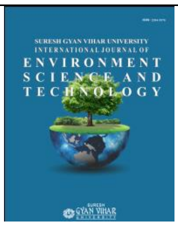

SGVU International Journal of Environment, Science and Technology

Journal homepage: https://www.gyanvihar.org/researchjournals/envirmental_science.php

E-ISSN: 2394-9570

Vol. 10 Issue 2 Page No

Review article

An Enhanced DCT-Based Technique for Image Fusion

Vipin Kumar Gupta*, Dr. Nagendra Kumar Swarnkar

Research Scholar, Professor

School of Engineering & Technology, Suresh Gyan Vihar University, Jaipur, India

Keywords

Multi-focus, Discrete Cosine Transform, Spatial frequency

Abstract

In Wireless Sensor Networks (WVSN), multi-focus image fusion (MFIF) is the procedure of integrating multiple Images in order to generate a new composite image which reflects the scene more precisely and more informative than the individual source images might have. This paper offers an approach that uses standards derived from the Discrete Cosine Transform (DCT) to efficiently fuse the images or movies within WVSN. By evaluating their frequency distributions, the segment blocks in the source images' contrast requirements are ascertained. Blocks containing greater spatial frequencies constitute the output image's DCT representation as an outcome. To assess the fusion performance, we carried out experiments on numerous combinations of true colour images encoded in the JPEG standard. The outcomes show that our fusion technique improves the entropy value as desired.

Introduction

Image fusion is a technique that involves combining pixel value information from two or more images of the identical scene to generate a single, more informative image. This technique is widely used in various applications such as medical imaging,remote sensing, surveillance, and computer vision. The objective of fusion process is to enhance the overall quality of the resulting image by preserving relevant information from each input image while minimizing redundancy and noise.

One popular transform domain for image fusion is the Discrete Cosine Transform (DCT). DCT-based fusion techniques exploit the frequency information of the input images to achieve better fusion results. Other transform domain methods include the Discrete Wavelet Transform (DWT), the Shift-Invariant DWT (SIDWT), and the Non-Subsampled Contourlet Transform (NSCT). These methods offer different advantages and are suitable for different types of images and applications.

The performance of an image fusion technique can be evaluated using various metrics, including subjective evaluation by human observers and objective evaluation using quantitative measures such as entropy and mutual information (MI). Selection of evaluation metrics depends on the requirements of the application and the statistical feature of the input images.

In recent years, there has been significant research interest in developing novel image fusion algorithms that can handle various types of input images and achieve superior fusion performance. These algorithms often incorporate advanced techniques from machine learning, optimization, and signal processing to improve fusion quality and efficiency.

Overall, image fusion is a fundamental process in image processing and computer vision, with numerous applications and ongoing research efforts aimed at advancing the state-of-the-art techniques for better fusion results in various domain. Techniques such as averaging, weighted averaging, and Laplacian pyramid fusion are commonly used in this domain. Transform domain methods, on the other hand, involve transforming the input images into a different domain (such as frequency or wavelet domain) before performing fusion. This allows for more effective separation and combination of image features.

The presented technique DCT, marks a notable progression in the field of image fusion, specifically leveraging the Discrete Cosine Transform (DCT). This approach aims to refine the fusion process, yielding outcomes that surpass those of existing methodologies. The algorithm adeptly integrates information from multiple images, resulting in an improved output that offers a more precise representation of the scene. Through experimental evaluations involving diverse image pairs encoded in the Joint Photographic Experts Group (JPEG) standard, our enhanced DCT-based image fusion method demonstrates superior performance.

