Novel Image Segmentation Methods Based on Improved Ncut Algorithm

Dao-cai FU¹, Huai-xin CHEN¹,*, Yun-zhi YANG², Zhou HUANG¹, Chang-yin WANG² and Bang-mei Li¹

¹School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China
²CETC Special Mission Aircraft System Engineering Co. Ltd, Chengdu 611730, China
*Corresponding author

Keywords: Image segmentation, MeanShfit algorithm, Graph theory Ncut algorithm, Entropy rate superpixel.

Abstract. Image segmentation is an important process from image processing to image analysis and scene understanding. Targets of different scales in an image may belong to different classes. Single scale image segmentation methods are prone to over-segmentation or under-segmentation. Combining MeanShfit and Ncut multiscale sampling image segmentation methods, the image is smoothed by MeanShift algorithm and the edges are preserved. Clustering using weights (similarities) between two pixels by Ncut method, this method can be applied to remote sensing image segmentation. In order to solve the problems of high time complexity and long calculation time of Ncut algorithm, firstly, the entropy rate superpixel segmentation algorithm is used to over-segment the image. The algorithm divides the image into a series of compact regions with good regional consistency, Combining Ncut method to solve the weights of similarity between two superpixels to determine the attribution of the region can reduce the segmentation time. Experiments show that this method can use multiscale information for image segmentation effectively. The first method can be applied to remote sensing image segmentation, and the second method can improve the efficiency of Ncut which is suitable for natural scene image segmentation.

Introduction

Image segmentation can understand the scene and extract the target or region of interest. The traditional image segmentation methods are mainly threshold segmentation and clustering merging. Typical threshold segmentation methods include bimodal method, iterative method and OSTU method, and representative clustering and merging methods include MeanShift algorithm, k-means algorithm and DBSCAN algorithm. Usually the scale of different objects in images is often different, which is more obvious in remote sensing images. In remote sensing images, the scales of tall buildings, highways and automobiles are more obvious. Single scale segmentation often results in over-segmentation or under-segmentation, which makes it difficult to extract regions of interests from remote sensing images.

The Ncut algorithm based on graph theory has good segmentation effect[1], and has been applied to remote sensing image segmentation, medical image segmentation, vector field segmentation and other image segmentation research. In graph theory, if there is a line between two vertices, that is, there are edges, each edge with weight. Pixel points in the image are regarded as vertices in the graph, and the weight between the pixels can reflect the similarity between the two pixels. For example, in a gray scale image, if the difference between the two points is large, their similarity is relatively small, that is, the edge weight connecting them will be relatively small, it can be judged that they do not belong to the same object, then the two points can be separated. Graph and image have great similarity, or image belongs to a branch of graph. It can be divided into disjoint sets from vertex set of graph to disjoint regions. Ncut algorithm for image segmentation can minimize the correlation between classes and maximize the correlation within classes[2].
Because Ncut algorithm may have errors in computing similarity matrix and clustering process, and image segmentation results will be over-segmentation and under-segmentation. In order to reduce the over-segmentation area in the multiscale segmentation results, firstly, MeanShift algorithm [3] used to preprocess the image. MeanShift algorithm is used for every pixel in the image, which will not change much for normal pixels. For a small number of noise points, the difference between the value of the pixel and that of the surrounding pixels is large. The algorithm will move the pixel to the mean of the eigenvalues of all the pixels in its neighborhood, and repeat until convergence. Therefore, after each pixel is iterated, the image is smoothed and the noise points are suppressed. In the subsequent Ncut algorithm for image segmentation, the influence of noise points on the clustering iteration process is reduced, and the edge of the target can be well preserved.

In recent years, superpixel has been widely used in the field of computer vision. It can aggregate single pixels into regions, and reduce the time complexity of subsequent processing. At present, the representative superpixel algorithms include Turbopixel, SLIC[4], SEEDS, Entropy Rate Superpixel(ERS) Segmentation[5] and so on. Because of the compact segmentation results and good regional consistency of the Entropy Rate Superpixel algorithm, the principle of the algorithm is not very complex, so this paper chooses the Entropy Rate Superpixel Pre-segmentation method. Firstly, the image is over-segmented by the entropy rate superpixel algorithm, and the superpixel regions are clustered and merged. After clustering, the region consistency is better, and the edge of the target can be well preserved.

