A Novel Block Toeplitz Matrix for DCT-based, Perceptually Enhanced Image Fusion

Fayadh Alenezi, Ezzatollah Salari
1Dept. of EECS, University of Toledo, Ohio, United States

Abstract
Image fusion aims at increasing the information content of the composite image. However, many existing transform-domain image fusion methods have failed to properly enhance the finer details in the input images, and the perceptual quality of the fused image often suffers as a result of fusion process. This paper describes and evaluates a novel image pre-processing technique based on the application of a block Toeplitz matrix as a part of a Discrete Cosine Transform (DCT) -based fusion method. The proposed DCT implementation seeks to enhance the finer details of all input images prior to fusion. A post-fusion stage aimed at adjusting image contrast and improving smoothness is then added in order to improve quality of fused image. The proposed method is applied to a set of medical images, and its results are evaluated based on objective performance measures such as entropy, standard deviation, and root-mean-squared error. These indicators are then compared to values obtained from other existing transform-based fusion methods, showing a significant performance improvement across all experiments.

Keywords
Block Toeplitz Matrix, Contrast Adjustment, Histogram Equalization, Bilateral Filter, Gaussian Kernel, Medical Image Fusion.

I. Introduction
Image fusion is the process of merging several information-bearing images into a single composite image [1]. The resulting fused image is more useful and suitable for many applications than any of the input images considered separately [2]. Image fusion techniques are generally chosen based on the application under consideration [3]. Image fusion aims at extracting all valuable information from the source images, while preventing the formation of artifacts in the final image. It also aims at enhancing information from the source images by jointly displaying all the relevant information they contain [4].

Numerous image fusion methods have been developed [5]. Pixel-averaging image fusion is the simplest technique; it is often associated with reducing the contrast of the fused image. In order to overcome the side effects of the pixel-averaging method, many fusion techniques have been developed based on multi-resolution [6-7], multi-scale [8] and statistical signal processing. Some researchers have also explored the fusion of images in a transformed domain such as those based on the Discrete Cosine Transform (DCT), which is aimed at enhancing the contrast in the resulting fused image [9-10].

DCT has been used extensively in image compression since it is not computationally expensive [11]. Image fusion within the DCT domain is also popular since the process is highly efficient and less time consuming than other methods [10-11]. The DCT also has better energy compaction capabilities and is therefore preferred to other transforms [12]. Energy compaction is quantified by the sum of pixel squared transform coefficients that are concentrated in a small fraction of the transformed image. DCT has a tendency to speed the decay of the image spectra, which results in better energy compaction. In contrast, the Discrete Fourier Transform (DFT), often used as a benchmark reference, offers relatively poor energy compaction, which leads to discontinuities at the signal borders.

Energy compaction in the DCT is enabled by a better image approximation, which is related to the eigenvectors of the Toeplitz matrices [13]. The block Toeplitz matrices in the DCT explain its computational efficiency [14]. These features of the DCT have made it suitable for a range of applications, much wider than those of other fast transforms like the DFT, the Walsh-Hadamard Transform (WHT), the Karhunen-Loeve transform (KLT) and the Haar Transform.

The 2D form of the DCT is the most common and widely used for block-based image processing [15-16]. Yet problems with smoothness and boundary discontinuities between adjacent blocks have motivated many researchers to come up with various solutions and approaches, including using a modified implementation of the 2D DCT to solve discontinuity issues along boundaries. One of the main variations is based on the incorporation of direction information in the DCT (Directional DCT or DDCT), which leads to the formation of smoother boundaries [17]. However, this approach has faced technical difficulties such as lack of coherence in the transform, leading to poorer results than those produced by the original DCT.

Another attempt to solve the directionality issue in the DCT led to the development of mode-dependent directional transforms (MDDT) from KLT, which uses prediction residuals from training video data. In this approach, a block Toeplitz matrix is used to reduce the number of transform matrices needed during coding [18]. The block Toeplitz matrix helps in giving spectral interpretation of channel diversity [19]. This means more features are extracted for visualization. Images with better visualization of features can be interpreted more easily and are therefore more useful in diagnosis.

