2011 International Conference on Information and Intelligent Computing
IPCSIT vol.18 (2011) © (2011) IACSIT Press, Singapore

An Improved Multi-Classifiers Classification Methodology using Local Contextual Information for Remote Sensing Images

Lin Yilun, Wu Jianwei(Corresponding author) and Zhao Rui

School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

allen_lin@whu.edu.cn, jianwei_wu@whu.edu.cn

Abstract. Combining multiple classifiers has been regarded as an effective solution applied in remote sensing image classification. However, generally by adopting common voting principle, it faces a dilemma when setting voting threshold to balance the validity of the classified pixels and the proportion of unclassified pixels. In this paper, a multi-classifiers classification methodology considering local contextual knowledge was proposed to solve the dilemma. The whole work consists of two stages: primary classification and post classification. In the primary classification, a conservative (or near-conservative) voting principle was adopted to ensure a strong validity of the classified pixels. In post classification, a new adjusted spatial and spectral minimum distance (ASSMD) was used, integrating local contextual information of classified pixels to resolve the rest unclassified pixels. Effectiveness of this methodology was demonstrated by experiments compared with several classic classifiers. The overall accuracy increased by 14% and Kappa statistic increased by 0.19 at most. Furthermore, a sensitivity analysis of ASSMD to two factors: mask size and percentage of contextual information, has been carried out to explore the best configuration.

Keywords: multiple classifiers system; voting principle; contextual classification; remote sensing image classification.

1. Introduction

The idea of multiple classifiers systems (MCS) originated from pattern recognition study and soon became an important topic. A variety of researches have been conducted on such topic and theory of combination strategies has been consummated. The basic methods to combine classifiers can be summarized into three categories: parallel, cascading and hierarchical [1],[2]. Practically, according to the type of output that different individual classifiers produced, the combining method can also be categorized as: on abstract level, on ranked list of classes and on measurement level [3]. The MCS framework has been widely applied to many areas such as handwritten words recognition [4],[5], face recognition [6] , medical diagnosis [7] and so on.

In the past two decades, MCS increasingly drew researchers' attention on remote sensing image classification. Among all combining strategies, voting rule is the most commonly used in remote sensing [8]. This is a combination strategy on abstract level based on parallel topology. Various researches have been carried out on application and amelioration on voting rule in this field [9],[10],[11].

However, all these methodologies encountered the same dilemma when setting voting threshold: high voting threshold leads to good reliability but poor recognition rate; on the contrary, low voting threshold leads to poor reliability but good recognition rate. With regard to this dilemma, most of the researchers adopted majority voting rule (MVR). This method sacrifices, to some extent, the reliability of the classified pixels for less unclassified pixels. Moreover, the post-processing of unclassified pixels deserves further concern: some assigned unclassified pixels according to the prior-knowledge of input classifiers' performance, which entails exhaustive accuracy assessment for each individual classifier [12],[9], some adopted complex and advanced algorithm derived from pattern recognition [13]; other even didn't set threshold [14]. All these post-processing ignored the contextual information provided by the classified items.

Contextual classification has been considered as a convincing strategy for remotely sensed image classification. Richards and Jia[15] summarized them into four categories: Preprocessing, Post processing, Probabilistic label relaxation and Markov random fields. Series of researches have been done on this topic and desirable results are available [15],[16],[17].

The classification methodology proposed in this paper tries to address the above problems of voting rule by incorporating the idea of contextual classification. The target is to reasonably resolve unclassified pixels from primary classification while ensuring best validity of classified ones. Distinct from other applications of voting principle in remote sensing image classification, this methodology accepted only complete consensus in voting process and the unclassified pixels would be addressed using the knowledge extracted from classified pixels which is ignored by others.

