Pixel- and Fetaure- Level Image Fusion Techniques

Venkatesh Chandra Naik M*, Shreedarshan* and VPS Naidu$

*Dept of ECE, MS Ramaiah Institute of Technology, Bangalore.

$MSDF Lab, FMCD, CSIR-National Aerospace Laboratories, Bangalore.

vpsnaidu@gmail.com

Abstract— **Image fusion is a process combining two or more images in to a single composite image that contains complete information for further processing. In this paper, feature level image fusion was developed and evaluated and the results were compared with pixel level image fusion algorithms using fusion quality evaluation metrics. It was concluded that feature level image fusion provides better fusion results at the cost of execution time. Among the pixel level image fusion algorithms, it was observed that DT-CWT based fusion algorithm provides good results.**

Keywords- Image Fusion, DTCWT, Feature-Based Fusion, Watershed transform, Region-Based Fusion, Segmentation.

I. INTRODUCTION

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.

There are three levels of image fusion viz., Pixel level, feature level and decision level. Pixel level image fusion is studied by many researchers [1,2]. Feature level image fusion is one level higher than pixel level image fusion. One technique for achieving feature level image fusion is with a region based fusion scheme. Initially an image is segmented to produce a set of regions. Various region properties can be calculated and used to determine which features from each original image 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. Hence it is better to incorporate the feature information in the process of fusion [3,4].

Segmentation algorithm plays a vital role in region based image fusion process. Preferably each feature should be segmented as single region in the image. But often some feature may split into more than one region; in that case each region has to be treated separately. If possible, less number of regions should be generated to reduce the computational burden. Segmentation can be done either separately (uni-model segmentation) or jointly (joint segmentation). Uni-model segmentation method may create many regions than joint segmentation since different images have different features. In this paper, joint segmentation using dual tree complex wavelet transform (DTCWT) is used to generate joint segmentation map.

It is assumed that the images to be fused are already registered. In this paper, feature level image fusion algorithm is implemented and studied and the results are compared to pixel level image fusion algorithms available in the literature [1,2].

II. PIXEL LEVEL IMAFE FUSION

Pixel level image fusion is performed using wavelets by many researchers. Wavelets are suffered with shift variant and edges are not properly produced in the fused image. Image fusion by stationary wavelet transform is studied. It is shift invariant but it does not produce the edge information properly on the fused image. It could be because it does not consider the directional edge information. These short comes are evaded using dual tree complex wavelet transform. 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 dual tree complex wavelet transform based fusion is as shown in Fig-10.Maintaining the Integrity of the Specifications

The source images I_1 and I_2 are decomposed into approximation and detailed coefficients at the required level using DT-CWT. The coefficients of both images are

