Image Fusion using Cross Bilateral Filter and Wavelet Transform Domain

Kapil Joshi, N.K. Joshi, Manoj Diwakar

Abstract: The actual goal of fused images is all about to reduce the complexity of visual data of any images using the process of merging two relevant visual images of the similar scene. In this paper, an algorithm on fusion concept through discrete wavelet transforms (DWT) and cross bilateral filter (CBF) is introduced. In this proposed detailed work, all existing images are dissolved into obscure frequency sub bands and advanced frequency sub bands with the help of discrete wavelet transform (DWT). After converting these images will be categorized into obscure frequency sub bands transformed images and advanced frequency sub bands transformed images. We have used pixel average method and weighted average method but pixel averaging method is for low frequency sub bands transformed images meanwhile weighted averaging method is applicable for high frequency sub bands transformed images. Further the weights are calculated by cross bilateral filter (CBF) on both types of images. At the end of process, DWT is applied to rebuild the fused resultant images over the fused some coefficients. The introduced work completely tested on various multi-focus images and other multi-sensors images. Existing methods results and proposed methods results are simultaneously compared with the various metrics just for qualitative measurement. For future prospective, the proposed results will be better than previous existing completed work in the form of qualitative as well as the quantitative parameters.

Index Terms: Image fusion, Discrete Wavelet Transform, Inverse Discrete Wavelet Transform, Cross Bilateral Filter.

1. INTRODUCTION

In current research field, image fusion is growing day by day. It has wide scope to do the research in image processing field. Image fusion is just a genuine process for merging the visual data from different kind of source images of the similar scene or to generate the relevant data into a separate fused image. We can say that the separate fused image will be considered as a individual enhanced image without any loss of data. The fused image may be more suitable for human being phenomenal intelligence and for imaging process tasks. Image fusion already developed in various zone like as satellites imaging, medical imaging fields, surveillance fields, robotics fields and military fields etc. Especially in medical field, many researchers got the achievement with the machine perception [1].

A single image having less appropriate information/data does not clear the particular scene. Consider, we take two similar images [16] where first image is focused with some part of left or right side and another portion is already in out of focusing then second image will be focused with defocused portion that is already in the first one image. In the whole process, left side of first image will be focused i.e. defocused (out of focus) and the right side of second image will be focused is also defocused (unable to focus) [33]. To complete the whole process within a individual image is very sensorial so it will be achieved through multi-focus. This process is known as multi-focusing [14] image fusion process and those images be knowing like multi-focusing [2] images. It is a challenging task for every researcher [3-5].

Every image can be captured with the help of no. of sensors [31]. It will be categorized into the two sections i.e. single sensor and multi sensor. In both sensors, commonly, there is a sequence of images in a real world but the gap is all about the sensors, one scene can be taken from single sensor apart from that many images of a single scene are captured through multiple sensors [15]. Multi-sensors remove all the barriers of single-sensor image fusion [2][12].

When we capture some images using single sensor or multi sensors and those images could be medical images. These medical images are of similar types like as CT (computed Tomography) images and MRI (Magnetic Resonance Imaging) images [8][21][34][35][36]. Such kind of images supplies various information of the similar organs to diagnose the illness [6][17].

CT types of images allow the useful data of bones, blood vessels as well as hard tissues where as MRI images [7] pass the essential knowledge of the soft tissues. Those multiple information’s are separated by the other kind of sensors having their own capability. So if we merge a couple of images into a specific [10] individual image then after the information about bones and soft tissues [17] of the particular organ which will generate more benefits for further diagnosis [19].

This whole process of merging the relevant information is known as multi sensors fusion techniques [33]. Further study, Image fusion is performed at different levels where fusion is classified into three levels and their features. These are Pixel level/signal level, features level/object level and other decision level/symbol level [22]. Many Applications [25-32] such as image denoising, image fusion follow the concept of DWT to perform the respective operations.

