JDCS
디지털콘텐츠학회논문지
Journal of Digital Contents Society
Vol. 18, No. 8, pp. 1603-1613, Dec. 2017

RGB 색상 기반의 실시간 영상에서 잡음에 강인한 손영역 분할

양 혁 진[1] · 김 동 현[2] · 서 영 건[2*]
[1]경상대학교 컴퓨터과학과
[2*]경상대학교 컴퓨터과학과, 대학원 컴퓨터과학과, 대학원 문화융복합학과

Noise-robust Hand Region Segmentation In RGB Color-based Real-time Image

Hyuk Jin Yang[1] · Dong Hyun Kim[2] · Yeong Geon Seo[2*]
[1]Dept. of Computer Science, Gyeongsang Nat'l University, Gyeongnam 52828, Korea
[2*]Dept. of Computer Science and CCBM, Graduate School, Gyeongsang Nat'l University, Gyeongnam 52828, Korea

[요 약]

본 논문은 널리 알려진 RGB 색상 기반의 웹캠을 사용한 손 영역을 효율적으로 분할하는 방법을 제안한다. 이 방법은 잡음을 제거하기 위하여 네 번의 경험적 전처리 방법을 수행한다. 먼저, 전체 영상 잡음을 제거하기 위하여 가우시안 평활화를 수행한다. 다음으로, RGB 영상은 HSV와 YCbCr 색상 모델로 변환되어, 각 색상 모델에 대해 통계적인 값에 기반하여 전역 고정 이진화가 수행된 후, 잡음은 bitwise-OR 연산에 의해 제거된다. 다음으로, 윤곽 근사화와 내부 영역 구멍 연산을 위해 RDP와 flood fill 알고리즘이 사용된다. 끝으로, 모폴로지 연산을 통하여 잡음을 제거하고 영상의 크기에 비례한 임계값을 결정하여 손 영역이 결정된다. 본 연구는 잡음 제거에 초점을 맞추고 있고 손 동작 인식 응용 기술에 사용될 수 있다.

[Abstract]

This paper proposes a method for effectively segmenting the hand region using a widely popular RGB color-based webcam. This performs the empirical preprocessing method four times to remove the noise. First, we use Gaussian smoothing to remove the overall image noise. Next, the RGB image is converted into the HSV and the YCbCr color model, and global fixed binarization is performed based on the statistical value for each color model, and the noise is removed by the bitwise-OR operation. Then, RDP and flood fill algorithms are used to perform contour approximation and inner area fill operations to remove noise. Finally, ROI (hand region) is selected by eliminating noise through morphological operation and determining a threshold value proportional to the image size. This study focuses on the noise reduction and can be used as a base technology of gesture recognition application.

색인어 : RGB 색상, 손 영역, 손 영역 분할, 가우시안 스무딩, 플러드 필
Key word : RGB Color, Hand Region, Hand Region Segmentation, Gausian Smoothing, Flood Fill

http://dx.doi.org/10.9728/dcs.2017.18.8.1603

Received 05 December 2017; **Revised** 21 December 2017
Accepted 25 December 2017

***Corresponding Author; Yeong Geon Seo**

Tel: +82-55-772-1392
E-mail: young@gnu.ac.kr

Ⅰ. Introduction

Microchip has evolved rapidly as represented by formulas based on empirical predictions of Moore's Law. On the other hand, the mouse and keyboard that have been invented in the 1960s and 1970s and widely used so far have stagnated when viewed from the interface performance, that is, User eXperience(UX). However, from the current point of view that the fourth industrial revolution has recently emerged and technologies such as Augmented Reality(AR) and Internet of Things(IoT). The need for a new and intuitive interface that enables a variety of objects and humans to interact conveniently is emerging in preparation for the future ubiquitous computing environment[1].

When we look at the history of development, fingers can express various forms in joint structure, and have played a very large role as a means of communication between human beings before natural language development. In other words, hand gesture has played a role as a basic communication means of humans, and it can be regarded as a historic and very intuitive interface which is still used as a means of communicating between human beings. The Natural User Interface(NUI) based on the recognition of the hand gesture has been continuously studied in the Human-Computer Interaction(HCI) field by focusing on the intuitive and original features of such hand gestures and breaking away from the physical input interface using the existing keyboard and mouse[1][2].

A gesture-based interface such as a hand gesture can be classified into a contact-type method that uses a user's movement by attaching a sensor or a device to a user's body, and a contactless method that acquires a user's movement with a camera and uses only the visual technique of the acquired image. The contact method is used to track and recognize relatively accurate user gestures using expensive equipment or sensors, but users must take the inconvenience of using unnatural equipment. In contrast, the non-contact method eliminates unnecessary equipment from the user and ensures natural movement, but it is difficult to track the details of the user's movement relatively.

The contactless gesture interface is divided into technical categories of hand region segmentation and hand movement recognition. The hand region segmentation method can be divided into hand region segmentation method based on color information, hand region segmentation method based on depth information, and a proper combination of the two methods when discriminating based on required digital information. In general, compared with the hand region segmentation method based on color information, the depth information based hand segmentation method is highly accurate in hand region segmentation because the degree of influence of the surrounding luminance is low. However, from the viewpoint of the user, there is a disadvantage that a separate sensor for measuring depth information is required. In addition, the color information based hand region segmentation methods that have been studied in the past have mostly limited the noise correction by dividing the hand region using the single color model and the statistical color range.

