文章编号：1001－9014(2016)04－0398－09

DOI:10.11972/j.issn.1001－9014.2016.04.004

A framework for classification of urban areas using polarimetric SAR images integrating color features and statistical model

LIU Hong-Ying, WANG Shuang, WANG Rong-Fang, SHI Jun-Fei,
ZHANG Er-Lei, YANG Shu-Yuan, JIAO Li-Cheng

(Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University, Xi'an 710071, China)

Abstract: The color features were exploited in a novel framework for the unsupervised classification of urban areas in this paper. Firstly, based on the recent four-component decomposition model of the polarimetric synthetic aperture radar (PolSAR) data, the common color spaces, such as YUV, RGB, HSI, and CIELab were calculated. The color feature was quantitatively selected from these color spaces by introducing the color entropy. Then together with the texture feature and the extended scattering power entropy, the adaptive mean-shift algorithm was used to segment the PolSAR data into clusters. Finally, the clusters were merged according to the G0 distribution-based distance measurement. The proposed framework was verified by the experiments on one AIRSAR L-band and two Radarsat-2 C-band PolSAR data. The classification accuracy indicates that the proposed method has superior discriminative ability for urban areas compared with existing works.

Key words: polarimetric synthetic aperture radar (PolSAR), image classification, color features, extended scattering power entropy.

PACS: 84.40.Xb

集成颜色特征和统计特征的极化合成孔径雷达图像城区分类框架

刘红英, 王 爽, 王蓉芳, 石俊飞, 张二磊, 杨淑媛, 焦李成
(西安电子科技大学 智能感知与图像理解教育部重点实验室,陕西 西安 710071)

摘要:针对城区分类,利用颜色特征构造一个新颖的无监督的分类框架.首先,基于最近提出的极化合成孔径雷达(PolSAR)数据的四分量分解模型,计算了常用的颜色空间:YUV,RGB,HSI和CIELab,通过引入颜色熵量化的选择颜色特征;然后,联合纹理特征和扩展的散射功率熵,用自适应的均值漂移算法分割PolSAR图像;最后,根据基于G0分布的距离测度合并聚簇为较为匀质的地物类别.通过L波段AIRSAR数据和C波段Radarsat-2的PolSAR数据验证了提出算法的有效性,分类正确率表明,相比于已有的工作,提出的算法对于城区有较好的区分能力.

关 键 词:极化合成孔径雷达(PolSAR);地物分类;颜色特征;扩展的散射功率熵

中图分类号:TP751.1 **文献标识码**:A

Introduction

With the breakthroughs in synthetic aperture radar (SAR) imaging techniques[1], abundant data is observed and needs for interpretation. Terrain classification plays an important role in image interpretation. As the prior knowledge for classification of Polarimetric SAR

Received date: 2015-10-29, **revised date**: 2016-04-01 **收稿日期**:2015-10-29, **修回日期**:2016-04-01

Foundation items:Supported by National Basic Research Program of China (2013CB329402), National Natural Science Foundation of China (61473215, 61472306, 61502369), the Fund for Foreign Scholars in University Research and Teaching Programs (B07048), the Major Research Plan of the National Natural Science Foundation of China (91438201, 91438103), the Program for Cheung Kong Scholars and Innovative Research Team in University (IRT_15R53), the foundation from Ministry of Education of China (BK16015020001), the National Science Basic Plan in Shaanxi Province of China (2016JQ6049), and the Fundamental Research Funds for the Central Universities (7215598901)

Biography:LIU Hong-Ying(1983-), female, Ph. D. Research field involvs image processing, machine learning, etc. E-mail: hyliu@xidian.edu.cn

(PolSAR) image is limited, the unsupervised classification attracts much attention. Numerous methods for the unsupervised classification of PolSAR data have been proposed. In general, these methods fall into two categories. One category is based on target decomposition and the statistical model of PolSAR data, such as the H/a-wishart[3] classification proposed in 1999 based on the Cloude's decomposition, and Wishart distribution of the polarimetric covariance matrix proposed by Lee[4] in 2004 based on the Freeman et al's decomposition[5]. There are also some improved methods. In the work[6], both the parameters from Freeman decomposition and Cloude decomposition were utilized to classify. And in the work[8] the scattering power entropy calculated from Freeman decomposition was introduced to identify terrain types. Then Wang et al. [9] in 2013 applies the scattering entropy and Copolarized ratio after the Freeman et al's decomposition for a detailed classification. This method exploits the physical scattering mechanisms and the statistical characteristics of terrain types. The different terrain types could be discriminated and segmented. It obtains relatively high classification accuracy 89.20% for the crop land of Flevoland. However, there are also flaws. It has difficulty in classifying urban areas due to the following reasons. 1) The Wishart distribution of the polarimetric covariance matrix is accurate for homogenous areas, but it is not appropriate for characterizing heterogeneous areas which contains various terrains. This is indicated by the study in the work[10]. 2) The Freeman decomposition has difficulty in discriminating complicated or man-made targets and natural scatters owing to the assumption of the reflection symmetry. This is confirmed by the four-component decomposition models[11-12]. 3) The boundaries of the segmentation are not preserved well since other features of the PolSAR image are not concerned.

The other category is the methods based on image segmentation[7] which partitions an image into homogeneous regions corresponding to objects or parts of objects. Among them the clustering techniques[2,17] are widely applied. It is a process of grouping the data in terms of its feature description based on certain similarity measure and objective criterions. The similarity measure and the features significantly influence the results. 1) As for the similarity measure, the methods, such as the fuzzy clustering proposed by Du et al. [13], the Expectation-maximization clustering used by Kersten et al. [14], and the work[15], adopt the Wishart distance or revised Wishart distance, since it is well admitted that the polarimetric covariance matrix has a complex Wishart distribution. These methods have good classification performance for common terrain types but not the urban areas because the Wishart distribution is not fit well for such heterogeneous areas. 2) As for the features, the works[13-14] only utilize the covariance matrix for clustering. Although the covariance matrix is one of the most important characteristics and it holds the total polarimetric information of the pixel, it is the lowest order operator for extracting polarimetric parameters of the distributed scatters and it could not discriminate complicated terrains. In the clustering based on mean-shift algorithm[18], only the logarithm of the SPAN image which is the total scattering power of the polarimetric SAR data is employed as features. In the work[16], the coefficients from Freeman decomposition are chosen as features. These features help to characterize the ordinary terrains but not the urban areas, and the final classification result is limited by the similarity measure.

