International Journal of Scientific Engineering and Research (IJSER)
ISSN (Online): 2347-3878
Index Copernicus Value (2015): 62.86 | Impact Factor (2015): 3.791

Decision level based Image Fusion using Wavelet Transform and Support Vector Machine

Nalini B. Kolekar[1], R. P. Shelkikar[2]

[1]PG Student, Department of E &TC Engineering, TPCT's College of Engineering, BAMU Unversity, Osmanabad, India

[2]Assistant Professor, M.E. (Electronics) Dept. of E&TC Engineering, TPCT's College of Engineering, Osmanabad, India

Abstract: ***Image Fusion is a process of combining the relevant information from a set of images, into a single image, wherein the resultant fused image will be more informative and complete than any of the input images. The fusion methods for combining infrared images with visible spectrum images concentrate heavily on the surveillance and remote sensing applications Decision level based image fusion using Shift Invariant Discrete Wavelet transform and SVM is presented here. Shift invariance of the wavelet transform is important in ensuring robust sub-band fusion. Support Vector Machine is trained to select the coefficient blocks with significant features, extracted from the SIDWT coefficients. The performance of the proposed scheme is evaluated by various quantitative measures like Mutual Information (MI), Standard deviation, and Entropy (EN). Visual and quantitative analysis show the effectiveness of the proposed scheme in fusing multimodality images.***

Keywords: Image fusion, Shift invariant discrete wavelet transform, SVM classifier

1. Introduction

With the recent rapid developments in the field of sensing technologies multisensory systems have become a reality in a growing number of fields such as remote sensing, medical imaging, machine vision and the military applications. Multi-sensor data often presents complementary information about the region surveyed, so image fusion provides an effective method to enable comparison and analysis of such data. With the recent fast developments in the field of sensor technology, there has been a growing interest in the use of multiple sensors to increase the capabilities of intelligenachines and systems in a number of fields such as surveillance, remote sensing, medical imaging, machine vision and military. Multi-sensor data often presents complementary information about the region surveyed, so image fusion provides an effective method to enable comparison and analysis of such data. Image fusion process effectively integrates the images taken by multiple sensors and produces a single image extracting all the relevant information from the source images. The use of multiple sensors result in large amount of data and image fusion reduces the amount of data and results in new images which are more suitable for human visual and machine perception. The accuracy of a system can be improved through image fusion exploiting the redundancy given by the multiple camera inputs. The different images to be fused may come from sensors of the same type or from different types of sensors, taken at the same time or at different times. the ability to combine complementary information from a range of distributed sensors with different modalities can be used to provide enhanced performance for visualization, detection or classification tasks [1]. Fusion of visible and infrared (IR) images and video sources is becoming increasingly important for surveillance purposes. The fused image obtained by combining features of visible and infrared source images has enhanced information which enables improved detection and unambiguous localization of a target. Image fusion techniques can be generally classified into Spatial and Transform domain techniques.

In spatial domain techniques, we directly deal with the pixel value of an image. The pixel values are manipulated to achieve desired result. The fusion methods such as averaging, Brovey method, principal component analysis (PCA) and IHS based methods fall under spatial domain approaches [1]. This often produces serious side effects such as reduced contrast. The multi resolution analysis has become a very useful tool for analyzing remote sensing images. The most commonly employed mutiresolution decomposition methods are the Pyramid Transform and the Discrete Wavelet Transform (DWT)[3]. A pyramid is a simple structure for representing an image at more than one resolution. The pyramid transform can provide information on the sharp contrast changes and human visual system is sensitive to these sharp contrast changes Some other fusion methods are also there such as Laplacian pyramid based, Curvelet transform based etc. There are different variations of pyramid transform that include Laplacian (LP) , Gradient , the Morphological [8] and Contrast pyramid [5]. These methods show a better performance in spatial and spectral quality of the fused image compared to other spatial methods of fusion.

