

Semantic Annotation Method of Clothing Image

Lu Zhaolao, Mingquan Zhou, Wang Xuesong^{*}, Fu Yan, and Tan Xiaohui

College of Information Science and Technology, Beijing Normal University, Beijing, China
`{lzl,mqzhou,wangxs,fuyan}@bnu.edu.cn, xiaohuitan@163.com`

Abstract. Semantic annotation is an essential issue for image retrieval. In this paper, we take the online clothing product images as sample. In order to annotate images, we first segment the image into regions, then remove the background and noise information. The illumination and light interference is considered too. Cloth position and region are determined by rules. Images are translated into some features. Visual words are prepared by human and calculate methods. Finally, Image features are mapped to different visual words. Pre-processing and post-processing steps which uses face recognition method and background rule analysis are applied. Finally, some segmentation and annotation results are given to discuss the method.

Keywords: Semantic annotation, Image segmentation, Graph cut.

1 Introduction

Nowadays, with the growing of Internet and digital camera equipment, the number of digital images steps into a geometric progression growth. It is estimated, images stored in the Web has reached ten billion orders of magnitude. On the one hand, the vast amounts of web images have result the disaster of information retrieval. It is difficult to obtain the desired image from the Internet. On the other hand, it is also a powerful way to spread the e-commerce rapidly. In order to meet people's need, semantic annotation is focused.

1.1 Image Segmentation for Semantic Annotation

The semantic annotation is the basis of the image semantic retrieval, and adds keywords to reflect the image semantics by a specific algorithm to the image, and use the marked images or other information automatically learn the semantic and visual feature space, and established relationships marked image. The main image annotation methods include two types: global features of image annotation (Global-Based Image Annotation) and by region-based image annotation methods (Region-Based Image Annotation). The global features of images based on global statistical characteristics marked the [1-5], does not require segmentation and object-oriented analysis, but provide only coarse-grained semantic description, cannot provide more image detail semantics; marked by region-based image by image segmentation, semantic

^{*} Corresponding author.

understanding can provide a low-level visual features of the regional level, more and more researchers focus, forming a series of effective methods [6-10].

In this paper, we take the online clothing product images as sample. The method establishes the global and region features of the images too. Pre-processing and post-processing steps which uses face recognition method and background rule are applied.

1.2 Related Works

Image segmentation is the basic work and an important part of [11-13] image annotation and analysis. It is one of the hot research fields in computer image processing and machine vision. Automatic segmentation has still many difficulties and problems, but has been made part of the research results. Segmentation method using graph theory can change complex image processing problems into computable problem.

Image segmentation method, but there is no common, universal segmentation method. Mainstream image segmentation method include: regional competition and merger-based segmentation method [3], segmentation method based on graph theory [14-16], atlas-based segmentation method [17], based on gray level threshold segmentation method [19,20], edge detection-based segmentation method, based on target feature model segmentation method [18], based on level set segmentation of the Level Set method [21,22] based on Active Contour (Snake) [23, 24] segmentation method.

1.3 Graph-Cut Based Image Segmentation

Based on graph theory, image segmentation technology has made great progress in recent years. It has become hot spots of image segmentation [25]. Fixed thresholds and local features of the earlier graph cut algorithm to partition Zahn [26] using the minimal spanning tree method of image segmentation, the weight of edges in the graph using the gray value of the pixel distance calculation, divided the criterion of a maximum weight of minimum spanning tree (MST) segmentation method.

Wu and Leahy [27] proposed a map segmentation method, using the minimal cut for image segmentation by finding the minimal cut minimizes the similarity between the segmented pixels as segmentation criteria. Shi and Malik [28] proposed an improved Normalized Cut segmentation method. The method is also a minimum cut, but is calculated for all the right side of the connection value in the calculation of connection in the entire map edge centralization value of the share component. Grady and Schwartz [29] calculated for Ncut method for complex image size. Felzenszwalb [30,31] consider the overall visual effect, the use of global information and efficient function of the weights and the segmentation criteria, map segmentation by region merging.

2 Image Representation and Regional Definition

2.1 Image HSI Model Represents

Segmentation depends on the expression of the color image, color models, including RGB, HSI, due to the complexity of natural images, simple gray-scale information, likely to cause over-segmentation or under segmentation. HSI model based on the

characteristics of human vision is used for it is closest to the human visual point of view to describe the image.

