Segmentation of Images Using Wavelet Packet Based Feature Set and Clustering Algorithm

Ying Li[1], Yanning Zhang[1], Xiaoyue Jiang[1], Rongchun Zhao[1], and Mengdao Xing[2]

[1] School of Computer, Northwest Polytechnical University, Xi'an, 710072, China
{lybyp, ynzhang, xyjing, rchzhao}@163.com
[2] Key Lab for Radar Signal Processing, Xidian University, Xi'an, 710071, China
mdxing@mail.xidian.edu.cn

Abstract

The presence of speckle in Synthetic Aperture Radar (SAR) images makes the segmentation of such images difficult A novel method for automatic segmentation of SAR images is proposed. Firstly, a wavelet packet based texture feature set is derived. It consists of the energy of the feature subimages obtained by the overcomplete wavelet packet decomposition of local areas in SAR image, where the downsampling between wavelet levels is omitted. Then an improved unsupervised clustering algorithm is proposed for image segmentation, which can determine the number of classes automatically. Segmentation results on real SAR image demonstrate the effectiveness of the proposed method.

Keywords: Image segmentation, overcomplete wavelet packet decomposition, texture feature, clustering algorithm.

1 Introduction

Segmentation is a fundamental process in digital image processing which has found extensive application in areas such as content-based image retrieval, medical image processing, and remote sensing image processing. Its purpose is to extract labeled regions or boundaries for targeted objects for subsequent processing such as surface description and object recognition.

A segmentation procedure usually consists of two steps. The first step is to choose a proper set of features which can identify the same-content regions and meanwhile different-content regions; the second step is to apply a segmentation method to the chosen features to achieve a segmentation map.

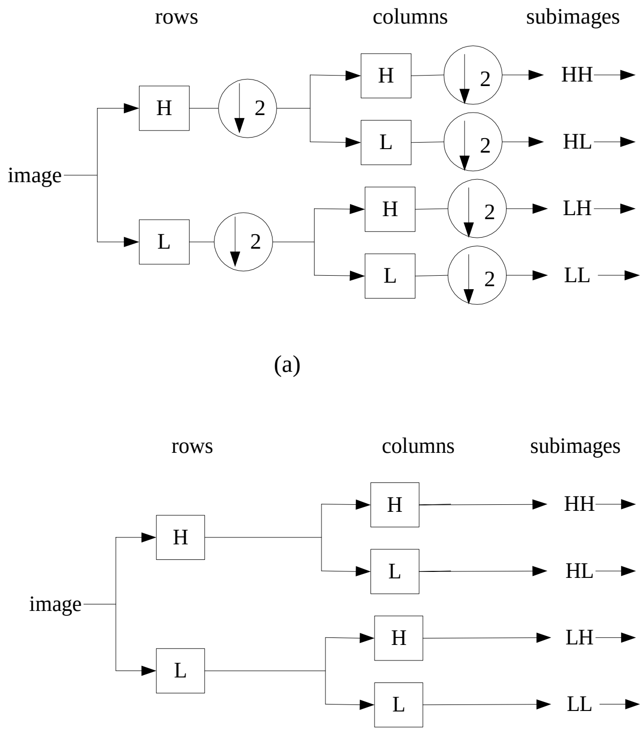
Texture is an essential key to the segmentation of SAR image. Texture originates from the scale-dependent spatial variability, and is observed as the fluctuation of gray levels in SAR images. It is well known, however, that another source of the fluctuation, speckle, exists in SAR images. Speckle is caused by the interaction of reflected waves from various independent scatters within a resolution cell. Here we consider texture separately from speckle, as Ulaby *et al.* did in [1].

The notion of scale gives an important hint to texture analysis. Recently, wavelet transform and wavelet packet transform have received much attention as a promising tool for texture analysis [2]-[6], because they have the ability to examine a signal at different scales. The texture feature set derived from the overcomplete wavelet decomposition where the downsampling between wavelet levels is omitted is superior to that derived from the standard (critically downsampled) wavelet decomposition [2], [3]. This is attributed to the property that the overcomplete wavelet decomposition can provide translation invariant features. The work by Chang and Kuo [5], however, indicates that texture features are most prevalent in intermediate frequency bands, thus that the octave band decomposition is not optimal. The trend probably therefore seems to be a concentration on the wavelet packet transform [4], [5], [6], which basically is the wavelet transform with subband decompositions not restricted to be dyadic. The discrete wavelet packet transform are critically sampled multi rate filter banks. However, critically sampled filter banks typically imply inaccurate texture edge localization [2]. Therefore, in this paper, overcomplete wavelet packet decomposition is employed to extract local texture features from SAR.