The actual fusion process can take place at different levels of information representation i.e. pixel, feature and decision level. Pixel level is a low level of fusion which is used to analyze and combine data from different sources before original information is estimated and recognized. Pixel level fusion works directly on the raw pixels obtained from imaging sensors. Feature level is a middle level of fusion which extracts important features from an image like shape, length, edges, segments and direction. Decision level is a high level of fusion. it comprises of sensor information fusion, after each sensor has processed an image and a preliminary determination has been made(entity's location, attributes, and identity).And merges the interpretations of different images obtained after image understanding. Decision level fusion uses the outputs of initial object detection and classification as inputs to the fusion algorithm. The image obtained after fusion should contain maximum required information of the input image and may not include any distortions or loss of information in fused image.

This paper proposes a multimodal image fusion method using Shift Invariant Discrete Wavelet Transform (SIDWT) [2] and Support Vector Machines (SVM). SIDWT are used for multiresolution decomposition and a trained SVM is used to select salient features in the image. Support Vector Machines are a set of related supervised learning methods that analyze data and recognize patterns, used for classification [6].

2. Image fusion using SIDWT

The Discrete Wavelet Transform (DWT) was successfully employed in the field of image processing with the introduction of Mallat's algorithm The discrete wavelet transform (DWT) is a powerful tool for multiresolution analysis [9]. In this the input image is decomposed into a set of wavelet decomposition levels. In image fusion using discrete wavelet transform, first the DWT of each source image are computed. For two source images I1, and I2, DWT based image fusion algorithm can be described by,

F= ω-1(φ (I1), ω (I2)) (1)

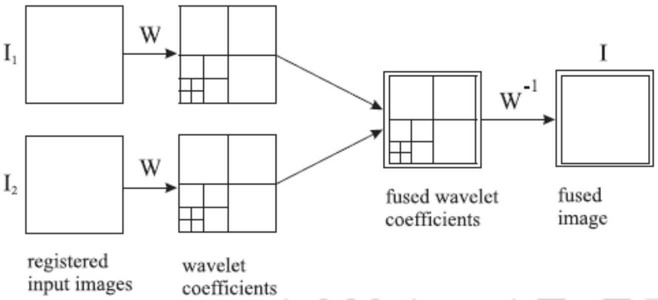


Figure 1: Image Fusion using discrete wavelet transforms

Where ω is the DWT, ω -1 the inverse DWT, φ the fusion rules, and F the fused image. The source images are transformed from spatial domain to wavelet coefficients by ω, wavelet coefficients are combined by rules φ, and the combined wavelet coefficients are transformed to fused image F by ω-1. In the first step DWT decomposition is applied on all input images, resulting in a multiscale edge representation of the input imagery. Then a composite multiscale edge representation is built from the selection of most salient wavelet coefficients of the input imagery. The selection scheme can be a simple choose-max of the absolute values or a more sophisticated area based energy computation. In the final step, the fused image is computed using inverse DWT on the composite wavelet representation. The discrete wavelet transform is not shift-invariant due to the underlying down-sampling process. Hence, in practice, their performance quickly degrades when there is slight object movement or when the source images cannot be perfectly registered [2]. To overcome this problem a Shift Invariant Discrete Wavelet Transform (SIDWT) can be used [13]. In this section, the decomposition of images using Shift Invariant Discrete Wavelet Transform (SIDWT) is discussed. The SIDWT uses an over complete wavelet decomposition by discarding DWT's down-sampling process [13]. The advantage of the method is the improved temporal stability and consistency of the fused sequence compared to other existing fusion methods [12].

