



Classification of Remote Sensing Images Based on Band Selection and Multi-mode Feature Fusion

Xiaodong Yu^{1,2}, Hongbin Dong², Zihe Mu³, and Yu Sun⁴(✉)

- ¹ College of Computer Science and Technology, Harbin Normal University, Harbin, China
- ² College of Computer Science and Technology, Harbin Engineering University, Harbin, China
{yuxiaodong, donghongbin}@hrbeu.edu.cn
- ³ College of Mathematical Sciences, Harbin Normal University, Harbin, China
mzh19961102@126.com
- ⁴ Heilongjiang Institute of Construction Technology, Harbin, China
donglin_2016@163.com

Abstract. As feature data in multimodal remote sensing images belong to multiple modes and are complementary to each other, the traditional method of single-mode data analysis and processing cannot effectively fuse the data of different modes and express the correlation between different modes. In order to solve this problem, make better fusion of different modal data and the relationship between the said features, this paper proposes a fusion method of multiple modal spectral characteristics and radar remote sensing image according to the spatial dimension in the form of a vector or matrix for effective integration, by training the SVM model. Experimental results show that the method based on band selection and multi-mode feature fusion can effectively improve the robustness of remote sensing image features. Compared with other methods, the fusion method can achieve higher classification accuracy and better classification effect.

Keywords: Remote sensing classification · Classification of features · Band selection · Multimodal feature fusion · SVM

1 Introduction

Remote sensing images can accurately and comprehensively describe the features of ground objects and are widely used in agriculture, geological exploration, environmental monitoring and military reconnaissance [1–4]. At present, many algorithms are used in identification classification, such as maximum likelihood method [5], minimum distance method [6], artificial neural network [7] (ANN), k-means clustering [8], and so on. Theoretically, the sample size of these methods must be large enough to ensure high-level, but in practice, it is not guaranteed to select enough classification samples. Support vector machine (SVM) method is a new type of data mining method [9], which

has outstanding advantages in small sample, nonlinear and high-dimensional pattern recognition. Therefore, it can be widely used in distributed classification research and has achieved good results.

At present, the rapid development of satellite technology and the increasing number of sensors not only improve the quality of remote sensing images, but also result in the coexistence of multiple types and resolutions of remote sensing images. Optical sensor and synthetic aperture radar image, as the most important two kinds of sensor data, have their own advantages and disadvantages. Although they both have excellent performance in ground object identification, they rarely cooperate with each other and fail to make full use of the existing surface information. The fusion technology of remote sensing image came into being. The image fusion technology is classified according to different standards. According to the level of fusion, the fusion method is divided into three categories: pixel level, feature level and decision level. In the research of multi-source remote sensing image fusion method, the pixel-level fusion is widely used in the field of remote sensing image fusion due to its advantages of high computing accuracy, small data change and less information loss.

In the past 20 years, many researchers at home and abroad have proposed many fusion algorithms of different strategies. Ramakrishnan et al. fused panchromatic images with multispectral images using the improved IHS transform, and also discussed the application of fused images in earth science [10]. According to the spectral characteristics and application prospects of multi-spectral bands in Ref. [11], PCA, IHS and HPF methods were compared in maintaining the spectral quality of fusion images. Dupas proposed using Brovery transform and IHS method to fuse SAR images and multi-spectral images for land cover classification [12]. Single sensor data and fusion data are used for classification, and the overall classification accuracy of fusion data is higher than that of pure multi-spectral data. Chen et al. used IHS transformation to merge hyperspectral images with SAR images at hyperspectral resolution to obtain enhanced features of urban areas [13]. The fused image was superimposed on the digital elevation model (DEM) to obtain a three-dimensional image. This integration helps to resolve the ambiguity of different types of land cover. Pal et al. proposed a PCA-based fusion method to improve geological interpretation [14]. FFT method was used to filter SAR images, PCA was used to fuse multi-spectral data, and then the principal component selection method based on features was used to conduct false color synthesis of the fused data. Chandrakanth et al. applied HPF to the frequency domain and fused high-resolution panchromatic and high-resolution SAR images [15]. This method extracted low-frequency details from panchromatic images and high-frequency details from SAR images. Battsengel et al. proposed the fusion of high-resolution optical images and SAR images to improve the accuracy of land cover classification [16]. The method of multiplication was used to fuse the source images, and the fusion method based on BT, PCA and IHS was adopted. The results were compared visually on the basis of classification accuracy. Yang et al. used GS orthogonalization method to integrate optical images and SAR images to improve the classification of coastal wetlands [17]. The above methods are not computationally complex and can produce a fusion product with rich spatial information. However, these methods are highly dependent on the correlation of image fusion, so that the image after fusion has obvious spectral distortion.

