Pre-detection Technology of Clothing Image Segmentation Based on GrabCut Algorithm

Lei Lei Deng¹

Published online: 28 November 2017

© The Author(s) 2017. This article is an open access publication

Abstract Image segmentation, as a key technology in digital image processing, serves as the basis for image processing and analysis and understanding. Its main purpose is to separate useful image information (i.e. foreground) from useless image information (i.e. background) in some way. With segmentation of clothing image in online shopping as the center, this paper studied image segmentation technology. Study and analysis of a large number of online clothing images found that image clothing can be roughly divided into two categories. The first is simply displayed clothing without human model, with several clothing displayed together. The other is image with human model. By distinguishing the two types of images with face detection algorithm and edge detection method, applying different location algorithms for the two types of images, and adjusting image location with iterative algorithm, ultimately, more accurate localization frames were obtained, which can replace the part in traditional classical GrabCut algorithm that requires manual participation, and realize automatic batch operation of image segmentation. The final test data proved effectiveness of the new improved algorithm which can be applied in retrieval system of mass images at the time of online clothing shopping.

Keywords Image segmentation · Clothing area location · Edge detection · Face recognition · GrabCut

1 Preface

With the rapid development of network information, use of image information transfer is becoming increasingly widespread, and there is a growing demand for analysis and retrieval of vast amounts of image information in network. How to quickly find needed information from tens of millions of image information, how to accurately extract useful

✉ Lei Lei Deng
deng_ll@163.com

¹ School of Information Technology, Jilin Agricultural University, Changchun, China

information from images, and faced with all kinds of existing algorithms with respective strength, how to open up new ideas based on the original foundation to meet the new demands for image processing have been the issues that researchers concern. Aiming at this problem, with image segmentation problem in clothing image retrieval of online clothing sales system as example, based on introduction of research status quo in first chapters and understanding of relevant basic knowledge, this paper made in-depth study on classical algorithm Crabcut, made improvement based on advantages and disadvantages of the algorithm, proposed clothing image segmentation algorithm based on pre-detection, and finally provided an objective evaluation of the new algorithm. The results show that the new algorithm achieves the desired effect [1].

1.1 Definition of Image Segmentation

Image segmentation technology has enjoyed many years of development, and scholars have provided different interpretations and expressions to define it. In general sense, image segmentation is the first step of image processing, which separates useful contents from useless contents, leaving useful image information called foreground; abandoning the part named as background. Only after this operation can subsequent higher level image processing operations be carried out. Put it more abstractly, image segmentation is to divide pixels in the image into different blocks according to certain characteristics, so that pixel property within the same block is similar, while pixel property of different blocks differs greatly. In computer programming technique, idea of set is preferred for description of image segmentation definition:

Suppose R is set of images to be segmented, if correct segmentation is $R_1, R_2, R_3, \dots, R_N$, it must satisfy the following five conditions:

- (1) $\bigcup_{i=1}^N R_i = R$;
- (2) For $\forall i, j, i \neq j$, there is $R_i \cap R_j = \emptyset$;
- (3) For $i = 1, 2, \dots, N$, there is $P(R_i) = T$;
- (4) For $\forall i, j, i \neq j$, there is $P(R_i \cap R_j) = F$;
- (5) For $i = 1, 2, \dots, N$, R_i are non-null;

Description 1 means all subsets equal to the original image set after combination. Description 2 means any two subsets are disjoint, that is, each division is independent. Description 3 means all elements of each subset are connected. Description 4 means elements of different subsets are not connected. Description 5 means each subset is non-null. Images can be basically segmented based on such definition. Application of expression of such definition in computer means great significance, but only such segmentation is insufficient for image segmentation which needs to mark the area people find useful and extract it. Usefulness of this area is entirely subjectively defined. Only when extraction of useful area is completed can image segmentation be completed in the true sense [2–4].