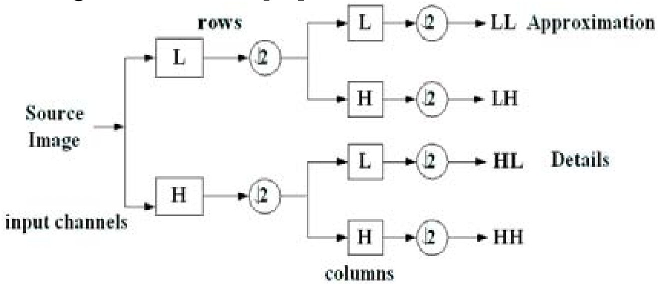


Figure 2: Two-dimensional sub band coding algorithm for DWT

In the decomposition phase of 2-D SIDWT, each row of the image is separately filtered by H and G filter. The resulting row- transformed image is then similarly filtered in the column direction, finally yielding four sub-bands at the first decomposition level (i=1). The three detail sub-bands, LH, HL, HH, contain the vertical, horizontal, and diagonal high frequency information, respectively, while the approximation sub-band LL is a low pass filtered version of the original image. This approximation sub-band is subsequently passed to the next level for further sub-band decomposition. The filter coefficients at each level are modified by up sampling the filter coefficients by a factor of 2(i-1) in the ith level of the algorithm. Thus, a SIDWT with 'n' decomposition levels will have a total of 3n+1 frequency sub-bands, all of them are of the same size. The resultant signal representation is both aliasing free and translation- invariant.

3. SVM Classifier

A SVM is a classification method that discriminates between two classes, by fitting an Optimal Separating Hyper plane (OSH) (Gunn, 1998). It actually maximizes the margin between two classes of training data.

The foundations of Support Vector Machine have been developed by Vapnik (1995) and are gaining popularity due to many attractive features and promising empirical performance. [6] The term SVM is typically used to describe classification and regression with support vector methods [6].The SVM performs a mapping of the training patterns to a feature space from the input space and constructs a hyper plane that separates the various classes with maximum margin. An SVM training algorithm builds a model that assigns new examples into one category or the other. The separation of data can be either linear or non-linear

k(xi,xj) is the SVM kernel function that performs the mapping, 's are support vectors, 0 $\leq \sigma_i$ C and $\sum_{i=1}^{n} \alpha i y i =$ for i = 1,2,...,n , C is a regularization parameter to be defined by the user. Higher the value assigned to C, higher the penalty assigned to the margin errors.

One of the popular kernels used in SVM is the Radial Basis Function (RBF) kernel, which has a parameter known as Gaussian width. In this study, Gaussian Radial Basis Function (GRBF) is used and is given by

K(xi,xj) = exp -(|xi-xj|2/2 σ (2)

Where, xi and xj denote the training patterns given.

To perform classification, the input data is divided into two sets: training and test. From the input data, several features or attributes are computed which form the input variables. The training set at each instance consists of a target value called class labels and the observed input variables. In the training phase, the SVM is trained using the training data generated. In the testing phase, the trained SVM determines the target values of the test data based on its attributes.

4. Image Fusion Using SIDWT and SVM

In this method, fusion of visual and infrared (IR) images is performed using Shift Invariant Wavelet transform and Support Vector Machines.

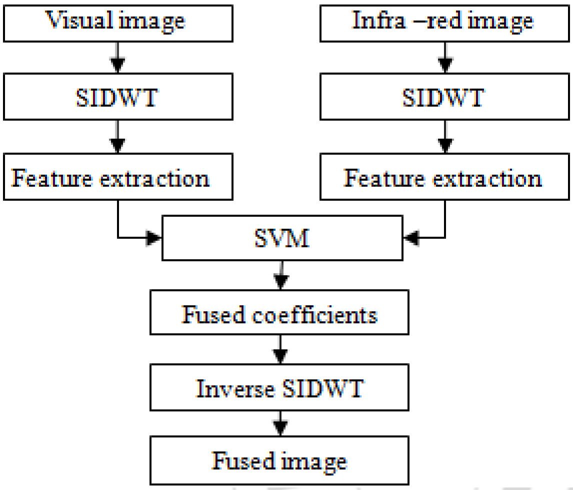


Figure 3: Block diagram

Fig.3 gives the block diagram of the proposed system.