2. Methodology

A multi-classifiers model incorporating both voting strategy and contextual classification was proposed in this paper (see Fig. 1). This methodology consists of two parts: a primary classification and a post classification. In primary classification, a high threshold voting method is adopted to

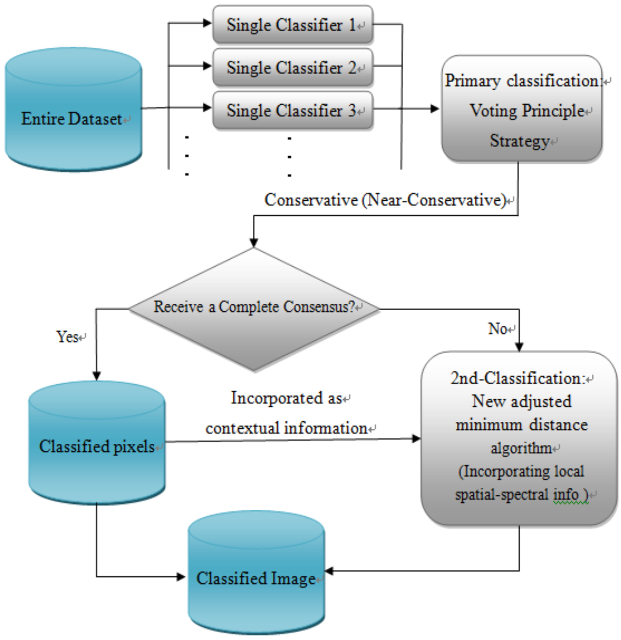

Figure 1. The multi-classifiers classification model.

combine several classic classifiers' outputs, which both guaranteed a strong authenticity of classified pixels and generated considerable amount of unclassified pixels. In the post classification, the output of voting principle was recognized as a contextual knowledge and utilized for classifying the remaining unclassified pixels. In this stage, a new adjusted spatial and spectral minimum distance (ASSMD) was devised to show spatial-spectral difference between the target pixel and its neighboring classified pixels.

2.1. Primary Classification

The primary classification combines several traditional single classifiers with conservative voting principle. The purpose of traditional classifiers is: 1). to provide several outputs on abstract level that can be used as input for the voting strategy; 2). to act as benchmark to compare with the multi-classifiers model proposed in this paper.

Voting principle is the simplest but most effective implementation method among all combination strategies. The fundamental philosophy behind voting principle is the resolution of a group is superior to that of any individual among it. According to this, a target pattern will be classified to the category that most classifiers agree on.

Assuming that a classification of remote sensing imagery involves M categories: $C_1 \cup C_2 \cup ... \cup C_i \cup ... \cup C_M$, where C_i, $\forall i \in A = \{1,2,3...M\}$ denotes a category. Each individual classifier $\varepsilon_k (m = 1,2,...K)$ assigns pixel (X) to a category, namely: $\varepsilon_k(X) = i$. The combination of multi-classifiers is to devise a abstract unified classifier E, whose output $E(X) = j$, $j \in A \cup \{M+1\}$.(M+1 is the label for pixels that E is unable to identify). Finally, the voting rule can be described as [4]:

$$E(X) = \begin{cases} j, & \text{if } T_E(X \in C_j) = \max_{i \in A} T_E(X \in C_i) \geq \alpha \times K \\ M+1, & \text{Else} \end{cases} \quad . \tag{1}$$

$$. \ T_E(X \in C_i) = \sum_{k=1}^{K} T_k(X \in C_i), \ \text{i=1,2...M} \tag{2}$$

where $T_k(X \in C_i)$ denotes the vote classifier k contributes to whether classifying sample X to category C_i (value 1 for agree, 0 for disagree), $0 < \alpha \leq 1$, $\alpha \times K$ specifies the voting threshold.

In this paper, since a conservative (or near conservative) voting method was applied, $\alpha = 1$ (or $\alpha \approx 1$)

Compared to majority voting method, although conservative voting strategy creates much more unclassified pixels, its employment has two obvious benefits:

- According to the classification model (Fig. 1), as long as a pixel was classified in primary classification, the assigned class value will be directly recognized as the final result. Conservative voting principle can therefore increase the accuracy of this part of pixels.
- If the primary classification is abstractly perceived as one classifier, the whole workflow can be regarded as a cascadingly combined multi-classifiers system according to Lu's summary [1]. Since a contextual classifiers was incorporated later which needed the result of primary classification, any flaw in the primary stage will be magnified and accumulated in next stage. Thus conservative voting rule also ensured a better performance of post classification.