Therefore in this paper, we attempted to combine the advantages of both these two methods for better classification of the urban areas, namely, using the clustering techniques based on the intrinsic and discriminative features from PolSAR target decomposition and the sophisticated statistical model. Specifically, a framework for unsupervised classification of the terrain for PolSAR data was proposed.

1) The color features of the pseudo-color image from the new four-component decomposition are integrated into the unsupervised classification to enhance the details of the image especially in urban areas. The color entropy is introduced for quantitatively selecting the transformed color spaces for classification.

2) The scattering power entropy is extended to the four-component decomposition case, and is also employed as one of the multi-features, which could lead to better discrimination of the urban areas.

3) G0 distribution[21]-based merging scheme was developed. It is superior to that of the Wishart distribution for describing the heterogeneous terrains as that in urban areas. The neighborhood-based parameter estimation was also proposed for small clusters, which contributes to the similarity measure between clusters.

The remaining part of this paper is organized as follows. The motivations and related work are briefly introduced in section 1. The proposed framework is presented in detail in section 2. The experiments on three real PolSAR data with urban terrains from different systems are demonstrated in section 3. The conclusions are drawn in section 4.

1 Motivations

The proposed framework which is based on the scattering mechanism, statistical model and the clustering techniques, aims at not only differentiating the common terrains but also providing a fine classification for the urban areas. It includes two main stages: the initial clustering based on multi-features and the cluster merging according to the similarity measure from statistical model.

For the initial clustering, firstly the color features are concerned. As it is known that color image not only has a better visual display, but also has richful information of the details of the image compared with gray-scale image. The color feature is one of the primary features for discriminating object by human. In fact, this feature has been exploited in the supervised classification for PolSAR data in a recent work[19]. The main color descriptors were extracted from the Pauli decomposition coded pseudo-color image. The better classification result was obtained. Unlike this work, we extracted color features from the new four-component decomposition with extended volume (FDEV) model[12] coded pseudo-color image through a selection process. As mentioned above,

the four-component decomposition improves the Freeman decomposition. The vegetation and the oriented buildings could be well distinguished. The color intensities represent specific scattering mechanisms. Thus, the pseudo-color image from this decomposition is the intrinsic features containing discriminative information for the urban areas. Moreover, learnt from the chromatics, the components in RGB color space are not independent from each other and not optimal for color image processing. The color spaces[20] commonly used in computer vision: HSI (Hue-Saturation-Intensity), HSV (Hue-Saturation-Value), HSL (Hue-Saturation-Lightness), CIELab, CIE-Luv, YIQ and YUV are also taken into consideration. We introduce the color entropy to select the transformed space that has the most abundant details which is significant for classifying urban areas.

Then other features are also concerned for classifying urban areas: texture and the scattering power entropy. 1) Texture: Since the co-occurrence matrices (GLCM)-based texture can be used to analyze the contents of a variety of terrain types, it is studied in this work. Moreover, as it is indicated by the work[13], the total power SPAN representing the scattering intensity is capable of maintaining the fine structures of the PolSAR image. The texture features are calculated from SPAN. 2) The scattering power entropy: this feature has been verified to be effective in identifying terrain types by the work[6-7]. To better classify the urban areas, we derived the extended scattering power entropy (ESPE) from the improved four-component decomposition[11-12], which reveals the randomness of the scattering mechanisms.

Thirdly the clustering algorithm: mean-shift algorithm is applied. As we mentioned before, various clustering algorithms or the revised versions have been used in PolSAR segmentation and classification. Compared with other clustering algorithms, the predominant advantage of the mean-shift algorithm is that it does not constrain the distribution of the feature space and does not require the prior knowledge of the number of clusters. This fits well for our high-dimensional feature space of the PolSAR data whose feature vectors come from diverse decompositions and the mixed distribution is complicated.

For the cluster merging, the G0 distribution[21]-based similarity measure was developed. For the terrain classification that contains heterogeneous terrains like urban areas, the conventional Wishart distribution is not proper. A recent work[10] shows that the covariance matrix follows K-wishart, G0 or KummerU distributions. Due to the accurate description ability of the urban areas and the low computational complexity compared with KummerU distribution, G0 distribution is considered for the basis of the similarity measure.

2 The proposed method

The key idea of proposed method is to cluster the PolSAR data based on the multiple features derived from the polarimetric target decomposition, as shown by the flowchart in Fig. 1. The preprocessing with filtering, such as the Lee filtering, to reduce the speckle noise contained in PolSAR data is helpful for the feature extraction and the following classification.

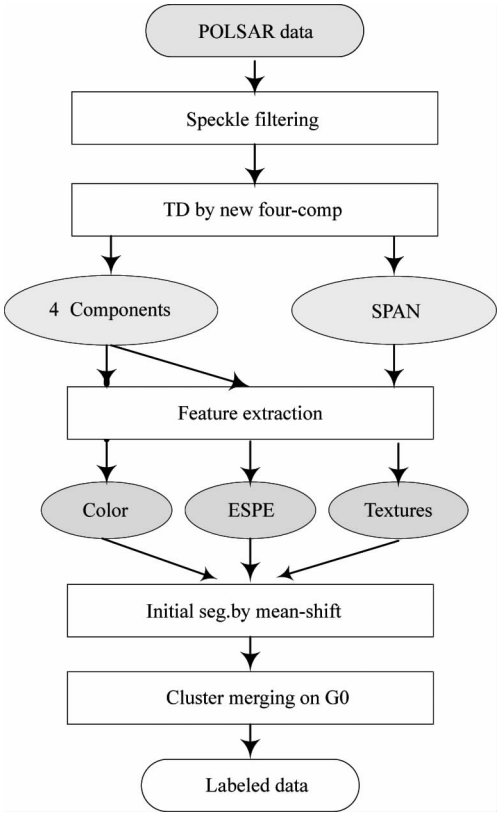

Fig. 1 The flowchart of the proposed framework (TD stands for target decomposition by new four-component model; ESPE denotes the extended scattering power entropy)
图1 提出方法的流程图(TD表示目标分解;ESPE表示扩展的散射功率熵)

