# Image Fusion using Morphological Pyramid Consistency Method

Neha Uniyal
Department of CSE
GBPEC, Pauri, Garhwal
Uttarakhand, India

S.K. Verma
Department of CSE
GBPEC, Pauri, Garhwal
Uttarakhand, India

## ABSTRACT

Image fusion is an imperative approach of integrating relevant information from the set of images that may be captured from different sensors, acquired at different time or having different spatial or spectral characteristics. The objective of image fusion technique is to acquire more enhanced, reliable, efficient vignette and minimize redundancy in the output fused image while maximizing relevant information specific to desired application or task. Image fusion play an important role in the area of medical imaging, disaster monitoring, satellite imaging, environmental monitoring, land use/cover change detection, surveillance etc.

This paper focuses on the development of an image fusion method using morphological operator like and, or, Erosion, Dilation operator. Consistent analysis of techniques will help in deciding the suitability of a particular technique towards the fusion of large number of images. The results show the proposed algorithm has a better visual quality than the base methods. Also the quality of the fused image has been evaluated using a set of quality metrics.

## Keywords
Dilation, erosion, image fusion, or operator, and operator

## 1. INTRODUCTION
Image fusion has been receiving increasing attention in the research community with the goal of investigating general formal solutions to a wide spectrum of image processing applications. In the remote sensing field, the increasing availability of space borne imaging sensors, operating in a variety of ground scales and spectral bands, undoubtedly provides strong motivations. Because of the trade-off imposed by the physical constraint between spatial and spectral resolutions, spatial enhancement of poor-resolution multispectral (MS) data is desirable. Image fusion is the process of integrating two or more source images into a single fused image with the help of certain algorithms which can give more information as compared to single source images and contain less signal to noise ratio. Image fusion aims to extract all the relevant information from the source images and provides better way for reducing increasing volume of information.[1] Image fusion creates new images that are more beneficial for the purpose of visual/machine perception and for further image processing tasks such as Segmentation, region characterization, morphological processing, object detection or target recognition in application such as medical imaging and satellite application. The main goal of image fusion is to create a new fused image that is more suitable for the purpose of human /machine perception, contain all possible relevant information contained in the source images ,reduce amount of data and retain all information. Image fusion can be classified into three categories: pixel level fusion, feature level fusion, and decision or symbol level fusion [2]. At pixel level, fusion is done only on the basis of the intensity value of the pixels at particular points. At times neighborhood property of pixel values may also be used for combining the images. Most colour-transform based fusion techniques such as IHS and HSV transform fall under this category. In feature level fusion [3], source images are segmented into regions and features like pixel intensities, edges, and textures, are used for fusion. Feature level fusion requires individual features to be identified by object recognition schemes. Decision or symbol level fusion [4] is a high-level fusion which is based on statistics, voting, fuzzy logic, prediction and heuristics. Spatial and transform domains are the two fundamental approaches for image fusion. In spatial domain fusion, the fusion rule is directly applied to the intensity values of the source images. Averaging, weighted averaging, principal component analysis (PCA) are a few examples of spatial domain fusion scheme. One of the major disadvantages of spatial domain fusion method is that it introduces spatial distortions in the resultant fused image and does not provide any spectral information. Since many images those are used in digital image processing application are generally of poor contrast, the spatial information should be preserved in the medical images without introducing any distortion or noise. These requirements of images are better preserved in transform domain fusion.

## 2. IMAGE FUSION TECHNIQUES
### 2.1 Average method
Average method remains one of the most effective, yet simple and easy to implement image fusion algorithm. However, the method is based on a local-Gaussianity assumption for wavelet subbands.
1. Decompose each input image into subbands
2. For each subband pair X, Y, except the lowpass residuals:
    a. Compute neighborhood saliency measures, $\sigma X$ and $\sigma Y$
    b. Compute neighbourhood matching coefficient
    c. Calculate the fused coefficients
3. Average coefficients in lowpass residuals.
4. .Reconstruct the fused image from the processed subbands and the lowpass residual.

### 2.2 Maximum selection
Maximum selection fusion schemes and the DWT multiscale and multiresolution fusion schemes are required for data correlation. This method allows us to easily correlate the data in a much more efficient way in cases of subject responses. One method widely used to combine the wavelet coefficients of the source images is maximum selection method. However, we can eventually apply window-based type of verification to test for consistency.

