

# **Image Fusion Using Discrete Cosine Transform (DCT) AND Singular Value Decomposition(SVD) of Two Multifocused Images**



**Indian Institute of Technology (ISM)  
Dhanbad**

**AKSHAY DOBARIYA**

**16JE001934**

**B. tech Computer Science And Engineering**

**Guided By: Dr. Arup Kr. Pal**

# Abstract

Multi-focus image fusion is a process that fuses several images from a scene with different focal lengths into a whole image in which all areas are focused on.

In this project, a low complexity multi-focus image fusion in DCT domain is presented which increases the output image quality. Proposed method makes it suitable for real-time applications because of its implementation in DCT domain. On the other hand, it is stable in noisy conditions. The proposed method uses the singular values of Singular Value Decomposition (SVD) of  $8 \times 8$  input blocks in DCT domain. The geometric mean of the 5 largest singular values (out of 8 singular values) is computed as a criterion of focused block detection. The blocks which have the highest geometric mean value among other corresponding blocks is selected as the focused block. These blocks are then used for constructing the output image. This method can be utilized both in DCT domain and in spatial domain. Various experiments and comparisons between the proposed method and the previous methods in noisy and noiseless conditions have been presented, which confirm the increase in image quality and stability in noisy images.

## INTRODUCTION

The image fusion process is defined as gathering all the important information from multiple images, and their inclusion into fewer images, usually a single one. This single image is more informative and accurate than any single source image, and it consists of all the necessary information. The purpose of image fusion is not only to reduce the amount of data but also to construct images that are more appropriate and understandable for the human and machine perception.

Various multi-focus image fusion researches have been done in the spatial domain. These methods are time and energy consuming, making them unsuitable for real-time applications. Multi-scale transforms include gradient, Laplacian, and morphological pyramid transform, Discrete Wavelet Transform (DWT), Shift-Invariant Wavelet Transform (SIDWT) and Discrete Cosine Harmonic Wavelet Transform (DCHWT). These create complexity in computation, making it not only unsuitable considering time consumption, but also inappropriate because of the ringing artefacts. Ringing artefacts at the edge of the images decrease the

quality of the output image. Recently, several methods have been proposed for fusing images by the properties and attributes of Singular Value Decomposition (SVD) such as Multiresolution Singular Value Decomposition (MSVD) and Higher Order Singular Value Decomposition (HOSVD).

**Spatial domain:** It refers to the image plane itself and methods in spatial domain are based on directly modifying the value of the pixels.

**Transform domain:** In this method first image is converted into some vectors like (fft, dct etc) then the fusion took place and then after taking inverse of transform gives a fused image.

## Multifocused Image Fusion

**Multifocus Image fusion** is process of combining information of two or more images of a scene and as a result has "*all-in-focus*" image.

When one scene contains objects in different distance, the camera can be focused on each object one after the other, creating set of pictures. Then, using image fusion technique, an image with better focus accross all area can be generated. There are many multifocus image fusion methods, today. One of them is Empirical Mode Decomposition based multifocus image fusion.

Feature extraction is a key procedure in image fusion which is usually employed to reduce the complexity of calculation, time and energy consumption. Therefore, scientists presented significant ways of implementing image fusion algorithms in Discrete Cosine Transform (DCT) domain for real-time applications; In other words, image fusion methods in DCT domain for images which are compressed in

JPEG format take less time and are more efficient . After dividing input images into smaller blocks (e.g. 8×8), DCT based methods utilize appropriate criteria to select focused blocks. Subsequently, the output fused image is merged using these selected blocks. Tang introduces DCT+Average and DCT+Contrast methods for multi-focus image fusion . In DCT+Average, the output image is computed by taking the mean of all DCT coefficients. Likewise, DCT+Contrast chooses the maximum AC value of DCT representation block coefficients for creating the output image. Throughout this method, the undesirable side effects such as blurring and blocking artifact have been noticed in the fused images.

