

Block DCT based Image Fusion Techniques

VPS Naidu

MSDF Lab, FMCD
CSIR – National Aerospace Laboratories
Bangalore-17, India, vpsnaidu@gmail.com

Abstract

Block DCT based image fusion techniques have been developed and studied using image fusion quality evaluation metrics. Five image fusion architectures such as feature (FDCT), resizing DCT (RDCT), wavelet structure DCT (WSDCT), subband DCT (SDCT) and Morphological DCT (MpDCT) based on block DCT have been demonstrated. It is observed that WSDCT based image fusion algorithm out performed. The Matlab code will be distributed.

Keywords: Discrete Cosine transform, Image fusion, Block DCT, Resizing DCT, Wavelet structured DCT, Morphological DCT, Sub band DCT

1. Introduction

Multi Sensor Data Fusion (MSDF) is a mathematical tool that increases the dimensionality of the data space to increase the quality of deduced information. MSDF reduces ambiguity and vulnerability of the system to be processed. Off late, MSDF has been a focus of research area. Still there is possible to improve the MSDF techniques, particularly in image fusion. Multi sensor image fusion (MSIF) has become vital signal processing tool in image processing for improving visual interpretation of various applications such as enhanced vision system, robotic vision & navigation, machine vision, medical diagnosis, surveillance, remote sensing and military. The objective of MSIF is to combine (integrate) the visual information from source images (images to be fused) without introducing any artifacts. Image fusion can be achieved by any one of three different processing levels: pixel (signal) level, feature level and decision level fusion. Pixel level image fusion is defining the process of combing the visual information at pixel level of the source images. Feature level image fusion fuses the features that have extracted from the images to be fused. Decision level image fusion represents the fusion of probabilistic decision information obtained by local decision makers operating on features computed from images to be fused. Off late, much research has been done for fusion of multi focus and multispectral images. There are standard image fusion techniques available in literature such those use PCA [4], IHS [1] and Brovery transforms. These fusion algorithms often produce poor visual quality of the fused images. Image fusion with Laplacian pyramid has been discussed in [2]. The energy compaction and multiresolution properties of wavelet transform for image fusion were exploited and demonstrated [3-7]. Multiresolution DCT based image has been proposed and demonstrated for image fusion application in [7,8]. Discrete cosine harmonic wavelets transform base multifocus and multispectral image fusion has been discussed in [9]. One of the prerequisite for multi sensor image fusion is that multi sensor images have to be aligned (registered) on a pixel by pixel basis. Image

registration methods (techniques) have been presented in open literature. In this paper, it is presumed that the images to be fused are already registered.

In this paper, five different fusion algorithms based on block DCT are presented and evaluated. Section 2 describes the discrete Cosine transforms, fusion algorithms and rules are given in section 3. Fusion quality evaluation metrics are given in section 4. Results are presented to demonstrate the performance of the proposed image fusion algorithms in section 5. Finally, the conclusions are presented in section 6.

2. Discrete Cosine Transform

Discrete cosine transform (DCT) is a very important transform in image processing and it is widely accepted by researchers. Large DCT coefficients are concentrated in the low frequency region; hence, it is known to have excellent energy compactness properties and edges may contribute high frequency coefficients. The signal energy due to smooth regions is compacted mostly into DC coefficients; hence edges in the spatial domain can only contribute energy to a small number of AC coefficients. The 2D discrete cosine transform $Z(u, v)$ of an image or 2D signal $z(x, y)$ of size $M \times N$ is defined as [10]:

$$Z(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} z(x, y) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right), \quad \begin{matrix} 0 \leq u \leq M-1 \\ 0 \leq v \leq N-1 \end{matrix} \quad (1)$$

$$\text{Where } \alpha(u) = \begin{cases} \frac{1}{\sqrt{M}}, & u = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq u \leq M-1 \end{cases} \quad \text{and} \quad \alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}, & v = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq v \leq N-1 \end{cases}$$

u & v are discrete frequency variables (x, y) pixel index

Similarly, the 2D inverse discrete cosine transform is defined as [10]:

$$z(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v)Z(u, v) \cos\left(\frac{\pi(2x+1)u}{2M}\right) \cos\left(\frac{\pi(2y+1)v}{2N}\right), \quad \begin{matrix} 0 \leq x \leq M-1 \\ 0 \leq y \leq N-1 \end{matrix} \quad (2)$$

3. Image Fusion Techniques

3.1 Feature DCT (FDCT)

The images ($I_i, i=1,2$) to be fused are divided into small blocks $\{b_{ij}(x, y), 0 \leq x, y \leq N-1\}$ of size $N \times N$ and then transformed into DCT domain $\{B_{ij}(u, v), 0 \leq u, v \leq N-1\}$, where subscripts i and j indicates i^{th} image and j^{th} block respectively. The framework of FDCT based image fusion algorithm is shown in Fig.1. The following fusion algorithms are developed using different features.

