
<tr>
<th>BTRPCA</th>
<th>Standard PCA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="BTRPCA MRI" /></td>
<td><img src="image" alt="Standard PCA MRI" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block-by-block PCA</th>
<th>BTR-BD</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Block-by-block PCA MRI" /></td>
<td><img src="image" alt="BTR-BD MRI" /></td>
</tr>
</tbody>
</table>

Fig. 2. The reconstructed MRI brain scans obtained by proposed BTRPCA and the comparing algorithms where \( bpp = 0.125 \).
The reconstructed images for all the algorithms compressed at \( \text{bpp} = 0.125 \) are displayed in Figure 7 and Figure 8. Although all features are still discernible, the blocking effects are seen on the edge of MRI brain image compressed using block-by-block and BTR-BD algorithms at \( \text{bpp} = 0.125 \). The mechanisms of block-by-block algorithm to work locally for each block rendered the algorithm to be highly sensitive to the slight change of intensity in the area. In block-by-block algorithm, when the block includes an image edge, the edge will be degraded such that the block boundary appears like the edge. On closer examination as seen in Figure 9, blocky artifacts are seen clearly on the MRI brain image compressed by block-by-block algorithm at \( \text{bpp} = 0.125 \). Visual inspection revealed that test images compressed by BTR-BD suffer more pronounced image degradation in which the compressed test images are blurred and contains noise structures.

Out of the four algorithms, the MRI brain images compressed by BTRPCA and Entire Image PCA still manage to uphold their image fidelity at high compression rate of \( \text{bpp} = 0.125 \). Edges, lines, textured areas and regions of constant brightness look pleasing although there are spurious texture seems to be distorted on MRI image compressed by Standard PCA. It is interesting also to take note that the tested algorithms work well on the retinal fundus image that has higher image resolutions and more uniform pixel distribution as seen in Figure 9. These observations are parallel to their higher reported PSNR value. It is also important to take note that the original and reconstructed images are virtually indistinguishable by human eye with PSNR score of 40 dB and above [19].
Table I. PSNR of Proposed BTRPCA and the Comparing Algorithms on the MRI Brain Image (Size = 512x512)

<table>
<thead>
<tr>
<th>Seq.</th>
<th>bpp</th>
<th>BTRPCA</th>
<th>Standard PCA</th>
<th>Block-by-block</th>
<th>BTR-BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>54.3597</td>
<td>52.1415</td>
<td>43.2585</td>
<td>38.0603</td>
</tr>
<tr>
<td>2</td>
<td>0.50</td>
<td>51.6431</td>
<td>49.2648</td>
<td>40.1457</td>
<td>34.4343</td>
</tr>
<tr>
<td>3</td>
<td>0.125</td>
<td>40.3697</td>
<td>36.6891</td>
<td>32.5588</td>
<td>29.8365</td>
</tr>
</tbody>
</table>

Table II. PSNR of Proposed BTRPCA and the Comparing Algorithms on the MRI Brain Image (Size = 1024x1280)

<table>
<thead>
<tr>
<th>Seq.</th>
<th>bpp</th>
<th>BTRPCA</th>
<th>Standard PCA</th>
<th>Block-by-block</th>
<th>BTR-BD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00</td>
<td>58.1327</td>
<td>56.1596</td>
<td>52.0369</td>
<td>42.6699</td>
</tr>
<tr>
<td>2</td>
<td>0.50</td>
<td>53.0019</td>
<td>49.3955</td>
<td>41.6484</td>
<td>38.3860</td>
</tr>
<tr>
<td>3</td>
<td>0.125</td>
<td>41.6789</td>
<td>39.4441</td>
<td>33.3464</td>
<td>32.2855</td>
</tr>
</tbody>
</table>

Table I and Table II show the comparison results for BTRPCA and the comparing algorithms for MRI and fundus images respectively. It can be seen from the tables that PSNR values obtained using BTRPCA algorithm is higher than BTR-BD, block-by-block and Standard PCA algorithm at the same configuration. The PSNR result is in agreement with the visual observation for the test images where image artifacts are clearly seen on MRI brain scans compressed with block-by-block and BTR-BD algorithm at bpp = 0.125. Taking 40 dB as the minimum acceptable PSNR for any compression algorithm, MRI image compressed using BTRPCA managed to achieve 40 dB at bpp = 0.125 while the comparing algorithms scores well below 40 dB at the same configuration. This is an important result as it indicates that only BTRPCA can compress image with suitable quantitative quality parameter at low bit rate. The PSNR improvement of BTRPCA over the Standard PCA, block-by-block and BTR-BD algorithms for MRI brain scan are 3.6806 dB (10.03%), 7.8117 dB (23.99%) and 10.5332 dB (35.3%) respectively.

