

# Low complexity multifocus image fusion in discrete cosine transform domain

ASNATH VICTY PHAMILA\*, R. AMUTHA

Department of Electronics and Communication Engineering, SSN College of Engineering, SSN Nagar, Chennai, India – 603110

\*Corresponding author: phamila@yahoo.com

This paper presents a low complex, highly energy efficient sensor image fusion scheme explicitly designed for wireless visual sensor systems equipped with resource constrained, battery powered image sensors and employed in surveillance, hazardous environment like battlefields *etc.* Here an energy efficient simple method for fusion of multifocus images based on higher valued AC coefficients calculated in discrete cosine transform domain is presented. The proposed method overcomes the computation and energy limitation of low power devices and is investigated in terms of image quality and computation energy. Simulations are performed using Atmel ATmega128 processor of Mica 2 mote, to measure the resultant energy savings and the simulation results demonstrate that the proposed algorithm is extremely fast and consumes only around 1% of energy consumed by conventional discrete cosine transform based fusion schemes. Also the simplicity of our proposed method makes it more appropriate for real-time applications.

Keywords: sensor image fusion, discrete cosine transform (DCT), energy consumption, computation complexity, fusion metrics.

## 1. Introduction

Image fusion is the process of combining multiple source images from sensor network into a single one, which contains a more accurate description of the scene, more informative and suitable for both visual perception and further processing [1]. In the multifocus image fusion technique, several images of a scene captured with focus on different objects are fused such that all the objects will be in focus in the resulting image. So far, several researches have been focused on image fusion which is performed on the images in the spatial and spectral domain [2–5]. The sharpness measure [3] is exploited to perform adaptive multifocus image fusion in wavelet domain and proved to give better fused results than other discrete wavelet transform (DWT) based multifocus image fusion schemes. In automated battlefields, where the robots collect image data from sensor network, since the computational energy is much less than the transmission energy, data are compressed and fused before transmission [6]. Hence, when the source images are to be coded in Joint Photographic Experts Group (JPEG) standard or when the resultant fused image is to be saved or transmitted in JPEG format, the fusion methods in discrete cosine transform (DCT) domain will be more efficient [4, 5].

But the standard JPEG-DCT and the mean, contrast and variance calculation [4, 5] involve complex floating point arithmetic operations which incur high energy consumption in resource constrained sensor nodes. In general, the transform (DCT and DWT) based image fusion methods are complex, time-consuming and highly energy-consuming which are not best suitable for real-time applications and in resource constrained battery powered sensors.

In this paper, a low-complexity and energy efficient multifocus image fusion scheme that is suitable for resource constrained (processing, bandwidth, memory space, battery power) image sensor network is proposed. Binary DCT is used for image transform [7, 8] and the proposed fusion rule is very simple and considerably reduces the computational complexity without compromising image quality. Since the proposed fusion scheme is performed in DCT domain, it is time-saving and simple when the fused image needs to be saved or transmitted in JPEG format. Simulations are performed using Atmel ATmega128 processor of Mica 2 mote to measure the resultant energy savings.

## 2. Image fusion

Figure 1 illustrates the general framework of our proposed image fusion scheme. The algorithm can be extended for more than two source images with the assumption that all the source images are aligned with some registration methods.

In the proposed scheme, the key step is to fuse the DCT representations of multifocus images into a single fused image [4, 5]. The input images are divided into blocks of size  $8 \times 8$  and the DCT coefficients of each block are computed. Then the fusion rule is applied wherein the transformed block with bigger number of higher valued AC coefficients is absorbed into the fused image.

### 2.1. Fusion criteria (MaxAC)

In [4] variance is used as the activity level for fusion criteria because in multifocus images, the focused region is more informative and the information details of that

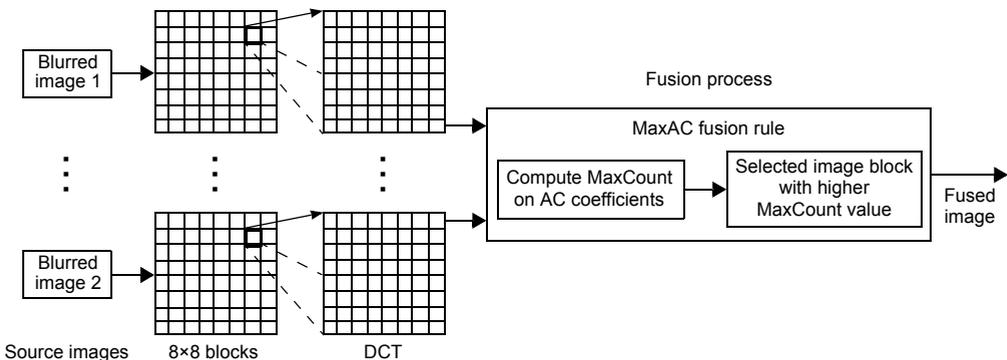


Fig. 1. General structure of the proposed fusion scheme.