Algorithm Principle

MeanShift Algorithm

MeanShift algorithm is an iteration algorithm based on fast statistics without parameters of nuclear density gradient estimation. At the beginning of the iteration, it starts with a pixel, and then iterates along the direction of the density gradient rising. Finally, it converges at the place where the density is the highest. MeanShift algorithm is mainly used in image segmentation and target tracking.

Entropy Rate Superpixel Algorithm

In image segmentation theory based on graph theory, the image is mapped to undirected graph G= (V, E), where V represents the vertex set of graph, E represents the edge set, and the weight of edge represents the similarity between vertices, which is expressed by weight function w. Thus, the problem of image segmentation is transformed into the problem of graph partition. The weights of undirected graph edges are symmetrical $w_{ij} = w_{ji}$, Define the random walk process of graph G is $X = \{X_t | t \in T, X_v \in V\}$, the weight W is not negative. The transition probability of random walk model is defined as:

$$
P_{ij} = \Pr(X_{t+1} = v_j | X_t = v_i) = \frac{w_{ij}}{w_i}$$

(1)

In the formula (1), i and j denote vertices of graphs, and $e_{ij}$ denotes edges connecting vertices I and j. $w_i = \sum_{k \in E, k \neq i} W_{ik}$ denotes the sum of all edge weights that are edgewise connected to node I. The normalized results are as follows:

$$
\mu = (\mu_1, \mu_2, \cdots, \mu_{|V|})^T = \left(\frac{W_1}{W_T}, \frac{W_2}{W_T}, \cdots, \frac{W_{|V|}}{W_T}\right)^T
$$

(2)
\( w_e = \sum_{i=1}^{|V|} w_i \) is a normalized constant. The goal is to select a subset \( A \subseteq E \) of edges so that the constructed result graph \( G = (V, A) \) contains \( k \) adjacent nodes.

(1) Entropy rate: If a particle moves randomly from one vertex \( i \) of undirected graph to another vertex \( j \), the transfer probability mapping function \( p_{ij} : 2^E \rightarrow R \) is expressed as follows. Entropy rate is used as a criterion to obtain compact and regional consistency.

\[
p_{ij}(A) = \Pr(X_{i+1} = j | X_i = i) = \begin{cases} \frac{w_e}{w_i} & \text{if } i \neq j \text{ and } e_{ij} \in A \\ 0 & \text{if } i \neq j \text{ and } e_{ij} \notin A \\ 1 - \frac{\sum_{e \in A} w_e}{w_i} & \text{if } i = j \end{cases}
\]

Therefore, the entropy rate of random walks on graph \( G = (V, A) \) as follows

\[
h(A) = -\sum_i \mu_i \sum_j p_{ij}(A) \log(p_{ij}(A))
\]  \hspace{1cm} (3)

(2) Equilibrium function: In order to make the clusters have similar size, the equilibrium function is used as the decision criterion. Let \( A \) set be the selected edge set and \( N_A \) is the number of connected subgraphs in the graph, if the result of graph segmentation is \( S = \{S_1, S_2, \ldots, S_{N_A}\} \). The distribution \( N_A \) formula of cluster members is

\[
P_{z_i} = \frac{|S_i|}{|V|}, i = \{1, 2, \ldots, N_A\}
\]  \hspace{1cm} (5)

The equilibrium function is

\[
B(A) \equiv h(Z_A) - N_A = -\sum_i P_{z_i}(i) \log(P_{z_i}(i)) - N_A
\]  \hspace{1cm} (6)

Entropy \( h(Z_A) \) can make the generated cluster have similar size, and \( N_A \) can control the size of the cluster.