In view of the above problems, this paper first selected the bands of Landsat8 images, selected the optimal three bands, and used HSV transform to effectively integrate the multi-spectral image features and sentinel-1 SAR image features in the form of vector or matrix according to the spatial dimension, so as to train the SVM model and obtain better ground object recognition effect.

2 Classification Principle of Support Vector Machines

The mechanism of support vector machine (SVM) is to find an optimal classification hyperplane that meets the classification requirements, so that the hyperplane can maximize the empty space on both sides of the hyperplane while ensuring the classification accuracy. Taking two types of data as examples, the given training set (x_i, y_i) , $i = 1, \dots, l$; $x \in R^n$, $y \in \{+1, -1\}$ and hyperplane are denoted as $(w \cdot x) + b = 0$. In order for the classifier to correctly classify all samples and have classification intervals, the following constraints need to be met:

$$y_i[(w \cdot x_i) + b] \geq 1 \quad i = 1, 2, \dots, l \quad (1)$$

We can calculate the classification interval $2/\|w\|$, so the problem of constructing the optimal hyperplane is transformed into:

$$\min \phi(w) = \frac{1}{2} \|w\|^2 = \frac{1}{2} (w' \cdot w) \quad (2)$$

In high-dimensional space, if the training sample is indivisible or it is not known whether it is linearly separable, it will allow the introduction of a certain number of misclassified samples, and a non-negative relaxation variable ζ_i will be introduced. The above problem is transformed into a quadratic programming problem with linear constraints:

$$\min \frac{1}{2} (w' \cdot w + C \sum_{i=1}^n \zeta_i) \quad i = 1, 2, \dots, n \quad (3)$$

However, in actual classification, it is difficult to ensure linear separability between categories. For the case of linear inseparability, SVM introduces a kernel function, maps the input vector to a high-dimensional feature vector space, and constructs the best classification surface in the feature space. Due to the good performance of the RBF kernel function, the radial kernel function is selected as the kernel function of the SVM in practical applications.

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (4)$$

3 Classification of Remote Sensing Images Based on Band Selection and Multi-mode Feature Fusion

3.1 Correlation-Based Feature Selection (CFS)

CFS [18] is a classic filtering type feature selection algorithm. It evaluates the influence of each feature on each classification separately by heuristics, so as to obtain the final feature subset. The evaluation method is shown in formula (5):

$$M_s = \frac{kr_{cf}}{\sqrt{k + k(k-1)r_{ff}}} \quad (5)$$

Among them, M_s represents the evaluation of the feature subset S containing k features; r_{cf} represents the average correlation between attributes and classification; r_{ff} represents the average correlation between attributes. The higher the computational correlation between each attribute and the classification attribute, or the lower the redundancy between the attributes, the more positive the evaluation. In CFS, information augmentation is used to calculate correlations between attributes. The method of computing information enhancement will be described below.

Assuming that the attribute is Y , γ represents each possible Y value, the entropy of Y is calculated from Eq. (6).

$$H(Y) = - \sum_{\gamma \in Y} P(\gamma) \log_2(P(\gamma)) \quad (6)$$

For the known attribute X , Y 's entropy is calculated by Eq. (7)

$$H(Y \setminus X) = - \sum_{x \in X} P(x) \sum_{\gamma \in Y} P(\gamma \setminus x) \log_2(P(\gamma \setminus x)) \quad (7)$$

