Medical Image Compression using DDCT, Ripplet transform and SPIHT

Waseem Abbas, Numan Khurshid, M. Haroon Shakeel

Abstract — Medical images contain very significant information so very high quality images are used in medical image domain. These high quality images demand very high cost and bandwidth for storage and transmission. Therefore, image compression is very important in medical image processing. Main objective of the paper is to achieve high compression ratio and preserve 2D discontinuities/singularities as well as provide better representation of texture. In this paper, a new image compression technique is proposed. We use Directional Discrete Cosine Transform (DDCT) followed by Ripplet Transform and Set Partitioning in Hierarchical Trees (SPIHT) coding. DDCT in combination with Ripplet provides better texture information and preserves 2D discontinuities/singularities. SPIHT provides better compression ratio. Experimental results show that proposed approach outperforms previous techniques in achievement of high compression ratio and texture representation.

Keywords — Medical Image Compression, Directional Discrete Cosine Transform (DDCT), Ripplet Transform, Set Partitioning in Hierarchical Trees (SPIHT)

1 INTRODUCTION

It is necessary to represent images in better quality for the processing of images in applications like image enhancement and fusion [1-6]. These applications assist innumerable fields which cover medical imaging. Image compression is an important field in digital image processing for reducing the storage cost and bandwidth requirements for transmitting an image [12]. Apart from preserving vital information, the high compression ratio is a major concern in medical image compression [13].

In the clinical tool, medical image processing has great importance because of express visualization and quantifying properties. Visual information gained from the medical images reduces the chances of the wrong diagnosis by enhancing the image quality of medical images like CT, X-Ray and MRI pictures. The visual quality of MRI, X-Rays, and tablets can be enhanced to extract valuable information.

A number of cells and its size in an image, morphology, and shape analysis are the targeted applications in image processing and image enhancement techniques are used for the aforementioned applications. For analysis purpose, image feature can be analyzed using transformation; a technique used to represent an image in its coefficients. For transformation, Fourier transform is used which transforms an image into its frequency components but the edges of the image cannot be represented efficiently by using Fourier transform. Singularities are the artifacts encountered in image processing on edges and bending lines. These singularities can be resolved using wavelet transformation where the image is decomposed into wavelet of frequency components. Limitation of wavelet transform occurs in 2D singularities and discontinuities. To resolve 2D contour and discontinuities, many other techniques have been proposed i.e. Curvelet, Ripplet, Contourlet, and Ridgelet transform. These transformation techniques have the advantage of the better transformation of the 2D signal with curves and bending lines discontinuities. Radon transform serves as the base of Ridgelet transform which makes it possible to resolve 1D discontinuities either in the vertical or horizontal direction [7]. Hence Ridgelet transform can solve singularity.
in only one direction and is not applicable to resolve singularities of 2D signals.

Wavelet transform decomposes the signal into its wavelet bases where the coarser areas of the image represented efficiently in transform domain but the areas with high-frequency components and discontinuities (edges and curves) affect all the frequency wavelets because of poor representation of discontinuities in high 2D singularities. Curvelet transform represents 2D functions along with smooth curve discontinuities. However, the discretization here is challenging and hence the algorithm is highly complicated [19]. Contourlet transform is suitable for image compression [20] and is less computationally intensive but it has less directional features which result in artifacts in compressed images. Ripplet transform was proposed to overcome these singularities and representation of images in different domains and different directions. Literature suggests that Ripplet transform can handle singularities with smooth curves but it is not good for the textures presented in images. Combined transformation using DCT and Ripplet can give better results in coarser areas and boundary (curves) areas of the image respectively.

To exploit the properties of Ripplet transform in terms of dimensionality, Set Partitioning in Hierarchal Trees (SPIHT) is used. SPIHT has inherent properties like self-adoptive and idempotent. Further, SPIHT supports image in multi-dimensions, modularity and multiresolution coding for the images that have bit depth larger than 8 bits. Combining DCT, Ripplet transforms and SPIHT can result in better compression. Our focus is to achieve high medical image compression as well as better representation of textures which is not achieved by using Ripplet alone. In this paper, we propose a technique of using DDCT (Directional Discrete Cosine Transform), Ripplet Transform and SPIHT coding. DDCT in combination with Ripplet can give a better representation of textures whereas SPIHT is used for compressing the resultant image.