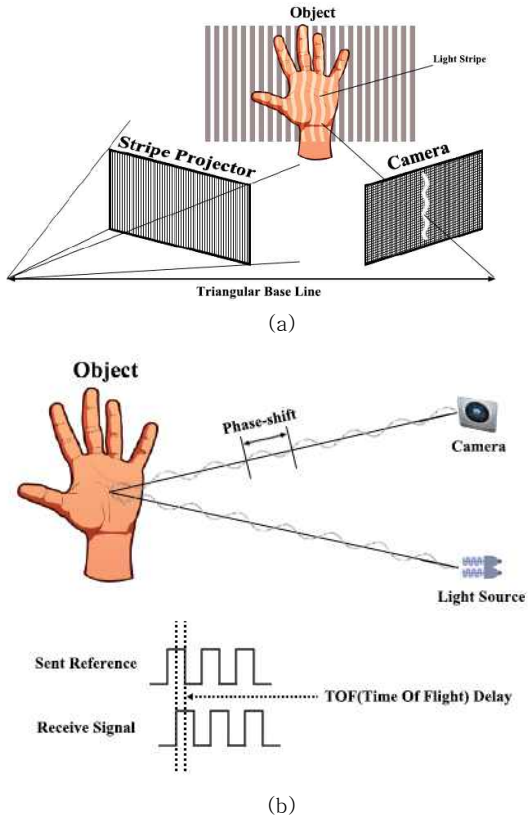


Fig. 1. (a) Structured light technique (b) TOF technique

In this paper, first, it is assumed that application technology using gesture recognition is applied to a webcam that is common in everyday life without requiring a separate sensor. And we aimed to improve the noise reduction and accuracy of the conventional single color model based hand region segmentation method. Therefore, we propose Noise-robust hand region segmentation method in RGB color-based real-time image which is the technical basis of such application technology.

Ⅱ. Related Studies

2-1 Color Information Based Hand Region Segmentation

The hand region segmentation method based on color information divides the hand region using information of each channel in a specific color model. Generally, the RGB color model, the HSV color model, and the YCbCr color model are used in many cases. This color information-based hand region segmentation method is adopted by many hand region segmentation studies in consideration of the economical efficiency of the real-world environment to which the technology is applied. The hand region segmentation based on the color information can be roughly divided into the static threshold-based hand region segmentation method and the dynamic threshold-based hand region segmentation method[3-5].

The static threshold-based hand region segmentation method divides the hand region using a statistical color value range of a specific skin color as a reference of hand region segmentation[6-10]. The dynamic threshold-based hand region segmentation method divides the hand region by calculating a difference image in units of pixels for an input image in real time, dynamically determining a threshold value, and dividing the hand region. Conventional static threshold-based hand region segmentation methods generally divide the hand region through a statistical color range for the skin color of a certain race on the basis of a single color model. Therefore, even if the skin color is the same race, there is a limit to divide the difference of the skin color slightly by each person. The dynamic threshold-based hand region segmentation method using the difference image is a method in which when the hardware (webcam) that is acquiring real-time video physically shakes or moves other areas except the hand, such as the background and the face. There is a disadvantage in that the accuracy is poor.

2-2 Depth Information-based Hand Region Segmentation

Depth information-based hand region segmentation method is dividing an object (hand) region by measuring depth information using a separate depth sensor and clustering the corresponding depth information. There are two types of depth sensing techniques : Structured Light technique and Time of Flight(ToF) technique(fi. 1). The structural optical technique is a technique of projecting the infrared rays of a specific pattern scattered through the scatter film of the IR projector and calculating the depth information by comparing the degree of IR pattern distortion of the object region with the parallax. The ToF technique is a method of projecting infrared rays from an IR projector to measure depth information by measuring the return time of reflected objects. The depth information based hand region segmentation method is generally more accurate than the color information based one which is highly interfered with the luminance change, but requires a separate sensor to measure depth information. Therefore, there is a drawback in that it is economically inefficient from a user point of the hardware configuration level of a small-sized video input device such as a widely-deployed webcam.

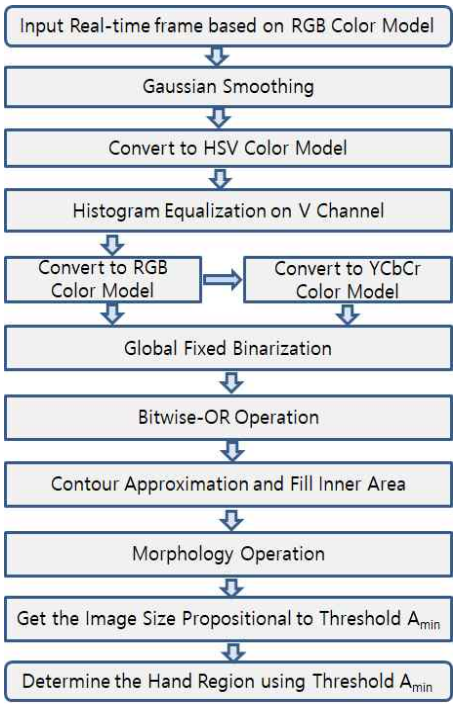


Fig. 2. Overall structure of proposed method

Ⅲ. Noise-robust Hand Region Segmentation

In this chapter, we describe the configuration of the whole system for segmenting the hand region and the interface and environmental preconditions for effectively performing the proposed technique.