The difference $H(Y) - H(Y \setminus X)$ (namely the reduction of feature Y 's entropy) reflects the additional information attribute X provides to attribute Y , also known as information enhancement. The information increment reflects the amount of information provided by X to Y , so the larger the information increment, the stronger the correlation. Since information enhancement is a symmetric measure, its disadvantage is that it tends to select those attributes with more potential values. Therefore, each attribute needs to be information enhanced and normalized to ensure that each attribute can be compared with other attributes so that different attribute selections can get the same result. The method of symmetric uncertainty is used to normalize it into $[0, 1]$.

$$U_{XY} = 2.0 \times \frac{H(Y) - H(Y \setminus X)}{H(Y) + H(X)} \quad (8)$$

3.2 HSV Transform Fusion Method

HSV transform is an inverted cone color space composed of Hue, Saturation and Value, which transforms the color space composed of RGB three colors in multi-spectral images [19]. The basic idea of HSV transformation is to replace the brightness Value in the original image with a high-resolution full-color remote sensing image, and then re-sample the Hue and Saturation into the high-resolution size image through interpolation (nearest neighbor method, bilinear interpolation method and cubic convolution interpolation method), and finally convert the image back to the RGB color space.

$$H = \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases} \quad (9)$$

$$S = 1 - \frac{3}{(R + G + B)} \min(R, G, B) \quad (10)$$

$$V = \frac{1}{3}(R + G + B) \quad (11)$$

$$\text{Among them : } \theta = \arccos \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{\left[(R - G)^2 + (R - B)(G - B) \right]^{\frac{1}{2}}} \right\}$$

4 Experiments and Analysis

4.1 Data Source and Preprocessing

Landsat8 remote sensing data from September 10, 2015 and sentinel-1 standard polarization model data were mainly used in this paper. The Sentinel-1 standard polarization model data included a scene on December 23, 2015, with a VV spatial resolution of 15 m and an incidence Angle of 37°. The data processing of this paper mainly includes the following parts:

Landsat8 remote sensing data were imported into ENVI5.3 software, and on the basis of 1:50,000 topographic map, the intersection points of rivers and roads were selected as control points for geometric correction of Landsat8 remote sensing data. Afterwards, the 1–7 bands of Landsat8 remote sensing data were fused with panchromatic bands, and the required study areas were intercepted according to the research needs.

Using the ENVI extension package SAR scape, the Sentinel-1 standard polarization pattern data was radiative scaled, and the amplitude intensity image was converted into the backscattering coefficient image. Then the Lee filtering algorithm of 7×7 filter window is used to remove the speckle noise of radar image. The filtered image was converted into a transverse Mercator (UTM) projection and resampled to 30 m. Due to the different sensor of acquired radar image and optical image, the imaging mode of

each satellite is different, so the space registration should be carried out. Using ENVI5.3 software, the radar images were registered based on Landsat8 image data dated 10 September 2015.

4.2 Experimental Data and Sample Selection

Study area selection, western songnen plains in western jilin province county territory of adlai was in charge of the town, momo, national nature reserve ($44^{\circ}51'N \sim 46^{\circ}18'N$, $123^{\circ}07'E \sim 124^{\circ}04'E$). According to the distribution of land features in this area, it is divided into 6 categories: unused land, grassland, marsh, wetland, water area, residential land and cultivated land. In this paper, a total of 16 160 pixel-level training samples were selected, among which 30% of each ground object was randomly selected as the verification set, and the remaining data samples were selected as the training set.

4.3 Classification and Accuracy Verification

SVM classifier using ENVI5.3 remote sensing software. The spectral characteristics of 1–7 bands of Landsat8 remote sensing data, spectral characteristics after band selection (CFS), and spectral characteristics after band selection (CFS) and radar fusion were used respectively to extract wetland information by integrating topographic assistance data and radar image backscattering. The results are shown in Fig. 1 (a), Fig. 1 (b) and Fig. 1 (c). The classification results of 578 measured sample points were selected for accuracy verification, and the error confusion Matrix was established by using the tool ENVI5.3 confusion Matrix. The accuracy evaluation results obtained are shown in Table 1, Table 2 and Table 3 respectively.

