

# Intelligent Image Segmentation and Synthesis

Qizhi Xu

*Beihang University, Beijing, China*

**ABSTRACT:** With the rapid development of the e-commerce, shopping online has gradually become one of the most important way of shopping in our life. But there are still some problems, such as size, color, quality and other issues. We can always find some comments that some clothes have color difference. The consumers shopping online can only see the picture from the Internet, without experiencing the goods in the picture by themselves. So many people often find the goods are different from their expectation after accepting them, which results in returning and other problems. To solve this problem, this study uses the MATLAB language and the knowledge of image segmentation, smoothing and the adjustment of brightness and contrast to design of the application. The function of this application is to automatically identify the central object in the original image, using segmentation method based on the edge to segment the image and cover the segmented objects into the new background. When the image is merged, the edge of the image is smoothed by low-pass filtering to adjust the brightness and the exposure of the image to achieve the best results. So the image can reflect the effect of the object under different scenarios, which can make consumers see the tie-in result with the goods. The application can be used in virtual try-on, which can combine various kinds of dress online with their own photos. It gives consumers more intuitive understanding of the goods before buying them, so that shopping will get more scientific.

**KEYWORD:** image segmentation; image smoothing; e-commerce

## 1 INTRODUCTION

With the development of electronic technology, computer software and hardware technology, the image processing technology and the technology related with human vision, the image processing technology of machine vision has been developed rapidly. Image processing technology has been widely used in many fields, such as manufacturing, industry, medical treatment and so on. In the industrial field, the image processing technology of machine vision can replace the manual test method whose detection has slow speed and not enough testing accuracy. In the medical field, the application of image processing is quite extensive. It can help doctors to analyze and detect the size of the body's pathogens and the pathogens, so they can judge whether the human body is different from the normal situation. However, image segmentation technology is an indispensable part of image processing technology. People are always interested in a certain part of the image, which we call it the target image and we call the other part of the image the background. That is the first step of image analysis. Therefore, the image segmentation technology is the most important part of the whole image processing.

With the rapid development of the e-commerce, shopping online has gradually become one of the most important way of shopping in our life. But there are still some problems, such as size, color, quality and other issues. We can always find some comments that some clothes have color difference. The consumers shopping online can only see the picture from the Internet, without experiencing the goods in the picture by themselves. So many people often find the goods are different from their expectation after accepting them, which results in returning and other problems. To solve this problem, I start my study, intelligent image segmentation and synthesis. This study can make consumers see the tie-in result with the goods, so that it can reduce the failure rate of the transaction and enhance the consumer's shopping experience.

Scott T. Acton proposed a method seeking thresholds for segmentation of grayscale in his study (Sang, Q. et al, 2016). And Boykov, Y. & Funkalea, G. proposed an efficient image segmentation technology which is suitable for a wide range of areas: Despite its simplicity, this application epitomizes the best features of combinatorial graph cuts methods in vision: global optima, practical efficiency, numerical robustness, ability to fuse a wide range of visual cues and constraints, unrestricted topological

properties of segments, and applicability to N-D problems (Boykov, Y. et al, 2006). Ramponi, G. offered the method for smoothing treatment in his study (Ramponi, G. 1996). What's more, in Kitchen, L.'s study (Kitchen, L. 1983), although the field was different, it also gave me some help to solve my problems. What gives me great help is Katuyuki Kozima & Suguru Uchida's study (Katuyuki Kozima et al, 1982). This paper led me to a great solution to solve one of the functions of this study. Not only in the technical level, image processing has a lot of research in the application, Masood, S. did the research in medical science (Masood, S. et al, 2015). And Xing, J.'study is the research in agriculture (Xing, J. et al, 2007).