Haghighat et al in the DCT+Variance method have calculated the variance for every block in DCT domain. Thus, the block with the maximum variance in DCT domain, has been chosen as the focused block. In DCT+Ac\_Max method, the block which has more quantity of the maximum value of AC coefficients in DCT domain has been chosen as the block for the output image. Considering Spatial Frequency criterion, DCT+SF has been introduced by Cao et al. In this method, Spatial Frequency has been calculated in DCT domain, while the block with the maximum SF value has been chosen for the output image. the focus criterion Sum-Modified-Laplacian (SML) is calculated in DCT domain. So the blocks with higher SML values are selected for the output image. On the other hand, the Consistency Verification (CV) process is considered in the main algorithm of DCT+SML. Whereas CV is a post-processing step and it should not be considered completely with DCT+SML method because this is DCT+SML+CV. So there is a lot of unsuitable block selection in DCT+SML. Due to their selective criteria, these methods (DCT+Variance, DCT+Ac-max, DCT+SF, and DCT+SML) create abundant errors during suitable block selection. Hence, these methods face quality loss and blocking artifacts. Amin-Naji and Aghagolzadeh introduced new methods for multi-focus image fusion in DCT domain. In DCT+Corr method, the images are passed through a low-pass filter to make artificial blurred images. Therefore the correlation coefficient between source blocks and the artificially blurred blocks used as a focus criterion. In DCT+Sharp method, the blocks are passed through the unsharp filter in order to make bigger distance of variances' values. In addition, two arbitrary thresholds have been considered to identify blocks suspicious of close variance values. Thus, this increases the image quality and decreases blocking artifacts in the output images. Although having minor errors throughout selecting suitable blocks, these methods (DCT+Corr and DCT+Sharp), have had great improvements in comparison to the previous DCT methods.

## PRELIMINARY MATHEMATICS

**Discrete Cosine Transform using vector processing** Two-dimensional DCT transform of  $N \times N$  blocks of the image  $a(m,n)$  are represented as.

$$A = C \cdot a \cdot C^T, \quad (1)$$

where the orthogonal matrix  $C$  consists of coefficients of the discrete cosine transform and superscripts  $T$  stands for the transpose operator. On the other hand,  $A$  represents the DCT coefficients for image's matrix of  $a$ . Furthermore  $C$ , is known as a Hermitian matrix, i.e.:

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

The DCT inverse of  $A$  is therefore defined as (3).

### B. Singular Value Decomposition (SVD)

In linear algebra, the singular-value decomposition (SVD) is a factorization of a real or complex matrix. It is the generalization of the Eigen decomposition of a positive semidefinite normal matrix (for example, a symmetric matrix with positive eigenvalues) to any  $(m,n)$  matrix via an extension of the polar decomposition. It has many useful applications in signal processing and statistics. Formally, the singular value decomposition of an  $(m,n)$  real or complex matrix  $\mathbf{M}$  is a factorization of the form  $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ , where  $\mathbf{U}$  is an  $(m,m)$  real or complex unitary matrix,  $\mathbf{\Sigma}$  is a  $(m,n)$  rectangular diagonal matrix with non-negative real numbers on the diagonal, and  $\mathbf{V}$  is an  $(n,n)$  real or complex unitary matrix. The diagonal entries of  $\mathbf{\Sigma}$  are known as the **singular values** of  $\mathbf{M}$ . The columns of  $\mathbf{U}$  and the columns of  $\mathbf{V}$  are called the **left-singular vectors** and **right-singular vectors** of  $\mathbf{M}$ , respectively.

The singular-value decomposition can be computed using the following observations:

- The left-singular vectors of  $\mathbf{M}$  are a set of orthonormal eigenvectors of  $\mathbf{M}\mathbf{M}^*$ .
- The right-singular vectors of  $\mathbf{M}$  are a set of orthonormal eigenvectors of  $\mathbf{M}^*\mathbf{M}$ .
- The non-zero singular values of  $\mathbf{M}$  (found on the diagonal entries of  $\mathbf{\Sigma}$ ) are the square roots of the non-zero eigenvalues of both  $\mathbf{M}^*\mathbf{M}$  and  $\mathbf{M}\mathbf{M}^*$

$$\mathbf{a} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (4)$$

$$\mathbf{a} = \sigma_1 u_1 V_1^T + \sigma_2 u_2 V_2^T + \dots + \sigma_n u_n V_n^T . \quad (5)$$

$$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_8. \quad (6)$$

$$\det(\lambda \mathbf{I} - \mathbf{a}\mathbf{a}^T) = 0, \quad (7)$$

which the eigenvalues are  $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_8$ .

Therefore:

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\lambda_3} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \sqrt{\lambda_8} \end{bmatrix}_{8 \times 8} . \quad (8)$$

## PROPOSED METHOD

The general structure of the proposed methods is depicted in Fig.1. In order to simplify the description of the proposed method, a fusion of two multi-focus images is considered. However, the proposed algorithm can easily be extended to more than two images. In addition, it is assumed that multi-focus images are already registered. In multi-focus images, areas which are in the focal length and are clear in one image, may be vague in another. This focused area has more contrast and clearness compared to the defocused area. Furthermore, it is expected that these focused areas have larger singular values  $\sigma_i$ ; therefore the largest value of  $\sigma_i$  is selected as a criterion of measurement which represents the amount of focus is applied there. The proposed method starts by dividing the input images into  $8 \times 8$  blocks similar to the encoding of JPEG image. Each block is transferred to the DCT domain by (1). By computing Singular Value Decomposition (SVD) in DCT domain, focused blocks in DCT domain are recognized.