2.2. Post Classification

Because a conservative voting strategy was applied in primary classification, it is possible that large proportion of unclassified pixels was generated. So an effective processing towards these pixels is the main concern in this stage.

The idea of the post classification was to use the local spatial and spectral information from primary classification to improve the labeling of unclassified pixels. A new adjusted spatial and spectral minimum distance (ASSMD) was devised as the similarity measurement in this step. Unlike the traditional measure which only considers the Euclidean distance between spectral vector of each pixel and mean spectral vector of each class from training dataset, ASSMD also incorporates local spatial relation between unclassified pixels and labeled pixels in its local mask. Luo and Mountrakis [16] developed a similarity measurement that use the summation of spatial and spectral distance between unclassified pixel and mean value of each class in local mask, and devise a ratio α to adjust the relative weight of spatial and spectral distance.

The inconvenience of their measurement is that it requires calibration for the optimum value of α, which is based on huge amount of trials beforehand. Furthermore, the value of α varies when it is applied to different image. In order to simplify the calculation of similarity measurement to increase its applicability and efficiency, we proposed an adjusted spatial and spectral minimum distance (ASSMD) similarity measurement as following:

$$Dist_{wi} = \frac{\sum_{m=1}^{\lambda_{wi}} (Dist_m^{spec} \times Dist_m^{spat})}{\lambda_{wi}} \tag{3}$$

where $Dist_{wi}$ denotes the ASSMD between pixel X (pixel to be classified) and class i, λ_{wi} is the number of pixels of class i in local mask, $Dist_m^{spec}$ and $Dist_m^{spat}$ respectively specifies the spectral and spatial distance between the pixel X and pixel that belongs to class i.

Specifically, Dist_m^{spec} was calculated as the Euclidean distance between spectral feature of the unclassified pixel and that of the classified pixel in given class; Dist_m^{spat} was calculated as the spatial distance between the unclassified pixel and the classified pixel.

To calculate ASSMD, a local proximity definition was applied to determine the mask size around an unclassified pixel. In this paper, we adopted fixed mask size strategy in each experiment.

Finally, target pixel was assigned to class w^*, which has the minimum ASSMD:

$$w^* = \min_{wi}(\mathrm{Dist}_{wi}) \tag{4}$$

3. Experiments Result and Analysis

A Landsat TM5 scene from Canon City, USA was used as trial dataset to assess the performance of the classification methodology in this paper. Fig. 2 is the standard pseudo-color image of this area. 5 classes were selected to represent

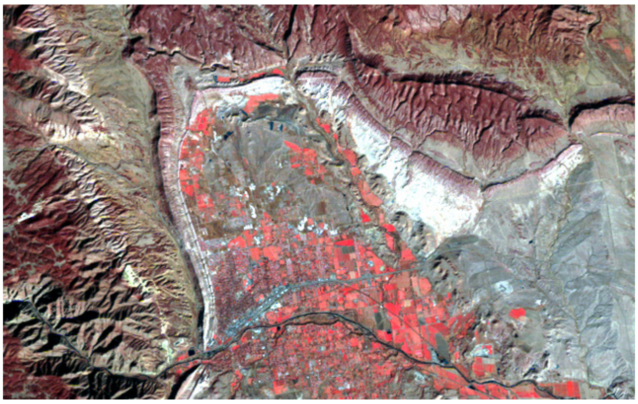

Figure 2. Experiment Dataset –Cannon's standard pseudo-color image of Landsat TM (R:4, G:3, B:2)

and classify the image: Rock, Crops, Trees, Soil and Shadow. Referenced data and training data were selected through visual interpretation.

3.1. Output of single classifiers' classification

To extend the diversity, classifiers that differ in principle were employed in this experiment. The employed classifiers include: Maximum Likelihood Classifier (MLC), Minimum Distance Classifier (MinD), Mahalanobis Distance Classifier (MahD), Neural Net Classifier (NNC), Support Vector Machine Classifier (SVM). The classified accuracy of these traditional classifiers is shown in TABLE II. Experiment result shows that, although overall accuracy of single classifier varies little, each has unique strength and weakness on distinguishing different classes.