Learning a lot about image segmentation technology improved, and their related papers in medicine, agriculture, manufacturing and application of industry, I start this study. The main idea of this study is to apply the image processing technology to Internet shopping by using the improved image segmentation technique, which is closely related to life shopping. The function of this study is to cover a major object in a picture to any place on the other, and then adjust the picture after covering, making the picture natural. It can reflect the effect of the object in different scenarios. For example, we can make the product such as clothes separated from the pictures provided by the electricity suppliers and combine the product with the picture provided by the consumers, and then smooth it by low-pass filtering and adjust the brightness and the exposure of the image. Therefore, consumers can have a direct impression of the clothes and judge whether or not the clothes and themselves are well matched. This study can give the consumers great shopping experience.

## 2 METHOD

### 2.1 The system structure

As shown in Figure 1, this study has a total of three major modules. These modules form a complete process, by which the main function of the system is completed. The system accepts two images as input, one of which contains the object to be isolated, and then cover the objects to another picture. Finally, adjust the picture, and output the synthesized results.

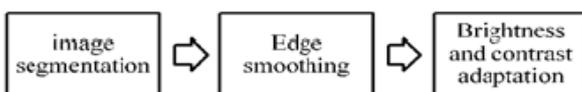


Figure 1. the system structure

## 2.2 Function module design

### 2.2.1 Image segmentation sub module

#### (1) Edge-based segmentation methods

Edge-based segmentation methods can be said to be one of the earliest methods of research. Because the change of pixel gray value in the region edge is often more severe, it tries to solve the problem of image segmentation by detecting the edges of different regions.

The simplest edge detection method is the image filtering method. It makes use of the properties that the pixel values in the adjacent region are not continuous, using first order or two order derivative to detect the edge. In recent years, many new methods are proposed, such as the method based on local image function, the method based on reaction diffusion equation, the method based on boundary curve fitting, the serial boundary search based on graph theory, the method based on deformation model, the multi-scale method and the multi-resolution method. Let's introduce the method of image filtering.

The image filtering method is based on the following properties of the convolution operation:

$$\frac{d}{dx}(f(x) \times g(x)) = \frac{d}{dx}f(x) \times g(x) + f(x) \times \frac{d}{dx}g(x) \quad (1)$$

Image filtering is to use a filter operator and the image for the convolution operation. By the above formula, the first derivative is obtained by the convolution of the filter operator and the image, which is equivalent to the first derivative of the operator and the image. So, as long as the first order or two order derivative of the operator, the image smoothing filter and the smoothing of the image after the first or two order derivative can complete fast. The template of several common operators is shown in figure 2:

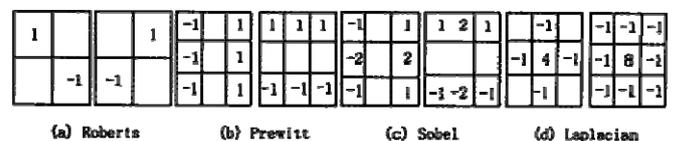


Figure 2. the template of several common operators

The key problem of this method is the design of the filter operator. Canny John has a deep research on this problem, and the three principles of the optimal filter are presented: Good detection, that is, to identify the larger probability of the real edge points to a smaller probability of the non-edge points as edge points; The accurate location of edge points; only one response to each edge point. He put these three principles into a mathematical form, and then use the variational method to find the optimal filter. Canny considers that the first derivative of Gauss's function is a good approximation of the

optimal filter obtained by him. In general, the Laplacian operator is used to find the two order derivative of Gauss's function to get the LOG filter operator. The operator is first proposed by Marr, the founder of computer vision (Kass, M. et al, 1988).

### 2.2.2 Edge smoothing sub module

#### (1) frequency division

For an image, the energy is mainly concentrated in the low frequency component, and the noise and image detail information is mainly concentrated in the high frequency component. In the process of image processing, the two components are divided into different operation and processing, which can achieve better treatment effect. The aim of frequency division is to separate the high frequency and low frequency components of the input image. When dividing the frequency, both the high pass filter and the low pass filter can be selected, and the purpose of frequency division can be achieved. We use the low pass filter in the development process. In many low pass filter, Butterworth low pass filter without "ringing" phenomenon, and can improve the details of the definition of the image.