2.1 Target decomposition

The FDEV model[12], which is the improved four-component decomposition, was applied to extract features for classification in this study, because it has a discriminative ability for urban areas compared with other decompositions. Four components P_s, P_v, P_d and P_h which represents the surface, volume, double-bounce and helix scattering power, respectively, are calculated from the covariance matrix of the PolSAR data. The total power TP or SPAN is expressed as

$$\mathrm{SPAN} = P_d + P_v + P_s + P_h \quad . \quad (1)$$

The FDEV improves the decomposition accuracy for distinguishing the urban areas by reducing the volume scattering power and enhancing the double-bounce scattering power within man-made structures. Thus it was applied to our study.

2.2 Feature extraction

For the unsupervised classification of PolSAR data, it is regarded as an image segmentation problem. Each pixel in the image is a vector $\boldsymbol{F}$ described by multiple features, expressed as

$$\boldsymbol{F}(F^C, F^T, F^O) \quad , \quad (2)$$

where F^C denotes the color features, F^T is the texture features, and F^O represents other features. The extraction of

the feature vectors is explained in the following.

2.2.1 Color feature selection

It is known that there are multiple color spaces for the digital image expression and processing. A proper color space is selected for the initial segmentation. Pseudo-color images are widely used for PolSAR data representation. For example, the PolSAR data after target decomposition, such as Pauli, Freeman, is mapped to the Red, Green, and Blue color channels in the RGB color space for display. Here, the three components P_s, P_d and from FDEV are mapped to the three color channels, respectively, shown as

$$\mathrm{RGB}(P_d,P_v,P_s)=f_{\mathrm{RGB}}(P_d,P_v,P_s) \quad . \quad (3)$$

In addition, it is noticed that the above mentioned color spaces can be divided into three groups due to their similarity: HSI(HSV, HSL), CIELab(CIELuv), and YUV(YIQ). The representative space in each group is used for the color space conversions, shown as

$$\mathrm{YUY}(P_d,P_v,P_s)=f_{\mathrm{RGB}\to\mathrm{YUY}}(P_d,P_v,P_s) \quad , \quad (4)$$

$$\mathrm{HSI}(P_d,P_v,P_s)=\mathrm{f}_{\mathrm{RGB}\to\mathrm{HSI}}(P_d,P_v,P_s) \quad , \quad (5)$$

$$\mathrm{CIELab}(P_d,P_v,P_s)=\mathrm{f}_{\mathrm{RGB}\to\mathrm{CIELab}}(P_d,P_v,P_s) \quad . \quad (6)$$

The colorfulness of a pseudo-color image both depends on the data and the color space. We defined an objective criteria: color entropy $\mathrm{Enp}_{\mathrm{spa}}$ to assess the colorfulness of the a pseudo-color image in certain space. $\mathrm{Enp}_{\mathrm{spa}}$ indicates the average colorfulness, and Enp is the information entropy of a single channel of the pseudo-color image. They are given by

$$\mathrm{Enp}_{\mathrm{spa}}=\sqrt{\frac{1}{M}\left(\sum_{i=1}^{3}\mathrm{Enp}_{\mathrm{spa}\_i}\right)} \quad , \quad (7)$$

$$\mathrm{Enp}_{\mathrm{spa}\_i}=-\sum_{j=0}^{255}p_j\log_2 p_j \quad , \quad (8)$$

where $M=3$, $i=1,2,3$. $j=0,1,\cdots,255$. P_j denotes the probability distribution of the j_{th} gray level in the single-channel spa_i. If there is only one gray-level, then the entropy of this channel is $\mathrm{Enp_spa}\_i=-\sum_{0}^{255}1\times\log_2 1=0$. So $\mathrm{Enp}_{\mathrm{spa}}$ measures the richness of a color image. For image in different color spaces, the values in the three channels are normalized and scaled to [0, 255]. The candidate color spaces are RGB, YUV, HSI, and as well as CIELab. spa is expressed as

$$\mathrm{spa}=\{\mathrm{RGB}(P_d,P_v,P_s),\mathrm{YUY}(P_d,P_v,P_s),\mathrm{HSI}(P_d,P_v,P_s),\mathrm{CIELab}(P_d,P_v,P_s)\} \quad . \quad (9)$$

Then the color space that has the maximum of color entropy from the calculated entropies $\mathrm{Enp}_{\mathrm{spa}}$ datasets is selected as the final color features for the classification. It is given by

$$F^{\mathrm{C}}=\max_{\mathrm{spa}}(\mathrm{Enp}_{\mathrm{spa}}) \quad . \quad (10)$$

In this selected color space, the colorfulness is the richest. The details, such as the boundary or the edges of the terrains might be well distinguished.

2.2.2 Texture feature

Texture shows the information about the spatial arrangement of the intensities in an image or selected region of an image. As the SPAN represents the scattering intensity, it is capable of maintaining the fine structures of the PolSAR image. In this work the GLCM is computed from the SPAN image within a sliding window due to the spatial dependence frequencies. It represents the spatial relationship of gray levels in it. For a SPAN image which is quantified into l gray levels, the GLCM is a matrix with size $l\times l$. It is the conditional joint probabilities of all pairwise combinations of the gray levels within interpixel distance du and orientation φ in a local spatial window w.

