Feature Level Image Fusion

S. Saleem Malik, B.J. Shivprasad and G.B. Maruthi

Abstract— Image fusion is the process of combining information from two or more sensed or acquired images into a single composite image that is more informative and becomes more suitable for visual processing or computer processing. Image fusion fully utilizes much complementary and redundant information of the original images. The aim of image fusion is to integrate complementary and redundant information from multiple images to create a composite image that contains a better description of the scene than any of the individual source images. The objective is to reduce uncertainty, minimize redundancy in the output, and maximize relevant information pertaining to an application or a task. This paper focuses on feature level image fusion based on dual-tree complex wavelet transform (DT-CWT). A dual-tree complex wavelet transforms and watershed transform is used to segment the features of the input images, either jointly or separately, to produce the region map. Characteristics of each region are calculated and a region-based approach is used to fuse the images, region by region. The images used are already registered. Misregistration is a major source of error in image fusion.

Keywords--- Image Fusion, DTCWT, Feature-Based Fusion, Watershed Transform, Region-Based Fusion, Segmentation

I. INTRODUCTION

ONE level higher than pixel level image fusion is feature level image fusion. One technique of achieving this is with a region based fusion scheme. Initially an image is segmented to produce a set of regions. Various region properties can be calculated. The properties can be used to determine which features from which images are used in the fused image. Feature level image fusion has some advantages over pixel level image fusion as more intelligent semantic fusion rules can be considered based on actual feature in the image rather than on single pixel. Feature is very important than the actual pixel. Hence it is better to incorporate the feature information in the process of fusion [1]. Segmentation algorithm plays a vital role in region based image fusion process. Features should be segmented as single regions. Feature may split into more than one region and each region has to be treated separately. If possible, less number of regions should be generated to reduce the computational burden.

S. Saleem Malik, Assistant Professor, KVGCE, Sullia. E-mail: baronsaleem@gmail.com
B.J. Shivprasad, Assistant Professor, KVGCE, Sullia. E-mail: shivprasad1607@gmail.com
G.B. Maruthi, Assistant Professor, KVGCE, Sullia. E-mail: maruthigb123@gmail.com

A. Objective of Image Fusion

- Extract all the useful information from the source image.
- To avoid inconsistencies this will distract human observers or other subsequent processing stages.
- Temporal stability-Gray level changes in the fused image sequence should only be caused by gray level changes in the input sequence, not by the fusion process.

II. SEGMENTATION

The information flow diagram of image segmentation algorithm is illustrated in Fig-1 [2, 3]. Segmentation can be performed by applying DTCWT and watershed transform. The DT-CWT provides the six subbands oriented at ±15, ±45, ±75. Complex wavelets are shift invariant and retain the properties of scale and orientation sensitivity. The detail coefficients of dual tree complex wavelet transform is used for texture process. Denoting the detail coefficients at level , orientation by , (,), (,), and retain the complex magnitude for further analysis.

Figure 1: Information Flow Diagram of Image Segmentation Algorithm

Simple gradient calculation of complex magnitude gives rise to a double edge in the gradient magnitude as shown in Fig.2. Application of watershed algorithm produces a spurious narrow region along the boundaries is as shown in Fig.3. It can be avoided using median filter. Median filter is edge preserving smoothing filter that can suppress isolated noise without blurring sharp edges. Specifically the median filter replaces a pixel by the median of all pixels in the neighborhood and it is computationally burdensome. The solution is separable median filter and it has to be chosen with care and is given by
The order of the median filter is chosen as \( (7 + 2n) \), where \( n \) is the current level of the wavelet transform and the constant term is equal to the size of wavelet filters.

\[
S_{x,y}(x,y) = \text{MedFilt}\left( \text{MedFilt}_{\theta = 0.5\sigma}(D_{x,y}(x,y)) \right)
\]

(1)

\( S_{x,y}(x,y) \) is the half wave rectification to suppress the negative exponents as:

\[
R_{\text{half}}(\zeta) = \begin{cases} 
0 & \zeta < 0 \\
\zeta & \zeta \geq 0 
\end{cases}
\]

(5)

Texture energy \( E_{\text{tex}} \) is calculated on up sampled subbands. Texture features respond slightly larger area than the desired because of the involved spatial integration. Morphological erosion \( E \) is used to overcome the problem and strel used in this function is a square neighborhood of nine pixels. The texture energy is computed as

\[
E_{\text{tex}} = \sum_{i=0}^{i} \text{int}\left( \text{erp} \left( \frac{S_{x,y}(x,y)}{2^i} \right) \right)
\]

(6)

Where \( 2^i \) is used to correct the DC gain of the wavelet filters. Finally, the weighted sum of texture and modulated intensity gradient is computed as

\[
\text{GS}(x,y) = \frac{|IG(x,y)| + TG(x,y)}{w_i} + w_i \cdot \text{activity}(x,y) \cdot \text{w} \cdot \text{activity}(x,y) \cdot \text{activity}(x,y)
\]

(7)

Watershed Transform: Separating touching objects in an image is a difficult task. Watershed transform is often used to solve this type of problem. Watershed transform detects "catchment basins" and "watershed ridge lines" in an image by treating it as a surface where light pixels are high and dark pixels are low.

