FUSION OF MULTIMODAL IMAGES USING NON-SUBSAMPLED CONTOURLET TRANSFORM AND ADABOOST SUPPORT VECTOR MACHINE CLASSIFIERS

D. Egfin Nirmala[1], *R.K.Vignesh*[2], *Dr.V.Vaidehi*[3]

[1,2] Department of ECE, B.S. Abdur Rahman University, Vandalur, Chennai-600048.
[3] Department of Information Technology, Madras Institute of Technology, Chromepet, Chennai-600044.
[1]egfin_nirmala@yahoo.co.in, [2]vigneshrk1109@gmail.com, [3]vaidehivijay@gmail.com

ABSTRACT

A novel region based multisensor image fusion technique using Non-Subsampled Contourlet Transform (NSCT) and AdaBoost Support Vector Machines (AdaBoostSVM) suitable for multimodal surveillance images is proposed in this paper. Recent studies show NSCT as an efficient technique for multisensor image fusion. In this paper, we improve NSCT fusion by using an AdaBoostSVM classifier. This technique uses NSCT for multiresolution decomposition as it can efficiently capture geometric structures such as smooth contour edges, than wavelets. Region based joint segmentation of the source images is performed in the spatial domain and various features of each region are computed in transform and spatial domains. An AdaBoostSVM classifier is trained using the region features to select regions from the source images with important features. The proposed method is tested for a number of multimodal images and found to perform better than the state-of-art methods both visually and in terms of various fusion metrics.

Index Terms— Image Fusion, Non-Subsampled Contourlet Transform, AdaboostSVM classifier, Joint Segmentation.

1. INTRODUCTION

With the recent development in the field of visual sensor technology, multiple sensors of different types are widely employed in a number of fields such as surveillance, remote sensing and military. Image fusion combines information captured by different sensors that enhances the perception of a scene [1], so that the fused images are more suitable for human perception and computer processing. The main advantages of image fusion are that information can be obtained with more accuracy, in less time and at a lower cost [2]. Image fusion can be performed at different levels of information representation namely signal level, pixel level, feature level and decision level [2]. The pixel level fusion works directly on the raw pixels. Feature level fusion, works on the various image features extracted from the source images. Decision level fusion, merges the interpretations of different images obtained after image understanding [3]. In recent years, multiscale transforms are found to be very useful for fusion [2]. The widely used multiresolution decomposition methods for image fusion are the Pyramid Transforms [4-7] and the Discrete Wavelet Transform (DWT) [8].

For one-dimensional signals, wavelets were proven as the right tool, because they provide an optimal representation for these signals [9]. But wavelets in two-dimension (2-D) cannot represent the 'line' and the 'curve' discontinuities effectively [9]. Also, wavelets capture only limited directional information and are not able to represent the directions of the edges accurately [10]. To overcome the drawbacks of traditional DWT, a 2-D image representation called Contourlet Transform (CT) [9] is introduced by Do and Vetterli. But, the Contourlet transform is not shift-invariant, and in 2006, Arthur Cunha, and Jianping Zhou introduced an over complete shift-invariant image representation called the Nonsubsampled Contourlet Transform (NSCT) [11]. This is a shift-invariant, multiscale, anisotropic and multidirection transformation with a fast implementation. Image fusion using multiresolution transforms are based on certain 'fusion rules' formulated to combine the coefficients of the source images in the transform domain.

Recently, region based methods have attracted the researchers as it is more meaningful to combine regions in images rather than pixels [12]. The objects in a scene and the region they represent are more important than the individual pixels. Hence, object and region information are incorporated in the fusion process so that it becomes more robust and reduces some of the problems in pixel level fusion such as high sensitivity to noise and blurring effects [12, 13]. Region based image fusion involves, features selection and extraction, a pattern classification problem. In such cases, the selection of a suitable classifier and effective use of multiple features of sensed data are very important [14]. A major development in machine learning in the past decade is the introduction of the Boosting method, AdaBoost [15], proposed by Freund and Schapire, that combines many weak classifiers to form a strong classifier. The proposed method uses SVM as a component classifier in AdaBoost, as SVM outperforms conventional weak classifiers in many applications [3]. AdaBoostSVM classifiers can perform much better than individual SVMs and the classification accuracy is greatly improved [16].

