

HYBRID COLOR IMAGE SEGMENTATION BASED FULLY AUTOMATIC CHROMA- KEYING SYSTEM WITH CLUTTERED BACKGROUND

Li Shijin, Zhu Yuelong, Yang Qian and Liu Zhe

*School of Computers and Information Engineering, HOHAI University, Nanjing, P. R. China,
210098*

Abstract: In this paper, we present a fully automatic digital chroma-keying system, which is based on the integration of color image segmentation algorithm and improved alpha estimation technique. Chroma-keying is a critical technology in virtual studio system. When used with cluttered background, it calls for much intelligence. According to the characteristics of frame images in the target application, a hybrid color image segmentation algorithm is put forward, which makes good use of both chromatic and luminance information. Then, refinement measures are further taken to deal with the color distribution in the neighborhood of the boundary through modified Ruzon-Tomasi alpha estimation algorithm. In contrast to the previously reported methods, our system needs no human interaction in the whole procedure. Experimental results on China sports lottery TV programs show that the proposed fully automatic keying system is viable and can be applied to the real program post production process of TV stations.

Key words: Virtual studio, color image segmentation, chroma-keying, alpha estimation.

1. INTRODUCTION

Chroma-keying is a critical technology in virtual studio system [1]. By using this technology, actors in the studio, who have performed in front of a blue screen, are shown on the TV with new background, which is composed by a computer using graphics or animations. Because of the composition of

the computer, the background can be rapidly changed, which not only improves the efficiency of the program production and the availability of the studio, but also makes the background much more splendid and attractive than the usual physical one.

Traditional chroma-keying adopts blue-screen technology [2], which means that the actor is recorded in front of a solid color background (blue, green etc). In recent years, chroma-keying in a hybrid color or a cluttered background becomes a hotspot in the research community [3-6](in some technical papers, it is also called matting [2,3]). When used with cluttered background, it calls for much intelligence. All these methods are based on the following basic composition equation:

$$C = \alpha F + (1 - \alpha) B$$

Where C represents the output signal of the composition of foreground and background signals, F is the foreground signal, B denotes the background signal, and α is the key signal, which is also known as alpha. So in some papers, it is also called alpha estimation technique [4].

For the unlimited color distribution of the background, alpha estimation algorithm loosens the strict requirement (the solid color background) in the blue-screen scenario, and the only thing needed is the boundary. The common features of all the current algorithms based on this idea are that first manually segment the image into three parts –the foreground, the background and the unknown region (i.e. the contour of the foreground), and then estimate values of foreground colors, background colors and the alpha respectively.

With this kind of algorithms, it will cost a lot of money and human resources to manually deal with every frame in the video, furthermore, too many human interactions may lead to the poor quality of the final result. So more effective methods are needed to improve this situation. If the actor keeps staying in the same place when broadcasting in the air, we can use image segmentation to pick up the rough contour of the anchor person automatically, which can reduce or even eliminate the human interactions.

The basic idea of the new algorithm proposed in this paper is to firstly exert color image segmentation to outline the rough contour of the foreground object, and then carefully refine the boundary using modified Ruzon-Tomasi alpha estimation algorithm [4]. Experimental results on China sports lottery TV programs show that the proposed fully automatic keying system integrates the advantages of the image segmentation and alpha estimation algorithms and greatly extend the application scenario of virtual studio keying system.

2. COLOR IMAGE SEGMENTATION

Along with the development of color imaging and its relevant techniques in recent years, color image segmentation techniques become a hot research topic in the image processing domain. Since the color image has more useful information (hue, saturation etc) than grayscale (only has brightness information), we can use statistical features of such information to segment it. Because the conclusions drawn from different color spaces are so much different that sometimes they are even contradictory, the selection of color spaces is not important, considering their unsuitability to all types of image. In this paper we choose the HSI color space for its similarity to the characteristics of human vision [7].

2.1 Multi-channel information fusion based hybrid color image segmentation

For the application in this paper, firstly, we have to refine the contour of the hostess in China sports lottery TV programs with a cluttered background, and then insert the virtual background to produce the virtual synthesis video shown on different TV stations. For the usually adopted H information can't separate the foreground object from the background effectively, the segmentation in this paper is mainly based on the S and I information in HSI color space. Figure.1 shows an original frame image in our application.