(3) Object function: It combines the entropy rate and equilibrium function to form a compact clustering result with good regional consistency. Clustering is obtained by optimizing the objective function of formula (7) on the edge set.

\[
\max_A h(A) + \lambda B(A) \quad \text{subject to } A \subseteq E \text{ and } N_A \geq K
\]  \hspace{1cm} (7)

In the formula (7), \( \lambda = \beta k \lambda \) is the number entered by the user, generally \( \lambda = 0.5 \). \( \beta \) is a value calculated automatically according to the following formula, that is, according to the gain of the entropy rate and the gain of the equilibrium term, the degree of influence on the objective function is adjusted automatically.

\[
\beta = \frac{\max_{e_{ij}} h(e_{ij}) - h(\emptyset)}{\max_{e_{ij}} B(e_{ij}) - B(\emptyset)}
\]  \hspace{1cm} (8)
Normalized Cut Algorithm

Normalized Cut Principle

Image is an undirected graph $G=(V,E)$, and its edge with weight. Delete some edges which total weights smaller in figure $G$, then divide the graph into $A$ and $B$ sets, and $A \cup B = V$, $A \cap B = \emptyset$, the two sets have low similarity. In formula (9), $w_{ij}$ denotes the edge weight between the point $i$ and the point $j$, also denote the similarity between two points. The sum of all edge weights is called cut:

$$Cut(A, B) = \sum_{i \in A, j \in B} w_{ij}$$ \hspace{1cm} (9)

Optimal segmentation in an image needs to minimize the edge weight of cut sets $A$ and $B$. If cut into several sets, it may produce many outliers. In order to avoid this phenomenon, Shi and Malik proposed the Normalized cut method, which can solve this problem effectively.

The normalized cut formula as below:

$$NCut(A, B) = \frac{Cut(A, B)}{assoc(A, V)} + \frac{Cut(A, B)}{assoc(B, V)}$$

$$assoc(A, V) = \sum_{i \in A, j \in V} w_{ij}$$

denotes the total degree of association between points in $A$ and points in the graph. $assoc(B, V)$ is similar with $assoc(A, V)$. $NCut(A, B)$ represents the similarity between two classes, the smaller the value, the better the cut result.

Solution of Ncut

The degree matrix represents sum of the points connected to the vertices’ weights, and is the sum of similarity matrix columns.

$$d_i = D(i, i) = \sum_j w_{ij}$$ \hspace{1cm} (11)

Compute minimum value of $NCut(A, B)$ is equaled to solve the following characteristic equation (12).

$$(D - W)y = \lambda Dy$$ \hspace{1cm} (12)

The problem of solving the normalized cut is transformed into the K smallest eigenvector corresponding to the K smallest eigenvalue of formula (12). The result of graph cut or image segmentation can be obtained by clustering the eigenvectors.

Algorithms Improvement and Results

Computational similarity matrix $W$ is a key part of Ncut algorithm. It is the decisive factor of clustering segmentation algorithm. In the method-MeanShift combined with Ncut algorithm, the method of calculating $W$ has not changed, so the calculation time and the result of remote sensing image segmentation have not been improved before the improvement. In method 2, calculating $W$ between superpixels by calculating the centroid of superpixels can reduce the computing time, and has better segmentation results in natural scene images, but for remote sensing images, the segmentation results have not been improved.

MeanShift and Ncut Image Multiscale Segmentation

Steps of This Method

Principle flow chart of image segmentation using MeanShift and Ncut algorithm as Fig.1 below, the image is preprocessed by MeanShift algorithm, and the similarity matrix $Ws$ of multiple scales is calculated. The similarity matrix $W$ of the image is obtained by adding the similarity matrices of each scale. Finally, the image is segmented by Ncut algorithm.
The multiscale method is used to construct the correlation matrix. Compared with the traditional method, it needs to calculate the similarity between each point, which not only reduces the amount of calculation, but also can carry out multiscale analysis, and the segmentation effect is better.

**Similarity Matrix of This Method**

The $w_{ij}$ method is defined as Formula (13), which consists of two parts $w_I(i, j)$ and $w_C(i, j)$, $\alpha$ as constants.

$$w_{ij} = \left\{ \begin{array}{ll} w_I(i, j) \times w_C(i, j) + \alpha w_I(i, j), & \text{if } |X_i - X_j| < r \\ 0, & \text{otherwise} \end{array} \right.$$  \hspace{1cm} (13)

In formula (13), $w_I(i, j)$ denotes intensity, as shown in formula (14).

$$w_I(i, j) = e^{-\frac{|x_i - x_j|}{\sigma_x}} \frac{|l_i - l_j|}{\sigma_l}$$  \hspace{1cm} (14)

$\sigma_x$, representing pixel brightness information variance.