Four texture features which are frequently used in PolSAR processing are calculated from GLCM in this study. They are the energy, contrast, correlation, and local homogeneity (or inverse difference moment) which are denoted by ENG, CON, COR and LHM, respectively. Then the texture features are expressed as

$$F^{\mathrm{T}}=(\mathrm{ENG},\mathrm{CON},\mathrm{COR},\mathrm{LHM}) \quad . \quad (11)$$

2.2.3 Other features

The scattering power entropy was proposed in[8], which is based on the Freeman et al.'s decomposition, and including the double-bounce, surface, and volume scattering power. We extended it into four-component-decomposition based scattering entropy. The extended scattering power entropy (ESPE) H_Y is defined as

$$H_Y=-\sum_{i=1}^{4}p_i\log_4 p_i \quad 0\leqslant H_Y\leqslant 1 \quad , \quad (12)$$

$$p_1=\frac{p_d}{p_d+p_v+p_s+p_h} \quad p_2=\frac{p_v}{p_d+p_v+p_s+p_h}$$

$$p_3=\frac{p_s}{p_d+p_v+p_s+p_h} \quad p_4=\frac{p_h}{p_d+p_v+p_s+p_h} \quad , \quad (13)$$

where P_d, P_s, P_v and P_h have been defined previously and denote the double-bounce, surface, volume, and helix scattering power, respectively. The ESPE represents randomness of the terrains in the scattering mechanism and can be used for classification. This feature is expressed as the feature vector

$$F^{\mathrm{O}}=H_Y \quad . \quad (14)$$

2.3 Initial segmentation by adaptive mean-shift algorithm

The mean-shift algorithm is a nonparametric estimator of density gradient and extensively used in feature analyses, smoothing, segmentation, target tracking for image and video processing. Here we adopted it to the initial segmentation of the PolSAR pseudo-color image. For the PolSAR image, we adopted the adaptive mean-shift, namely the variable bandwidth mean-shift algorithm for PolSAR data segmentation. The feature vectors in the searching space is $F_i(F_i^{\mathrm{C}},F_i^{\mathrm{T}},F_i^{\mathrm{O}})\,i=1,2,\cdots,n$

The procedure is as follows.

1) Estimate the local variable bandwidth $h_c(F_i)$, $h_T(F_i)$ and $h_O(F_i)$ for the color feature, texture feature, and other features, respectively, using the "sample point estimator" method.

2) Randomly select an initial pixel from the unprocessed pixel set, its feature vector is $\boldsymbol{F}$.

3) Compute the vector $m(\boldsymbol{F})$, which is given by

$$m(F)=\frac{\sum_{i=1}^{n}\frac{F_i}{h_{\mathrm{C}}^3(F_i)h_{\mathrm{T}}^2(F_i)h_{\mathrm{O}}(F_i)}g\left(\left\|\frac{F^{\mathrm{C}}-F_i^{\mathrm{C}}}{h_{\mathrm{C}}(F_i)}\right\|^2\right)g\left(\left\|\frac{F^{\mathrm{T}}-F_i^{\mathrm{T}}}{h_{\mathrm{T}}(F_i)}\right\|^2\right)g\left(\left\|\frac{F^{\mathrm{O}}-F_i^{\mathrm{O}}}{h_{\mathrm{O}}(F_i)}\right\|^2\right)}{\sum_{i=1}^{n}\frac{1}{h_{\mathrm{C}}^3(F_i)h_{\mathrm{T}}^2(F_i)h_{\mathrm{O}}(F_i)}g\left(\left\|\frac{F^{\mathrm{C}}-F_i^{\mathrm{C}}}{h_{\mathrm{C}}(F_i)}\right\|^2\right)g\left(\left\|\frac{F^{\mathrm{T}}-F_i^{\mathrm{T}}}{h_{\mathrm{T}}(F_i)}\right\|^2\right)g\left(\left\|\frac{F^{\mathrm{O}}-F_i^{\mathrm{O}}}{h_{\mathrm{O}}(F_i)}\right\|^2\right)} \quad , \quad (15)$$

where $h_{\mathrm{C}}(F_i)$, $h_{\mathrm{T}}(F_i)$, $h_{\mathrm{O}}(F_i)$ denote the local variable

bandwidth for the color feature, texture feature, and other features, respectively.

4) If $\|m_h(\boldsymbol{F}) - \boldsymbol{F}\| < \varepsilon$, where ε is a very small constant to stop the iteration, then this iteration stops. Otherwise, go on.

5) Set the value of current feature vector $\boldsymbol{F}$ as $\boldsymbol{F} = m_h(\boldsymbol{F})$, and go to (1).

With the initial segmentation, a number of clusters which have similar features are obtained.

2.4 Cluster merging

The main idea for cluster merging is that clusters having similar parameters of the G0 distribution can be gathered in one group. Unlike the Wishart distribution, the G0 distribution has more parameters. Therefore, the parameters are estimated and then the merging criterion is developed in the following.

2.4.1 Parameter estimation based on neighborhood extension

Firstly, the texture parameter λ of a certain cluster need to be estimated. It is shown in Algorithm 1, and explained as follows.

For large clusters (the number of the pixel $N \geqslant Th$), the Method of matrix log-cumulants (MoMLC)[22] is adopted to estimate the parameter. Compared with the general moments-based estimation method as that used in[10], it simplifies the parameter estimation and improves the accuracy. This method is based on the Mellin transform which uses Mellin convolution to estimate the PDF of a product of random variables and is a natural analytical tool to deal with the distribution of products and quotients of independent random variables.

For small clusters (the number of the pixel $N < Th$), the neighborhood-based estimation is developed. The eight-neighbors of the each pixel in the cluster are also involved into the estimation of the texture parameter in order to extend the number of samples. This is reasonable when the pixels in the image are correlated in a local spatial scope.

2.4.2 Merging criterion

Similar to the merging criterion using Wishart distance, the G0-based distance criterion can be derived. The dispersion within clusters and the distance between clusters are calculated for merging.

From the initial segmentation map, the center V_i, namely the average coherency matrices of each cluster, is computed from pixels in each cluster as

$$\overline{V}_i = \frac{1}{N_i}\sum_{k=1}^{N_i} T_k \quad , \quad (16)$$

where i denotes the class i, and V_i is the class center of class I, $k = 1, 2, \cdots N_i$, N_i denotes the number of pixels in class i. T_k denotes the coherency matrix of pixels in class j.