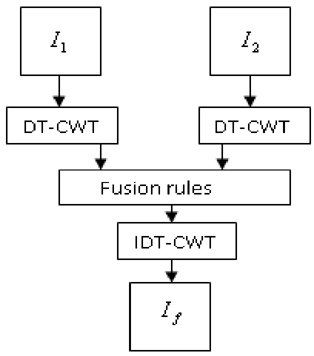


Fig-1: Information flow diagram of pixel level image fusion algorithm

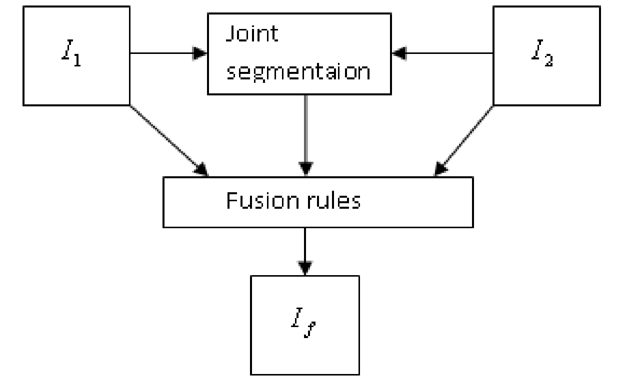


Fig-2 Information flow diagram of feature level image fusion algorithm

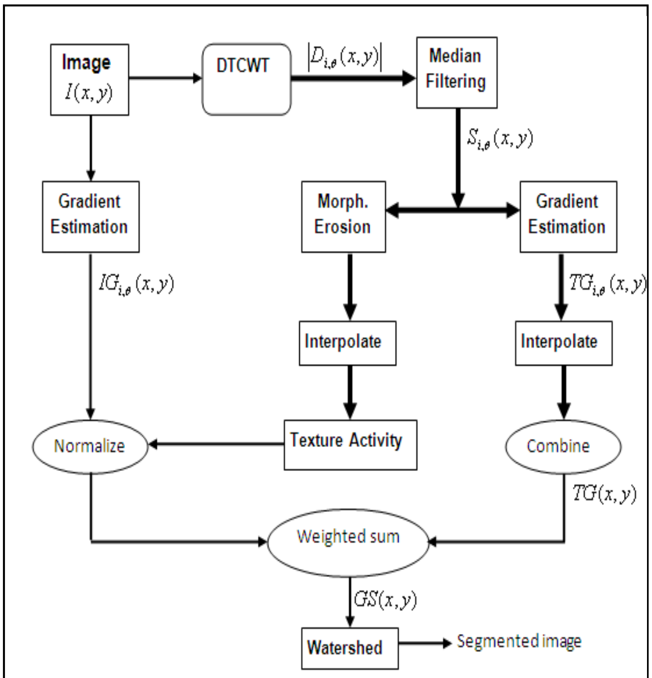


Fig-3 Information flow diagram of image segmentation algorithm

subsequently combined using the fusion rule. The fused image I_f is then obtained by taking the inverse DT-CWT to the fused coefficient as [6].

$$I_f = T^{-1}[\phi\{T(I_1), T(I_2)\}], \text{ where } T \text{ is the DT-CWT}$$

The fusion rule (ϕ) used here is, approximation coefficients are fused by average the low frequency coefficients and for detailed coefficients are fused by taking the largest magnitude of the high frequency coefficients.

III. Feature Level Image fusion

The information flow diagram of feature level image fusion algorithm is shown in Fig-2. The registered images are used to generate the joint segmentation. This segmentation map is used to extract the segments from the images to be fused and salient features are calculated [5,6]. Based on salient feature, a particular segment is selected and placed in the fused image. The details are presented in successive sections.

A. Image Segmentation

The information flow diagram of image segmentation algorithm is illustrated in Fig-1 [3-6].

Texture representation: Generally, Gabor filter had been used for texture representation because of facts from psychophysical experiments. Human visual system decomposes the visual field into perceptual channels. These channels are evenly spaced in angle. Gabor filter representation is computationally burdensome. Complex wavelets are alternative to Gabor functions for texture analysis. 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 i, orientation θ by $D_{i,\theta}(x, y)$ and retain the complex magnitude $|D_{i,\theta}(x, y)|$ for further analysis.

Texture post processing: Simple gradient calculation of complex magnitude gives rise to a double edge in the gradient magnitude. Application of watershed algorithm produces a spurious narrow region along the boundaries. It can be avoided using median filter before gradient operator. Median filter is edge preserving smoothing filter and it is computationally burdensome. The solution is separable median filter and it has to be chosen with care

$$S_{i,\theta}(x, y) = MedFilt_i\left(MedFilt_{(\theta+0.5\pi)}\left(|D_{i,\theta}(x, y)|\right)\right)$$

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.

Texture gradient: Gaussian derivative function is used as gradient operator. The texture gradient magnitude of each subband is

$$TG_{i,\theta}(x, y) = \sqrt{\left(S_{i,\theta}(x, y) * G_x^{'}\right)^2 + \left(S_{i,\theta}(x, y) * G_y^{'}\right)^2}$$

Where $G_x^{'}$ and $G_y^{'}$ are the Gaussian partial derivative filters in x and y directions and *denotes convolution.