3.1.1 Average DCT coefficients (FDCTav)

3.1.2 The fused j^{th} block image (I_{fj}) is computed using eq. 2 and it is repeated for all blocks to get the final fused image (I_f).

$$I_{fj} = \frac{IDCT(B_{1j}) + IDCT(B_{2j})}{2} \quad (3)$$

3.1.3 Energy Probability (FDCTep)

The energy probability is computed for i^{th} image and j^{th} block as:

$$ep_{ij} = \frac{|DCT(B_{ij})|^2}{\sum (|DCT(B_{ij})|^2)} \quad (4a)$$

The fused j^{th} block image (I_{fj}) is computed using eq. 4 and it is repeated for all blocks to get the final fused image (I_f).

$$I_{fj} = \begin{cases} IDCT(B_{1j}), & ep_{1j} \geq ep_{2j} \\ IDCT(B_{2j}), & ep_{1j} < ep_{2j} \end{cases} \quad (4b)$$

3.1.4 Variance (FDCTva)

Variance is computed for each block using:

$$\sigma_{ij}^2 = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \frac{B_{ij}^2(u,v)}{N^2} - B^2(0,0) \quad (5a)$$

The fused j^{th} block image (I_{fj}) is computed using eq. 6 and it is repeated for all blocks to get the final fused image (I_f).

$$I_{fj} = \begin{cases} IDCT(B_{1j}), & \sigma_{1j}^2 \geq \sigma_{2j}^2 \\ IDCT(B_{2j}), & < \sigma_{2j}^2 \end{cases} \quad (5b)$$

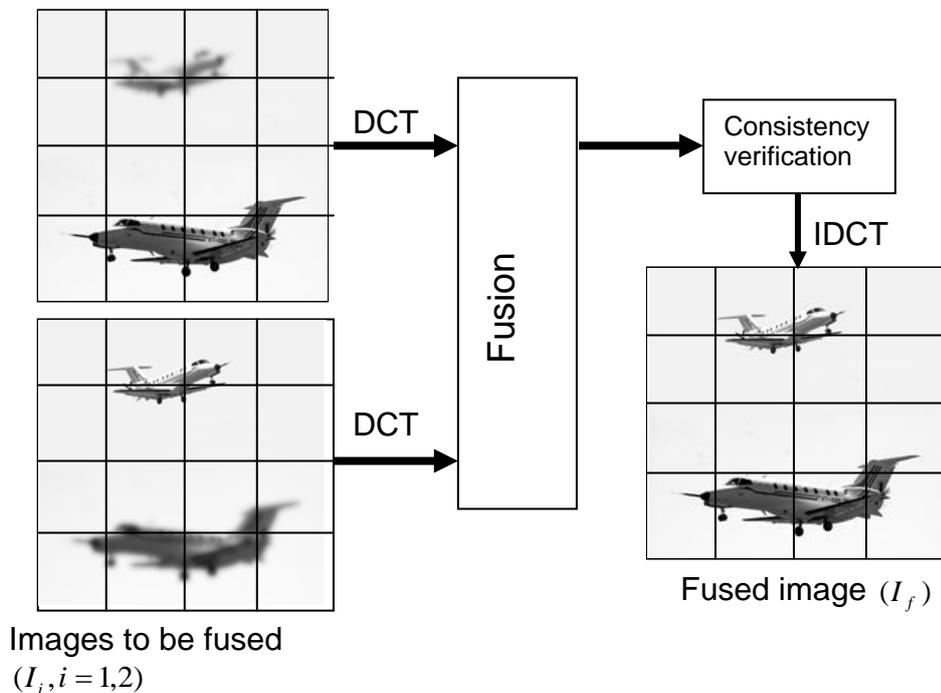


Fig-1 Framework of FDCT based image fusion algorithm

Consistency verification procedure is applied to reduce the erroneous selection blocks from images to be fused. Majority filter [11] is used for consistency verification. If the center block comes from I_1 image, while the majority of the surrounding blocks come from image I_2 , then the center block is replaced by the block from I_2 .

3.2 Resizing DCT (RDCT)

Image resize can be done in either spatial domain or DCT domain. Image resizing in spatial domain is computationally complex than transform domain. In DCT domain, high frequency (HF) coefficients are truncated for down sampling and assuming HF coefficients to be zero for up sampling. This approach has significant drawbacks such as blocking artifacts and ringing effects in the resized image. Image resizing method presented in ref [12] is followed in this paper for image fusion.

3.2.1 Down Sampling: To down sample the image by a factor of two, the following procedure is carried out as shown in Fig-2a. The image is divided by non-overlap blocks of size 8x8. Each block is then transformed into DCT domain. The 4x4 low frequency (LF) coefficients out of each 8x8 DCT block $\{I_i(m,n), 0 \leq m,n \leq 3, i = 1,2,\dots,4\}$ as shown in Fig-2a. A 4x4 IDCT is applied on the LF coefficients to get down sampling. In this way, four consecutive 8x8 blocks become four consecutive 4x4 blocks in spatial domain. This image in spatial domain can be down sampled by repeating the same procedure. This procedure is called reduction function.