$\sigma_l$, representing pixel coordinate information variance.

$x_i, x_j$ represent pixel coordinates, and $|x_i - x_j|$ represents the Euclidean distance of two pixels.

$l_i, l_j$ represents pixel brightness, and $|l_i - l_j|$ represents the difference between two pixels.

In formula (13), $w_C(i, j)$ the interference contour, i.e. the weight of the edge, is represented as formula (15).

$$w_C(i, j) = e^{-\frac{\max_{\text{slice}(i, j)} \|\text{Edge}(x)\|}{\sigma_c}}$$  \hspace{1cm} (15)

$\sigma_c$ is a constant term. \(\max_{\text{slice}(i, j)} \|\text{Edge}(x)\|\) calculates the point with the strongest edge intensity on the line between I and J pixels, that is, the magnitude of the edge gradient. For example, on the i, j connection edge, there is a point a whose coordinates $(m, n)$ have the largest gradient magnitude in the X and Y directions, such as formula (16).

$$\max_{\text{slice}(i, j)} \|\text{Edge}(x)\|^2 = \sqrt{\nabla_x f(m, n)^2 + \nabla_y f(m, n)^2}$$  \hspace{1cm} (16)

**Entropy Rate Superpixel and Ncut Image Segmentation Method**

**Steps of This Method**

Principle flow chart of image segmentation using ERS and Ncut algorithm as Fig.2 and Fig.3.
Figure 1. MeanShift and Ncut algorithm’s flow chart.

Figure 2. ERS and Ncut algorithms’s flow chart.

Figure 3. Ncut’s flow chart.

Firstly, the image is segmented by the entropy rate superpixel algorithm. The centroid of the over-segmented region represents the whole superpixel, and the centroid of the two superpixels calculates the similarity. Finally, the image is segmented by Ncut algorithm.

**Algorithmic efficiency:** as below V is the number of vertices in an image, and K(K<<V) is the number of classifications.

**Entropy rate superpixel complexity:** $O(V \log V)$

**Time complexity of Ncut algorithm:** $O(V^2)$; After ERS: $O(K^2)$

**Total time complexity:** $O(V \log V) + O(K^2) = O(V \log V) < O(V^2)$

So the time complexity is reduced.

### 3.2.2 Similarity matrix of this method

Similar to Section 3.1.2, here we add LUV color space information. So define $W_{ij}$ as follows:

$$W_{ij} = e^{\frac{||x_i - x_j||^2}{\sigma_x} + \frac{||c_i - c_j||^2}{\sigma_c}}$$  \(17\)

$\sigma_x$, representing pixel LUV color space information variance.

$\sigma_c$, representing pixel coordinate information variance.

$x_i, x_j$ represent pixel coordinates, and $||x_i - x_j||$ represents the Euclidean distance of two pixels.

$c_i, c_j$ represents pixel LUV color value, and $||c_i - c_j||$ represents the difference between two pixels.

**Ncut Solves Segmentation Results**

According to Section 3.1.2 and 3.2.2, the similarity matrix is obtained, and then the degree matrix can be obtained according to formula (14),(17). The segmentation results are realized by kmeans clustering method.

The calculation method of kmeans is as follows:

1) Select K centers randomly.
2) Traverse all data and divide each data into the nearest central point.
3) Calculate the average value of each cluster and act as a new central point.
4) Repeat 2-3 steps until the K centers are unchanged (convergent) or enough iterations are performed.

**Steps of Ncut**
The similarity matrix \( W \) is normalized into a Laplacian matrix, and then the eigenvectors of the Laplacian matrix are solved. K-means clustering is performed on the K-means minimum eigenvectors, so that the original image can be divided into \( k \) disjoint subgraphs, i.e. \( K \) (far less than \( V \) (number of vertices)). The flow chart show in above Fig. 4.

