Journal of The Korea Convergence Society
Vol. 12. No. 10, pp. 1-8, 2021

I-SSN 2233-4890 / e-ISSN 2713-6353
DOI : https://doi.org/10.15207/JKCS.2021.12.10.001

Automatic Video Chromakeying Generation Technology Using Background Modeling

Gil-Sang Yoo
Research Professor, College of Informatics, Korea University

배경 모델링을 이용한 비디오 크로마키 생성기법

유길상
고려대학교 정보창의교육연구소 연구교수

Abstract In online meetings and classes using webcams, the chromakey technique is a very necessary part to produce content. We proposed a technology that enables background synthesis without using a cloth for chromakey. The proposed method consists of three steps: an HSI image conversion step, a step of detecting a region changed from a background, and a step of replacing the background region with a chromakey and applying it. In the input video, the block average image of each frame is calculated, and the difference between the block average image of the background image and the block average image of the input image is used to detect the change area. The developed chromakey effect technology uses a technique of acquiring a background image without an object from a single camera and extracting only an object by distinguishing the moving object and the background. The proposed method is not only capable of processing even if the background has a variety of colors, but also has the seamless processing of the boundary lines of objects.

Key Words : Image convergence, Background, HSI, Webcam, Chromakey, Blue screen, Image processing

요 약 웹캠을 이용한 온라인 회의 및 수업에서 크로마키 기법을 이용한 콘텐츠 제작은 중요한 기법중의 하나이다. 본 연구에서는 크로마키 배경을 사용하지 않고 배경 합성이 가능한 기술을 제안하였다. 제안하는 방법은 HSI 이미지 변환 단계, 배경에서 변경된 영역을 감지하는 단계, 배경 영역을 크로마키로 대체하여 적용하는 단계의 3단계로 구성된다. 입력 영상에서 각 프레임의 블록 평균 영상을 계산하고, 배경 영상의 블록 평균 영상과 입력 영상의 블록 평균 영상의 차이를 이용하여 변화 영역을 검출한다. 개발된 크로마키 효과 기술은 하나의 카메라에서 물체가 없는 배경 이미지를 획득하고 움직이는 물체와 배경을 구분하여 물체만 추출하는 기술을 사용하였다. 실험결과, 제안한 방법은 배경색이 다양한 경우에도 처리가 가능할 뿐만 아니라 물체의 경계선을 매끄럽게 처리할 수 있어서 현장에서 쉽게 적용할 수 있을 것으로 기대할 수 있다.

주제어 : 영상융합, 가상배경, HSI, 웹캠, 크로마키, 블루스크린, 영상처리

*This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1I1A1A01064580).
*Corresponding Author : Gil-Sang Yoo(ksyoo@korea.ac.kr)
Received August 20, 2021 Revised September 4, 2021
Accepted October 20, 2021 Published October 28, 2021

1. Introduction

In online meetings and classes using webcams, the chromakey technique is a very necessary part to produce content. To get the chromakey effect, use a blue or green cloth for a professional studio or small indoor background. In this paper, we proposed a technology that enables background synthesis without using a cloth for chromakey so that the background image can be automatically replaced in the image acquired from the camera.

chromakey has been an important part of filming for broadcasts. Broadcasters have used a chromakey by setting a background space of a single color for chromakey and excluding the frequency of the color from the data input from the camera. This method of chromakey is very sensitive to lighting, so when a shadow occurs, a chromakey error has occurred. Complete lighting and a single background studio were essential for chromakey[1,2].

The background extraction system using video finds and divides moving objects in a continuous image, and predicts and tracks the movement, and provides the necessary analysis. The core of such a system is to detect an object to be tracked, and an image difference method is usually used for object detection[3]. A background difference method is used that generates a changing background image and uses the difference from the consecutive frames[4,5]. The frame difference method compares the difference between the previous frame and the current frame and considers a part with a large difference as a moving object. It requires relatively little memory and requires little computation because it does not require special background extraction. However, in the case of a slowly changing object, it is not possible to extract a moving object, and even if a rapidly changing object, only an area corresponding to the difference before and after the movement of the object remains in the difference image. For this reason, various post-processing is generally required to find the entire object. The background difference method is a method of obtaining an unchanged background image from a series of video images obtained from a fixed camera, and then segmenting the background image by considering it as a moving object that is remarkably different from the background image. The extraction of a background image that is harmonious with is very important. For this, the background image must be updated online, and the background must be extracted while considering the change in lighting over time.