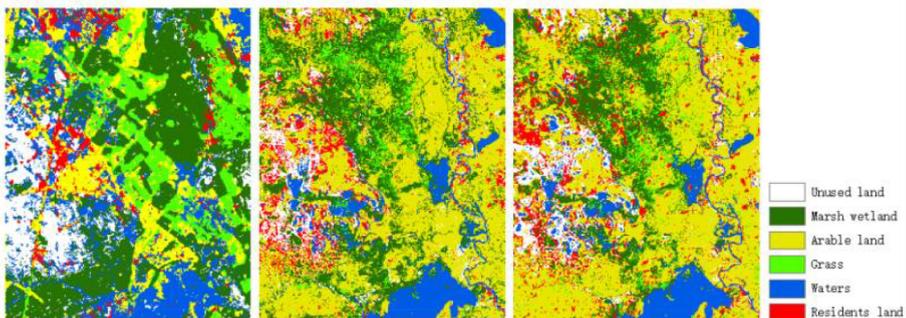


Fig. 1. (a), (b), (c) Results of extracting wetland information using SVM with three different features

Table 1. Confusion matrix of SVM classification results for spectral characteristics of 1–7 bands of Landsat8 remote sensing data

The sample points	Arable land	Grass	Marsh wetland	Waters	Residents land	Unused land	Total	User accuracy %
Arable land	82	10	25	4	3	0	124	66.13
Grass	0	39	21	14	0	0	74	52.70
Marsh wetland	4	9	93	0	0	0	106	87.74
Waters	17	3	31	60	14	10	135	44.44
Residents land	3	1	0	0	85	0	88	96.59
Unused land	0	0	0	0	0	51	51	100.00
Total	106	61	170	78	102	642	578	
Mapping accuracy %	77.36	63.93	54.71	76.92	83.33	83.61		

The overall accuracy = 70.93% Kappa = 0.64

Table 2. Confusion matrix of SVM classification results for spectral characteristics after band selection (CFS)

The sample points	Arable land	Grass	Marsh wetland	Waters	Residents land	Unused land	Total	User accuracy %
Arable land	94	0	14	0	12	2	122	77.05
Grass	0	70	14	0	0	0	84	83.33
Marsh wetland	6	19	89	0	1	0	115	77.39
Waters	0	0	0	107	0	0	107	100.00
Residents land	0	0	0	0	48	39	87	55.17
Unused land	0	0	0	0	19	44	63	69.84
Total	100	89	117	107	80	85	578	
Mapping accuracy %	94.0	78.65	76.07	100.00	60.00	51.76		

The overall accuracy = 78.20% Kappa = 0.73

Table 3. Confusion matrix of SVM classification results for band selection (CFS) after the spectrum and radar fusion characteristics

The sample points	Arable land	Grass	Marsh wetland	Waters	Residents land	Unused land	Total	User accuracy %
Arable land	102	0	13	0	0	4	125	81.60
Grass	0	62	13	0	6	0	75	82.67
Marsh wetland	1	14	101	0	0	0	116	87.07
Waters	0	0	0	106	0	0	106	100.00

(continued)

Table 3. (continued)

The sample points	Arable land	Grass	Marsh wetland	Waters	Residents land	Unused land	Total	User accuracy %
Residents land	6	0	1	0	78	4	88	88.64
Unused land	0	0	0	0	14	54	68	79.41
Total	109	76	127	106	98	62	578	
Mapping accuracy %	93.58	81.58	79.53	100.00	79.59	87.10		

The overall accuracy = 87.02% Kappa = 0.84

5 Conclusions

Based on the principle of image fusion, from a practical perspective, this paper puts forward a fusion method of multiple modal spectral characteristics and radar remote sensing image. It implemented the image Landsat8 band selection, chose the best three bands, and reused of HSV transform the multispectral image and effective Sentinel-1 SAR radar image fusion, and trained the SVM model, extract the feature information of wetland. Experimental results show that the method based on band selection and multi-mode feature fusion can effectively improve the robustness of remote sensing image features. Compared with other methods, the fusion method can achieve higher classification accuracy and better classification effect.

