



Novel Image Fusion Techniques using DCT

V. P. S. Naidu

MSDF Lab, CSIR-NAL
Bangalore

ABSTRACT

Six different novel image fusion techniques using discrete Cosine transform were presented and studied. It is concluded from the study that DTMDCT (Dual tree multi-resolution discrete Cosine transform) provides better fusion results followed by Laplacian pyramid based image fusion technique. DTMDCT is simple and computationally efficient algorithm. Matlab code for these algorithms has been provided.

Keywords

Image fusion, Multi-resolution analysis, Laplacian pyramid, Distributed multi-resolution DCT and Dual tree multi-resolution DCT

1. INTRODUCTION

Multi Sensor Image Fusion (MSIF) is a process that integrates a set of images to produce a composite (fused) image that is more suitable for the purpose of human visual perception, object detection and target recognition. The fused image should contain more information than any individual images to be fused. MSIF can take place on pixel level, feature level and decision level. Pixel level image fusion is the basic algorithm among them. Off late, many researchers have attempted to develop different pixel level image fusion algorithms. Most algorithms are based on wavelet transform [1,2], pyramid transform [3], statistical signal processing [4], principal component analysis [1,4], fuzzy logic [5] and so on. In this paper, novel image fusion algorithms based on discrete Cosine transform (DCT) are presented. DCT is well proven for image compression and well accepted for many image preprocessing applications.

One of the prerequisite for MSIF is that multi sensor images have to be registered (aligned) on a pixel by pixel basis. Image registration methods have been presented in open literature [6]. It is assumed that the images to be fused are already aligned.

In this paper, six different DCT based novel multi sensor image fusion algorithms are presented and evaluated. First three fusion algorithms use 1D DCT where the image is converted into vector data and rest two fusion algorithms use 2D DCT. 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 algorithms in section 5. Finally, the conclusions are presented in section 6.

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 1D discrete cosine transform $Z(u)$ of an vector or 1D signal $z(x)$ of size M is defined as [7-9]:

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

Similarly, the 1D inverse discrete cosine transform is defined as:

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

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 [6-8]:

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

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

$$\alpha(v) = \begin{cases} \frac{1}{\sqrt{N}}, & v = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq v \leq N-1 \end{cases}$$

u & v are discrete frequency variables

(x, y) pixel index

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

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



2. IMAGE FUSION TECHNIQUES

2.1 Using 1D DCT

The image $z(x, y)$ of size $M \times N$ is divided into rows and concatenates these rows to form a 1D vector data $z(x)$ whose size would be MN as shown in Fig.1a. The Matlab code for this is:

```
function[R] = c2dt1d(R,m,n)
% conversion from 2D array to 1D vector
% input=> R: 2D image/array, m: no. of rows and n: no. of columns
% output=> R: 1d vector data
R(2:2:end,:)=R(2:2:end,end:-1:1);
R = reshape(R',1,m*n);
```

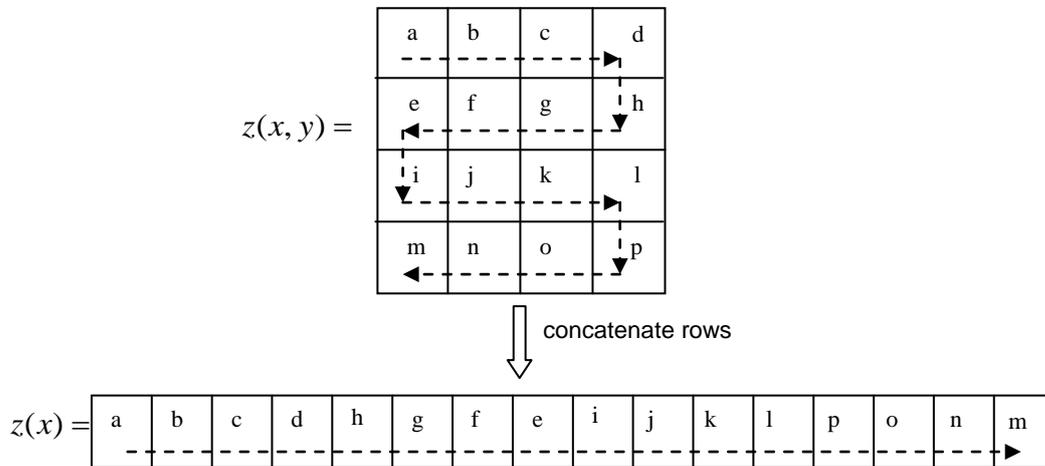


Figure 1a: Concatenation of rows to convert 2D image into 1D vector data

The 2D image can be reconstructed from the 1D vector data by reversing the procedure outlined in Figure 1a and the corresponding Matlab code is:

```
function[R] = c1d2d(R,m,n)
% conversion from 1D vector to 2D array
% input=> R: 1d vector data, m: no. of rows and n: no. of columns
% output=> R: 2D image/array
R = reshape(R,n,m)';
R(2:2:end,:)=R(2:2:end,end:-1:1);
```

Similarly, the image $z(x, y)$ of size $M \times N$ is divided into columns and concatenates these columns to form a 1D vector data $z(x)$ whose size would be MN as shown in Fig.1b. The Matlab code is: $R = c2dt1d(R',m,n)$;