## 2.3 Minimum selection

Like Maximum selection we can apply Minimum selection fusion schemes and the DWT multiscale and multiresolution fusion schemes for efficient correlation of data. In this method the objective metric results for the data correlate even better to subject responses with correct classification. The method used for the combination of wavelet coefficients of the source images is Minimum selection technique. Afterwards, we have to apply window-based type of verification for testing consistency .

## 2.4 PCA Method

Principal component analysis (PCA) [5] is a vector space transform often used to reduce multidimensional data sets to lower dimensions for analysis. PCA is a mathematical tool which transforms a number of correlated variables into a number of uncorrelated variables PCA method is quite similar to IHS except that, instead of the IHS transformation, the Principal Components (PC) of the MS bands are calculated and the first PC is replaced by the PAN band. The advantage of PCA over IHS is that it does not  have the three band limitation and can be applied to any number of bands at a time. However, this technique also introduces spectral distortion in the pan sharpened image like the IHS method

## 2.5 Laplacian Pyramid

1.  Considering the pair of input image matrices as M1 and M2 respectively.
2.  A single dimensional filter mask is generated as W = [1/16 4/16 6/16 4/16].
3.  The level of fusion (decomposition and recomposition) is decided upon. Both the decomposition part and the recomposition part are iteratively executed "level" number of times.[4]
4.  Decomposition
    a.  The image matrices are filtered (convolved) vertically and horizontally with filter mask generated, producing the filtered image matrices G1 and G2 respectively
    b.  An additional level of filtering is performed over G1 and G2 producing filtered matrices G11 and G22.
    c.  The difference matrix is calculated for both the (G11 and G22). The differences between the original image matrices and the second level of filtered image matrices.
    d.  The Pyramid of this level of decomposition is generated using any of the three algorithms below:
        i.    Select Maximum
        ii.   Burt's Method
        iii.  Lis Method
    e.  The pyramid thus formed is retained for the level as E[level].
    f.  The images are decimated to half the size and the decomposition steps are iterated "level" number of times.
5.  The finally decimated pair of images M1 and M2 is manipulated as one of the following, producing X matrix.
    a.  Average M1 and M2
    b.  Select Maximum in M1 and M2
    c.  Select Minimum in M1 and M2
6.  Recomposition
    a.  Matrix X obtained in step 5 is undecimated by alternatively padding zero columns and rows.
    b.  The undecimated matrix is filtered (convolved) with the doubly scaled filter mask W.

c.  The filtered matrix is added upon with the retained pyramid of the level E[level].
d.  The matrix generated in step c will act as the input matrix X to the next level of recomposition.
e.  Recomposition steps are performed "level" number of time, eventually undecimating  the matrices in each level, obtaining the fused image matrix of the original size at the end of the final level.

## 2.6 FSD

1.  Consider a pair of input image matrixes as M1 and M2.
2.  A single dimensional mask is generated as W= [1/16,4/16,6/16,4/16,1/16].
3.  The level of fusion (decomposition and recomposition) is decided upon. Both the decomposition part and the recomposition part are iteratively executed "level" number of times.
4.  Decomposition
    a.  The image matrices are filtered (convolved), vertically and horizontally, with the filter mask generated, producing the filtered image matrices G1 and G2 respectively.
    b.  The difference matrix is calculated for both the images as (M1-G1) and (M2-G2).The difference between the original image matrices with the filtered image matrices.
    c.  The Pyramid of this level of decomposition is generated using any of the three algorithms below:
        i.    Select Maximum
        ii.   Burts' Method
        iii.  Lis Method
    d.  The pyramid thus formed is retained for the level as E[level].
    e.  The images are decimated to half the size and the decomposition steps are iterated "level" number of times.
5.  The finally decimated pair of images M1 and M2 is manipulated as one of the following, producing X matrix.
    a.  Average M1 and M2
    b.  Select Maximum in M1 and M2
    c.  Select Minimum in M1 and M2
6.  Recomposition
    a.  Matrix X obtained in step 5 is undecimated by alternatively padding zero columns and rows.
    b.  The undecimated matrix is filtered (convolved) with the doubly scaled filter mask W.
    c.  The filtered matrix is added upon with the retained pyramid of the level E[level].
    d.  The matrix generated in step c will act as the input matrix X to the next level of recomposition
    e.  Recomposition steps are performed "level" number of time, eventually undecimating the matrices in each level, obtaining the fused image matrix of the original size at the end of the final level.