This paper proposes a specific moving object tracking algorithm using color information from a fixed webcam. The proposed method converts the RGB (red, green, blue) color coordinate system of the image input from the camera into the HSI (Hue, saturation, intensity) color coordinate system, and then the background image and the object exist with only the color tone area of the image. The chromakey function was implemented using the differential image technique and the virtual projection technique. Calculate the difference between the block average image of the input video image frame and the block average image of the background image to calculate the change value from the background image, and calculate the difference image to detect changes in the background image due to object movement. By replacing the detected area with a new background image, it is possible to perform the chromakey function.

The proposed technique can produce the chromakey effect only with a personal PC and camera without any chromakey equipment. In an era where un-contact is emphasized, chromakey technology can be expected to greatly contribute to the growing fourth

industry.

This paper is organized as follows. Chapter 2 describes related research and proposed methods, Chapter 3 explains experimental results, and Chapter 4 describes conclusions and future research directions.

2. Materials and Methods

Video motion detection and segmentation are one of the most important steps to detect and segment the motion of an object to be processed[6]. The motion detection and segmentation method in video image processing introduces how to use frame difference, how to use background subtraction from background images, and how to calculate optical flow[7-10].

Object detection methods are largely divided into a background removal method and a direct detection method. In the background removal technology, the background is first removed, the moving object is detected, and then the object is classified. The techniques used for object classification are classified using human shape model-based, wavelet-based functions, contour curve components, and human function vectors. Detection technology that does not use a background recognizes a person directly from an image or video. As features used to classify whether a person is a given input, it detects a person based on shape and movement features, skin color detection, and feature vectors. In direct detection, there is a method of detecting faces at high speed by combining several Haar-like features[11]. However, since a webcam requires relatively fast image processing for real-time processing, a method using only one feature was used.

2.1. The HSI color coordinate system

Background extraction is an important technique for finding moving objects in a chromakey system. In this paper, we propose a background extraction method using the difference between vector images. In the proposed method, color pixel values and vector average values of consecutive frames are used, taking advantage of the fact that the background has a higher cumulative occurrence frequency than the object.

The HSI color coordinate system expresses color with three characteristics: Hue, Saturation, and Intensity. The reason why the HSI color coordinate system is widely used in the image processing field is that when a change in color tone occurs in the RGB color coordinate system, all three RGB parameters change. On the other hand, when the hue changes in the HSI color coordinate system, only the hue H angle changes[12,13].

Color represents the primary color of that color. Saturation indicates the purity of color and indicates how much white is mixed with the primary color. The HSI color model facilitates color conversion. For example, to turn green into a yellow-green color, you simply need to adjust the saturation. If you want to brighten the dark, you can increase the brightness. For this reason, an image processing system uses a lot of HSI models when processing color images.

In some image processing, only brightness is used. In this processing, it is more effective to use the HSI model than the RGB model. It is also effective to use the HSI model when classifying objects by color in computer vision. In the HSI color model, color and brightness are expressed separately, but in RGB color, information on the brightness of light is not separated and the color of an object is affected by the brightness of light, so it is difficult to know whether it was the original color.

