

# Multi-Source Remote Sensing Classification Based on Mallat Fusion and Residual Error Feature Selection

Dongdong Cao<sup>1</sup>, Qian Yin<sup>2</sup>, Ping Guo<sup>3</sup>

<sup>1,2,3</sup> Image Processing and Pattern Recognition Laboratory,  
Beijing Normal University, Beijing 100875  
China.

<sup>3</sup>School of Computer Science  
Beijing Institute of Technology  
Beijing 100081. China.

pguo@ieee.org



Journal of Digital  
Information Management

Uncorrected  
Proof

**ABSTRACT:** Classification of multi-source remote sensing images has been studied for decades, and many methods have been proposed or improved. Most of these studies focus on how to improve the classifiers in order to obtain higher classification accuracy. However, as we know, even if the most promising method such as neural network, its performance not only depends on the classifier itself, but also has relation with the training pattern (i.e. features). On consideration of this aspect, we propose an approach to feature selection and classification of multi-source remote sensing images based on Mallat fusion and residual error in this paper. Firstly, the fusion of multi-source images can provide a fused image which is more preferable for classification. And then a feature-selection scheme approach based on fused image is proposed, which is to select effective subsets of features as inputs of a classifier by taking into account the residual error associated with each land-cover class. In addition, a classification technique base on selected features by using a feed-forward neural network is investigated. The results of computer experiments carried out on a multi-source data set confirm the validity of the proposed approach.

## Categories and Subject Descriptors

F.1.1[Models of Computation]; Self-modifying machines C.2.1 [Network Architecture and Design] I.4 [Image Processing and Computer Vision]

## General Terms

Neural Networks, Remote sensing classification, Mallet fusion

**Keywords:** Mallet fusion, Multi-source classification, Feature selection, Residual error, Neural network

Received 30 November 2006; Revised 15 February 2007; Accepted 12 March 2007

## 1. Introduction

In recent years, remote sensing images have been used more and more widely in many application fields. These images are used to identify the objects from the scenes, namely remote sensing data classification. Traditionally, classification is implemented based on the single image. However, the features provided by the single resource are incomplete and imprecise [1]. Consequently, the classification accuracy is not satisfied. To improve the problem, it is necessary to acquire more useful information from the source image. Now, advanced technology has made it possible to take detailed measurements about the Earth surface by providing multi-source images collected by different remote sensing platform. Given such multi-source data, it is mandatory to develop effective and suitable fusion method to obtain

more information from obtained images. Besides, how to improve classification accuracy further needs to take not only feature selection but also classification algorithm into account [2] [3].

In the literature, many papers have been published to address fusion method [4 - 10]. A major observation in previous research on multi-source fusion is to use conventional parametric statistical pattern recognition methods. A kind of simplest approaches to the fusion of multi-source data is to concatenate the data from different sources as if they were from one single source, and use the multivariate normal distribution to describe the fused data [1][4][6]. However, in many cases, the assumed convenient multivariate statistical model does not fit to real data set. Consequently, this kind of approaches is not suitable when a common distribution model can not describe the various sources considered. Benediktsson and Swain [11] modified the approach by including "reliability factors" to weigh the importance of sources according to their reliability. Zhao and Chen [7] adopted it and their experiments show that it is superior to the traditional methods. However, the fusion result is also not preferable.

Other well-known approaches are (Hue, Intensity, Saturation) (HIS) transform and Principal Component Analysis (PCA) [13]. The image is transformed from RGB color space to HIS color space by HIS method. Then, the Intensity component is filled by the high resolution image. At last, fused image is acquired by the inverse HIS transform. The principle of PCA analysis is same to that of HIS. Although the two methods may enhance spatial resolution, they will introduce inconsistencies which results in spectral information distraction [14]. Thus the fused image is not suitable to implement classification.

Generally speaking, for a good remote sensing fusion method, it is required not only improving spatial details but also preserving the spectral information fidelity. Here, we are interested in the Mallat fusion method for multi-source images. By fusion, a new image that is more suitable for classification will be acquired. Compared to other fusion methods, it can provide more information on the sharp contrast changes, which are especially important for classification, and provide both spatial and frequency domain localization, which can avoid introducing artifacts.

As for the feature selection and classification algorithm, similarly, many methods have been proposed for classifications of multi-source images. They are mainly based on statistical, symbolic and neural network approaches. Bayesian statistical theory has been widely used as a theoretically robust foundation for the classification of remote sensing data. However, the statistical classification cannot be used satisfactorily in processing multi-source data since in most cases they cannot be modeled by a

<sup>3</sup> To whom all correspondence should be addressed

\*This work was supported by the grants from the National Natural Science Foundation of China (Project No. 60275002, 60675011)

convenient multivariate statistical model [8] [14].