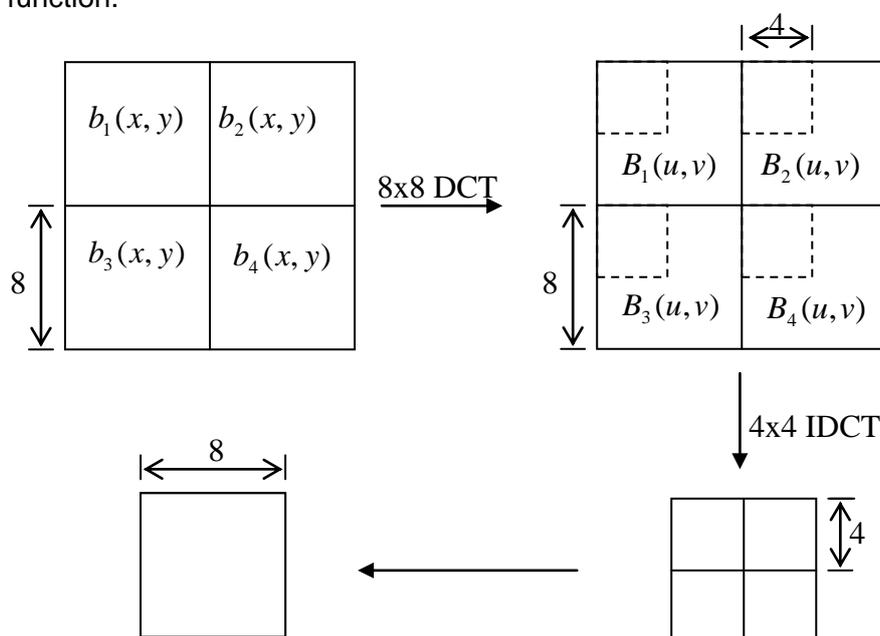


Fig-2a Down sampling process

3.2.2 Up Sampling: Up sampling the image by a factor of two can be done by reversing the above procedure described in section 3.2.1. The image to be up sampling a factor of two is divided into 4x4 blocks. Four consecutive 4x4 blocks are transformed into DCT domain as shown in Fig-2b. These are treated as LF coefficients and used as the LF components in the 8x8 blocks and the remaining HF coefficients are assumed to be zero. Then consecutive 8x8 blocks in DCT domain

are converted into spatial domain (up sampling) by applying 8x8 IDCT. This procedure is called expand function.

3.2.3 Laplacian Pyramid: The procedure for Laplacian pyramid construction and reconstruction is illustrated in Fig-3 [13]. The image at the 0th level z_0 of size $M \times N$ is reduced (down sampling) to obtain next level z_1 of size $0.5M \times 0.5N$ where both spatial density and resolution are reduced. Similarly, z_2 is the reduced version of z_1 and so on. The level to level image reduction is performed using the function reduce R .

Reduction Function R :
$$z_k = R(z_{k-1}) \tag{6}$$

The reverse of function reduce is expand function E (Up sampling). Its effect is to expand the image of size $M \times N$ to image of size $2M \times 2N$ by.

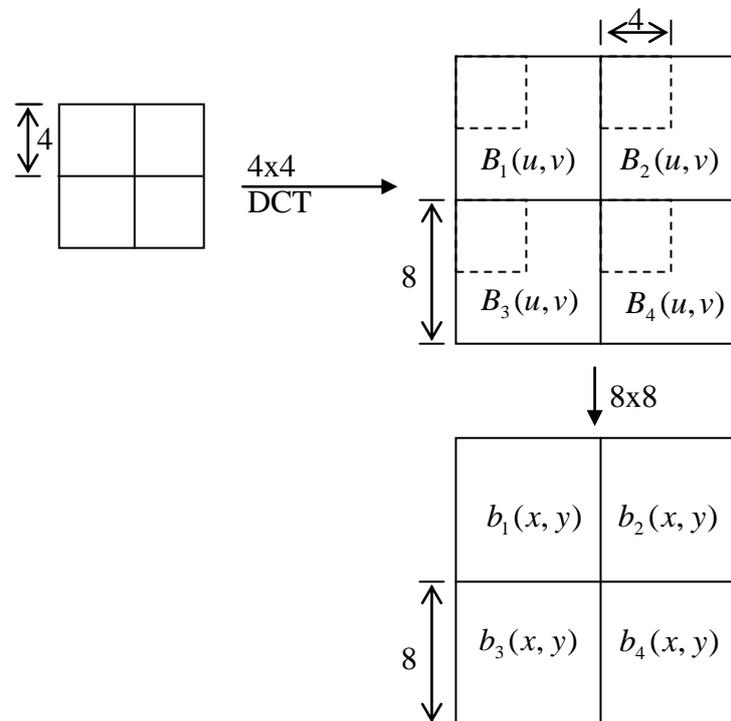


Fig-2b Up sampling process

Expand Function E :
$$\hat{z}_k = E(\hat{z}_{k+1}) \tag{7}$$

Construction of pyramid is done using (Fig-3):
$$z_{k+1} = R(z_k) \tag{8a}$$

$$l_k = z_k - E(z_{k+1}) \tag{8b}$$

where l_0, l_1, \dots, l_{k-1} are Laplacian image pyramids that contain band pass filtering images and keeping these records to utilize on reconstruction process and z_k is the coarser level image. The k levels of image pyramid are

represented as $P_k \rightarrow \{z_k, l_0, l_1, \dots, l_{k-1}\}$. At coarser level $\hat{z}_k = z_k$, since there is no more decomposition beyond this level.

$$\hat{z}_{k-1} = l_{k-1} + E(\hat{z}_k) \quad (9)$$

3.2.4 Fusion: Let, there are two images (I_1 & I_2) to be fused. Pyramid construction is done for each image and keeping the error records. Denote the constructed k levels of Laplacian image pyramid for 1st image is ${}^1P_k \rightarrow \{z_k, {}^1l_0, {}^1l_1, \dots, {}^1l_{k-1}\}$ and similarly for of 2nd image is ${}^2P_k \rightarrow \{z_k, {}^2l_0, {}^2l_1, \dots, {}^2l_{k-1}\}$.

Then the image fusion rule is as follows:

$$\text{At } k^{\text{th}} \text{ level, } {}^f z_k = \frac{{}^1\hat{z}_k + {}^2\hat{z}_k}{2} \quad (10a)$$

$$\text{For } k-1 \text{ to } 0 \text{ levels } {}^f z_{k-1} = {}^f l_{k-1} + E({}^f z_k) \quad (10b)$$

Where ${}^f l_{k-1} = \begin{cases} {}^1l_{k-1} & |{}^1l_{k-1}| \geq |{}^2l_{k-1}| \\ {}^2l_{k-1} & |{}^1l_{k-1}| < |{}^2l_{k-1}| \end{cases}$ and the magnitude comparison is done on corresponding pixels.