region correspond to high variance. It is inferred that the variance of an  $N \times N$  block of pixels can be exactly calculated from its DCT coefficients by computing the sum of the squared normalized AC coefficients of the DCT block [4]. Two dimensional DCT transform of an  $N \times N$  image block  $f(x, y)$  is given as

$$F(u, v) = \frac{2}{N} C(u)C(v) \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x, y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \tag{1}$$

where  $u, v = 0, 1, \dots, N-1$  and

$$c(u) = \begin{cases} 1/\sqrt{2} & \text{if } u = 0 \\ 1 & \text{if } u \neq 0 \end{cases} \tag{2}$$

The inverse transform is defined as

$$f(x, y) = \frac{2}{N} \sum_{v=0}^{N-1} \sum_{u=0}^{N-1} C(u)C(v)F(u, v) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \tag{3}$$

where  $x, y = 0, 1, \dots, N-1$ .

Here  $F(0,0)$  is the DC coefficient and it represents the mean value of that image block. Remaining coefficients are AC coefficients and the normalized transform coefficients are defined as

$$\hat{F}(u, v) = \frac{F(u, v)}{N} \tag{4}$$

Variance  $\sigma^2$  of the image block can be inferred from the transformed coefficients as follows

$$\sigma^2 = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \frac{F^2(u, v)}{N^2} - \hat{F}^2(0, 0) \tag{5}$$

where  $\hat{F}(0, 0)$  is the normalized DC coefficient and other  $\hat{F}(u, v)$  are the normalized 63 AC coefficients ( $A_1, A_2, \dots, A_{63}$ ). Equation (5) implies that the variance of a block is given by the sum of the squares of the normalized AC coefficients.

$$\sigma^2 = \sum_{i=1}^{63} \hat{A}_i^2 \tag{6}$$

The advantage of DCT is that the energy of the original data is concentrated in only a few low frequency components of DCT depending on the correlation in the data. Also the low-frequency components usually contain the most of the image information. Higher value of AC coefficients implies finer image information. Because of the energy compaction property of AC coefficients, only few coefficients towards the top

left submatrix of the DCT transformed matrix have larger values and the contribution of these coefficients to variance is more compared to other AC coefficients. Hence, if the AC coefficient value is high, then the variance value is also high. Our proposed fusion criterion (MaxAC) is based on this fact. Here instead of computing variance using all the transformed AC coefficients, which involves floating point multiplication and additions, the proposed algorithm checks only the number of higher valued AC coefficients that contributes to larger variance.

Hence, in our proposed method we absorb the block with a bigger number of higher valued AC coefficients for two reasons. First is that a higher AC component value implies more fine details of the image. Secondly, from Eq. (6) it is inferred that a higher AC component value results in higher variance. Thus the energy needed for computation is drastically reduced. The quality of the fused image is significantly high because only the blocks where more image details are stored (bigger number of high valued AC coefficients) are selected for fusion. Here, instead of computing variance using all the transformed AC coefficients which involves floating point multiplication and additions, the proposed algorithm checks only the number of higher valued AC coefficients that contributes to larger variance. Thus the energy needed for computation is drastically reduced.

**2.2 MaxAC fusion method**

Let  $Y = \{y_{i,j}\}$  ( $i = 0, \dots, N - 1$  and  $j = 0, \dots, M - 1$ ) be an image and it is divided into  $Q$  number of  $8 \times 8$  blocks. Let  $X_n = \{x_{n,k,l}\}$  ( $k = 0, \dots, 7; l = 0, \dots, 7; n = 0, \dots, Q - 1$ ) be the  $n$ -th  $8 \times 8$  block and the corresponding DCT output of the block  $X_n = \{x_{n,k,l}\}$  be  $D_n = \{d_{n,k,l}\}$ . Then the set  $D = \{D_0, D_1, D_2, \dots, D_{Q-1}\}$  denotes the DCT representation of image  $Y = \{y_{i,j}\}$ . Let  $D^t = \{D_0^t, D_1^t, D_2^t, \dots, D_{Q-1}^t\}$  be the DCT representation of the  $t$ -th input image and let  $B$  be the number of input source images to be fused. Then the fused image is represented by  $D^F = \{D_0^F, D_1^F, D_2^F, \dots, D_{Q-1}^F\}$ . The fusion criterion is that the block with the majority of maximum valued DCT AC coefficients is absorbed into the fused image since it contributes more significant signal information to the fused image.