**Segmentation Results**

**Evaluation Criteria**
In this paper, image segmentation time and Region Consistency are selected as criteria for evaluation results. The smaller the values of these two indicators, the better the effect.

**MeanShift and Ncut Remote Sensing Image Segmentation Results**
The remote sensing image segmentation results show in Fig. 4 and Fig. 5. The experimental data are from NWPU-RESISC45 dataset[6].

As can be seen from Fig. 4 and Fig. 5, the combination of MeanShift and Ncut algorithm has better segmentation results and better edge preservation than Ncut algorithm. The former reduces the phenomenon of over-segmentation and under-segmentation compared with the latter (shown in the yellow box in Fig. 4 and Fig. 5). After smoothing the image with mean shift algorithm, the adjacent and similar pixel values in the image are equal. According to formula (14), the similarity of such pixels will be higher, which is equivalent to giving similar pixels larger weights.

So Table 1 and Table 2 will show a comparison of the regional consistency of the segmentation results of the segmentation algorithm. It can be seen that the improved algorithm has better regional consistency than the Ncut algorithm.

![Figure 4. Residual segmentation results.](image1.png)  
![Figure 5. Farmland segmentation results.](image2.png)

**Table 1. Comparisons of farmland image segmentation algorithms.**

| Segmentation classes | \( k=15 \)   | \( k=20 \)   | \( k=30 \)   |
|----------------------|---------------|---------------|---------------|
| Regional consistency | Ncut          | 0.3865        | 0.3812        | 0.3805        |
|                      | MeanShift+Ncut| 0.3797        | 0.3780        | 0.3754        |
Table 2. Comparison of image segmentation algorithms in residential areas.

| Segmentation classes | $k=10$ | $k=20$ | $k=30$ |
|----------------------|--------|--------|--------|
| Regional consistency | Ncut   | 0.6948 | 0.6457 | 0.6399 |
|                      | MeanShift+Ncut | 0.6837 | 0.6403 | 0.6155 |

**ERS and Ncut Natural Scene Image Segmentation Results**

The natural scene image segmentation results show in Fig.6, the data from BSDS500 dataset.

Table 3. Segmentation time of BSDS500 dataset[7].

| Image name | 41004 | 8068 | 196027 | 24004 |
|------------|-------|------|--------|-------|
| Segmentation time(s) | Ncut     | 43.7580 | 36.1974 | 27.2292 | 32.2611 |
|              | ERS+Ncut | 1.5495   | 1.4714   | 1.6597   | 1.6123   |

It is obvious from Table 3 that the improved algorithm takes much faster time than the Ncut algorithm. Because the super-pixel region is used instead of a single pixel to calculate the weight matrix, the calculation time is less. As can be seen from Fig.6, the improved algorithm achieves better segmentation results. Because the Entropy Rate Superpixel can preserve the edge of the target well, after clustering by Ncut algorithm, the segmentation result is better than that by Ncut algorithm.

![Figure 6. Ncut and ERS combine Ncut segmentation result.](image)

**Acknowledgments**

This work is supported by a grant from the Sichuan Major Science and Technology Special Project (No. 2018GZDZX0017).
References

[1] Felzenswalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59(2): 167-181.

[2] Shi J, Malik J. Normalized cuts and image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22 (8): 888-890.

[3] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. TPAMI, 24(5):603–619, 2002

[4] Achanta R, Shaji A, Smith K, et al. Slic superpixels, EPFL 149300[R]. [S.l.]:cole Poly technique Fédéral de Lausssanne (EPFL), 2010.

[5] Liu Ming-Yu, Tuzel O, Ramalingam S, et al. Entropy rate superpixel segmentation [C]/Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2011:2097-2104.

[6] NWPU-RESISC45 dataset, remote sensing image scene created by Northwest Polytechnic University, web site: http://www.escience.cn/people/JunweiHan/NWPU-RESISC45.html.

[7] Contour Detection and Hierarchical Image Segmentation. P. Arbelaez, M. Maire, C. Fowlkes and J. Malik. IEEE TPAMI, Vol. 33, No. 5, pp. 898-916, May 2011.