References

1. Zhao, M., et al.: A robust delaunay triangulation matching for multispectral/multidate remote sensing image registration. *IEEE Geosci. Remote Sens. Lett.* **12**(4), 711–715 (2015)
2. Izadi, M., Saeedi, P.: Robust weighted graph transformation matching for rigid and nonrigid image registration. *IEEE Trans. Image Process.* **21**(10), 4369–4382 (2012). A Publication of the IEEE Signal Processing Society
3. Shahdoosti, H.R., Ghassemian, H.: Fusion of MS and PAN images preserving spectral quality. *IEEE Geosci. Remote Sens. Lett.* **12**(3), 611–615 (2014)
4. Akhavan-Niaki, H., et al.: Evaluation of spatial and spectral effectiveness of pixel-level fusion techniques. *IEEE Geosci. Remote Sens. Lett.* **10**(3), 432–436 (2013)
5. Chen, F., Wang, C., Zhang, H.: Remote sensing image classification based on an improved maximum-likelihood method: with SAR images as an example. *Remote Sens. Land Resour.* **28**(1), 75–78 (2008)
6. Liu, J., Zhang, C., Wan, S.: The classification method of multi-spectral remote sensing images based on self-adaptive minimum distance adjustment. In: Li, D., Chen, Y. (eds.) CCTA 2012. IAICT, vol. 393, pp. 430–437. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36137-1_50
7. Yu, X., Dong, H.: PTL-CFS based deep convolutional neural network model for remote sensing classification. *Computing* **100**(8), 773–785 (2018). <https://doi.org/10.1007/s00607-018-0609-6>

8. Wu, T., Chen, X., Xie, L.: An optimized K-means clustering algorithm based on BC-QPSO for remote sensing image. In: IGARSS 2017 - 2017 IEEE International Geoscience and Remote Sensing Symposium. IEEE (2017)
9. Yu, X., Dong, H., Patnaik, S.: Remote sensing image classification based on dynamic co-evolutionary parameter optimization of SVM. *J. Intell. Fuzzy Syst.* **35**(1), 343–351 (2018)
10. Ramakrishnan, N.K., Simon, P.: A bi-level IHS transform for fusing panchromatic and multispectral images. In: Proceedings of 5th International Conference on Pattern Recognition and Machine Intelligence, pp. 367–372 (2011)
11. Rokhmatuloh, R., Tateishi, R., Wikantika, K., et al.: Study on the spectral quality preservation derived from multisensor image fusion techniques between JERS-1 SAR and landsat TM data. In: Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS), pp. 3656–3658 (2003)
12. Dupas, C.A.: SAR and LANDSAT TM image fusion for land cover classification in the Brazilian atlantic forest domain. In: Proceedings of 19th International Congress for Photogrammetry and Remote Sensing, pp. 96–103 (2000)
13. Chen, C.M., Hepner, G.F., Forster, R.R.: Fusion of hyperspectral and radar data using the IHS transformation to enhance urban surface features. *ISPRS J. Photogram. Remote Sens.* **58** (1), 19–30 (2015)
14. Pal, S.K., Majumdar, T.J., Amit, K.: ERS-2 SAR and IRS-1C LISS III data fusion: a PCA approach to improve remote sensing based geological interpretation. *ISPRS J. Photogram. Remote Sens.* **61**(5), 281–297 (2007)
15. Chandrakanth, R., Saibaba, J., Varadan, G., et al.: Fusion of high resolution satellite SAR and optical images. In: International Workshop on Multi-Platform/Multi-Sensor Remote Sensing and Mapping, pp. 1–6 (2011)
16. Battsengel, V., Amarsaikhan, D., Bat-erdene, T., et al.: Advanced classification of lands at TM and Envisat images of Mongolia. *Adv. Remote Sens.* **2**(2), 102–110 (2013)
17. Yang, J.F., Ren, G.B., Ma, Y., et al.: Coastal wetland classification based on high resolution SAR and optical image fusion. In: International Geoscience and Remote Sensing Symposium, pp. 886–889 (2016)
18. Yu, L., Liu, H.: Feature selection for high-dimensional data: a fast correlation-based filter solution. In: Proceedings of the Twentieth International Conference Machine Learning (ICML 2003), Washington, DC, USA, 21–24 August 2003. AAAI Press (2003)
19. Zhu, Q., Liu, B.: Multispectral image fusion based on HSV and red-black wavelet transform. *Comput. Eng.* (2012)