This paragraph presents the details on how our paper is related to prior work in this field. In [3] a pixel based method for fusing multifocus images using DWFT and SVM classifier is proposed, in which the SVM is trained to select the source image that has the best focus at each pixel location. Shaohui Chen *et al* [17] proposed a combination of Empirical Mode Decomposition (EMD) and SVM as an improved pixel based method for merging multifocus images. In [3, 17] the authors demonstrated an improvement in fusion performance by using a SVM classifier. According to [10], the NSCT is more suitable for image fusion, as more information for fusion can be obtained by using NSCT and the impacts of mis-registration on the fused results are reduced effectively. NSCT is applied for image fusion by Yang et al [18] in which the transform's coefficients are combined using different fusion rules and results showed better performance than wavelet based schemes. In [19] Shutao Li and Bin Yang proposed a pixel based hybrid multiresolution method by combining the Stationary Wavelet Transform with NSCT. Improved performance over NSCT fusion is obtained with more decomposition levels and directions, which consumed more time. Huang Qingqing et al [20] introduced NSCT into infrared and visible remote sensing image fusion and proposed a multi-scale analysis method based on region energy. Nikolaos Mitianoudis et al. [1] have proposed a region based approach for image fusion using Independent Component Analysis (ICA) and found to outperform wavelet based methods. In [12], Piella presented a robust region based multiresolution image fusion algorithm that claimed improved performance over pixel based multiresolution fusion.

It is evident from the related work, that the combination of NSCT and AdaBoostSVM are more suitable for image fusion applications. In this paper, a new region-based image fusion method is proposed for multimodal images, that incorporates the local statistical characteristics of regions, with a better decision making process by the AdaBoostSVM. This work is different from the related work in the sense that multiple region features computed in both spatial and transform domain are used in order to determine the importance of a region and a trained AdaBoostSVM replaces the role of fusion rules in selecting the coefficients.

In the proposed method, first the source images are decomposed to obtain the NSCT coefficients. The source images are segmented into regions and various features are computed for each region in spatial and NSCT domain. The trained AdaBoostSVM classifier is used to select regions with significant features from the segmented source images.

The remaining sections of this paper are organized as follows: Section 2 briefly introduces the NSCT, Segmentation Process and AdaBoostSVM, Section 3 gives the proposed fusion scheme. Section 4 presents experimental results, and Section 5 summarizes this paper.

2. BACKGROUND

2.1. Non Subsampled Contourlet Transform

NSCT is a flexible multiscale, multidirection, and shift-invariant image decomposition scheme implemented using nonsubsampled pyramids (NSP) and nonsubsampled directional filter banks (NSDFB) [22]. NSP ensures the multiscale property and NSDFB provides directionality. More details about NSP and NSDFB are given in [11].

2.2. Segmentation

In region based fusion, the source images are segmented to produce a set of regions. Some important statistical properties of each region can be computed and these are used to determine from which source image the particular region is to be chosen in the fused representation. In the proposed work, joint segmentation of source images is performed using the combined morphological-spectral unsupervised image segmentation algorithm [21], as it is suitable for multimodal images. Jointly segmented images work better for fusion, as the segmentation map will contain a minimum number of regions of same size to represent all the features in the scene. More details about this algorithm are available in [21].

2.3. AdaBoost SVM

The Support Vector Machine (SVM) performs classification by constructing an *N*-dimensional hyperplane that optimally separates the data into different categories [23]. The SVM initially maps the training samples from input space to some feature space and then separates the different classes by constructing a maximum margin hyperplane. A popular kernel used in SVM is the Radial Basis Function (RBF) kernel, which has a parameter known as Gaussian width σ . The performance of RBFSVM can be modified by simply adjusting σ. AdaBoostSVM uses a set of trained RBFSVM as weak learners for AdaBoost. Initially, a large value is set to the step size σ , corresponding to a RBFSVM classifier with very weak learning capability. The σ values are gradually decreased as the boosting iteration proceeds. The process continues until σ is decreased to the given minimal value. The AdaBoostSVM algorithm is discussed in detail in [16].