If it is possible to identify or mark the foreground objects and background locations then the segmentation using watershed transform works well. Marker-controlled watershed segmentation has the following steps [4]:

- Compute a segmentation function. This is an image whose dark regions are the objects to be segmented.
- Compute foreground markers. These are connected blocks of pixels within each of the objects.
- Compute background markers. These are pixels that are not part of any object.
• Modify the segmentation function so that it has minima at the foreground and background marker locations.
• Compute the watershed transform of the modified segmentation function.

Here we used the watershed segmentation using the distance transform. The distance transform of a binary image is a simple concept. It is the distance from every pixel to the nearest non-zero valued pixel. Note that 1-valued pixels have a distance transform value of 0. The distance transform can be computed by using the matlab function 'bwdist'.

**Joint Segmentation:** Let the weighted sum of the texture and modulated intensity gradients of the infrared image is \( GS_{ir}(x, y) \) and of the visible image is \( GS_{vi}(x, y) \). These two individual gradients are combined as:

\[
GS = \frac{GS_{ir}(x, y)}{\text{median}(GS_{ir}(x, y))} + \frac{GS_{vi}(x, y)}{\text{median}(GS_{vi}(x, y))}
\]

(8)

The marker controlled watershed algorithm is applied on this combined gradient image, Fig 6(a) and 6(b) shows the segmentation of vis image and ir image. Fig 6(c) shows the ridge lines produced from watershed segmentation. Fig 6(d) and Fig 6(e) shows ridge lines superimposed on vis and ir image and Fig 6(f) shows the joint segmented image.

![Figures 6(a) to 6(f)](image)

III. **FEATURE LEVEL IMAGE FUSION**

The block diagram of feature level image fusion is as shown in Fig-7. The input images are joint segmented by using DTCWT. The joint segmented image is shown in Fig-6 (f) is used as the segmentation map. By using the segmentation map we are calculating the salient feature like standard deviation, if the standard deviation of the segmented part of the input image \( I_1 \) is greater than the standard deviation of the segmented part of the input image \( I_2 \) then the fused image part comes from input image \( I_1 \), otherwise it is from input image \( I_2 \).

IV. **WAVELET BASED FUSION**

The Dual Tree Complex Wavelet Transform (DT-CWT) provides both good shift invariance and directional selectivity. It has ability to differentiate positive and negative frequencies and produce six subbands oriented in ±15, ±45, ±75. The DT-CWT gives the perfect reconstruction as the filters are chosen from a perfect reconstruction bi-orthogonal set. The block diagram of wavelet based fusion is as shown in Fig-8.After applying the DT-CWT to the input images produces the wavelet coefficients i.e. approximation and detailed coefficients.

The approximation coefficients are fused by the average fusion rule and for the detailed coefficients are combined using the maximum-selection fusion rule to produce a single set of coefficients corresponding to the fused image. The maximum selection scheme selects the largest absolute wavelet coefficient. After applying the Inverse Dual Tree Complex Wavelet Transform (IDT-CWT) to the fused coefficients produces the fused image. The fused image has the more information than the input images.

![Figures 8](image)

V. **PERFORMANCE METRICS**

**Entropy:** Entropy is used to measure the information content of the image. Entropy is sensitive to noise and other
unwanted rapid fluctuations. An image with high information content have high entropy. We calculated the entropy by using the matlab function ‘entropy’.

**Standard Deviation:** Standard deviation composed of signal and noise parts. It is more efficient in the absence of noise. It is used to measure the contrast in the fused image. An image with high contrast would have a high standard deviation. We are calculating the standard deviation by using the matlab function ‘std2’.

**Spatial Frequency:** The frequency in the spatial domain indicates the overall activity level in the fused image and is given by

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

Where \( RF\) - Row frequency.
\( CF\) - Column frequency.

\[
RF = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [I_f(x,y)-I_f(x,y-1)]^2}
\]

\[
CF = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [I_f(x,y)-I_f(x-1,y)]^2}
\]

**Cross Entropy:** Cross-entropy evaluates the similarity in information content between input images and fused images. Fused image and input image containing the same information would have low cross entropy. The overall cross entropy of source image \(I_1, I_2\) and the fused image \(I_f\) is

\[
CE(I_1, I_2; I_f) = \frac{CE(I_1; I_f) + CE(I_2; I_f)}{2}
\]

where \( CE(I_1; I_f) = \sum_{i,j} h_{i,j}(I_1) \log \left( \frac{h_{i,j}(I_1)}{h_{i,j}(I_2)} \right) \) and \( CE(I_2; I_f) = \sum_{i,j} h_{i,j}(I_2) \log \left( \frac{h_{i,j}(I_2)}{h_{i,j}(I_1)} \right) \)