Weighted sum of the magnitudes is

$$TG(x, y) = \sum_{i,\theta} \operatorname{int}erp\left(w_{i,\theta} \mathrm{x} \mathrm{TGH}_{i,\theta}(\mathrm{x}, \mathrm{y})\right)$$

Where $TGH_{i,\theta}(x, y) = \dfrac{TG_{i,\theta}(x, y)}{\max\left(TG_{i,\theta}(x, y)\right)}$

$$w_{i,\theta} = \frac{N_i}{\sum_{x,y} TGH_{i,\theta}(x, y)^2}$$

N_i : number of pixels in subband at level i

$\operatorname{int}erp()$: interpolation function and performed separably

Gradient Combination: Texture and intensity gradients are combined to get final gradient capturing all perceptual edges in the image. The combined gradient will be dominated by intensity gradient in smooth regions and texture gradient in textured regions. The activity measure is

$$activity(x, y) = \exp\left(R_{half}\left(\frac{E_{tex}(x, y)}{\alpha} - \beta \right) \right)$$

Where $\alpha = 2$ & $\beta = 7$ are chosen based on intuition

$R_{half}(\varsigma)$ is the half wave rectification to suppress the negative exponents as:

$$R_{half}(\varsigma) = \begin{cases} 0: & \varsigma < 0 \\ \varsigma: & \varsigma \geq 0 \end{cases}$$

Texture energy E_{tex} is calculated on up sampled sub-bands. 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

$$E_{tex} = \sum_{i,\theta} \operatorname{int} erp\left(E\left(\frac{S_{i,\theta}(x, y)}{2^i} \right) \right)$$

The denominator 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

$$GS(x, y) = \frac{|IG(x, y)|}{activity(x, y) \mathrm{x} w_I} + \frac{TG(x, y)}{w_T}$$

Where w_I is to be four times the median intensity gradient

w_T is the median value of the texture gradient

Watershed transform[7]: 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.

B. Joint Segmmentation

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:

$$GS = \frac{GS_{ir}(x, y)}{median(GS_{ir}(x, y))} + \frac{GS_{vi}(x, y)}{median(GS_{vi}(x, y))} \qquad (8)$$

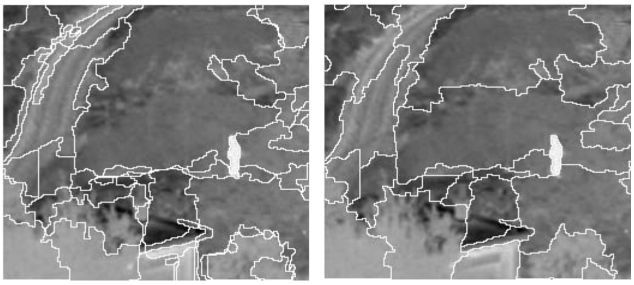
The marker controlled watershed algorithm is applied on this combined gradient image.

C. Fusion

Using the segmentation map, salient feature like standard deviation is computed for all the segments. If the standard deviation of the segmented part of the image $\mathbf{I_1}$ is greater than the standard deviation of the corresponding segmented part of the image $\mathbf{I_2}$, then the fused image part comes from input image $\mathbf{I_1}$, otherwise it is from image $\mathbf{I_2}$. This complete process is done

(a) (b)

Fig-4: Images to be fused (a. visival and b.infrared)



(a)Uni-model segmentation (b) Joint segmentation

Fig-5: Segmented maps

in spatial domain. Instead of taking the segmented images, the corresponding wavelet coefficients are considered. IDT-CWT is applied on the fused coefficients to get the fused image.

IV. RESULTS AND DISCUSSIONS

The images to be fused are shown in Fig-4. Fig-4a is a visual image. Bushes and fencing are very clear in this image but man is not visible. Fig-4b is an infrared image. Man is very clear but bushes & fencing are not clearly visible. The uni-model segmentation and joint segmentation maps are shown in Fig-5. It is observed that uni-model segmentation generates more number of segments as expected. Fused image using DT-CWT is shown in Fig-6. The man and fencing & bushes are clearly visible in the fused image. The fused images using feature level image fusion algorithm are shown in Fig-7. The fusion quality evaluation metrics [8] are shown in Table-1. The entries in the table with bold is the parameter. It is observed that feature level image fusion done in spatial domain provides good results at the cost execution time. Fusion quality is improved by considering the number of decomposition levels in the fusion

Fig-6: DT-CWT Fused image

(a). (b)

Fig-7 Fused images (Fusion is done in a. spatial domain and b. transform domain)

process. As expected, DWT [1] and SWT [2] based image fusion algorithms are not performed as DT-CWT as shown in Table-1.