Figure 10 illustrates the average computation time taken to compress an image including the encoding and decoding process. All algorithms generally take longer time to process the fundus image due to its higher resolution. In comparison, the proposed algorithm achieves the shortest average computation time due to the time saved in encoding and decoding the eigenvectors of the covariance matrix in the bit streams. BTRPCA process slightly faster than Standard PCA due to the time saved in calculating the covariance matrices for all images since the eigenvectors for the model image is pre-determined once and stored for each compression process.

When using the paired t-test to compare mean PSNR values of the images compressed by BTRPCA with block-by-block and BTR-BD algorithms, it is found that the mean PSNR in the images compressed by BTRPCA increased
significantly at all tested compression levels \((p < 0.05)\). By comparing BTRPCA with Standard PCA, it is found that the mean PSNR for

BTRPCA increased significantly at \(bpp = 0.5\) (at the 5% significant level).

The presented results show that BTRPCA algorithm is the best compression scheme among all the evaluated algorithms and it is the most optimal scheme for MRI brain image. It is expected that BTR-BD produces highest bit-rate and computation time among all testing algorithms because it requires greater coefficient set for image representation. The performance of block-by-block PCA is restricted largely due to low compression efficiency of applying PCA on each block. In this case, the correlation within the image blocks is not examined. Besides, block-by-block PCA generates additional overhead information to include the index that is used to tag the groups and the number of principal components for each group. In this study only two groups with same number of principal components have been considered.

The BTRPCA without the trade-off to store the additional statistical elements together with the compressed data enables the algorithm to outperform its PCA variants, particularly its conventional counterpart - Standard PCA. Unlike Standard PCA that must encode and transmit the eigenvectors of the input data covariance matrix, the improvement of BTRPCA comes from deriving the compression data from the eigenvectors of a model image that belongs to the same image modality. The statistics of the model image at the level of small block tend to be similar to the test images and these are the exact data that are used to compress and reconstruct the image. The results in this phase demonstrates that same set of eigenvectors of the same image modality can be conveniently applied in PCA transformation, thereby reducing the main limitation of PCA algorithm in having overwhelming side information and computation overhead. It is therefore safe to say a model image generalizes well to the other image in PCA algorithm as long as they are acquired from the same image modality. Furthermore, row vectorization approach utilized in BTRPCA contributes to the performance gain by reducing the correlation within the image blocks.

4. CONCLUSION

The need to transfer and store hundreds to thousands of radiological images per day drives the development of an efficient compression scheme. The answer to its efficiency involves the amount of data that can be compressed and how well information is retained. The efficiency performance of PCA algorithm in data compression for single image is known to be hampered by the need to store the eigenvectors of the covariance matrix in the bit streams. The capability of PCA on single frame medical image can be explored based on the fact that medical images from the same modality tend to have similar image statistics. To circumvent this limitation, this study developed a BTRPCA that adapts the transform with the use of eigenvectors basic sets from a model image coupled with a row vectorization approach, eliminating the need to encode and transmit each covariance matrix over a communication link. Experimental results show that BTRPCA can achieves PSNR improvements of up to 10 dB compared to its PCA counterparts at compression ratio of 64:1.

It is intuitive to understand that the results might be affected if different model image were to be employed in this study. Hence the future work will probe on how different model image affects the performance of the proposed algorithm and to look for other defining criteria in choosing a model image, if there is any. On top of that, the true contribution of this approach shall also be tested in a larger scale on a different image dataset and compare with different state-of-the-art compression algorithms such as JPEG and JPEG2000. One may also be interested of how well the proposed algorithm is to be applied to non-medical images. The proposed method is readily incorporated with other image compression technique such as a region-based image compression algorithm to yield better rate-distortion algorithm.

ACKNOWLEDGMENT

This study was supported in part by grant number FRGS/2012/FKEKK/TK02/03/1/F0012 and PRGS/2012/Tk02/FKEKK/01/1/T0002 from the Ministry of Higher Education of Malaysia.
REFERENCES


[17] E-Health Lab, Department of Computer Science, University of Cyprus, 2011. MRI Lesion Segmentation in Multiple Sclerosis Database.


DOI: https://doi.org/10.15379/ijmst.v10i2.3002

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.