In pixel based fusion level, [15] fused image is generated through respective pixels and it’s a bottom level fusing process.
Features level fusion shows the complete features of any image like its edges, texture pattern and different color shades that have been fied through original image. Lastly the decision based fusion level determines the top level fusion that modulates the final result using many algorithms approaches for capturing the conclusive images. In between, pixel based fusion level is fully utilized due to the less time consumption. Fusion methods are of two type’s domains. Spatial [24] and transform/frequency based domain [21]. The spatial domain processes the fusion progress directly and it is very easy to implement. But transform domain works on many transformation methods like Discrete Wavelet Transform (DWT) [8], Stationary Wavelet Transform (SWT) [23], curvelet [18] Transform [9] and contourlet Transform, Non sub sampled contourlet transform (NSCT) [13] & multi-resolution singular value decomposition (MSVD). In image processing, pixel to pixel is the simple fusion operation and it has already attracted by many researchers. The problem of this method is about the undesirable many effects. To overcome this situation, multi-resolution methods have been proposed and having three following steps. First of all, source image are basically decomposed into low and high frequency information & it contains some transformed coefficients. Secondly, these transformed based coefficients are merging side by side with few fusion rules. Finally, the inverse transformation method is executed over the coefficient to rebuild the fused images. Further the multi-resolution [11] method provides the better resolution and visual results. Apart from this, other transformation methods preserve some major characteristics of image namely edges, texture, line and color used it fusion progress. Discrete Cosine Transformation (DCT) [16] is also introduced for the concept of image fusion. So afterwards a maiden multi-resolution DCT decomposition concept has been given the solution of computational factors in absentia any other harm of information. Various fusion schemes have been invented using the different methods as max, min, averaging, Principal Component Analysis (PCA) [13], Intensity Hue Saturation (IHS) [16][22]. In these methods, averaging method gives the regular data and it produce the better fusion results. Other methods are also useful in image fusion better performance [24].

II. THEORIES OF DWT AND CBF

Some basic methods are discussed in below subsections which are used for proposed methodology.

A. DWT

A DWT technique [16] comes under the transform domain and it has many types of property such as multi-resolution, multi-spectral, multi-focus, multi-temporal etc. In between multi-resolution property is generally used in digital image processing. DWT is the core concept of frequency/transformation domain. The initial steps of DWT are just for converting any image from spatial (pixel) domain to frequency domain [20]. In DWT [17], decomposition is the actual process where images clearly converting and separating into low and high frequency based sub bands. Low frequency bands compare with the nearest approximation part which consists of usual data of the whole image and it is defined as low-low sub bands. As well as, the high frequency founded sub bands are treated as detail part. It contains the glaring images. The detail part is structured with three other bands where as low-high (LH), high-low (HL), and high-high (HH). In second level, low low (LL) part will be again decompose into four sub bands and rest of other three frequency sub bands. Low-high, low-high and high-high (LH, HL, and HH) will be remain same as figure 1.

![Figure 1](image)

**Fig 1.** (a) First level-based DWT decomposition (b) Second level-based DWT decomposition

B. CBF

In image processing, filters are used to stifle the high frequency in any image. It means clear and sleeking the images. Filters process is also used to enhance the quality of image as well as detecting edges. There is no. of filters techniques used in image processing and in between the Gaussian filter is one of plausible filter providing sharpness and isolates unnecessary noisy information from the existing images but the drawback of this method is that it ignores the accurate details. To track this kind of problem, cross bilateral filter was proposed by the scientists naming tomasi and manduchi. Bilateral filter is called as a non-linear filtering technique used to preserve the edges and it’s also for reducing the noise of an image. It protects the sharper details. Further we extend this bilateral filter technique using two different filter kernels. These are spatial based filter kernel and range based filter kernel. Spatial based filter kernel is considered as low passing filter used to gain geometric pattern closeness amid the neighborhood pixels now those range based filter kernel is just an edges stopping function applicable for the scaling of gray color.

Both type of filter kernels are depends on particular gaussian distribution process weighted acquired by means of Euclidian distance and gray or color spaces.

For single image naming X, the following output of bilateral filter at individual pixel location p is defined as:

\[
Y_{CBF}(p) = \frac{1}{W} \sum_{q \in S} G_{x}(||p - q||) \times G_{y}(||X(p) - X(q)||) Y(q)
\]

Where,

\[
W = \sum_{q \in S} G_{x}(||p - q||) \times G_{y}(||X(p) - X(q)||)
\]
is a normalization factor and

\[ G_{\alpha}(X(p) - X(q)) = e^{-\frac{|X(p) - X(q)|^2}{2\alpha^2}} \]

is a level of gray scale similarity or a function of edge stopping.