Figure 1. One frame color image from the video and its corresponding HSI components

In figure 1, both the original color image of the hostess and the corresponding HSI component pictures illustrate that S and I information are useful in our image segmentation task.

At present, color image segmentation algorithms can be classified into three sorts of methods [8]: fusion of chromatic based on gray image segmentation; multidimensional gradient thresholding and color vector

analysis. Having adopted the first segmentation methodology in this paper, we employ the optimal edge detection operator, which has been put forward by Shen in [9], to binarize the triple channel images in HSI space before the integration of the result of each part. Before giving the specific realization of our segmentation strategy, a brief introduction will be made about the optimal edge detection operator theory and its recursive implementation.

Based on Gaussian function, D. Marr [10] proposed the LOG operator whose precision of edge detection is varied following the change of the variance of Gaussian function. After carefully analyzing the previous boundary detection filters, upon the symmetric exponential filter, Shen [9] worked out an optimal edge detection operator which can not only effectively reduce the noise by an infinitely large size window, but also makes the central part of the filtering function sharp enough to enhance the boundary location precision.

The symmetric exponential filter has the following form:

$$F(x, y) = C_1^2 \times C_2^{|x|+|y|} \quad (1)$$

where $C_1 = a_0 / (2 - a_0)$; $C_2 = 1 - a_0$, $a_0 \in (0, 1)$. When a_0 is

near 1, $F(x,y)$ becomes narrower, and the anti-noise capability is lower, while the precision of edge localization is higher; when a_0 becomes smaller, the anti-noise capability is higher, but more detailed edges will be lost. Assuming $f(i,j)$ denotes the original image, and it is filtered by the above filters, Shen [9] has proved that the filtering process can be decomposed into the cascading of two-pass filtering in row and column directions respectively and can be implemented by two recursive filters in two passes. Firstly, the filtering is done in row direction:

$$\begin{aligned} f_1(i, j) &= f_1(i, j-1) + a_0 \times [f(i, j) - f_1(i, j-1)] \quad j = 1, 2, \dots, n \\ f_2(i, j) &= f_2(i, j+1) + a_0 \times [f_1(i, j) - f_2(i, j+1)] \quad j = n, n-1, \dots, 1 \end{aligned} \quad (2)$$

Where $f_2(i, j)$ is the output of the filtering in row direction. Then, the filtering in column direction is conducted based on the previous results.

$$\begin{aligned} f_3(i, j) &= f_3(i-1, j) + a_0 \times [f_2(i, j) - f_3(i-1, j)] \quad i = 1, 2, \dots, n \\ f_4(i, j) &= f_4(i+1, j) + a_0 \times [f_3(i, j) - f_4(i+1, j)] \quad i = n, n-1, \dots, 1 \end{aligned} \quad (3)$$

Hence,

$$f_4(i, j) = f(i, j) \times C_1^2 \times C_2^{|i|+|j|} \quad (4)$$

$$f_4(i, j) - f(i, j) \approx \frac{1}{2C_1 \ln C_2} \times \nabla^2 F(x, y) \times f(i, j) \quad (5)$$

And the last output results are the Laplacian image of the original one [9].

In our application scenario, we separate the two-pass filtering and conduct it using different parameters, i.e., the 'a0' in Equation 2) and 3) can take different values, and we use the parameter pair (a0, b0) to characterize

the procedure. So we can configure different parameters to detect different oriental edges in images.

We divide the image into three sections (the section above the head, the section with head and the section with body) according to the image characteristics in this specific application, and choose diverse segmentation strategies in different channels of each section. By the way, through the vertical projection, we can get the positions of the two dividing lines (the up line and the down line), which separate the image into three parts.

Because the brightness in the section above the head is high and the color information is reduced by the glaring for the highlighting in the studio, we segment it in the brightness channel. The edge of head is mostly in vertical direction, so we mainly consider the vertical orientation filtering in Shen optimal edge detector, and in operation we only filter in vertical direction. The parameters take the value of $(a_0, b_0) = (0.0625, 1)$.

