World Essays Journal / 1 (2): 77-82, 2016

Image fusion Using Invariant Moments and DCT Transform

M. Shahrezaee

Department of Mathematics, Faculty of Science, University of Imam Hussein (AS), Tehran, Iran

Abstract: The purpose of image fusion is to combine two or more images into a single image by retaining the important features. One of the most efficient methods in image fusion is Discrete cosine transform based methods of image fusion. Practically, these methods are time-saving. One of the main step in image fusion is to pixel by pixel detection of the regions with more accuracy. This can be performed by utilizing different features. invariant moments are a set of feature vectors which are invariant under shifting, rotation and scaling. This approach can be utilized to extract the global features for pattern recognition due to its discrimination power and robustness. In this paper, we proposed a new DCT based image fusion method based on combining the invariant moments and variance features. Experimental results verify the proposed method has a superior efficiency output quality in comparison with recent proposed techniques.
Keywords: Image fusion; DCT transform, Invariant moments, Variance

INTRODUCTION

Image fusion is the process of combining two or more relevant information from images into a single image (Haghighat et al., 2011). The achieved image will be more accurate descriptor of the scene toward any of the individual source images (Haghighat et al., 2011).

main purpose of image fusion is to reduce the amount of data in network transmissions to provide more suitable and informative result in both further computer processing and visual perception processes (Drajic and Cvejic, 2007). Other important applications of the fusion of images include, microscopic imaging, medical imaging, remote sensing, robotics and computer vision (Hill et al., 2002). There are some generic requirements to have a good image fusion (Rockinger, 1997):

All relevant information of the input imagery should be preserved in the composite output image.

Output fused image should have no artifact or inconsistency to preserve the human observer without distraction

The considered fusion technique should be invariant toward the shifting and should not depend on the location of an object in the input imagery

The output fused image should be temporal consistent and stable with the input sequences.

Nowadays, there are several methods to image fusion depending on whether the images are fused in the spatial domain or they are transformed into another domain, and their transforms fused. In the recent years, multiresolution approaches like image pyramids (Burt et al., 1993) and the 2d wavelet transform (Singh and Khare, 2014) have been employed for pixel-level image fusion. Generally, most of the spatial domain image fusion techniques are complicated and time consuming which makes them not to profitable in the practical applications.

Discrete Cosine Transform (DCT) is a popular image transform process which is especially used in the image compression applications. So, the DCT domain image fusion can be efficient and less time consuming (Naidu and Bindu, 2013).

Because the explained introduction, in this paper a DCT based image fusion technique is proposed to simplify the method for using in the real-time applications and also enhance its quality in the output image. In this study, the original image is divided into 8*8 blocks and then a combination of the variance and the mean value of the invariant moments are performed into these 8*8 DCT coefficients.

The proposed method's accuracy is illustrated in the experimental results and comparisons show the system's considerable improvement in the quality of the output image.

Block analysis in the DCT domain

There exists a great deal of DCT applications in the image and video compression (Richardson, 2002). The input image is in the range of 0:255, so the coding process begins with a *level shifting*. i.e. the input data should be shifted into the range -128:127 in order to reduce the amount of DC coefficient. Since, the data becomes distributed about zero. In order to apply DCT into the image, we first divide the input image into 8×8 blocks and then two-dimensional DCT is performed on each blocks. After that, DCT coefficients are quantized by using a definite quantization table in every block.

In the quantization process, small coefficients are quantized to zeros and therefore the quantization converted into a lossy process. Afterwards, the quantized coefficients of blocks get rearranged into a zigzag order and the high frequency coefficients with zero values are classified together at the end of the rearranged array.

Huffman encoder and run-level procedure are then employed to encode the coefficients. In this part, all of the compressed data blocks are decoded and then reconstructed using quantization table. Finally, by using an inverse DCT, the reconstructed coefficients are transformed back into the image.