Another approach to multi-source classification based on Dempster-Shafer evidence theory is reported in [15], where an unsupervised method is proposed. The method is not successful in the context of multi-source data, because they are not common in classification of various data. It is necessary to know probabilistic distribution of different data classes and so results of classification are dependant on user's knowledge.

Neural network classifiers provide an effective classification of different types of data. The nonparametric approach they implement allows the aggregation of different data sources into a stacked vector without need of assuming a specific probabilistic distribution of the data. Several studies on the classification of multi-source remote-sensing data by neural networks have been reported [1] [16]. Some of them investigated and compared the performances of neural network classifiers with those of both parametric and nonparametric statistical methods. Results point out the effectiveness of neural network approaches for the classification of multi-source data. However, due to the characteristics of neural networks, the complexity and performance of network also depends on appropriate and proper training patterns. It is necessary to further investigate the problem in order to develop methodologies capable of outperforming the current available method. The choice of a set of features that can be best discriminate among land-cover classes to be recognized by a classifier is one of main problem involved in the development of a classification system.

Based on above theories, in this paper, a new approach to feature selection and classification based on residual error is presented after fusion. In particular, a feature-selection approach is proposed that selects effective subsets of features to be given as input to a classifier by taking into account the residual error associated with each land-cover class. In addition, a classification technique base on selected features by using a neural network is implemented. Such a classification technique, which is of the non-parametric type, is suitable to process the multi-source data. The experiments are implemented to confirm the validity of the proposed approaches.

The paper is organized into five sections. The Mallat fusion theory is briefly described in section II. A feature selection and neural network classification based on residual error is proposed in section III. In section IV, the data set used for experiments, the process of fusion and the classification pre-processing applied to data and experiments' results are reported and discussed. Finally, the conclusions are drawn in Section V.

## 2. Background

Recently, the wavelet transform approach has been used for fusing data and becomes hot topic in research. However, little work has been done to implement multi-level decomposition and explore the effects of different wavelet coefficients to the fused image in both spectral and spatial features.

### A. Basic principle of wavelet transforms

The principle of the fusion method is to merge the wavelet decompositions of the multiple original images based on fusion rules, which is applied to approximations coefficients and details coefficients [8].

As for remote sensing data, a filtering approach of wavelet decomposition is implemented. This is decomposed into a

low pass approximation and three high pass detail images (wavelet coefficients) which correspond to the three directions: horizontal, vertical, and diagonal.

In that way four under images (HH, GH, HG, and GG) can be obtained from one full resolution image [9]. HH is the context image at the inferior resolution (the approximation), GH is the image of the horizontal details, HG is the image of the vertical details and GG is the image of the diagonal details (see Figure 1).

Perfect reconstruction of the original can also be achieved through the inverse wavelet transform [10].

### B. Mallat Method

Mallat, enlightened by the decomposition and reconstruction pyramid algorithm proposed by Burt and Adelson, developed the algorithm based on multiple resolution analysis of wavelet transformation [12] [13].

The procedure of mallat fusion can be described as the following four steps.

#### (1) Wavelet transformation

The process can be expressed in the following equation(1)-(4), where  $H$  and  $G$  represent the low-pass and high-pass filtering;  $H_r$  and  $G_r$  are the expression of filtering operated on the row;  $H_c$  and  $G_c$  are the expression operated on the column, respectively.

$$C_{j+1} = H_r H_c C_j \quad (1)$$

$$D_{j+1}^1 = H_r G_c C_j \quad (2)$$

$$D_{j+1}^2 = G_r H_c C_j \quad (3)$$

$$D_{j+1}^3 = G_r G_c C_j \quad (4)$$

#### (2) Multi-level decomposition

Wavelet decomposition is implemented on the low frequency segmentation of upper level decomposed, and then the pyramid structure of wavelet transformation is obtained. The Figure 2 shows the pyramid structure.

#### (3) Feature selection in the highest level

In the highest level of transformation, it is mandatory to select or fuse data information. The fusion rules play a very important role during the fusion process.

#### (4) Reconstruction

The fused image can be achieved by reconstruction of information in frequency domain. The algorithm of reconstruction can be modelled by the following equation (5).

$$C_r = H_r^* H_c^* C_{j+1} + H_r^* G_c^* D_{j+1}^1 + G_r^* H_c^* D_{j+1}^2 + G_r^* G_c^* D_{j+1}^3 \quad (5)$$

It is worthy noting that the choice of the wavelet basis and the number of decomposition level do affect the fused image. In the experiment, we can compare the fused images with each other and obtain the best parameters.