1.2 Initial Location Method

This paper proposed that we should start from face detection results, take detected face rectangular box as the reference position, follow certain human body proportion, make rough frame selection of clothing position in image, thereby making initial location of image segmentation. Face recognition technology is now one of technologies with focused

research by various industries, while face detection technology is its core technology, whose development also marks the development of face recognition technology. Its main process is to read images to be detected via algorithm program, detect whether face exists in images, and then judge position, size, dimensions, etc. [5].

Figure 1 shows simple process of face detection. After image input, the first is to extract facial features, which is a key step of the testing process that concerns face detector configuration. Face detector is to judge where is face in the image. Its output results are generally not unique, there will be overlaps, so integration of result needs to be set up to integrate and process output result of the detector, thus making face detection results more accurate.

According to the theory of human body proportion, as well as costume design principles, by combining characteristics of online clothing images, we summarize that position right below the face is mostly clothing area, whose size can be set in proportion according to size of face rectangle. For women, clothing size is substantially three times of length and twice of width of the head. For men, there is a relative increase in number of times, as clothing area size is about four times of length and three times of width of the head. For program algorithm, it is basically impossible to separate men from women via the image, and this is only a rough location. Therefore, in the paper, men's standard was selected for design. Suppose a face rectangular box is obtained through face detection algorithm, its length and width are $a * b$, and center coordinate of the rectangular box is (x_0, y_0) . According to conclusion of the above summary, it can be known that rectangular box size in initial clothing location is $4a * 3b$, abscissa of center coordinate of the rectangular box remains unchanged, vertical coordinate by $b/2 + 2b$, i.e., center coordinate of the clothing rectangular box is $(x_0, y_0 - 5a/2)$. In this way, initial clothing location of face image is completed [6–9] as shown in Fig. 2.

2 Clothing Edge Detection Method Research

People tend to notice the area in image where an object intersects with the other object when looking at an image, which is a physiological characteristic of human vision. People will unconsciously extract main information after image segmentation in the brain. Area where an object intersects with the other object is area with significant pixel gray value changes in image. The area information provides important information for location of the main body in image, and provides an important basis for segmentation of target object. Studies found that area with dramatic change in pixel gray value is often the edge of an object. If an algorithm can be devised to determine boundary of the object through changes in gray value, accuracy and efficiency of image segmentation will be greatly enhanced.

Fig. 1 Simple process of face detection

Fig. 2 Initial location result of clothing area with face

2.1 Edge Detection Method Based on Clothing

Image edge is an important useful information in image segmentation. Near the edge, pixel is big and value variation is discontinued, showing dramatic bounce. This characteristic provides a great help for to find the edge of objects in images. Researchers found that, there are basically three types of presentation forms of signal obtained along the edge, namely, ladder form, roof form and linear form. In ladder form, its characteristic is that there is sudden up or down change in gray value at a steady state, afterwards, it maintains equilibrium; in roof form, its characteristic is that gray value remains creeping up or down, then an inflexion point suddenly appears, afterwards, gray value is suddenly in creeping down or up; in linear edge form, its characteristic is that after sudden drastic upward or downward change following a steady state, gray value restores to its original state not after long [10]. Figure 3 shows edge of the three forms.

In the famous “Mach band effect”, the human eye will automatically enhance and adjust the portion with light intensity mutation. Normally, for position with light intensity mutation in image, gray value also changes violently. This area is the location of edge. Figure 4 is a result of clothing image processing with Canny algorithm.

Fig. 3 Three forms of edge

Fig. 4 Results of image edge detection with Canny algorithm

If there are images with face, there are certainly images without face. In fact, there are many images without face. Then how to make initial location of images without face? Clothing image features are mentioned in the second chapter and it can be judged from merchants' purpose of clothing photo shooting that to better display goods, merchants usually display the main body of clothing in the center of the image, which occupies most area of the image. Therefore, we specify that center of initial location box in images without face is in the center of the image, while length and width of the rectangular box respectively account for a certain proportion of length and width of the image itself. Through abundant image measurement, two average values α , β were measured. Wherein, $\alpha = 0.7$ is proportionality coefficient of image length, $\beta = 0.8$ is proportionality coefficient of image width. Figure 5 shows initial location results of the images [11, 12].

