Recognition of Autumn Crop Based on PolSAR Data and Feature Selection

Li T Q\textsuperscript{1,2}, Chen J\textsuperscript{*}, Liu J Y\textsuperscript{1,2}, Lian Z\textsuperscript{1} and Li J\textsuperscript{1}

\textsuperscript{1}China Aero Geophysical Survey & Remote Sensing Center for Natural Resources, Beijing 100083, P.R.C
\textsuperscript{2}Key Laboratory of Airborne Geophysics and Remote Sensing Geology, Ministry of Nature and Resources, Beijing, P.R.C

\textsuperscript{*}E-mail: chenj@mail.cgs.gov.cn; 6592296@qq.com

Abstract. Optical remote sensed images have been intensively used to map global and regional agriculture information. However, few optical images could be collected due to cloud contamination during the crop growth period. Therefore, synthetic aperture radar (SAR) images could be used to extract crop distribution since it is capable of acquiring data without regard to bad weather conditions. Although numerous studies have been successfully carried out to highlight the potential of SAR images in crop monitoring, there are still some problems to be solved. The high dimensionality of multi-temporal SAR images remains a major issue, classification using all features is limited to efficiency. In this study, a method of autumn crop recognition based on PolSAR data and feature selection was proposed. Using Radarsat-2 PolSAR images acquired during the autumn crop growing season, the optimal subset of features was obtained by 3 feature selection methods and 15 polarimetric target decomposition methods. With optimal feature subset and train samples, crops and other ground objects were classified by SVM. The classification result was assessed by test samples.

1. Introduction

The autumn crop is an important part of grain crops in China. It is essential to recognize the type of crop accurately, which is meaningful for the estimation of the crop area and yield. The growth cycle of autumn crop is mainly concentrated in the summer of the same period of rain and heat. Influenced by factors such as clouds and rain, it is difficult to obtain crop information in critical growth periods timely and accurately. SAR could be used to observe the growth of crops since it is capable of acquiring data without regard to bad weather conditions.

PolSAR data contains abundant information and classification features of ground objects, which has been widely used in the recognition of targets or the classification [1-3]. However, in the existing research, there is little research on the recognition of autumn crop, and there is no related research on the selection of crop growth periods or polarimetric target decomposition methods for recognition.

In this study, a method of autumn crop recognition based on PolSAR data and feature selection was proposed. Using Radarsat-2 PolSAR images acquired during the autumn crop growing season, the optimal subset of features was obtained by 3 feature selection methods and 15 polarimetric target decomposition methods. With optimal feature subset and train samples, crops and other ground objects were classified by SVM. The classification result was assessed by test samples.
2. Data

2.1. Study area
The study area is located in Shenzhou County in North China and belongs to temperate continental monsoon climate, as shown in Figure 1. Corns and cottons mainly growing from June to October are the main autumn crop in the study area.

![Figure 1. The location of the study area.](image)

2.2. Data and preprocessing
In this study, five Radarsat-2 Fine Quad-Pol images (Single Look Complex), acquired from June to October in 2014, were used for the autumn crop recognition, with a resolution of 5.2 $\times$ 7.6 m, as shown in Figure 2.

![Figure 2. Radarsat-2 Fine Quad-Pol images.](image)
The classes in the study area can be summarized into five categories, and 3547 samples were obtained by manual interpretation of high-resolution images, as shown in Table 1.

Data preprocessing mainly included converting the raw data into Sinclair matrix, filtering with J. S. Lee Refined Filter with the window size of 9[4], and geocoding. The final data format was coherence matrix T3.

Table 1. The number of samples for each class.

| Class            | Samples |
|------------------|---------|
| Corn             | 719     |
| Cotton           | 700     |
| Tree             | 706     |
| Artificial building | 712    |
| Water            | 710     |

3. Method

The technical flowchart of this research is shown in Figure 3. The polarization features were mainly obtained by polarimetric target decomposition. Three feature selection methods and SVM classifier were combined to construct the autumn crop recognition model, and the accuracy of the result was assessed.

![Figure 3. The technical flowchart of autumn crop recognition.](image)

3.1. Polarization features

The polarization scattering mechanism of autumn crop mainly includes three types: the surface scattering of soil, secondary scattering of soil and crop stalks, and volume scattering of crop canopy. The scattering characteristics of the crop were influenced by soil surface roughness, soil water content, crop above-ground biomass, and geometry. In this study, 15 polarimetric target decomposition methods were used: Huynen, Barnes1, Barnes2, Cloude, Holm1, Holm2, H/A/Alpha, Freeman 2
Components, Freeman 3 Components, Van Zyl, Yamaguchi 3 Components, Yamaguchi 4 Components, Neuman 2 Components, Krogager and Touzi. A total of 390 polarimetric components and other polarimetric parameters (SPAN, T3, and polarimetric correlation coefficient) were used for autumn crop recognition.

