ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

Multimodal Image Fusion Using Discrete Wavelet Transform and Support Vector Machine

Anushka B. Jagtap, M.P.Satone
Department of E&TC Engineering, K.K.Wagh College Of Engineering & Research Centre, Nasik, India
Department of E&TC Engineering, K.K.Wagh College Of Engineering & Research Centre, Nasik, India

Abstract— ***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. Image fusion provides an effective way of reducing volume of information while at the same time extracting all the useful information from the source images. For image fusion Discrete Wavelet Transform (DWT) is used for feature extraction and Support Vector Machines (SVM) is used for the classification of the coefficients of DWT. DWT reduces resolution and reduces the computation time. It also decreases the storage space required. The proposed fusion method incorporates feature extraction of the wavelet transform sub bands such as energy, entropy, mean and standard deviation then classification of these features of can be done using classifiers such as Support Vector Machines.***

Index Terms —**multiple sensor images, Image fusion, Support vector machine (SVM), Wavelet coefficients, Feature extraction, Discrete Wavelet transform (DWT).**

I. INTRODUCTION

Multi sensor data is widely available in many fields, such as remote sensing, medical imaging or machine vision. Several images of the same scene provide different information as image is captured with different sensors. If we merge different information we can obtain new improved image. This is called fusion scheme. Multisensor image fusion is an integral part of any surveillance, weapon detection or tracking system. In the last few years, there has been a rising interest in surveillance applications due to the increasing availability of lost cost and low power visual sensors. In recent years there has been fast development in the use of multiple sensors to increase the capabilities of systems in a number of fields such as surveillance, remote sensing, medical imaging and military. 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 provides more information. 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 image fusion process can be done at three levels like pixel level, decision level and feature level [4]. Pixel level fusion works directly on the raw pixels and it then generates a fused image in which each pixel is obtained from a set of pixels in each source image. Feature level fusion implements feature extraction and works on image features extracted from the source images. Decision level fusion works at higher level and merges the interpretations of different images obtained after image understanding [5]. In both feature level and decision level fusion loss of information may take place in the information extraction process and it may consequently lead to less accurate fusion results [6]. The fusion process should preserve all relevant information of the source images in the fused image and it must not include any disturbances or loss of information.

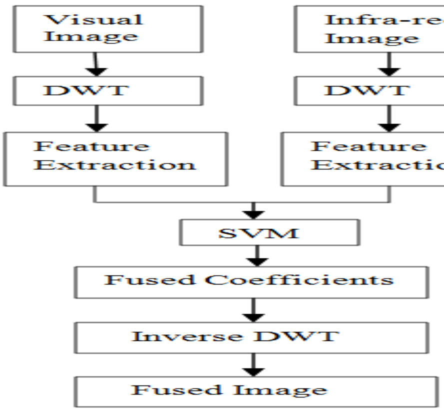
II. RELEVANCE

The simplest pixel level image fusion method takes the pixel-by-pixel intensity weighted average of the given source images. This often produces serious side effects such as reduced contrast [5]. Other popular methods used for image fusion are IHS based method, Brovey's method, Principal Component Analysis (PCA) and Independent Component Analysis (ICA). In recent years, many researchers have recognized that multiscale transforms (MST) are very useful for analyzing the information content of images for the purpose of fusion [4]. The most commonly employed mutiresolution decomposition methods are the Pyramid Transform and the Discrete Wavelet Transform (DWT). 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. There are different variations of pyramid transform that include Laplacian pyramid (LP) [6], Gradient Pyramid [7], the Morphological pyramid [8] and Contrast pyramid [9]. In fusion based on Averaging the simplest way to fuse two images is to take the mean value of the corresponding pixels. Maybe that for some applications this may be enough, but there will always be one image with poor lighting and thus the quality of an averaged image will obviously decrease. Principal components analysis is statistical technique that converts multivariate data with correlated variables into one with uncorrelated variables and obtains new variables which are linear combination of the original variables. The implementation process include steps like images size checking, source images must have the same size then the input images (images to be fused) are arranged in two column vectors. After that compute the eigenvector and eigenvalues for this resulting vector and the eigenvectors corresponding to the larger eigenvalue obtained, and finally

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

the normalized components are computed from the obtained eigenvector and then fused image obtained. In image fusion using Pyramid Transform techniques a pyramid structure can be described as a collection of images at different scales that together represent the original image. One image can be represented as a pyramid structure via pyramid transform. The Laplacian Pyramid implements a "pattern selective" approach to image fusion, so that the composite image is constructed not a pixel at a time, but a feature at a time. The basic idea is to perform a pyramid decomposition on each source image, then integrate all these decompositions to form a composite representation, and finally reconstruct the fused image by performing an inverse pyramid transform.