As can be seen from Figs. 1 and 5, clothing initial location in the eight images is inaccurate. Some has collar outside the rectangular box, and some has sleeves outside the

Fig. 5 Initial location results of clothing area without face

rectangular box. Such location results are not satisfactory, so subsequent precise adjustment is needed.

2.2 Accurate Re-location Method

2.2.1 Accurate Re-location of Images with Face

Combining edge detection algorithm, location box of target area of image is further refined in this paper so that the four sides are closer to clothing edges. For example, by gradually expanding outward, that is, making outward translation of the border for some distance, there will be a new expansion area between the resulting border after translation and the original border. Suppose the area is D as shown in Fig. 6. In this area, number of contained edge information was calculated by iterative computation according to edge detection algorithm and then saved. This process can be repeated several times. After completion of predetermined number of times, edge information amount obtained after each operation was compared. That containing the most edge information amount is its optimal solution, so exact border can be determined finally.

Calculation of translation step of left, right, and lower borders is shown in Fig. 6:

- (1) Specify that vertex of lower left corner of the original image locates at origin of coordinate axis, left edge of the original image coincides with vertical axis, the lower edge coincides with abscissa axis. Suppose length of the original image is L , its width is W .
- (2) If length and width of face rectangular box is $a * b$, the center coordinate is (x_0, y_0) , then it can be known that rectangular box size in initial clothing location is $4a * 3b$, abscissa of center coordinate of the rectangular box remains unchanged, vertical coordinate by $b/2 + 2b$, i.e., center coordinate of the clothing rectangular box is $(x_0,$

Fig. 6 Schematic diagram of accurate re-location border expansion

$y_0 - 5a/2$). Then distance between the left border and vertical axis is $w_L = -x_0 - 3b/2$, distance between the right border and right edge of the original image is $w_R = W - x_0 - 3b/2$, distance between lower edge and abscissa axis is $w_B = -y_0 - 9a/2$. In this paper, suppose expansion times is 3, then step size of the three borders after one expansion can be calculated. Suppose step size of the left border is d_L , step size of the right border is d_R , step size of the lower border is d_B , then, $d_L = (x_0 - 3b/2)/3$, $d_R = (W - x_0 - 3b/2)/3$, $d_B = (y_0 - 9a/2)/3$ [13–15].

2.2.2 Accurate Re-location Results of Image

To verify effectiveness of accurate relocation method for clothing image, experiment was undertaken in this paper. A thousand pieces of clothing images were downloaded from Vipshop, Taobao and other large online shopping platforms, to be received with clothing foreground location. Figure 7 shows part of experiment results with good effect. As can be seen from the figure, rectangular box after accurate location basically covers all the edges of clothing, and the borders are along the outer edge of clothing, which indicates that the algorithm filters out background noise as much as possible based on accurate location, so as to lay a solid foundation for further segmentation.

Among the numerous image location results, some are not satisfactory with error detection and leak detection. Among the 1500 images, there are 1315 images with correct location, 129 images with error detection and 56 images with leak detection, each accounting for 87.67, 8.6 and 3.73%.

3 Clothing Image Segmentation Algorithm Based on Pre-detection

This paper proposed an innovative automatic image segmentation algorithm—clothing image segmentation algorithm based on pre-detection. The algorithm adds pre-detection to classical algorithm Grabcut, thereby replacing manual participation in initialization in Grabcut algorithm. The so-called pre-detection is to make an edge detection and location

Fig. 7 Accurate re-location results of image

of image before algorithm segmentation. The process will automatically generate a rectangular box that contains foreground, which is consist with the effect of manually drawing separation box in Grabcut algorithm. Hence, improved algorithm by combining pre-detection technology and classical algorithm Grabcut will certainly sustain advantages of efficiency and accuracy of classic algorithm, and will also enjoy automatic image segmentation performance.