While this approach has had some success, it also has drawbacks: it is time-consuming, and optimal results are specific to the data sets used. A recently developed approach where each image pixel is viewed as a node in a graph, and edges are seen as connectivity relations, has resulted in the design of more efficient edge-aware transforms [20-21]. However, the results of this approach suffer from distorted and low information content images when compared to other initial transforms [15, 20-21].

In this paper, a new variation of the DCT is introduced, in which the smoothness and boundary discontinuity issues associated with the previous methods have been mitigated by using a novel block Toeplitz matrix. The resulting DCT coefficients are fused in the transformed domain using the standard built-in MATLAB® function `imfuse`. Contrast adjustment aimed at reducing contrast inhomogeneities that may have resulted from DCT processing follows the fusion process. The final image is also smoothed and processed through a Gaussian-kernel bilateral filter for noise suppression. These later stages of the algorithm aim at improving the finer image details.
The rest of this paper is organized as follows: Section II introduces the proposed method, giving a detailed explanation of the novel block Toeplitz matrix used; Section III presents and discusses the simulation results, and Section IV provides conclusions and points at directions of future research.

II. Proposed Method

A. Overview

Properties of linear discrete transforms such as scaling, shifting and convolution help achieve image transformation [22], which in turn contributes to the removal of redundancy between neighboring pixels [23]. The action of the DCT on images is based on approximations of eigenvectors of Block Toeplitz [24] matrices.

Block matrices have wide range of application such as in communication, information theory and Signal Processing [25]. Images are non-infinite and non-periodic since the boundaries of the image have no correlation among themselves [23]. A 2D image may contain \((512)^2\) pixels and the grey level of a pixel at position \((i,j)\) is represented by amplitude \((i,j)\) between 0 and 255, using 8 bits per pixel. Most existing efforts around DCT-based image processing have concentrated on improving the computation of the blocks while disregarding the role played by vector \(h\) in improving the end results of the processed image, though the method investigated in [26] does note that using an irreducible singular nonzero block Toeplitz matrix in the DCT increases the correlation between the pixels. It is well known that a higher correlation within pixels during transformation enhances the complexity and accuracy trade-offs within pixels intensities [27]. The method proposed in this paper is expected to increase the finer details of the image by maximizing entropy of the image which may have been reduced by the DCT processing. The final image is also smoothed and passed through a Gaussian-kernel bilateral filter to reduce noise. These steps are illustrated in Fig. 1: first, the proposed implementation of the DCT is applied to the input images in order to perform image fusion in the transformed domain; second, the contrast adjustment seeks to improve image quality [28]; and finally, the Gaussian-kernel bilateral filtering aims at smoothing the final image and mitigating the effect of noise [29].

![Fig. 1: Schematic Representation of the Proposed Fusion Method.](image)

B. Block Toeplitz Matrix for DCT-based Fusion

A block Toeplitz matrix is a special kind of matrix in which blocks are repeated along its diagonal [30]. Block Toeplitz matrices provide a good image approximation due to their eigenvectors, which to date has been the sole motivation behind their application to DCT [31]. A Toeplitz matrix constitutes the autocovariance matrix of a first-order stationary Markov process, and it has better performance than earlier transforms such as Discrete Fourier Transform (DFT) [32]. On the other hand, existing implementations of the DCT are based on a matrix with entries \(\rho^{i\pm j}\), representing the covariance matrix of a useful class of signals, where coefficient \(\rho\) measures the correlation between nearest neighbors. The image resulting from these DCT implementations is lossy and often displays a lower quality than the input image [23]. The method proposed in this work seeks to overcome this.