The results not only showcase enhanced visual quality but also outperform prior DCT-based techniques and contemporary methodologies when subjected to objective evaluations. It is tricky to adequately convey complicated events with a single image as optical lenses have a narrow range of focus [1]. An individual scene has been recorded by numerous sensors in wireless visual sensor networks. The initial images from these sensors are subsequently assembled into a single image by a centralised fusion centre, strengthening its compatibility for machine and human vision [2]. The fused image that generates is then transmitted to a superior node. While the spatial domain of fusion of images has received a significant amount of the research's focus in earlier studies [3][4][5], multi-scale transform methods are currently gaining popularity. Examples of these techniques include the Discrete Wavelet Transform (DWT) [3], Shift Unaffected Discrete Wavelet Transform (SIDWT) [4], and Non-Subsampled Contourlet Transform (NSCT) [5].

Nevertheless, a lot of these methods are convoluted and require a lot of execution time which restricts their implementation in wireless network for visual sensors, limiting their application in wireless visual sensor networks with constrained resources. It is standard procedure in Wireless Visual Sensor Networks (WVSN) to compress images in advance of transmitting them to other nodes. Computational complexity might be greatly reduced by employing strategies established in the DCT sector when using DCT-based standards for storing or transferring source pictures [6]. Overall, image fusion is a fundamental process in image processing and computer vision, with numerous applications and ongoing research efforts aimed at advancing the state-of-the-art techniques for better fusion results in various domain

Recently, a number of image fusion approaches in the DCT region have been proposed. The DCT + Average and DCT + Contrast is two DCT domain approaches that were introduced by Tang et al. [7]. Nevertheless, the image quality has been adversely affected by these procedures' undesirable side effects, which include distorting or obstructing artefacts. Due to the amount of high valued AC coefficients is an insufficient criteria—especially after maximum of the coefficients are quantized to null value—the method suggested by [8], referred to as DCT+AC – Max, can end up resulting in errors when identifying the appropriate JPEG coded blocks. In [9], an additional strategy investigated variance as a criterion for contrast in fusion. On the other hand, variance performs less effectively than other focus metrics, according to experimental data [10].

This work proposes a universal picture fusion method in the DCT domain. High spatial frequency picture blocks are incorporated into the fused image in this case. The process of verification is conducted consistently to improve the output image quality. Experiments conducted on multiple JPEG-coded databases indicate that our strategy significantly raises the level of quality of the merged image. This is exactly the remaining part of the letter is planned: Section II presents the fundamental ideas of our algorithm. In Section III, the recommended method for image fusion is explained. The investigational results are analysed in Section IV, and conclusions are presented in Section V.

DCT analysis

The discrete cosine transform (DCT) stands as a pivotal transformation in image compression, finding extensive application [11]. Many prevalent commercial standards, including the JPEG which is still a coding standard [12], Motion-JPEG, MPEG, and the H263 video coding standards [13], rely on the DCT. Employing vector processing, the resulting matrix of a two-dimensional 8x8 DCT operation on a source matrix is as follows:

$$F = C.f.C^t \quad (1)$$

Here C is a matrix having orthogonal properties which contains cosine coefficients as elements of the matrix. It has the property that inverse of the matrix is the transpose coefficients.

$$C^{-1} = C^t \quad (2)$$

The inverse DCT (IDCT) is also expressed as:

$$f = C^t.F.C \quad (3)$$

According to [10],

$$trace(ff^t) = trace(FF^t) \quad (4)$$

Here trace(x) means sum of diagonal elements of matrix x. The Row Frequency (RF) and Column Frequency (CF) of a segment of size 8 x 8 block are given by:

$$RF^2 = \frac{1}{8\times 8}\sum_{a=0}^{7}\sum_{b=1}^{7}\big(f(a,b) - f(a,b-1)\big)^2 \quad (5)$$

$$CF^2 = \frac{1}{8\times 8}\sum_{a=1}^{7}\sum_{b=0}^{7}\big(f(a,b) - f(a-1,b)\big)^2 \quad (6)$$