Hence in our fusion method, the  $n$ -th block of the fused image  $D_n^F$  is obtained by:

$$D_n^F = Y_n^T \tag{7}$$

$$Y_n^T = \arg \max_t \{C_n^t\}, \quad t = 1, \dots, B \tag{8}$$

where  $C_n^t$  specifies the number of maximum valued transformed AC coefficients found in the  $n$ -th block of  $t$ -th image when compared with the respective blocks in other source images. For example, the fused block  $D_n^F$  of two source images  $Y^t$  where  $t = 1, 2$  is obtained as follows:

1. Initialize  $\{C_n^1 = 0; C_n^2 = 0;\}$
2. For all the AC coefficients repeat step 3

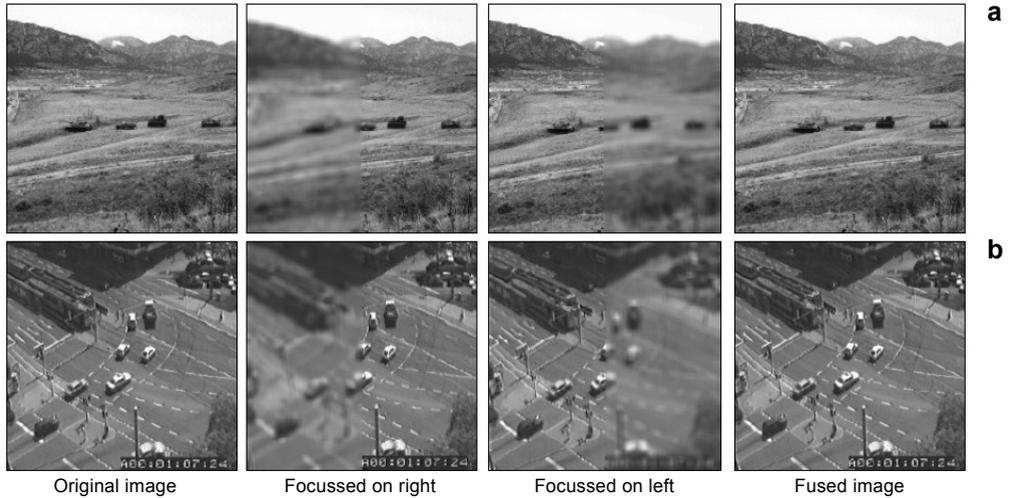


Fig. 2. Fusion results of multifocus images: battlefield (a), and traffic (b).

3. If  $(D_n^1(i, j) > D_n^2(i, j) + \text{threshold})$  then  
     Increment  $C_n^1$  by 1;  
     Else If  $(D_n^2(i, j) > D_n^1(i, j) + \text{threshold})$  then  
         Increment  $C_n^2$  by 1;
4. If  $(C_n^1 > C_n^2)$  then  
      $D_n^F = Y_n^1$  and  $\text{Prev} = 1$   
     Else If  $(C_n^2 > C_n^1)$   
          $D_n^F = Y_n^2$  and  $\text{Prev} = 2$   
     Else  
          $D_n^F = Y_n^{\text{Prev}}$   
     EndIf

This is repeated for all  $Q$  blocks to fuse the DCT representations of multi-images into a fused image. The *threshold* value is set between 0 and 10. Figure 2 depicts the fusion result of the proposed method on multifocused traffic and battlefield images.

### 3. Simulation results

The proposed fusion algorithm is applied on a set of non-referenced and a set of referenced images and the results are evaluated. The first experiment is conducted using an extensive set of artificially generated images with different focus levels. Standard test images like traffic, Lena, battlefield, Barbara, bird, *etc.* are taken as ground truth images [9, 10]. Two blurred artificial images are generated for each test image by convolving the test image with a  $9 \times 9$  averaging filter centered at the left and right part, respectively [3]. The second experiment is conducted on sets of standard non-referenced multifocus test images [11]. The fusion result of Lena, Pepsi and clock standard test images by applying various fusion algorithms in DCT domain [4, 12] (DCT + average, and DCT + variance), DWT domain (DWT [13] and SIDWT [14])