III. IMAGE FUSION METHOD USING DISCRETE WAVELET TRANSFORM (DWT) AND SUPPORT VECTOR MACHINE (SVM)



"Fig 1. Image Fusion Method"

In this method, fusion of visual and infrared (IR) images is performed using Discrete Wavelet transform and Support Vector Machines. Here only two source images are considered for fusion, but the algorithm can be extended to handle more input images.

The various steps involved in fusion using this method are as follows:

1. Using Discrete Wavelet transform (DWT), decompose the source images to be fused.
2. Arrange the sub bands in non overlapping blocks.
3. To train the SVM, extract the features energy, entropy and standard deviation for training blocks.
4. Train the SVM using the features extracted in step 3.
5. In fusion process, load the extracted features of blocks of sub bands and determine whether the wavelet coefficient block from visual image or infrared image is to be used.
6. If the classes value is 1 corresponding coefficient block from Visual image will be selected otherwise corresponding coefficient block from IR image will be selected.
7. The fused image is obtained by performing inverse DWT from the selected coefficients.

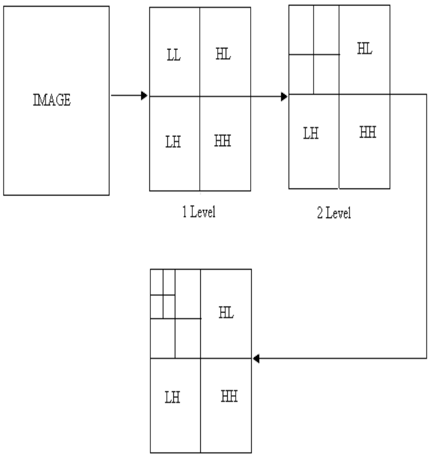
A. Discrete Wavelet Transform

In wide applications, we have to analyze a function in both time and frequency. Wavelet transform decomposes a signal into a set of basis functions (wavelets). It provides time-frequency representation. Like sines and cosines in FT, wavelets are used as basis functions in Wavelet transform. Wavelets are generated from a single basic wavelet $\psi_{s,\tau}(t)$, the so-called mother wavelet, by scaling and translation:

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \qquad (1)$$

Where s is the scale factor, τ is the translation factor. Scaling a wavelet simply means stretching (or compressing) it. Low scale means high frequency and high scale means low frequency. Wavelets are a powerful statistical tool which can be used for a wide range of applications, namely Signal processing, Data compression, Speech recognition. For a periodic function the classical method is Fourier transform. But the main drawback of Fourier transform is that we lose our time information which is very important. In the wavelet transform we do not lose the time information, which is useful in many contexts [13]. One of the main advantages of wavelets is that they offer a simultaneous localization in time and frequency domain. The second main advantage of wavelets is that, using fast wavelet transform, it is computationally very fast. Wavelets have the great advantage of being able to separate the fine details in a signal. Very small wavelets can be used to isolate very fine details in a signal, while very large wavelets can identify coarse details. A wavelet transform can be used to decompose a signal into component wavelets. On image matrix first 1-D filter bank is applied on the rows of the image and then it is applied on the columns of each channel of the result. So, we obtain 3 high pass channels which correspond to vertical, horizontal, diagonal, and one approximation image. The approximations are the high-scale, low-frequency components of the signal. The details are the low-scale, high-frequency components. We can iterate the above procedure on the low pass channel.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014



"Fig 2. 2 D DWT for image"