Two-dimensional DCT transform on an N*N block of an input image x(m, n) can be illustrated as below:

$$d(\mathrm{k},\mathrm{l}) = \frac{2\alpha(\mathrm{k})\alpha(l)}{N} \times \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x(\mathrm{m},\mathrm{n}) \times \cos\left[\frac{(2\mathrm{m}+1)\pi\mathrm{k}}{2N}\right] \times \cos\left[\frac{(2n+1)\pi\mathrm{k}}{2N}\right] \quad (1)$$

$$\alpha(\mathrm{k}) = \begin{cases} \frac{1}{\sqrt{2}}, & if\ k=0 \\ 1, & O.W. \end{cases} \quad (2)$$

here *k* and *l*= 0, 1, …, *N*-1, *d* (*0, 0*) defines DC coefficient and *d* (*k, l*) describes the AC coefficients of the block. The inverse DCT can be also described as below:

$$x(m,n) = \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} \frac{2\alpha(\mathrm{k})\alpha(l)}{N} \times d(\mathrm{k},\mathrm{l}) \times \cos\left[\frac{(2\mathrm{m}+1)\pi\mathrm{k}}{2N}\right] \times \cos\left[\frac{(2n+1)\pi\mathrm{k}}{2N}\right] \quad (3)$$

Feature extraction

Image processing and computer vision are typical important fields of information science and technology. Several areas of mathematics have contributed to essential progress of these fields (Razmjooy et al., 2012). In image processing, feature extraction technique is utilized for extracting an informative and non-redundant property from the considered image. This technique is necessary in image fusion for achieving good results. In features extraction, some definite features should be defined. This process is performed in order to compare the input image blocks and to select the best one to illustrate in the output image. In this paper, we used two different types of these features: variance and invariant moments. A short definition of these features is illustrated in the following.

Variance

Variance of an N×N block in the spatial domain is described as below:

$$\sigma^2 = \frac{1}{N^2}\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x^2(m,n) - \mu^2 \quad (4)$$

The variance of the block from DCT coefficients can be considered as:

$$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x^2(m,n) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x(m,n).x(m,n) \quad (5)$$

$$= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x(m,n) \times \left(\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} \frac{2\alpha(\mathrm{k})\alpha(l)}{N} \times d(\mathrm{k},\mathrm{l}) \times \cos\left[\frac{(2\mathrm{m}+1)\pi\mathrm{k}}{2N}\right] \times \cos\left[\frac{(2n+1)\pi\mathrm{k}}{2N}\right] \right)$$

Geometric invariant moment

Geometric moment (GM) invariants were first introduced by Hu (1962). This technique was derived from the algebraic invariant theory. In this study, GM method is employed to extract image features. This technique makes the generated features as Rotation Scale Translation (RST)-invariant.

Two-dimensional moments of an N×N block that has gray function *f (x, y), (x, y* = 0, *..N* − 1*)* is given as:

$$m_{pq} = \sum_{x=0}^{N-1}\sum_{n=0}^{N-1} (x)^p(m,n).(y)^p(m,n) \quad (6)$$

Where *p* and *q*=0, 1, 2,… .

Geometric invariant moment of the block from DCT coefficients can be defined as below:

$$\sum_{m=0}^{N-1}\sum_{n=0}^{N-1}(x)^p(m,n)(y)^p(m,n) = \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x(m,n).x(m,n).y(m,n).y(m,n)$$
$$= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x(m,n)\times y(m,n)\times\left(\sum_{k=0}^{N-1}\sum_{l=0}^{N-1}\frac{2\alpha(\mathrm{k})\alpha(l)}{N}\times d(\mathrm{k},\mathrm{l})\times\cos\left[\frac{(2\,\mathrm{m}+1)\pi\,\mathrm{k}}{2N}\right]\times\cos\left[\frac{(2n+1)\pi\,\mathrm{k}}{2N}\right]\right)^2 \quad (7)$$

Generally, Hu moments describes seven values which are achieved by normalizing central moments through order three and are invariant to object scale, position and orientation.

Simulation Results

variance and GM based image fusion in DCT domain

As it is explained in the above, the main purpose in image fusion is to achieve a fused image with relatively more information than any of the source images. Each of these images has some informative feature which should be extracted for generating the fused image. In here, we employed variance and mean value of the geometric moments for extracting informative features. GMs and Variance value are usually considered as two suitable contrast measure applications in image processing. From the previous section, it was shown that the variance and GM value in DCT domain is easy. Since, variance and GM can be utilized as the activity level measure of the 8×8 blocks of the input images.