Image segmentation can be divided into two catalogs: supervised and unsupervised techniques. Supervised segmentation holds both advantages and drawbacks. One advantage is that one knows in detail how the algorithm works. One of the major drawbacks is that one needs to know beforehand how certain phenomena will appear in the images and this information is not always available. Even though the human interpreter often is superior in identifying strange details and phenomena in images, there is still a need to automate this process. This leads to the concept of unsupervised segmentation or classification of images. Clustering analysis is the main component of unsupervised techniques. Unfortunately, most clustering algorithms, for example the fuzzy C-means (FCM) method [8]-[10], suffer several difficulties: a) sensitive to the initialization; b) inability to find a global minimum and; c) difficulty of deciding how man clusters exist. In this paper, an improved unsupervised clustering algorithm is proposed for image segmentation, which can determine the number of classes automatically.

2 Wavelet Packet Based Feature Set

From the subband filtering (filter bank) point of view, the difference between the wavelet packet transform and the conventional wavelet transform is that the former also recursively decomposes the high-frequency components, thus constructing a tree-structured multiband extension of the wavelet transform. A two-dimensional (2-D) wavelet packet transform decomposes an image into four subimages. Fig.1(a) shows the standard decomposition. H and L in Fig. 1 denotes a highpass and lowpass filter respectively. $\downarrow 2$ denotes the downsampling by a factor of two. The approximated image LL is obtained by lowpass filtering in both row and column directions. The detail images, LH, HL, and HH, contain high frequency components. By decomposing the approximated image and the detail images of each level into four subimages iteratively, a complete image quad-tree is acquired.



(a)

(b)

Fig. 1. Wavelet packet decomposition of an image into four subimages, LL, LH, HL, and HH. (a) Standard decomposition. (b) Overcomplete decomposition. H and L denote a highpass and lowpass filter respectively. ↓2 denotes the downsampling by a factor of two. In the overcomplete decomposition, the downsampling is omitted.

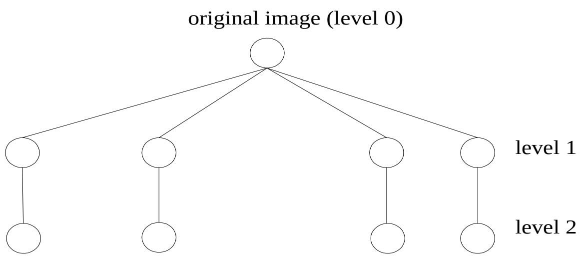


Fig. 2. The structure of the feature subimges

In the “standard” wavelet packet decomposition, the filtered outputs are downsampled by a factor of two, and are a quarter of the size of the previous image. There is neither loss nor redundancy of information between the levels, since the dyadic decomposition is an orthogonal representation of the image. The decomposition is therefore useful for applications such as image coding. The downsampling has, however, a drawback that the decomposition is not translation invariant. The translation-invariance property is indispensable for texture analysis. A solution of this problem is the overcomplete wavelet packet decomposition without the downsampling [see Fig. 1(b)]. The filtered outputs are the same size as the previous images. The overcomplete wavelet packet decomposition provides robust texture features at the expense of redundancy.

Feature extraction using overcomplete wavelet packet transform can extract all bandpass information about texture, but accompanying with the default that with the increasing of decomposition level, the number of subbands increasing exponentially, e.g. for 4 level complete tree structured wavelet transform, there are 341 leave nodes. This will generate high feature dimension, and cause great different in the following classification. Therefore, from the point of decreasing feature dimension and maintaining the stability of the feature dimension, we use a 2 level overcomplete wavelet packet transform. At level 1, four subimages are chosen to be the feature subimages, and at level 2, only the one with the maximum variance in each subchannel is chosen to be the feature subimages among all of the subimages. Thus, the number of the feature subimages is 8. Fig. 2 shows the structure of the eight feature subimages.