The statistical characteristics of the clothing image are relatively stable and very suitable for image segmentation as a graph theory basis for the weights. Red, green and blue components of the image pixels, respectively, for the R, G, B, and H, S, I can define the image RGB to the HSI conversion functions.

2.2 Image with a Weighted Undirected Graph

The image data is a two-dimensional matrix of pixels, in the context of graph theory, the figure is the number of nodes and connections between different nodes, the edge of the geometry of. The graph of the graph theory used to represent relationships between objects, the object itself as a graph node v_i , the connection between the object node to the edge e_i . Graph theory algorithms used in image processing pixel map mapping [34, 35].

To image P , Each pixel as a vertex $v_i \in V$, Set V is a set of all vertices; any two pixels constitute an edge $e(v_i, v_j) \in E$, marked e_{ij} , Set E is the collection of all edges ; the whole image corresponding to the undirected graph $G = (V, E)$.

If each side of the figure corresponds to a relative weight $w(v_i, v_j) \in W$, marked as w_{ij} , Set W is the collection of all edge weights set ; ;;;; whole image corresponding to a weighted graph is expressed as $G = (V, E, W)$.

2.3 Image Map on the Partition Defined

Graph theory, image segmentation method, in fact, the threshold to complete the fuzzy transformed into optimization to solve problems, form the point of clustering results under certain conditions.

In a variety of graph theory of image segmentation method, the pixel as a node, as a side to the correlation between the pixels, the image is converted to weight undirected graph processing. Our image segmentation using graph theory to obtain the best segmentation of images.[32]

The graph theory of image segmentation, in fact transformed into a weighted diagram $G = (V, E, W)$, calculation for the minimal cut sets. Vertices of the graph corresponds to each pixel in the image P , the image P segmentation is equivalent to the set of vertices V divided into disjoint subsets. The definition of Graph Cut can be expressed as[36]:

$$V = \bigcup_{i=1}^k V_i, V_i \cap V_j = \emptyset, i \neq j, i, j = 1, 2, \dots, k, S(V_1, V_2, \dots, V_k) \text{ is a segmentation of } V, k \text{ is the number of blocks.}$$

The goal of the vertex set V partitioning is to split a subset of nodes within the same class difference between pixels, the pixels of different classification between node dissimilarity. If all nodes of the graph is divided into m subsets V_1, V_2, \dots, V_m , segmentation criteria [36]: (1) Partition of the set V_i of internal vertices relevance and similarity; (2) The relevance and similarity between V_i and V_j different subsets is low.

Image segmentation, the graph G 'belongs to G , G' is divided into A and B , two-part cost function:

$$S(A, B) = \sum_{u \in A, v \in B} w(u, v)$$

The smallest division of the cost function S corresponding graph G is optimal binary divide, the split graph G , corresponding to the best split of the original image.

2.4 Image Feature Representation

Image segmentation to extract the appropriate low-level visual features, said to facilitate follow-up semantic annotation and similarity calculation. Contour and texture of the clothing commodity characteristics with the figure standing, layouts, block, etc., are very different, therefore the characteristics of the image area is the most stable color characteristics, combined with the macro issues the analysis is complete.

The entire image overall features, you can target prospects after removing the background area, quantified HSI space, computing 128-dimensional histogram of the overall feature of the dress pictures. $L = \{l_1, l_2, \dots, l_{128}\}$, 其中 $i = 1, 2, \dots, 128$

3 Image Segmentation Method

3.1 Image Preprocessing

Picture of online apparel goods into two categories, one category is the single items of clothing of the online store, usually the better picture quality, the background is relatively simple; pictures of people with complex background and dress models to deal with relatively complex.

Targeted to deal with the clothing image, the first category of images need to remove the background, the second category of images need to remove the interference of the face, hair, etc. For clothing commodities image, we set two types of rules:

- **Rule 1** simple background rules

Assume that the commodity picture for a simple background image, then the edge of the area of the larger connected region as the background area, the central region of the prospects for apparel region;

- **Rule 2** figure background rules

Assumed that the image has a character background, character clothing worn should be consistent with the rules of human proportion, costumes position and the position of the face with a specific topology.