B. Feature Extraction

The input data is very large in some cases so it needs to be represented and reduced in set of features and hence this procedure is called as feature extraction. The features extracted should be carefully chosen and those extracted features should extract the relevant information from the input data to perform the desired task using this reduced representation instead of the full size input. Feature extraction simplifies the amount of resources that are required to describe a large set of data accurately. While performing analysis of complex data major problem is with large number of variables present. For analysing a large number of variables large amount of memory and computation power is required. Also classification algorithm is required which involves training and testing phase. Feature extraction is a method which constructs combinations of variables and describes the data with sufficient accuracy. Fig. 1 shows the block diagram of the system. Each source image is decomposed by DWT. 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

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

Entropy

$$H_c = -\sum_{i,j=1}^{N} C(i,j)\log_2 C(i,j) \qquad (3)$$

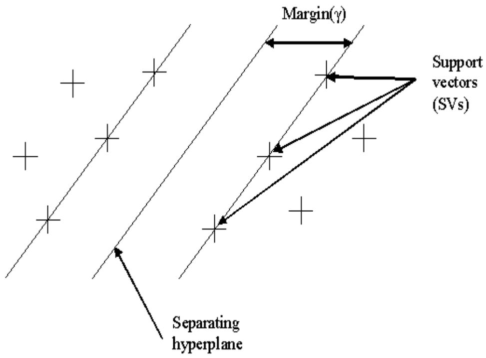
Standard Deviation

$$SD = \sqrt{\frac{1}{N^2}\sum_{i,j=1}^{N}\left[C(i,j)-\overline{C}\right]^2} \qquad (4)$$

Where $C(i,j)$ are the DWT coefficients and $\overline{C}$ is the mean of each coefficient block.

C. Support Vector Machine

SVM stands for "Support Vector Machine" it uses different planes in space to divide data points or samples using planes. It is viewing input data as two sets of vectors in an n-dimensional space, a SVM will construct a separating hyper plane in that space, which maximizes the margin between the two data sets [15]. In case of support vector machine, an object is viewed as a n-dimensional vector and we want to separate such objects with a n-1 dimensional hyper plane. This is called a linear classifier. Given a set of training examples, each marked as belonging to one of two categories, a SVM training algorithm builds a model that assigns new examples into one category or the other. A SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. In addition to performing linear classification, SVMs can efficiently perform non-linear classification, implicitly mapping their inputs into high-dimensional feature spaces. Support vector machine constructs a hyper plane or set of hyper planes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyper plane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

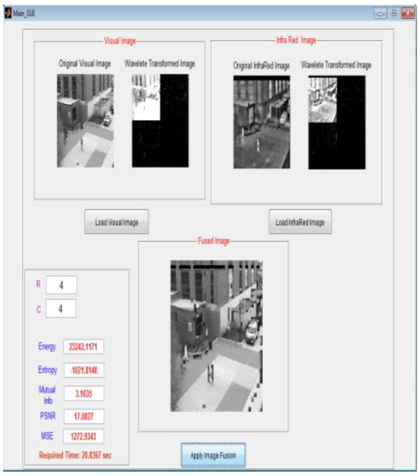


"Fig 3. Classification of Points using SVM"

The original optimal hyper plane algorithm proposed by Vapnik in 1963 was a linear classifier. However, in 1992, Bernhard E. Boser, Isabelle M. Guyonand Vladimir N. Vapnik suggested a way to create nonlinear classifiers by applying the kernel trick (originally proposed by Aizerman) to maximum-margin hyper planes. The resulting algorithm is formally similar, except that every dot product is replaced by a nonlinear kernel function. This allows the algorithm to fit the maximum-margin hyper plane in a transformed feature space. The transformation may be nonlinear and the transformed space high dimensional; thus though the classifier is a hyper plane in the high-dimensional feature space, it may be nonlinear in the original input space.