The pyramid $I_f = {}^f z_0$ is the fused image.

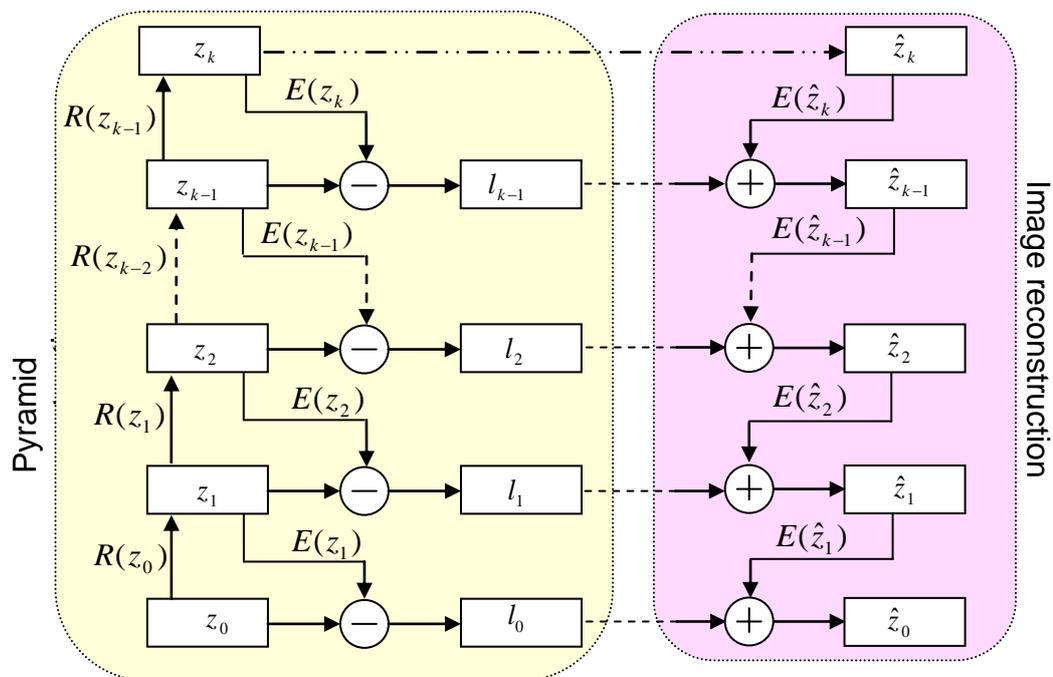


Fig-3 Laplacian pyramid construction

3.3 Wavelet Structured DCT (WSDCT)

It is motivated by the recognition that a DCT on $N \times N$ block of an image can be viewed as wavelet decomposition. Fig-4 illustrates the relationship between wavelet and DCT coefficients structure [14]. Let us consider a DCT that is applied on 2×2 non-overlapping block of an image. Each transformed block contains four DCT coefficients and the coefficients are labeled as a (low frequency) to d (high frequency) as shown in Fig-4a.

As shown in Fig-4b, grouping of the lowest frequency coefficients of each DCT transformed block forms a LL band of Wavelet transform. Similarly grouping other coefficients of each DCT block forms LH, HL and HH bands of Wavelet transform. The next level of decomposition can be achieved by repeating the same procedure on LL band.

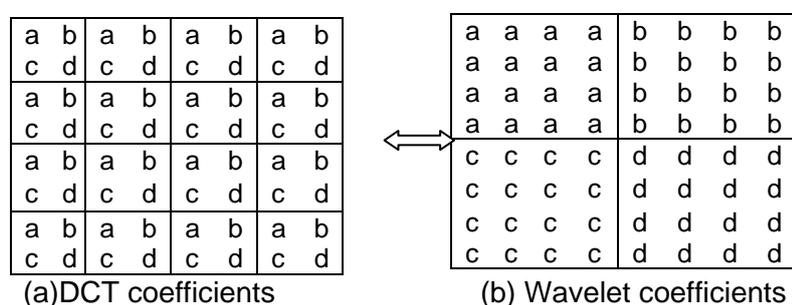


Fig-4 Relationship between DCT and wavelet coefficients

The images to be fused I_1 and I_2 are decomposed into D ($d = 1, 2, \dots, D$) levels using WSDCT. The resulted decomposed images from I_1 are $I_1 \rightarrow \{^1LL_D, \{^1LH_d, ^1HH_d, ^1HL_d\}_{d=1,2,\dots,D}\}$ and from I_2 are $I_2 \rightarrow \{^2LL_D, \{^2LH_d, ^2HH_d, ^2HL_d\}_{d=1,2,\dots,D}\}$. At each decomposition level ($d = 1, 2, \dots, D$), the fusion rule will select the larger absolute value of the two WSDCT detailed coefficients, since the detailed coefficients are corresponds to sharper brightness changes in the images such as edges and object boundaries etc. These coefficients are fluctuating around zero. At the coarsest level ($d = D$), the fusion rule take average of the WSDCT approximation coefficients since the approximation coefficients at coarser level are the smoothed and subsampled version of the original image. The fused image I_f can be obtained using :

$$I_f \leftarrow \{^fLL_D, \{^fLH_d, ^fHH_d, ^fHL_d\}_{d=1,2,\dots,D}\} \quad (11)$$

$$\text{Where } ^fLL_D = \frac{^1LL_D + ^2LL_D}{2}$$

$$^fLH_d = \begin{cases} ^1LH_d & |^1LH_d| \geq |^2LH_d| \\ ^2LH_d & |^1LH_d| < |^2LH_d| \end{cases}, d = 1, 2, \dots, D$$

$${}^f HH_d = \begin{cases} {}^1 HH_d & |{}^1 HH_d| \geq |{}^2 HH_d| \\ {}^2 HH_d & |{}^1 HH_d| < |{}^2 HH_d| \end{cases}, d = 1, 2, \dots, D$$

$${}^f HL_d = \begin{cases} {}^1 HL_d & |{}^1 HL_d| \geq |{}^2 HL_d| \\ {}^2 HL_d & |{}^1 HL_d| < |{}^2 HL_d| \end{cases}, d = 1, 2, \dots, D$$