### C. Feature Selection and Quantitive Assessment

In the process of images fusion, it is critical to select appropriate fusion rule. Since the useful features in the images usually are larger, the pixel-by-pixel selection rule

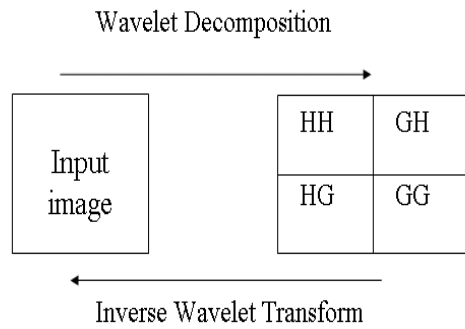


Figure 1. One level diagram of wavelet

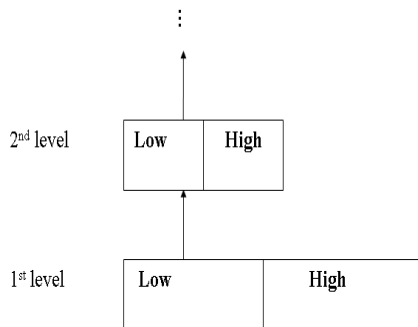


Figure 2. The pyramid structure of decomposition

may be appropriate method. In this paper the feature selection or fusion rule we adopt is the pixel-based fusion rule [10].

The quality assessment of fused image can be performed by quantitative method and by visual interpretation. Although it is direct and simple to implement the visual analysis, this assessment is subjective. In order to evaluate the improvement of the fused images, we should take into account both the spectral and spatial features. In this paper, the assessment is performed by statistical quantitative method. Three kinds of statistical parameters can be taken to analyse and evaluate the fused image. The first kind of factors reflects the grey information, such as mean, standard deviation. The second reflects the spatial information, such as entropy. The last can be used to provide spectral features, such as distortion, bias index [9].

### 3. Feature selection and classification based on residual error

After fusion, we can implement feature selection and classification based on fused image to confirm the effectiveness of fusion and to obtain higher classification accuracy.

#### A. Residual Error

Let us consider a remote sensing image in which a generic pixel, described by an  $n$ -dimensional feature vector  $x = (x_1, x_2, \dots, x_n)$  in the feature space  $X$ , is to be assigned to one of  $p$  different land-cover classes  $\Omega = (w_1, w_2, \dots, w_p)$  characterized by the a priori probabilities  $P(w_i)$ ,  $w_i \in \Omega$ . Let  $P(x/w_i)$  be the conditional density function for the feature vector  $x$  given the class  $w_i \in \Omega$ . It is well known that a classifier based on BRME assigns the pixel characterized by the feature vector  $x$  to the

class  $w_k$  if the posterior probability  $P(w_k/x)$  is the highest one [17] as in (6)

$$x \in w_k, \text{ if } P(w_k/x) = \max_{w_i \in \Omega} \{P(w_i/x)\}. \quad (6)$$

The optimality of this rule guarantees to obtain the minimum possible error probability for the considered classification problem. In real remote sensing classification problems the work we need to do is to get the posterior probability. This is particularly true when advanced classification method such as neural network, if proper learning techniques and appropriate training set, can provide an approximation of the conditional posterior probabilities of classes. However, the accuracy of the estimation process strongly depends on the training phase of the neural classifier, which usually is carried out according to the optimization of the global cost functions [18]. In the context of a large number of land-cover classes or problem with high complexity, this learning phase may result in the estimation of posterior probabilities that are not sufficiently accurate. As a consequence, in real applications often the error results to be a theoretical limit to the error probability and cannot be reached by standard classifiers. The error between the conditional posterior probability of class obtained by the classifier and real conditional posterior probability can be described by residual error, i.e.

$$r_i = P(w_i/x) - \hat{P}(w_i/x) \quad (7)$$

Where  $\hat{P}(w_i/x)$  denotes the value of the conditional posterior probability of class  $w_i$  obtained by the classifier. The residual error of each class ( $RE_i$ ) and average residual error ( $ARE$ ) can be described by the following equations (8) (9), Where  $c_i$  represents the number of samples in class  $i$ , and  $p$  is the number of land-cover classes.

$$RE_i = \frac{\sum_{j=1}^{c_i} r_i / x_j}{c_i} \quad (8)$$

$$ARE = \frac{\sum_{i=1}^{\Omega} RE_i}{p} \quad (9)$$

From the above description,  $ARE$  can be regarded as an index used to assess the classification accuracy. If  $ARE$  is lower, the classification accuracy will be higher correspondingly.