In this study, a block Toeplitz matrix is proposed, in which the correlation coefficient is \(\rho = \Delta x\), where \(\Delta x\) denotes the small change in pixel values, and the block size is \(N = 8\). The consideration of small changes in pixel intensities is viewed as the main contribution of this research effort. Modeling small variations in pixel intensities enhances the totality of image contrast, hence improving human visual perception of the final image. Incorporating small pixel intensities in the Toeplitz matrix design also increases the smoothness of the intensity trends in the final image and improves the image finer details. The 8 by 8 block Toeplitz matrix used is of the form represented by

\[
A = \begin{bmatrix}
\beta & \alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha \\
\alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha & \beta \\
\beta & \alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha \\
\alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha & \beta \\
\beta & \alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha \\
\alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha & \beta \\
\beta & \alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha \\
\alpha & \beta & \alpha & \beta & \alpha & \beta & \alpha & \beta
\end{bmatrix}
\]

The output coefficients from the block Toeplitz matrix DCT operation are fused via ‘imfuse’, and the results transformed back via inverse DCT and post-processed, as explained next.

C. Contrast Adjustment Using Adaptive Histogram Equalization

Contrast enhancement using adaptive histogram equalization has been used due to its simplicity and efficiency [33]. Adaptive histogram equalization has performed excellently on both natural and medical images [34]. Contrast enhancement through adaptive histogram equalization uses probability distribution to adjust the gray level of an image.

In the proposed algorithm presented in Fig. 1, the pixels of the infused image are gray mapped through the use of gray operations and transformed into a histogram which is smooth and uniform [33]. The mapping and transformation of infused image pixels ensures contrast enhancement is achieved.

Assuming the gray value of the infused image is \(r\) such that \(0 < r < 1\) and the probability density function of \(r\) is given by \(p(r)\), the gray value of the pixel of the output image (the image that will be used for smoothing by the bilateral filter using Gaussian Kernel) is \(s\) such that \(0 < s < 1\). The probability density function of the output image is given by \(p(s)\) and the mapping function is represented by \(s = T(r)\). If every bar of the histogram has equal height, then

\[
p_s(s)ds = p_r(r)dr\tag{2}
\]

Suppose the mapping function \(s = T(r)\) is a monotonically increasing function in the interval of the histogram heights and its inverse function \(r = T^{-1}(s)\) is also a monotonic function, from (2) we deduce that [33]

\[
p_s(s) = \left[\frac{p_r(r)1}{ds \frac{dr}{ds}}\right]_{r=T^{-1}(s)} = p_r(r)\frac{1}{p_r(r)} = 1\tag{3}
\]
If we assume based on (3) that the entropy of the final image is given by [33]

\[ E = \sum_{r=0}^{n-1} e(r) = -\sum_{r=0}^{n-1} p_r \log p_r \]  

(4)

The maximum entropy of the whole image is achieved by ensuring the histogram of the image has uniform distribution, that is, when \( p_0 = p_1 = \cdots = p_{n-1} = \frac{1}{n} \). The image with maximum entropy from this stage is then taken to the next level for smoothing and edge completeness.

**D. Smoothing by bilateral filter using Gaussian Kernel**

Image edges represent high frequency components, and as such, they can be smoothed by lowpass filtering. In the proposed algorithm, images are processed post-fusion in order to suppress noise and promote uniform smoothing by a bilateral filter that is defined by the combination of a domain kernel and a range kernel, as described next.

Let us assume that the image created from section II-C has a pixel of interest \( p \) with intensity \( y_p \). Pixel averaging is set to operate within a \( \Omega \)-neighborhood of \( p \). If \( q \) is a pixel within the \( \Omega \)-neighborhood of \( p \) with intensity \( y_q \), the output of the filter obtained by using the kernel \( \phi_{p,q} \) is represented as [29]

\[ \hat{x}_p = h_p^{-1} \sum_{q \in \Omega} \phi_{p,q} y_q \]  

(5)

where \( h \) is a normalizing factor given by \( h_p = \sum_{q \in \Omega} \phi_{p,q} \) and the filter kernel \( \phi_{p,q} y_q \) is formed by combining the domain and range kernels as

\[ \phi_{p,q}(y_p, y_q) = w_{p-q} r(y_p - y_q) \]  

(6)

where \( (y_p - y_q) \) is the range kernel and \( w_{p-q} \) represents the domain kernel.