2 RELATED WORK

With the development in technology and processing speed of compilers, many researchers have proposed algorithms for image compression and have compared the results of their proposed techniques with state of the art techniques. Venugopal et al. proposed a Huffman coding based Ripplet transform for compression of color medical images. The Ripplet transform breaks the inherent limitations of the wavelet transform. In order to provide high quality compressed images, it represents the image in different scales and directions. In their method, Wavelet transform is applied to the three bands (R, G, B) of input medical image separately which is decomposed into the set of subbands which are partially reconstructed to Ripplet subbands. Next, the low-frequency subbands are directly encoded using Huffman encoding algorithm whereas the high-frequency ones are dissected into small portions by multiplying it with the smooth window function that produces a smooth dissection of the function into squares as shown in Fig. 1. Further renormalization is applied to the squares.

Fig. 1 Decomposed image in different frequency levels

Each pixel in the renormalized square is analyzed in Ripplet domain, which is capable of capturing singularities along the curves. In the Ripplet system, the effective region is analyzed and the resulting discrete Ripplet coefficients are further encoded using a Huffman encoder. The compressed image is obtained and the compression ratio is calculated. After it, Huffman decoding is performed and inverse Ripplet transform is applied over the decoded Ripplet coefficients that cause fusing of the high and low-frequency coefficients together. These fused coefficients are further correlated by performing inverse wavelet transform. Finally, the reconstructed image is obtained whose performance metrics are compared with the original image and with Ripplet transformed image. They have then applied this
method on color images and have observed that PSNR (Peak Signal to Noise Ratio) is optimum with reduced MSE (Mean Square Error), the compression ratio is better and bits per pixel is less [8]. However, the proposed approach lacks the focus towards reducing the latency in transmitting the images.

Ripplet transform provides a hierarchical representation of images. It can effectively approximate images from coarse granularity to a fine granularity. Ripplets can be pointed to arbitrary directions as well. Ripplets are well localized in both spatial and frequency domains and allow scaling with an arbitrary degree. The degree can take any real value. Anisotropy can be achieved through flexible scaling and arbitrary support range. Ripplet transform can provide a more efficient representation of images with singularities along smooth curves. Ripplets have the capability of representing the shape of an object, but they are not good at representing textures. It is promising to combine Ripplet and other transforms such as DCT to represent the entire image, which contains object boundaries and textures [9].

As far as textures are concerned Ripplet is weak to represent textures. So to represent the whole image it is crucial to combine Ripplet with another transform such as DCT which is capable to represent textures [11].

S. Juliet et. al. in their approach used a Wavelet transform to find high and low-frequency components. The high-frequency components are fed into the Ripplet transform (the process described earlier). The Ripplet coefficients and the coefficients in the coarsest subband are further coded using SPIHT algorithm which exploits the dependencies between the location and value of the coefficients across subbands [12]. So, in contrast to [8], they have used SPIHT coding instead of Huffman coding technique. In [16] A. Dogra et al. has proposed six various experiments for medical image compression, using Ripplet transform, directional DCT and their several combinations. They have claimed that the use of DDCT first, followed by Ripplet is best suitable for representation of the shape of the object as well as texture details. However, this technique does not use any further compression method such as Huffman or SPIHT used in the previously mentioned approaches.

3 PROPOSED MODEL

From the literature review, we got to know that wavelet and Fourier transform are unable to resolve edge and contour discontinuities in 2D signal and Ripplet can be used to resolve it. Ripplet properties are multiresolution, good-localization, high directionality, scaling and support in arbitrary degrees, anisotropy and fast decay of coefficients. Ripplet provides efficient representation of image along smooth curves and provides better representation of edges in images.