3-1 Proposed Method Configuration and Premise

The overall configuration is shown in fig. 2. The proposed method can be classified as 1st order noise removal by normal distribution[11], global fixed binarization by statistical value (2nd order noise removal), 3rd order noise removal by using contour approximation and flood fill, 4th noise removal through morphology operation', and 'hand region selection by determining image size proportional threshold value'. First, a Gaussian smoothing technique is applied to remove an overall noise of an image. Next, global fixed binarization is performed using HSV and YCbCr color model, and bitwise-OR operation is performed to remove noise. Then, the contour approximation is performed using the RDP (Ramer Douglas Peuker) algorithm and the noise is removed by filling the approximated contour with the flood fill algorithm. Finally, the noise is removed by morphology operation, and the hand region is selected by determining the image size proportional threshold value.

This method is to effectively divide the hand region using the RGB color-based webcam, as mentioned earlier in the introduction. Therefore, the RGB color-based webcam must be connected to PC, and the webcam must be located in front of the user and able to acquire frames containing the user's hand in real time. And this technique using only RGB color information obtained from a webcam without a separate sensor is influenced by the capability of securing the light quantity of the hardware(webcam). In other words, depending on the hardware performance of the webcam, the degree of response to the change in the brightness of the surrounding environment varies. Most of the newer webcam products on the market today come with software technology that detects ambient lighting and automatically adjusts frame rate and exposure, which can interfere with achieving consistent performance results. Therefore, it is recommended not to use frame rate and exposure compensation function as much as possible to obtain stable image information. In addition, the method proposed in this paper includes a global fixed binary technique using statistical numerical values. The statistical value referred to here is a statistical value for skin color based on east Asian people[12]. Therefore, the hand region segmentation method proposed in this paper is based on the skin color of east Asian people.

3-2 1st Order Noise Removal by Normal Distribution

First, the overall noise is removed for the real-time frame obtained from the webcam. The Gaussian smoothing technique is applied to this technique, and the noise is randomly distributed, and the Gaussian distribution is observed. That is, a convolution is performed on the input image using a 15 × 15 Gaussian kernel generated through a two-dimensional Gaussian function such as eq. 1.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \tag{1}$$

(a) (b)

Fig. 3. (a) Source image (b) Image that is convoluted with a 15x15 Gaussian kernel

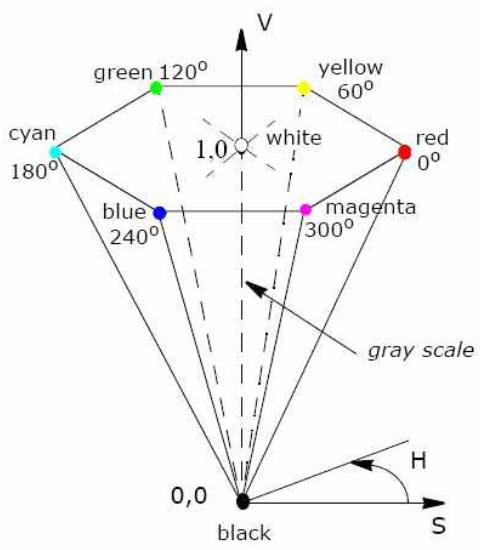
3-3 Global Fixed Binarization by Statistical Numbers

As mentioned earlier, conventional RGB color model based images are highly responsive to changes in luminance of the surrounding environment. As a generalized solution to this problem, there is a method of avoiding the influence of the brightness information as much as possible using the normalized RGB, and expressed by eq. 2 [2][5].

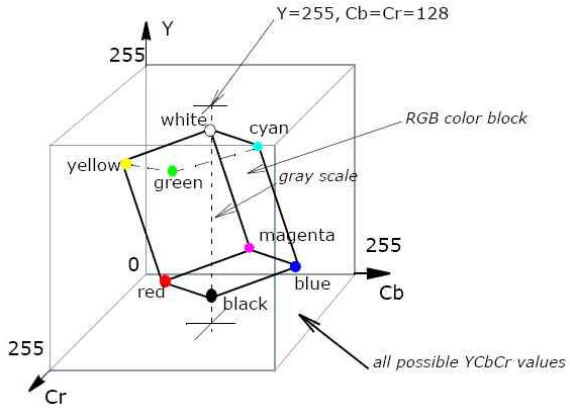
$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \tag{2}$$

However, the method using the normalized RGB has a problem that the color information becomes very unstable when the overall brightness of the input image is dark. For example, when the RGB values of one pixel are R = 2, G = 0, B = 0 and the RGB values of adjacent pixels are R = 0, G = 0, B = 3, RGB values of the two pixels normalized by using R, G, and B become RGB (1, 0, 0) and RGB (0, 0, 1), respectively. That is, the values of two pixels, which appear almost unchanged as a black color when viewed by the human eye, can be expressed as noise appearing in the opposite colors after the normalization process. Therefore, to pursue the robustness against the luminance change, we examined different color models other than the RGB color model, and it is appropriate to use HSV model and YCbCr model as shown in fig. 4 (a) and (b). Unlike the RGB color model, the HSV color model and the YCbCr color model separately represent brightness components as one channel. V channel in HSV color model, Y channel in YCbCr color model. By using the configuration of such a color model and performing binarization

by considering only the other components except for the channel representing the brightness component in each color model, the binarization can be performed with less influence of the luminance change. In this section, we describe how to perform global fixed binarization using HSV color model and YCbCr color model, ie, two color models.