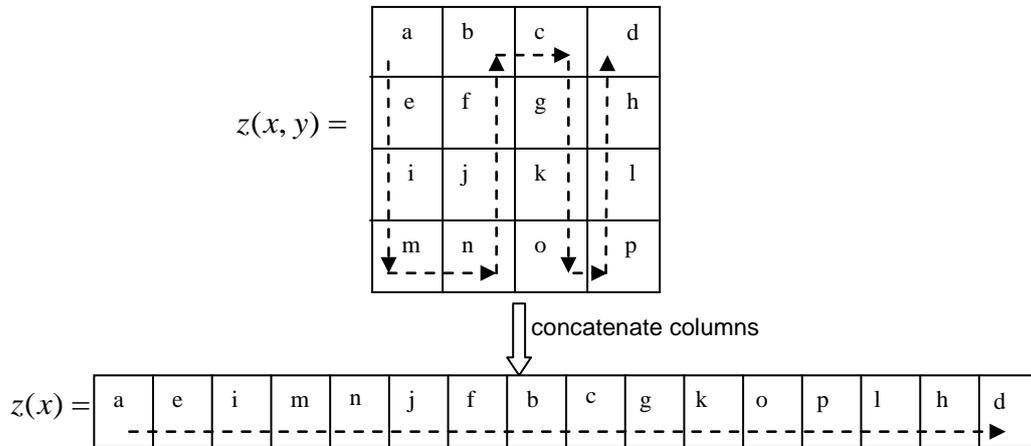


Figure 1b: Concatenation of columns to convert 2D image into 1D vector data

Fusion process is done on both row and column images separately and averaged together to avoid any noise or distortion introduced in the fusion process.

2.1.1 Frequency Partition (FP)

One dimensional (1D) DCT is applied on vector data $z(x)$ and partition the DCT coefficients into low frequency (LF) and high frequency (HF) components with a partition factor f as shown in Figure 2 using energy compaction property of DCT coefficients.

$$Z(u) = DCT(z(x)), \quad x, u = 0, 1, 2, \dots, MN - 1 \quad (3a)$$

$$ZL(u) = Z(u), \quad u = 0, 1, 2, \dots, MNf - 1 \quad (3b)$$

$$ZH(u) = Z(u), \quad u = MNf, MNf + 1, \dots, MN - 1 \quad (3c)$$

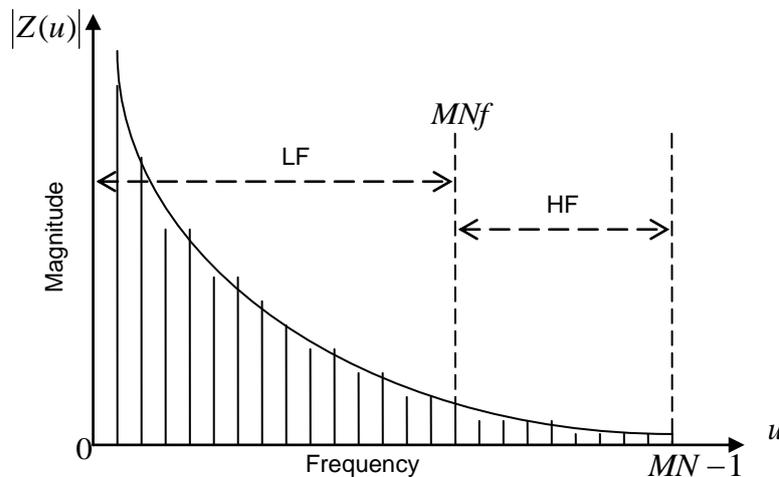


Figure 2: Separation of LF & HF coefficients



Let the images to be fused are $z_1(x, y)$ & $z_2(x, y)$ and the image fusion process is as follows:

$$z_1(x) = c2dt1d(z_1(x, y), M, N) \quad (4a)$$

$$z_2(x) = c2dt1d(z_2(x, y), M, N) \quad (4b)$$

$$Z_1(u) = DCT(z_1(x)) \quad (4c)$$

$$Z_2(u) = DCT(z_2(x)) \quad (4d)$$

Using eq. 3, the fused coefficients are:

$$ZL_f(u) = 0.5(ZL_1(u) + ZL_2(u)), u = 0, 1, \dots, MNf - 1 \quad (5a)$$

$$ZH_f(u) = \begin{cases} ZH_1(u) & \text{if } |ZH_1(u)| \geq |ZH_2(u)| \\ ZH_2(u) & \text{if } |ZH_1(u)| < |ZH_2(u)| \end{cases}, u = MNf, MNf + 1, \dots, MN - 1$$

(5b)

$$Z_f(u) = [ZL_f(u) \quad ZH_f(u)] \quad (5c)$$

$$z_f(x) = idct(Z_f(u)), x, u = 0, 1, 2, \dots, MN - 1 \quad (5d)$$

The fused image is:

$$I_f = c1dt2d(z_f(x), M, N) \quad (5e)$$

Where the subscript 1 or 2 or f indicates 1st or 2nd or fused image respectively.

