

Natural image segmentation based on mutual information

Yiwei Zhu*

School of Electronic and Information Engineering, Zhejiang Business Technology Institute, 315012, China

Keywords: Natural image segmentation, Mutual information, Superpixel.

Abstract. Natural image segmentation plays an important role in the fields of image processing and computer vision. Image segmentation based on clustering is an important method in unsupervised image segmentation algorithms. But there are two problems with this type of approach. First, feature extraction is generally pixel-based, which results in poor segmentation results and boundary fitting. In order to solve this problem, it is proposed to introduce super pixel to be segmented image preprocessing. Second, the number of partitions is difficult to determine. Aiming at this problem, an energy difference based on mutual information is proposed, which can automatically determine the number of partitions. The experimental results on the standard database show that the proposed algorithm overcomes the above problems and achieves better experimental results.

1 Introduction

Natural image segmentation is a classic problem in computer vision [1-4]. A good segmentation algorithm should classify pixels into partitions with similar statistical features (color or texture), and the boundaries of these partitions are relatively simple and spatially accurate. Commonly used segmentation methods include clustering, thresholding, and region growing. The Gaussian Mixture Models (GMM) clustering method has been widely used because of its extensive research and excellent results in machine learning, and has become the de facto standard in such methods [5, 6]. However, from a statistical perspective, the GMM method has at least two difficulties in natural image segmentation applications:

1. The statistical properties of local features of natural images (such as colors, textures, edges, contours, etc.) usually do not show the same consistency at the same spatial scale. This is not the case with different natural images, as is often the case with different regions of the same image. Therefore, the segmentation result is not unique. Instead, it should be a hierarchical result of segmentation at different scales.

2. Even if the segmentation results at different scales are not considered, the textures of different regions still have different intrinsic complexity. This makes determining the

* Corresponding author: Zyw@zjbti.net.cn

correct number of partitions and dimensions a difficult statistical problem. For example, if GMM is used to model different texture features, the dimension of a Gaussian component with a simple texture will be lower than the dimension of a Gaussian component with a complex texture.

This paper focuses on unsupervised image segmentation. Existing common methods include feature-based methods such as Mean-Shift, graph-based methods, region-based split-merge methods, and global optimization methods based on energy functions or shortest description lengths. In recent years, research has focused on how to fuse texture information at different scales. For example, regional growth or split-merge techniques use Markov random fields to model textures or other image cues.

Aiming at the existing problems, this paper adopts super pixel and mutual information to maximize mutual information, and uses mutual information energy difference as a new criterion to determine the number of partitions. Based on this, a new design is designed. Image segmentation algorithm to overcome the shortcomings of GMM algorithm.

2 Mutual information for image segmentation

2.1 Shannon entropy and mutual information

Given two random variables X, Y . If they are discrete, then we remember

$$\begin{aligned} p_i(X) &= p(X = x_i) \\ p_i(Y) &= p(Y = y_i) \\ p_{ij} &= p(X = x_i, Y = y_i) \end{aligned} \tag{1}$$

Their entropy is defined as

$$\begin{aligned} H(X) &= -\sum_i p_i(X) \log p_i(X) \\ H(Y) &= -\sum_i p_i(Y) \log p_i(Y) \\ H(X, Y) &= -\sum_{i,j} p_{ij} \log p_{ij} \end{aligned} \tag{2}$$

Conditional entropy is defined as

$$H(X | Y) = H(X, Y) - H(Y) = -\sum_{i,j} p_{ij} \log p_{i|j} \tag{3}$$

The base of the logarithm determines the unit of information measured. In particular, if the base 2 is used, the information is measured in bits. In this article, we use natural logarithms. The definition of mutual information is

$$\begin{aligned} MI(X; Y) &= H(X) + H(Y) - H(X, Y) \\ &= \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{p_i(X) p_j(Y)} \end{aligned} \tag{4}$$

It can be shown that $MI(X;Y)$ is non-negative, if and only if X and Y are completely unrelated, $MI(X;Y)=0$.

For continuous variables, the container is typically introduced first, and then the above formula is applied to the variable corresponding to the container. There are many ways to estimate the mutual information and entropy of continuous variables. This paper uses the estimation method proposed in the paper [7].