In the range kernel, pixels with relatively similar intensities are assigned higher weights, and vice versa. Therefore, the range kernel quantifies the similarity between pixel intensity \( y_p \) and neighborhood pixel intensity \( y_q \),

\[ r(y_p - y_q) = \exp \left( -\frac{|y_p - y_q|^2}{2\sigma_r^2} \right) \]  

(7)

where \( \sigma_r \) controls the rates of Gaussian decay and directly determines degree of smoothing. At the same time, the domain kernel depends only on the geometric distance between the pixel of interest \( p \) and neighboring pixel \( q \), following also a Gaussian shape represented by

\[ w_{p-q} = \exp \left( -\frac{|p-q|^2}{2\sigma_d^2} \right) \]  

(8)

where \( |p-q| \) is the geometrical distance between the pixels \( p \) and \( q \), and \( \sigma_d \) is the Gaussian decay parameter. Pixels that are closer to \( p \) are assigned higher weights; domain kernel coefficients decay according to \( \sigma_d \) as the distance increases. This kernel acts only on the domain between pixels \( p \) and \( q \), which explains why it is referred to as domain kernel. Pixels outside such domains are not considered for the weighted average defined in (5).

**III. Simulation Results**

**A. Experimental Setup**

Three different examples are used to evaluate the proposed method. Each example aims at fusing one Computer Tomography (CT) image and one Magnetic Resonance image (MRI) into a single composite image. The values proposed for the block Toeplitz matrix defined in (1) are listed in Table 1. The final results were evaluated using three objective quantitative measurements: standard deviation (SD), root-mean-squared error (RMSE) and entropy. These bases for evaluation were selected based on their use as descriptors of image human perceptual quality, information content and signal-to-noiseratio (SNR). In addition, these criteria were chosen because they are available for comparison with previously published fusion methods. SD is a measure of contrast. Images with high standard deviation show better human visual properties. RMSE measures the amount of change per pixel due to processing. Lower RMSE is an indicator of lower noise levels in the image and thus improved visual features and finer image details. Entropy measures image quality in terms of information content and visibility of finer details, including textural uniformity [35].

The example images and results are shown in Figs. 2, 3, and 4; the corresponding measurements are presented in Table 2. The best results were then compared in Table 3 and Figure 5 with measurements from previously published (existing) techniques.

**Table 1: List of Proposed Parameters of the Block TOEPLITZ MATRIX Used in Experiments 1, 2 and 3.**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta )</td>
<td>1</td>
</tr>
<tr>
<td>( \xi )</td>
<td>0.1</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Fig. 2: Example 1. Input images are the CT (above) and MRI (below) images. The result of the final fusion is also shown (right).

Fig. 3: Example 2. Input images are the CT (above) and MRI (below) images. The result of the final fusion is also shown (right).

Fig. 4: Example 3. Input images are the CT (above) and MRI (below) images. The result of the final fusion is also shown (right).
Table 2 shows the performance of the proposed algorithm in relation to other existing algorithms. The first column of Table 2 lists the examples used in the algorithm, the second column contains the entropy values produced for each example by the proposed and existing algorithms, the third column contains RMSE values, and the last column lists the sources of the existing algorithms. Table 2 and Fig. 5 describe the performance of the proposed algorithm in comparison with some other methods including Contourlet Transform (CT) [36], Discrete wavelet transform (DWT) [37], and Shearlets and Human Feature Visibility (SHFV) [38]. The results presented in Table 2 are also plotted for convenient visualization in Fig. 6, Fig. 7 and Fig. 8. The numbers clearly show the improvements the proposed method offers in comparison to the benchmarked methods. The entropy values of the proposed method were greatly improved due to the use of the DCT matrix and the smoothing process after the fusion. Similarly, the SD values are higher than in all the other methods, which proves that the fused results have high contrast, which is desirable. Finally, the low RMSE values of the proposed method suggest that the final fused results demonstrate improved noise suppression.