In the section with the head, since the color similarity of hair and the background Chinese characters leads to the unapparent distinction of hue, we segment the image in S channel, which is very high in the hostess for their densely make-up. The edge of this part is in mostly horizontal direction, so we mainly consider the horizontal orientation filter in Shen optimal edge detection, and in practice we only filter in horizontal direction. The parameters take the value of $(a_0, b_0) = (0.5, 0.125)$.

As the body section is concerned, for the distinction of colors in hostess and background is apparent and the background is nearly the solid yellow, we adopt the classical Otsu histogram thresholding [11] in H channel to accomplish the segmentation.

After the segmentation of these three parts individually, we compose them into a whole image. The results are given in Fig.2 (a) ~ (d), and the 2(d) image is the result of initial integration.

2.2 Post-processing

Getting the binary image after segmentation, we go on with the post processing, which includes the removal of the small noise region, the combination of the segmented regions belonging to the hostess and the brim fairing using the mathematical morphology operations.

Firstly, we take the edge tracing in the binary image after the segmentation. Among the connected regions in the image, the area containing the body of hostess is the largest one, so we set the lower limit area of the bound rectangles, which surround the connect region, 100 to eliminate the small connect regions. After computing the area of the largest bordering rectangle, in order to remove the noise interference, we eliminate

the connected regions far from the center of the rectangle, and incorporate the connected region near the contour. Then trace this largest region and go on with mathematical morphology processing to smooth the edge and filling the holes. Fig. 3 illustrates the aforementioned procedure.

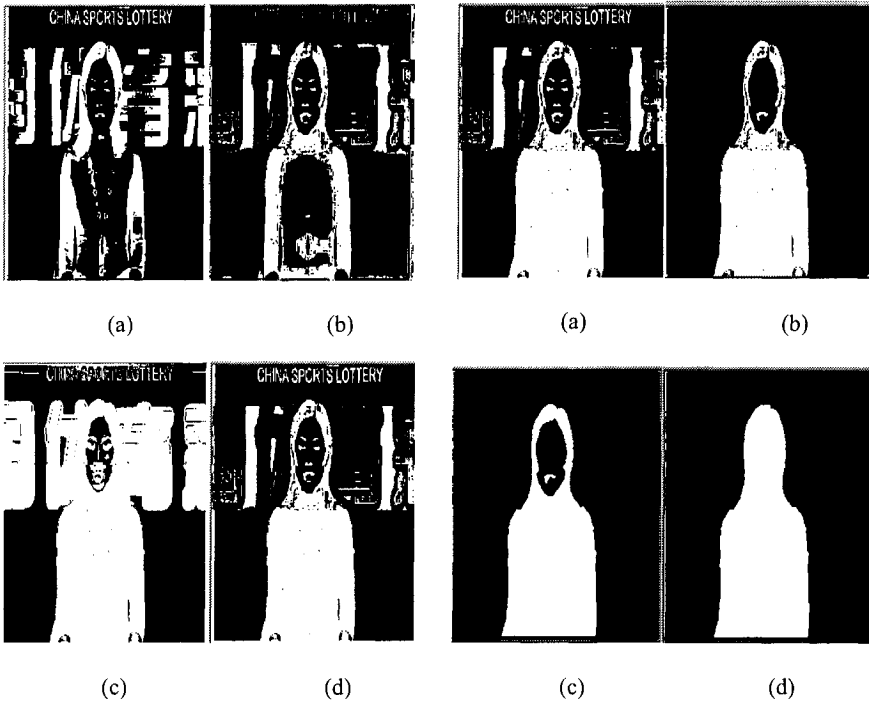


Figure 2. Multi-channel image segmentation and the initial integration result

Figure3. The result of post-processing

Compared with the original image, result of single frame image segmentation is relatively satisfying. Although the foreground object is almost completely extracted, the zigzag and unsmooth hair brim is unable to show the nicety of the hair, and some ends of the hair around the shoulder still can not be refined. If the segmented frames are composed into one AVI file, the resulting video makes the audiences visually uncomfortable. To improve this situation, we employ the alpha estimation technique after the traditional image segmentation process is conducted.

3. ALPHA ESTIMATION

Alpha channel or α channel in short, determines the transparency of each pixel besides tricolor channels in digital image. α channel is also called opacity or coverage percentage. Put forward by Porter and Duff in 1984[12], the conception of α channel is that one pixel can be described in four channels (r, g, b, α) and the distributions of colors (red, green and blue) in each pixel can be measured by the product of α and the value of each color.