Mathematically, Spatial Frequency (SF) of the block is computed as:

$$SF^2 = RF^2 + CF^2 \quad (7)$$

After performing a brief calculation, it becomes possible to determine the SF of the block based on the AC coefficients within the DCT domain. We represent Δx and Δy as the matrix with difference between rows and columns, respectively:

$$\Delta x = \begin{matrix} f(0,1) - f(0,0) & \dots & f(0,7) - f(0,6) \\ \vdots & \vdots & \vdots \\ f(7,1) - f(7,0) & \dots & f(7,7) - f(7,6) \end{matrix} \quad (8)$$

$$\Delta y = \begin{matrix} f(1,0) - f(0,0) & \dots & f(1,7) - f(0,7) \\ \vdots & \vdots & \vdots \\ f(7,0) - f(6,0) & \dots & f(7,7) - f(6,7) \end{matrix} \quad (9)$$

It’s clear to say that

$$\Delta x = fb = C^tFCC^tBC = C^tFBC \quad (10)$$

$$\Delta y = b^tf = C^tBCC^tFC = C^tB^tFC \quad (11)$$

$$
\text{where b} = \begin{matrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{matrix}
$$

where B is the DCT representation of b.

From(3),(10)and(11),it is possible to calculate FB and B^tFwhich are the DCT representations of Δx and Δy, respectively. Therefore, RF and CF are obtained as:

$$
RF^2 = \frac{1}{8\times 8}\sum_{x=0}^{7}\sum_{y=0}^{7}\Delta x^2(x,y) = \frac{1}{8\times 8}trace(\Delta x(\Delta x)^t)
$$

$$
= \frac{1}{8\times 8}trace(FB(FB)^t)
$$

$$
= \frac{1}{8\times 8}trace(FBB^tF^t) \qquad (12)
$$

$$
CF^2 = \frac{1}{8\times 8}\sum_{x=0}^{7}\sum_{y=0}^{7}\Delta y^2(x,y) \quad = \frac{1}{8\times 8}trace((\Delta y)^t\Delta y)
$$

$$
= \frac{1}{8\times 8}trace(FB(FB)^t)
$$

$$
= \frac{1}{8\times 8}trace(F^tBB^tF) \qquad (13)
$$

Let matrix D is obtained by multiplication of B and B^t. It is also possible to compute that D is a diagonal matrix shown in equation (15). Then, SF becomes:

$$
SF^2 = RF^2 + CF^2 = \frac{1}{8\times 8}[trace(DF^tF) + trace(DFF^t)]
$$

$$
= \frac{1}{8\times 8}\sum_{c=0,d=0}^{7}\big(D(c,c) + D(d,d)\big)\times F^2(c,d)
$$

$$
= \frac{1}{8\times 8}\sum_{c=0,d=0}^{7}[E(c,d)\times F^2(c,d)] \qquad (14)
$$

Concludingly, it can be said that SF of an 8 x 8 block of pixels can be efficiently computed by (14).

$$
D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3.4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3.9 \end{bmatrix} \qquad (15)
$$

Proposed Methodology

There are several methods for performing image fusion, which is broadly classified into spatial domain methods and transform domain methods. In spatial domain methods, fusion is carried out directly through computation process on pixel values of the input images. Techniques such as averaging, weighted averaging, and Laplacian pyramid fusion are commonly used in this domain. Transform domain methods, on the other hand, involve transforming the input images into a different domain (such as frequency or wavelet domain) before performing fusion. This allows for more effective separation and combination of image features.

The concept of spatial frequency, originating from research on the human visual system, provides a comprehensive indication of the global activity level within a matrix of an image [14]. While fully understanding the human visual system remains a challenge using current physiological methods, spatial frequency serves as a perfect criterion for fusion of images[15]. As detailed in Section II, calculating spatial domain frequency (SDF) in the domain of DCT is straightforward. Consequently, SDF value can be employed as aeffective measure for the 8x8 segments of the input images. To simplify, two input images, A and B, have been illustrated, though the method is extendable to accommodate more input images. The process comprises the steps which are as follows:

1) Begin by decoding, thereafter. de-quantizing the input images, followed by dividing them into segments of size 8x8. Refer to the block pair at location (a, b) as Aa,b and Ba,b, respectively.
2) Calculate the SDF of each block using equation (14), and represent the results for $A_{a,b}$ and $B_{a,b}$ as $SFA_{a,b}$ and $SFB_{a,b}$, respectively.
3) Compare the SDFs of the respective blocks to determine which must contribute to constructing the fused image. Establish a decision matrix, denoted as W, to document the outcomes of the feature comparisons based on a predetermined selection rule.

$$W_{a,b} = \begin{cases} 1 & SFA_{a,b} > SFB_{a,b} + U \\ -1 & SFA_{a,b} < SFB_{a,b} - U \\ 0 & therwise \end{cases} \tag{16}$$

Where, U is a user dependent constant.

4) Implement a verification process to check the consistency of the algorithm so as to enhance the quality of the fused image. Utilize a 3x3 filter [3] to derive aadvanced decision matrix denoted as R:

$$R_{x,y} = \sum_{x=a-1}^{a+1} \sum_{y=b-1}^{b+1} w_{x,y} \tag{17}$$

5) Subsequently, acquire the Discrete Cosine Transform (DCT) representation of the output image, with reference to the refined decision map R.

$$F_{a,b} = \begin{cases} A_{a,b} & R_{a,b} > 0 \\ B_{a,b} & R_{a,b} < 0 \\ \frac{(A_{a,b}+B_{a,b})}{2} & R_{a,b} = 0 \end{cases} \quad (18)$$

The execution of the MATLAB code for the fusion methods were done using an Intel i7-12700 processor with 16GB of RAM. All images utilized in the MATLAB simulations were reformatted to JPG files.

Experiments & Results

In this section, assessment of the effectiveness of the proposed method against established image fusion techniques operating in the DCT domain, including DCT + Average, DCT + Contrast, DCT + AC – Max, DCT + Variance, and DCT + Variance + CV. Additionally, we consider multi-scale fusion methods like Discrete Wavelet Transform (DWT), Shift Invariant DWT (SIDWT), and Non-Subsampled Contourlet Transform (NSCT) as state-of-the-art approaches.

Objective evaluation was performed using Entropy and MI quantitative analysis. The entropy values for the experimental images are tabulated, along with the runtime required for fusion using DCT-based methods.

The proposed fusion method, particularly without coefficient variation (CV), exhibits superior performance compared to various DCT-based along with DWT-based methods. Moreover, incorporating CV improves the performance even surpassing the NSCT-based algorithm, albeit with slightly increased complexity.Notably, observed fusion results reveal undesired blurring effects and blocking artifacts. Additionally, certain methods like DCT+AC-Max and DWT exhibit distinct errors in block selection and undesirable artifacts, respectively.

The images used are shown in Fig. 1(a), Fig. 1(b) and Fig. 1(c).

Fig. 1(a)

Fig. 1(b)

Fig. 1(c)

Table 1. Entropy Values of proposed method

Image 1	Image 2	Image 3
7.1585	7.4042	7.4910

Further performance comparison is conducted using metrics such as Localized Mutual Information (LMI), Piella metric (QW), and Feature Mutual Information (FMI). These quantitative measures gauge the transfer of local regions into the output image, with higher values indicating better quality. Performance analyses for images "Ball.jpg" and "wpeppers.jpg" are tabulated in Table 1, demonstrating the superiority of the proposed technique over conventional methods across these metrics. In Table 1., Image 1 denotes ball.jpg, image 2 denotes wpeppers.jpg and image 3 denotes fused image.

Conclusions

The paper introduces a novel approach for fusing multi-focus images utilizing spatial frequency within the Discrete Cosine Transform (DCT) domain, departing from conventional spatial domain methods. Through rigorous evaluation employing various metrics, our method demonstrates superior fusion performance in the DCT domain compared to established techniques relying on DCT and state-of-the-art methods such as DWT, SIDWT, and NSCT. It can be seen in Table 1. that the proposed method provides image with better entropy value than the input images. This superiority is evident in both quantitative parameters andvisual quality. Additionally, the method proposed boasts simplicity in implementation and computational efficiency, particularly advantageous when dealing with source images encoded in JPEG format, notably beneficial in contexts like wireless visual sensor network.