Table 2. Comparison of Image Quality Metrics for Various Fusion Algorithms, Including the Proposed One

<table>
<thead>
<tr>
<th>Example #</th>
<th>Algorithm</th>
<th>Entropy</th>
<th>Standard Deviation</th>
<th>RMSE</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Proposed</td>
<td>8.1634</td>
<td>82.9216</td>
<td>0.0179</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>7.1332</td>
<td>54.1504</td>
<td>0.1662</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>DWT</td>
<td>6.9543</td>
<td>47.2304</td>
<td>0.2703</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>SHFV</td>
<td>7.6572</td>
<td>56.7993</td>
<td>0.1164</td>
<td>[38]</td>
</tr>
<tr>
<td>2</td>
<td>Proposed</td>
<td>8.1403</td>
<td>78.3379</td>
<td>0.0129</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>6.9351</td>
<td>46.6294</td>
<td>0.2538</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>DWT</td>
<td>6.6997</td>
<td>41.4623</td>
<td>0.2889</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>SHFV</td>
<td>7.3791</td>
<td>55.8533</td>
<td>0.2410</td>
<td>[38]</td>
</tr>
<tr>
<td>3</td>
<td>Proposed</td>
<td>6.9895</td>
<td>94.2736</td>
<td>0.0177</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>CT</td>
<td>6.8824</td>
<td>43.1963</td>
<td>0.2422</td>
<td>[36]</td>
</tr>
<tr>
<td></td>
<td>DWT</td>
<td>6.5198</td>
<td>42.0087</td>
<td>0.3142</td>
<td>[37]</td>
</tr>
<tr>
<td></td>
<td>SHFV</td>
<td>6.9467</td>
<td>44.2937</td>
<td>0.2133</td>
<td>[38]</td>
</tr>
</tbody>
</table>

A summary of the performance of the proposed method in relation to some existing medical image fusion methods based on other transforms is also shown in bar graphs in Fig. 6, Fig. 7 and Fig. 8. The bar graphs demonstrate the improvements in the key objective measures in the comparison between the proposed method in relation to the benchmark methods used in this paper. The orange bar represents standard deviation values, and it shows clearly that the proposed method yields results up to 123% greater than the existing methods used in the comparison. The blue bar illustrates entropy values, and it shows that the proposed method yields the highest values, with an improvement of 24%. Lastly, the gray bar represents the RMSE values and it shows that the proposed method yields the lowest values of these, with a reduction of up to 96%.
B. Discussion

The results presented in Table 2 and in fig. 6, fig. 7 and fig. 8 show that the proposed method exhibits better quality indicators than any of the reference methods, as measured by entropy, standard deviation and RMSE. Further analysis from this performance comparison indicates that entropy values for the proposed algorithm are greater than for any other technique. Images with higher entropies depict visibly finer details and maximum entropy in turn suggests better image quality, increased smoothness and edge completeness and preservation. Higher entropy values are due to better extraction and enhancement of image finer details, which can be attributed to the use of the block Toeplitz DCT matrix, resulting in an improvement in the amount of information contained in the final images as compared to input images.

The proposed method also yielded results with better standard deviation than any other reference method. Post-fusion processing aimed at enhancing edge completeness, smoothing and reduction of noise, and is the probable reason behind these very high standard deviation values. A significant reduction in RMSE is also observed in these indicators, which could also be associated with suppressed noise in the final image in the proposed technique.

IV. Conclusion

Existing state-of-the-art medical image fusion techniques have failed to address the loss of fine details and the lack of image contrast. This paper presents new transform-domain image fusion technique based on a block Toeplitz DCT matrix, followed by a post-processing stage aimed at reducing noise, improving edge completeness and smoothing of the fused image. The proposed method has produced fused images with lower RMSE, higher SD and better entropy than any of the selected benchmark methods. Future studies should be done to assess the performance of the proposed technique in the absence of contrast adjustment and smoothing by bilateral filter using Gaussian Kernel techniques.

References