III. PROPOSED WORK

The proposed method has been demonstrated in figure.2. In the given chart two blurred images are accessed for input. These images are fused using wavelet transform where approximation and detail part is obtained. Over the both approximation part, cross bilateral filter is applied and over the detail part, weighed fusion is performed.

The following steps of the given proposed method are mentioned as below:-

Step 1: First we take two images i.e. \(i_1\) & \(i_2\).

Step 2: Discrete Wavelet Transform will be applied on two given images and decompose them into low and high frequency sub bands.

Step 3: After completing decomposition, we found approximation and detail parts to perform further task.

Step 4: The Cross Bilateral Function is used to create the sharpness and clear all the characteristics of images so it will be tested on approximation part of both images i.e. \(i_1\) and \(i_2\).

Step 5: Weighed fusion is carried through the detail part of the decomposed images. The weight value is obtained by estimating value of gray level similarity in normalized form. Weighted fusion using correlation (Corr) is performed on detail parts using following relationship:

\[ R_{\text{final}} = \frac{1}{i} \sum_{i=1}^{k} \beta_i T_i \]

where, \(\beta_i = \frac{\text{Var}^{-1}(T_i)}{\sum_{i=1}^{k} \text{Var}^{-1}(T_i)}\)

\(T_i\) is detail part of both input images, \((R_i)\) with block size b.

The function \(\text{Var}^{-1}\) can be estimated as inverse variance.

Step 6: Integrate these two section such as fuse and weighted fusion using Inverse Discrete Wavelet Transform (IDWT) and get the new coefficients \((L_{\text{new}}, L_{\text{new}}, H_{\text{new}}, H_{\text{new}})\) as well as the fused image.

IV. RESULT AND DISCUSSION

This new proposed method is executed on particular data set 1 of pilot images. The further result are shown in clock images i.e. fig. 3. Every image has the length of 512x512.

Left and right side images are blurred in fig 3(a) & 3 (b) and top and bottom side images are also blurred in fig 4(a) & 4(b).

The parameters \(\sigma_s=1.8\)and \(\sigma_r=25\) are used in proposed method with the window size of 5x5.

Fig.3. Image Fusion results for clock image data set: (a) left portion blurred; (b) right portion blurred; (c) fused image using DWT method; (d) fused image using PCA method; (e) fused image using SWT method; (f) fused image through CBF method; (g) fused image is achieved through proposed method.
Image Fusion using Cross Bilateral Filter and Wavelet Transform Domain

Fig. 4. Fusion results for Airplane image data set: (a) upper portion blurred; (b) lower portion blurred; (c) fused image using DWT method; (d) fused image through PCA method; (e) fused image by SWT method; (f) fused image by CBF method; (g) fused image using given proposed method.

The following metrics are used to evaluate the results:

A. **Average Pixel Intensity (API) or mean (µ):** It is used to measure an index of contrast, which is represented as

\[ API = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} f(i,j)}{mn} \]  

Whereas \( f(i,j) \) represents the fused image and its size will be considered as \( mn \).

B. **Standard Deviation (SD):** Its is described as

\[ SD = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (f(i,j) - \mu)^2}{mn}} \]  

It meditates the expansion in respective data and the fusion result is represented by higher value.

C. **Average Gradient (AG):** It shows the actual clarity with sharpness so it is defined as

\[ AG = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (|f(i,j) - f(i+1,j)| + |f(i,j) - f(i,j+1)|)^2}{mn} \]  

D. **Entropy (EN):** It scales the number of information available in fused image so it is allocated as below:

\[ EN = - \sum_{k=0}^{255} p_k \log_2(p_k) \]  

Where \( p_k \) called as the effective probability of its intensity value defined as \( k \) in any image. The higher values define preferable results.