REFERENCES

Djamel, S., & Mouldi, B. (2006). Image compression via embedded coder in the transform domain. Asian Journal of Information Technology, 5(6), 633–639.

Haghighat, M. B. A., Aghagolzadeh, A., & Seyedarabi, H. (2010). Real-time fusion of multi-focus images for visual sensor networks. In 6th IEEE Iranian Machine Vision and Image Processing (MVIP).

Haghighat, M. B. A., Aghagolzadeh, A., & Seyedarabi, H. (2011). A non-reference image fusion metric based on mutual information of image features. Computers & Electrical Engineering, 37(5), 744–756.

Haghighat, M. B. A., Aghagolzadeh, A., & Seyedarabi, H. (2011). Multi-focus image fusion for visual sensor networks in the DCT domain. Computers & Electrical Engineering, 37(5), 789–797.

Hossny, M., Nahavandi, S., Creighton, D., & Karmakar, C. (2010). Image fusion performance metric based on mutual information and entropy-driven quadtree decomposition. Electronics Letters, 46(18), 1266–1268.

Huang, W., & Jing, Z. (2007). Evaluation of focus measures in multi-focus image fusion. Pattern Recognition Letters, 28(4), 493–500.

Image Fusion. (n.d.). Retrieved from http://imagefusion.org/

Jinshan, T. (2004). A contrast-based image fusion technique in the DCT domain. Digital Signal Processing, 14(3), 218–226.

Li, H., Manjunath, B. S., & Mitra, S. K. (1995). Multisensor image fusion using the wavelet transform. Graphical Models and Image Processing, 57(3), 235–245.

Li, S., & Yang, B. (2008). Multifocus image fusion using region segmentation and spatial frequency. Image and Vision Computing, 26(7), 971–979.

Li, S., Kwok, J. T., & Wang, Y. (2001). Combination of images with diverse focuses using the spatial frequency. Information Fusion, 2(3), 169–176.

Nirmala, D. E., Vignesh, R. K., & Vaidehi, V. (2013). Multimodal image fusion in visual sensor networks. In 2013 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) (pp. 1–5). IEEE.

Phamila, Y., & Amutha, R. (1995). Discrete cosine transform-based fusion of multi-focus images for visual sensor networks. Signal Processing, 161–170.

Piella, G., & Heijmans, H. (2003). A new quality metric for image fusion. In Proceedings of the 2003 International Conference on Image Processing (ICIP), Vol. 3 (pp. 173–176). IEEE.

Rao, D. S., Babu, C. R., Kumar, N., Rajasekhar, N., & Ravi, T. (2019). Medical image fusion using transform-based fusion techniques. International Journal of Recent Technology and Engineering (IJRTE), 8(2).

Richardson, I. E. (2002). Video codec design. Wiley.

Rockinger, O. (1997). Image sequence fusion using a shift-invariant wavelet transform. In Proceedings of the 1997 International Conference on Image Processing (ICIP), Vol. 3 (pp. 288–291). IEEE.

Shalima, & Virk, R. (2015). Review of image fusion techniques. International Research Journal of Engineering and Technology (IRJET), 2(3), 333–339.

Shreyamsha Kumar, B. K., Swamy, M. N. S., & Ahmad, M. O. (2013). Multiresolution DCT decomposition for multifocus image fusion. In 2013 26th Annual IEEE Canadian Conference on Electrical and Computer Engineering (CCECE) (pp. 1–4). IEEE.

Wallace, G. K. (1991). The JPEG still picture compression standard. Communications of the ACM, 34(4), 30–44.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600–612.

Zhang, Q., & Guo, B.-L. (2009). Multifocus image fusion using the non-subsampled contourlet transform. Signal Processing, 89(7), 1334–1346.