1) Here only two source images visual and infrared are considered for fusion, each source image is decomposed by SIDWT to ‘n’ levels into one approximation sub band and 3n details sub band(3n+1).The approximation subband after ‘n’ level decomposition will not contain sharp image details and hence not used for feature extraction. In the proposed work, four decomposition levels are used.
2) To train the SVM, extract the features. For extracting features, all detail sub band coefficients are divided into non-overlapping blocks of fixed size and three features energy, entropy and standard deviation are computed for each block using the following equations

Energy $$E = \sum_{i,j=1}^{N} I^2(i,j) \qquad (3)$$

Entropy $$E= -\sum I(i,j)\log 2I(i,j) \qquad (4)$$

Standard deviation

$$SD = \sqrt{\frac{1}{N*M}\sum_{i=1}^{N}\sum_{j=1}^{M}|I(i,j)-\mu|^2} \qquad (5)$$

Where I (i, j) are the SIDWT coefficients and μ is the mean of each coefficient block.

3) Train the SVM using the features extracted in step 2 to determine whether the wavelet coefficient block from visual image or infrared image is to be used. The SVM target output is taken as +1 if the coefficient block from visual image has significant features and -1 otherwise. The images are trained with different block sizes.
4) If the classes value is +1 corresponding coefficient block from Visual image will be selected and for -1 corresponding coefficient block from IR image will be selected.
5) The fused image is obtained by performing inverse SIDWT from the selected coefficients and the Parameters of fused image are calculated.

5. Result and Discussion

In this proposed method, visual and infrared images are fused using SIDWT and SVM. This experiment tested on different sets of multimodal images. The images selected have different illumination levels, various object sizes and different distances to the sensors [16]. As SVM is used for classification, this proposed image fusion works at decision level. For SVM we use the MATLAB Bioinformatics toolbox with the Gaussian Radial Basis function (GRBF) kernel.GUI of the proposed system is shown in fig 4.This GUI diagram shows the visual and infrared image with its wavelet transformed image. In this work, the experiment is carried out with coefficient blocks of sizes 4x4, 8x8, 16x16 and 32x32. The SVM is trained with patterns taken from both visual and IR images. Then the trained SVM is used to perform fusion. Classification is performed at all levels, for all coefficient blocks. Figures 5 and 6 shows the results of fusion using the proposed method for different block size. The performance of the proposed method is tested using the fusion metrics Energy, Entropy, Mutual Information (MI), PSNR and MSE presented [14].

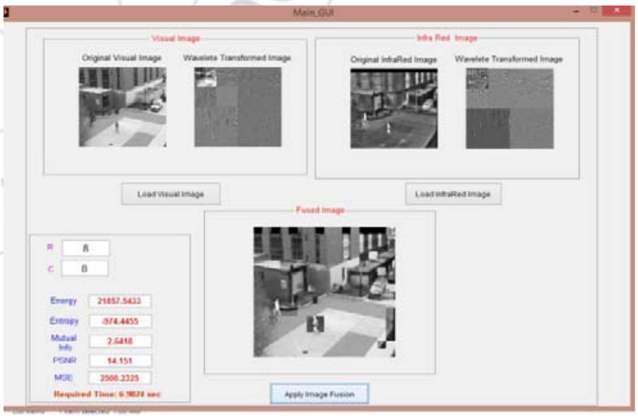


Figure 4: Fusion result for visual and IR image

(a) (b)

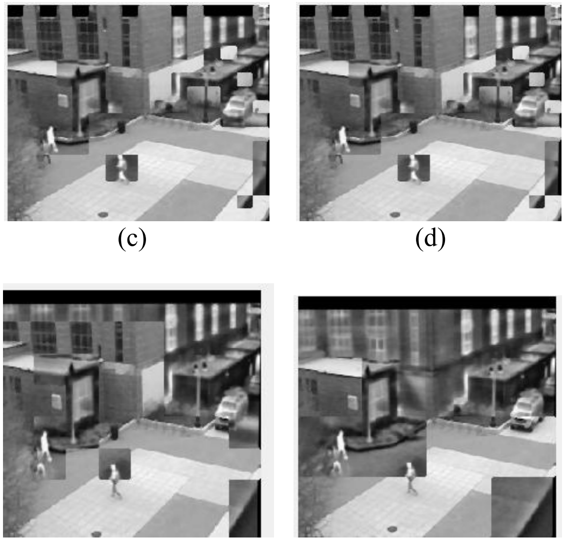


Figure 5: Fusion results of road image (a) source image (visual) (b)Source image (IR) (c)-(f) fused images of different block size (c) 4*4 block size (d) 8*8 block size (e) 16*16 block size(f)32*32 block size