3. PROPOSED FUSION METHOD

The proposed fusion method generates a composite fused image, from a pair of registered visible and IR source images. Firstly, the images to be fused are jointly segmented into different regions using combined morphological-spectral unsupervised image segmentation algorithm. The following features are computed for all regions in spatial and NSCT domain.

3.1. Features in Spatial domain:

The following four features are computed for all regions in the spatial domain as these are generally a good measure of a region's importance:

1. Energy: $$E=\sum_{i=1}^{M}\sum_{j=1}^{N}R^2(i,j) \tag{1}$$

   where $R(i,j)$ is the $(i,j)^{th}$ entry in each region.
2. Standard Deviation:
   $$SD=\sqrt{\frac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[R(i,j)-\overline{R}\right]^2} \tag{2}$$
   where $\overline{R}$ is the mean of the pixels in a region.
3. Sharpness: $$S=\frac{\sqrt{G_x^2+G_y^2}}{M*N} \tag{3}$$
   where G_x and G_y are the gradients of the input image.
4. Contrast: $$C_o=\sum_{i=1}^{M}\sum_{j=1}^{N}(i-j)^2\,R(i,j) \tag{4}$$

3.2. Features in Transform domain:

Texture is one of the important characteristics used in identifying objects and the regions of interest in images [24]. From the Gray

Level Co-occurrence Matrix (GLCM) of an image various textural features can be computed. Each entry (i, j) in GLCM corresponds to the number of occurrences of the pair of gray levels i and j which are a distance 'd' apart in the image, which is taken as (0,1) in this work. The features calculated from the GLCM of the input images in NSCT domain are:

5. Entropy: $$H = -\sum_{i=1}^{M}\sum_{j=1}^{N} C(i,j)\log_2 C(i,j) \quad (5)$$

6. Local homogeneity:
$$LH = \sum_{i=1}^{M}\sum_{j=1}^{N} \frac{1}{1+(i-j)^2} C(i,j) \quad (6)$$

7. Cluster Shade:
$$CS = \sum_{i=1}^{M}\sum_{j=1}^{N} \left(i - P_x(i) + j - P_y(j)\right)^3 C(i,j) \quad (7)$$
where $C(i,j)$ is the $(i,j)^{th}$ entry in the GLCM. $P_x(i)$ and $P_y(j)$ are the i^{th} and j^{th} entry in the marginal probability matrix obtained by summing the rows and columns of the GLCM respectively.

8. Cluster Prominence:
$$CP = \sum_{i=1}^{M}\sum_{j=1}^{N} \left(i - P_x(i) + j - P_y(j)\right)^4 C(i,j) \quad (8)$$

9. Correlation:
$$CR = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N} ijC(i,j) - \mu_x\mu_y}{\sigma_x\sigma_y} \quad (9)$$
where μ_x , σ_x and μ_y , σ_y are the mean and standard deviation of P_x and P_y respectively.

3.3 Training and Fusion:

To train the AdaBoostSVM a set of training images both visual and IR are required. The training images are segmented into regions using the segmentation algorithm [21]. Some regions of different sizes, from the segmented images are randomly selected and the features discussed in 3.1 and 3.2 are computed. The combination of all these features is represented as a 9-dimensional feature vector and is given to the AdaBoostSVM for classification. The AdaBoostSVM target output is positive (+1) if the region is chosen from visual image and negative (-1) if chosen from IR image. Thus the AdaBoostSVM is trained to select a particular region with significant features from the visual or IR source image. The trained AdaBoostSVM is used to perform fusion of the source images.