IV. SIMULATION RESULTS

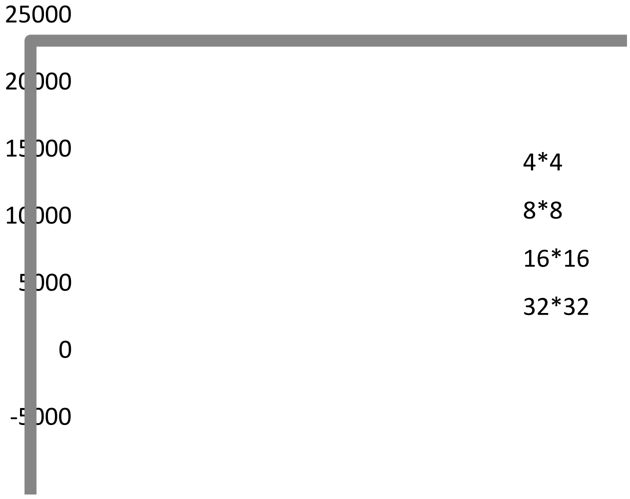
In results image fusion which has main two stages like training and then testing as classification includes two phases training of machine and then testing based on the trained model. Fig. 4 shows GUI with images in panel which include original visual image and its wavelet transformed image and infrared image panel include original infrared image and its wavelet transformed image. R and C is for selecting size of blocks to be fused from infrared image onto visual image. On application of image fusion algorithm shows the fused image. If R and C are changed then the fused image changes. Also fusion metrices like energy, entropy, mutual information(MI), peak signal to noise ratio(PSNR), mean square error (MSE) are calculated for different block sizes. The metrices were recorded for different block sizes like 4*4, 8*8, 16*16, 32*32 and are shown in tabular form and also graphical form so as to see variation.



"Fig 4. Image Fusion result for Visual and IR Image Blocks of 4*4"

"Table I: Fusion metrices for visual and IR image pair"

Block size	4*4	8*8	16*16	32*32
Energy	23242.12	21857.54	19455.38	18383.73
Entropy	-1021.81	-974.446	-888.802	-855.563
MI	3.1635	2.6418	2.2139	1.7917
PSNR	17.0827	14.151	11.9482	11.2954
MSE	1272.934	2500.233	4151.997	4825.446
Time (sec)	18.7793	5.0055	1.3577	0.48027



"Fig 5. Fusion metrices graphical representation"

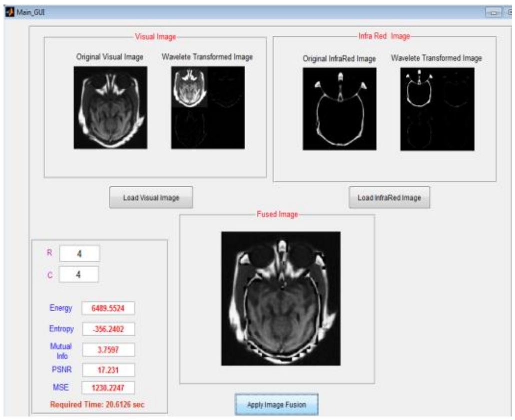
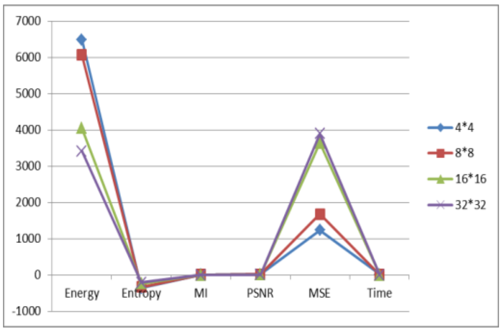


Fig 6. Image Fusion result for MRI and CT Image"

"Table II: Fusion metrices for CT and MRI image pair"

Block size	4*4	8*8	16*16	32*32
Energy	6489.552	6065.069	4045.95	3416.956
Entropy	-356.24	-329.36	-237.005	-204.515
MI	3.7597	3.4585	2.4167	2.3832
PSNR	17.231	15.8928	12.5188	12.2244
MSE	1230.225	1674.165	3640.84	3896.201
Time (sec)	18.8607	5.0348	1.364	0.46728

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014



"Fig 7. Fusion metrices graphical representation"

IV. APPLICATIONS

In remote sensing applications, the increasing availability of space borne sensors gives a motivation for different image fusion algorithms. Several situations in image processing require high spatial and high spectral resolution in a single image. Most of the available equipment is not capable of providing such data convincingly. The image fusion techniques allow the integration of different information sources. The fused image can have complementary spatial and spectral resolution characteristics. In satellite imaging, two types of images are available. The panchromatic image acquired by satellites is transmitted. At the receiver station, the panchromatic image is merged with the multispectral data to convey more information. Image fusion has become a common term used within medical diagnostics and treatment. Fused images may be created from multiple images from the same imaging modality or by combining information from multiple modalities, such as magnetic resonance image (MRI), computed tomography (CT), positron emission tomography (PET), and single photon emission computed tomography (SPECT). For accurate diagnoses, radiologists must integrate information from multiple image formats. Fig. 6 shows the image fusion result for the MRI and CT image. Instead of visual image we have loaded MRI image and instead of infrared image we have loaded CT image. In the fused image CT image is fused onto MRI image and thus explores soft tissue information of MRI image and bony information of CT image in the same image.