3.2. Output of classification adopting voting method.

In this step, in order to explore diverse possibilities, we tested two conservative voting methods and one near-conservative voting method. Two conservative voting strategies are: (1) 5/5 voting rule, where MalD, MLC, MinD, NNC, SVM were incorporated and only when they reach a complete consensus wills the pixel be classified. (2) 4/4 voting, where MLC, MinD, NNC, SVM were incorporated (eliminating MalD) and only when they reach a complete consensus will the pixel be classified. The near-conservative voting method is (3) 4/5 voting where 4 out of 5 classifiers' agreement can resolve a pixel's category. The result of this stage is a partially classified image. The overall contextual information statistic was presented in TABLE I.

3.3. Output of Post Classification using ASSMD

The three inputs with different proportion of contextual information mentioned above were calculated according to the ASSMD (see (3)) respectively. Here, a fixed local mask (7*7) was applied to define the neighboring area. TABLE III illustrates the accuracy assessments of this methodology based on different voting configurations. It is obvious that overall accuracy all increased significantly compared with each used classifier (TABLE IV).

4. Further Experiments on the Sensitivity of Assmd Similarity Measurement

Assuming that authenticity of the contextual information has been assured by same means (e.g. Conservative voting method), ASSMD calculation is influenced by the following two factors: local mask size and proportion of contextual information.

4.1. On the variation of local mask size

To test the influence of local mask size, a set of subset

TABLE I. OVERALL CONTEXTUAL INFORMATION STATISTIC

Proportion	4/5	5/5	4/4
Unclassified Pixels(%)	16.895	41.909	26.814

images were selected as the trial data. We calculated ASSMD with different local mask sizes varying from 3*3, 5*5, up to 11*11. The overall accuracy and kappa statistic could be seen in Fig. 3. (It is worth noting that because the trial data was subset rather than the whole image, the specific accuracy value might be different from the accuracy assessment result in TABLE III.) A trend could be figured out easily: when using a comparatively small mask size, the increase of mask size contributes considerably positively to the increase of accuracy and kappa statistic, however, when mask size reach a certain extent, accuracy and kappa statistic remains stable, even drop slightly. In this trial, setting 7*7 or 9*9 could get optimum classification result.

4.2. On the variation of percentage of contextual information

This experiment tries to explore the relation between the percentage of contextual information and the result of ASSMD calculation. We selected a series of subset images that contain different percentage classified pixels as the trial data. All subset images were tailored from the classification result that is generated by applying 7*7 mask size and 4/4 voting rule. Percentage of contextual information varies from 70% 75% up to 90%. Accuracy Assessments were shown in Fig. 4. (Also, note that because the trial data was subset rather than the whole image, the accuracy in this trial might be different from what is in TABLE III.) It is evident that, when contextual information is highly reliable, less

TABLE II. ACCURACY ASSESSMENT FOR INDIVUDUAL CLASSIFIERS

Accuracy%	MahD	MinD	MLC	NNC	SVM
Rock					
Producer’s	66.67	85.71	100	87.50	100
User’s	28.57	54.55	40.00	43.75	66.67
Crops					
Producer’s	87.50	83.33	80.00	92.31	100
User’s	87.50	90.91	88.89	85.71	100
Trees					
Producer’s	60.00	87.50	92.31	73.68	92.31
User’s	52.94	43.75	57.14	58.33	34.29
Soil					
Producer’s	64.86	57.58	65.71	61.54	65.22
User’s	77.42	90.48	95.83	86.49	97.83
Shadow					
Producer’s	87.50	100	60.00	92.31	83.33
User’s	87.50	90.91	100	85.71	100
Overall accuracy	69.01	74.29	74.63	73.33	74.77
Kappa	0.5632	0.6649	0.6475	0.6394	0.6195

TABLE III. ACCURACY ASSESSMENT FOR THIS METHODOLOGY ON VARIOUS VOTING CONFIGURATION

Accuracy%	4/5	5/5	4/4
Rock			
Producer’s	87.50	71.43	87.50
User’s	87.50	83.33	87.50
Crops			
Producer’s	75.00	70.00	87.50
User’s	100.00	87.50	100.00