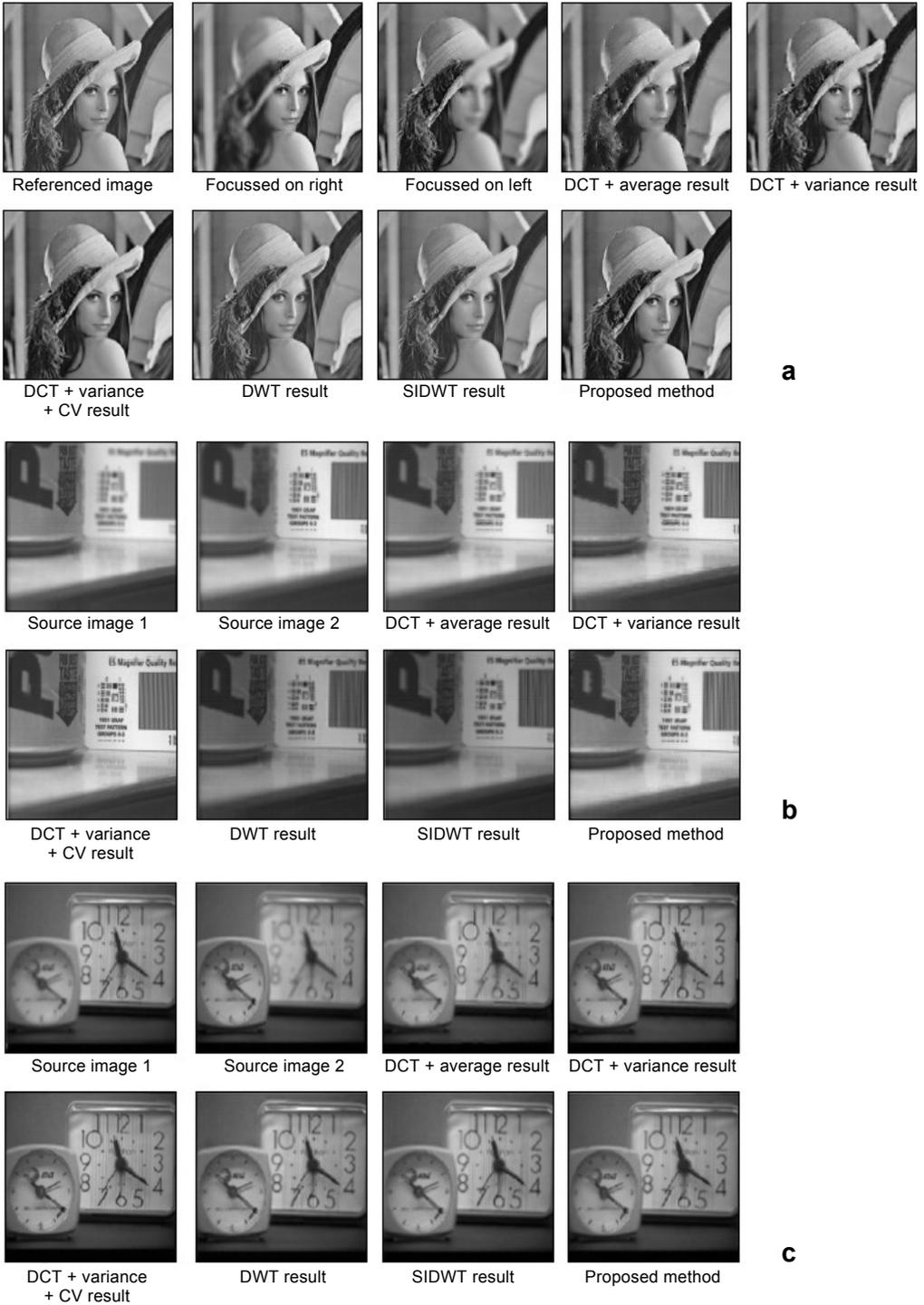


Fig. 3. Fusion results of Lena (a), Pepsi (b), and clock (c) images.

and our proposed method is shown in Figs. 3a, 3b and 3c, respectively. For the wavelet based methods (the DWT with DBSS (2,2) and the SIDWT with Haar basis), the simulation was carried out with the *Image Fusion Toolbox*, kindly provided by Rockinger [15]. Here the comparison of fusion methods is done similar to [4, 13].