V. CONCLUSION

Image fusion reduces the large amount of data. Also it provides more suitable fused image for human/machine perception. Image fusion has tremendous scope in surveillance, medical imaging, and remote sensing. In proposed image fusion method first of all it is important to transform the input images and then arrange their sub bands in non-overlapping blocks and then do the feature extraction of each sub band. After that those features of sub bands should be classified and selected appropriately so as to get the best possible results of fusion. We can conclude from tables I and II that as block size is increased energy, entropy, mutual information (MI), peak signal to noise ratio(PSNR) are reduced and mean square error (MSE) is increased. So we are getting better fusion results at smaller block size. But time increases 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 experimental results show that this method gives the more informative fused image.

REFERENCES

[1] D. Egfin Nirmala, Bibin Sam Paul, Dr.V.Vaidehi, "A Novel Multimodal Image Fusion method using Shift Invariant Discrete Wavelet Transform and Support Vector Machines" IEEE-International Conference on Recent Trends in Information Technology, ICRTIT 2011 ©2011 IEEE MIT, Anna University, Chennai. 2011

[2] Nedeljko Cvejic, Tapio Seppänen, and Simon J. Godsill, "A No reference Image Fusion Metric Based on the Regional Importance Measure", IEEE Journal of Selected Topics in Signal Processing, Vol. 3, No. 2, April 2009.

[3] Nedeljko Cvejic, Member, "Adaptive Fusion of Multimodal Surveillance Image Sequences in Visual sensor networks",IEEE Transactions on Consumer Electronics, Vol. 53, No. 4, November 2007.

[4] R. Blum and Z. Liu, Multi-sensor Image Fusion and Its Applications, CRC Press, London, United Kingdom, 2005.

[5] Shutao Li, James Tin Kwok, Ivor W. Tsang, Yaonan Wang, "Fusing Images with Different Focuses Using Support Vector Machines", IEEE Transactions on Neural Networks, Vol.15, No.6, Nov.2004

[6] 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.

[7] OTCBVS Benchmark Dataset Collection Burt, P. J., A gradient pyramid basis for pattern-selective Image fusion, SID Tech. Digest, 16, 467 – 470, 1992

[8] Toet, A., A morphological pyramid image decomposition, Pattern Recognition, Letters, 9, 255 – 261, 1989

[9] A. Toet, L.J. van Ruyven, and J.M. Valeton, "Merging thermal and visual images by a contrast pyramid," Optical Engineering, vol. 28, No. 7, pp. 789–792, July 1989.

[10] P.K.Varshney, "Multisensor data fusion", Electron. Commun. Eng., 245-253, Dec.1997.

[11] H. Li, B.S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," Graphical Models and Image Processing, vol. 57, no. 3, pp. 235–245, May 1995.

[12] S.G. Mallat, "A Wavelet Tour of Signal Processing", Academic Press, San Diego, 2nd edition, 1999

[13] O. Rockinger. Image sequence fusion using a shift invariant wavelet transform, IEEE Transactions on Image Processing, 3:288–291.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 3, Issue 12, June 2014

[14] Steve R.Gunn , "Support Vector Machines for classification and Regression", Technical Report University of Southampton

[15] Nikolaos Mitianoudis and Tania Stathaki, "adaptive image fusion using ICA Bases", Imperial College London, Exhibition Road, SW7 2AZ London, UK

[16] V. Petrovic, C. Xydeas, "Objective evaluation of signal-level image fusion performance," Optical Engineering, SPIE, Vol. 44(8), 087003, 2005.

[17] R.C.Gonzalez, R.E.Woods, "Digital Image Processing", 2nd ed., Prentice Hall.