## 2.7 Gradient Pyramid

1.  Considering the pair of input image matrices as M1 and M2 respectively.
2.  Two dimensional filters W = [1/16 4/16 6/16 4/16 1/16] ,V = [1/4 2/4 1/4] and four directional filters
    a.  Horizontal filter
$$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$
    b.  Vertical filter

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

c. Diagonal filter

$$\begin{bmatrix} 0 & 0 & 0.5 \\ 0 & -1 & 0 \\ 0.5 & 0 & 0 \end{bmatrix}$$

d. Diagonal filter

$$\begin{bmatrix} 0.5 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}$$

3. The level of fusion (decomposition and recomposition) is decided upon. Both the decomposition part and the recomposition part are iteratively executed "level" number of times.
4. Decomposition
   a. The image matrices are filtered (convolved) vertically and horizontally, with the filter mask(W) generated, producing the filtered image matrices G1 and G2 for M1 and M2 respectively
   b. An additional level of filtering is performed over M1 and M2 with the filter mask(V) producing filtered matrices Z1 and Z2.
   c. Compute directional derivatives for the filtered matrices Z1 and Z2 using four the directional filters. Directional derivative using the first filter will be input for second derivative using second filter and so on.
   d. The Pyramid of this level of decomposition is generated using any of the three algorithms below:
      i. Select Maximum
      ii. Burts' Method
      iii. Lis Method
   e. Divide the coefficient by 8 before retaining.
   f. The pyramid thus formed is retained for the level as E[level].
   g. The images are decimated to half the size and the decomposition steps are iterated "level" number of times.
5. The finally decimated pair of images M1 and M2 is manipulated as one of the following, producing X matrix.
   a. Average M1 and M2
   b. Select Maximum in M1 and M2
   c. Select Minimum
6. Recomposition
   a. Matrix X obtained in step 5 is undecimated by alternatively padding zero columns and rows.
   b. The undecimated matrix is filtered (convolved) with the doubly scaled filter mask W.
   c. The filtered matrix is added .
   d. The matrix generated in step c will act as the input matrix X to the next level of recomposition.
   e. Recomposition steps are performed "level" number of time, eventually undecimating the matrices in each level, obtaining original size at the end of the final level.

## 2.8 DWT with DBSS(2,2) wavelet

Two sets of wavelet coefficients are obtained, including approximation (LL) and detail (HL, LH, and HH) signals of the original data. A decision fusion strategy usually establishes the rule by which spatial details are injected. The approximation coefficients LLHi and LLMS are considered for computing local correlation coefficients (LCC) over a square sliding window An LCC map is computed between the approximations LLMS of each of the MS bands, and that of the image, LLHi , both at the scale of the fused image.

## 2.9 Proposed Method
## Morphological Pyramid consistent method

An effective algorithm which is suitable for combining multi-focus image sets of a scene is proposed in this section. The spatial frequency of a pixel's neighbour block is used to judge its sharpness and morphological opening and closing are used for post processing. Finally, the fused image is obtained by combining the focused regions To correct for these defects, morphological opening and closing constructed by combining dilation and erosion are employed.

As an assumption opening, denoted as $Z \circ B$, is simply erosion of Z by the structure element B, followed by dilation of the result by B. It removes thin connections and thin protrusions. Closing, denoted as $Z \bullet B$, is dilation Followed by erosion. It joins narrow breaks and fills long thin gulfs. Holes larger than B cannot be removed simply using opening and closing operators. In practice, small holes are always judged incorrectly therefore, a threshold, is set to remove the holes smaller than the threshold. Opening and closing are again performed to smooth object contours.

Proposed algorithm

1. Considering the pair of input image matrices as M1 and M2 respectively.
2. The level of fusion (decomposition and recomposition) is decided upon. Both the decomposition part and the recomposition part are iteratively executed "level" number of times.
3. Decomposition
   a. As a standard morphological pyramid procedure, the input images M1 and M2 are filtered at two levels, namely, open filtering and closed filtering. The same are also known as image opening and image closing.
   b. Image opening is a procedure consisting of the following operations,
      i. Image erosion
      ii. Image dilation
   c. Image closing procedure, complementing the image opening process, consists of

the following operations,
      i. Image dilation
      ii. Image erosion
   d. Step b and c are believed to get rid of noise in the image very effectively.
   e. The image erosion process is performed using "order filtering" procedure which is similar to median filtering, except for the order of filtering is explicitly mentioned as the minimum in the 5x5 neighbourhood.
   f. In image dilation procedure, similar to image erosion, the order of filtering is explicitly mentioned as the maximum in the 5x5 neighbourhood.
   g. The thus obtained matrices are finally filtered with image dilation process over a 7x7 neighbourhood of each pixel.
   h. The difference matrices are calculated as the difference between the input image matrix of the level and the filtered image matrix.
   i. The Pyramid of this level of decomposition is generated using any of the three algorithms below:
   i. Select Maximum
   ii. Burt's Method
   iii. Lis Method
   j. The pyramid thus formed is retained for the level as E[level].

k.    The images are decimated to half the size and the decomposition steps are iterated

"level" number of times.