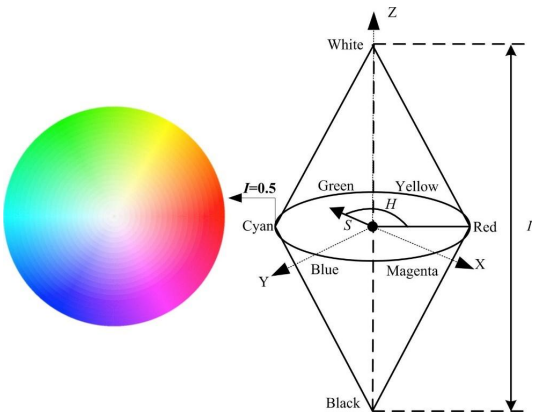

Fig. 1. HSI color model

As shown in Fig. 1, the HSI color model is represented by a conical coordinate system. Colors are expressed as angles ranging from 0 to 360 degrees along the perimeter of the cone. 0 degrees is red, 120 degrees is green, and 240 degrees is blue. Saturation has a value from 0 to 1 and is expressed as the horizontal distance from the center of the cone. The saturation value at the center of the cone is 0, and white becomes 100%, and the saturation at the edge of the cone is 1, resulting in a pure primary color with no white mixture at all. The brightness corresponds to the vertical axis, with the lowermost being 0 and black, and the uppermost being 1 and white.

Also, the HSI color model can more accurately describe the characteristics of color information compared to the RGB color model. As shown in Fig. 1, the angle around the vertical axis corresponds to a pure color and a hue that describes its changing tendency. The distance from the axis corresponds to the saturation, which gives a measure of the degree to which it is diluted by white light. From black to white, the saturation is low and is on the vertical axis. Pure color saturation is much greater than black and white.

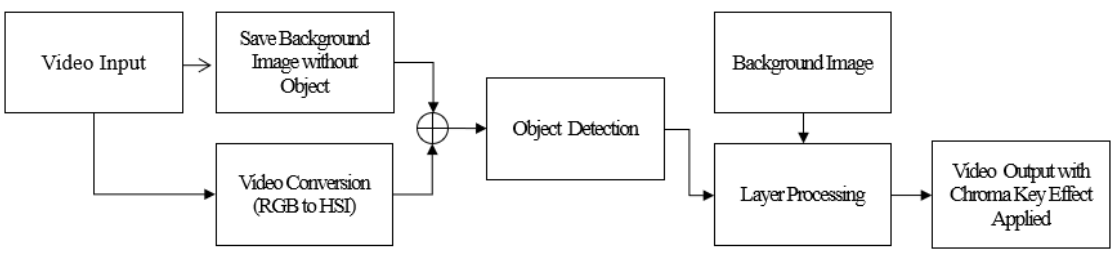

Fig. 2. Proposed automatic video chroma king system

2.2 HSI conversion

Conversion from RGB color model to HSI color model can be performed as following equations (1) - (3) [14,15]. Equations (1) show that the intensity values of the color channels, red, green and blue, have an anticosin relationship with the hue component of the HSI model.

$$H = \cos^{-1}\left[\frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}}\right] \quad (1)$$

$$I = \frac{1}{3}(R+G+B) \quad (2)$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \quad (3)$$

2.3 Background Detection

A simple method of detecting the movement or change of an object is to use the difference in image frames. Equation (4) was used to use the difference in image frames and calculates the difference image Diff(x, y, t) as the difference between the background image and the image frame I(x,y,t) acquired at time t.

$$Diff(x,y,t) = |B(x,y,t) - I(x,y,t)| \quad (4)$$

A pixel with a value greater than the threshold t is divided into a pixel with a change, and a binary image BinImg(x, y, t) divided into a region with and without motion or change is calculated by equation (5).

$$BinImg(x,y,t) = |I(x,y,t) - I(x,y,t-k)| \quad (5)$$

In the next step, since the binary image divided by the threshold value contains motion information for each pixel, pixels having motion are clustered using the connection information between the pixels to find a region having motion. This method can effectively obtain the background image B(x,y), it is suitable when the camera is fixed like a webcam.

In order to detect a person or object, an algorithm that extracts an object through the difference between two backgrounds and replaces the background information with an alternative background image for classification was proposed. The composition of the proposed process is shown in Fig. 2.

3. Results and Discussion

In order to verify the performance of the background detection algorithm proposed in this paper, we implemented and experimented with python in Intel 9900 CPU, 3.6GHz and windows 10 environment. For the experimental video, a background image taken from 30cm in front of a Logitech BRIO 4K camera and a video image of 1920 x 1080 x 24 bits acquired at 30 fps were used.

