J. Vis. Commun. Image R. 30 (2015) 242–251

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Background replacement using chromatic adaptation transform for visual communication ☆,☆☆

Tatsuki Murakami *, Yoichi Kageyama, Makoto Nishida

Akita University, 1-1, Tegata Gakuen-Machi, Akita 010-8502, Japan

ARTICLE INFO

Article history:
Received 4 October 2014
Accepted 15 April 2015
Available online 6 May 2015

Keywords:
Visual communication
Background replacement
Chromatic adaptation transform
Illuminant color estimation
Gray world assumption
Dichromatic reflection model
Video chat
Subjective evaluation

ABSTRACT

Replacing a video chat background with a landscape image can generate the realism of a user actually being in the landscape. To enhance this realism, we proposed in our previous study a background replacement method that uses a chromatic adaptation transform. This method can enhance the realism of video chat by fitting the color of the foreground image to an illuminant color of a landscape, which is used as the new background image. However, if an incorrect color of the landscape illuminant is obtained through this method, which estimates the illuminant color on the basis of a gray world assumption, the method might not enhance the realism. This is because it converts the foreground color to an incorrect color. In this paper, we therefore propose a method to estimate illuminant color on the basis of the dichromatic reflection model, which improves background replacement using the chromatic adaptation transform. We perform a subjective evaluation using 13 subjects to examine the effects of the proposed method. The results indicate that the proposed method can effectively enhance the realism of the background replacement video.

1. Introduction

With recent developments in information technology, visual communication, such as video chat, has become widely used in daily life. Video chat enables interactive communication using voice and video, including videos of users themselves. When video chatting, however, private information leakage may occur because the background information behind the user includes personal information (e.g., the contents of a room or office). Therefore, replacing the background with another image can effectively eliminate this leakage. In addition, when a video background is replaced with a landscape image, the replacement should generate a sense of realism, as if the person is actually in the landscape. Nevertheless, the difference of illuminant colors between the foreground image (i.e., the user) and the new background image (i.e., the replacing landscape image) might cause viewer discomfort; the viewer may perceive that the person does not truly exist in the landscape when the illuminant color of the foreground differs from that of the landscape.

To reduce this effect and enhance the realism, we proposed in a previous study a background replacement method that uses the chromatic adaptation transform (CAT) [1]. CAT processing converts the color of the camera image so that the illuminant of the scene appears to be replaced by another illuminant. Accordingly, the proposed method fits the color of the foreground image to an illuminant color of the landscape that serves as the new background image, which can enhance the sense of realism, as shown in Fig. 1.[1] Results of our previous study suggested that the proposed method could reduce viewer discomfort caused by a background replacement video that employs a natural landscape image as the new background.

In our previous study, the illuminant color of the natural landscape for the background replacement was estimated based on a gray world assumption (GWA) [2]; the mean color of the landscape was obtained as its illuminant color. However, if the illuminant color was estimated from both the natural landscape and an artificial landscape (e.g., the landscape image including some buildings), an incorrect illuminant color could be obtained by the GWA. This was because the mean color of the artificial landscape tended to approximate the mean color of buildings in the landscape, not the color of the landscape illuminant. To improve illuminant

☆ This document is a collaborative effort.

☆☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author.
E-mail address: kageyama@ie.akita-u.ac.jp (Y. Kageyama).

[1] For interpretation of color in Figs. 1–11, the reader is referred to the web version of this article.

http://dx.doi.org/10.1016/j.jvcir.2015.04.007
1047-3203/

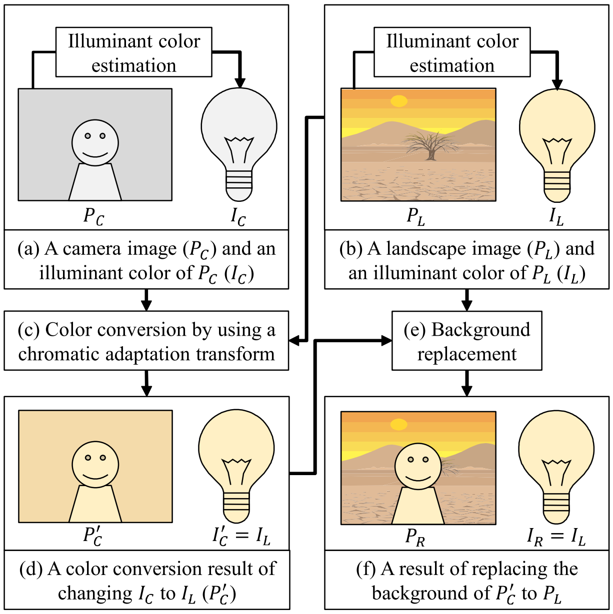

Fig. 1. Background replacement feature using a chromatic adaptation transform.

estimation accuracy for the background replacement, we proposed a method for estimating the illuminant colors of the natural and artificial landscape images based on the dichromatic reflection model (DRM) [3,4]. Our evaluation results suggested that the illuminant color obtained by the proposed method tended to effectively approximate the landscape illuminant color rather than the building color. Although we studied illuminant color estimation for background replacement, we have yet to subjectively evaluation it, such as through user tests.