The energy (i.e., the averaged l_2-norm) of each feature subimages can be a favorable feature of texture because it indicates dominant spatial-frequency channels of the original image. Our texture feature set is also made up of the energy of feature subimages by the overcomplete wavelet packet decomposition. After an image of a square local area is decomposed, the following averaged energy is calculated, and is assigned to the component of the feature vector of the central pixel in the area.

$$e = \frac{1}{N \times N} \sum_{i=1}^{N} \sum_{j=1}^{N} [s(i, j)]^2 \tag{1}$$

where $s(i, j)$ denotes the wavelet coefficients of a feature subimage in the $N \times N$ window centered at pixel (i, j).

3 Segmentation Based on Clustering

Due to the drawback of FCM algorithm, here we adopt a clustering algorithm [7] by which the clusters are automatically generated and the data points are appropriately classified. It is worth emphasizing that the adopted clustering method does not need any prior assumption about the structure of the data. The clustering algorithm is described as follows.

Let $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ be a set of n vectors in a p-dimensional feature space, where $\boldsymbol{x}_i = \{x_i(1), x_i(2), \ldots, x_i(p)\}$ is a vector, then

1. Define n movable vectors $\boldsymbol{v}_i$ ($i=1,2,\ldots,n$) and let $\boldsymbol{v}_i=\boldsymbol{x}_i$, that is, $\boldsymbol{x}_i$ is the initial value of $\boldsymbol{v}_i$.
2. Calculate the relational grades between the reference vector $\boldsymbol{v}_i$ and the comparative vector $\boldsymbol{v}_j$ by

$$r_{ij}=\exp\left(-\frac{\left\|\boldsymbol{v}_i-\boldsymbol{v}_j\right\|^2}{2\sigma^2}\right),\quad i=1,2,\ldots,n,\quad j=1,2,\ldots,n \tag{2}$$

   where $\left\|\boldsymbol{v}_i-\boldsymbol{v}_j\right\|$ represents the Euclidean distance between $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$; and σ the width of the Gaussian function.
3. Modify the relational grades between the reference vector $\boldsymbol{v}_i$ and the comparative vector $\boldsymbol{v}_j$ by the following formula

$$r_{ij}=\begin{cases}0, & if\ r_{ij}<\xi\\ r_{ij}, & otherwise\end{cases} \tag{3}$$

   where ς is a small constant.
4. Calculate $\boldsymbol{v}'_i=\left\{v'_i(1),v'_i(2),\ldots,v'_i(p)\right\}$ by the following equations

$$\boldsymbol{v}'_i=\frac{\sum_{j=1}^{n}r_{ij}\boldsymbol{v}_j}{\sum_{i=1}^{n}r_{ij}},\quad i=1,2,\cdots,n \tag{4}$$

5. If all the vectors $\boldsymbol{v}'_i$ are the same as $\boldsymbol{v}_i$, $i=1,2,\ldots,n$, then go to Step 6; otherwise let $\boldsymbol{v}_i=\boldsymbol{v}'_i$, $i=1,2,\ldots,n$, and go to Step 2.
6. Based on the final results $\boldsymbol{v}_i$, $i=1,2,\ldots,n$, we can determine that the number of clusters is equal to the number of convergent vectors, the original data with the same convergent vector are grouped into the same cluster, and the convergent vector is the cluster center.

From the above clustering algorithm, we notice that a $n^2\times n^2$ relationship matrix is computed in each iterative step. When the number of input vectors is not too many, the memory and computing cost problem is not obvious. But when it is applied to image segmentation which needs to classify every pixel, the large memory space demanded always goes beyond the real physics memory which will cause the computation impossible. From Eq. (5), we can see that the computation of $\boldsymbol{w}_i$ is only related with the corresponding $\boldsymbol{r}_i$. Thus we can only compute the *i*-th row vector in the $n^2\times n^2$ relationship matrix other than the whole matrix in each iterative step. That is, only the *i*-th reference vector other than all reference vectors is computed from the above step2 to step5. The improved clustering algorithm will reduce the memory space demanded and computation cost greatly. An example is given to substantiate the effectiveness of the improved method. Fig.3 (a) shows a synthetic data set containing three clusters of various shapes and sizes. Fig.3 (b), (c), (d) show the obtained results by the original clustering method, the improved clustering method and the FCM algorithm respectively, where the “ * ” signs indicate the location of the resulting cluster centers. From Fig.3, we can see that the data set is classified correctly by both the original and the improved clustering method, but many data points are misclassified by the FCM method. Table 1 lists the running time of the original and the improved clustering algorithms for the different image size, where “／” sign denotes that the system runs out of memory which causes the computation impossible. So the original clustering method can only segment the images with few sizes. Besides, as the number of input vectors increases, the improved method takes less computing time than the original one.