(a)



(b)

Fig. 4. (a) HSV color model (b) YCbCr color model

1) Global fixed binarization using HSV model

First, an HSV color model is transformed using eq. 3 for an RGB color model-based image.

$$
\begin{aligned}
&(R', G', B') = (R/255, G/255, B/255) \\
&C_{\max} = \max(R', G', B') \\
&C_{\min} = \min(R', G', B') \\
&\triangle = C_{\max} - C_{\min}
\end{aligned} \tag{3}
$$

$$
H = \begin{cases} 0^\circ & \text{if } \triangle = 0 \\ 60^\circ \times \left(\dfrac{G' - B'}{\triangle} mod 6 \right) & \text{if } C_{\max} = R' \\ 60^\circ \times \left(\dfrac{B' - R'}{\triangle} + 2 \right) & \text{if } C_{\max} = G' \\ 60^\circ \times \left(\dfrac{R' - G'}{\triangle} + 4 \right) & \text{if } C_{\max} = B' \end{cases}
$$

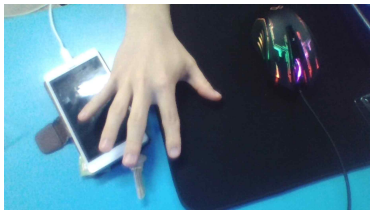
$$
S = \begin{cases} 0 & \text{if } C_{\max} = 0 \\ \dfrac{\triangle}{C_{\max}} & \text{if } C_{\max} \neq 0 \end{cases}
$$

$$
V = C_{\max}
$$

As shown in fig. 5, histogram smoothing is performed on the V channel using eq. 4 to correct the uneven brightness distribution after conversion to HSV color model.

$$
\begin{aligned}
&p_x(i) = p(x = i) = \frac{n_i}{n} \; (0 \le i \le L) \\
&cdf_x(i) = \sum_{j=0}^{i} p_x(j) \\
&cdf_y(i) = iK \\
&y' = y \bullet (\max\{x\} - \min\{x\}) + \min\{x\}
\end{aligned} \tag{4}
$$

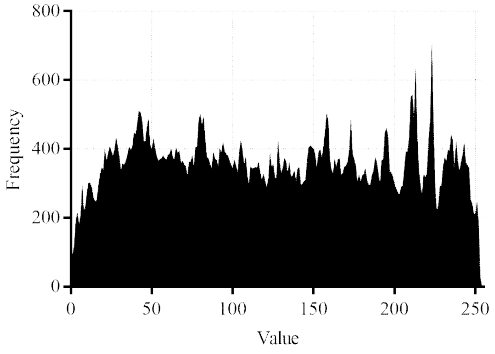
$$
B_{hsv}(x, y) = \begin{cases} 255 & \text{if } 0 \le H \le 50 \\ & \text{and } 20 \le S \le 255 \\ 0 & o.w \end{cases} \tag{5}
$$



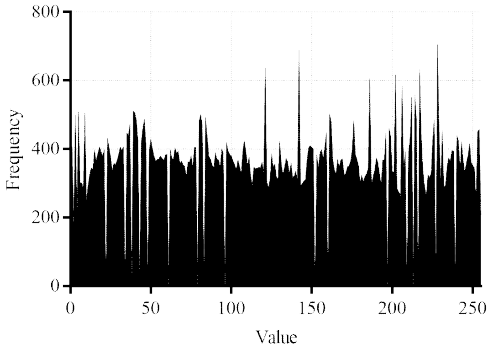
(a)

(b)

Fig. 5. (a) Source Image (b) Image transformed to HSV color model

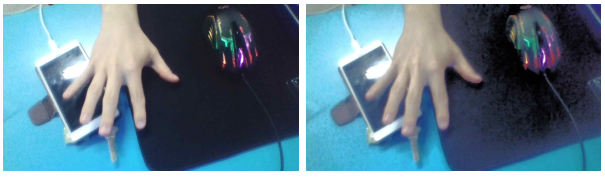


(a)



(b)

Fig. 6. Histogram of V-channel values (a) before and (b) after performing histogram smoothing



(a) (b)

Fig. 7. (a) Source image (b) Image after histogram smoothing, and converting into the RGB color model on V channel of an HSV image

Next, histogram smoothing is performed to combine the V-channel in fig. 6 with uniformly distributed brightness components and the existing H and S channels that are not subjected to any processing, and then convert them into an RGB color model image. Fig. 7 shows the histogram smoothing of the original RGB image and the V channel of the transformed HSV color model, and the result of conversion into the RGB color model again. Finally, we perform global fixed binarization using different threshold values for H and S channels for the RGB image normalized by the brightness distribution as shown in eq. 5.