E. **Mutual Information (MI):** It declares the actual sum of fraternal details between original images and fused images and scales the higher degree of dependence of both those images. The formula is given as follow:

\[ MI = \sum_{l=1}^{m} \sum_{j=1}^{n} h_{xf}(i,j) \log_2 \frac{h_{xf}(i,j)}{h_x(i)h_f(j)} \]  

Where \( h_{xf} \) as the joint histogram between \( x \) and \( f \).

In \( MI_x \), \( Y \) will be as source images and \( F \) as the fused images. Similarly the reciprocal knowledge between two source image \( X,Y \) and fused image as \( F \) is represented as:

\[ MI_{\text{total}} = MI_{xf} + MI_{fy} \]  

\( MI_{\text{total}} \) denote the preferable fusion results.

F. **Fusion symmetry (FS):** It shows that the large number of symmetric details of the fused image in terms of original image so it’s given as:

\[ FS = 2 - \left| \frac{MI_{xf}}{MI_{\text{total}}} - 0.5 \right| \]  

G. **Correlation Coefficient (CC):** It calculates the contingency of fused image to main original image, higher value shows the relevant fused results. This is considered as follow:

\[ CC = \frac{r_{xf} + r_{fy}}{2} \]  

Where,

\[ r_{xf} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (X(i,j) - \bar{x})(f(i,j) - \mu)}{\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (X(i,j) - \bar{x})^2} \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (f(i,j) - \mu)^2}} \]  

And

\[ r_{fy} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (Y(i,j) - \bar{y})(f(i,j) - \mu)}{\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (Y(i,j) - \bar{y})^2} \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (f(i,j) - \mu)^2}} \]  

H. **Spatial Frequency (SF):** It also calculates the complete relevant information that is represents as:

\[ SF = \sqrt{RF^2 + CF^2} \]  

Where RF considered as row frequency and CF is defined as column frequency. So these are denoted as:

\[ RF = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (f(i,j) - f(i,j-1))^2}{mn}} \]  

\[ CF = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (f(i,j) - f(i+1,j))^2}{mn}} \]
\[ CF = \sqrt{\frac{\sum_{i=1}^{n} \sum_{j=1}^{m} (f(i,j) - f(i-1,j))^2}{mn}} \]

The highest value of spatial frequency shows the major amount of useful data in a particular image.

I. Q^\text{xy/f} - It estimates the complete information that is shifted from source image to fused image so it is explained in mathematical form.

\[ Q^{xy/f}(i,j) = Q^x(i,j) + Q^y(i,j) \]

Where \( x \) and \( y \) are the source images and \( f \) is defined by fused image. The definition of \( Q^x \) and \( Q^y \) is defined as follow:

\[ Q^x(i,j) = Q^x(i,j) - Q^x(i-1,j) \]
\[ Q^y(i,j) = Q^y(i,j) - Q^y(i-1,j) \]

where \( Q^{xy/f}(i,j) \) and \( Q^x(i,j) \) shows the strength of edges.