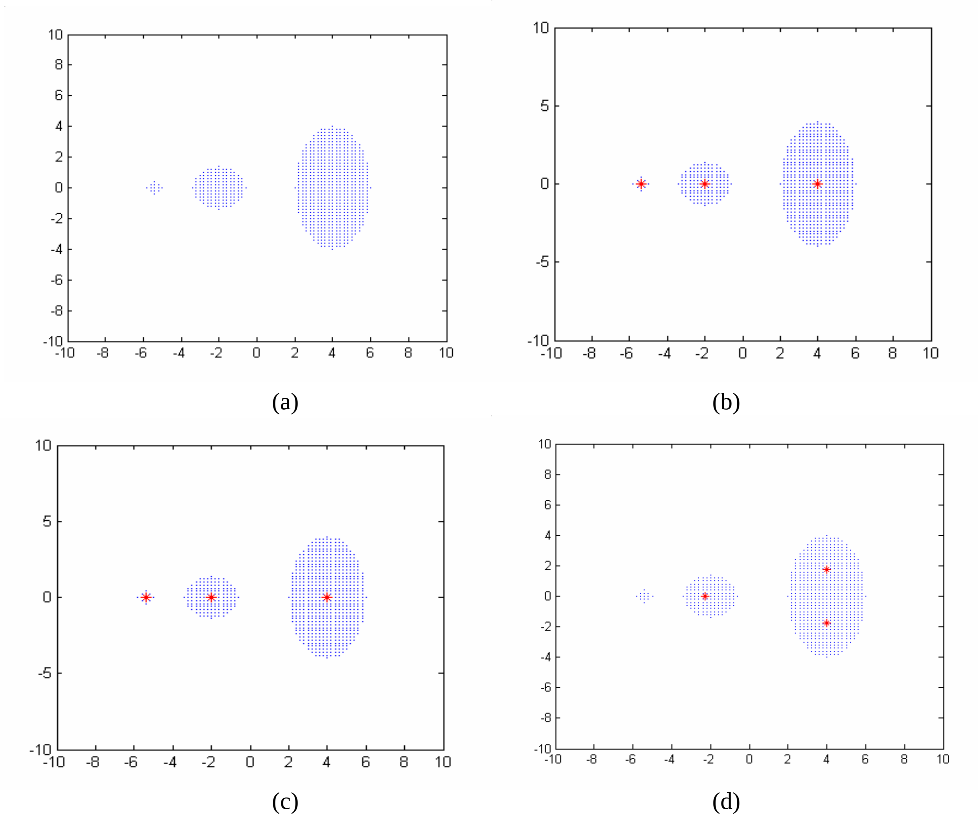


Fig. 3. (a) Synthetic data set (b) Result obtained by the original clustering algorithm (c) Result obtained by the improved clustering algorithm (d) Result obtained by the FCM algorithm

Table 1. Run-time comparison between the original and the improved clustering method

Size of images	Run-time (seconds)	
	Unimproved algorithm	improved algorithm
40×40	24.47	11.45
50×50	39.91	28.37
60×60	75.16	53.84
70×70	194.66	120.75
80×80	401.76	221.41
90×90	/	420.67

4 Experimental Results

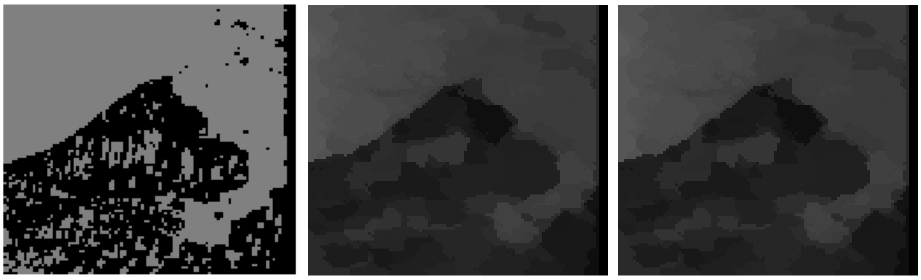
We apply the proposed method to SAR image segmentation and compare the results with other algorithms.