## 4. Performance metrics

Extensive experiments are performed to demonstrate the superior performance of the proposed algorithm using six performance metrics. The structural similarity measure (SSIM) [4, 16], mean-square error (MSE) and peak signal-to-noise ratio (PSNR) are used as quality measures for objective evaluation of results of referenced images. If MSE equals zero, it implies that the fused image is exactly the same as that of the original referenced image. If  $MSE = 0$ , then the PSNR becomes infinity which implies ideal fusion, where the fused image and the referenced image are exactly identical. The higher the values of PSNR and SSIM, the better is the quality of the fused image. To evaluate our proposed algorithm on non-referenced multifocus images, the spatial frequency (SF) metric [2, 17], the state-of-the-art fusion performance metric Petrovic [4, 18, 19], the metric  $Q^{AB/F}$ , the mutual information (MI) [20], the mean gradient (MG) [17] and the feature mutual information (FMI) [21] are used.

### 4.1. Structural similarity index (SSIM)

The structural similarity measure (SSIM) [4, 16], as a quality criterion, is used for objective evaluation of a fused image. The general form of the metric that is used to measure the structural similarity between two signal vectors  $x$  and  $y$  is given by the following equation:

$$SSIM(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (9)$$

where  $\mu_x$  and  $\mu_y$  are the sample means of  $x$  and  $y$ , respectively,  $\sigma_x^2$  and  $\sigma_y^2$  are the sample variances of  $x$  and  $y$ , respectively, and  $\sigma_{xy}$  is the sample cross-covariance between  $x$  and  $y$ . The default values for  $C_1$  and  $C_2$  are 0.01 and 0.03. The average of the SSIM values across the image (mean SSIM or MSSIM) gives the final quality measure. The MSE, PSNR and SSIM performance comparison of various fusion methods is presented in Tab. 1, where one can see that the proposed approach performs better than the other conventional approaches by producing the best metric values.

### 4.2. Petrovic metric $Q^{AB/F}$

This measure was proposed by XYDEAS and PETRVIĆ [18, 19]. This metric is based on the assumption that the fusion algorithm that transfers input gradient information into a resultant image more accurately performs better. In this case, a pixel wise measure of information preservation is obtained between each input image ( $A$  and  $B$ ) and

Table 1. The MSE, PSNR and mean SSIM comparison of various image fusion approaches on reference images.

Lena			
Fusion method	MSE	PSNR [dB]	SSIM
DCT + average	39.11	32.20	0.9383
DCT + variance	8.70	38.73	0.9854
DCT + variance + CV	0.01	68.07	0.9999
DWT with DBSS(2,2)	5.62	40.63	0.9906
SIDWT with Haar	5.48	40.74	0.9899
Proposed (DCT + MaxAC)	0	$\infty$	1
Traffic			
Fusion method	MSE	PSNR [dB]	SSIM
DCT + average	112.6	27.62	0.8997
DCT + variance	4.2	41.89	0.9920
DCT + variance + CV	0.03	63.78	0.9998
DWT with DBSS(2,2)	12.81	37.05	0.9901
SIDWT with Haar	10.33	37.99	0.9905
Proposed (DCT + MaxAC)	0	$\infty$	1
Bridge			
Fusion method	MSE	PSNR [dB]	SSIM
DCT + average	28.83	33.53	0.9465
DCT + variance	6.69	39.87	0.9870
DCT + variance + CV	0.32	53.11	0.9990
DWT with DBSS(2,2)	4.49	41.61	0.9895
SIDWT with Haar	4.54	41.55	0.9893
Proposed (DCT + MaxAC)	0.03	62.91	0.9998

the fused image ( $F$ ) of size  $M \times N$  to compute  $Q^{AB/F}$  using simple local perceptual importance factors. It is calculated by:

$$Q^{AB/F} = \frac{\sum_{n=1}^N \sum_{m=1}^M \left[ Q^{AF}(n, m) w^A(n, m) + Q^{BF}(n, m) w^B(n, m) \right]}{\sum_{n=1}^N \sum_{m=1}^M \left[ w^A(n, m) + w^B(n, m) \right]} \quad (10)$$

where  $Q^{AF}$  and  $Q^{BF}$  are computed using edge information preservation values [2, 18];  $w^A(n, m)$  and  $w^B(n, m)$  are the weighted importance factors for  $Q^{AF}$  and  $Q^{BF}$ , respectively;  $Q^{AB/F}$  is in the range [0, 1] where 0 means complete loss of information and 1 means ideal fusion.