2.1.2 Multi-resolution DCT (MSDCT)

The multi-resolution analysis is illustrated in Figure 3. The vector data is passed through DCT. Consider first half of the DCT coefficients as LF and the remaining as HF coefficients. The LF coefficients are passed through IDCT to get the vector data for next level of decomposition as shown in Figure 3. Let $z_k(x) = z(x)$ at $k = 1$ level and at each k^{th} level:

$$Z_k(u) = dct(z_k(x)), x, u = 0, 1, \dots, \frac{MN}{2^{k-1}} - 1 \quad (6a)$$

$$ZL_k(u) = Z_k(u), u = 0, 1, 2, \dots, \frac{0.5MN}{2^{k-1}} - 1 \quad (6b)$$

$$ZH_k(u) = Z_k(u), u = \frac{0.5MN}{2^{k-1}}, \frac{0.5MN}{2^{k-1}} + 1, \dots, MN - 1 \quad (6c)$$

$$z_{k+1}(x) = idct(ZL_k(u)) \quad (6d)$$

Where the subscript k indicates multi resolution decomposition level.

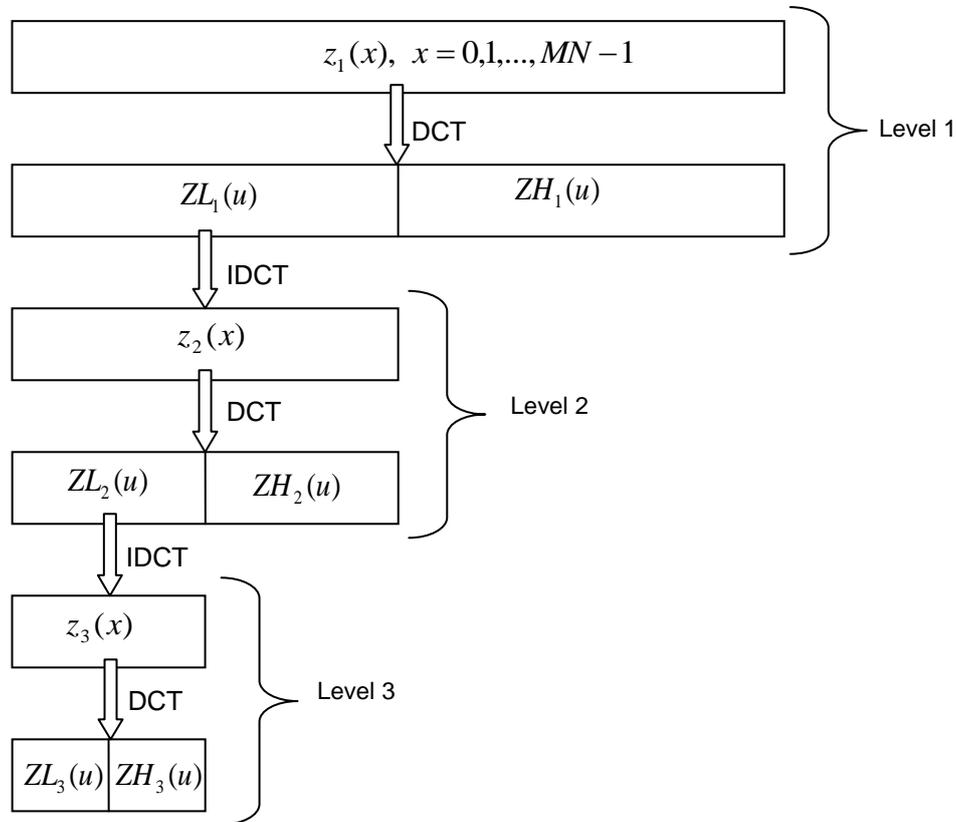


Figure 3: Multi resolution image analysis using 1D DCT

Let the images to be fused are $z_1(x, y)$ & $z_2(x, y)$ and the image fusion process is as follows. Using eq.4 and 6, the fused coefficients are:

$$ZL_{kf}(u) = 0.5(ZL_{k1}(u) + ZL_{k2}(u)), \quad u = 0, 1, \dots, 0.5 \frac{MN}{2^{k-1}} - 1$$

$$k = K \quad (7a)$$

$$ZH_{kf}(u) = \begin{cases} ZH_{k1}(u) & \text{if } |ZH_{k1}(u)| \geq |ZH_{k2}(u)| \\ ZH_{k2}(u) & \text{if } |ZH_{k1}(u)| < |ZH_{k2}(u)| \end{cases}, \quad u = 0.5 \frac{MN}{2^{k-1}}, 0.5 \frac{MN}{2^{k-1}} + 1, \dots, \frac{MN}{2^{k-1}} - 1 - 1$$

$$k = K, K - 1, \dots, 1 \quad (7b)$$

Using eq.7, the fused image can get by doing the procedure outlined in Figure 3 reversely.

2.1.3 Laplacian Pyramid (LPID)

Laplacian pyramid provides information on the sharp contrast changes to which human visual system is principally sensitive to. It gives both spatial and frequency domain localization. The procedure for Laplacian pyramid



construction and reconstruction is illustrated in Figure 4. The vector data $z(x)$ at the 0th level $z_0(x)$ of size MN is reduced to obtain next level $z_1(x)$ of size $0.5MN$ where both spatial density and resolution are reduced. Similarly, $z_2(x)$ is the reduced version of $z_1(x)$ and soon. Image reduction is done by taking the DCT and applying the IDCT on first half of coefficients. The level to level image reduction is performed using the function reduce R .

Reduction Function R :

$$z_k = R(z_{k-1}) \tag{8a}$$

The reverse of function reduce is expand function E . Its effect is to expand vector data of size $0.5MN$ to data of size M by taking IDCT after padding the $0.5MN$ zeros.

