# Region Extraction with Chromakey using Striped Backgrounds

Atsushi Yamashita*, Toru Kaneko*, Shinya Matsushita* and Kenjiro T. Miura*
Department of Mechanical Engineering
Shizuoka University

## Abstract

In this paper, we propose a new region extraction method with chromakey using a two-tone striped background. The conventional chromakey uses the unicolored blue background, and has a problem that one's clothes are regarded as the background if their colors are same. Therefore, we utilize the adjacency condition between two-tone striped areas on the background, and extract the foreground regions whose colors are same with the background. The validity of proposed method is shown through experiments.

## 1 Introduction

The composition of images[1] has significant consequences to creative designs such as cinema films, magazine covers, promotion videos, and so on. This technique can combine actors or actresses in a studio and scenery that is taken in another place. The foreground must be segmented from the background when compositing images, and there are a lot of studies about image segmentation[2, 3]. The snake[4] is a popular technique based on the edge detection. However, the automatic methods for the arbitrary background have not been developed to a practical level. Additionally, robust methods are needed especially for a live program on TV[5, 6].

Therefore, an image composition method called chromakey in which human objects are segmented from a uniform color background and superimposed with another background has been proposed. The conventional chromakey technique uses background whose color is blue or green, and extracts the region in images whose colors are not same with the background. Although this can be successfully used, several problems associated with this method still need a solution. Shimoda et al. deal with the camera motion that the background image alters accordingly as the foreground image is altered by panning, tilting, zooming and focusing operations of the camera[7]. Mitsunaga et al. propose the human assisted segmentation method for an arbitrary background that leaves little work to operators[8]. Debevec et al. consider the lighting condition and achieve realistic composites between a foreground actor and a background environment with a large sphere of inward-pointing light focused on the actor[9]. One of the most critical and unsolved problems is that if one's clothes and the background have the same color, the former is regarded as the latter (Figure 1(a)). Smith and Blinn propose the blue screen matting method that allows the foreground object to be shot against two backing colors[10]. However, this multi-background technique cannot be used for live actors or moving object because of the requirement for repeatability.

*Address: 3-5-1 Johoku, Hamamatsu-shi, Shizuoka 432-8565, Japan. E-mail: yamashita@ieee.org

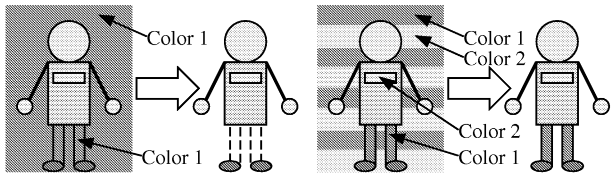


Figure 1: Region extraction with chromakey. (a) Unicolor background. (b) Striped background.

In this paper, we propose a new method that makes it possible to segment objects from a background precisely even if objects have the same color as the background, by using two-tone striped background (Figure 1(b)). Our method does not need two shots, and threshold values for a color extraction can also be adjusted automatically to improve the robustness against the change of the lighting conditions.

## 2 Region Extraction

The foreground cannot be extracted by only using the color segmentation approach when the colors of foreground objects are same with those of the striped background. Therefore, the boundaries where the color of the striped background changes are utilized for detecting the foreground region whose color is same as the background.

The procedure of our method consists of 4 steps: 1) background color extraction, 2) striped region extraction, 3) foreground extraction, and 4) image composition.

### 2.1 Background Color Extraction

After obtaining an image with the two-tone striped background, a median filter is used for noise removal. The difficulty of setting threshold values for backgrounds' colors can be reduced to remove noises in the background.

Candidate regions of the background are extracted by using a color space approach. Let $C_1$ and $C_2$ be the colors of the two-tone background in acquired images from the camera, and $R_1$ and $R_2$ be the region whose color is $C_1$ and $C_2$, respectively. $R_i$ $(i = 1, 2)$ is defined as:

$$R_i = \{(x, y) | F(x, y) \in C_i\}, \qquad (1)$$

where $F(x, y)$ is the color of an image at a pixel $(x, y)$.