### 4.3. Spatial frequency (SF)

The row and column frequencies of the fused image  $F$  of size  $M \times N$  are computed [2, 17] as:

$$RF = \sqrt{\frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=1}^{N-1} [F(m, n) - F(m, n-1)]^2} \tag{11}$$

$$CF = \sqrt{\frac{1}{MN} \sum_{n=0}^{N-1} \sum_{m=1}^{M-1} [F(m, n) - F(m-1, n)]^2} \tag{12}$$

Then the total spatial frequency of the fused image which is based on edge information is computed as follows:

$$SF = \sqrt{RF^2 + CF^2} \tag{13}$$

Higher the spatial frequency, higher is the clarity of the image.

#### 4.4. Mutual information (MI)

This metric gives the amount of information that the fused image  $F$  has from input source images ( $A$  and  $B$ ) [20]. The mutual information between the source image  $A$  and the fused image  $F$  is computed as:

$$I_{AF} = \sum_{a, f} p_{AF}(a, f) \log \left[ \frac{p_{AF}(a, f)}{p_A(a)p_F(f)} \right] \tag{14}$$

where  $p_{AF}$ ,  $p_A$  and  $p_B$  are computed by normalisation of the joint and marginal histograms of  $A$  and  $F$ . Similarly mutual information  $I_{BF}$  is computed between the source image  $B$  and the fused image. Then the mutual information between the source images ( $A, B$ ) and the fused image ( $F$ ) is given as follows:

$$MI = I_{AF} + I_{BF} \tag{15}$$

#### 4.5. Mean gradient (MG)

Mean gradient is used to measure the edge details contained in the gradient image [17]. MG of the fused image  $F$  of size  $M \times N$  will be high if the edges are well preserved in the fused image. MG of the fused image  $F$  is calculated as:

$$MG = \frac{1}{(M-1)(N-1)} \sum_{x=1}^{N-1} \sum_{y=1}^{M-1} \sqrt{\frac{[F(x, y) - F(x-1, y)]^2 + [F(x, y) - F(x, y-1)]^2}{2}} \tag{16}$$

#### 4.6. Feature mutual information (FMI)

The feature mutual information (FMI) metric calculates the amount of image feature information transferred from the source images to the fused image by means of MI. Since the gradient map contains information about the pixel neighbourhoods, edge

strength and directions and texture and contrast, the normalized values of the gradient magnitude feature images are used as marginal distributions [21]. The FMI metric is given by

$$FMI = I_{FA} + I_{FB} \tag{17}$$

where  $I_{FA}$  and  $I_{FB}$  are the amount of feature information [21], which the fused image  $F$  contains about the source images  $A$  and  $B$ . They are individually measured by means of MI using a gradient map.

The performance metric comparison is given in Tabs. 1 and 2 for referenced and non-referenced images. From Table 1, it is clear that the proposed method reproduces a very high quality fused image and in most cases, the fused image is exactly identical to the referenced image (MSE = 0). It performs better than all the other compared approaches.

Table 2. The  $Q^{AB/F}$ , spatial frequency (SF), mutual information (MI), mean gradient (MG) and feature mutual information (FMI) of various image fusion approaches on non-referenced images.

Clock					
Fusion method	$Q^{AB/F}$	SF	MI	FMI	MG
DCT + average	0.65	11.56	7.13	0.7970	5.02
DCT + variance	0.73	18.39	9.02	0.8660	7.45
DCT + variance + CV	0.74	18.40	9.08	0.8750	7.32
DWT with DBSS(2,2)	0.67	18.78	6.50	0.8139	7.37
SIDWT with Haar	0.71	17.24	6.75	0.8567	7.35
Proposed (DCT + MaxAC)	0.73	18.46	9.03	0.8690	7.43
Pepsi					
Fusion method	$Q^{AB/F}$	SF	MI	FMI	MG
DCT + average	0.63	10.57	6.84	0.7314	3.6
DCT + variance	0.76	13.90	8.31	0.8599	5.38
DCT + variance + CV	0.78	13.91	8.67	0.8675	5.49
DWT with DBSS(2,2)	0.73	14.18	6.35	0.8207	5.45
SIDWT with Haar	0.74	13.39	6.60	0.8431	5.40
Proposed (DCT + MaxAC)	0.77	13.96	8.39	0.8633	5.46
Lab					
Fusion method	$Q^{AB/F}$	SF	MI	FMI	MG
DCT + average	0.55	7.72	7.08	0.6912	2.72
DCT + variance	0.73	13.17	8.49	0.8336	4.38
DCT + variance + CV	0.74	13.11	8.82	0.8452	4.52
DWT with DBSS(2,2)	0.66	13.10	6.52	0.7716	4.37
SIDWT with Haar	0.68	12.38	6.93	0.8146	4.36
Proposed (DCT + MaxAC)	0.74	13.41	8.65	0.8385	4.52

The performance metric comparison for fusion of non-referenced images is given in Table 2 and from the table it is inferred that the proposed approach performs slightly lower than that of the DCT + variance + CV method but better than the other four approaches by producing a good objective performance. The performance of the proposed method can be further enhanced by performing the consistency verification (CV) check as proposed in [13]. But since the focus is mainly on low complexity image fusion explicitly designed for low power resource constrained sensor nodes, CV is not performed in the proposed method.