In addition, given a specific training set, if effective features are extracted, the  $ARE$  will be smaller correspondingly and the classifier is able to obtain optimized estimations. On these considerations, we can focus the attention on the choice of effective features, which will bring the lower residual error, to improve the classification accuracy.

#### B. Feature Selection based on Residual Error

In remote sensing, besides the features related to the spectral channels acquired by sensors, other features extracted by the processing of the information contained in these spectral channels (e.g., texture features) are often considered. Even though these features may increase the capability to distinguish land-cover classes, the resulting feature set often contains redundant information. Consequently, in the phase of the system design, it is

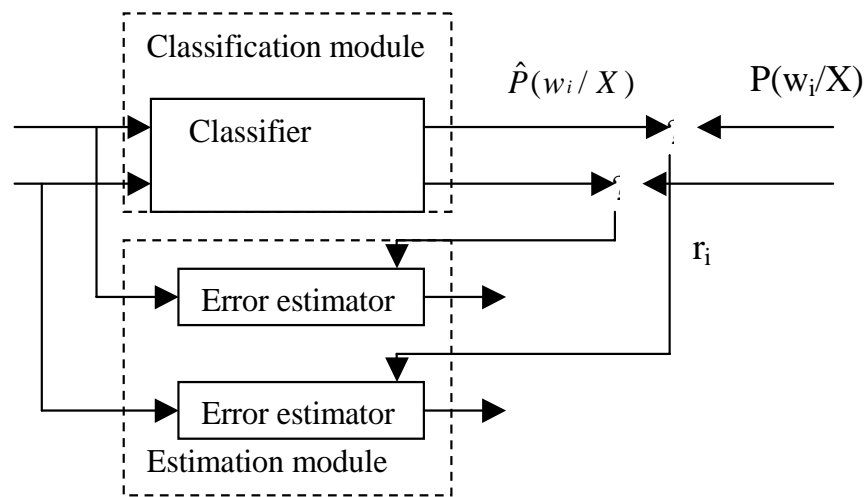


Figure3 Architecture of the proposed residual-based approach

lowest residual error. Then we can obtain most effective features according to the behavior of residual error.

The application of a feature selection based on residual error requires the definition of residual error, the training of the classifier. The computation of residual error is a critical step. The architecture is shown in the following Figure 3.

The architecture is composed of two modules: classification module and estimation module. Firstly, the proposed architecture requires training the classification module. The learning of the classifier can be carried out in a standard way to the use of a training set and of a proper learning technique. The only constraint we have is that this module should estimate the conditional posterior probabilities of class. The learning of the residual error is more critical. We can implement the learning according the conditional posterior probability of class obtained by the classifier and real conditional posterior probability (see Equation 7). Average residual error is trained according to the use of a set of samples (training set), which should be as more uncorrelated with the following test data. This is necessary in order to reduce the risk of overfitting the test data.

In our experiments, we implement the classifier according to an artificial neural network module: error back-propagation (BP) neural network. This choice depends on the following consideration: 1) this neural model, if properly trained [7][11], provide in output an approximation of conditional posterior probabilities of classes; 2) it is distribution free classifier that can be applied to multisource data set ; 3) it is proved to be effective in many remote sensing classification problems.

#### C. BP Neural Network Classifier

Corresponding to multi-source remote sensing images classification, we apply the extracted features from remote sensing images to the inputs of network. After trained by some rule and allow the signals to propagate through the network, we can implement classification in output layer. The

number of nodes in input layer and output layer should equal to the number of features and classes, respectively. The training process of network is crucial. We can obtain weights and thresholds of network by training samples. The detailed information can refer to [20]. The classifier is not only used to select features but also applied to classify. So, the training set should be irrelevant to test set to avoid the risk of overfitting.

## 4. Experiments

### A. Fusion

In the experiment, a pair of TM4 and TM5 images of the same area is adopted. Figure 4 shows the original images and resulting fused images, and the corresponding statistical measure is stated in the Table 1.

From the view of visual interpretation, the TM5 is clearer and has more information compared with TM4. The image obtained by HIS fusion is displayed in Figure 4(c) and it can be seen that it contains many finer features that are not present in the original image and the contrast is increased. But some unique information such as the textual patterns with sharp contrast changes is reduced. This problem is well resolved by using the Mallat fusion as shown in Figure 4(d). Compared to the image in Figure 4(c), the fused image looks quite different in many regions and contains more features, for the implementation of domain localization. In addition to enhance information, it preserves high fidelity.