### A. Singular Value Decomposition in DCT domain

The previous discussions are applicable for the spatial domain. By transforming the discussion into DCT domain and comparing the SVD matrix, it is concluded that matrix  $\mathbf{\Sigma}$  in (4) has the exact data in both spatial and DCT domains; so

$$\mathbf{b}_{8 \times 8} = \mathbf{U}\mathbf{\Sigma}_{8 \times 8}\mathbf{V}^T \quad (\text{Spatial domain}), \quad (9)$$

$$\mathbf{B}_{8 \times 8} = \mathbf{U}'\mathbf{\Sigma}'_{8 \times 8}\mathbf{V}'^T \quad (\text{DCT domain}). \quad (10)$$

All parameters are different; however,  $\Sigma$  and  $\Sigma'$  are identical. In order to prove this case:

**Remark:**

Defining:

$$b = U \Sigma V^T \quad (11)$$

$$B = U' \Sigma' V'^T \quad (12)$$

where:

$$B = C b C^T \quad (13)$$

$$U' = C U C^T \quad (14)$$

$$\Sigma' = C \Sigma C^T \quad (15)$$

$$V'^T = C V^T C^T, \quad (16)$$

**Proof:**

then:

$$B = C b C^T = U' \Sigma' V'^T = C U C^T C \Sigma C^T C V^T C^T \quad (17)$$

where  $C^T C = I$ , then

$$B = C U \Sigma V^T C^T. \quad (18)$$

It can be observed that  $C U = U'$  and  $V^T C^T = V'^T$ , hence:

$$B = U' \Sigma V'^T \quad (19)$$

whereas it was known that

$$B = U' \Sigma' V'^T. \quad (20)$$

Therefore

$$\Sigma' = \Sigma. \blacksquare \quad (21)$$

In the equations shown above, it has been proven that  $\Sigma'$  (DCT Domain) and  $\Sigma$  (Spatial Domain) matrices are exactly equal. To put it in other words, eigenvalues of  $b b^T$  and  $B B^T$  are the same.

## B. Image fusion Based on SVD in DCT domain

By computing SVD and eigenvalues of  $B B^T$  for 8x8 blocks in DCT, singular values  $\sigma_i$  ( $\sigma_1, \dots, \sigma_8$ ) are extracted for each block. The lower indices are for low frequencies and the higher indices are for high frequencies. Singular values  $\sigma_6, \sigma_7$ , and  $\sigma_8$  have smaller values, near to zero, which block noise has great impact on changing their values. Therefore, in order to prevent selecting unsuitable blocks, it is better to use of  $\sigma_1, \dots, \sigma_5$  only. Geometric mean of 5 largest singular values is chosen as a criterion for the focused blocks:

$$SVD_{DCT}^5 = \sqrt[5]{\sigma_1 \sigma_2 \sigma_3 \sigma_4 \sigma_5}. \quad (22)$$

Consequently, the block with the most value of  $SVD_{DCT}^5$  is selected as the focused block. The output fused image is constructed by the decision map (M) of selected suitable focused blocks. The decision map of two multi-focused images (imA and imB) is as follows:

$$M(i, j) = \begin{cases} 1 & \text{if } SVD_{DCT}^5(\text{imA}) > SVD_{DCT}^5(\text{imB}) \\ -1 & \text{if } \text{else} \end{cases} \quad (23)$$

where  $i$  and  $j = 1, 2 \dots \frac{\text{size of the input image}}{8}$ .

Accordingly for the  $(i, j)^{th}$  block of the output image, if  $M(i, j) = 1$ , imA is used and if  $M(i, j) = -1$ , imB is used. This stage of the proposed method is called DCT+SVD.

### C. Consistency Verification Step:

Consistency Verification (CV) which has been presented in [18], is a common step that is being used for post-processing in multi-focus image fusion algorithm to improve the output fused image quality. In this case, by considering an area of the decision map which the middle block is from imA while the majority blocks are from imB, the middle block must be from imB. This is possible by applying the majority filter (e.g 3×3 averaging filter) as follows:

$$W(i, j) = \frac{1}{9} \sum_{k=-1}^{+1} \sum_{l=-1}^{+1} M(i + k, j + l) \quad (24)$$

For  $W(i, j) > 0$ , the selected block for the output image is selected from imA and for  $W(i, j) < 0$ , the selected block for the output image is selected from imB. This stage has been named as DCT+SVD+CV.