The negative logarithmic distance measure d_{ML} calculated from PDF of G0 is given by

$$\begin{aligned} d_{ML} = & -Ld\ln L - (L-d)\ln|T| - \ln\Gamma(Ld+\lambda) \\ & -\lambda\ln(\lambda-1) + (Ld+\lambda)\ln[LTr(V^{-1}T) + \lambda - 1] \\ & + \ln\Gamma_d(L) + L\ln|V| + \ln\Gamma(\lambda) \quad . \quad (17) \end{aligned}$$

Algorithm 1 Texture parameter estimation using neighborhood extension

Input: Cluster Q with the number of pixels N, and label q,

Threshold parameter Th

Initialize:

Eight-direction vector $D = [-1,1;0,1;1,1;1,0;1,-1;0,-1;-1,-1;-1,0;-1,1]$,

Extended cluster $Q' = Q$,

if $N \geqslant Th$

$\overline{\lambda}$ = MoMLC ParaEstimation(cluster Q);

else

for each pixel $p(x,y)$ in cluster Q'

for k from 1 to 8

if GetLabel ($p(x,y) + D(k,:)) \neq q$

SetLabel ($p(x,y) + D(k,:)) = q$

$\overline{\lambda}$ = MoMLC ParaEstimation(cluster Q');

end

end

Output: estimated texture parameter $\overline{\lambda}$

where Tr is the trace of a matrix. V_i^{-1} denotes the inverse of matrix V_i. By removing the class-unrelated items, it is further simplified as

$$\begin{aligned} d(T,V,\lambda) = & L\ln|V| + \ln\Gamma(\lambda) - \ln\Gamma(Ld+\lambda) \\ & -\lambda\ln(\lambda-1) + (Ld+\lambda)\ln[LTr(V^{-1}T) \\ & + \lambda - 1] \quad . \quad (18) \end{aligned}$$

The dispersion within class i, which means the average distance from all the pixels in class i to the estimated class center $(\overline{V}_i, \overline{\lambda}_i)$, is defined as

$$S_{ii} = \frac{1}{N_i}\sum_{k=1}^{N_i} d(T_k; \overline{V}_i, \overline{\lambda}_i) \quad . \quad (19)$$

The distance between two classes i and j, which denotes the averaging sum of the distance from all the pixels in class i to the class center $(\overline{V}_j, \overline{\lambda}_j)$ and the distance from all the pixels in class j to the class center $(\overline{V}_i, \overline{\lambda}_i)$, is defined as

$$S_{ij} = \frac{1}{2}\left[\frac{1}{N_i}\sum_{k=1}^{N_i} d(T_k; \overline{V}_j, \overline{\lambda}_j) + \frac{1}{N_j}\sum_{k=1}^{N_j} d(T_k; \overline{V}_i, \overline{\lambda}_i)\right] \quad . \quad (20)$$

The merging index Ind_{ij} is given by

$$\mathrm{Ind}_{ij} = \frac{S_{ii} + S_{jj}}{S_{ij}} \quad . \quad (21)$$

The two classes i and j are merged if the merging index Ind_{ij} is the largest among the indexes from all the classes.

3 Experiments and discussions

Three real PolSAR datasets which are from different systems were used to verify the proposed method. They are the NASA/JPL AIRSAR L-band PolSAR data of San Francisco area from airborne system, the Canadian Space Agency RADARSAT-2 C-band of San Francisco area from spaceborne system, and also the PolSAR data of city XiAn in China from RADARSAT-2.

In order to show the differences between Wishart distance and G0 distance measure, two results are demonstrated. They are the results using : 1) color feature and Wishart distance measure; 2) color feature and G0 distance measure. They are briefly denoted as C-Wishart, and C-G0, respectively. Then the proposed framework is represented by C-T-ESPE-G0 in which the T stands for texture feature, and ESPE is the extended scattering power entropy.

Moreover, the proposed method was compared with two other types of related methods: The SPAN-based mean-shift algorithm[18] which is based on the clustering technique, and the Wang et al's[9], which is based on the scattering power entropy(SPE) from Freeman decomposition and the Wishart statistical model. They are termed as SPAN-MS and SPE-Wishart, respectively. In addition, all these data is preprocessed by Lee filter with 5×5 window to reduce the speckle noise.

3.1 Results on dataset 1: AIRSAR of San Francisco

The San Francisco area from AIRSAR L-band, which is 4-look fully polarimetric SAR data with a dimension of 900×700 pixels was used. The pseudo color image is shown in Fig. 2(a). The RGB color channels are the three components P_d, P_v, P_s which are from the new four-component decomposition.

As reported by other studies on this dataset, there are several categories of terrain, such as forestland, sea water, and urban areas. For more detailed classification, 8 main classes can be identified. The sea water may have three subclasses according to its density: the near-sea area, the medium-dense sea, and the far sea area. They have different colors due to the differences from the decomposition as shown in Fig. 2(a). The urban areas include four subclasses as the built-ups, streets, other large-sized man-made objects (e. g. golf course), and the mixture of vegetation (e. g. in parks).

The parameters for this experiment are explained. For the color feature selection, the color entropies for the RGB space and other converted spaces were computed and compared. The maximum one: CIELab (P_s, P_d, P_v) whose value is 2.70 is chosen as F^C. For the texture feature, the gray level $l = 32$, the distance $du = 1$, a sliding window of size 7×7 with 4 orientations was used to compute the texture vector for each component of F^T. This setting can capture the texture content of the SPAN. Then in order to reduce other noise, the average value for the 4 orientations was calculated as the descriptor for F^T, that is 4 dimensions of feature vector F^T is extracted for each pixel within the local spatial window. There is no parameter involved for the feature F^O. For the segmentation using adaptive mean-shift, the initial value for $h_C(F_i)$, $h_T(F_i)$ and $h_O(F_i)$ need to be assigned to start the iterative bandwidth estimation process. Generally, a large bandwidth brings coarse segmentation and a small bandwidth causes plenty of clusters. Since each color channel is constrained within $[0, 255]$, the $h_C(F_i)$ is set around 40 which brings more than a hundred clusters. The $h_T(F_i)$ and $h_O(F_i)$ are assigned within $[0.001, 0.01]$ after the normalization. For the cluster merging, the threshold Th for the neighborhood extension is set as 7×7. According to the MoMLC method, generally within this threshold, the estimation accuracy for the PolSAR image degrades. Thus it requires a local expanded estimation to enhance the accuracy. All the classifications are set 8 classes from the concerned methods here. For the method of SPAN-MS[18], the bandwidth of mean-shift algorithm is set around 10 to obtain 8 classes. For the SPE-Wishart[7], the iteration is set as 50, and the four threshold parameter are $x_1 = 0.5$, $x_2 = 0.9$, $x_3 = -3$, and $x_4 = 3$.