Trees			
Producer's	94.74	88.00	91.67
User's	54.55	59.46	61.11
Soil			
Producer's	81.82	79.45	86.36
User's	93.10	90.63	95.00
Shadow			
Producer's	55.56	80.00	44.44
User's	100	80.00	100
Overall accuracy	81.82	80.00	84.35
Kappa	0.7062	0.6673	0.7489

TABLE IV. COMPARISON BETWEEN BENCHMARKS AND METHODOLOGY IN THIS PAPER ON OVERALL ACCURACY AND KAPPA STATISTICS

Accuracy	4/5	5/5	4/4	MahD	MinD	MLC	NNC	SVM
Overall(%)	81.82	80.00	84.35	69.01	74.29	74.63	73.33	74.77
Kappa	0.7062	0.6673	0.7489	0.5632	0.6649	0.6475	0.6394	0.6195

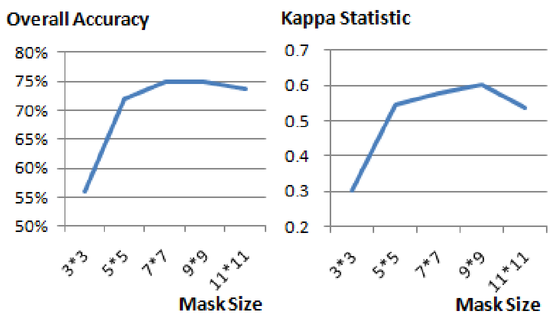

Figure 3. Accuracy Assessment of ASS\MD algorithms with different local mask size

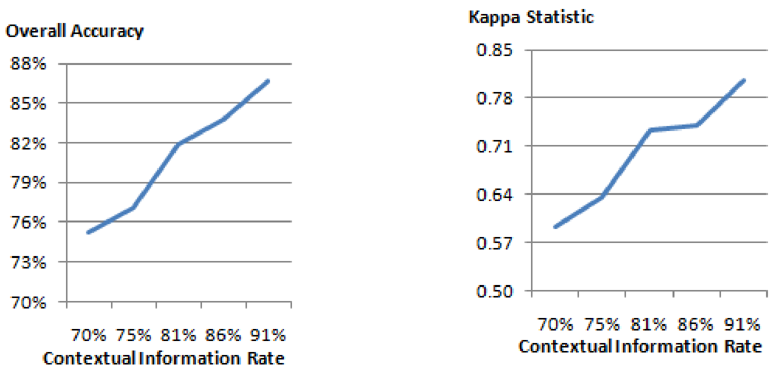

Figure 4. Accuracy Assessment of ASSMD algorithm with different contextual information proportion

amount of contextual information leads to the decrease of overall accuracy.

5. Discussion and Conclusion

The novelty of this paper consists in:

- Proposed a practically feasible classification methodology to utilize conservative voting method, which both ensures the validity of the classification result and makes the most of it to address the unclassified items.
- Put forward ASSMD similarity measurement used in contextual classification and providing examination and specification to its characteristics.

The initial idea of proposing this methodology is to find a practical way to apply conservative voting principle in order to take advantage of its reliability. Hence, this methodology could be regarded as improvement of voting method which has wide application in remotely sensed image classification. In another perspective, since the primary classification used a parallel combination topology and its output was utilized as the input by the post classification, this methodology could be considered as a hybrid MCS of parallel and cascading. Considering the parallel and cascading configuration have complementary advantages [2], this hybrid can make up for the disadvantage of cascading classifiers by offering a high quality input.

It is important to notice that from the discussion in 3.4.2 we predicted a trend that greater amount of contextual information leads to better performance of ASSMD calculation. However, when result of 4/5 and 4/4 strategies in Table 3 were referred to, it might be curious to find out that greater proportion of known contextual knowledge does not ensure a higher accuracy. This is probably because among the classified pixels, the accuracy of conservative method (4/4) is higher than that of the majority method (4/5) hence providing more authentic contextual information to be input to post classification. As a result, it might be reasonable to claim that as far as the contextual information is concerned, both authenticity and proportion are factors that contribute to the final accuracy of this methodology.