The proposed image fusion approach is described as follows:

1) The source images are jointly segmented using combined morphological-spectral unsupervised image segmentation algorithm and for every region the features discussed in section 3.1 and 3.2 are computed.
2) Each of the source images is decomposed by the NSCT to L levels to obtain the corresponding coefficients.
3) The trained AdaBoostSVM is used to select segmented regions from image A or B. If the classifier target output is positive (+1) for a region, the corresponding NSCT coefficients for that region from visual image will be selected and for negative (-1) corresponding coefficients from IR image will be selected. The selected coefficients of all the regions form the fused NSCT representation.
4) The fused image is obtained by performing inverse NSCT to the fused coefficients.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment is done using a 2.4 GHz, i3 CPU, 4 GB RAM, with MATLAB 7.10. In this work, we use three scales of decomposition for NSCT. This method uses RBFSVM as weak learner for AdaBoost. Before the fusion process, the AdaBoostSVM was trained using 200 regions randomly taken from a set of visual and IR images from the image fusion site [25]. The nine features are computed for all regions in the visible and IR source images. The classification performance of SVM is affected by the RBF kernel model parameter the Gaussian width σ . A larger σ lowers the classification accuracy and a smaller σ increases the complexity. A set of RBFSVM component classifiers is obtained by adaptively adjusting their σ values in steps σ_{step} .The $\sigma_{\min}$ value is chosen as minimal distance between any two training regions and the initial value σ_{in} is chosen as the scatter radius of the training samples. Fig. 1 shows the variation of the classification error for different values of σ_{step} . The number of learning cycles in AdaBoostSVM changes with the value of the σ_{step} , but final test error is almost stable.

The proposed method is tested on different sets of visible and IR images. The fusion results are compared with other fusion algorithms such as Laplacian Pyramid (LP), DWT, Dual Tree Complex Wavelet Transform (DT-CWT), NSCT and Hybrid NSCT. In this work, the objective fusion metrics Entropy, Mutual Information (MI), Petrovic metric (Q^f_{AB}) [26], Q_w and Q_E [27] are compared to evaluate the quality of the fused images. These metrics for evaluation of image fusion are based on the amount of information transferred from input images to the fused output.

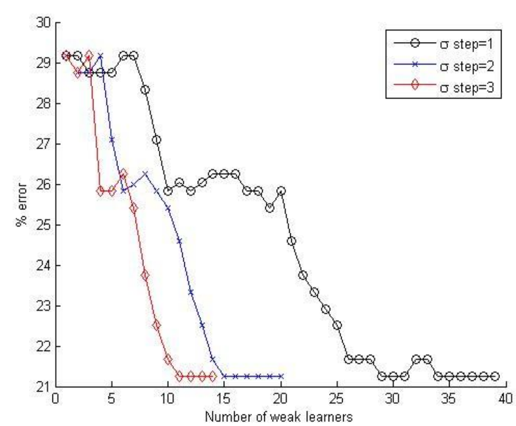

Fig.1 Performance of AdaBoostSVM for different σ_{step} values

Fig. 2 shows a pair of Visual and IR source images, and the fused results obtained by the different methods. Results show that the fused image obtained by the proposed method is with the best visual quality and almost all the important information in the source images have been transferred to the fused image. Only the target (the gun) information from the IR image is transferred to the fused image whereas the background is taken from the visual image. Fig. 3 shows the UN camp source images and the fused images. In the fused image produced by the proposed method, the person appears brighter and the trees in the visual image are clearly seen than in any other method. The fine details from the visual image are transferred to the fused image in the proposed method. The performance evaluated is presented in Table I and II.

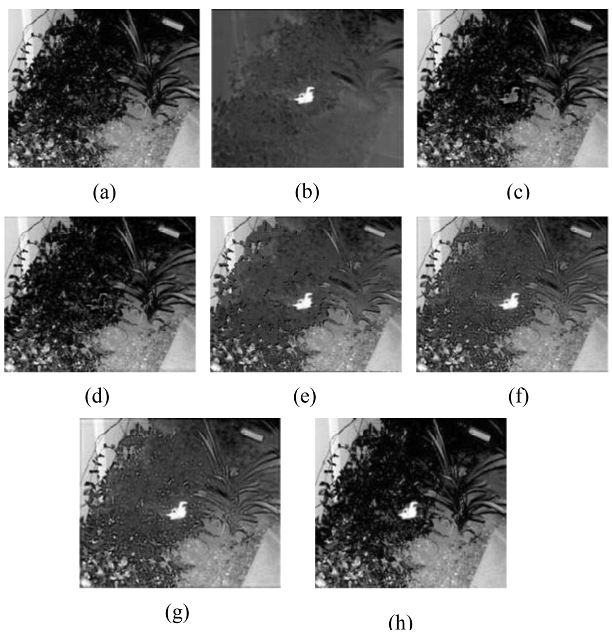

Fig.2 GUN source Images (a) Visual image (b) IR image. Images fused using (c) LP (d) DWT (e) DTCWT (f) NSCT (g) Hybrid NSCT (h) Proposed Method