The classification results with different features and two other methods are shown in Fig. 2(b) ~ (f) respectively, and Fig. 2(d) is the classification results using the proposed framework. They indicate the effectiveness of the proposed framework on the L-band AIRSAR PolSAR data. In Fig. 2(b) and Fig. 2(c), since the color features are used, the details of the sea water are distinguished. However, the C-Wishart does not have a satisfactory class merging. The middle sea water is merged into the urban class. The near-sea region is merged into the sea water which is in contrast with C-G0. In Fig. 2(f) C-T-ESPE-G0, the sea water is classified into 4 or 5 classes, and the urban areas are classified into 3 subclasses at least by all these methods due to the application of the color features and G0-based merging scheme. Moreover, the golf course (in the middle) and the reservoir (near the bottom) can be correctly classified compared with (b) and (c), which is due to the combination of multiple features. In Fig. 2(e) although the golf course and reservoir are not kept in fragment, the urban areas are almost completely mis-classified as the class of vegetation, which may be caused by the application of the single SPAN-based feature. In Fig. 2(f) although the near-sea region is clearly identified, the variation in the sea water is not distinguished, which is likely owing to the Wishart-based merging criterion that only works for relatively homogeneous region.

3.2 Results on dataset 2: Radarsat-2 of San Francisco

The setting of the parameters is the same as that of dataset 1 since the number of colors and the size of the image is similar. The CIELab (P_s, P_d, P_v) also has the highest color entropy 2.68. Thus the color features are extracted from this color space.

The results with different feature combinations are shown in Fig. 3(b) ~ (d), respectively. Fig. 3(d) is the classification results using the proposed framework. They demonstrate that the proposed framework works well even for the 'rotated' image in spaceborn PolSAR system. Fig. 3(a) show that the colors in this image is quite abundant. In Fig. 3(b) and Fig. 3(c), based on the same initial segmentation by adaptive mean-shift algorithm using color features, the G0-based merging criterion keeps the diversity of the heterogeneous area. This is indicated by the results of the large part of the urban area. Meanwhile the results preserve the uniformity of the homogeneous region, which is represented by the sea water on the left. Fig. 3(d) indicates the effectiveness the combination of the color, texture, and ESPE features. The sea water (in blue) is classified into a more homogenous class than that in (e). In Fig. 3(d) the urban areas is merged into relatively homogeneous category, which is in contrast to (e) and (f). This may be resul-

ted from the coarse texture features.

3.3 Results on dataset 3: Radarsat-2 of XiAn

The PolSAR data of city XiAn is also adopted for classification. The original image is in dimension of 3 572 × 5 767 with coverage of 25 km × 25 km. The pixel size is 8 m in the slant range direction and 8 m in the azimuth direction. A selected region of western XiAn with 512 × 512 pixels was used. The setting of the parameters is similar to that of the dataset 2. The color entropies were computed. The CIELab (P_s, P_d, P_v) has the highest color entropy 2.72. Thus the color features are extracted from CIELab space.

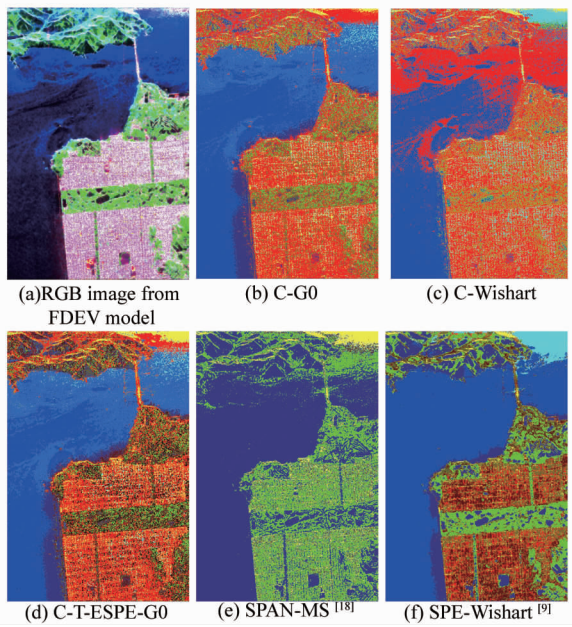

Fig. 2 Classification results with different features and methods for the L-band AIRSAR of San Francisco area. (a) RGB pseudo-colored image from FDEV model, (b) C-G0, (c) C-Wishart, (d) C-T-ESPE-G0, (e) SPAN-MS[18], (f) SPE-Wishart[9]

图2 在L波段AIRSAR系统San Francisco地区图上用不同特征和方法的分类结果 (a) 用FDEV模型的RGB伪彩图, (b) C-G0, (c) C-Wishart, (d) C-T-ESPE-G0, (e) SPAN-MS[18], (f) SPE-Wishart[9]

The RGB image is shown in Fig. 4(a). The urban areas (with high-dense built-ups) mainly sit on the up left part of the image. Several streets and roads nearby are identified. In the middle of the image, there is the River Weihe which is one primary branch of Yellow River from the west to the east of China, and the Weihe Bridge which spans the River Weihe. They appear dark due to the weak reflection. A three-line railway also spans the river. There is a sub-branch of the River Weihe in the right part of the image. The vegetation, both the low-dense wetland and the high-dense forestland (trees), are around the rivers.