4.    The finally decimated pair of images M1 and M2 is manipulated as one of the following,

producing X matrix.

    a.    Average M1 and M2
    b.    Select Maximum in M1 and M2
    c.    Select Minimum in M1 and M2

5.    Recomposition

    a.    Matrix X obtained in step 5 is undecimated by alternatively padding zero columns and rows.

    b.    The undecimated matrix is filtered as image dilation over a 7x7 neighbourhood of each pixel.

    c.    The filtered matrix is added upon with the retained pyramid of the level E[level].

    d.    The matrix generated in step c will act as the input matrix X to the next level of recomposition.

    e.    Recomposition steps are performed "level" number of time, eventually undecimating the matrices in each level, obtaining the fused image matrix of the original size at the end of the final level.

# 3.    QUANTITATIVE IMAGE QUALITY MEASURES

Image quality is a characteristic of an image that calculates the perceived image degradation (typically, compared to an ideal or perfect image). Imaging systems may introduce some amounts of distortion or artifacts in the signal, so the quality assessment is an important problem.[8] with the help of image quality measures we can know which image fusion technique is better as compared to others. The  fused image will be compared with all other methods in terms of present parameters as described above.

## 3.1. Average Difference

Average difference is the average value of the difference between the actual/ideal data and the obtained/resultant data . This metric helps in providing the overall average difference between the corresponding pixels of the two images proving us a value that specifies, how much different is the fused image from the perfect image[8].

$$AV = \frac{1}{mn}\sum_{i=0}^{m}\sum_{j=0}^{n}(A_{ij} - B_{ij}) \qquad (1)$$

## 3.2 Maximum difference

Maximum Difference is a very simple metric that gives us the information of the largest of the corresponding pixel error .

$$MD = \max(A_{ij} - B_{ij}), i = 0,1,2\ldots m; j = 0,1,2\ldots n \quad (2)$$

## 3.3 Mean square error  (MSE)

Mean square error is one of the most commonly used error projection method where, the error value is the value difference between the actual data and the resultant data[8]

The mean of the square of this error provides the error or the actual difference between the expected/ideal results to the obtained or calculated result.

$$MSE = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}(Aij - Bij)^2 \qquad (3)$$

## 3.4. Normalised Cross-correlation

Cross correlation is performed between the expected data and the obtained data and normalized with respect to the expected data. This  metric is calculated as the ratio between the net sum of the multiplication of the corresponding pixel densities of the perfect and the fused images and the net sum of the squared values of the pixel densitied of the perfect images. The Normalized Cross Correlation value would ideally be 1 if the fused and the perfect images are identical.

$$NCC = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n}(A_{ij}*B_{ij})}{\sum_{i=0}^{m}\sum_{j=0}^{n}(A_{ij})^2} \qquad (4)$$

## 3.5  Normalise Absolute Error

This is a metric where the error value is normalised with respect to the expected or the perfect data. The Normalized Absolute value will be zero (0) if both the fused and the perfect images are identical.

$$NAE = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n}(A_{ij} - B_{ij})}{\sum_{i=0}^{m}\sum_{j=0}^{n}A_{ij}} \qquad (5)$$

## 3.6 Peak signal to noise ratio

PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation [2][9]. The PSNR measure is given by:-

$$PSNR = 10 * \log_{10}\left(\frac{peak^2}{MSE}\right) \qquad (6)$$

## 3.7 Structural Content

It is the ratio between the net sum of the square of the expected data and the net sum of square of the obtained data.

$$SC = \frac{\sum_{i=0}^{m}\sum_{j=0}^{n}(A_{ij})^2}{\sum_{i=0}^{m}\sum_{j=0}^{n}(B_{ij})^2} \qquad (7)$$

Where *m* is the height of the Image implying the number or pixel rows

*n* is the width of the image, implying the number of pixel columns.

*A(i,j)* being the pixel density values of the perfect image.