3.2 Face Detection and Location and Background Removal

Apparel goods pictures often include models, how to extract the pure clothing area, there will be conducive to the subsequent segmentation and labeling. The most effective method to detect figures is to detect face in the image. It can help to determine the human faces' location, size, and numbers in the given images.

The face detection is using a template matching algorithm with default template. In order to improve the reliability and validity, we limit the minimum pixel scale and the largest proportion of the face images.

3.3 Weight Calculation of Graph Cuts

After the split of a subset, the degree of similarity need to quantify the calculation and description, usually expressed as a weight function w. The similarity degree function between the different pixel nodes, usually calculated by the distance of node.

Weight function and similarity function can take many forms and methods of calculation. The theory is mapping actually pixel to a feature space to calculate the quantitative distance in the space. The common form is based on the pixel gray value of the weight function. The function using local feature analysis of similarity:

Analysis of the commodity image pixel distance and color information is very important. It needs a variety of information in order to achieve effective segmentation. Each pixel mapping feature point (x, y, and h, s, i), where (x, y) pixel location information, (h, s, i) is the hue brightness of the pixel message.

HUE functions between pixel values and adjacent from the constructor the right functions, consider the global and local features [36], constructing an image pixel chrominance distance function:

$$w_{ij} = \begin{cases} \exp\left(-\frac{\|F_i - F_j\|_2^2}{\sigma_f^2}\right) \times \exp\left(-\frac{\|X_i - X_j\|_2^2}{\sigma_X^2}\right) & \text{if } \|X_i - X_j\|_2^2 < r \\ 0 & \text{others} \end{cases} \quad (1)$$

Where F is the quantization of the chrominance hue brightness function used to represent the color properties of pixels; X represents the pixel location, r is the associated pixel radius of the specified pixel, the two values σ is the color and distance of the Gaussian smoothing parameter. The function of the weights between the two pixels are color values the closer the greater the similarity between two pixels, the closer the distance between two pixels.

3.4 Graph Segmentation Method

Usually graph theory, image segmentation using split larger regional, and gradually obtained the spin-off results, the discriminated function to determine the minimal cut sets to make splitting the cost minimum. The method due to uncertainty in the segmentation process, it is difficult to optimize the partition.

We adopt a bottom-up merging the idea [31], to initialize a smaller area, gradually merged to form suitable region segmentation. An image corresponding to a weighted graph $G = (V, E, W)$, suppose there are n nodes and m connections, to get the corresponding image segmentation results.

4 Image Region Analysis Rule

4.1 Product Image Segmentation Implementation

The core algorithm of graph theory have been listed in the front, the original image processing and analysis need to pretreatment, post-region coloring, format conversion and a series of auxiliary work. The overall process of the algorithm is summarized as follows: (1) Read the given image, convert the data format, structure weighted directed graph G.(2)The right value for a given function, the weights for calculation of weight W values;(3)according to the weight W value, the consolidation of the computational domain to obtain the final image region segmentation results;(4)Segmentation second merger absorption of discrete points of the regional;(5)Reasonable detection of the image area, remove the background and foreground region;(6)Get the image segmentation results of the analysis and feature extraction, the output feature vector.

4.2 Graph Theory Segmentation of Short-Range Absorption Algorithm

For apparel merchandise picture image segmentation, focus on the analysis of the target area for example .we determine the location and area of the clothing. As part of the dress patterns and designs, easy to region embedded within the texture segmentation. To eliminate too small parts, we use large area contains a small region to merge.

The region merging post-processing criteria of the need to maintain the original segmentation results, while meeting the basic criteria: (1) to guarantee the value of the similarity between the different degrees and classes in the class, meet the threshold of the threshold L requirements on a smaller area to absorb; (2) the size of the region of smaller area to less than 1/10 of the larger size of the area.

4.3 Rules Limit Divided by Graph Theory

Image segmentation forms a series of block regional. According to the rules described in the chapter 2.1 template methods for authentication. Images contain faces, according to four times the face width, 12 times of face height to determine the clothing active area for the region in the color segmentation block. if its center is located outside the region, discarded the block as a complex background.

For color segmentation of the image area, if the image area that contains the edge pixels, and the central location of the regional block located on the edge of the 1/9 of the region, background removal happened for the face image rely on edge detection.

Image feature calculation includes the overall characteristics of the image and the regional characteristics. Feature description of the main block is a 128-dimensional feature for each sub-block together with the center block and the number of pixels, as well as the main color values can be further described.