### Table 1: Fusion Metrics

<table>
<thead>
<tr>
<th>Input Images</th>
<th>Fusion Methods</th>
<th>SD</th>
<th>CC</th>
<th>API</th>
<th>AG</th>
<th>EN</th>
<th>MI</th>
<th>FS</th>
<th>SF</th>
<th>Q^\text{xy/f}</th>
<th>L^\text{xy/f}</th>
<th>N_{xy/f}</th>
<th>N_{x/f}</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOCK</td>
<td>[8]</td>
<td>49.316</td>
<td>0.9721</td>
<td>97.038</td>
<td>3.879</td>
<td>3.8869</td>
<td>5.728</td>
<td>1.8904</td>
<td>6.2841</td>
<td>0.8341</td>
<td>0.1659</td>
<td>0.0017</td>
<td>0.0016</td>
</tr>
<tr>
<td></td>
<td>[26]</td>
<td>49.316</td>
<td>0.9754</td>
<td>97.037</td>
<td>3.878</td>
<td>4.3445</td>
<td>5.875</td>
<td>1.8722</td>
<td>6.2817</td>
<td>0.8338</td>
<td>0.1662</td>
<td>0.0021</td>
<td>0.0013</td>
</tr>
<tr>
<td></td>
<td>[12]</td>
<td>49.409</td>
<td>0.9872</td>
<td>97.035</td>
<td>4.472</td>
<td>4.4882</td>
<td>6.492</td>
<td>1.9021</td>
<td>7.4981</td>
<td>0.8700</td>
<td>0.1299</td>
<td>0.0118</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>[24]</td>
<td>49.893</td>
<td>0.9869</td>
<td>96.548</td>
<td>5.5265</td>
<td>7.2755</td>
<td>7.3415</td>
<td>2.1960</td>
<td>7.1042</td>
<td>0.8295</td>
<td>0.0995</td>
<td>0.0014</td>
<td>0.0011</td>
</tr>
<tr>
<td>PROPOSED</td>
<td>50.417</td>
<td>0.9890</td>
<td>97.080</td>
<td>4.4044</td>
<td>7.2761</td>
<td>6.8687</td>
<td>1.9934</td>
<td>8.0545</td>
<td>0.8583</td>
<td>0.1406</td>
<td>0.0056</td>
<td>0.0025</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[8]</td>
<td>51.872</td>
<td>0.9746</td>
<td>221.44</td>
<td>3.1869</td>
<td>2.5333</td>
<td>1.1918</td>
<td>1.9878</td>
<td>10.2225</td>
<td>0.7938</td>
<td>0.2042</td>
<td>0.0296</td>
<td>0.0072</td>
</tr>
<tr>
<td></td>
<td>[26]</td>
<td>51.874</td>
<td>0.9798</td>
<td>221.43</td>
<td>3.1871</td>
<td>2.8173</td>
<td>5.0368</td>
<td>1.9590</td>
<td>10.2257</td>
<td>0.7899</td>
<td>0.2041</td>
<td>0.0279</td>
<td>0.0079</td>
</tr>
<tr>
<td></td>
<td>[12]</td>
<td>52.176</td>
<td>0.9876</td>
<td>221.42</td>
<td>4.2307</td>
<td>3.5673</td>
<td>5.3640</td>
<td>1.9526</td>
<td>13.5788</td>
<td>0.8698</td>
<td>0.1299</td>
<td>0.0036</td>
<td>0.0052</td>
</tr>
<tr>
<td></td>
<td>[24]</td>
<td>52.562</td>
<td>0.9886</td>
<td>220.58</td>
<td>4.9215</td>
<td>4.2667</td>
<td>5.1454</td>
<td>1.9635</td>
<td>16.9262</td>
<td>0.9284</td>
<td>0.0635</td>
<td>0.0345</td>
<td>0.0082</td>
</tr>
<tr>
<td>PROPOSED</td>
<td>55.242</td>
<td>0.9904</td>
<td>221.45</td>
<td>5.8073</td>
<td>4.7166</td>
<td>5.0203</td>
<td>1.9766</td>
<td>18.6560</td>
<td>0.8452</td>
<td>0.1474</td>
<td>0.0286</td>
<td>0.0076</td>
<td></td>
</tr>
</tbody>
</table>

Visually we can see the fused images from new proposed scheme are much better than previous existing methods. To improve the quality of fused image only visual overview is not enough. The following metrics are used in quantitative measurement. After experimental of all tested work, result of previous existing methods and new proposed methods are in tabular form in Table 1. In most of the cases, our proposed work gives better outcomes in compare to existing methods.

V. CONCLUSION

In this paper, new algorithm is described based on DWT n CBF using multi-sensor and multi-focus-based image fusion. In the given proposed work, essential detail of existing images has been integrated into a individual fused image with the accuracy.

The execution of introduced fusion method is tested on various data sets. By means of all experimental work, it is discovered that the offered methods protect much more accurate details during deleting artifacts and reported improved results than other existing image fusion methods visually.

REFERENCES


International Journal of Engineering and Advanced Technology (IJEAAT)
ISSN: 2249-8958, Volume-8 Issue-4C, April 2019

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication

Retrieval Number: D24300484C19190BEIESP
35. Sangam, Nirajan Lal, Manoj Diwakar “Texture Extraction in Video Tracking with Comparison of LBP and DWT” Presented in World Conference on Advances in Communication & Control System (CACCs) at DIT University on 6th – 8th April 2013, (pp. 369-374).