It occurs frequently that the color of the background is similar to that of the foreground object, though they can be divided in the RGB color space. For example, the color of the blue background is very similar to that

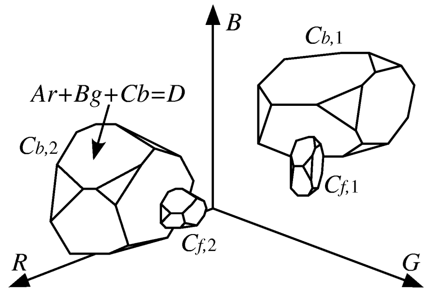

Figure 2: RGB color space

of blue denim pants or jackets that a number of people usually wear. Therefore, to set $C_i$ systematically, we define $C_{f,i}$ $(i = 1, 2)$ explicitly as the foreground color that is similar to that of the background. Accordingly, $C_i$ is expressed as follows:

$$C_i \quad = \quad C_{b,i} \cap \overline{C_{f,i}}, \tag{2}$$

where $C_{b,i}$ $(i = 1, 2)$ is the color of the background. The human operator can decide the threshold values easily to configure both $C_{b,i}$ and $C_{f,i}$.

$C_{b,i}$ and $C_{f,i}$ $(i = 1, 2)$ are extracted by thresholding process in the RGB color space, respectively. These colors are expressed as the region surrounded by the planes in the RGB color space (Figure 2).

$$C_{b,i} = \{F | F \in \bigcap_{j=1}^{m_i} P_{i,j}\}, \tag{3}$$

$$C_{f,i} = \{F | F \in \bigcap_{j=1}^{n_i} Q_{i,j}\}, \tag{4}$$

where $P_{i,j} = \{(r, g, b) | a_{i,j} \leq b_{i,j}r + c_{i,j}g + d_{i,j}b \leq e_{i,j}\}$, $Q_{i,j} = \{(r, g, b) | p_{i,j} \leq q_{i,j}r + r_{i,j}g + s_{i,j}b \leq t_{i,j}\}$, and $m_i$ and $n_i$ $(i = 1, 2)$ are the total numbers of the planes, respectively.

$b_{i,j}$, $c_{i,j}$, $d_{i,j}$, $q_{i,j}$, $r_{i,j}$, $s_{i,j}$ are decided by the human operator in advance by using the images in which only the striped background exists. $a_{i,j}$, $e_{i,j}$, $p_{i,j}$, $t_{i,j}$ can be adjusted the lighting condition. However, the change of the lighting condition cannot be detected exactly from the image. Therefore, the average brightness of two regions in the background is utilized. From the result of the rough grouping of $C_1$ and $C_2$ in the HSI color space, the average values of $(r, g, b)$ in $R_1$ and $R_2$ can be calculated respectively except in the middle area of the image where the probability that the foreground objects exist is very high. To compare the initial values of $(r, g, b)$ with the present values, $a_{i,j}$, $e_{i,j}$, $p_{i,j}$, $t_{i,j}$ changes adequately.

From these procedure, $R_1$ and $R_2$ are extracted, respectively (Figure 3(a)(b)).

## 2.2 Striped Region Extraction

The background mainly consists of the three regions: $R_1$, $R_2$, and $R_3$ which is the intermediate region between $R_1$ and $R_2$ (Figure 3(b)). It is difficult to extract $R_3$ by a color space approach because the color of $R_3$ may be a composite of $C_1$ and $C_2$ and not either $C_1$ or $C_2$. Let $C_3$ be the color defined as follows:

$$C_3 = \{F | F \notin (C_1 \cup C_2)\}, \tag{5}$$

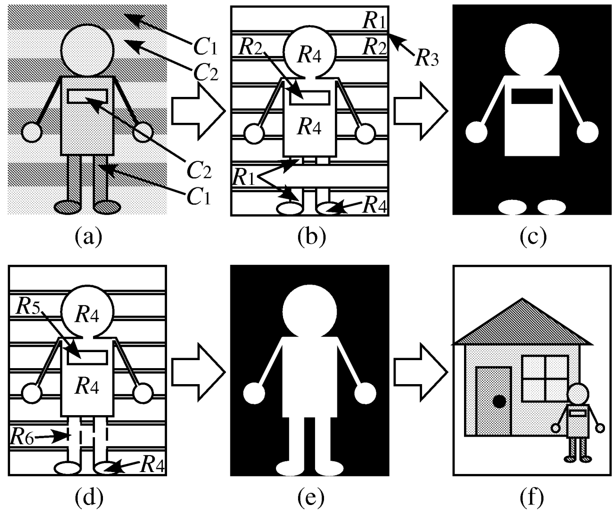

Figure 3: Region extraction. (a) Original image. (b) Region segmentation without using the information of the striped area. (c) Foreground extraction without using the information of the striped area. (d) Region segmentation by using the information of the striped area. (e) Foreground extraction by using the information of the striped area. (f) Image composition.

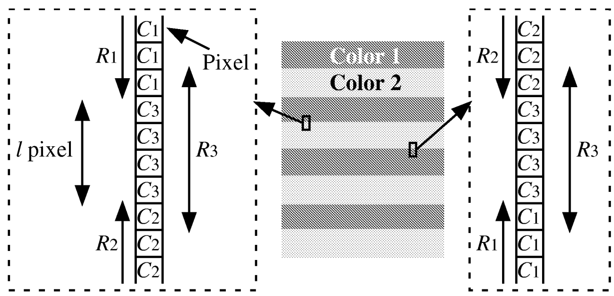

Figure 4: Detection of $R_3$ region

$C_3$ is the color of $R_3$ or that of the foreground. The difference between $R_3$ and the foreground is that $R_3$ contacts with both $R_1$ and $R_2$ according to the following rule. In the case of horizontal stripe, the color of the upper and lower region of $R_3$ differs from each other. Therefore, $R_3$ is extracted from adjacency conditions expressed as follows:

$$\begin{aligned} R_3 \quad = \quad & \{(x, y), (x, y+1), \ldots, (x, y+l+1) | \\ & (x, y+1) \in C_3, \ldots, (x, y+l) \in C_3, \\ & ((F(x, y) \in C_1, F(x, y+l+1) \in C_2) or \\ & (F(x, y) \in C_2, F(x, y+l+1) \in C_1))\}, \end{aligned} \tag{6}$$

where $l$ is the total number of the pixel whose color is $C_3$ in the vertical direction (Figure 4). The size of $l$ mainly depends on the resolution of the image.

Consequently, the background region of the image $R_{bg}$ is obtained.

$$R_{bg} = \{(x, y) | (x, y) \in (R_1 \cup R_2 \cup R_3)\}, \tag{7}$$

After searching $R_3$, the foreground region $R_4$ is determined as follows:

$$\begin{aligned} R_4 \quad & = \quad \{(x, y) | F(x, y) \in C_3, (x, y) \notin R_3\} \\ & = \quad \{(x, y) | (x, y) \notin R_{bg}\}, \end{aligned} \tag{8}$$

In the foreground region $R_4$, there exist the areas whose colors are not $C_1$ nor $C_2$, but are actually the background region. This is because the setting of the background colors $C_i$ $(i = 1, 2)$ cannot be executed impeccably and the image noise unavoidably exists. These areas are isolated points or isolated areas whose size is very small in most cases. Therefore, isolated areas whose size are under $s_{max}$ pixels in $R_4$ are removed and regarded as the background.

In this procedure, all regions whose colors are same as the background are extracted as the background. An example of extraction error is shown in Figure 3(c). They must be corrected in the next step.

### 2.3 Foreground Extraction

The boundary between the foreground and the background is detected to recheck the foreground region whose color is same as the background, and the background region whose color is same as the foreground. Let $R_5$ be the region whose color is same as the background and is inside the foreground, and $R_6$ be on the surface of the foreground (Figure 3(d)). To detect $R_5$ and $R_6$, the adjacency conditions with $R_3$ are utilized.