First, we use a 2 level overcomplete wavelet packet transform to obtain eight feature subimages as mentioned in Sec. 2, and compute the energy of each feature subimages. Here we use Daubechies D3 for image decomposition.

Then, the improved clustering algorithm is applied to the chosen features to achieve a segmentation map. Based on experience, the width (σ) of the Gauss function in (2) is set to 0.2, and the constant ξ in (3) is set to 0.01. Fig.4 (a) shows the original SAR image, and fig.4 (b) shows the segmentation result by the proposed algorithm. Fig.4 (c) ~ (f) show the results by the FCM [8], [9, [10], the OSTU [11], the split-and-merge method [12] and the MRF model [13] respectively. As Fig.4 illustrates, the result of the proposed method is better than that of the others.

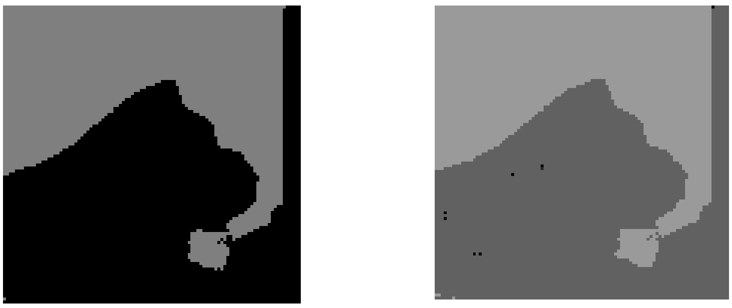
In many clustering algorithms, the width (σ) of the Gauss function affects the segmentation results greatly. From the proposed clustering algorithm, we notice that the specified value of σ affects the relationship of data set, and affects the number of clusters (as fig.5 illustrates). Besides, due to the difference of gray-level distribution of different image, the same σ may obtain different results. So the choosing of σ is very important, which need further research.

(a) (b) (c)

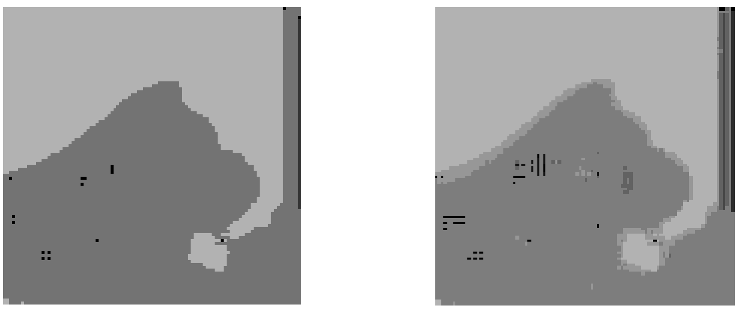


(d) (e) (f)

Fig.4. (a) Original SAR image (b) segmented result by the proposed algorithm (c) segmented result by the FCM method (d) segmented result by the OSTU method (e) segmented result by the split-and-merge method (f) segmented result by the MRF model



(a) (b)



(c) (d)

Fig.5. Affection of σ to the segmentation result, here c denotes number of classes. (a) $\sigma = 0.2$, c=2 (b) $\sigma = 0.14$, c=6 (c) $\sigma = 0.11$, c=10 (d) $\sigma = 0.1$, c=18.

5 Conclusion

In this paper, a wavelet packet based texture feature set has been introduced. The texture feature set is made up of the energy of feature subimages by the overcomplete wavelet packet decomposition of a local area in an image, where the downsampling between wavelet levels is omitted. The overcomplete wavelet packet decomposition provides translation-invariant features at the expense of redundancy. In addition, the proposed clustering algorithm can automatically determined the number of classes, which is much better than the FCM method. Segmentation results on the real SAR image demonstrate the effectiveness of the proposed method.