As far as ASSMD is concerned, it has been demonstrated as a simple-applicable contextual classifier with only one parameter (local mask size) to calibrate and configure. Because ASSMD calculation requires only a partially classified result with multi-spectral data, primary classifier can be substituted by other classifiers as long as they generate the same kind of result. This suggests ASSMD's broad application as a similarity measurement. The idea of ASSMD is also consistent with Tobler's first law of geography claiming "everything is related to everything else, near things are more related than distanced things".

The proposed classification methodology established a framework that combines multi-classifiers by voting method and integrates classified pixels' spatial and spectral information to address the unclassified pixels. Experiment result shows that overall accuracy increase 14%, Kappa statistics increase 0.19 at most compared with traditional classifier. A new adjusted spatial and spectral minimum distance (ASSMD) has been used and validated as a similarity measurement. Future work may focus on the refining of the ASSMD's organization which involves a more physically sound model(e.g. Gauss Diffuse Model) and further exploration on the influence of mask size and contextual information.

6. References

[1] Y. Lu, "Knowledge integration in a multiple classifier system," Appl. Intell.6(2) (1996), 75–86

[2] R. Ranawana and V. Palade, "Multi-Classifier Systems: Review and a roadmap for developers," International Journal of Hybrid Intelligent Systems, v.3 n.1, p.35-61, January 2006.

[3] C. Suen and L. Lam, "Multiple classifier combination methodologies for different output levels," First International Workshop, MCS 2000, vol. 1857, pages 52–66, 2000.

[4] L. Xu, A. Krzyzak and C.Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," IEEE Trans. Systems, Man, and Cybernetics, vol. 22, pp. 418-435, 1992.

[5] X. Zhu, Y. Shi and S. Ma, "Research on handwritten character recognition," PR & AI, vol. 13, no. 2, 174-180, 2000.

[6] X. Jing and D. Zhang, "Face recognition based on linear classifiers combination," Neurocomputing, vol. 50, pp. 485-488

[7] X.C. Yin, C.P. Liu and Z. Han, "Feature combination using boosting," Pattern Recognition Letters 26(16): 2195-2205, 2005.

[8] F. Roli, G. Giacinto, and G. Vernazza, "Comparison and combination of statistical and neural network algorithms for remote-sensing image classification," Neurocomputation in Remote Sensing Data Analysis, pp. 117-124, 1997

[9] X. Zhang, X. Feng and W. Liu, "Urban vegetation categories recognition by multiple classifier system from IKONOS imagery," Journal of Southeast University(Natural Science Edition), 5: 399-403, 2007.

[10] L. Bruzzone, R. Cossu, and G. Vernazza, "Combining parametric and non-parametric algorithms for a partially unsupervised classification of multitemporal remote-sensing images," Inf. Fusion, vol. 3, no. 4, pp. 289-297, Dec. 2002

[11] Lv, Y., Shi, P.F., Zhao, Y.M., "Voting Principle for Combination of Multiple Classifiers," Journal of Shanghai Jiao Tong University 34(5), 680–684 (2000).

[12] X.H. Liu, A.K. Skidmore and H. Van Oosten, Integration of classification methods for improvement of land-cover map accuracy, ISPRS Journal of Photogrammetry and Remote Sensing **56** (4) (2002), pp. 257–268

[13] H. Zhou, Q. Lin, Y. Wu and Q. Wang, "Application of combining classifiers based on AdaBoost to remote sensing classification," Application Research of Computers, vol. 24, no. 10, pages: 181-184.

[14] Y. Bo and J. Wang, "Combining multiple classifiers for thematic classification of remotely sensed data," Journal of Remote Sensing, vol.9, no. 5, pp. 555-563. 2005

[15] J. A. Richards, Remote Sensing Digital Image Analysis: An Introduction, Springer-Verlag New York, Inc., Secaucus, NJ, 1993.

[16] Luo, L. and Mountrakis, G., "Converting local spectral and spatial information from a priori classifiers into contextual knowledge for impervious surface classification," ISPRS J. Photogram. Remote Sensing (2011), doi:10.1016/j.isprsjprs.2011.03.002

[17] De Jong, S.M., Hornstra, T., Maas, H.G., "An integrated spatial and spectral approach to the classification of Mediterranean land cover types: the SSC method," International Journal of Applied Geoscience 3 (2), 176–183, 2001.