3.4 Subband DCT (SDCT)

An N-point DCT transform can be deduced as an N-band filter bank whose filters are the transform's basis functions [15]. The DCT coefficients can be rearranged into a subband structure (based on their frequencies). Subband ij contains all coefficients at the position ij from every DCT block transform. These coefficients characterize the same frequency component of the whole image. For example, subband 00 collects the coefficients at position 00 from all DCT block transforms; subband 01 collects the coefficients at position 01 from all DCT block transforms and so on as shown in Fig-5. Fig-5b illustrates the subbands formation of the image shown in Fig-5a.

A ₀₀	A ₀₁	A ₀₂	A ₀₃	B ₀₀	B ₀₁	B ₀₂	B ₀₃
A ₁₀	A ₁₁	A ₁₂	A ₁₃	B ₁₀	B ₁₁	B ₁₂	B ₁₃
A ₂₀	A ₂₁	A ₂₂	A ₂₃	B ₂₀	B ₂₁	B ₂₂	B ₂₃
A ₃₀	A ₃₁	A ₃₂	A ₃₃	B ₃₀	B ₃₁	B ₃₂	B ₃₃
C ₀₀	C ₀₁	C ₀₂	C ₀₃	D ₀₀	D ₀₁	D ₀₂	D ₀₃
C ₁₀	C ₁₁	C ₁₂	C ₁₃	D ₁₀	D ₁₁	D ₁₂	D ₁₃
C ₂₀	C ₂₁	C ₂₂	C ₂₃	D ₂₀	D ₂₁	D ₂₂	D ₂₃
C ₃₀	C ₃₁	C ₃₂	C ₃₃	D ₃₀	D ₃₁	D ₃₂	D ₃₃

A ₀₀	B ₀₀	A ₀₁	B ₀₁	A ₀₂	B ₀₂	A ₀₃	B ₀₃
C ₀₀	D ₀₀	C ₀₁	D ₀₁	C ₀₂	D ₀₂	C ₀₃	D ₀₃
A ₁₀	B ₁₀	A ₁₁	B ₁₁	B ₁₂	B ₁₂	A ₁₃	B ₁₃
C ₁₀	A ₁₀	C ₁₁	D ₁₁	C ₁₂	D ₁₂	C ₁₃	D ₁₃
A ₂₀	B ₂₀	A ₂₁	B ₂₁	A ₂₂	B ₂₂	A ₂₃	B ₂₃
C ₂₀	C ₂₀	C ₂₁	D ₂₁	C ₂₂	D ₂₂	C ₂₃	D ₂₃
A ₃₀	B ₃₀	A ₃₁	B ₃₁	A ₃₂	B ₃₂	A ₃₃	B ₃₃
C ₃₀	D ₃₀	C ₃₁	D ₃₁	C ₃₂	D ₃₂	C ₃₃	D ₃₃

(a) (b)

Fig-5a 4x4 DCT block transformed image and Fig-5b Reordered DCT block coefficients into subband structure

Let the subband decomposition of images to be fused (I_1 & I_2) are ${}^1 S_{ij}$ and ${}^2 S_{ij}$. Then the fusion process is:

$${}^f S_{11} = \frac{{}^1 S_{11} + {}^2 S_{11}}{2} \tag{12a}$$

$${}^f S_{ij} = \begin{cases} {}^1 S_{ij} & |{}^1 S_{ij}| \geq |{}^2 S_{ij}| \\ {}^2 S_{ij} & |{}^1 S_{ij}| < |{}^2 S_{ij}| \end{cases}, i, j \neq 0 \tag{12b}$$

Fused image (I_f) is obtained by reversing the process on ${}^f S_{ij}$.

3.5 Morphological DCT (MpDCT)

Image is partitioned into $m \times m$ sized blocks, where $m = 2^K$, $K > 0$. DCT is computed on each block and the coefficients are taken as a K scale tree of coefficients with $3K + 1$ sub-band decomposition as shown in Fig-6a for $K = 2$ [16]. The same sub bands for all blocks are grouped and placed onto their corresponding position as shown in Fig-6b where G0 denotes group of sub

band 0 and G1 denotes group of sub band 1 and so on. It is observed from Fig-6b that we have better energy clustering in the transformed domain. Coefficients within sub-bands tend to be better clustered. The characteristics of clustered coefficients are used in image fusion.

0	1	4	5
		6	7
2	3		
8	9	12	13
10	11	14	15

Fig-6a DCT coefficients of 4x4 image block as two scale tree with seven sub band decomposition

G0	G1	G4
G2	G3	
G5		G6

Fig-6b Reorganization of DCT coefficients of 4x4 blocks into a single DCT clustering entity

MpDCT construction is done for each image to be fused. Denote the constructed k levels of MpDCT for $z_1(x, y)$ image is $I_1 \rightarrow \{G_{k1}\}$, $k = 0, 1, 2, \dots, 3K$ and similarly for of $z_2(x, y)$ image is $I_2 \rightarrow \{G_{k2}\}$, $k = 0, 1, 2, \dots, 3K$.