Expand Function E :

$$\hat{z}_k = E(\hat{z}_{k+1}) \tag{8b}$$

Construction of pyramid is done using (Figure 4):

$$z_{k+1} = R(z_k) \tag{9a}$$

$$l_k = z_k - E(z_{k+1}) \tag{9b}$$

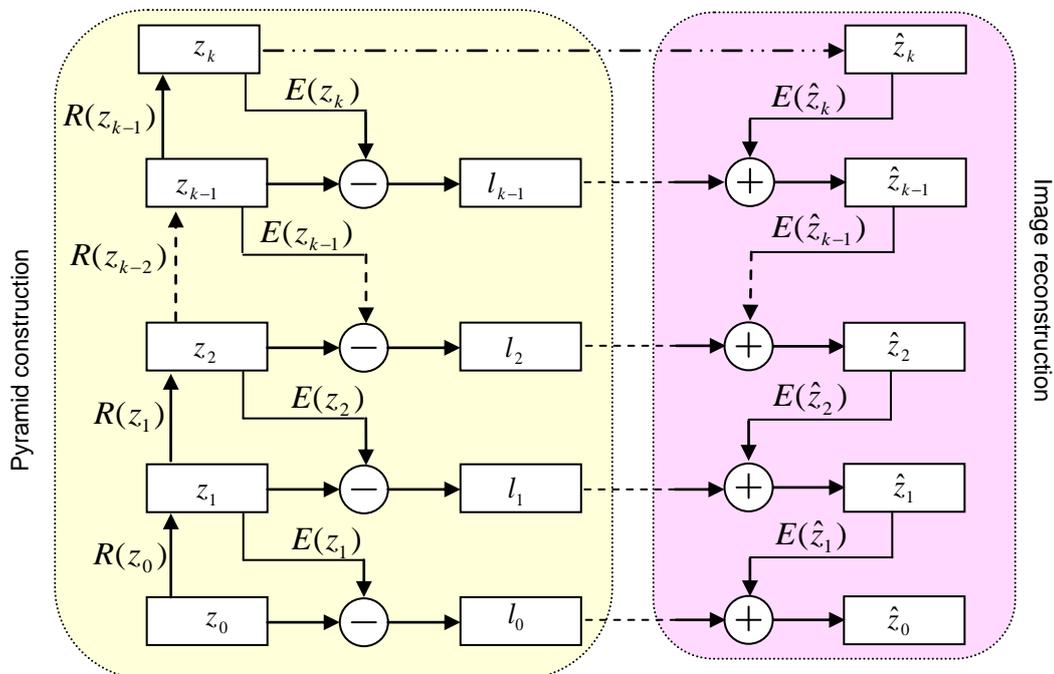


Figure 4: Laplacian pyramid construction



Pyramid construction is done for each image to be fused using eq. 9. Denote the constructed k levels of Laplacian pyramid for 1st image is $P_1 \rightarrow \{z_{K1}, l_{01}, l_{11}, \dots, l_{(K-1)1}\}$ and similarly for of 2nd image is $P_2 \rightarrow \{z_{K2}, l_{02}, l_{12}, \dots, l_{(K-1)2}\}$.

Then the image fusion rule is as follows:

$$\text{At } k^{\text{th}} \text{ level, } z_{Kf} = 0.5(\hat{z}_{K1} + \hat{z}_{K2}) \quad (10a)$$

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

Where $l_{(k-1)f} = \begin{cases} l_{(k-1)1} & |l_{(k-1)1}| \geq |l_{(k-1)2}| \\ l_{(k-1)2} & |l_{(k-1)1}| < |l_{(k-1)2}| \end{cases}$ and the magnitude comparison is

done on corresponding pixels.

The pyramid $I_f = z_{0f}$ is the fused image.

2.2 Using 2D DCT

In this section, the images to be fused are not converted into vector data. The fusion techniques presented in this section are directly operated on 2D array (image) data.

2.2.1 DMDCT (Distributed Multi-resolution DCT)

Distributed multi resolution DCT (DMDCT) [10] is localized in both spatial and frequency domain and it could be applicable to non-stationary signals. The DMDCT decomposition process of an image size 4x4 is shown in Figure 5. The spatial domain image is transformed DCT domain using 2D DCT. Addition and subtraction of coefficients as shown in Figure 5 is done and segregated the processed coefficients as four sub bands such as LL , LH , HL and HH (as in Wavelet decomposition). The procedure is repeated on the LL band for next level of decomposition. Similarly, image can be reconstructed by reversing the procedure depicted in Figure 5.

The images to be fused are decomposed into K ($k = 1, 2, \dots, K$) levels using DMDCT. The resultant decomposed images from $z_1(x, y)$ are

$$I_1 \rightarrow \left\{ LL_{K1}, \{ LH_{k1}, HH_{k1}, HL_{k1} \}_{k=1,2,\dots,K} \right\} \text{ and from } z_2(x, y) \text{ are}$$

$$I_2 \rightarrow \left\{ LL_{K2}, \{ LH_{k2}, HH_{k2}, HL_{k2} \}_{k=1,2,\dots,K} \right\}.$$