**Mutual Information:** It measures the degree of dependence of the two images. A larger measure implies better quality. If the joint histogram between \(I_1(x,y)\) and \(I_f(x,y)\) is defined as \(h_{i,j}(I_1, I_f)\) and \(h_{i,j}(I_f, I_2)\). Then the mutual information between source and fused image is given by

\[
MI = MI_{I_1} + MI_{I_2}
\]

where \( MI_{I_1} = \sum_{i} \sum_{j} h_{i,j}(I_1) \log \left( \frac{h_{i,j}(I_1)}{h_{i,j}(I_1)h_{i,j}(I_f)} \right) \)

\[
MI_{I_f} = \sum_{i} \sum_{j} h_{i,j}(I_f) \log \left( \frac{h_{i,j}(I_f)}{h_{i,j}(I_1)h_{i,j}(I_2)} \right)
\]

VI. RESULT

The image to be fused (vis.bmp and ir.bmp) is shown in Fig-9. The intensity gradient \(IG(x,y)\) is shown in Fig-2. Edges are highlighted than smoothing areas. The segmented image using the watershed algorithm is shown in Fig-3. Modulated intensity image after suppressing the edges within the textured region is shown in Fig-5 and the textured gradient image is shown in Fig-4. The edge of the textured region is high contrast than other non-textured regions. The man is highlighted. Fig-6 shows the weighted combination of texture and modulated gradients. The watershed algorithm directly often results in over segmentation. This over segmentation can be avoided using marker controlled watershed algorithm. There should be a single segmentation map for both the images. It could be possible by combining the individual (uni-model) segmentation maps. The segmentation map of the infrared image is shown in Fig-6 (a) and the segmentation map of visible image is shown in Fig-6 (b). The combined segmentation map is shown in Fig-6 (c). One can observe that there are many segments and hence it take more computational time. As mentioned earlier, this problem can be solved by joint segmentation, the joint segmented image shown in 6 (f). Figure 9 (c) shows the fused image by the feature level image fusion and figure 9 (d) gives the wavelet based fused image. After applying the linear spatial filter by using the matlab function imfilter to the fused image gives the better result, table 1 gives comparison of performance metric with and without filter, for the other set of images results are shown in Fig10.
<table>
<thead>
<tr>
<th>Input Image</th>
<th>Performance metric</th>
<th>Wavelet based fusion (using SWT)</th>
<th>Wavelet based fusion (using DT-CWT)</th>
<th>Feature level fusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without filter</td>
<td>With filter</td>
<td>Without filter</td>
<td>With filter</td>
</tr>
<tr>
<td>Vis and IR</td>
<td>Entropy</td>
<td>6.489</td>
<td>6.478</td>
<td>6.918</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>22.54</td>
<td>995.9</td>
<td>27.25</td>
</tr>
<tr>
<td></td>
<td>SF</td>
<td>6.653</td>
<td>152.3</td>
<td>9.983</td>
</tr>
<tr>
<td></td>
<td>Cross Entropy</td>
<td>0.1434</td>
<td>1.899</td>
<td>0.136</td>
</tr>
<tr>
<td></td>
<td>Mutual Information</td>
<td>2.1309</td>
<td>2.0</td>
<td>2.119</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>35.25</td>
<td>1542</td>
<td>47.21</td>
</tr>
<tr>
<td></td>
<td>SF</td>
<td>11.84</td>
<td>314.4</td>
<td>19.32</td>
</tr>
<tr>
<td></td>
<td>Cross Entropy</td>
<td>1.7204</td>
<td>2.816</td>
<td>2.477</td>
</tr>
<tr>
<td></td>
<td>Mutual Information</td>
<td>2.487</td>
<td>2.078</td>
<td>2.198</td>
</tr>
<tr>
<td></td>
<td>Standard deviation</td>
<td>40.01</td>
<td>1753.8</td>
<td>51.13</td>
</tr>
<tr>
<td></td>
<td>SF</td>
<td>18.02</td>
<td>204.2</td>
<td>23.91</td>
</tr>
<tr>
<td></td>
<td>Cross Entropy</td>
<td>0.977</td>
<td>1.959</td>
<td>0.976</td>
</tr>
<tr>
<td></td>
<td>Mutual Information</td>
<td>2.254</td>
<td>2.0</td>
<td>2.196</td>
</tr>
</tbody>
</table>

**VII. CONCLUSION**

Images are fused by using feature level image fusion and wavelet based fusion. In feature level image fusion images are segmented and the corresponding segmented parts are fused. In wavelet based fusion the wavelet coefficients are fused. Calculated the entropy, standard deviation, spatial frequency, Cross entropy and Mutual information for various images, feature level image fusion gives better results than the wavelet based fusion.

REFERENCES