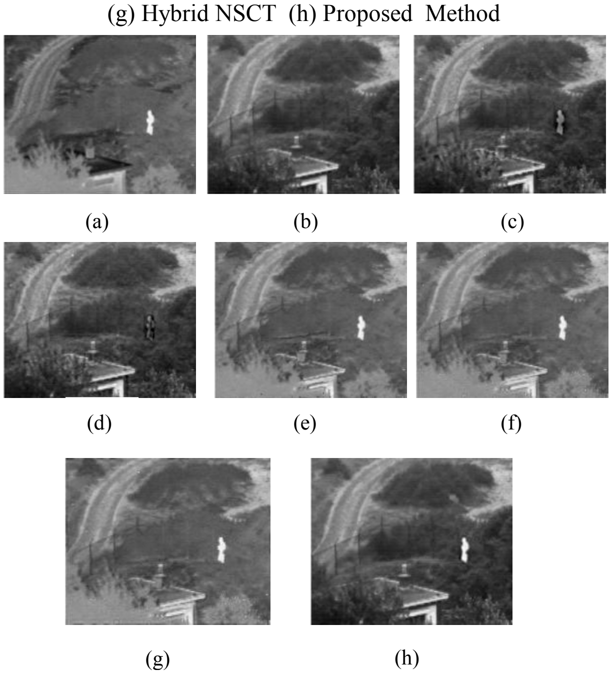

Fig.3 UN CAMP source Images (a) Visual image (b) IR image, Images fused using (c) LP (d) DWT (e) DTCWT (f) NSCT (g) Hybrid NSCT (h) Proposed Method

Table.I Performance Metrics of Various Fusion Methods for Gun Image

Fusion Methods	Performance Metrics				
	Entropy	MI	Q^{f}_{AB}	Q_w	Q_E
Laplacian Pyramid	7.5	4.98	0.64	0.85	0.76
DWT	7.49	4.5	0.61	0.83	0.68
DT-CWT	7.31	4.06	0.45	0.72	0.59
NSCT	7.41	4.48	0.64	0.80	0.70
Hybrid NSCT	7.43	4.6	0.65	0.81	0.71
Proposed Method	**7.53**	**7.96**	**0.75**	**0.89**	**0.81**

Table.II Performance Metrics of Various Fusion Methods for UN Camp Image

Fusion Methods	Performance Metrics				
	Entropy	MI	Q^{f}_{AB}	Q_w	Q_E
Laplacian Pyramid	**7.16**	2.96	0.44	0.68	0.57
DWT	7.15	3.1	0.40	0.64	0.47
DT-CWT	6.73	3.49	0.43	0.57	0.43
NSCT	6.72	3.12	0.43	0.58	0.44
Hybrid NSCT	6.75	2.84	0.42	0.57	0.46
Proposed Method	7.14	**7.41**	**0.60**	**0.75**	**0.64**

It is evident from the fusion results that, the proposed method scores high in terms of the fusion metrics than the individual and hybrid multiresolution methods.

5. CONCLUSION

In this work, we have developed a novel region based image fusion method using NSCT and AdaBoostSVM classifier. Experimental results on several multimodal images show that the NSCT with AdaboostSVM are superior to those of the individual multiresolution based methods like LP, DWT, NSCT and also hybrid NSCT, both visually and objectively. The proposed method may appear as to increase the computational complexity since the AdaBoostSVM is to be trained. But this task is performed only once prior to the fusion process and hence does not introduce additional computation cost. The fusion results can be improved further by increasing the number of NSCT decomposition levels and directions, at the cost of increased computational complexity.

6. REFERENCES

[1] Nikolaos Mitianoudis, Tania Stathaki, "Pixel-based and Region-based Image Fusion schemes using ICA bases," *Information Fusion.l* 8(2): 131-142, 2007.