5 Image Annotation

5.1 Product Image Main Color Annotation

To each image, we use to clear the background after the first hue and the second hue representative of the color properties, due to the presence of Person class apparel pictures in color feature extraction, on the picture, in accordance with the upper and lower portions to be divided. The the calculated color color label said, can form a multi-level representation. Clothing color accuracy requirements are not very high, you can select a less standard color expression.

5.2 Product Image Classify and Attribute Annotation

Different goods belonging to a specific category, such as:T-shirt, jacket, men windbreaker, etc. Through extracting and validating the tag words and further image content verification, we get the vocabulary to annotate online commodity pictures. The result should be a relatively accurate one.

In addition to the generic commodities, garments, due to the design style and color, style difference, the formation of a range of potential properties, such as: leisure, sports, etc. described in words. These words are easy online analytical process, to provide more personalized search and shopping guide service.

6 Image Segmentation and Annotation Results Analysis

6.1 Character Image Segmentation Results

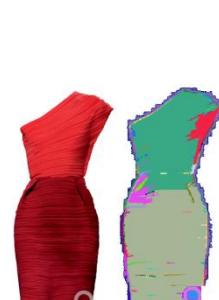
The experimental results are the Gaussian smoothing parameter 0.5.[30] In our algorithm contains the threshold of the independence of the small area in order to reduce over-segmentation problem. For product images with human, the original image in Fig.1(a) Picture of experimental results in Fig.1(b).



(a) Original image



(b) result image



(a) original image



(b) result image

Fig. 1. Image with human Results**Fig. 2.** Pure clothing image Results

Pure Clothing Image segmentation Results are shown in Fig.2. Segmentation results in Fig.2.(a) of the original image, Fig.2.(d) the results of this article HSI distance function graph cut.

6.2 Image Annotation Results

We download images from the web, the color annotation results. We got color vocabulary {red} for Fig.1 image and {Pink, Light brown} for Fig.2 image. Each image, we got a classify annotation results. In our example, the classify vocabulary is respectively named {women dress} and {Tight skirt}.

After we got the classify words, attribute annotation results are reached by image content and tag word analysis. We got some attribute for above samples, {beautiful, charming, and young} for Fig.1 and {professional, fresh, sleeveless} for Fig.2.

7 Conclusion

Graph theory for image segmentation can achieve good results. Experimental results show that the HSI model can be very effective with such image analysis needs. Using the machine learning methods, we can get the statistic modal characteristics of the image HIS features. Segmentation of the image results based on statistical knowledge to judge divided region need further research. Image semantic annotation should be treated in more detail.

Acknowledgements. This paper is supported by “the Fundamental Research Funds for the Central Universities”. Some example data is provided by BlueBerrytech inc., China.

References

1. Mori, Y., Takahashi, H., Oka, R.: Image-to-word transformation based on dividing and vector quantizing images with words. In: Proc. of Intl. Workshop on Multimedia Intelligent Storage and Retrieval Management (MISRM 1999), Orlando (October 1999)
2. Oliva, A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. Int. J. Computer Vision 42(3), 145–175 (2001)
3. Hare, J.S., Lewis, P.H.: Saliency-based models of image content and their application to auto-annotation by semantic propagation. In: Proceedings of the Second European Semantic Web Conference (ESWC 2005), Heraklion, Crete (May 2005)
4. Yavlin斯基, A., Schofield, E., Rüger, S.M.: Automated Image Annotation Using Global Features and Robust Nonparametric Density Estimation. In: Leow, W.-K., Lew, M., Chua, T.-S., Ma, W.-Y., Chaisorn, L., Bakker, E.M. (eds.) CIVR 2005. LNCS, vol. 3568, pp. 507–517. Springer, Heidelberg (2005)
5. Lavrenko, V., Manmatha, R., Jeon, J.: A model for learning the semantics of pictures. In: Proc. of Advances in Neural Information Processing Systems (NIPS 2003) (2003)