As far as textures are concerned Ripplet is weak to represent textures. So to represent the whole image it is crucial to combine Ripplet with another transform such as DCT which is capable to represent textures [11].

S. Juliet et. al. in their approach used a Wavelet transform to find high and low-frequency components. The high-frequency components are fed into the Ripplet transform (the process described earlier). The Ripplet coefficients and the coefficients in the coarsest subband are further coded using SPIHT algorithm which exploits the dependencies between the location and value of the coefficients across subbands [12]. So, in contrast to [8], they have used SPIHT coding instead of Huffman coding technique. In [16] A. Dogra et al. has proposed six various experiments for medical image compression, using Ripplet transform, directional DCT and their several combinations. They have claimed that the use of DDCT first, followed by Ripplet is best suitable for representation of the shape of the object as well as texture details. However, this technique does not use any further compression method such as Huffman or SPIHT used in the previously mentioned approaches.

Fig. 2 General block diagram of proposed technique

Moreover, Conventional discrete cosine transform performs well on horizontal and vertical edges which causes some defects in other directions like diagonal directions whereas Directional Discrete Cosine transform operates on edges in different directions within a specified block size [16].

We are to take forward the approach of [16] by using directional DCT with Ripplet transform and modify it in a way by applying SPIHT coding after the Ripplet transform as done in [12]. As Ripplet is involved, 2D singularities of the image can be resolved well. DDCT will provide the texture information. Further, SPIHT will deliver a good level of compression. State of the art techniques are mentioned in related work and considering these techniques, our proposed model for medical image compression is shown in Fig. 2. Hence, we are introducing a novel approach in sense
of combined properties of Cosine transform, Ripplet transforms and SPIHT for medical image compression.

DDCT is applied on the 8x8 blocks of the input image which results in high frequency and low-frequency subbands. At next step, low-frequency components are submitted directly to SPIHT encoding and Ripplet transform is applied on high-frequency components.

Ripplet Transform is applied in three steps including smooth partitioning, renormalization, and analysis in Ripplet domain. After application of Ripplet transformation, we get Ripplet coefficients for high-frequency components. Then low-frequency components and resultant Ripplet coefficients are passed to the SPIHT algorithm as input for further coding. This algorithm has two stages: the sorting stage and the refinement stage. SPIHT algorithm returns a compressed image in the result. The compressed image is decoded and then inverse Ripplet transform is applied over the decoded Ripplet coefficients that cause the fusing of high and low-frequency coefficients together which are further correlated by performing inverse DDCT. Finally, the reconstructed image is obtained which can be compared with the original image for performance metrics. Detail of each block is given below:

### 3.1 Input Image

Standard datasets are available for medical image compression as benchmark images. The medical image \( f(x, y) \) of size 256 x 256 is grayscale image and used as input to the system on which further processing is done.

### 3.2 DDCT (Directional Discrete Cosine Transform)

Discrete cosine transform is the technique used to transform data into frequency domain where only the cosine components exist so that imagery components not present dislike Fourier transform. Compared with the conventional DCT, the directional DCT framework is able to provide a better coding performance for image blocks that contain directional edges. The input image is divided into a block of size 8x8 and DDCT (Directional Discrete Cosine Transform) is applied which operates on edges in different directions within a specified block size.

In H.264, there are in total eight directional prediction modes (the dc mode—Mode 2—is not counted) for blocks of size 4 x 4. Among these modes, one is the vertical prediction (Mode 0), one is the horizontal prediction (Mode 1), and the remaining are named as diagonal down-left (Mode 3), diagonal down-right (Mode 4), vertical-right (Mode 5), horizontal-down (Mode 6), vertical-left (Mode 7), and horizontal-up (Mode 8), respectively. Mode 4 can be obtained by flipping Mode 3 either horizontally or vertically; Mode 6 can be obtained by transposing Mode 5, and Mode 7/8 can be obtained by flipping Mode 5/6 either horizontally or vertically [10] which can be analyzed from Fig. 3.