2.2 Coherent clustering based on mutual information

Mutual information itself can be used to measure the similarity of two random variables, and a smaller value indicates a larger distance. Based on this, we can design the following hierarchical clustering algorithm:

- (1) Calculate a similarity matrix based on mutual information, and allocate one cluster per vector;
- (2) Find the two closest clusters i and j ;
- (3) Combine i and j to form a new cluster (ij) ;
- (4) Delete the rows and columns indexed i and j from the similarity matrix, add new rows and columns for the new cluster (ij) , including the mutual information of the new cluster (ij) and all other clusters;
- (5) If the termination condition is not met, go to step (2); otherwise, end.

3 Image segmentation based on mutual information x2

This section describes how mutual information is applied to natural image segmentation. First we give the image segmentation target. Secondly, the texture features extracted in this paper are described. Then introduce the use of super pixels to improve the segmentation effect. Finally, the flow of the whole algorithm is given.

3.1 Image segmentation target

Since the distance between two random variables is defined, it is natural that in the image segmentation problem, we can define the following image energy function as the segmentation target.

$$E_n = MI(I, S_n) \tag{5}$$

Here I is the original image, and S_n is the segmentation result containing the number of n categories. However, in the study, we found that the distance between the image I to be segmented and its segmentation result S_n decreases as the number of classes in the segmentation result S_n increases, and converges to its maximum value $MI(I, I)$. In order to compare between different images, we redefine the segmentation target.

$$E_n = \frac{MI(I, S_n)}{MI(I, I)} \tag{6}$$

It is known that this is equivalent to the definition of equation (5), from which we can redefine the energy difference.

$$\Delta E_n = E_n - E_{n-1} \quad (7)$$

We find that as the number of classes in the S_n is reduced, the energy difference ΔE_n increases or oscillates (depending on the image segmentation algorithm and the image itself).

How to determine the number of categories or the number of partitions in clustering and image segmentation has always been a difficult problem in research, and has been extensively studied. A number of methods have been proposed to estimate the number of classes, such as AIC (Akaike Information Criteria), BIC (Bayesian Information Criteria), etc. [8]. Unfortunately, their theoretical basis and the effects in practice are unsatisfactory. In this paper, a new method based on mutual information energy difference measure is proposed to determine the algorithm. The algorithm can be briefly described as follows: For image I , the number of partitions n of the segmentation result starts from the number of superpixels, then decreases one by one, and ΔE_n is calculated every time. When $\Delta E_n > \varepsilon$, stop decreasing, and n is the number of divided blocks. ε is a given threshold and is obtained through training.

3.2 Feature extraction

First convert the image from the *RGB* color space to the *LAB* color space. The *LAB* color space is closest to human perception of color and is better able to represent textures through Gaussian mixture models.

In the literature, there are two main methods to extract local texture features. The first uses a two-dimensional filter bank to extract textures [9, 10]. The second method intercepts a $w \times w$ rectangular window around each pixel, and then stores the pixel values in this window as a vector as a local feature of that pixel [11]. Studies have shown that although the features extracted by the first method are more complex, the effect of using image segmentation is equivalent to the second method [12]. In this paper, the second method is adopted, which takes a 7×7 window around each pixel and then uses *PCA* to drop to 8 dimensions as the texture feature of the corresponding pixel.

3.3 Superpixel

After extracting the texture features from the image, the conventional idea is to directly divide the image into small blocks according to the clustering result using the algorithm in Section 2.2. Figure 1 shows the results of two standard images, baboon and lena, split using this method. We can see that this method produces many small split blocks. The edges of these small segments are not themselves true boundaries. This situation obviously should not happen for a good segmentation algorithm.

In order to eliminate such small noise-like segments, an image is first segmented by simple cues such as color boundaries by a low-level segmentation algorithm. That is, we first split the image into hundreds of small, compact areas called superpixels. Reference [13] suggests that all regional fusion algorithms use this pre-processing step. This paper uses the SLIC (Simple Linear Iterative Clustering) algorithm proposed by the literature [14]. Compared with other existing algorithms, this algorithm is fast and effective. Figure 2 shows the results of the SLIC superpixel segmentation algorithm for the separation of two standard images baboon and lena.