Acknowledgment

This works is supported by the National Natural Science Foundation of China (No. 60472072) and the Foundation of the National Key Lab for Radar Signal Processing, Xidian University, China (No. 51431020204HK0302).

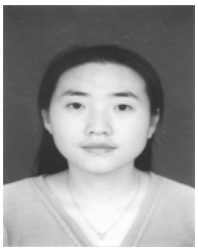
References

[1] Ulaby, F.T., Kouyate, F., Brisco, B., Williams L.: Textural Information in SAR Images. IEEE Trans. Geosci. Remote Sensing. Vol. 24 (3) (1986), pp. 235-245.
[2] Unser, M.: Texture Classification and Segmentation Using Wavelet Frames. IEEE Trans. Image Processing. Vol. 4 (11) (1995), pp. 1549-1560.
[3] Fukuda, S., Hirosawa, H.: A Wavelet-based Texture Feature Set Applied to Classification of Multifrequency Polarimetric SAR Images. IEEE Trans. Geosci. Remote Sensing. Vol. 37 (1999), pp. 2282-2286.
[4] Laine, A., Fan, J.: Texture Classification by Wavelet Packet Signatures. IEEE Trans. Pattern Anal. Mach. Intel. Vol. 15 (11) (1993), pp. 1186-1190.
[5] Chang T., Kuo, C.-C.J.: Texture Analysis and Classification with Tree-Structured Wavelet Transform. IEEE Trans. Image Processing. Vol. 2 (4) (1993), pp. 429-441.
[6] Saito, N., Coifman, R.R.: Local Discriminant Bases and Their Applications. J. Mathematical Imaging and Vision. Vol. 5 (4) (1995), pp. 337-358.
[7] Wong, C.C., Chen, C.C.: A Hybrid Clustering and Gradient Descent Approach for Fuzzy Modeling. IEEE Trans. on SMC-Part B. Vol. 29 (6) (1999), pp. 686-693.
[8] Canno, R.L., Dave, J.V., Bezdek, J.C.: Efficient Implementation of the Fuzzy C-Means Clustering Algorithm. IEEE Trans. Pattern Anal. Mach. Intel. Vol. 8 (1986), pp. 248-255.
[9] Bezdek, J.C., Ehrlich, R., Full,W.: FCM: The Fuzzy C-Means Clustering Algorithm. Computers and Geosciences. Vol. 10 (1984), pp. 191-203.
[10] Kamel, M.S., Selim,S.Z.: New Algorithms for Solving the Fuzzy Clustering Problem. Pattern Recognition. Vol. 27 (1994), pp. 421-428.
[11] Otsu, N.: A Threshold Selection Method from Gray Level Histograms. IEEE Trans. Syst., Man, Cybern. Vol. 9 (1979), pp. 62-66.
[12] Wu, X.: Adaptive Split-and Merge Segmentation Based on Piecewise Least Square Approximation. IEEE Trans, Patt. Anal. Machine Intel. Vol. 25 (8) (1993), pp. 808-815.

[13] Panjwani, D.K., Healey, G.: Markov Random Field Models for Unsupervised Segmentation of Textured Color Images. IEEE Trans. Pattern Anal. Mach. Intel. Vol. 17 (10) (1995), pp. 939-954.

Ying Li is an associate professor of School of Computer, Northwest Polytechnical University, Xi'an. Her interests include fuzzy logic, neural network, computation intelligence, and signal processing. She is the author or co-author of more than 40 scientific papers.

Yanning Zhang is a professor of School of Computer, Northwestern Polytechnical University. Her research interests include pattern recognition and computer vision, virtual realization and science visualization.



Xiaoyue Jiang is the PH.D candidate in Northwestern Polytechnical University. Her research interests include sequences image processing, pattern recognition and computer vision.

Rongchun Zhao is a professor of School of Computer, Northwestern Polytechnical University. His research interests include speech, image signal processing and understanding.

Mengdao Xing is a professor of Key Lab for Radar Signal Processing, Xidian University. His research interests include radar signal and image processing.