The difference between $R_{1,2}$ and $R_5$ is whether it contacts with $R_3$ or not, while the color of $R_{1,2}$ and $R_5$ is same. This is because there exists no two-tone striped region whose colors are same as the background inside the foreground objects. Therefore, the regions that do not contact with $R_3$ among $R_1$ and $R_2$ can be judged as $R_5$.

The distinction is executed at each pixel in the case of $R_1$, $R_2$, $R_3$, and $R_4$. However, $R_5$ is searched at each conjunct region (Figure 5(a)). Here, let $R_{i,j}$ be a subset of $R_i$. $\partial(R_{i,j})$ and $a\vec{d}j(R_{i,j})$ denote the boundary (the edge pixel) of $R_{i,j}$ and the adjacent region of $R_{i,j}$ which is a neighbor pixel of $\partial(R_{i,j})$ outside $R_{i,j}$, respectively (Figure 5(b)). Then, $R_5$ is defined as follows:

$$\begin{aligned} R_5 &= \{R_{i,j} | R_{i,j} \subset (R_1 \cup R_2), \\ &\quad \forall (x,y) \in R_{i,j} (a\vec{d}j(R_{i,j}) \notin R_3)\}, \end{aligned} \tag{9}$$

To determine $R_6$ region, endpoints of $R_3$ are utilized. $R_3$ regions become horizontal lines if the stripes and the scanlines of the camera are perfectly horizontal. However, this does not always occur. Therefore, the approximate line of each $R_{3,j}$ region is calculated (Figure 6(a)). After that, the endpoints of the approximate lines are detected (Figure 6(b)). The neighbor areas of the endpoints that are adjacent to $R_6$ have same color as $C_1$ or $C_2$, although the foreground next to the endpoints has different color from $C_1$ or $C_2$. After judging the endpoints next to $R_1$ or $R_2$, these points are connected smoothly with interpolation lines (Figure 6(b)). To detect the boundary of $R_6$, each pixel which is on the interpolation line is regarded as $R_4$ (Figure 6(c)). As the result of these procedure, the region that is surrounded by the interpolation lines and $R_4$ becomes the same region as $R_5$ which do not contact with $R_3$. Therefore, $R_6$ can be obtained by the same way of detecting $R_5$ (Figure 6(c)).

$$\begin{aligned} R_6 &= \{R_{i,j} | R_{i,j} \not\subset R_5, R_{i,j} \subset (R_1 \cup R_2), \\ &\quad \forall (x,y) \in R_{i,j} (a\vec{d}j(R_{i,j}) \notin R_3)\}, \end{aligned} \tag{10}$$

After that, the image is divided into 6 regions (Figure 3(d)) and the extraction procedure finishes. Accordingly, the foreground region of the image $R_{fg}$ is obtained as follows (Figure 3(e)):

$$R_{fg} = \{(x,y) | (x,y) \in (R_4 \cup R_5 \cup R_6)\}, \tag{11}$$

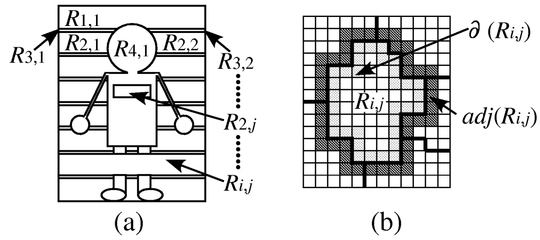


Figure 5: Conjunct region. (a) $R_{i,j}$. (b) $\partial(R_{i,j})$ and $a\vec{d}j(R_{i,j})$.