The ground-truth map is shown in Fig. 4(c) and (d). Reference data was collected from river inventory maps, aerial photographic interpretation, Google satellite images and the fieldwork. The 6 classes are labeled and the legend is provided too. 7 classes are set for the final

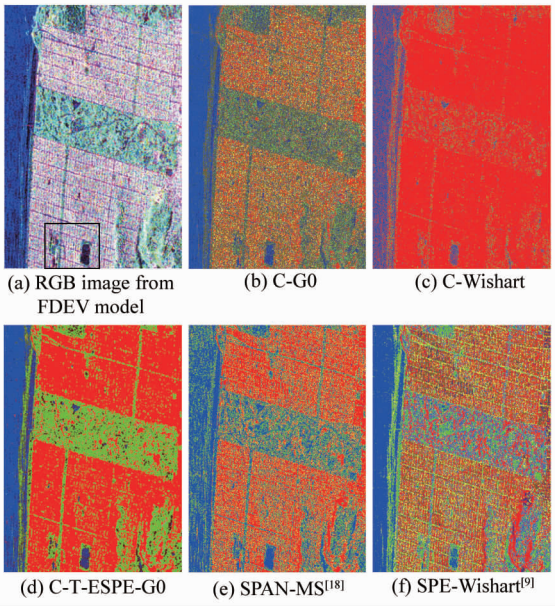

Fig. 3 Classification results with different features and methods for the sub region of San Francisco area from C-band Radarsat-2. (a) RGB pseudo-colored image from FDEV model, (b) C-G0, (c) C-Wishart, (d) C-T-ESPE-G0, (e) SPAN-MS[18], (f) SPE-Wishart[9]

图3 在C波段Radarsat-2系统San Francisco地区图上用不同特征和方法的分类结果 (a) 用FDEV模型的RGB伪彩图, (b) C-G0, (c) C-Wishart, (d) C-T-ESPE-G0, (e) SPAN-MS[18], (f) SPE-Wishart[9]

merging for all the concerned methods. This setting allows the unclear terrain type, such as the class in light blue shown in the classification results, which denotes other man-made targets.

Similar to the previous experiment, Fig. 4(e) and (f) indicate that the G0-based merging scheme preserves the diversity of the urban area. A combination of texture and ESPE features leads to the detailed classification as shown in Fig. 4(g). The texture feature may have a superior discriminative ability for the built-ups and other man-made targets in the urban area. Fig. 4(h) shows that the SPAN-based classification has poor discrimination in River Weihe and the surroundings due to the similar intensity of SPAN. In Fig. 4(i) although the bridge is better identified than in other figures, the three-line railway is mis-classified into vegetations. The rectangle on each figure highlights this result.

The overall classification accuracy of all the methods are listed in Table 1. It shows that the proposed framework (namely C-T-ESPE-G0) has the highest overall accuracy of 86.31%. Among the accuracy for various terrains, it has the highest accuracy 90.79% for the railway/built-ups. What's more, it can obtain 64.56% accuracy for the wetland/grass, which mainly results from the color and texture feature. The method on SPAN-MS[18] has 98.76% accuracy for the river because that it mis-classifies other types of terrains. It is noted that all these methods could not correctly classify the streets. One reason is that the reflection from street is weak and mixed in the complex urban areas. The other one is probably because that the color and texture features are

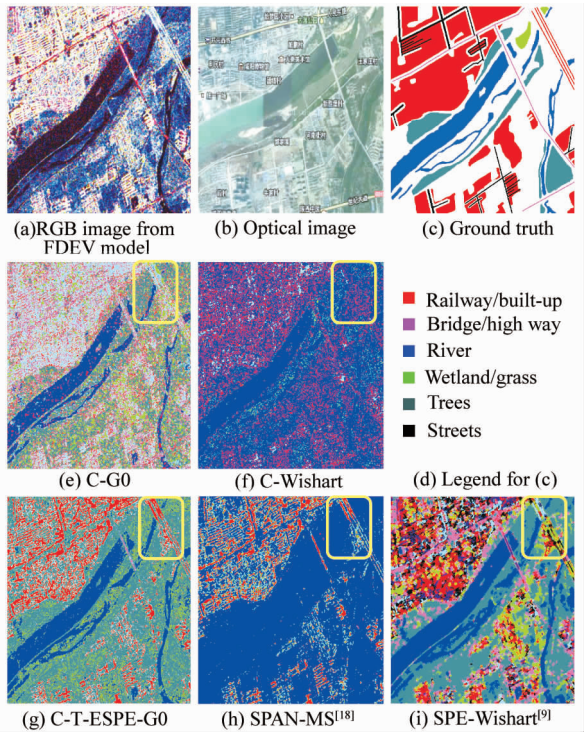

Fig. 4　Classification results with different features and methods for the western XiAn from C-band Radarsat-2. (a) RGB pseudo-colored image from FDEV model, (b) Optical image, (c) Ground truth, (d) The legend for the ground truth, (e) C-G0, (f) C-Wishart, (g) C-T-ESPE-G0, (h) SPAN-MS[18], (i) SPE-Wishart[9]

图 4　在 C 波段 Radarsat-2 系统西安西部地区图上用不同特征和算法的分类结果 (a) 用 FDEV 模型的 RGB 伪彩图, (b) 光学遥感图, (c)标记图, (d)图例, (e) C-G0, (f) C-Wishart, (g) C-T-ESPE-G0, (h) SPAN-MS[18], (i) SPE-Wishart[9]

similar to that of trees thus it is classified as trees in the proposed method.

Table 1　Overall classification accuracy (%) of westerm XiAn area form radarsat-2C-band PolSAR data

表 1　用极化 SAR 数据 C 波段 Radarsat-2 系统西安西部地区图的总体分类精度(%)

Methods \ Terrain	Railway/ built-up	Bridge/ highway	Wetland/ grass	River	Trees	Overall accuracy
C-G0	28.41	48.76	37.26	86.95	78.96	68.53
C-Wishart	41.25	0.00	19.93	96.14	0.00	53.69
C-T-ESPE-G0	90.79	25.66	64.56	93.62	77.32	86.31
SPAN-MS[18]	58.23	0.00	0.00	98.76	0.00	62.09
SPE-Wishart[9]	51.24	56.94	0.00	98.03	89.80	74.28

4　Conclusion

A novel framework for the unsupervised classification that integrates the color features was proposed based on the scattering mechanism and the clustering techniques of the PolSAR image. The color features are extracted from the four-component decomposition, which inherits the discriminative ability on the vegetations and the buildings in the urban areas. The color space conversions and selection allows further improvement of the color feature according to the introduced color entropy. Then the texture feature which is computed from the SPAN of the PolSAR data preserves fine structure of the image. Additionally, the extended power scattering entropy considers the randomness of the scattering mechanism. All these contribute to the discrimination of various targets in the urban areas. Moreover the G0 distribution-based merging criterion obtains refined classification results on Radarsat-2 C-band XiAn data since it characterizes the urban areas. The classification accuracy also indicates the superiority of the proposed framework to the existing methods especially on the built-up regions.