At each decomposition level ($k = 1, 2, \dots, K$), the fusion rule will select the larger absolute value of the two DMDCT 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 ($k = K$), the fusion rule take average of the DMDCT approximation coefficients since the approximation coefficients at coarser



level are the smoothed and subsampled version of the original image. The complete set of fusion rules are:

$$LH_{kf} = \begin{cases} LH_{k1} & |LH_{k1}| \geq |LH_{k2}| \\ LH_{k2} & |LH_{k1}| < |LH_{k2}| \end{cases} \quad (11a)$$

$$HH_{kf} = \begin{cases} HH_{k1} & |HH_{k1}| \geq |HH_{k2}| \\ HH_{k2} & |HH_{k1}| < |HH_{k2}| \end{cases} \quad (11b)$$

$$HL_{kf} = \begin{cases} HL_{k1} & |HL_{k1}| \geq |HL_{k2}| \\ HL_{k2} & |HL_{k1}| < |HL_{k2}| \end{cases} \quad (11c)$$

$$LL_{Kf} = 0.5(LL_{K1} + LL_{K2}) \quad (11d)$$

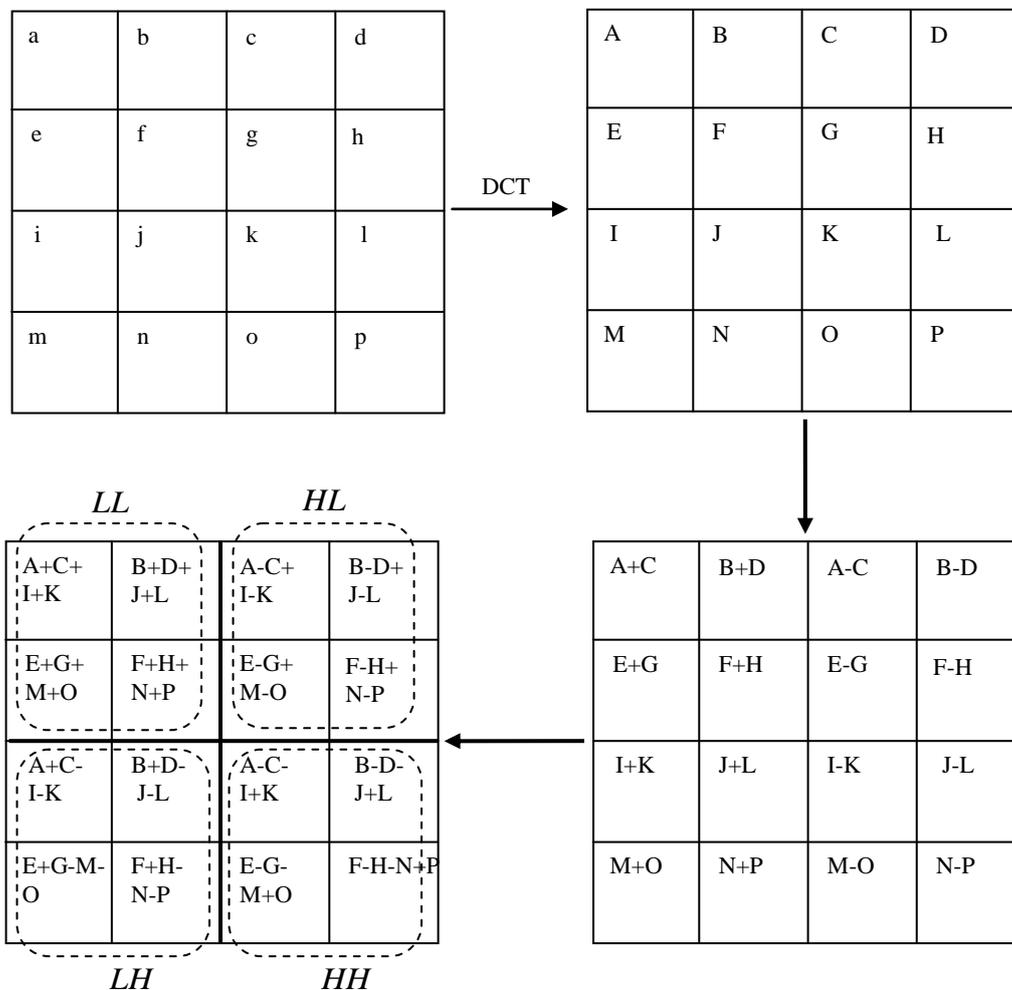


Figure 5: Distributed multi resolution DCT of a 4x4 image



The fused image I_f can be obtained by revercing the procedure explained with Figure 5 :

$$I_f \leftarrow \{LL_K, \{LH_k, HH_k, HL_k\}_{k=1,2,\dots,K}\} \quad (11e)$$

2.2.2 DTMDCT (Dual tree MDCT)

Dual Tree MDCT (DTMDCT) uses two multi resolution discrete Cosine transform (MDCT) trees. One MDCT [7] tree is used for real part of the analytic signal and the other MDCT tree is used for imaginary part of the analytical signal. The analytical signal is achieved with Hilbert transform. The information flow diagram of DTMDCT is shown in Figure 6. More details on MDCT can be found in [7].