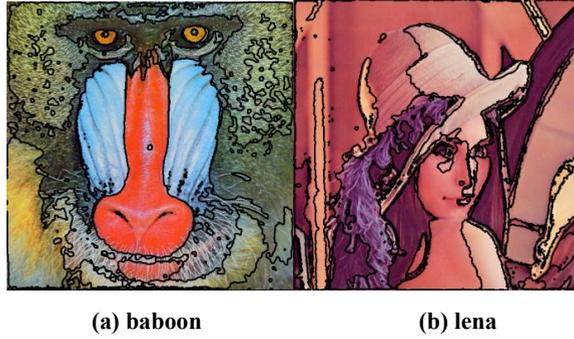


Fig. 1. General idea segmentation results.

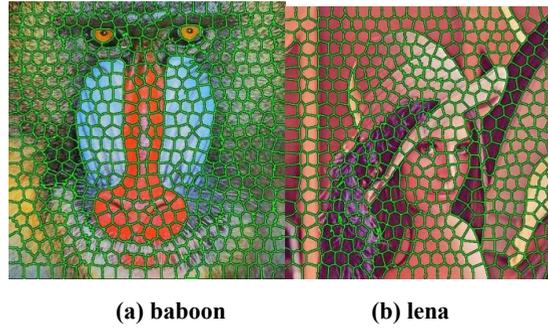


Fig. 2. SLIC superpixel segmentation results.

Because the superpixel segmentation algorithm has a good segmentation effect on those sharp boundaries, the small segmentation block with unclear boundaries as shown in Figure 1 is significantly reduced. The features of all the pixels in each superpixel are initialized into the same segmentation block, so that those pixels close to the boundary are always together with the pixels at the center of the segmentation block during the subsequent segmentation block merging process. Another benefit of preprocessing a segmented image with a superpixel algorithm is a significant drop in computational effort. Using superpixels as the initial partitioning block, the algorithm only needs to merge hundreds of regions instead of combining hundreds of thousands of pixels.

3.4 Algorithm

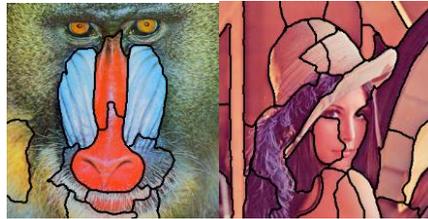
The image segmentation algorithm proposed in this paper takes the maximization of mutual information as the segmentation target, and uses the energy difference as the basis for judging the number of segments. The main steps of the algorithm are as follows:

- (1) Given the image to be segmented I ;
- (2) Segmented into n superpixels by *SLIC* as n initial partitions;
- (3) Calculate E_n using equation (6), one cluster per partition, and calculate a similarity matrix;
- (4) Find the two closest partitions i and j ;
- (5) Calculate their combined energy E_{n-1} ;
- (6) If $E_n - E_{n-1} < \varepsilon$, go to step (7); otherwise end;
- (7) Combine i and j to form a new partition (ij);
- (8) Update the similarity matrix. The rows and columns indexed i and j are deleted

from the similarity matrix, and new rows and columns are added for the new partitioning block (ij), including the mutual information of the new partitioning block (ij) and all other partitioning blocks.

4 Experiment

In order to test the algorithm proposed in this paper, the standard images baboon and lena are first tested. *Figures 3 (a) and (b)* show the segmentation results of the algorithm in this paper. *Figures 3 (c) and (d)* show the segmentation results of GMM using 5 Gaussian components, and *Figures 3 (e) and (f)* show The segmentation result of GMM using 10 Gaussian components. Since the segmentation result using the GMM method is too thin, the clear segmentation blocks cannot be divided by lines like the algorithm in this paper, so their average color is used here to represent that class. It can be seen from the comparison that although the segmentation result in this paper is not perfect, basically several clear segments with certain meanings are given, and there is no need to specify the number of partitions. On the contrary, although the segmentation result using the GMM method can also divide the pixels into manually designated categories according to the clustering result, it is not possible to form a clear segmentation block.



(a) Algorithm of this paper baboon (b) Algorithm of this paper lena



(c) GMM5 segmentation baboon (d) GMM5 segmentation lena



(e) GMM10 segmentation baboon (f) GMM10 segmentation lena

Fig. 3 Comparison of standard image segmentation results.