Butterworth low-pass filter transfer function:

$$H(u, v) = \frac{1}{1 + \left[ \frac{D(u, v)}{D_0} \right]^{2n}} \quad (2)$$

In the formula,  $D_0$  works as the cut-off frequency,

$$D(u, v) = \sqrt{(u)^2 + (v)^2} \quad (3)$$

in the  $D(u, v) = D_0$ ,  $H = \frac{1}{2} H_{\max}$ ;  $n$  is for the order, taking a positive integer, used to control the speed of decay.

For an input image  $f(x, y)$ , it can be decomposed into high frequency component  $f_h(x, y)$  and low frequency component  $f_l(x, y)$ . As shown in the following:

$$f(x, y) = f_h(x, y) + f_l(x, y) \quad (4)$$

The image  $f(x, y)$  is passed through the 2 order Butterworth low-pass filter  $g(x, y)$ , which means making  $f(x, y)$  and  $g(x, y)$  for convolution. And we can get the low frequency component of image  $f(x, y)$ , as follows:

$$f_l(x, y) = f(x, y) \times g(x, y) \quad (5)$$

Among the image  $f(x, y)$ ,

$$g(x, y) = \frac{1}{1 + \left[ \frac{D(x, y)}{D_0} \right]^2} \quad (6)$$

( $D_0$  works as the cut-off frequency) minus the low frequency component  $f_l(x, y)$  to get the high frequency component:

$$f_h(x, y) = f(x, y) - f_l(x, y) \quad (7)$$

#### (2) Low frequency component equalization

Image energy is mainly concentrated in the low frequency component  $f_l(x, y)$ , which plays a decisive role in the overall visual effect of the image. So it is necessary to carry on the histogram equalization to this component. Use the traditional algorithm on the low frequency component  $f_l(x, y)$  for histogram equalization. And we can get  $t_l(x, y)$ :

$$t_l(x, y) = HE(f_l(x, y)) \quad (8)$$

Noise exists in the high frequency component. If the histogram equalization is performed directly, the noise will be enhanced, and the image quality will be affected. So, take  $f_h(x, y)$  to a weighted coefficient  $a$  to get  $t_h(x, y)$ :

$$t_h(x, y) = a \times f_h(x, y) \quad (9)$$

The value of  $a$  depends on the specific circumstances: If the image is more serious noise, the value of  $a$  should be less than 1, so as to avoid amplification of noise; if the image is less affected by noise, the value of  $a$  should be greater than 1. The output image  $f(x, y)$  is the sum of the high frequency component  $f_h(x, y)$  and low frequency component  $f_l(x, y)$ :

$$f(x, y) = f_h(x, y) + f_l(x, y) \quad (10)$$

So, the output image  $f(x, y)$  is the linear combination of  $f_h(x, y)$  and  $HE(f_l(x, y))$ :

$$t(x, y) = a \times f_h(x, y) + HE(f_l(x, y)) \quad (11)$$

$t(x, y)$  is the image that it wants to get.

By dividing frequency, the input image is decomposed into high frequency component and low frequency component. Do different operations on the two component and then make the linear combination of them. We can get the final enhanced image which has high image clarity, rich details of the information and obvious reinforcing effect (Zhang, T. W. et al, 2012).

### 2.2.3 Brightness and contrast adaptation sub module

#### (1) Image brightness and contrast adjustment

As for image brightness and contrast adjustment: taking 24 color images as an example, each color can be used 0-255 for a total of 256 kinds of depth. If we draw it in a two-dimensional coordinate system, it is a straight line. For example, we will pixel color deep as abscissa, the output color deep paintings as ordinate, just a pass through the origin (0, 0) of the 45-degree diagonal.

Then it's easy to write its linear equation:

$$Out = In \times 1 \quad (12)$$

Coefficient 1 is the concept of contrast. If a line is added with an offset to B, the linear equation becomes:

$$Out = In \times 1 + (ab) \quad (13)$$

Offset (ab) is the increment of brightness.