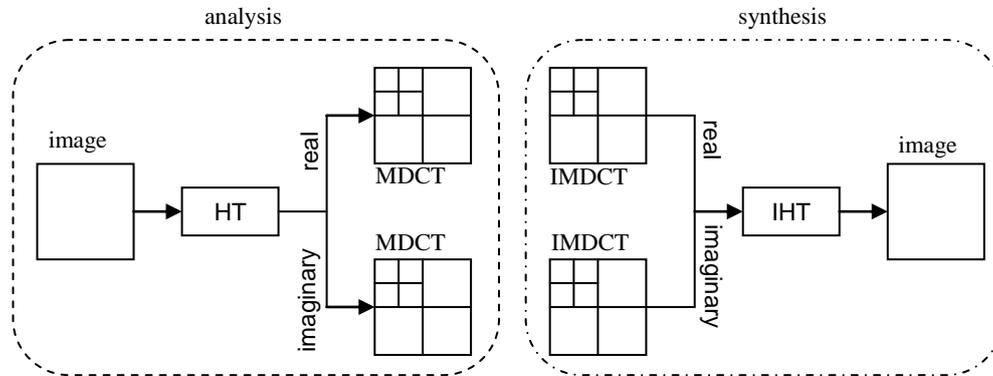


Figure 6: Dual tree MDCT analysis and synthesis

The registered images $z_1(x, y)$ and $z_2(x, y)$ to be fused are decomposed into K ($k = 1, 2, \dots, K$) levels using DTMDCT. The resultant decomposed images from $z_1(x, y)$ are $I_1 \rightarrow \{LL_{TK1}, \{LH_{Tk1}, HH_{Tk1}, HL_{Tk1}\}_{k=1,2,\dots,K} \& T=1,2\}$ and from $z_2(x, y)$ are $I_2 \rightarrow \{LL_{TK2}, \{LH_{Tk2}, HH_{Tk2}, HL_{Tk2}\}_{k=1,2,\dots,K} \& T=1,2\}$. The complete set of fusion rules are:

$$LH_{Tkf} = \begin{cases} LH_{Tk1} & |LH_{Tk1}| \geq |LH_{Tk2}| \\ LH_{Tk2} & |LH_{Tk1}| < |LH_{Tk2}| \end{cases} \quad (12a)$$

$$HH_{Tkf} = \begin{cases} HH_{Tk1} & |HH_{Tk1}| \geq |HH_{Tk2}| \\ HH_{Tk2} & |HH_{Tk1}| < |HH_{Tk2}| \end{cases} \quad (12b)$$

$$HL_{Tkf} = \begin{cases} HL_{Tk1} & |HL_{Tk1}| \geq |HL_{Tk2}| \\ HL_{Tk2} & |HL_{Tk1}| < |HL_{Tk2}| \end{cases} \quad (12c)$$

$$LL_{TKf} = 0.5(LL_{TK1} + LL_{TK2}) \quad (12d)$$



Where the subscript T indicates tree ($T = 1$ for real and $T = 2$ for imaginary).

The fused image I_f can be obtained by revercing the procedure explaied with Figure 6 :

$$I_f \leftarrow \left\{ LL_{TKf}, \{ LH_{TKf}, HH_{TKf}, HL_{TKf} \}_{k=1,2,\dots,K} \& T=1,2 \right\} \quad (12e)$$

2.2.3 Laplacian Pyramid (LP2D)

The procedure for Laplacian pyramid construction and reconstruction is illustrated in Figure 1. The image at the 0th level z_0 of size $M \times N$ is reduced 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. Image reduction is done by taking the DCT and applying the IDCT on first half of coefficients in both directions. The level to level image reduction is performed using the function reduce R .

Reduction Function R :

$$z_k = R(z_{k-1}) \quad (13)$$

The reverse of function reduce is expand function E . Its effect is to expand the image of size $M \times N$ to image of size $2M \times 2N$ by taking IDCT after padding the M zeros in horizontal and N zeros in vertical directions.

Expand Function E :

$$\hat{z}_k = E(\hat{z}_{k+1}) \quad (14)$$

Construction of pyramid is done using (Figure 4):

$$z_{k+1} = R(z_k) \quad (15a)$$

$$l_k = z_k - E(z_{k+1}) \quad (15b)$$

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. Reconstruction of the pyramid is done using (Figure 1b):

$$\hat{z}_{k-1} = l_{k-1} + E(\hat{z}_k) \quad (16)$$

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 $P_k^1 \rightarrow \{z_k^1, l_0^1, l_1^1, \dots, l_{k-1}^1\}$ and similarly for of 2nd image is $P_k^2 \rightarrow \{z_k^2, l_0^2, l_1^2, \dots, l_{k-1}^2\}$.



Then the image fusion rule is as follows:

$$\text{At } k^{\text{th}} \text{ level, } z_k^f = \frac{\hat{z}_k^1 + \hat{z}_k^2}{2} \quad (17)$$

$$\text{For } k-1 \text{ to } 0 \text{ levels } z_{k-1}^f = I_{k-1}^f + E(z_k^f) \quad (18)$$

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

corresponding pixels.

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

3. 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.