3.2. Optimal feature subset
In this study, distance metrics and information metrics were used to evaluate the importance of polarization features. Feature selection based on distance metrics was implemented by Relief-F method and feature selection based on information metrics was implemented by BIF and mRMR methods [5-7]. Mutual information was used as an evaluation index in BIF and mRMR. Compared with BIF, mRMR considers redundancy between the feature to be selected and the current feature subset.

In order to ensure efficiency and comparability among feature subsets, the maximum number of feature subsets was set to 8. Feature subsets were used to train the classifier and the optimal feature subset was selected according to the classification accuracy of the model.

3.3. Classification and accuracy assessment
In this study, two hundred samples in each class were randomly selected as test samples and the rest were used as training samples. SVM classifier based on radial basis function kernel was used for classification, and 10-fold cross-validation method was used in model training. The study area was classified using optimal feature subset and training samples. The accuracy of classification was assessed by test samples.

4. Result and discussion
The results of the feature selection and accuracy of corresponding models are presented in Table 2. The accuracy of the feature subset based on information metrics is higher. Due to the great difference in the scattering characteristics of the same ground objects, multiple cluster centers were formed in the same class when using distance metrics, which limits the use of Relief-F method. The result of mRMR is better than the result of BIF. As shown in Table 2, the features ranked 2nd to 5th in the result of BIF all dropped in the result of mRMR, and the 4th feature ranks 9th in the result of mRMR, which is not shown in Table 2. The features ranked 6th to 8th in the result of BIF are not included in the result of mRMR, because there is a strong similarity between TSVM components and the alpha component [8, 9].

| Importance | Relief-F Features | BIF Features | mRMR Features |
|------------|-------------------|-------------|---------------|
| 1          | 0721Neumann_psi   | 0603span_db | 0603span_db   |
| 2          | 1001Neumann_psi   | 1001span_db | 1001CCC       |
| 3          | 0814Neumann_psi   | 0814span_db | 0603alpha     |
| 4          | 0907Neumann_psi   | 0907span_db | 0603CCC       |
| 5          | 0721Neumann_delta_pha | 0721span_db | 1001span_db   |
| 6          | 0603CCC           | 0603TSVM_alpha_s1 | 0814CCC     |
| 7          | 0603TSVM_phi_s1   | 1001TSVM_alpha_s1 | 0814span_db |
| 8          | 0603span_db       | 0603TSVM_psi1  | 0721span_db   |

Model accuracy | 77.76% | 79.80% | 82.65%

Table 2. Feature subsets and corresponding accuracy.
Table 3. Classification accuracy assessment.

| Class              | PA    | UA    |
|--------------------|-------|-------|
| Corn               | 84.5% | 82.0% |
| Cotton             | 78.5% | 80.0% |
| Tree               | 75.0% | 69.0% |
| Artificial building| 91.0% | 90.0% |
| Water              | 86.5% | 96.0% |
| OA                 | 83.10%|
| Kappa              | 0.66  |

The optimal feature subset obtained by mRMR method was used for classification, and the result is shown in Figure 4 and Table 3. Corns and cottons were recognized accurately by the method proposed in this study, but the speckle noise in PolSAR data causes serious salt-and-pepper noise in the classification results.

Figure 4. Classification result.
5. Conclusion
In this study, a new method was presented for the recognition of autumn crop by using PolSAR data. The method using the optimal feature subset selected by mRMR from the polarization features for SVM classification. As the method is feasible and efficient, it is suggested for a wide application in the recognition of autumn crop in North China.

References
[1] Lee J S and Grunes M R 1994 *International Journal of Remote Sensing* 15 2299-2311
[2] Ferro-Famil L, Pottier E and Lee J S 2001 *IEEE Transactions on Geoscience & Remote Sensing* 39 2332-42
[3] Lee J S, Grunes M R, Ainsworth T L, Pottier E, Krogager E and Boerner W M 2000 *IGARSS* 20 1101-03
[4] Lee J S 1981 *Computer Graphics & Image Processing* 15 380-389
[5] Kononenko I 1994 *ECML* 7 171-182
[6] Jain A K, Duin R P W and Mao J 2000 *IEEE Computer Society* 22 4-37
[7] Peng H C, Long F H and Ding C 2005 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27 1226-38
[8] Touzi R 2007 *IEEE Transactions on Geoscience & Remote Sensing* 45 73-84
[9] Che M Q, Samat A, Du P J, Luo J Q and Bao R 2016 *Journal of Remote Sensing* 20 303-314