## 5. Energy consumption analysis

In the proposed scheme, the key step is to fuse the DCT representations of multifocus images into a single fused image based on the AC coefficients values. The computation cost includes the discrete cosine transform and the fusion rule. Since the fusion rule does not involve any complex arithmetic floating point operations like mean or variance calculations, it is extremely simple, fast and efficient and hence is suitable for real-time applications. Also the transform using binary DCT (BinDCT) [8] can be used instead of floating point standard DCT which will further reduce the computation complexity and make it more appropriate for resource constrained image sensor nodes and low powered devices for energy efficient fusion and subsequent compression.

Hence the input images are divided into blocks of size  $8 \times 8$ , and the lifting scheme-based multiplierless approximation of BinDCT based on Chen's factorization [7] is applied for each block instead of the floating point DCT. BinDCT is fast multiplierless approximation of the DCT with the lifting scheme. The BinDCT has a fast, elegant implementation of the forward and inverse transforms utilizing only shift-and-add operations. The multiplierless property of the BinDCT allows efficient VLSI implementations in terms of both chip area and power consumption. The BinDCT has reasonably high coding performances. For fusion the MaxAC rule is applied on the BinDCT AC coefficients. The fusion results of applying the Bin.DCT + MaxAC rule are compared with those of the floating point DCT + MaxAC rule in Table 3.

From Table 3 it is understood that BinDCT + MaxAC fusion results are slightly different from that of DCT + MaxAC. But when the computational energy is concerned, the BinDCT consumes significantly very less energy when compared to that of conventional DCT.

For energy consumption analysis, the ATmega128 processor of Mica 2 mote at 8 MHz with an active power consumption of 22 mW is used as the target platform [22]. Compilation is performed via WinAVR with "-O3" optimization setting. The computation cycles and energy consumption for DCT, DWT and BinDCT on an  $8 \times 8$  image block are given in Tab. 4. The computation cycles and the energy needed for fusing two  $8 \times 8$  image blocks using the fusion schemes in DCT domain are given in Tab. 5. From Tab. 5 it is evident that the proposed method is extremely fast, highly

T a b l e 3. Fusion performance metrics for the proposed fusion method with DCT and binary DCT.

Images	Fusion method	Fusion performance metrics				
		MSE	PSNR [dB]	SSIM	$Q^{AB/F}$	MI
Lena	DCT + MaxAC	0	$\infty$	1	0.77	14.04
	BinDCT + MaxAC	0.0037	72.43	1	0.77	14.04
Traffic	DCT + MaxAC	0	$\infty$	1	0.76	25.29
	BinDCT + MaxAC	0	$\infty$	1	0.76	25.29
Bridge	DCT + MaxAC	0.0333	62.91	0.9998	0.74	11.50
	BinDCT + MaxAC	0.0340	62.81	0.9998	0.74	11.50
Barbara	DCT + MaxAC	0	$\infty$	1	0.79	29.46
	BinDCT + MaxAC	0	$\infty$	1	0.79	29.46
Bird	DCT + MaxAC	0.1065	57.86	0.9998	0.75	11.05
	BinDCT + MaxAC	0.0625	60.17	0.9998	0.75	11.05

T a b l e 4. Energy consumption for various transforms on an 8×8 block in ATmega128.

Transform	Transform cycles for an 8×8 block	Execution time [ms]	Energy [ $\mu$ J]
DCT	580106	72.513	1595.29
DWT	1428057	178.51	3927.15
BinDCT	4549	0.568	12.51

T a b l e 5. Energy consumption for fusing two 8×8 blocks in DCT domain.

Fusion scheme	Transform	Transform cycles for two 8×8 blocks	Fusion cycles	Execution time [ms]	Energy [ $\mu$ J]
DCT + average	DCT (float)	1160212	118692	159.86	3516.98
DCT + variance	DCT (float)	1160212	477768	204.75	4504.44
Proposed	BinDCT	9098	4397	1.69	37.11

energy efficient and consumes only around 1% of energy needed by other DCT based fusion methods.