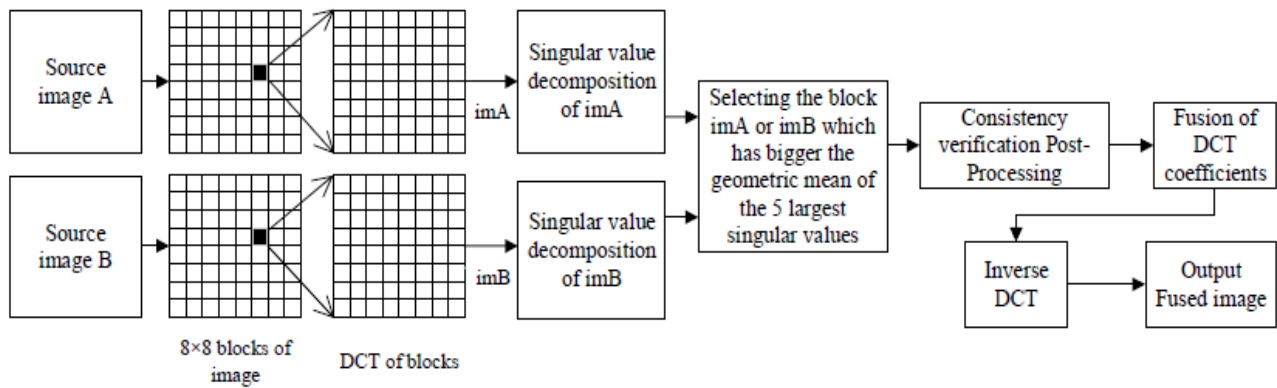
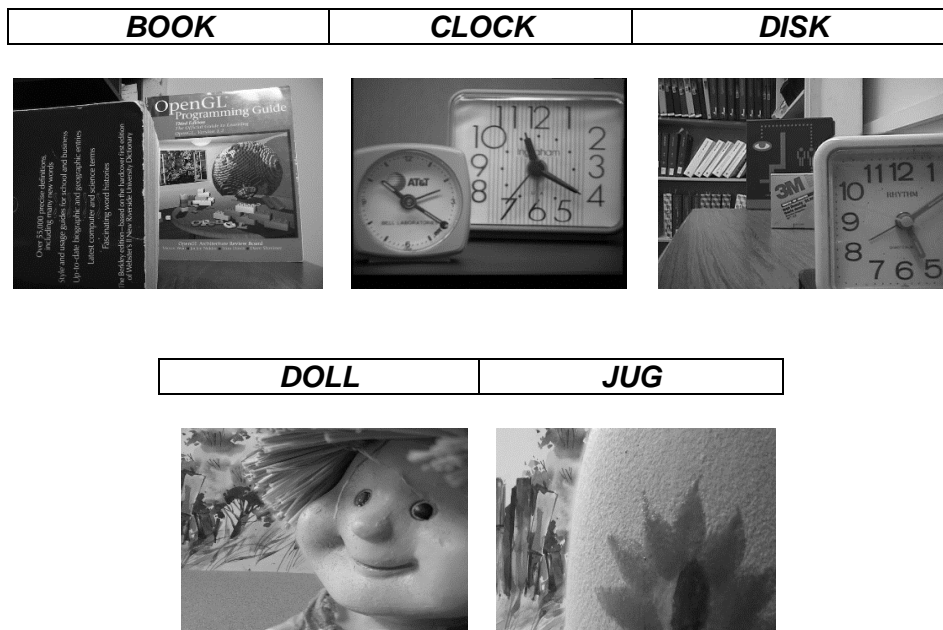


Figure 1. The general structure of the proposed method.

## EXPERIMENTAL RESULTS

A. Here Are some images used to demonstrate our Algorithm.:



**TABLE (1): Peak Signal to Noise Ratio of fused images (Using DCT+SVD) with respect to reference image.:**

**TABLE (2): Peak Signal to Noise Ratio of fused images (Using DCT+SVD+CV) with respect to reference image.:**

IMAGES		DCT+SVD	
S.NO		PSNR	SSIM
1.	BOOK	27.2714	0.9215
2.	CLOCK	25.8800	0.9063
3.	DISK	26.5168	0.8759
4.	DOLL	26.5912	0.8736
5.	JUG	24.6678	0.8483

IMAGES		DCT+SVD+CV	
S.NO		PSNR	SSIM
1.	BOOK	27.7403	0.9278
2.	CLOCK	26.3435	0.8883
3.	DISK	26.0989	0.8702
4.	DOLL	25.5642	0.8545
5.	JUG	26.2304	0.8343