2) Global fixed binarization using YCbCr model

In the previous section, the histogram smoothing is performed on the V channel of the HSV color model, and then converted to the RGB color model. This time, we convert the improved RGB image to the YCbCr color model and perform global fixed binarization. Similarly, global fixed binarization is performed using different threshold values for each channel. The process of converting RGB color model to YCbCr color model is shown in eq. 6.

$$\begin{aligned} Y &= 0.299 \times R + 0.587 \times G + 0.114 \times B \\ Cr &= 0.7132 \times (R - Y) + 128 \\ Cb &= 0.5643 \times (B - Y) + 128 \end{aligned} \tag{6}$$

The modified RGB image is transformed into the YCbCr color model through eq. 6, and global fixed binarization is performed on Cr and Cb channels using eq. 7.

$$B_{yuv}(x,y) = \begin{cases} 255 & \text{if } 133 \le Cr \le 173 \\ & \text{and } 77 \le Cb \le 127 \\ 0 & o.w \end{cases} \tag{7}$$

(a)

(b)

(c)

Fig. 8. (a)B_{hsv} (b) B_{yuv} (c) B_{sum}

3) 2nd order noise removal

When the two binary images generated in the previous two chapters are assumed to be B_{hsv}, B_{yuv} respectively in fig. 8, we obtain a binary image with noise removed by performing bitwise-OR operation using eq. 8.

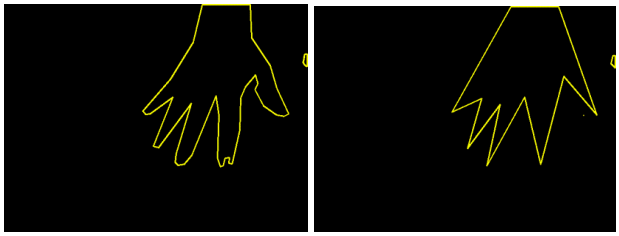
$$B_{sum}(x,y) = B_{hsv}(x,y) \mid B_{yuv}(x,y) \tag{8}$$

3-4 3rd Order Noise Removal by Contour Approximation and Interior Region Filling

After detecting the contour of the object, we use the RDP algorithm to compensate for the unclear boundary due to noise. RDP algorithm determines the value of ε(Epsilon) in a free curve (contour), reduces the number of free-running points, and then approximates it to a number of straight lines. Fig. 9 shows how the contour of the hand region approximates when ε is set to a low value and when ε is set to a high value. In the proposed method, ε of the RDP algorithm is calculated and applied by eq. 9.

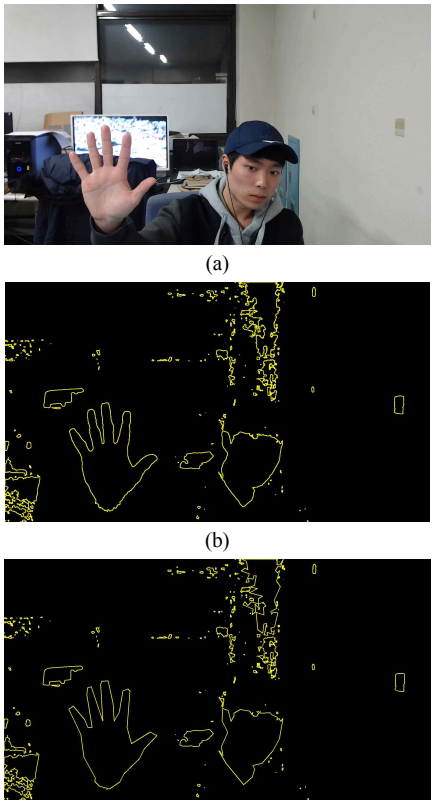
$$\varepsilon = L_{\text{arc}} \times 0.003 \tag{9}$$

L_{arc} is the object's contour length in pixels, and 0.003 is the empirical critical constant. The result of calculating ε through the equation and applying it to the actual image is shown in fig. 10. After correcting the contour of the object through its approximation, flood fill algorithm, which is one of the link element analysis algorithms, searches in 4 directions and fills in the inner region to remove the noise in the third order. The schematic principle of the algorithm is shown in fig. 11. And the result of the algorithm is shown in fig. 12.



(a) (b)

Fig. 9. (a) Low threshold (b) High threshold of ε



(a) (b) (c)

Fig. 10. (a) Source image (b) Contour detected before performing contour approximation (c) Corrected contour after performing contour approximation

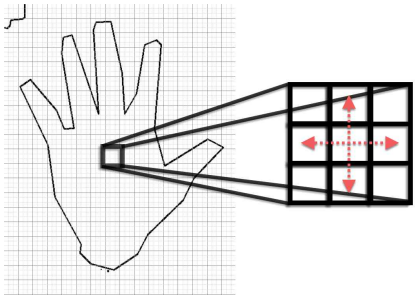
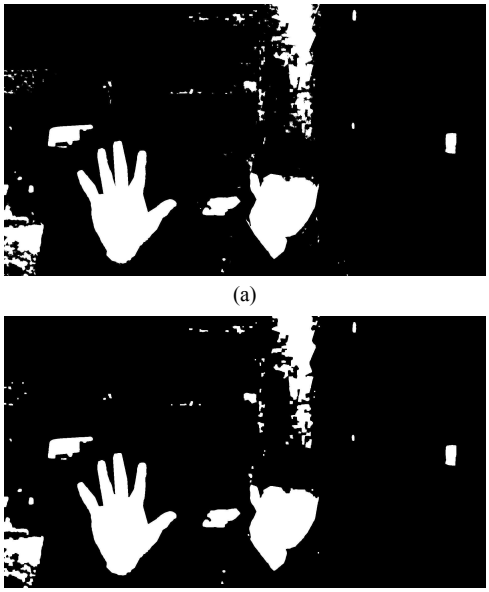