3.1 Percentage Fit Error (PFE) [11]

PFE is computed as the norm of the difference between the corresponding pixels of reference and fused image to the norm of the reference image. This will be zero when both reference and fused images are exactly similar and it will be increased when the fused image is deviated from the reference image. The PFE is computed as:

$$PFE = \frac{\text{norm}(I_r - I_f)}{\text{norm}(I_r)} * 100 \quad (19)$$

where *norm* is the operator to compute the largest singular value

3.2 Peak Signal to Noise Ratio (PSNR) [1,11]

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

$$PSNR = 20 \log_{10} \left(\frac{L^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (I_r(i, j) - I_f(i, j))^2} \right) \quad (20)$$

where *L* in the number of gray levels in the image

3.3 Standard Deviation (SD) [11]

It is known that standard deviation is composed of the signal and noise parts. This metric would be more efficient in the absence of noise. It measures the contrast in the fused image. An image with high contrast would have a high standard deviation.



$$\sigma = \sqrt{\sum_{i=0}^L (i - \bar{i})^2 h_{I_f}(i)}, \quad \bar{i} = \sum_{i=0}^L i h_{I_f}(i) \quad (21)$$

where $h_{I_f}(i)$ is the normalized histogram of the fused image $I_f(x, y)$ and

L number of frequency bins in the histogram.

3.4 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:

Row frequency:

$$RF = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=1}^{N-1} [I_f(i, j) - I_f(i, j-1)]^2} \quad (22a)$$

Column frequency:

$$CF = \sqrt{\frac{1}{MN} \sum_{j=0}^{N-1} \sum_{i=1}^{M-1} [I_f(i, j) - I_f(i-1, j)]^2} \quad (22b)$$

$$\text{Spatial frequency: } SF = \sqrt{RF^2 + CF^2} \quad (22c)$$

4. RESULTS AND DISCUSSION

The fusion algorithms developed in section 3 are evaluated using the images shown in Figure 7. The ground truth image is shown in Figure 7a and the images to be fused are shown in Figure 7b&c. Both Figure 7b and Figure 7c 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 Figure 7b and vice versa in image I_2 as shown in Figure 7c. The out of focus has been created by blurring the portions of the reference image with a Gaussian mask using diameter of 12 pixels. The fused (left side) along with the error (right side) images are shown in Figure 8 to 13. 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.



Figure 7a: Ground truth image



Figure 7b&c: Images to be fused

Table 1: Fusion quality evaluation metrics for frequency partition

	Partition scale factor (f)					
	0.0	0.2	0.4	0.6	0.8	1.0
PFE	3.0913	3.6284	3.8935	3.9818	4.0095	4.0194
PSNR	39.5879	38.8922	38.5859	38.4886	38.4584	38.4478
SD	48.9726	46.3338	46.0018	45.9041	45.8739	45.8628
SF	15.4065	12.7195	11.0493	9.9335	9.3618	9.0953



Table 2: Fusion quality evaluation metrics for 1D multi-resolution DCT

	No. of decomposition levels (K)				
	1	2	3	4	5
PFE	3.9508	3.7245	3.3793	3.0959	3.0593
PSNR	38.5225	38.7786	39.2011	39.5815	39.6332
SD	45.9385	46.2037	46.6796	47.2891	47.8383
SF	10.4128	12.2229	13.4712	14.2075	14.7721

Table 3: Fusion quality evaluation metrics for LP1D

	No. of pyramid levels (K)				
	1	2	3	4	5
PFE	3.8937	3.4647	2.7186	1.8675	1.4672
PSNR	38.5858	39.0927	40.1458	41.7767	42.8246
SD	45.9673	46.3366	46.9828	47.8254	48.5444
SF	10.8488	13.1805	14.7043	15.5537	16.1210

Table 4: Fusion quality evaluation metrics for distributed multi-resolution DCT

	levels				
	1	2	3	4	5
PFE	3.4946	3.8218	3.8018	3.8276	3.8274
PSNR	39.0554	38.6667	38.6895	38.6601	38.6603
SD	48.3188	48.7771	48.9340	48.9260	48.9381
SF	16.2207	17.4085	17.6379	17.7290	17.7415

Table 5: Fusion quality evaluation metrics for dual tree MDCT

	levels				
	1	2	3	4	5
PFE	3.8779	3.4505	2.4963	1.3788	1.0743
PSNR	38.6029	39.1101	40.5159	43.0940	44.1775
SD	46.1263	46.8081	47.8777	49.0998	49.6608
SF	12.4863	15.4813	16.5735	16.8939	16.9356



Table 6: Fusion quality evaluation metrics for LP2D analysis

	No. of pyramid levels (K)								
	1	2	3	4	5	6	7	8	9
PFE	3.88	3.31	2.31	1.05	0.60	0.53	0.52	0.52	0.52
PSNR	38.6	39.3	40.8	46.9	44.2	46.5	47.0	47.0	47.0
SD	46.1	46.8	47.9	49.1	49.7	50.0	50.1	50.1	50.1
SF	12.5	15.5	16.6	17.0	17.0	17.0	16.9	16.9	16.9



Figure 8: Fused and error image using frequency partition technique

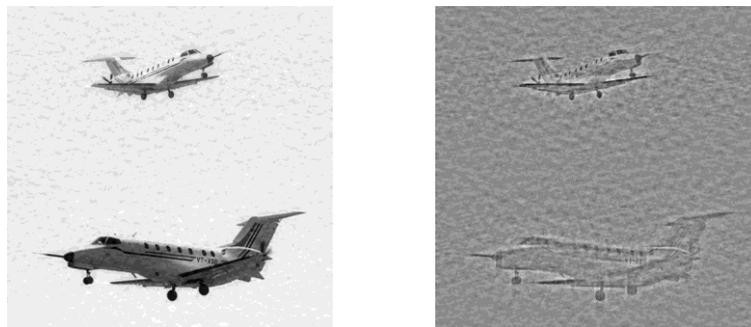


Figure 9: Fused and error image using frequency 1D multi-resolution analysis (MSDCT)