Then the image fusion rule is as follows:

$$G_{0f} = 0.5(G_{01} + G_{02}) \quad (13a)$$

$$G_{kf} = \begin{cases} G_{k1} & |G_{k1}| \geq |G_{k2}| \\ G_{k2} & |G_{k1}| < |G_{k2}| \end{cases}, \quad k = 1, 2, \dots, 3K \quad (13b)$$

The fused image can get by doing the reverse procedure explained with Fig-7.

$$I_f \leftarrow \{G_{kf}\}, \quad k = 0, 1, 2, \dots, 3K \quad (13c)$$

4. Fusion Quality Evaluation Metrics

Fusion quality can be evaluated visually. Human judgment decides fusion quality. Human object evaluators give grade to corresponding image (fused) and average the grade. This type of evaluation has some drawbacks such as the grade is not accurate and it depends on the human observer's ability. To

avoid these drawbacks, quantitative measures are used for accurate and meaningful assessment of fused images.

4.1 Root Mean Square Error (RMSE)

This metric is computed as the root mean square error of the corresponding pixels in the reference image I_r and the fused image I_f . This metric will be nearly zero when the reference and fused images are similar. This will increase when the dissimilarity increases.

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_r(i, j) - I_f(i, j))^2} \quad (14)$$

where M : number of rows
 N : number of columns
 (i, j) : pixel index
 I : given image
 $I(i, j)$: gray value at pixel (i, j)

4.2 Peak Signal to Noise Ratio (PSNR)

PSNR will be high when the fused and reference images are alike. Higher value means better fusion. It is computed as:

$$PSNR = 10 \log_{10} \left(\frac{L^2}{RMSE} \right) \quad (15)$$

where L is the number of gray levels in the image

4.3 Average Pixel Intensity (API)

It measures the index of contrast and it is computed as:

$$\mu = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N I_f(i, j) \quad (16)$$

4.4 Standard Deviation (SD)

Important index to weight the information of image, it reflects the deviation degree of values relative to mean of image. The greater the SD, more dispersive the gray grade distribution is. Standard deviation would be more efficient in the absence of noise [30]. An image with high contrast would have a high standard deviation. It is calculated using the formula

$$SD = \sqrt{\frac{1}{MN} \sum_{i,j=1}^{M,N} (I_f(i, j) - \mu)^2} \quad (17)$$

4.5 Contrast Visibility (CV)

Contrast visibility is computed as:

$$\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \frac{|I_f(i, j) - \mu|}{\mu} \quad (18)$$

4.6 Spatial Frequency (SF) [4,11]

SF indicates the overall activity level in the fused image. The spatial frequency for the fused image I_f of dimension $M \times N$ is defined as follows:

$$\text{Row frequency: } RF = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=1}^{N-1} [I_f(i, j) - I_f(i, j-1)]^2} \quad (19a)$$

$$\text{Column frequency: } CF = \sqrt{\frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=1}^{M-1} [I_f(i, j) - I_f(i-1, j)]^2} \quad (19b)$$

$$\text{Spatial frequency: } SF = \sqrt{RF^2 + CF^2} \quad (19c)$$

4.7 Entropy

Entropy is a mono modal measure of information content of an image.

$$H = -\sum_{i=0}^L p_i \log_2 p_i \quad (20)$$

where p_i is the estimated probability density function (normalized pixel intensity histogram).

4.8 Difference Entropy (ΔH)

It reflects the difference between the average amounts of information they contained. Its definition is,

$$\Delta H = |H(I_r) - H(I_f)| \quad (21)$$

Where $H(I_r)$ and $H(I_f)$ are the entropy of reference and fused images

5. Results and Discussion

The fusion algorithms developed in section 3 are evaluated using the images shown in Fig-8. The ground truth image is shown in Fig-8a and the images to be fused are shown in Fig-8b&c. Both Fig-8b and Fig-8c are complementary to each other. In first image (I_1) upper side aircraft is out of focused and the other aircraft is in focus as shown in Fig-8b and vice versa in image I_2 as shown in Fig-8c. The fused (left side) along with the error (right side) images are shown in Fig-9 to 14. The error image is the difference between reference I_r and fused I_f images. One can observe that the fused image preserves all the useful information from the two source images. The fusion quality evaluation metrics are shown in Table 1 to 6. The best results are shown in bold. In case of 1D DCT based image fusion techniques, the Laplacian pyramid based image fusion gives better results with high levels of decomposition. It could be because of series of quasi-band passed images which are localized in both space and spatial frequencies. In case of 2D DCT based image fusion techniques, DTMDCT provides better results with high levels of decomposition because of its shift invariant property. DTMDCT based image fusion technique provides better fusion results among six fusion

techniques presented this paper. Moreover, DTMDCT is simple and computationally efficient while compared to dual other tree analytical algorithm. Higher level of decomposition used in fusion process gives better performance. Fusion by Laplacian pyramid and DTMDCT with higher level of decomposition showed small PFE and high PSNR that shows the fused image is very similar to the ground truth image. Similarly, SD and SF are high that shows that the fused image having good contrast and retain the overall activity.