Fig. 1 illustrates the block diagram of the proposed image fusion technique. In this study, for simplicity, the main process includes two source images *A* and *B*. note that the proposed algorithm can be extended for more than two input images by repeating the result image with next images since the final fuse image get achieved.

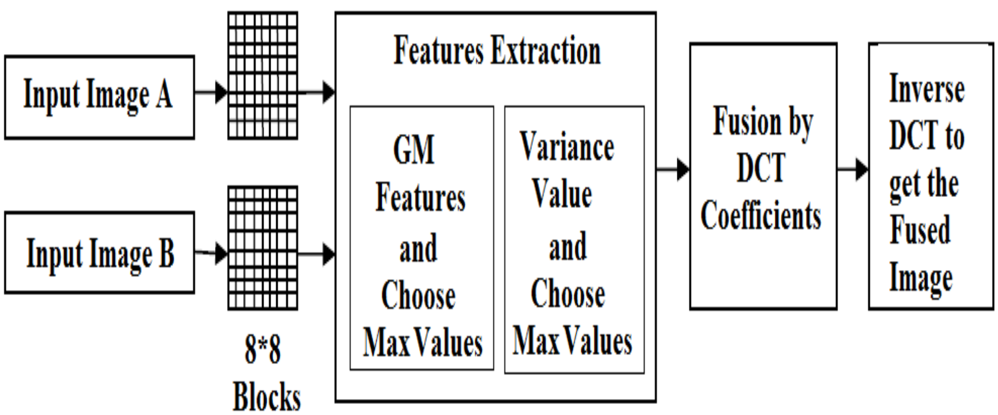


Figure 1. Block diagram of image fusion technique

From Fig. 1, it is obvious that after dividing the two input images into 8×8 pixel blocks and achieving the DCT coefficients for each block, the fusion algorithm is performed. This fusion technique is performed based on variance values and the mean value of the invariant moments of the corresponding blocks from input images. After this process, the highest activity level of the blocks is selected as the suitable block for the fused image. Finally, DCT coefficients of the fused image is achieved and the larger value of the variances and the mean value of the invariant moment is employed in the fused image.

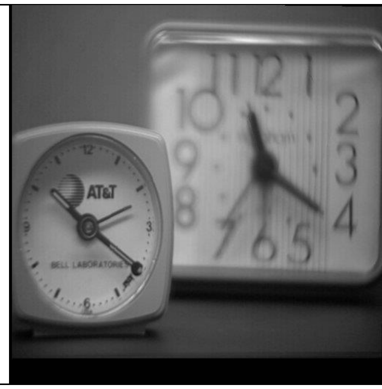
Quantitative Comparisons

In image processing, there are a great deal of metrics for comparison the methods (Zhang and Blum, 1999). The quantitative metric which is employed to analysis the proposed method for fusion purpose was taken from (Li et al., 1995):

$$\rho = \sqrt{\frac{\sum_{i=1}^{N}\sum_{j=1}^{N}[\mathrm{I}_{gt}(\mathrm{i},\mathrm{j}) - \mathrm{I}_{fd}(\mathrm{i},\mathrm{j})]^2}{N^2}} \quad (8)$$


AT&T
Ingraham

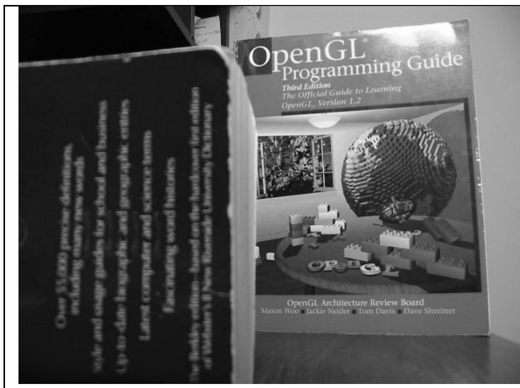
(A)