It is easy to see that it satisfies a linear equation:

$$Y = A \times X + B \quad (14)$$

However, the situation we have to deal with here is slightly different. In image processing, contrast and brightness should be treated separately. Because we used to put the gray(127,127) as a center point, we mapped the point (127,127) to the origin of the coordinate system. Then we need to modify the original linear formula:

$$Y = (X - 127) \times A + B \quad (15)$$

A represents the contrast and B represents the brightness increment. Let's verify it: as long as the brightness increment B=0, no matter how to change the contrast A, the straight line always passes through the center point (127,127), which is to say when we change the contrast, the brightness is not changed. From this, we can deduce the color contrast brightness calculation formula:

$$NewRed = (OldRed - 127) \times A + 127 + B \quad (16)$$

$$NewGreen = (OldGreen - 127) \times A + 127 + B \quad (17)$$

$$NewBlue = (OldBlue - 127) \times A + 127 + B \quad (18)$$

### 3 RESULTS

The effect of image segmentation and synthesis is shown in Figure 3 to 7:

Figure 3 is the original, first we will image gray in figure 5. Then use image segmentation algorithm to segment the goods. And then restore the gray image to color image, the background set to black. In the image synthesis process, after the segmentation effect figure 4 is superimposed on the background figure 6, we take smooth processing and contrast and

brightness adjustment on it, during this process we can use MATLAB to zoom in and out of the original image and select the position. Figure 7 is the synthesis effect.



Figure 3. the original goods      Figure 4. The background



Figure 5. the gray image      Figure 6. the effect of segmentation



Figure 7. the effect of synthesis

### 4 CONCLUSION

The study uses image segmentation, edge smoothing and other techniques to realize the automatic recognition of the central object in the original image and cover it on the new background. When the image is merged, the edge of the image is smoothed and the brightness of the whole image is adjusted, which aims to achieve the best effect. The application can be used in virtual try-on, which can combine various kinds of dress online with their own photos. It gives consumers more intuitive understanding of the goods before buying them, so that shopping will get more scientific. Secondly, what I am doing is the object and the background color has the obvious color difference, and asks the picture not to be very complex. I think it should be applied in a broader and more complex picture. And what I do is on a two-dimensional plane. If it can be extended to three-dimensional level in the future, we can construct a three-dimensional image out through the different directions of the image of the product, so that users can feel the full range and the effect will be much better and the significance will be greater. And I think that this application is not only applied to clothing and other goods, but also can be used in the field of home decorating.

## REFERENCES

- Boykov, Y., & Funka-Lea, G. (2006). Graph cuts and efficient n-d image segmentation. *International Journal of Computer Vision*, 70(2), 109-131.
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: active contour models. *International Journal of Computer Vision*, 1(4), 321-331.
- Katuyuki Kozima & Suguru Uchida. (1982). Noise smoothing in image improvement. *Journal of Modern Optics*, volume 29(4), 371-376.
- Kitchen, L., Pietikainen, M., Rosenfeld, A., & Wang, C. Y. (1983). Multispectral image smoothing guided by global distribution of pixel values. *IEEE Transactions on Systems Man & Cybernetics*, smc-13(SMC-13), 626-631.
- Masood, S., Sharif, M., Masood, A., Yasmin, M., & Raza, M. (2015). A survey on medical image segmentation. *Current Medical Imaging Reviews*, 11(1), 3-14.
- Ramponi, G. (1996). The rational filter for image smoothing. *IEEE Signal Processing Letters*, 3(3), 63-65.
- Sang, Q., Lin, Z., & Acton, S. T. (2016). Learning automata for image segmentation ☆. *Pattern Recognition Letters*, 74, 46-52.
- Xing, J., Saeys, W., & Baerdemaeker, J. D. (2007). Combination of chemometric tools and image processing for bruise detection on apples. *Computers & Electronics in Agriculture*, 56(1), 1-13.
- Zhang, T. W., Chen, H., & Wang, L. (2012). Design of digital butterworth low-pass filter based on matlab. *Journal of Anshan Normal University*.