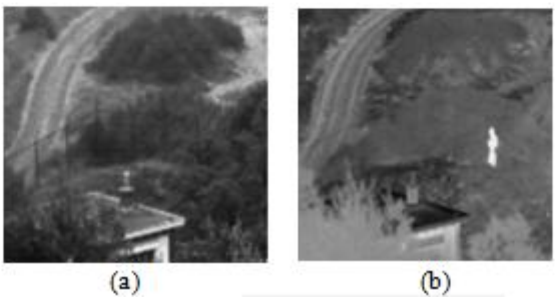
Table 1: Fusion results of road image

Block size	*4*4*	*8*8*	*16*16*	*32*32*
Energy	2342.11	21857.54	8342.11	8379.37
Entropy	-1021.81	-974.44	-538.81	-543.84
MI	3.1635	2.6418	2.2139	1.7917
PSNR	17.08	14.151	22.377	21.874
MSE	1272.93	2500.23	376.1632	422.3582
Time(sec)	28.871	6.9824	1.9522	0.6382

Fig. 7 shows a plot of fusion metric for different coefficient block sizes for the UN camp image. Fig. 8 shows a plot of execution time versus coefficient block sizes. Experiments were carried out with different block sizes for the source images shown. It is observed that as the block size increases the quality metric increases and the execution time decreases.

Table 2: fusion metrics for UN camp image

-	*4*4*	*8*8*	*16*16*	*32*32*
Energy	2342.11	21857.54	19455.3	18383.7
Entropy	-1021.81	-974.44	-888.80	-855.56
MI	3.1635	2.6418	2.2139	1.7917
PSNR	17.08	14.151	11.9482	11.29
MSE	1272.93	2500.23	4151.99	4825.44
Time(sec)	28.871	6.9824	1.9105	0.6626



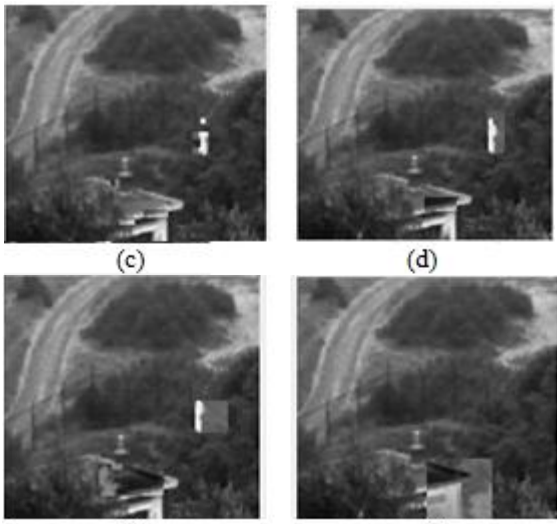


Figure 6: Fusion results of UN camp image (a) source image (visual) (b)Source image (IR)(c)-(f) fused images (c) 4*4 block size (d) 8*8 block size (e)16*16 block size (f)32*32 block size