In the real world, the boundary between objects area is not very sharp, but a gradual transition. Since each boundary pixel receives the light from many objects and the color penetration, the exact segmentation could not be achieved just depending on the use of image segmentation or boundary extraction. To solve the problem alpha estimation technique emerges.

Ruzon-Tomasi alpha estimation algorithm [4] is employed in this paper. Firstly, we automatically segment the boundary region into many sub-regions, then build a box to surround each sub-region containing those pixels already known in the foreground object and background. The surrounding pixels in the foreground are from the sampling function of $P(F)$, while the pixels in the background are with the sampling function of $P(B)$. All these pixels are divided into different color clusters – each has a Gaussian distribution in the color space, so we can regard the foreground color distribution as a mixture of Gaussian distributions. Like the foreground, all the background pixels can also be divided into different clusters according to their colors, and each cluster in the background is related to one foreground color cluster.

After building the color cluster distribution network, which map the foreground cluster to the background one, we treat the observed color C coming from an intermediate distribution $P(C)$, which is located between foreground and background. The intermediate distribution is also defined as a mixture of Gaussian distributions, each of which has a new expectation located at a certain position in the diagonal, connecting a pair of foreground and background color clusters. An optimum alpha is claimed with the intermediate distribution, which makes the observed color have the maximum probability. The foreground and background colors are estimated as a post-process using a weighted sum of the foreground and background cluster means. These values are then perturbed, so that a line segment with endpoints at F and B , goes through C and also satisfies the basic composition equation. The detailed algorithms can be found in [4].

In this paper, we make some modifications to the color quantization in the algorithm in [4]. Because the foreground and the background are manually picked out in [4], the color quantization is global, i.e. the colors of

all the pixels in the image are quantized. That may lead to some disadvantages. Since there may be millions of colors in an image, the to-be-segmented region in the image may only have very few color samples and they may be ignored by the global quantization algorithm, which will lead to the mismatch of the color codebook and the actual color in the boundary region. In this paper, because the preceding image segmentation can tell the sub-region of the boundary, only the pixels nearby the boundary are quantized and the precision and color fidelity are further guaranteed. In the implementation, after the four times dilatations and erosions of the segmentation outcome, we “XOR” the two resulting images to get the boundary region needed in alpha estimation algorithm. Also because the color quantification only deals with the original image in the boundary region above, the precision of alpha estimation algorithm has been further improved.

4. **EXPERIMENTAL RESULTS AND DISCUSSIONS**

Proven by a great deal of experiments on many samples, the automatic chroma-keying method which integrates color image segmentation and alpha-estimation is able to pick-up the fine and translucent complex region in the foreground from the background without the limitation that the background must be in purely colors. The system proposed in this paper can achieve the high quality of refined image and greatly extend the application scenarios of virtual studio. The procedures of the algorithm, and the corresponding results are presented in Fig.4.

The experiments demonstrate the feasibility of using automatic chroma-keying method which is based upon the color segmentation and alpha-estimation in the field of video keying technology, and they also demonstrate that the apparent improvement of the object refinement upon the simply image segmentation algorithm is meaningful to extend the application scenario of virtual studio. Furthermore, the proposed technology can be applied in other similar situations, such as the weather forecast program production. However, there are still some limitations in the system proposed in this paper. In various foregrounds and backgrounds, we have to adopt different segmentation algorithms without a uniform solution. Snake or active contour models [13] are currently being considered as an alternative. How to apply the technique proposed in this paper to more complex situations and how to further improve the capability of the algorithm are the main research directions in the future.

ACKNOWLEDGEMENTS

We are grateful to Mr. Wang Hui for the experimental videos he provided and the inspiring discussions with him are also appreciated.