In this paper, we propose a background replacement method that enhances the realism of replaced video chat backgrounds while protecting user privacy. The proposed method is comprised of the following steps. First, a matrix for CAT processing is generated with illuminant colors estimated from the camera image and new background image as a preprocessing phase. Next, the human region is extracted from the camera image by background subtraction [5]; CAT processing is then applied to this region to obtain the new human region. Finally, the background region of the landscape image is obtained using the background subtraction result, which was previously obtained from the logical disjunction of the new human region and background region. Evaluation results based on the feedback of 13 subjects suggest that the proposed method can produce a greater sense of realism than the GWA for illuminant color estimation. Our experimental results additionally indicate that the proposed method can enhance the realism of the background replacement video whose background was replaced by a landscape image having different illuminant color from a camera image.

2. Color conversion details

The proposed method includes a function that converts the color of a human region to that of a landscape illuminant. The color conversion is implemented by the von Kries model, which expresses chromatic adaptation techniques in terms of matrix transformations [6]. In addition, the proposed method includes an illuminant color estimation function to provide the illuminant colors of the camera and landscape images required by the chromatic adaptation transform.

2.1. Chromatic adaptation transforms

Objects are illuminated by variously colored lighting, such as by the white light of typical daylight, bluish light from overcast skies, and reddish light from sunsets. However, a human visual systems capability called chromatic adaptation adjusts to the color of these varying illumination in order to perceive approximately inherent colors of objects [6]; an orange, for example, always appears orange, no matter if viewed in day or evening light. Color conversion based on chromatic adaptation transform (CAT) enables the prediction of colors under varying illuminant conditions. In Fig. 1(c), the system converts the color of a camera image into the illuminant color of the landscape using CAT based on the von Kries model [6]. Specifically, an output color $(R'', G'', B'')^T$ is calculated from an input color $(R', G', B')^T$ using a 3×3 matrix, $\boldsymbol{M}$, as

$$(R'', G'', B'')^{\mathrm{T}} = \boldsymbol{M}(R', G', B')^{\mathrm{T}}. \tag{1}$$

$\boldsymbol{M}$ is a matrix converting W_S, an illuminant color of an input image, to a target illuminant color W_D. The matrix is previously calculated with $(L_{W_S}, M_{W_S}, S_{W_S})^T$, an LMS cone response of W_S, $(L_{W_D}, M_{W_D}, S_{W_D})^T$, an LMS cone response of W_D, and two 3×3 matrices, $\boldsymbol{A}$ and $\boldsymbol{B}$, of the color space conversion, as

$$\boldsymbol{M} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}\begin{pmatrix} \frac{L_{W_D}}{L_{W_S}} & 0 & 0 \\ 0 & \frac{M_{W_D}}{M_{W_S}} & 0 \\ 0 & 0 & \frac{S_{W_D}}{S_{W_S}} \end{pmatrix}\boldsymbol{AB}. \tag{2}$$

$\boldsymbol{A}$ is a Bradford transform matrix used in the color conversion from a CIE XYZ 1931 color space into an LMS color space [7], as in

$$\boldsymbol{A} = \begin{pmatrix} 0.8951 & 0.2664 & -0.1614 \\ -0.7502 & 1.7135 & 0.0367 \\ 0.0389 & -0.0685 & 1.0296 \end{pmatrix}. \tag{3}$$

Eq. (4) is used to convert color from a CIE XYZ 1931 color space into an LMS color space.

$$(L, M, S)^{\mathrm{T}} = \boldsymbol{A}(X, Y, Z)^{\mathrm{T}}. \tag{4}$$

$\boldsymbol{B}$ is a transform matrix from transforming an RGB color into a CIE XYZ 1931 color [8], as in

$$\boldsymbol{B} = \begin{pmatrix} 0.4124 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix}. \tag{5}$$

Eq. (6) is used to convert color from an RGB color space into a CIE XYZ 1931 color space.

$$(X, Y, Z)^{\mathrm{T}} = \boldsymbol{B}(R, G, B)^{\mathrm{T}}. \tag{6}$$

In this paper, the input and target illuminant colors are respectively estimated from the camera and landscape images to create $\boldsymbol{M}$.

2.2. Illuminant color estimations

In our previous study, W', which is an illuminant color of a landscape image, was estimated based on the GWA, which presumes that the average spectral reflectance for a visual field is equal to a medium gray [2]. Specifically, $(R_{W'}, G_{W'}, B_{W'})^T$, an RGB color of W', was calculated using $(R_i, G_i, B_i)^T$, which is the color of the i-th pixel of the image constructed with N pixels, as in