6. Yang, C., Dong, M., Hua, J.: Region-based image annotation using asymmetrical support vector machine-based multiple-instance learning. In: Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR 2006), New York, USA, pp. 2057–2063 (June 2006)
7. Feng, S.-H., Xu, D.: Transductive Multi Instance Multi Label Learning Algorithm with Application to Automatic Image Annotation. Expert Systems with Applications 37(1), 661–670 (2010)
8. Lu, H., Zheng, Y.-B., Xue, X.-Y., et al.: Content and context-based multi label image annotation. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Miami, FL, USA, pp. 61–68 (2009)
9. Liu, J., Li, M., Liu, Q., et al.: Image annotation via graph learning. Pattern Recognition 42(2), 218–228 (2009)
10. Wang, Y., Mei, T., Gong, S.G., et al.: Combining global, regional and contextual features for automatic image annotation. Pattern Recognition 42(2), 259–266 (2009)
11. Noble, J.A., Boukerroui, D.: Ultrasound Image Segmentation: A Survey. IEEE Transactions on Medical Imaging 25(8), 987–1010 (2006)
12. Lucchese, L., Mitray, S.K.: Color Image Segmentation: A State-of-the-Art Survey. Image Processing, Vision, and Pattern Recognition 67(2), 207–221 (2001)
13. Comaniciu, D., Meer, P.: Robust analysis of feature spaces: color image segmentation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, pp. 750–755 (1997)
14. Urquhart, R.: Graph theoretical clustering based on limited neighborhood sets. Pattern Recognition 15(3), 173–187 (1982)
15. Shi, J., Malik, J.: Normalized cuts and image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, pp. 731–737 (1997)
16. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 15(11), 1101–1113 (1993)
17. Weiss, Y.: Segmentation using Eigenvectors: A Unifying View. In: Proceedings of the IEEE International Conference on Computer Vision, Kerkyra, Greece, vol. 2, pp. 975–982 (1999)
18. Comaniciu, D., Meer, P.: Mean shift analysis and applications. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, pp. 1197–1203 (1999)
19. Ng, H.-F.: Automatic thresholding for defect detection. Pattern Recognition Letters 27(14), 1644–1649 (2006)
20. Zhang, Y.J., Gerbrands, J.J.: Transition region determination based thresholding. Pattern Recognition Letters 12(1), 13–23 (1991)
21. Osher, S., Sethian, J.A.: Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton–Jacobi Formulations. Journal of Computational Physics 79(1), 12–49 (1988)
22. Adalsteinsson, B., Sethian, J.A.: The fast construction of extension velocities in level set methods. Journal of Computational Physics 148(1), 2–22 (1999)
23. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active Contour Models. International Journal of Computer Vision 1(4), 321–331 (1988)
24. Li, P.-H., Zhang, T.-W.: Review on active contour model (Snakemodel). Journal of Software 11(6), 751–757 (2000)

25. Boykov, Y., Jolly, M.: Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. In: Proc. IEEE Conf. Computer Vision, Vancouver, Canada, vol. 1, pp. 105–112 (2001)
26. Zahn, C.T.: Graph-theoretic methods for detecting and describing gestalt clusters. *IEEE Transactions on Computing* 20(1), 68–86 (1971)
27. Wu, Z., Leahy, R.: An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(11), 1101–1113 (1993)
28. Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(8), 888–905 (2000)
29. Grady, L., Schwartz, E.L.: Isoperimetric Graph Partitioning for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28(3), 469–475 (2006)
30. Felzenszwalb, P., Huttenlocher, D.: Image segmentation using local variation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, California, USA, pp. 98–104 (1998)
31. Felzenszwalb, P.: Efficient Graph-Based Image Segmentation. *International Journal of Computer Vision* 59(2), 167–181 (2004)
32. Boykov, Y., Funka-Lea, G.: Graph Cuts and Efficient N-D Image Segmentation. *International Journal of Computer Vision* 70(2), 109–131 (2006)
33. Cooper, M.C.: The tractability of segmentation and scene analysis. *International Journal of Computer Vision* 30(1), 27–42 (1998)
34. Sarkar, S., Soundararajan, P.: Supervised learning of large perceptual organization:graph spectral and learning automata. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(5), 504–525 (2000)
35. Ding, C.H.Q., He, X., Zha, H., et al.: A min-max cut algorithm for graph partitioning and data clustering. In: ICDM 2001, pp. 107–114. IEEE Computer Society, Los Alamitos (2001)
36. Li, X.-B., Tian, Z., Liu, M.-G.: Weighted Cut Based Image Segmentation. *Acta Electronica Sinica* 36(1), 76–80 (2008)