3.1 Extract objects using background modeling

A typical background removal algorithm extracts the object area by subtracting the current image from the reference image. However, this method does not guarantee accurate object extraction, including changes in the environment, such as lighting from the time difference between the background and the current image. Therefore, this paper assumes a fixed webcam environment, and looks at the environmental factors that may occur in real time and uses a background modeling method to respond to it. Environmental factors in the real-time chromakey background is a fluorescent light, the overall brightness change due to the natural light, the partial brightness change of the camera light source due to the sudden distance difference between the camera and the user, in the case of the night when the overall brightness change due to the lighting and when the person is off the light is off. The background image is stored for use as an initial image that is entered because the moving object does not exist at the beginning. Subsequent background image recurs the image in pixels at certain time intervals.

The detection of a change area is an important detection step as a part that is covered by a background image. When an object or person moves, there is a difference from the background image before and after it is reflected on the camera. In this paper, we use block-by-block difference images to detect the change region. For the block average image, the block average image is calculated by dividing

the input video image frame into a certain square size and then calculating average from each block. Fig. 3 is the result of the image with the background removed from the input video image.

Fig. 3. Input video image (up), background removed image (down)

3.2. Improved background image

chromakey was a broadcasting technology that was possible only by setting an environment that was not possible with a simple image processing system. Lighting and single-color processing space is a burdensome environment for individual broadcasters. We have developed a system that processes complete chromakey from a single camera without any additional equipment. The developed chromakey effect technology uses a technique of acquiring a background image without an object from a single camera and extracting only an object by distinguishing the moving object and the background. Fig. 4 shows the result of synthesizing the background image to be replaced from the original video image with the background removed. The image below in Fig. 4 shows that the input image and the background image match well. In addition, the picture below in Fig. 4 can be output in the inverted form of the original background image.

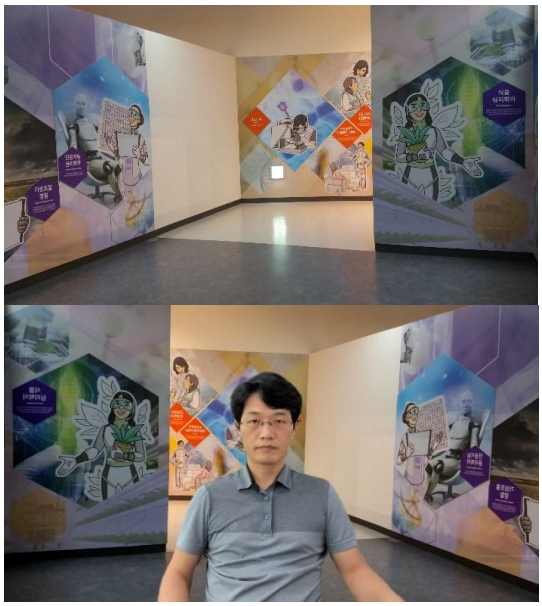

Fig. 4. Alternative background image (up), chromakeying frame (down)

4. Conclusion

Chromakey has been known as a technique that runs with expensive broadcasting equipment and environments. Advanced technology has miniaturized many equipment's and has maximized the performance of multimedia devices. chromakey should also be a system that is perfectly configured even in small-sized equipment and small environments. As such, chromakey was a broadcasting technology that was possible only by setting an environment that was not possible with a simple image processing system. The lighting and single-color processing space creates a

burdensome environment for individual broadcasters. We have developed a system that processes complete chromakey from a single camera without any additional equipment. The developed chromakey uses a technique of acquiring a background image without an object from a single camera and extracting only the object by distinguishing the next moving object and the background. Experimental results show the optimal effect in an indoor environment. However, it has a limitation that it is not suitable for an external environment with severe lighting changes or an environment with background movement. In a future study, it will be studied to show the optimal effect even in a background environment where various movements are frequent.

Chromakey technology which is processed without a separate background tent, can greatly contribute to the development of the un-contact economy.

REFERENCES

[1] A. Yamashita, H. Agata & T. Kaneko. (2008). Every color chromakey. *2008 19th International Conference on Pattern Recognition.* 1-4.
DOI : 10.1109/ICPR.2008.4761643.

[2] S. R. Rodrigues, A. C. Sementille, I. Aparecido & J. R. Ferreira. (2007). The Generation of Scenes in Mixed Reality Environments using the Chromakey Technique. *17th International Conference on Artificial Reality and Telexistence (ICAT 2007).* 296-7.
DOI : 10.1109/ICAT.2007.30.