From the view of quantitive assessment, the fused image by Mallat includes the most information, which can be seen form the mean, standard deviation and entropy in Table 1. The bigger these coefficients are, the more information is included in the image. And it preserves the least distortion, which also can be seen from the distortion in Table 1.

From the above descriptions, we can see that the fused image by Mallat fusion is the best from the perspectives of both the visual interpretation and quantitive result. Thus, it provides a good foundation for the following classification work.

Fusion	Mean	Standard deviation	Entropy	Distortion
TM4	34.698	6.9516	4.5	
TM5	52.352	13.666	5.0	
Conventional fusion	42.383	12.508	5.2	27.2
HIS fusion	125.29	56.004	5.4	24.8
Mallat fusion	128.47	61.191	6.1	15.3

Table.1 The statistical results of fused image



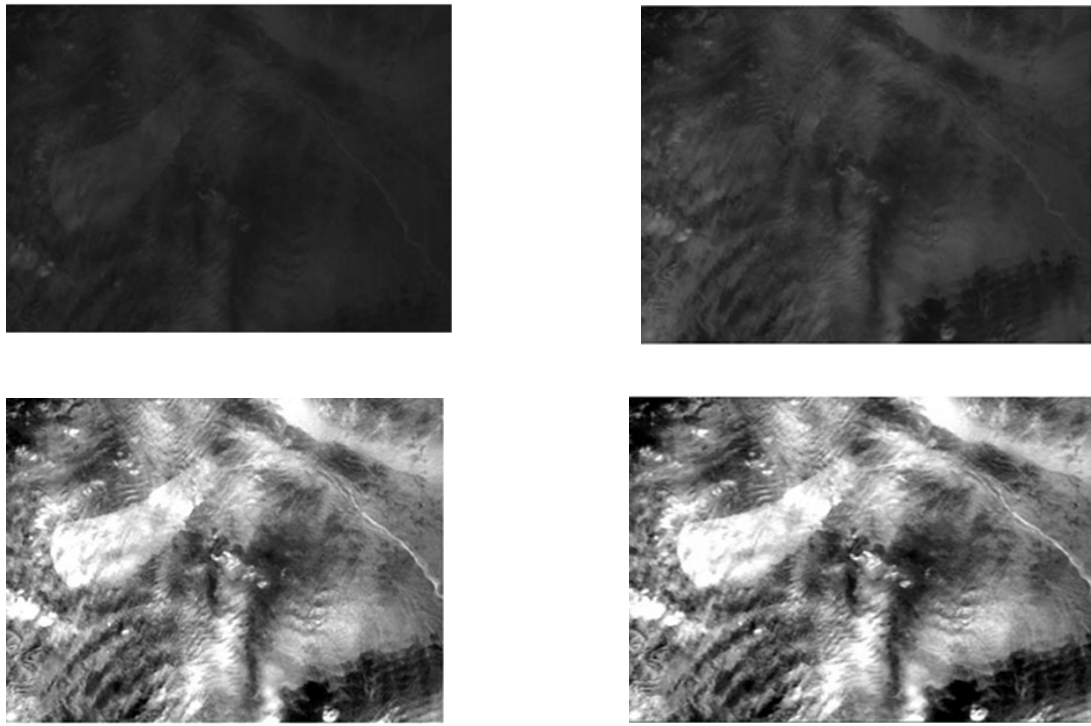


Fig. 4 (a) and (b) are a pair of TM4 and TM5 images, (c) and (d) are resulting fused image by the HIS and by the Mallat fusion algorithm, respectively.

### B Classification Preprocessing

For the experiments reported in this paper, the 10 land-cover classes were chosen. In all, 4411 pixels were selected. The pixels were randomly subdivided into two sets: 2202 training pixels were taken and 2209 test pixels (see Table 2). The training set was used for feature selection and to train classifiers, and the test set was used for performance evaluations and comparisons.

The considered TM images contain significant textual information that can be used to increase the separability of land-cover classes. In the literature, many techniques have been proposed to characterize remote-sensing image textures. For the present study, the texture features computed from the grey-level co-occurrence (GLC) matrix [21] were utilized.