Fig. 11. 4-way flood fill algorithm

Fig. 12. The result performing flood fill algorithm



(a)

(b)

Fig. 13. (a) Before (b) After performing morphology

3-5 4rd Order Noise Removal Through Morphology Operation

Morphology operation is a technique used to transform the shape of an object existing in the image[13-15], and to view and access the image from a morphological point of view. It includes erosion and expansion, and open and close operations that combine the two operations in meaningful order. Open operation has the effect of removing erosion by performing erosion operation first and expansion operation. On the other hand, the close operation is effective in removing the holes by performing the expansion operation first and performing the erosion operation. Therefore, in the proposed method, the open and close operations are performed two times to obtain the noise reduction effect as shown in fig. 13.

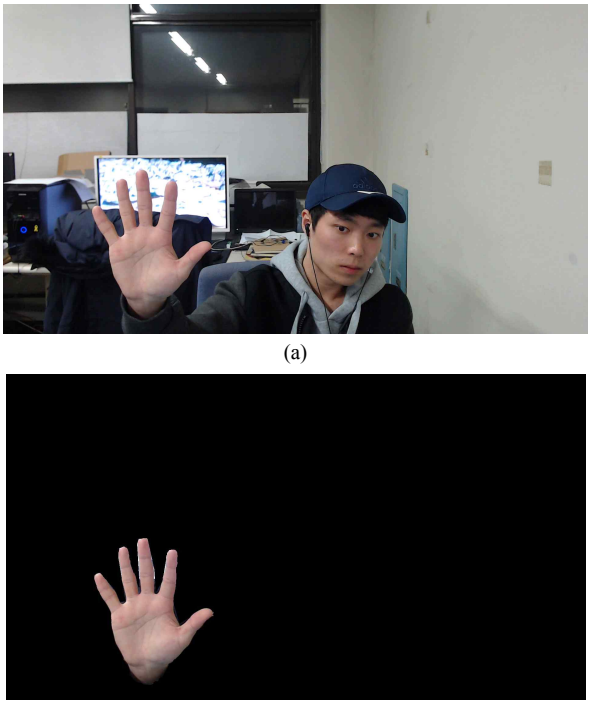
3-6 Selection of Hand Region by Determining Threshold Value Proportional to Image Size

Finally, the hand region is selected for the noise-removed binary image. According to the assumptions described in section 3-1 (webcam should be taken from the front of the user and capable of acquiring images with the hand), the most convenient condition for the user to take a gesture is that the hand is in front of the face to be. Considering these assumptions, we can conclude that the hand region is highly likely to be the link element that occupies the largest region in the binary image. However, if the image does not include the hand at a certain point in time, the noise that occupies the largest region among the various noises may be determined as the hand region.

Therefore, in the method, minimum area threshold $A_{\min}$, which is a criterion for selecting as the hand candidate region, is determined through empirical numerical values through eq. 10. $A_{\min}$ is determined by multiplying the total size ($w \times h$) of the image by a critical constant r representing a specific rate. In other words, the area size (in pixel unit) occupied by each link element is calculated by analyzing the connection element of the binary image, and if the size of the region is equal to or greater than $A_{\min}$, it is regarded as the hand candidate region.

$$r = 1.0/24.0$$
$$A_{\min} = w \times h \times r \tag{10}$$

Of the hand candidate regions selected like above, the region occupying the largest area is finally selected as the hand region. Fig. 14 shows the original image and the final selected hand region.



(a)

(b)

Fig. 14. (a) Source image (b) The result image

Table 1. PC Environment for experiment

object	specification
OS	Windows 10 Home
CPU	Intel(R) Core(TM) i5-6500
Memory	8192MB
VGA	NVIDIA GeForce GTX 1060 6GB
Webcam	Logitech HD Pro Webcam C920

Ⅳ. Experimentation and Evaluation

Table 1 shows the PC environment used for the experiment of the proposed method. We also used Microsoft Visual Studio 2017 Community version as the IDE to implement the proposed system and OpenCV v3.1.0 version which is the open source library. First, 105 frame images were randomly extracted from RGB color based real time images. The extracted 105 frames are shown in fig. 15.