[3] J. T. Glenn, M. D. Raymond, A. C. Jaime, L. L. Danielle. (2020). Activity analysis of thermal imaging videos using a difference imaging approach. *Journal of Thermal Biology, 91,* 102611.
DOI : 10.1016/j.jtherbio.2020.102611.

[4] C. Yang, J. Gao & F. Chen. (2012). Embedded moving image detection based on background subtraction and finite difference method. *2012 5th International Congress on Image and Signal Processing, Chongqing,* 68-71.
DOI : 10.1109/CISP.2012.6469790.

[5] A. Ajam, M. Forghani, M. M. AlyanNezhadi, H. Qazanfari & Z. Amiri. (2019). Content-based Image Retrieval Using Color Difference Histogram in Image Textures. *2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS).* 1-6.
DOI : 10.1109/ICSPIS48872.2019.9066062.

[6] L. Liu & J. Ding. (2018). Target Detection and Extraction Based on Motion Attention Model. *2018 Chinese Automation Congress (CAC). 2018,* 1919-22.
DOI : 10.1109/CAC.2018.8623780.

[7] G. A. Khaskheli, Z. A. Rajper, Q. A. Mangi, R. S. Hussain & A. H. Shar. (2020). A mediation analysis of social media marketing between the relationship of entrepreneurial marketing strategies and the performance of small & medium enterprises in Pakistan. *Indian Journal of Science and Technology, 13(29),* 2024-34.
DOI : 10.17485/IJST/v13i29.1035

[8] B. Naik, P. B. Dash & J. Nayak. (2020) Extra-tree learning based Socio-economic factor analysis and multi-class adaptive boosting meta-estimator for prediction of agricultural productivity. *Indian Journal of Science and Technology, 13(29),* 2081-101.
DOI : 10.17485/IJST/v13i29.839

[9] V. Ivan, G. P. Plöger & D. S. Johannes. (2019). Virtual reality for animal navigation with camera-based optical flow tracking. *Journal of Neuroscience Methods, 327,* 108403.
DOI : 10.1016/j.jneumeth.2019.108403.

[10] M. Piccardi. (2004). Background subtraction techniques : a review/ Systems, Man and Cybernetics. *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583).* 4:3099-104.
DOI : 10.1109/ICSMC.2004.

[11] K. Singh, A. K. Verma & M. Singh. (2020). Higher order Emden-Fowler type equations via uniform Haar Wavelet resolution technique. *Journal of Computational and Applied Mathematics, 376,* 112836.
DOI : 10.1016/j.cam.2020.112836.

[12] Y. Kong, X. Wang & Y. Cheng. (2018). Spectral-Spatial Feature Extraction for HSI Classification Based on Supervised Hypergraph and Sample Expanded CNN. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensin,*

11(11), 4128-4140.
DOI : 10.1109/JSTARS.2018.2869210.

[13] S. Lingfei, F. Ying, H. Hua & C. Yufeng. (2019). Fast HSI super resolution using linear regression. *IET Image Processing, 13(10),* 1671-79.
DOI : 10.1049/iet-ipr.2018.5475.

[14] A. R. Smith. (1978). Color Gamut transform pairs, *ACM SIGGRAPH Computer Graphics, 12(3),* 12-19.
DOI : 10.1145/965139.807361.

[15] G. S. Yoo & H. C. Kim. (2016). Embedding QR code in the wavelet domain of image for metadata hiding. *Indian Journal of Science and Technology, 9(29),* 1-10.
DOI : 10.17485/ijst/2016/v9i29/94759

유 길 상(Gil-Samg Yoo) [정회원]

· 2003년 8월 : 중앙대학교 영상공학과 (석사)
· 2010년 2월 : 중앙대학교 영상공학과 (박사)
· 2010년 1월 ~ 현재 : (사)한국컴퓨터게임학회 이사
· 2011년 3월 ~ 현재 : 고려대학교 컴퓨터학과 연구교수
· 관심분야 : 영상처리, 이러닝, 디지털영상, 게임
· E-Mail : ksyoo@korea.ac.kr