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

[3] Shutao Li, James Tin Kwok, Ivor W. Tsang, Yaonan Wang, "Fusing Images with Different Focuses Using Support Vector Machines," *IEEE Transactions on Neural Networks,* 15(6): 1555-1561, Nov.2004.

[4] T.Burt and E.H.Andelson. "The Laplacian pyramid as a compact Image code", *IEEE Transactions on Communications*. 31 : 532–540, 1983.

[5] Burt, "P. J. A gradient pyramid basis for pattern-selective Image fusion," *SID Tech. Digest*, 16 :467 – 470, 1992.

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

[7] A. Toet, L.J. van Ruyven, and J.M. Valeton. "Merging thermal and Visual images by a contrast Pyramid," *Optical Engineering. 28(7):* 789–792, July 1989.

[8] H. Li, B.S. Manjunath, and S.K. Mitra. "Multisensor image fusion using the wavelet transform." *Graphical Models and Image Processing.* 57(3): 235–245, May 1995.

[9] M.N. Do, M. Vetterli, The contourlet transform. "An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing.* 14 (12): 2091–2106, 2005.

[10] Qiang Zhang , Bao-long Guo. "Multifocus image fusion using the Nonsubsampled Contourlet transform," *Elsevier, Signal Processing* 89 :1334–1346, 2009.

[11] A.L. da Cunha, J.P. Zhou, M.N. Do. "The Nonsubsampled contourlet transform: theory, design, and applications." *IEEE Transactions on Image Processing.* 15 (10): 3089-3101, 2006.

[12] Gemma Piella. "A region based multiresolution image fusion algorithm," *Proceedings of Fifth International Conference in Image Fusion.* 2 :1557-1564, 2002.

[13] John J. Lewis, Robert J. O_Callaghan, Stavri G. Nikolov, David R. Bull, Nishan Canagarajah. "Pixel- and region-based image fusion with complex wavelets," *Information Fusion 8*: 119-130, 2007.

[14] D. Lu & Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *International Journal of Remote Sensing*,28:5, 823 – 870

[15] Freund, Y., Schapire, R.E. "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences.*, 55 (1), 119-139,1997.

[16] Xuchun Li, Lei Wang, Eric Sung. "AdaBoost with SVM-based component classifiers," *Elsevier, Engineering Applications of Artificial Intelligence.,* 21 :785-795, 2008

[17] Shaohui Chen, Hongbo Su , Renhua Zhang, Jing Tian and Lihu Yang, "Improving Empirical Mode Decomposition Using Support Vector Machines for Multifocus Image Fusion", *Sensors*, 8, 2500-2508, 2008

[18] B. Yang, S. T. Li, and F. M. Sun. "Image fusion using Nonsubsampled contourlet transform," *Proc. IEEE 4th Int. Conf. Image Graphics,Chengdu,* China :719-724, 2007.

[19] Shutao Li, Bin Yang," Hybrid Multiresolution Method for Multisensor Multimodal Image Fusion", *IEEE Sensors Journal,* Vol. 10, No. 9, September 2010.

[21] Robert J. O'Callaghan, David R. Bull, "Combined morphological-spectral unsupervised image segmentation," *IEEE transactions on Image Processing.* 14(1): 49-63, 2005.

[22] R.H. Bamberger, M.J.T. Smith. "A filter bank for the directional decomposition of images: theory and design." *IEEE Transactions on Signal Processing.* 40 (4): 882-893, 1992.

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

[24] R. M. Haralick, K. Shanmugam, and I.Dinstein.Textural "Features for Image Classification." *IEEE Trans. on Systems, Man,and Cybernetics.*,3(6): pp.610-621, November 1973.

[25] http://www.imagefusion.org

[26] C. S. Xydeas and V. Petrovic. "Objective image fusion performance measure." *Electron. Lett.*, 36(4): 308–309, 2000.

[27] G. Piella and H. Heijmans. "A new quality metric for image fusion." *Proc. IEEE Int. Conf. Image Process.*, Barcelona, Spain, 173-176.