To get the hand region segmentation accuracy for each 105 frame image extracted randomly, the position of the hand region is designated directly for each frame image and converted into a binary image as shown in fig. 16(b). As results, fig. 17 shows 3 images of all test images for objective view. When the proposed method is applied to F_s, 105 original frame images and a set of binary frame images is defined as F_c, which is a comparison reference by directly designating a hand region, Can be calculated by eq. 11.

$$
\begin{aligned}
&F_s = \{F_{s1}, F_{s2}, F_{s3}, \cdots, F_{s103}, F_{s104}, F_{s105}\} \\
&F_c = \{F_{c1}, F_{c2}, F_{c3}, \cdots, F_{c103}, F_{c104}, F_{c105}\} \\
&t = 0, 1, 2, \cdots, 104, 105 \\
&n_c = count(F_{ct}) \text{ if } F_{ct}(x, y) = 255 \\
&n_i = count(F_{st} \cap F_{ct}) \\
&Accuracy = \frac{n_i}{n_c} \; (0.0 \leq Accuracy \leq 1.0)
\end{aligned}
\tag{11}
$$

Through the eq. 11, the results of experiment with 105 frame images in a relatively uniform illumination environment are shown in Table 2. The test was conducted in an environment with a small change in luminance and uniform illumination, showing an accuracy of 86.8517%. The accuracy of the test for most frame images was more than 98%, but the accuracy was reduced by about 70 ~ 80% due to the reflected light from the hand or the afterglow when the hand was moved quickly. Therefore, unlike the experimental data presented in Table 2, the actual arithmetic mean of 86.7896% accuracy is analyzed by the experimental data in this case.

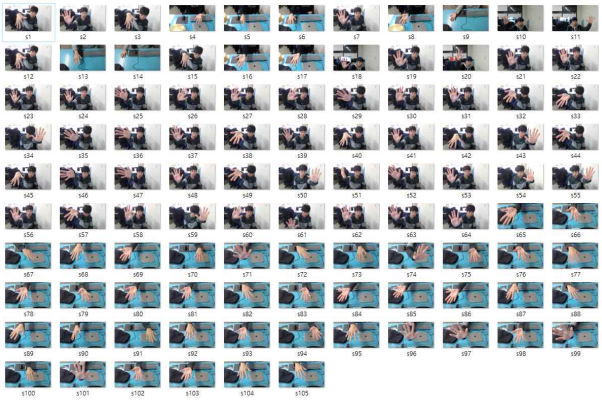
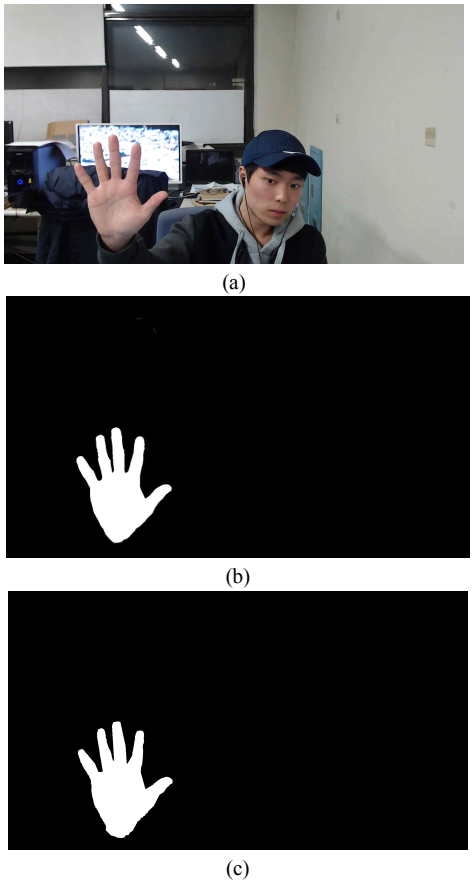


Fig. 15. 105 frame images randomly extracted from RGB color-based real-time image



(a)

(b)

(c)

Fig. 16. (a) Source image (b) Image binarized by specifying the hand regions directly by human (c) Image of divided hand region using proposed method

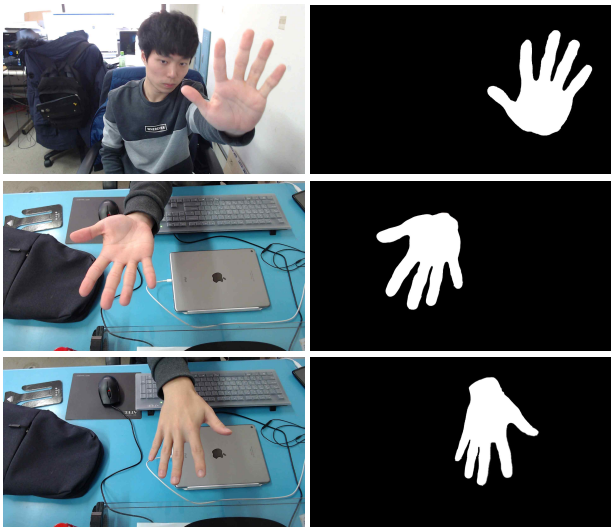


Fig. 17. Left-Source Image, Right-Result Image

Table 2. Experimental results

Value / Frame Number	experimental data		
	n_c	n_i	$Accuracy$
Frame 001	281775	277948	98.6418%
Frame 002	164308	158355	96.3769%
Frame 003	131845	131115	99.4463%
Frame 004	207086	205675	99.3186%
Frame 005	169803	169366	99.7426%
Frame 006	137432	134962	98.2027%
...			
Frame 100	268721	223282	83.0906%
Frame 101	155094	153711	99.1083%
Frame 102	225790	215126	95.2770%
Frame 103	144566	143477	99.2467%
Frame 104	166762	165554	99.2756%
Frame 105	173152	171598	99.1025%