Fig. 3. Modes of DDCT (Directional Discrete Cosine Transform)

To make our results general enough, we will use a truly directional DCT for the diagonal down-left mode and can extend it to other modes depending on the time available.

In directional DCT diagonal down-left mode (Mode 3), the first 1-D DCT is performed along the diagonal down-left direction, i.e., for each diagonal line with. There are in total diagonal down-left DCTs to be done, whose lengths are All of the coefficients after these DCTs are expressed into a group of column vectors:

\[
A_k = [A_{0,k}, A_{1,k}, \ldots, A_{N_k-1,k}]^T, \quad k = 0, 1, \ldots, 2N - 2
\]

Each column \( A_k \) has a different length \( N_k \) with the DC component placed at top, followed by the first AC component and so on. Next, the second 1-D DCT is applied to each row that can be expressed as \( [A_{uv}]_v = u:2N-2-u \) for \( u = 0, 1, \ldots, N-1 \). The coefficients after the second DCT are pushed horizontally to the left and denoted as \( [\tilde{A}_{uv}]_v = 0:2N-2-u \) for \( u = 0, 1, \ldots, N-1 \). A zigzag scanning is used to convert the 2-D coefficient block into a 1-D sequence so as to facilitate the
run length-based VLC. Next, extension to other directional modes is straightforward.

By using DDCT, we will decompose the image into low and high-frequency subbands. denote the lowest frequency component and denote the high-frequency ones. After performing DDCT, we are extracting the low and high-frequency components based on the technique mentioned in [17, 18]. It states that a common way, which is also used in JPEG encoding, proceeds diagonally from zero-frequency down to the maximum at low frequencies of the image are largely located in the top left corner. A cut-off value is set far above half of the coefficients where the high-frequency part still contains some relevant image information. If this cut-off is set to 0 or below, the only noise of small amplitude is left. Once separated, the low-frequency subbands are directly fed into the SPIHT encoding module. For the high-frequency subbands, we first take the Ripplet transform and then decode it. Ripplet is applied in a hierarchical fashion on the high-frequency components obtained from DDCT.

### 3.3 Ripplet Transform

Ripplet transform is applied to the high-frequency components of the image ($\Delta_B f(x, y)$) which consists of the following three steps [8,12].

#### 3.3.1 Smooth Partitioning

The high-frequency subbands are dissected into small partitions by multiplying with the smooth window function. This produces a smooth dissection of the function into squares. The windowing and the filtering $A_s$ are constructed to ensure that all the steps result in perfect reconstruction. The squares that do not intersect the edge or a Ripplet fragment have no energy and hence can be ignored as shown in Fig. 4.

#### 3.3.2 SPIHT Encoding

Renormalization: Each dyadic square is centered to the unit square in order to have a system of elements at all lengths and all finer widths. It gives an aspect ratio of width $= \text{length}^2$.

Ripplet Domain: Each pixel in the renormalized square is analyzed in Ripplet domain. The major axis denotes the effective length and minor axis denotes the effective width, which is orthogonal to the major axis, represents the effective region, which satisfies the property width $\approx c \times \text{length}$ where $c$ denotes the support of Ripplets and $d$ determines the degree of Ripplet transforms. Through this property, the Ripplets can capture the singularities along the curves.

Ripplet transform actually replicates the wavelet transform and overcomes the weak point of wavelet and represents edges of the image more efficiently. Using Ripplet transform images can be represented in a different direction with different scales.

Using the algorithm of 2D Curvelet transform, Ripplet function can be expressed as:

$$\rho_{a,b,d}(x) = \rho_{a,b}(R_{d}(x - \vec{b}))$$

where $\rho_{a,b}(\cdot)$ is the mother function of Ripplet in frequency domain and $R_{d}$ is rotation matrix which rotates $\theta$ radians. In frequency domain the element function of Ripplet can be represented as:

$$\rho_{d}(r, \omega) = \frac{1}{\sqrt{c}} e^{i \pi r a} W(a, r) V \left( \frac{1}{c}, \frac{\pi a}{\omega}, \omega \right)$$

where are Fourier transform of, $W(r)$ is the radial window on $[0.5,1]$ and $V(w)$ is an angular window on $[-1,1]$. They also fulfill the conditions of admissibility [14,15].