AT&T
BELL LABORATORIES

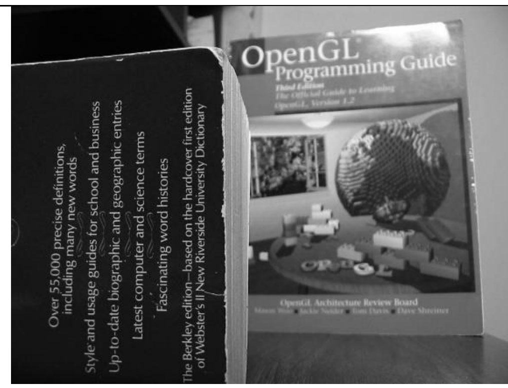
(B)


AT&T
BELL LABORATORIES

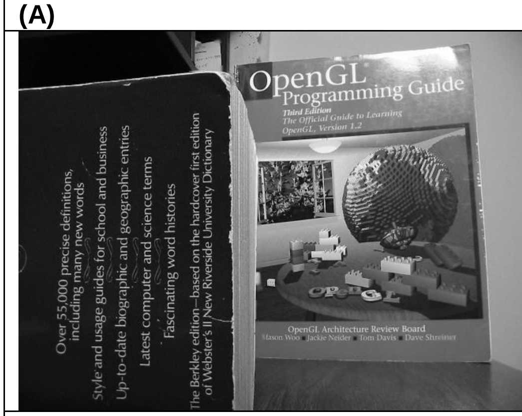
(C)


OpenGL
Programming Guide

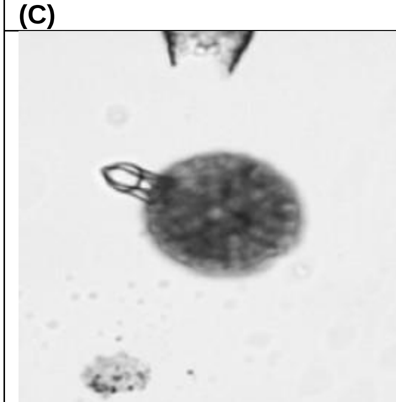
(A)


OpenGL
Programming Guide
Over 55,000 precise definitions, including many new words
Style and usage guides for school and business
Up-to-date biographic and geographic entries
Latest computer and science terms
Fascinating word histories
The Berkley edition—based on the hardcover first edition of Webster's II New Riverside University Dictionary

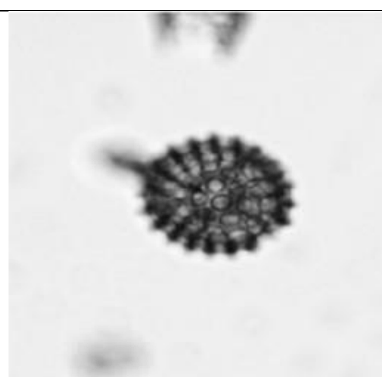
(B)


OpenGL
Programming Guide
Third Edition
The Official Guide to Learning OpenGL, Version 1.2
OpenGL Architecture Review Board
Over 55,000 precise definitions, including many new words
Style and usage guides for school and business
Up-to-date biographic and geographic entries
Latest computer and science terms
Fascinating word histories
The Berkley edition—based on the hardcover first edition of Webster's II New Riverside University Dictionary

(C)

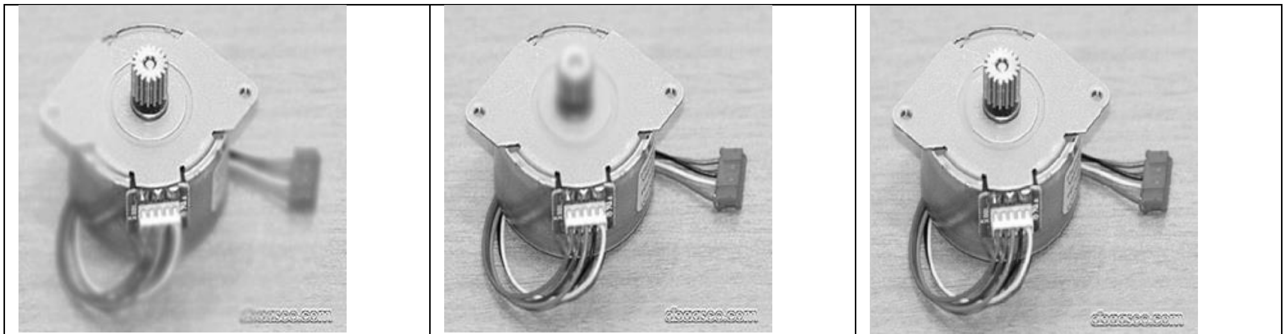


(A)