## 6. Conclusions

In this paper, a very simple, fast and efficient DCT based multifocus image fusion scheme is proposed which outperforms other DCT based fusion methods as verified in our extensive experiments. Since the fusion rule does not involve any complex arithmetic floating point operations like mean or variance calculations, it is extremely simple and energy efficient, making it more suitable for real time applications and resource constrained battery powered sensors for energy efficient fusion and subsequent compression.

In future it is planned to validate our approach on a sensor network tested.

## References

- [1] DRAJIC D., CVEJIC N., *Adaptive fusion of multimodal surveillance image sequences in visual sensor networks*, IEEE Transactions on Consumer Electronics **53**(4), 2007, pp. 1456–1462.
- [2] SHUTAO LI, BIN YANG, *Multifocus image fusion using region segmentation and spatial frequency*, Image and Vision Computing **26**(7), 2008, pp. 971–979.
- [3] JING TIAN, LI CHEN, *Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure*, Signal Processing **92**(9), 2012, pp. 2137–2146.
- [4] MOHAMMAD BAGHER AKBARI HAGHIGHAT, ALI AGHAGOLZADEH, HADI SEYEDARABI, *Multi-focus image fusion for visual sensor networks in DCT domain*, Computers and Electrical Engineering **37**(5), 2011, pp. 789–797.
- [5] JINSHAN TANG, *A contrast based image fusion technique in the DCT domain*, Digital Signal Processing **14**(3), 2004, pp. 218–226.
- [6] HOSSNY M., NAHAVANDI S., CREIGHTON D., BHATTI A., *Towards autonomous image fusion*, [In] *11th International Conference on Control Automation Robotics and Vision (ICARCV)*, 2010, pp. 1748–1754.
- [7] JIE LIANG, TRAN T.D., *Fast multiplierless approximations of the DCT with the lifting scheme*, IEEE Transactions on Signal Processing **49**(12), 2001, pp. 3032–3044.
- [8] LECUIRE V., MAKKAOUI L., MOUREAUX J.M., *Fast zonal DCT for energy conservation in wireless image sensor networks*, Electronics Letters **48**(2), 2012, pp. 125–127.
- [9] <http://decsai.ugr.es/cvg/CG/base.htm>
- [10] <http://links.uwaterloo.ca/Repository.html>
- [11] <http://www.imagefusion.org>
- [12] <http://www.mathworks.com/matlabcentral/fileexchange/40861>
- [13] LI H., MANJUNATH B.S., MITRA S.K., *Multisensor image fusion using the wavelet transform*, Graphical Models and Image Processing **57**(3), 1995, pp. 235–245.
- [14] ROCKINGER O., *Image sequence fusion using a shift-invariant wavelet transform*, [In] *International Conference on Image Processing – Proceedings*, Vol. 3, 1997, pp. 288–291.
- [15] <http://www.metapix.de/toolbox.htm>
- [16] ZHOU WANG, BOVIK A.C., SHEIKH H.R., SIMONCELLI E.P., *Image quality assessment: from error visibility to structural similarity*, IEEE Transactions on Image Processing **13**(4), 2004, pp. 600–612.
- [17] XIANGZHI BAI, FUGEN ZHOU, BINDANG XUE, *Edge preserved image fusion based on multiscale toggle contrast operator*, Image and Vision Computing **29**(12), 2011, pp. 829–839.

- [18] XYDEAS C.S., PETROVIC V., *Objective image fusion performance measure*, Electronics Letters **36**(4), 2000, pp. 308–309.
- [19] PETROVIC V., COOTES T., *Objectively optimised multisensor image fusion*, [In] *Proceedings of 9th International Conference on Information Fusion*, 2006, pp. 1–7.
- [20] GUIHONG QU, DALI ZHANG, PINGFAN YAN, *Information measure for performance of image fusion*, Electronics Letters **38**(7), 2002, pp. 313–315.
- [21] MOHAMMAD BAGHER AKBARI HAGHIGHAT, ALI AGHAGOLZADEH, HADI SEYEDARABI, *A non-reference image fusion metric based on mutual information of image features*, Computers and Electrical Engineering **37**(5), 2011, pp. 744–756.
- [22] DONG-U LEE, HYUNGJIN KIM, RAHIMI M., ESTRIN D., VILLASENOR J.D., *Energy-efficient image compression for resource-constrained platforms*, IEEE Transactions on Image Processing **18**(9), 2009, pp. 2100–2113.

*Received March 28, 2013  
in revised form July 6, 2013*