The aforementioned functions are defined as the Ripplet function because they possess shapes like ripples in the frequency domain. In the equation of Ripplet function $\rho$, the element $c$ determines the support of Ripplet and the degree of Ripplet is presented by $d$ (as mentioned before) by changing the parameters as support=1 and degree=2, we can obtain Curvelet transform.

### 3.3.2.2 SPIHT Encoding
The resulting Ripples coefficients and the ones in the coarsest subband are further coded using SPIHT algorithm which exploits dependencies between location and value of coefficients across subbands. This algorithm orders the resulting coefficients according to the significance test and stores the information in three separate sets of lists: list of insignificant sets (LIS), list of insignificant pixels (LIP) and list of significant pixels (LSP).

After the initialization, this algorithm takes two stages for each level of the threshold: sorting stage and refinement stage. During the sorting stage, the pixels in LIP are tested using a significance test and those that become significant are moved to the LSP. The sets are sequentially evaluated following the LIS order and when the set is found to be significant, it is removed from the list and partitioned. The new subsets with more than one elements are added back to the LIS, while the single-coordinate sets are added to the end of LIP or LSP, depending upon whether they are insignificant or significant, respectively. LSP now contains the coordinates of the pixels that are visited in the refinement pass, which outputs the nth most significant bit. The value of n is decreased by 1 and the sorting and refinement stages are repeated. When all the coefficients are processed completely, the compressed image is taken as the output.

The inherent properties of Ripples transform in conjunction with the coding of coefficients using SPIHT algorithm provide efficient representation of edges in images [12].

4 RESULTS AND DISCUSSION

We implemented our approach and tested it on the set of 8 medical images. We used different metrics to measure the effectiveness of our approach that is described below.

4.1 Experimental Setup

We used MATLAB 2015 to implement our proposed approach and the system used was MAC OS, 2.4 GHz Intel Core i5 processor, 8 GB 1600 MHz DDR3 RAM. The dataset consisted of 8 medical images in gray scale as used in [12].

4.2 Performance Evaluations

We have assessed the quality of the compressed images in terms of Peak Signal-to-Noise Ratio (PSNR) vs. File Size and Bitrate, SSIM vs. Bitrate and compression ratio. Result images are shown in figure 5.

**PSNR vs. File Size:** PSNR is one of the most adequate parameters to measure the quality of compression. If the PSNR values are higher, the quality of compression is better and vice versa. It is defined in [12] and can be formulated as,

\[
\text{PSNR} = 10 \log_{10} \left( \frac{255^2}{\text{MSE}} \right)
\]

where MSE represents the means squared error which is given by:

\[
\text{MSE} = \frac{1}{M \times N} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x,y) - F(x,y))^2
\]

where M x N represents the size of the image and F(x,y) is the compressed image. The proposed method is compared to several other techniques for each of the images in the dataset and the PSNR is calculated with respect to the file size which is shown in Fig. 5. It is evident that our proposed image outperforms the other methods as it has the highest PSNR value in case of each of the 8 images.