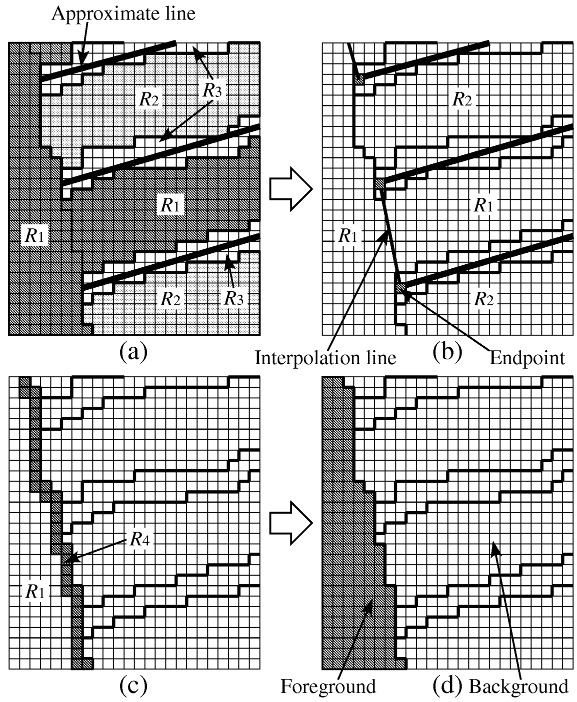


Figure 6: Detection of $R_6$ region

### 2.4 Image Composition

The extracted foreground and another background are combined. The process that smoothes the color of the foreground's boundary against the new background is carried out to create natural images. The human operator decide the position and the size of the foreground objects in compliance with the new background that the relationship between objects do not become unnatural (Figure 3(f)).

## 3 Experiments

In our experiment, horizontal stripes were adopted to extract the boundary precisely because human object is vertically long. The colors of the background are set blue and yellow from the trial and error results that red is similar with the color of brown hairs, and so on. The background is made of blue and yellow paper that are commonly sold in the stationery. The width of the stripe is set as 35mm to interpolate the boundary of $R_6$ region precisely.

When the resolution of the image is $640 \times 480$ pixel, the size of the mask of a median filter for noise removal is set as $3 \times 3$. $l$ (the length of $R_3$ region) is set as 3, and $s_{max}$ (the threshold value for isolated point removal) is set as 4. The value $b_{i,j}$, $c_{i,j}$, $d_{i,j}$, $q_{i,j}$, $r_{i,j}$, $s_{i,j}$ are set as 0 or $\pm 1$. At first, the initial values of $a_{i,j}$ and $p_{i,j}$ are set as 0, and $e_{i,j}$ and $t_{i,j}$ as 255. The human operator changes each value for investigating the color of the striped background. The values that do not change drastically when the lighting condition changes are set as the constant value. The values that change drastically when the lighting condition changes are set as the function of the average $(r, g, b)$ value of the background in the image.

An example of the experimental result is shown in Figure 7. In this figure, a sheet of blue paper that is same paper as the background is rolled around the left arm of the foreground person and a sheet of yellow paper is put on his shirt (Figure 7(a)). The result of the background color extraction and the detection of the endpoints of $R_6$ are shown in Figure 7(b) and (c), respectively. The extracted foreground is shown in Figure 7(d). This result shows that the foreground regions whose colors are same with the background are extracted without fail. The result of the image composition of the extracted foreground and another background is shown in Figure 7(f).

This method has been verified with over a hundred of people whose clothes are different in an indoor environment where the lighting condition changes as time goes on. Although extraction error occurred infrequently because some threshold values were not adjusted for the lighting condition, our method could work successfully with large proportion of cases. From experimental results, the effectiveness of our proposed method has been verified.