3.1 Segmentation Algorithm Flowchart

Figure 8 shows flowchart of clothing image segmentation algorithm based on pre-detection [16].

3.2 Segmentation Results

After the algorithm implementation, experiments were carried out for 1500 images in the paper. In order to meet scientificity, universality requirement of experiment samples, we selected clothing images in actual network sales system, including images with face, images without face, simple background images, complex background images. And clothing colors and styles are also as diverse as possible. After experiment results of the new algorithm were obtained, to facilitate reference and comparison, the 1500 images were segmented with classical Crabcut algorithm in this paper. Table 1 shows pictures of segmentation results with good effect, as well as picture of comparison between original image and results obtained through Crabcut algorithm processing [17, 18].

4 Result Evaluation

There are many evaluation criteria for image segmentation results. The two assessment indicators of recall ratio and precision ratio are the most widely used and most mature. Wherein, recall ratio shows how many correct segmentation results are completely segmented, whereas precision ratio shows how many obtained results are accurate, i.e. proportion of accurate part of segmentation results in total accurate segmentation results. The specific evaluation method is to first correctly segment all test samples, then indicate correct segmentation result of each image as R_{magic} , segmentation result obtained in the algorithm experiment as P_{magic} , recall ratio as $(R_{\text{magic}} \cap P_{\text{magic}})/R_{\text{magic}}$ and precision ratio as $(R_{\text{magic}} \cap P_{\text{magic}})/P_{\text{magic}}$.

Table 2 shows average recall ratio and precision ratio of segmentation results of 1500 images with new algorithm and classical algorithm. As can be seen from the table, recall ratio and precision ratio of new algorithm is slightly lower than that of classical algorithm, but new algorithm greatly exceeds classical algorithm in terms of efficiency, and achieves the goal of complete batch image segmentation [19, 20].

5 Conclusion

In this paper, pointing at the situation that existing image segmentation algorithm can not adapt to massive image data processing, clothing images in online clothing sales system were selected as research objects and a lot of preparation work was done for innovation of

Fig. 8 Flowchart of clothing image segmentation algorithm based on pre-detection

Table 1 Comparison of two segmentation results

Original images	Crabcut results	Results of this algorithm

Table 2 Average recall ratio and precision ratio of new algorithm and classical algorithm

Algorithm	Recall ratio	Precision ratio
Crabcut algorithm	0.911	0.877
New algorithm	0.868	0.858

new algorithms. Clothing image segmentation algorithm based on pre-detection was proposed, algorithm was improved based on classical algorithm Crabcut, and original manual operation was replaced by location box automatically generated by algorithm. In this paper, algorithms were described in detail, experiments were done on 1500 images, and the experiment results were verified with the two authoritative assessment criteria of recall ratio and precision ratio. The results show that although it falls behind classical algorithm in accuracy, its time efficiency is greatly improved, and image batch processing can be basically realized.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Shoudong, H., Yong, Z., Wenbing, T., & Nong, S. (2011). Gaussian super-pixel based fast image segmentation using graph cuts. *Acta Automatica Sinica*, 37(1), 11–20.
2. Chen, L., Fengxia, L., & Yan, Z. (2009). An interactive object cutout algorithm based on graph cut and generalized shape prior. *Journal of Computer-Aided Design & Computer Graphics*, 21(12), 1753–1760.