*B(i,j)* being the pixel density values of the fused image

# 4.    EXPERIMENTAL-RESULT

**Table 1 : Table of quality image set metrics**

| Techniques ↓ | MSE | AD | Mean absolute error | Normalised cross correlation | Peak signal to noise ratio | Maximum difference | SC |
|---|---|---|---|---|---|---|---|
| Average | 56.8620 | .2490 | .0406 | .9938 | 20.5826 | 127 | 1.0077 |
| Maximum | 9.796464 | -128.1431 | 1.366 | 1.8265 | 4.8420 | 135 | .2256 |
| Minimum | 9.57927 | -119.6736 | 1.3305 | 1.7525 | 5.0080 | 227 | .2371 |
| PCA | 9.6551135 | -123.9051 | 1.3430 | 1.7895 | 4.9494 | 176 | .2318 |
| Laplacian pymd | 9.585401 | -124.0821 | 1.3356 | 1.7919 | 4.9933 | 173 | .2325 |
| Gradient pymd | 9.598029 | -124.0922 | 1.3357 | 1.7921 | 4.9934 | 173 | .2325 |
| DWT with DBSS | 9.902477162 | -130.6245 | 1.3831 | 1.8476 | 4.7633 | 134 | .2219 |
| Morphological pyramid consistency method | 8.3609 | 3.0873 | .0608 | .9813 | 26.7839 | 94 | 1.0268 |

Fig 2: pair of input images

(i)   (ii)   (iii)

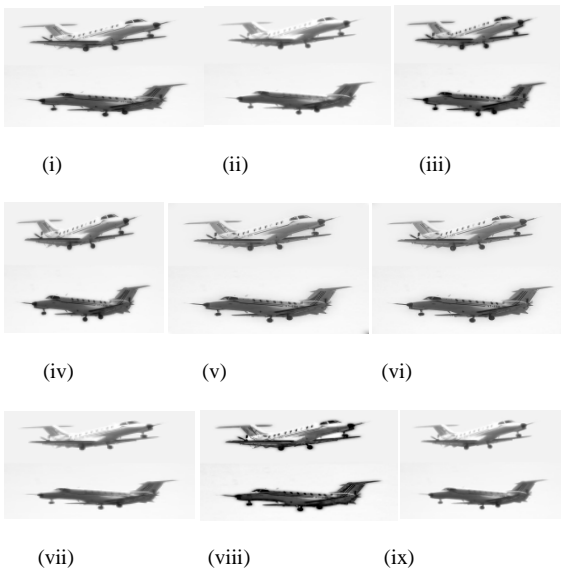(iv)   (v)   (vi)

(vii)   (viii)   (ix)

**Fig 3: Fused image using (i).average method (ii)maximum selection(iii)minimum-selection(iv)PCA(v)laplacian pyramid(vi)FSD(vii)gradient-pyramid(viii)DWTwith DBSS(ix)morphological pyramid**

## 5. CONCLUSION

From the table of Quality metrics for different images it was concluded that the proposed algorithm morphological pyramid algorithm having better value of PSNR when compared to other fusion methods, which shows it as the better fusion technique. Similarly if we check the MSE value, it is the least in proposed. Thus it proves proposed method performs the best.

## 6. REFERENCES

[1] Petrovic V, Xydeas C, Feb 2004. Gradient based Multiresolution Image Fusion, IEEE Transaction on Image Processing, Vol. 13, No .

[2] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*,vol. 11, pp. 674–693, July 1989.

[3] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, pp. 14–38, Apr. 1991.

[4] A.Akerman, "Pyramid techniques for multisensor fusion," *Proc. SPIE*, vol. 1828, pp. 124–131, 1992.

[5] X. Liu andW.Yang, "Enhanced visualization of images through fusion," *Proc. SPIE*, vol. 4231, pp. 340–345, Oct. 2000.

[6] P. Burt and R. Kolczynski, "Enhanced image capture through fusion,"in *Proc. 4th Int. Conf. Computer Vision*, Berlin, Germany, 1993, pp.173–182.

[7] H. L. B. Munjanath and S. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp.

[8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Transactions on Image Processing, Vol. 13, no. 4, pp.600-612, Apr.2004.

[9] Florence Laporterie, Guy Flouzat, "*Morphological Pyramid Concept as a Tool for Multi Resolution Data Fusion in Remote Sensing"*, Integrated Computer-Aided Engineering, pages 63-79, 2003.

[10] Peter J Burt, Edward Adelson, "*Laplacian Pyramid as a Compact Image Code"*, IEEE Transactions on Communications, Vol Com-31, No. 4, April 1983.