The GLC matrix constitutes a statistical approach to texture computation that has been successfully tested on remote-sensing images for land-cover mapping. In theory, the 12 GLC texture features could be computed for each of the TM-channels of the selected scene, thus obtaining a set of 84 textures

In practice, in order to reduce the computational cost, the 12 texture features were computed for only one channel. In particular, we implement classification based on fused image. Therefore, the GLC texture features were computed by using such a channel included in fused image. And, among the 12 texture features of each channel, Baraldi [22] pointed out that only four texture features can represent remote sensing images best. They are Energy (E), Homogeneity (HOM), Entropy (H) and Dissimilitude (DIS), described in the equation (10)-(13), respectively, where  $L$  represents the image width or height,  $P_{\delta}(i, j)$  denotes the pixel value and  $\delta$  is the angle.

(10)

$$E = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{\delta}(i, j)^2$$

$$HOM = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{\delta}(i, j) / [1 + (i - j)^2] \quad (11)$$

$$H = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{\delta}(i, j) \cdot \log P_{\delta}(i, j) \quad (12)$$

$$DIS = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} |i - j| \cdot P_{\delta}(i, j) \quad (13)$$

In order to reduce the computation load further, only these four features are computed among the 12 texture features in our experiments. The computation of the GLC texture features requires the choice of a given number of parameter values for the computation of the GLC matrix (i.e. interpixel distance, window size and orientation). Taking into account the fine textures of the considered TM images, the GLC matrix was computed by using an interpixel distance equal to 1 pixel and a window size of 9\*9 pixels. The texture was assumed to be isotropic, and then it was computed for an angle of 0 only. The original 256 grey levels were mapped into 64 levels in order to reduce the time required by the computation of the GLC matrix and to make the estimates of the terms of the GLC matrix more reliable.

In all, features were chosen to form a feature vector for each pixel. In particular, the 6 TM channels and the 4 GLC texture features described above were considered.

C. Feature Selection and Classification based on Residual Error

Preliminary feature-selection trials were performed to find the number  $m$  of features to be given as input to the classifiers in all the carried out experiments. In particular, average residual error was used to select the best subsets of  $k$  features (with  $k=1\dots n-1$ ) from among the  $n=10$  available ones. Fig.6 shows the behaviors of average residual error, versus the numbers of selected features. By analyzing the behaviors in the diagram, one can see that, for six or seven features, the value of residual error reaches the minimum. Consequently, six or seven features were selected to carry out the experiments described in this paper.

The six best features selected by using average residual error are channel 1, 3 and 4 of the TM, 3 texture features. These six features were given as input to a fully connected BP with six input neurons, 16 neurons in the hidden layer, and 10 output neurons, in order to estimate the  $P(w_i/x)$ . The learning procedure was used to train the network, which was initialized with random weights. As a convergence criterion, a MSE smaller than 0.015 is required. Classification is performed on the test set.

To analyze the obtained results, the classification matrices achieved by the approach based on Mallat fusion and re

sidual error and original features are shown in Table 3 and 4, respectively. The terms on the diagonals of the matrices

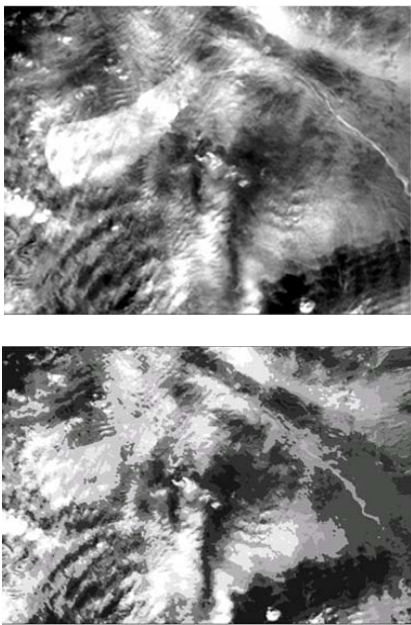


Fig.5 Fused image and classified image

Classified class	True Class										Classification accuracy
	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	$w_6$	$w_7$	$w_8$	$w_9$	$w_{10}$	
$w_1$	240	0	0	10	0	0	6	0	12	0	89.55%
$w_2$	0	318	9	0	0	0	0	0	0	0	97.24%
$w_3$	3	23	116	0	0	17	6	0	16	0	64.08%
$w_4$	23	2	0	190	0	0	15	0	13	0	78.18%
$w_5$	3	0	0	0	161	0	3	0	5	0	93.60%
$w_6$	0	0	24	0	0	152	0	0	19	0	77.94%
$w_7$	0	0	0	17	0	0	223	0	0	0	92.91%
$w_8$	11	0	0	7	0	0	0	164	9	0	85.86%
$w_9$	5	42	17	2	0	0	0	6	141	0	66.19%
$w_{10}$	10	0	12	0	0	4	0	0	7	146	81.56%