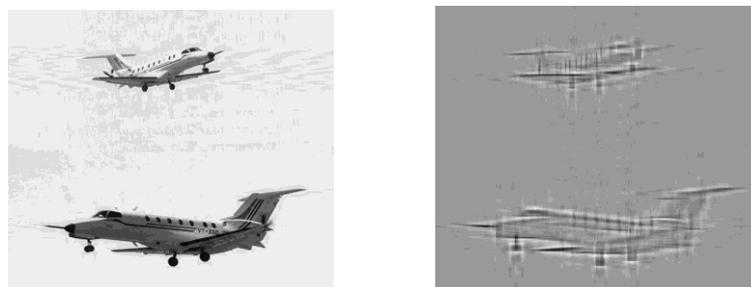


Figure 10: Fused and error image using Laplacian pyramid analysis

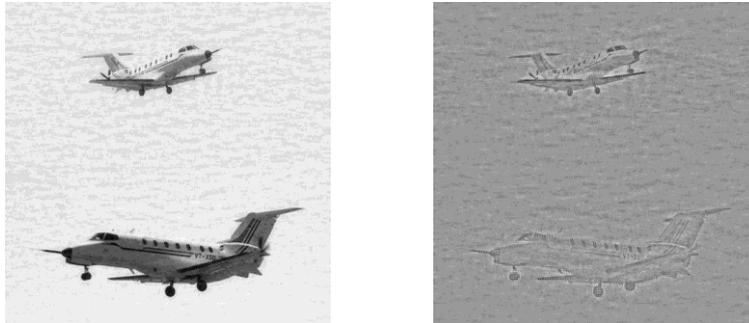


Figure 11: Fused and error image using distributed multi-resolution DCT

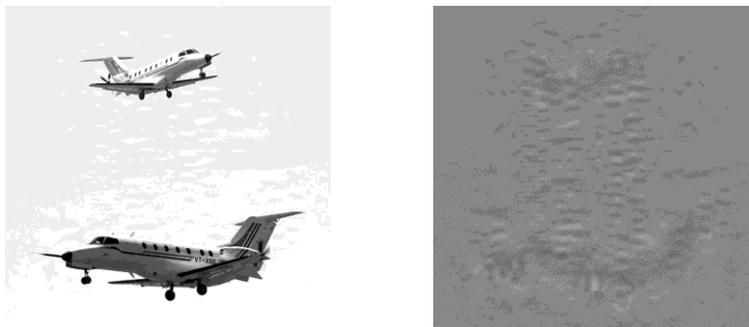


Figure 12: Fused and error image using Dual tree MDCT

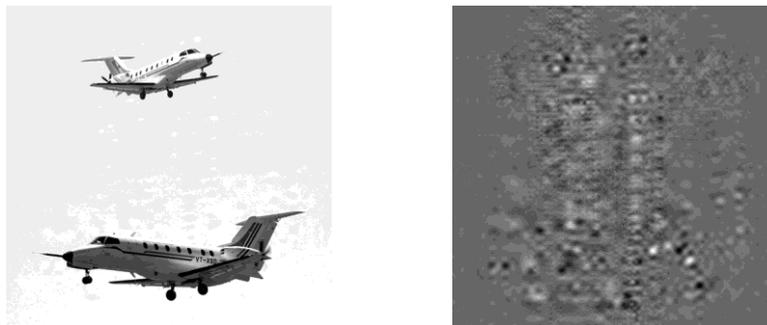


Figure 13: Fused and error images using eight level LP2D

4. CONCLUSIONS

Six different novel image fusion techniques using DCT are presented and studied. Based on the results, it is concluded that DTMDCT provides better fusion results followed by Laplacian pyramid based image fusion technique. DTMDCT is a simple and computationally efficient algorithm.



REFERENCES

- [1] 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.
- [2] Li H, Manjunath B.S and Mitra S.K, "Multisensor image fusion using the wavelet transform", *Graphical Models and Image Processing*, 57(3), pp.235-245, 1995.
- [3] Toet A, "Image fusion by a ratio of low-pass pyramid", *Pattern Recognition letters*, 9(4), pp.245-253, 1989.
- [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] Abdolhossein Nejatali and Loan R. Ciric, "Novel image fusion methodology using fuzzy set theory", *Opt. Eng*, 37(2), pp.485-491, Feb. 1998.
- [6] Barbara Zitova and Jan Flusser, "Image registration methods: a survey", *Image and Vision Computing*, 21, pp.997-1000, June 2003.
- [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.
- [9] VPS Naidu, "Multi Focus Image Fusion using the Measure of Focus", *Journal of Optics*, published by Springer (DOI 10.1007/s12596-012-0071-3), April 2012.
- [10] Gaurav Bhatnagar and Balasubramanian Raman, "Robust watermarking using distributed MR-DCT and SVD", *7th International conference on advances in pattern recognition*, 2007, pp.21024, retrieved may 11, 2009, from IEEEExplore database.
- [11] VPS Naidu, "Multi-Resolution Image Fusion by FFT", *ICIIP-2011*, 3-5 Nov. 2011. (IEEE DoI: 10.1109/ICIIP.2011.6108862).

Matlab code for all algorithms is provided at:

<http://www.mathworks.in/matlabcentral/fileexchange/43566-novel-image-fusion-techniques-using-dct>