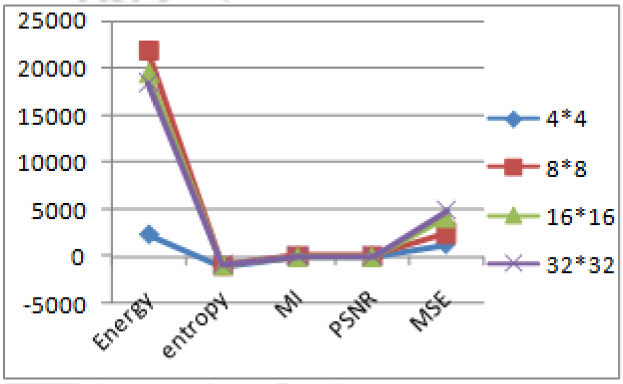


Figure 7: Graphical representation of fusion metrics

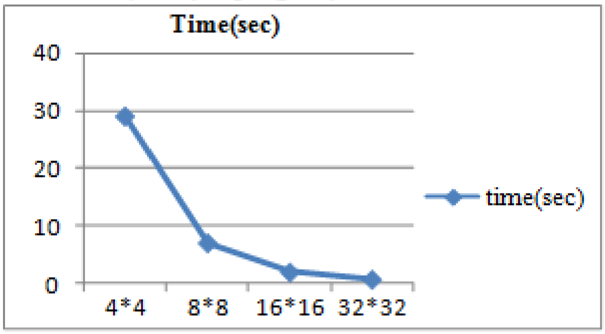


Figure 8:.Block size vs. execution time

6. Conclusion

Multisensor image fusion combines different sources of image information into a composite image which is more suitable to human perception and computer analysis. This paper presents, image fusion method suitable for multimodal images using SIDWT and SVM The wavelet image fusion technique can improve the spatial resolution while preserve the spectral characteristics at a maximum degree and a SVM are trained to select coefficient blocks with salient features. We can conclude from table 1 and 2; we are getting better fusion results at smaller block sizes. The proposed method seems to increase the computational time complexity since a support vector Machine is used. But SVM training is performed only once hence this does not introduce any extra computational cost. The SIDWT based fusion is reported to produce significantly better results in terms of the temporal stability in fused multi sensor sequences compared to conventional multire solution DWT fusion. The experimental results show that this method gives the more informative fused image.

References

[1] S. G. Nikolov, "Image fusion: A survey of image methods, applications, systems and interfaces." University of Bristol, Technical report, Jul. 1999
[2] O. Rockinger. Image sequence fusion using shift invariant wavelet transformation, IEEE Transaction on Image Processing, 3:288-291
[3] C. Pohl and J. L. Van Genderen, "Multisensory image fusion in remote sensing: concepts, methods and applications," International Journal of Remote Sensing, vol. 19, no. 5, pp. 823–854,1998.[Online]. Available: http://dx.doi.org/10.1080/014311698215
[4] RajaKumari K1, Anu Priya S,"Survey on contourlet based fusion techniques for multimodality image fusion", International Journal of Computer Engineering and Applications, Volume IX, Issue III, March 15 www.ijcea.com ISSN 2321-3469
[5] P.T. Burt and E.H. Andelson, "The Laplacian pyramid as a compact Image code," IEEE Transactions on Communications, vol. 31, pp. 532– 540, 1983
[6] N. Cristianini and J. Shawe-Taylor, "An Introduction to Support Vector Machines."Cambridge University Press, Cambridge, UK, 2000
[7] Zheng Liu Ottawa, Canada,"Investigations on multi-sensor image system and its surveillance applications"Canada, September 2007.
[8] R. Blum and Z. Liu, "Multi-sensor Image Fusion and Its Applications", CRC Press, London, United Kingdom, 2005.
[9] R.C.Gonzalez, R.E.Woods, Digital Image Processing, 2nd ed., Prentice Hall
[10] Y. Zeng, J. Zhang, and J. Van Genderen, "Comparison and analysis of remote sensing data fusion techniques at feature and decision levels," vol. XXXVI. Enschede, the Netherlands: ISPRS,Mar.2006. [Online].Available:http://www.isprs.org/proceedings/XXXVI/part7/PDF/014.pdf
[11] J. R. Jensen, Remote Sensing of the Environment: An Earth Resource Perspective. Prentice Hall, Jan. 2000.
[12] Chun-Lin, A Tutorial of the Wavelet Transform, Liu February 23, 2010
[13] Andrew P. Bradley, "Shift-invariance in the Discrete Wavelet Transform"' Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 1012
[14] V. Petrovic, C. Xydeas, "Objective evaluation of signal-level image fusion performance," Optical Engineering, SPIE, Vol. 44(8), 087003, 2005.
[15] Steve R Gunn "Support Vector Machines for Classification and Regression"Technical Report
[16] www.image fusion.org