Classified class	True Class										Classification accuracy
	$w_1$	$w_2$	$w_3$	$w_4$	$w_5$	$w_6$	$w_7$	$w_8$	$w_9$	$w_{10}$	
$w_1$	249	0	0	6	0	0	10	0	3	0	92.91%
$w_2$	0	314	7	0	0	0	0	0	6	0	96.02%
$w_3$	0	9	138	0	0	21	0	0	13	0	76.24%
$w_4$	21	0	0	196	0	0	17	0	9	0	80.65%
$w_5$	0	0	0	0	167	0	0	0	5	0	97.09%
$w_6$	0	0	13	0	0	182	0	0	0	0	93.33%
$w_7$	7	0	0	23	0	0	208	0	2	0	86.66%
$w_8$	0	0	0	0	0	0	0	187	4	0	97.90%
$w_9$	0	34	16	3	0	0	0	2	158	0	74.17%
$w_{10}$	6	0	7	0	0	0	0	0	1	165	92.17%

	Average classification accuracy	Overall classification accuracy
BP	82.71%	83.79%
the proposed method	88.71%	88.91%

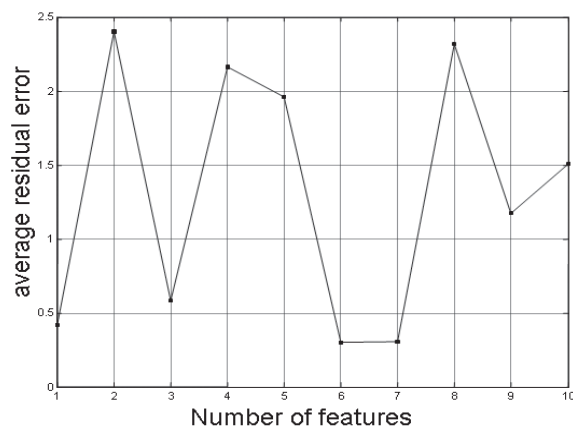


Figure.6 Behavior of the feature-selection based on residual error versus the number of selected features

give correctly recognized classes, while the other terms give the errors incurred on the pairs of classes. The classification accuracies of each class are given in the last column of the matrices. By comparing the accuracies given in the two matrices, one can see that, the approach based on Mallat fusion and residual error feature selection brings sharp increase for most of classes, although two classes have slight reduction.

To better assess the validity of the proposed approach, the total classification accuracy and average classification accuracy related to the original features are compared with one resulting from the features selected by average residual error as input to a classifier (see Table 5). In the results below, the average accuracy is defined as the average of the classification accuracy of each class, regardless of the number of samples in each class. The overall accuracy is defined as the number of correctly classified samples regardless of which class they belong to, divided by the total number of samples. In addition, a BP with the same architecture is employed in two experiments. And the same learning procedure, the same initial weights, and the same convergence criterion are adopted.

From the results in Table 5, we can deduce that the average and overall accuracy are both improved.

## 5. Conclusion

In this paper, the Mallat fusion for multi-source remote sensing images is investigated firstly, and then an approach to feature selection and classification of fused images based on residual error has been presented. From the experiments, we can find that the method performs quite well in terms of accuracies. As expected, the fused image retains more unique features, and it preserves high fidelity, which provides a more preferable image to classify. Therefore, a feature selection method which takes into account the residual error of each class based on the fused image have obtained a better feature subset. In addition, a classification technique based on selected features by using a BP neural network is implemented. The results confirm that the proposed feature selection and classification method based on fused image, as compared with the method based on original features, provide classification results in which overall and average classification accuracy are improved.

In the experiment described in this paper, there are two crucial steps. Firstly, the fusion for classification is vital and it is worthy noting the pyramid decomposition and choice of coefficient in the fusion process. Then, the proposed feature

selection method was applied by considering residual error based on fused image. Many preliminary feature selection trials are performed to find  $m$  features from  $n$  features. It's worth noting to analyze the behavior of average residual error versus features. Concerning the proposed classification technique, in this paper, the use of BP neural network to estimate posterior class probabilities has been suggested. By using such a network, one can perform a nonparametric estimation of posterior probabilities so that it is possible to process multi-source remote sensing data. However, one can utilize other parametric or nonparametric techniques to estimate these probabilities.

As a final remark, the pyramid decomposition and choice of coefficient in the fusion process and the training in the process to estimate residual error are particularly critical. The fusion result and selected features have direct relation to the performance of classifier. In the future, these aspects still need more attentions.