V. CONLUSION

Feature level image fusion (FLIF) algorithms (both in spatial and frequency domain) were developed and evaluated using fusion quality evaluation metrics. The images to be fused are passed through joint segmentation algorithm to get the common segmentation map. Salient feature viz., standard deviation is computed for corresponding segments (both the images) and the segment was chosen based on best salient feature. It was done for all the segments. To compare the performance of this algorithm, three different pixel level image fusion algorithms viz., DWT, SWT and DT-CWT were also implemented and evaluated.

From this study, it is concluded that FLIF provides a good fused image at the cost of execution time and also it requires a good segmentation map. Most of the time DT-CWT provides good fusion results since it considers the edge information in six directions. In all cases, the DWT based pixel level image fusion algorithm does not provide good results since it does not consider the edge information and lack of shift invariant. SWT based image fusion algorithm provides good results in some cases where there are no much edges in the images to be fused, it is shift invariant and it does not consider the directional edge information.

Only standard deviation is used as a salient feature to select the best segment. More salient features can be used along with some fuzzy logic or neural networks to choose the best segment. It can also be extended for color image fusion. There is a scope to develop robust joint segmentation algorithm and hence better feature level image fusion.

REFERENCES

[1] VPS Naidu and J.R. Raol "Pixel-Level Image Fusion using Wavelets and Principal Component Analysis A Comparative Analysis", Defense Science Journal,Vol 58, No.3, pp.338-352, May 2008.

[2] http://www.mathworks.fr/matlabcentral/fileexchange/28923-image-fusion

[3] J.J. Lewis et al, "Region-based Image Fusion using Complex Wavelets", Proceeding of the 7th International Conference on Information Fusion, p.555-562, 2004.

[4] Robert J. O' Callaghan and David R. Bull, "Combined Morphological-Spectral Unsupervised Image Segmentation", IEEE Trans. On Image Proc., Vol. 14, No.1, pp. 49-62, Jan 2005.

[5] VPS Naidu, "A Study on Joint Segmentation Technique", MSDF Report No.: 1018/ATR10, 7th Nov. 2010.

[6] http://www.mathworks.fr/matlabcentral/fileexchange/28418-multi-modal-image-segmentation

[7] http://www.mathworks.com/products/imageb/indexb.html?sec=tour&subsec=demo&file=/products/demos/shipping/images/ipexwatershed.html

[8] VPS Naidu, "Discrete Cosine Transform –Based Image Fusion", Defense Science Journal, vol. 60, No 1,January 2010.

Table-1: Fusion quality evaluation metrics

Fusion Method		entropy	Standard deviation	Spatial frequency	Cross entropy	Mutual information	Execution time (sec)
Algorithm	No. of levels						
DWT [1]	1	6.446	22.21	5.509	0.158	2.132	**0.40**
	3	6.420	22.04	5.72	0.156	2.09	0.422
	3	6.462	22.04	7.14	0.146	2.142	0.424
	4	6.475	22.46	7.461	0.140	2.13	0.468
SWT [2]	1	6.483	22.44	6.352	0.145	2.131	0.546
	2	6.488	22.53	6.652	0.143	2.130	0.665
	3	6.495	22.60	6.861	0.142	2.130	0.796
	4	6.497	22.67	7.021	0.141	2.129	1.236
DT-CWT	1	6.471	22.40	6.216	0.147	2.132	0.333
	2	6.503	22.92	7.712	0.144	2.126	0.371
	3	6.510	24.61	9.228	0.165	2.122	0.382
	4	6.477	27.24	**9.970**	0.136	2.119	0.402
FL	SD	**6.918**	**31.35**	8.709	0.254	**2.598**	32.76
	WD	6.621	23.90	6.521	**0.126**	2.120	26.8