There are also flaws for the proposed framework. When the colors of the terrains are quite similar, such as dark colors for both the streets and the wetland, it is difficult in identifying them with the method and the classification accuracy degrades. For the future work, more features will be investigated to further distinguish the terrains in similar colors.

Acknowledgments

The authors would also like to thank the anonymous reviewers for their valuable comments, and also thank the providers of the PolSAR data sets. This work was supported by the National Basic Research Program of China (No. 2013CB329402), the National Natural Science Foundation of China (No. 61473215, No. 61472306, No. 61502369), the Fund for Foreign Scholars in University Research and Teaching Programs (No. B07048), the Major Research Plan of the National Natural Science Foundation of China (No. 91438201, No. 91438103), the Program for Cheung Kong Scholars and Innovative Research Team in University (IRT_15R53), the foundation from Ministry of Education of China (No. BK16015020001), the National Science Basic Research Plan in Shaanxi Province of China (2016JQ6049), and the Fundamental Research Funds for the Central Universities (No. 7215598901).

References

[1] Cong X, Gui G, Long K, *et al*. Quadratic compressed sensing based SAR imaging algorithm for phase noise mitigation [J]. *IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences*, 2016, **99**(6): 1233－1237.

[2] Wu J, Wang F, Xiang P. Automatic network clustering via density-constrained optimization with group operator[J]. *Applied Soft Computing*, 2016, **38**: 606－616.

[3] Lee J S, Grunes M R, Ainsworth T L. Unsupervised classification using polarimetric decomposition and the complex Wishart classifier[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1999. **37**(5): 2249－2258.

[4] Lee J S, Grunes M R, Pottier E, *et al*. Unsupervised terrain classification preserving polarimetric scattering characteristics [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2004. **42**(4): 722－731.

[5] Freeman A, Durden S L. A three-component scattering model for polarimetric SAR data[J]. *IEEE Trans. Geosci. Remote Sens.*, 1998, **36**(3): 963－973.

[6] Zhao L W, Zhou X G, Jiang Y M, *et al*. Iterative classification of po-

larimetric SAR image based on the freeman decomposition and scattering entropy[C]. In Proc. 1st Asian Pac. Conf. Synthetic Aperture Radar, Huangshan, China, Nov. 2007:473-476.

[7] Ma X L, Jiao L C. SAR image segmentation based on watershed and spectral clustering[J]. *J. Infrared Millim. Waves*, 2008,**27**(6):452-456.

[8] Yang W, Zou T Y, Sun H, *et al*. Improved unsupervised classification based on Freeman - Durden polarimetric decomposition[C]. In Proc. 7th Eur. Conf. Synthetic Aperture Radar, Friedrichshafen, Germany, Jun. 2008:1-4.

[9] Wang S, Liu K, Pei J, *et al*. Unsupervised classification of fully polarimetric SAR images based on scattering power entropy and copolarized ratio[J]. *IEEE Geoscience and Remote Sensing Letters*, 2013,**10**(3):622-626.

[10] Freitas C C, Frery A C, and Correia A H. The polarimetric G distribution for SAR data analysis[J]. *Environmetrics*, 2005, **16**(1):13-31.

[11] Yamaguchi Y, Moriyama T, Ishido M, *et al*, Fourcomponent scattering model for polarimetric SAR image decomposition[J]. *IEEE Trans. Geosci. Remote Sens.*, 2005,**43**(8):1699-1706.

[12] Sato A, Yamaguchi Y, Singh G, *et al*. Four-component scattering power decomposition with extended volume scattering model[J]. *IEEE Geoscience and Remote Sensing Letters*, 2012, **9**(2):166-170.

[13] Du L, Lee J S. Fuzzy classification of earth terrain covers using complex polarimetric SAR data[J]. *International Journal of Remote Sensing*, 1996, **17**(4):809-826.

[14] Kersten P R, Lee J S, Ainsworth T L. Unsupervised classification of polarimetric synthetic aperture radar images using fuzzy clustering and EM clustering[J]. *IEEE Transactions on Geoscience and Remote Sensing*,2005, **43**:519-27.

[15] Cao F, Hong W, Wu Y, *et al*. An unsupervised segmentation with an adaptive number of clusters using the SPAN/H/α/A space and the complex Wishart clustering for fully polarimetric SAR data analysis [J]. *IEEE Transactions on Geoscience Remote Sensing*, 2007, **45**(11):3454-3467.

[16] Tan L, Yan R. Unsupervised classification of PolSAR data using Freeman decomposition and fuzzy clustering, Synthetic Aperture Radar [C]. 2009. APSAR 2009. 2nd Asian-Pacific Conference on, On page(s): 489-493.

[17] Wu J, Zhang L, Li Y, *et al*. Partition signed social networks via clustering dynamics[J]. *Physica A*. 2016,**443**:568-582.

[18] He W, Jager M, Hellwich O. Comparison of three unsupervised segmentation algorithms for SAR data in urban areas[C]. In: Proceedings of the International Geoscience and Remote Sensing Symposium. Boston, USA, 2008:241-244.

[19] Uhlmann S, Kiranyaz S. Integrating color features in polarimetric SAR image classification[J]. *IEEE Trans. Geosci. Remote Sens.*, 2014, **52**(4):2197-2216.

[20] Castleman K R. *Digital image processing*[M]. Prentice Hall, Englewood Cliffs, NJ, USA, 1996(Sec. 21.3.4).

[21] Frery A C, Muller H J, Yanasse C D, *et al*. A model for extremely heterogeneous clutter[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1997,**35**(3),648-659.

[22] Anfinsen S N, Eltoft T. Application of the matrix-variate Mellin transform to analysis of polarimetric radar images[J]. *IEEE Trans. Geosci. Remote Sens.* 2011, **49**(6):2281-2295.