3. Peng, T., Lin, G., & Peng, S. (2009). Infrared target extraction algorithm based on dynamic shape. *Journal of Optoelectronics Laser*, 20(8), 1049–1052.
4. Xiuli, Ma., & Licheng, J. (2008). SAR image segmentation based on watershed and spectral clustering. *Journal of Infrared and Millimeter Waves*, 27(6), 452–456.
5. Yang, C. W., Lu, Y. H., & Hwang, I. S. (2013). Imaging surface nanobubbles at graphite-water interfaces with different atomic force microscopy modes. *Journal of Physics Condensed Matter*, 25(18), 184010.
6. Liu, F., Dai, Q., Shi, X. B., & Liu, J. L. (2012). Fast infrared pedestrian image segment algorithm using MRF based on super-pixel. *Computer Simulation*, 29(10), 26–305.
7. Wang, Y., Wang, H., Bi, S., & Guo, B. (2015). Automatic morphological characterization of nanobubbles with a novel image segmentation method and its application in the study of nanobubble coalescence. *Beilstein Journal of Nanotechnology*, 6(1), 952.
8. Jia, L., & Hongqi, W. (2012). An interactive image segmentation method based on graph cuts. *Journal of Electronics and Information Technology*, 29(4), 1420–1424.
9. Wenbing, T., & Hai, J. (2007). A new image threshold segmentation method based on spectral graph theory. *Chinese Journal of Computers*, 1(1), 1-605–I-608.
10. Walczyk, W., & Schönherr, H. (2013). Closer look at the effect of AFM imaging conditions on the apparent dimensions of surface nanobubbles. *Langmuir the ACS Journal of Surfaces & Colloids*, 29(2), 620–632.
11. Li, F., Peng, J., & Zheng, X. (2004). Object-based and semantic image segmentation using MRF. *EURASIP Journal on Advances in Signal Processing*, 6, 1–8.
12. Caron, L. C., Filliat, D., & Geppert, A. (2014). Neural network fusion of color, depth and location for object instance recognition on a mobile robot. *European Conference on Computer Vision*, 8927(2), 791–805.
13. Karadağ, Ö. Ö., & Vural, F. T. Y. (2013). *MRF based image segmentation augmented with domain specific information* (Vol. 8157, pp. 61–70). Berlin: Springer.
14. Bi, Y., Qiu, T., Li, X., & Guo, Y. (2004). Automatic image segmentation based on a simplified pulse coupled neural network. *International Symposium on Neural Networks*, 3174, 405–410.
15. Arteagasalas, J. M., Zuzan, H., Langdon, W. B., Upton, G. J. G., & Harrison, A. P. (2008). An overview of image-processing methods for Affymetrix GeneChips. *Briefings in Bioinformatics*, 9(1), 25.
16. Chang, F. L., Liu, J., & Qiao, Y. Z. (2005). Self-adaptive threshold segmentation for color image using two-dimensional entropy method based on genetic algorithm. *Control & Decision*, 20(6), 674–678.
17. Soria-Frisch, A. (2006). Unsupervised construction of fuzzy measures through self-organizing feature maps and its application in color image segmentation. *International Journal of Approximate Reasoning*, 41(1), 23–42.
18. Wang, Y., Wang, H., Bi, S., & Guo, B. (2015). Automatic morphological characterization of nanobubbles with a novel image segmentation method and its application in the study of nanobubble coalescence. *Beilstein Journal of Nanotechnology*, 6(1), 952.
19. Ma, T., & Latecki, L. J. (2013). Graph transduction learning with connectivity constraints with application to multiple foreground cosegmentation. *Computer Vision & Pattern Recognition*, 9(4), 1955–1962.
20. Du, Y., Li, F., & Liu, R. (2015). Fast interactive image segmentation using bipartite graph based random walk with restart. In: *Pacific-rim symposium on image & video technology* (pp. 344–354).

Lei Lei Deng born in 1979, member of CPC, Associate Professor, master tutor, teacher of School of Information Technology, Jilin Agricultural University. She presided over the project of Jilin Provincial Science and Technology Department and the scientific research project of the Education Department of Jilin Province, and presided over the teaching and research projects of Education in Jilin Province. Won the second prize of provincial teaching achievement, second prize of education technology achievements, second prize of education scientific achievements, second and third prize of higher education society. I have completed more than 30 essays on provincial excellent courses. Published 1 department, edited 2 textbooks.