Fig-7 Ground truth image (I_r)



(a)



(b)

Fig-8a&b Images to be fused (I_1 & I_2)

Table-1a: Fusion quality evaluation metrics - FDCTav

Block size	RMS E	PSN R	API	SD	CV	EOG	SF	H	dH	Time sec.
2x2	0.043	61.84	0.868	0.203	0.161	0.0016	0.0012	4.6531	0.3184	26.2731
4x4	0.043	61.84	0.868	0.203	0.161	0.0016	0.0012	4.6531	0.3184	6.5949
8x8	0.043	61.84	0.868	0.203	0.161	0.0016	0.0012	4.6531	0.3184	1.7466
16x16	0.043	61.84	0.868	0.203	0.161	0.0016	0.0012	4.6528	0.3180	0.4962
32x32	0.043	61.84	0.868	0.203	0.161	0.0016	0.0012	4.6531	0.3171	0.1759

Table-1b: Fusion quality evaluation metrics - FDCTep

Block size	RMS E	PSN R	API	SD	CV	EOG	SF	H	dH	Time sec.
2x2	0.0698	59.727	0.8752	0.1966	0.1542	0.0044	0.0032	4.5502	0.2154	27.7939
4x4	0.0558	60.700	0.8747	0.2032	0.1579	0.0057	0.0042	4.5721	0.2374	7.2153
8x8	0.0412	62.017	0.8719	0.2108	0.1634	0.0057	0.0043	4.6845	0.3498	1.8984
16x16	0.0265	63.932	0.8698	0.2169	0.1681	0.0056	0.0042	4.7402	0.4054	0.5393
32x32	0.0153	66.317	0.8688	0.2211	0.1713	0.0055	0.0041	4.7157	0.3809	0.1934

Table-1c: Fusion quality evaluation metrics – FDCTva

Block size	RMS E	PSN R	API	SD	CV	EOG	SF	H	dH	Time sec.
2x2	0.0430	61.835	0.8684	0.2034	0.1611	0.0016	0.0012	4.6482	0.3134	19.4279
4x4	0.0430	61.835	0.8684	0.2034	0.1611	0.0016	0.0012	4.6479	0.3132	4.9857
8x8	0.0430	61.835	0.8684	0.2034	0.1611	0.0016	0.0012	4.6472	0.3124	1.3053
16x16	0.0430	61.835	0.8683	0.2034	0.1611	0.0016	0.0012	4.6480	0.3133	0.3705
32x32	0.0430	61.834	0.8683	0.2034	0.1611	0.0016	0.0012	4.6485	0.3137	0.1305



Fig-9a. Fused and error image using FDCTav



Fig-9b Fused and error image using FDCTep



Fig-9c Fused and error image using FDCTva

Table-2a: RMSE with different levels of decomposition using RDCT

Level	Block size				
	2x2	4x4	8x8	16x16	32x32
1	0.0397	0.0404	0.0411	0.0414	0.0415
2	0.0335	0.0333	0.0348	0.0354	0.0357
3	0.0269	0.0237	0.0244	0.0251	0.0252
4	0.0236	0.0168	0.0142	0.0129	0.0116
5	0.0221	0.0145	0.0102	0.0082	0.0061

Table-2b: PSNR with different levels of decomposition using RDCT

Level	Block size				
	2x2	4x4	8x8	16x16	32x32
1	62.1788	62.1050	62.0261	61.9947	61.9792
2	62.9099	62.9383	62.7514	62.6792	62.6414
3	63.8657	64.4129	64.2933	64.1710	64.1471
4	64.4341	65.8999	66.6563	67.0487	67.5239
5	64.7285	66.5518	68.0915	69.0215	70.3161

Table-2c: Computational time (sec.) with different levels of decomposition using RDCT

Level	Block size				
	2x2	4x4	8x8	16x16	32x32
1	72.6120	23.2600	6.0525	1.6596	0.5332
2	91.4707	29.0595	7.5408	2.0695	0.6629
3	95.8999	30.5414	7.9202	2.1677	0.6943
4	96.8721	30.9027	8.0228	2.1992	0.7023
5	97.2741	30.9824	8.0382	2.1975	0.7043

Table-3: Performance metrics with block size of 32x32 and 5 levels of decomposition- RDCT

RMSE	PSNR	API	SD	CV	EOG	SF	H	dH
0.0061	70.3161	0.8683	0.2222	0.1722	0.0055	0.0041	4.8069	0.4721



Fig-10. Laplacian with block size of 32x32 and 5 levels of decomposition

Table-4: Performance metrics using (WSDCT) with different levels of decomposition

levels	RMSE	PSNR	API	SD	CV	EOG	SF	H	dH	Time (sec.)
1	0.0395	62.204	0.8684	0.2057	0.1619	0.0035	0.0026	4.6540	0.3192	26.7
2	0.0322	63.092	0.8684	0.2100	0.1640	0.0050	0.0037	4.6603	0.3256	32.6
3	0.0222	64.706	0.8684	0.2150	0.1670	0.0054	0.0040	4.6487	0.3140	33.9
4	0.0127	67.115	0.8684	0.2191	0.1699	0.0055	0.0041	4.6346	0.2998	34.4
5	0.0077	69.283	0.8684	0.2217	0.1719	0.0055	0.0041	4.6061	0.2714	34.5



Fig-11. Fused and error image using WSDCT with 5 Levels of decomposition

Table-5: Fusion quality evaluation metrics using SDCT

Block size	RMSE	PSNR	API	SD	CV	EOG	SF	H	dH	Time sec.
2x2	0.0397	62.178	0.8684	0.2057	0.1620	0.0035	0.0026	4.6540	0.3192	26.7515
4x4	0.0334	62.925	0.8684	0.2101	0.1644	0.0051	0.0038	4.6540	0.3193	7.4114
8x8	0.0256	64.075	0.8684	0.2154	0.1676	0.0055	0.0041	4.6604	0.3257	2.4357
16x16	0.0192	65.331	0.8684	0.2197	0.1707	0.0055	0.0041	4.6578	0.3230	1.1717
32x32	0.0131	66.996	0.8684	0.2224	0.1724	0.0055	0.0041	4.6364	0.3017	0.8846