## References

- [1] Richards, J.A., Swain, P.H (1987). Probabilistic and evidential approaches for multisource data analysis, *IEEE Trans.Geosci.Remote Sensing*, Vol, GE-25, p.283-293, 1987.
- [2] Tian, Yan., Qin Guo, Ping., Lyu, M.R (2005). Comparative Studies on Feature Extraction Methods for Multispectral Remote Sensing Image Classification", *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, 2005. p. 1275-1279.
- [3] Bao, Qian and Guo, Ping (2004). Comparative Studies on Similarity Measures for Remote Sensing Image Retrieval", *Proceedings of the IEEE Conference on Systems, Man and Cybernetics*, 2004, p.1112~1116.
- [4] Leckei, D.G. (1990). Synergism of synthetic aperture radar and visible data for forest type discrimination", *Photogramm.Eng.Remote sens.*, 56, 1237-1246.
- [5] Tian, Yanqin., Guo,Ping., Lu, Hanqing (2004). Texture Feature Extraction of Multiband Remote Sensing Image Based on Gray Level Co-occurrence Matrix, *Computer Science*, 31 (12) 62-163 (195) (In Chinese).
- [6] Jiang, Qingjuan., Tan, Jingxin (2003). The Methods of the Pixel Level Image Fusion and the Choice of the Methods", *Journal of Computer Engineering and Application*, p.116-120, February.
- [7] Zhao, Zhigang., Chen, Xuequan (2000). Image classification Based on Data Fusion, *In: Proceedings of the 3rd World Congress on Intelligent Control and Automation*, June 28-July 2, 2000, Hefei, P.R China
- [8] Briem, Gunnar Jakobx ., Benediktsson, Gunnar Jakob., Sveinsson, Johannes, R (2002). Multiple Classifiers Applied to Multisource Remote Sensing Data, *IEEE Trans. Geosci, Remote Sensing*, 40 (10) October.
- [9] Li, Hui., Manjunath, B.S., Mitra, Sanjit K (1994). Multi-sensor Image Fusion Using the Wavelet Transform, *In: Proceedings of IEEE International Conference on Image Processing*, Vol.1, p. 51-55.
- [10] Benediktsson, Jon. A., Swain P. H., Ersoy, O.K (1990). Neural Network Approaches versus Statistical Methods in Classification of Multisource Remote Sensing Data, *IEEE Trans. Geosci and Remote Sensing*, 28 (4) July.

- [11] Mallat, S.G (1989). A theory for multiresolution on signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Machine Intel*, 11. 674-693, July.
- [12] Aggarwal, J (1993). Multisensor Fusion for Computer Vision, *Springer-Verlag*.
- [13] Liu, Li., Wang (2004). Comparison of Two Methods of Fusing Remote Sensing Images with Fidelity of Spectral Information, *Journal of Image and Graphics*, 9 (11) Nov. 2004.
- [14] Lee, T., Richards, J. A., Swain, P. H (1987). Probabilistic and evidential approaches for multisource data analysis, *IEEE Trans. Geosci Remote Sensing*, GE-25, p. 283–293.
- [15] H'egar-Masclé, S., Lel. Bloch, I., Vidal-Madjar, D (1997). Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing, *IEEE Trans. Geosci. Remote Sens.*, 35, 1018–1031, July.
- [16] Serpico, S. B., Roli, F (1995). Classification of multisensor remote-sensing images by structured neural networks, *IEEE Trans. Geosci. Remote Sens.*, 33, 562–578, May.
- [17] Lorenzo (2000). An Approach to Feature Selection and Classification of Remote Sensing Images Based on the Bayes Rule for Minimum Cost, *IEEE Trans. Geosci. Remote Sensing*, 38 p. 429–438, January.
- [18] Bruzzone, Carlin., Melgani (2004). A Residual-Based Approach to Classification of Remote Sensing Images, *IEEE Trans. Geosci. Remote Sensing*, 417-423.
- [19] Fukunaga, K (1990). Introduction to statistical pattern recognition, 2<sup>nd</sup> Ed, New York: Academic.
- [20] Li, Zuoyong (1998). Supervised Classification of Multispectral remote sensing images using BP Neural Network, *Journal of infrared Millim Waves*, Vol.11, April 1998.
- [21] Haralick, R.M., Shanmugan, K., Dinstein (1973). Textual Features for image classification, *IEEE Trans. Syst. Man Cyber.*, 3. 610-621.
- [22] Bardidi, A., Parmingian, F (1995) . An Investigation on the Texture Characteristics Associated with Gray Level Co-occurrence Matrix Statistical Parameters, *IEEE Transaction on Geoscience and Remote sensing*, 32 (2) 293-303.