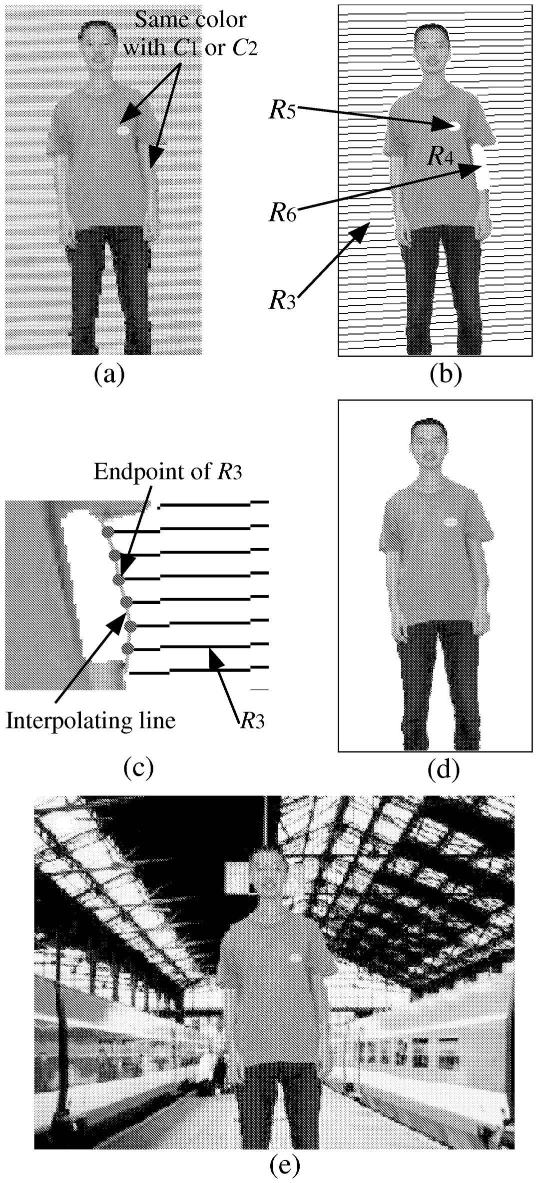


Figure 7: Result of region extraction. (a) Original image. (b) Background color extraction. (c) Detection of $R_6$. (d) Foreground extraction. (e) Image composition.

## 4 Conclusions

In this paper, we propose a new region extraction method with chromakey using two-tone striped backgrounds. We utilize the adjacency condition between two-tone striped areas on the background, and extract the foreground regions whose colors are same with the background. Experimental results show that the foreground region is constantly extracted from the background despite of the colors of the foreground objects.

As the future work, the thresholding process of the background colors must be executed more precisely to adopt the approach based both on the pixel-based segmentation and on the area based segmentation. The width of the stripe can be used for estimating the sizes of the foreground objects to improve the quality of the composed images.

## References


[1] Thomas Porter and Tom Duff: "Compositing Digital Images," *Computer Graphics* (*Proceedings of SIGGRAPH* 84), Vol.18, No.3, pp.253–259, 1984.

[2] King-Sun Fu and Jack Kin Yee Mui: "A Survey on Image Segmentation," *Pattern Recognition*, Vol.13, pp.3–16, 1981.

[3] Wladyslaw Skarbek and Andreas Koschan: "Colour Image Segmentation - A Survey," *Technical Report* 94–32, *Technical University of Berlin, Department of Computer Science*, 1994.

[4] Michael Kass, Andrew Witkin and Demetri Terzopoulos: "SNAKES: Active Contour Models," *International Journal of Computer Vision*, Vol.1, pp.321–331, 1988.

[5] Simon Gibbs, Constantin Arapis, Christian Breiteneder, Vali Lalioti, Sina Mostafawy and Josef Speier: "Virtual Studios: An Overview," *IEEE Multimedia*, Vol.5, No.1, pp.18–35, 1998.

[6] Andrew Wojdala: "Challenges of Virtual Set Technology," *IEEE Multimedia*, Vol.5, No.1, pp.50–57, 1998.

[7] Shigeru Shimoda, Masaki Hayashi and Yasuaki Kanatsugu: "New Chroma-key Imagining Technique with Hi-Vision Background," *IEEE Transactions on Broadcasting*, Vol.35, No.4, pp.357–361, 1989.

[8] Tomoo Mitsunaga, Yaku Yokoyama and Takashi Totsuka: "AutoKey: Human Assisted Key Extraction," *Computer Graphics* (*Proceedings of SIGGRAPH* 95), pp.265–272, 1995.

[9] Paul Debevec, Andreas Wenger, Chris Tchou, Andrew Gardner, Jamie Waese and Tim Hawkins: "A Lighting Reproduction Approach to Live-Action Compositing," *Computer Graphics* (*Proceedings of SIGGRAPH* 2002), pp.547–556, 2002.

[10] Alvy Ray Smith and James F. Blinn: "Blue Screen Matting," *Computer Graphics* (*Proceedings of SIGGRAPH* 96), pp.259–268, 1996.