Ⅴ. Conclusion

In this paper, we propose a noise-robust hand region segmentation method that is the technical basis of the contactless gesture interface. The hand region is divided using only the RGB color information of the real-time image input from the webcam without requiring a separate sensor. Compared with the conventional single color model based hand area segmentation method, noise reduction and hand region segmentation accuracy are improved. Experimental results of 105 frame images extracted randomly from RGB color based real time images show 86.8517%. Therefore, it can be used as a base technology of a contactless gesture interface using a RGB color-based webcam in a general real-life environment. However, due to the characteristics of this method, which performs color-based hand area segmentation without a separate sensor, it has the disadvantage that a certain level of interface and environmental preconditions must be provided as in other studies using color-based hand region segmentation. This is a hardware limitation considering the economical efficiency of the user. If the level of the hardware to be supplied later is leveled up through price and proper compromise, there is room for improvement in accuracy. Therefore, we will ultimately study finger recognition and gesture recognition considering future environment, aiming at improvement of gesture recognition accuracy of non-contact gesture interface.

References

[1] D. Hong and W. Woo, "Recent Research Trend of Gesture-based User Interfaces", *Telecommunications review*, Vol. 18(3), pp. 403-413, 2008.

[2] S. Park and E. Lee, "Hand Gesture Recognition Algorithm Robust to Comple Image", *Journal of Korea Multimedia Society,* Vol. 13, pp. 1000-1015, 2010.

[3] V. Vezhnevets, V. Sazonov and A. Andreeva, "A Survey on Pixel-Based Skin Color Detection Techniques", *Proc. Graphicon*, Vol. 3, pp. 85-92, 2003.

[4] M. Na, H. Kim and T. Kim, "An Illumination and Background-Robust Hand Image Segmentation Method Based on the Dynamic Threshold Values", *Journal of Korea Multimedia Society,* Vol. 14(5), pp. 607-613, 2011.

[5] F. Gasparini and R. Schettini, "Skin Segmentation using multiple thresholding", *Proceedings of SPIE,* Vol. 6061, pp. 1-8, 2006.

[6] M. Jones and J. M. Rehg, "Statistical Color Models with Application to Skin Detection", *International Journal of Computer Vision,* Vol. 46, pp. 81-96, 2002.

[7] S. L. Phung, A. Bouzerdoum and D. Chai, "Skin Segmentation Using Color Pixel Classification:Analysis and Comparison", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 27, pp. 148-154, 2005.

[8] J. Han, G. M. Award, A. Sutherland and H. Wu, "Automatic Skin Segmentation for Gesture Recognition Combining Region and Support Vector Machine Active Learning", *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*, pp. 57-64, 2006.

[9] M. Storring, T. Kocka, H. J. Andersen and E. Granum, "Tracking Regions of Human Skin through Illumination Changes", *Elsevier Pattern Recognition Letters,* Vol. 24, pp. 1715-1723, 2003.

[10] P. Kakumanu, S. Makrogiannis and N. Bourbakis, "A Survey of Skin-Color Modeling and Detection Methods",

Elsevier Pattern Recognition, Vol. 40, pp. 1106-1122, 2007.

[11] G. D. Finlayson, B. Schiele, and J. L. Crowley, "Comprehensive Colour Image Normalization", *Proc. European Conference on Computer Vision,* Vol. 1, pp. 475-490, 1998.

[12] S. Cho, J. Bae and S. Lee, "Adaptive Skin Segmentation based on Region Histogram of Color Quantization Map", *Journal of KIISE: Software and Applications,* Vol. 36(1), pp. 54-61, 2009.

[13] J. Kim and M. Kang, "A Study on Hand Shape Recognition using Edge Orientation Histogram and PCA", *Journal of Digital Contents Society,* Vol. 10, pp. 319-326, 2009.

[14] S. Kang, M. Nam and P. Rhee, "Color Based Hand and Finger Detection Technology for User Interaction", *International Conference on Convergence and Hybrid Information Technology*, pp. 229-236, 2008.

[15] J. G. Back and etc, "Extracting the Slope and Compensating the Image Using Edges and Image Segmentation in Real World Image", *Journal of DCS,* Vol. 17, No. 5, pp. 441-448, Oct. 2016

Hyuk Jin Yang

2015~now : Department of Computer Science of Gyeongsang National University

※ Research Interests : AI, Hand Gesture Recognition, Virtual Reality

Dong Hyun Kim

2016. 2 : Department of Computer Science of Gyeongsang National University (B.S Degree)

2018. 2 Department of Computer Science, Graduate School of Gyeongsang National University (M.S Degree Candidate)

※ Research Interests : AI, Deep Learning, Image Segmentation

Yeong Geon Seo

1987.2 : Department of Computer Science of Gyeongsang National University (B.S Degree)

1997.2 : Department of Computer of Soongsil University (Ph.D Degree)

1989~1992 : Trigem Computer Inc.

1997~now : Department of Computer Science of Gyeongsang National University, Professor

2014~now : Department of CCBM, Graduate School of Gyeongsang National University, Professor

※ Research Interests : AI, Deep Learning, Medical Imaging, IT Convergence, Computer Network