Fig-12. Fused and error image using SDCT with 32x32 block size

Table-6: Performance metrics using (MpDCT) with different levels of decomposition

Scale	RMSE	PSNR	API	SD	CV	EOG	SF	H	dH	Time (sec.)
1	0.0397	62.1780	0.8684	0.2057	0.1620	0.0035	0.0026	4.6540	0.3192	26.817
2	0.0334	62.9248	0.8684	0.2101	0.1644	0.0051	0.0038	4.6540	0.3193	8.172
3	0.0256	64.0749	0.8684	0.2154	0.1676	0.0055	0.0041	4.6604	0.3257	2.414
4	0.0192	65.3308	0.8684	0.2197	0.1707	0.0055	0.0041	4.6578	0.3230	0.729
5	0.0324	63.0640	0.8684	0.2238	0.1731	0.0083	0.0062	5.1327	0.7979	0.257
6	0.0673	59.8844	0.8684	0.2293	0.1790	0.0127	0.0095	5.8798	1.5450	0.136
7	0.0941	58.4268	0.8684	0.2342	0.1863	0.0121	0.0094	6.0204	1.6856	0.099



Fig-13 Fused and error image using MpDCT

6. Conclusion

Block DCT based image fusion algorithms have been developed and evaluated using image fusion quality evaluation metrics. Feature DCT (FDCT), Resizing DCT (RDCT), wavelet structure DCT (WSDCT), Subband DCT (SDCT) and Morphological DCTM (MpDCT) were proposed and studied.

WSDCT based image fusion algorithm out performed. Matlab code for these algorithms will be provided.

References

1. Carper, J.W., Lilles, T.M., Kiefer, R.W. "The use of intensity-hue saturation transformations for merging SPOT panchromatic and multi-spectra image data", *Photogr. Eng. Remote Sens.* **56**, 459–467 (1990).
2. Toet, A. "Hierarchical image fusion", *Mach. Vis. Appl.* **3**(1), 1–11(1990)
3. Daubechies, I. "Ten Lectures on Wavelets", CBMS-NSF Regional Conference on Series in Applied Mathematics, 61, 2nd edn. SIAM, Philadelphia (1992).
4. VPS Naidu and J.R. Raol, "Pixel-Level Image Fusion using Wavelets and Principal Component Analysis – A Comparative Analysis" *Defence Science Journal*, Vol.58, No.3, pp.338-352, May 2008.
5. Zhi-guo, J., Dong-bing, H., Jin, C., Xiao-kuan, Z. "A Wavelet based Algorithm for Multi-focus Micro-image Fusion", In: *Proceedings of International Conference on Image and Graphics (ICIG)*, pp. 176–179, Dec (2004).
6. Qu, G., Zhang, D., Yan, P. "Medical image fusion by wavelet transform modulus maxima", *Opt. Express* **9**(4), 184–190 (2001).
7. VPS Naidu, "Discrete cosine transform-based image fusion", *Special Issue on Mobile Intelligent Autonomous System*, *Defence Science Journal*, Vol. 60, No.1, pp.48-54, Jan. 2010.
8. VPS Naidu, "Discrete Cosine Transform based Image Fusion Techniques", *Journal of Communication, Navigation and Signal Processing*, 1(1), pp.35-45, Jan. 2012. (2010).
9. B. K. Shreyamsha Kumar, "Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform", *J. SIViP*, DOI 10.1007/s11760-012-0361-x, 14 July 2012
10. VPS Naidu, "Discrete Cosine Transform-based Image Fusion", *Special Issue on Mobile Intelligent Autonomous System*, *Defence Science Journal*, Vol. 60, No.1, pp.48-54, Jan. 2010.
11. VPS Naidu and J.R. Raol, "Fusion of Out Of Focus Images using Principal Component Analysis and Spatial Frequency", *Journal of Aerospace Sciences and Technologies*, Vol. 60, No. 3, pp.216-225, Aug. 2008.
12. Sang-Jun Park and Jechang Jeong, "Hybrid Image Upsampling Method in the Discrete Cosine Transform Domain", *IEEE Transactions on Consumer Electronics*, Vol. 56, No. 4, pp.2615-2622, November 2010.
13. VPS Naidu and Bindu Elias, "A Novel Image Fusion Technique using DCT based Laplacian Pyramid", *International Journal of Inventive Engineering and Sciences (IJIES)* ISSN: 2319–9598, Vol.1(2), pp.17-24, Jan. 2013.
14. Yeong-An Jeong and Cha-Keon Cheong, "A DCT-based Embedded Image Coder using Wavelet Structure of DCT for very low bit rate video CODEC", *EEE Transactions on Consumer Electronics*, Vol. 44, No. 3, pp.500-507, AUGUST 1998.
15. Y. Wongsawat, K.R. Rao, S. Orintara, "Multichannel SVD-Based Image De-Noising", in *Proc. of IEEE Int. Symp. Circuits and Systems*, 2005, pp.5990-93, Vol. 6, 2005.
16. Debin Zhao, Wen Gao and Y.K. Chan, "Morphological representation of DCT coefficients for image compression", *IEEE Tran. On Circuits and Systems for Video Technology*, 12(9), pp.819-823, Sep. 2002.