**Bitrate, SSIM vs. Bitrate and compression ratio:** Result images are shown in figure 5.
TABLE I. Compression ratio achieved for different medical images at 1.2 bpp (bits per pixel)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) TIWI-MRI Ankle</td>
<td>12.22</td>
<td>10.45</td>
<td>12.25</td>
<td>11.35</td>
<td>11.55</td>
<td>12.03</td>
<td>12.25</td>
<td>11.03</td>
<td><strong>12.33</strong></td>
</tr>
<tr>
<td>(b) T2WI-axial-1 vies of brain</td>
<td>12.04</td>
<td>12.44</td>
<td>11.34</td>
<td>10.46</td>
<td>10.67</td>
<td>12.08</td>
<td>12.10</td>
<td>10.12</td>
<td><strong>12.16</strong></td>
</tr>
<tr>
<td>(c) T2WI-axial-2 vies of brain</td>
<td>11.75</td>
<td>11.28</td>
<td>11.73</td>
<td>10.84</td>
<td>10.60</td>
<td>11.78</td>
<td>11.73</td>
<td>10.8</td>
<td>11.72</td>
</tr>
<tr>
<td>(e) T1-weighted MRI lungs</td>
<td>10.23</td>
<td>9.57</td>
<td>9.17</td>
<td>10.59</td>
<td>9.92</td>
<td>12.27</td>
<td>10.42</td>
<td>10.04</td>
<td><strong>12.32</strong></td>
</tr>
<tr>
<td>(f) MRI Abdomen</td>
<td>10.39</td>
<td><strong>12.22</strong></td>
<td>10.4</td>
<td>10.28</td>
<td>11.74</td>
<td>10.89</td>
<td>10.53</td>
<td>9.92</td>
<td>11.96</td>
</tr>
<tr>
<td>(g) CT-axial view of pancreas</td>
<td>8.06</td>
<td>9.73</td>
<td>8.81</td>
<td>9.22</td>
<td>9.69</td>
<td><strong>9.76</strong></td>
<td>8.60</td>
<td>9.18</td>
<td>9.70</td>
</tr>
<tr>
<td>(h) Sagittal stir of head</td>
<td>12.36</td>
<td>12.16</td>
<td>11.78</td>
<td>11.26</td>
<td>10.91</td>
<td>12.42</td>
<td><strong>12.70</strong></td>
<td>10.97</td>
<td>12.67</td>
</tr>
</tbody>
</table>

PSNR vs. Bitrate: Bitrate (bpp) is the ratio of the size of the compressed image in bits to the total number of pixels. The PSNR for each of the image is found for varying bitrate and the different compression methods are used for comparison. As shown in Fig. 6, our proposed approach outperforms other compression techniques for almost all of the bitrates.

SSIM vs. Bitrate: The SSIM is an objective image quality metric used to measure the similarity between two images based on the characteristics of the human visual system. It measures the structural similarity rather than error visibility between two images. SSIM is defined as [12]:

\[
SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}
\]

where \( x \) and \( y \) are spatial patches (windows) and are mean intensity values of \( x \) and \( y \) respectively and are standard deviations. \( C_1 \) and \( C_2 \) are constants. Fig. 7 shows the SSIM values for different bitrates against different techniques and here again, our proposed approach outperforms the remaining methods.

Compression Ratio: Compression ratio is used to enumerate the minimization in image representation size produced by the compression algorithm. It is defined as the ratio of the number of bits in the original image to that of the compressed image [12]. Table I shows the compression ratios achieved for different medical images at 1.2 bpp for different techniques. It can be seen that only the proposed approach has a higher compression ratio for four images whereas none of the other methods has achieved it for more than an image or two.

Fig. 6. PSNR (dB) vs. Bitrate (bpp) using different compression techniques.

Fig. 7. SSIM values at different bit rates using different techniques.
5 Conclusion and Future Work

In this paper, we provide a novel technique by combining the properties of DDCT, Ripplet transform, and SPIHT encoding. Purpose of the technique is to achieve a high compression ratio with preservation of 2D singularities/discontinuities and better representation of texture information. We used DDCT for better texture information, Ripplet transforms to resolve 2D singularities and SPIHT for better compression. Experimental results show that the proposed approach provides better PSNR as compared to different compression techniques in terms of file size and bitrate and provides highest SSIM at different bitrates.

REFERENCES

[16] Ayush Dogra, Sunil Agrawal and Bhawna Goyal: Efficient representation of texture details in medical images by fusion of Ripplet and DDCT
transformed images; *Trop J Pharm Res, September 2016*


[18] https://www.researchgate.net/post/How_can_I_decompose_an_image_into_high_and_low_frequency_components_using_dct_discrete_cosine_transform