(B)

(C)



(A) **(B)** **(C)**

Figure 2. (A) the first input image, (B) the second input image and (C) the fused image

where I_{gt} defines the cut and paste *ground truth* image, I_{fd} describes the fused image, N defines the image size. Lower values of ρ illustrate greater similarity between the images I_{gt} and I_{fd} and hence more successful fusion in terms of quantitatively measurable similarity. The results of the proposed method and some other methods is shown in Table 1.

Table 1. Quantitative results for different fusion methods.

Fusion Method	ρ
Proposed Method	7.12
Simple DCT	7.23
PCA based	8.37

As it can be seen from the above results, the proposed method has the best quantitative value. This illustrates the importance of the proposed hybrid method in DCT transform fusion.

CONCLUSION

In this paper, a new DCT based image fusion method is proposed. The method is a combination of two different features based on the variance and the mean value of the invariant moments definitions in DCT domain. Results showed that this type of combination for extract specific features from the source images is so profitable for image fusion. To analysis the system efficiency, some experiments on evaluating the fusion performance are performed and the results show that the proposed method outperforms the ordinary DCT based method and also PCA method both in quality and complexity reduction.

REFERENCES

Burt, P. J.; Kolczynski, R. J.: Enhanced image capture through fusion, Proceedings of the 4th International Conference on Computer Vision, 1993, S. 173-182

Drajic, Dejan, and Nedeljko Cvejic. "Adaptive fusion of multimodal surveillance image sequences in visual sensor networks." IEEE Transactions on Consumer Electronics 53.4 (2007): 1456-1462.

H. Li, B.S. Manjunath, and S.K. Mitra. Multisensor image fusion using the wavelet transform. Graphical Models and Image Processing, 57:235–245, 1995.

Haghighat, Mohammad Bagher Akbari, Ali Aghagolzadeh, and Hadi Seyedarabi. "Multi-focus image fusion for visual sensor networks in DCT domain." *Computers & Electrical Engineering* 37.5 (2011): 789-797.

Haghighat, Mohammad Bagher Akbari, Ali Aghagolzadeh, and Hadi Seyedarabi. "A non-reference image fusion metric based on mutual information of image features." Computers & Electrical Engineering 37.5 (2011): 744-756.

Hill, Paul R., Cedric Nishan Canagarajah, and David R. Bull. "Image Fusion Using Complex Wavelets." In BMVC, pp. 1-10. 2002.

Hu, M.K., 1962. Visual pattern recognition by moments invariants. IRE Trans. Information Theory, 8: 179- 87.

Naidu, V. P. S., and Bindu Elias. "A novel image fusion technique using DCT based Laplacian pyramid." International Journal of Inventive Engineering and Sciences (IJIES) ISSN (2013): 2319-9598.

Razmjooy, Navid, B. Somayeh Mousavi, and Fazlollah Soleymani. "A real-time mathematical computer method for potato inspection using machine vision." Computers & Mathematics with Applications 63.1 (2012): 268-279.

Richardson IEG. Video codec design. John Wiley & Sons Ltd.; 2002.

Rockinger, Oliver. "Image sequence fusion using a shift-invariant wavelet transform." Image Processing, 1997. Proceedings., International Conference on. Vol. 3. IEEE, 1997.

Singh, Rajiv, and Ashish Khare. "Fusion of multimodal medical images using Daubechies complex wavelet transform–A multiresolution approach." Information Fusion 19 (2014): 49-60.

Z. Zhang and R. Blum. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proceedings of the IEEE, pages 1315–1328, August 1999.