In order to study the performance of the algorithm in this paper, we tested it with the Berkeley image segmentation database BSDS500 [15]. The BSDS500 contains 500 color natural images with a resolution of 481×321 or 321×481 , of which 100 images are

used for training and 400 images are used for testing. Each image has 4 to 7 expert hand segmentation results used as a real segmentation result compared to the results given by the image segmentation algorithm. Figure 4 shows some of the results. It can be seen that the algorithm in this paper can also achieve good segmentation results in large standard databases.



Fig.4. Image segmentation results in the BSDS500 database.

5 Conclusions

The segmented image can be viewed as a simplified representation of the original image. When the mutual information between the original image and the segmentation result is maximized as the segmentation target, it can be considered that the segmentation result obtained is the segmentation containing the most information of the original image. The energy difference based on mutual information is a good description of the degree of reduction of the original image information contained in the segmentation result as the number of segmentation blocks decreases. When the energy difference based on mutual information is used as a criterion for determining the number of divided blocks, a balance is actually obtained between the number of divided blocks and the original image information.

In this paper, a large number of theoretical analysis and experimental results show that the energy function and energy difference based on mutual information reveals an intrinsic relationship between the segmentation result and the image to be segmented, and the image segmentation algorithm based on this can be very good. Make up for the shortcomings of the GMM algorithm. At present, this paper only gives the implementation and results of

threshold segmentation of images with a certain feature. The next research direction: try to fuse other clues such as contours, boundaries, etc., combined with other segmentation algorithms, in order to obtain more Good segmentation effect.

References

1. ALPERT S, GALUN M, BRANDT A, et al. Image segmentation by probabilistic bottom-up aggregation and cue integration [J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2012, 34(2): 315-27.
2. BROWN E S, CHAN T F, BRESSON X. Completely convex formulation of the Chan-Vese image segmentation model [J]. *International journal of computer vision*, 2012, 98(1): 103-21.
3. SCHOENEMANN T, KAHL F, MASNOU S, et al. A linear framework for region-based image segmentation and inpainting involving curvature penalization [J]. *International journal of computer vision*, 2012, 99(1): 53-68.
4. FORSYTH D, PONCE J. *Computer vision: a modern approach* [M]. 2 ed. New Jersey: Prentice Hall, 2013.
5. JI Z, XIA Y, SUN Q, et al. Fuzzy local Gaussian mixture model for brain MR image segmentation [J]. *Information Technology in Biomedicine, IEEE Transactions on*, 2012, 16(3): 339-47.
6. SEFIDPOUR A, BOUGUILA N. Spatial color image segmentation based on finite non-Gaussian mixture models [J]. *Expert Systems with Applications*, 2012, 39(10): 8993-9001.
7. KRASKOV A, ST GBAUER H, GRASSBERGER P. Estimating mutual information [J]. *Physical Review E*, 2004, 69(6): 99-138.
8. POSADA D, BUCKLEY T R. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests [J]. *Systematic biology*, 2004, 53(5): 793-808.
9. MALIK J, BELONGIE S, LEUNG T, et al. Contour and texture analysis for image segmentation [J]. *International journal of computer vision*, 2001, 43(1): 7-27.
10. ZHU S-C, GUO C-E, WANG Y, et al. What are textons? [J]. *International Journal of Computer Vision*, 2005, 62(1-2): 121-43.
11. OLIVA A, TORRALBA A. Modeling the shape of the scene: A holistic representation of the spatial envelope [J]. *International journal of computer vision*, 2001, 42(3): 145-75.
12. VARMA M, ZISSERMAN A. Texture classification: Are filter banks necessary?; proceedings of the Computer vision and pattern recognition, 2003 Proceedings 2003 IEEE computer society conference on, F, 2003 [C]. IEEE.
13. KANUNGO T, DOM B, NIBLACK W, et al. A fast algorithm for MDL-based multi-band image segmentation [M]. Springer, 1996.
14. ACHANTA R, SHAJI A, SMITH K, et al. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods [J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2012, 34(11): 2274-82.
15. ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation [J]. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2011, 33(5): 898-916.