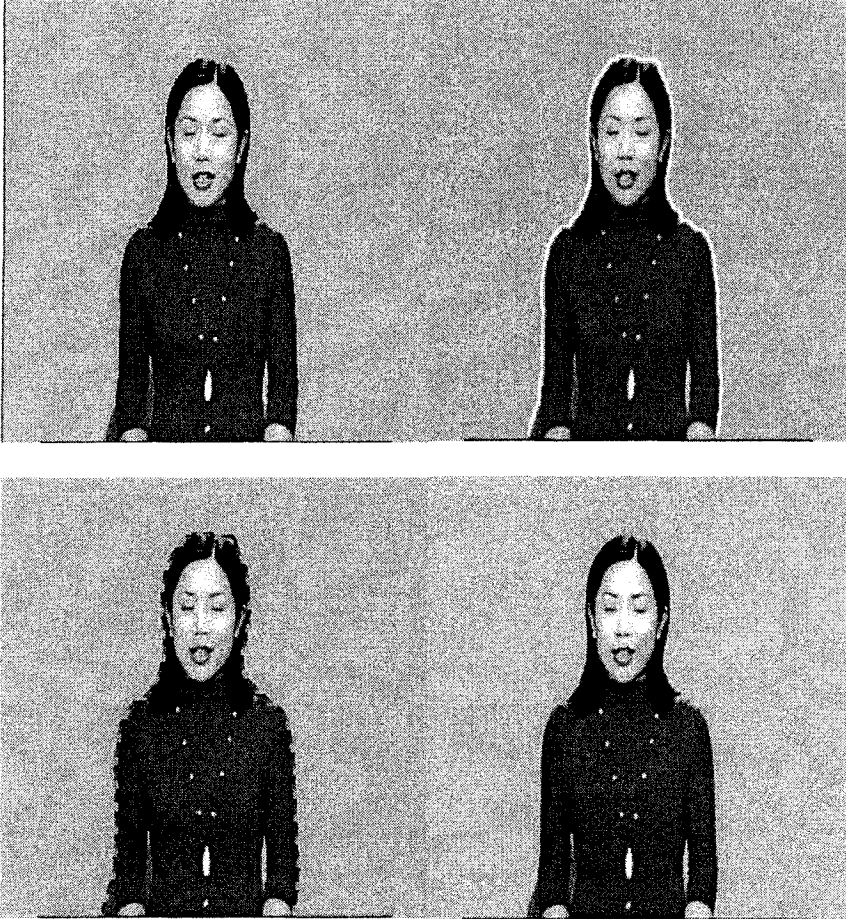


Figure 4. Alpha estimation procedure illustration and the last result

REFERENCES

1. Luo Yuhua, Virtual studio system: an overview, China Journal of Image and Graphics, 1996, Vol. 1, No.3, pp220-224.
2. Smith A. and Blinn J. Blue screen matting. In Proc. SIGGRAPH'96, pp.259-268, 1996.

3. Chuang Y., Curless B., and Salesin D., et al, A bayesian approach to digital matting, in Proceedings of IEEE Conference on CVPR 2001, Vol.2, pp.264-271.
4. Ruzon M. and Tomasi C., Alpha estimation in natural images, in Proceedings of IEEE Conference on CVPR 2000, Vol.1, pp. 18-25.
5. Mitsunaga T., Yokoyama T. and Totsuka T., AutoKey: Human assisted key extraction, in Proceedings of SIGGRAPH'95, pp. 265–272, 1995.
6. Qian R. J. and Sezan M. I. Video background replacement without a blue screen. in Proceedings of ICIP 1999, pp. 143–146, October, 1999.
7. Wei Baogang, Li Xiangyang, Lu Dongming, Survey of the segmentation of color images, China Journal of Computer Sciences, 26(4): 59-62.
8. Ruzon M., Early vision using distributions, Ph.D thesis, Computer Science Dept., Stanford Univ., Stanford, Calif., Apr. 2000.
9. Shen J., The optimal linear edge detection operator, China Journal of Pattern recognition and Artificial intelligence, 1987, Vol. (1): 86-103.
10. Marr D. and Hildreth E., Theory of Edge detection, Proc. Royal Soc. London, Vol. B207, pp.187-217, 1980.
11. Otsu N., A threshold selection method from gray-level histograms, IEEE Trans. Systems Man Cybernetics, SMC-9, 62-66, 1979.
12. Porter T. and Duff T. Compositing digital images. In *SIGGRAPH 1984*, pages 253–259, July 1984.
13. Kass, M., Witkin, A., and Terzopoulos, D., “Snakes: active contour models,” International Journal of Computer Vision, Vol.1, No. 4, pp. 321-331, 1987.