

Hand Segmentation Using Depth Information and Adaptive Threshold by Histogram Analysis with color Clustering

Rabia Fayyaz[†], Eun Joo Rhee^{††}

ABSTRACT

This paper presents a method for hand segmentation using depth information, and adaptive threshold by means of histogram analysis and color clustering in HSV color model. We consider hand area as a nearer object to the camera than background on depth information. And the threshold of hand color is adaptively determined by clustering using the matching of color values on the input image with one of the regions of hue histogram. Experimental results demonstrate 95% accuracy rate. Thus, we confirmed that the proposed method is effective for hand segmentation in variations of hand color, scale, rotation, pose, different lightning conditions and any colored background.

Key words: Hand segmentation, Hand area, Adaptive threshold, Clustering, Hue histogram analysis, Depth information.

1. INTRODUCTION

Segmentation of hand from an image is an essential step for motion tracking and gesture recognition. The purpose of hand segmentation is to detect the orientation and position of hands or fingers for better human computer interaction. However, it is a step, as it acquires a general skin color model for hands of all human races which has large variation caused by various skin color, lightning conditions and colored background.

Many methods have been developed to segment hand using color information [1-6]. A Heuristics approach is used to select the predefined threshold for segmenting hand in YCbCr color space [2]. This same approach is used by experimenting different color spaces, where hand segmentation method worked for static background as well as dependent on lightning condition [3]. These meth-

ods have some difficulties in segmenting hands for all human races in any colored background and lightning condition. Similar problems are found in work [4], where the background subtraction technique is used to segment hand from background in varying illumination that are controlled by grey world algorithm [5]. And skin is detected using Bayes method to estimate the skin and non-skin probability of pixels given in RGB color model. Anyhow, this method for hand detection is not enough to cover skin color in different illumination and colored background. Powar et al. [6] experimented YCbCr and RGB color spaces for face detection. It produced good segmentation result using YCbCr color space, and canny edge detector is used to separate the skin and non-skin region. But this method fails to discriminate the face when background color matches with skin color.

Instead, Kang [7] removed background using

* Corresponding Author : Eun Joo Rhee, Address: (305-719) Dongseodae-ro 125, Yuseong-gu, Daejeon, Korea, TEL : +82-42-821-1205, FAX : +82-42-821-1595, E-mail : ejrhee@hanbat.ac.kr

Receipt date : Jan. 23, 2014, Revision date : Mar. 19, 2014
Approval date : Apr. 1, 2014

** Dept. of Computer Eng., Graduate School of Information and Communications, Hanbat National University
E-mail : rabiafyz@hotmail.com

** Dept. of Computer Eng., Graduate School of Information and Communications, Hanbat National University
* This research was supported by the research fund of Hanbat National University in 2011.

depth information and extracted hand based on distance. This method of hand detection is designed in RGB color model [8] which is dependent on illumination. In [9], background is eliminated by depth information, and threshold to segment hand is decided by matching the color of hand center point with the regions in hue histogram. Since the threshold is decided by one point of hand center, it often can not represent the whole hand area. Likewise, Tara [10] segments hand from depth image by applying threshold using anthropometric approach but this method has some limitation of static centroid of all hand. Moreover, this method fails when exact hand location of one image region is overlapping another image region.

In [11], Avinash et al. segment hand by using a hybrid approach based on four components H, S, Cb and Cr of two color spaces, HSI and YCbCr, and morphological operations with labeling are applied to improve the segmentation result of hand from the complex background. However, this approach works with the constant threshold range of H, S, Cb and Cr defined by histogram model. In [12], hand and face are segmented using stereo information to remove background and elliptical skin color model in YCbCr is used to obtain the high skin color detection by using heuristic approach to select the predefined threshold. Pham [13] extracted hand using depth information to remove background and skin-color model is used in LUV color space, which is trained using Gaussian Mixture Model (GMM) but the training process of the model parameter estimation requires high computational complexity. In addition to that, this method is limited to the specific operational environment.

In this study, we propose a method for hand segmentation adaptively to solve the above mentioned problems for all colored races of images captured in any colored background and lightning condition. In order to solve the problem of colored background, we use depth information taken by

stereo camera to extract hand area as nearer to the camera than background, and further objects are excluded as background [13–15].

We approach lightning condition and segmentation of any colored hand by taking the threshold of hand color adaptively. The adaptive threshold is determined by taking the maximum color cluster on hand area, where clusters are determined by matching the color of well distributed points of hand area on the input image, with one of the regions in hue histogram. The histogram is segmented into regions by the shape analysis of histogram on hand area, which is estimated by depth information. By using the adaptive threshold and depth information, this proposed method provides reliable segmentation and detection of hands of all colored races from input image that is captured in any colored background and lightning condition.

The reason of using hue in HSV color model for hand segmentation is to exclude the influence of the brightness and detect hand by merely the color as suggested in our research [16]. But any color model may be used in our method, because it adaptively decides the threshold for hand segmentation.

Fig. 1 shows the overall process of the proposed adaptive hand segmentation method. It consists of three major parts; hand area extraction using depth information by stereo vision, hue regions division by histogram shape analysis, and getting the maximum color cluster to segment hand adaptively.

The organization of this paper is as follows. The proposed algorithm is described in section 2. The usefulness of the proposed technique is demonstrated through simulation in section 3. The conclusion is in section 4.

2. HAND AREA EXTRACTION

Hand area extraction is a challenging task in any colored background, where hand area consists of hand itself and some part of background around hand. To extract the hand area from any colored

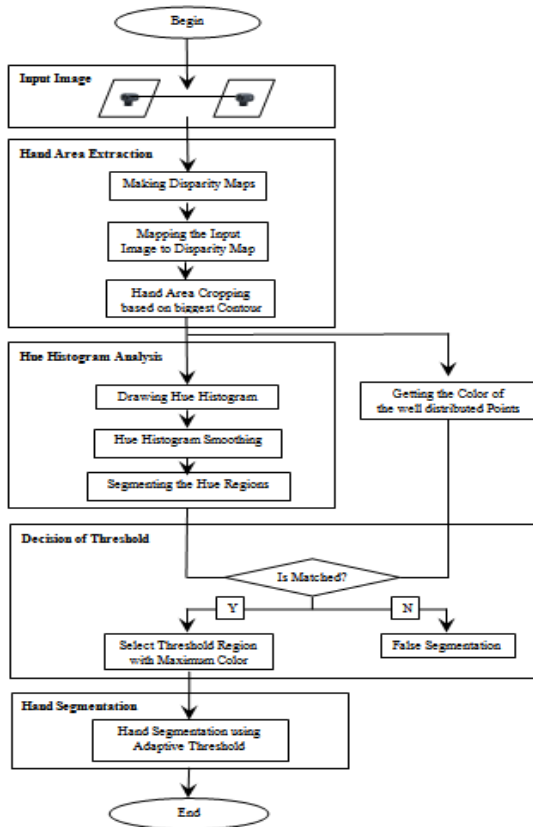


Fig. 1 The process of proposed method.

background, we use depth information taken by stereo camera, where the hand area is the nearer object to the camera than background in almost all images. So the area except hand area is easily excluded as background [13,14].

The Vision from two eyes is defined as Binocular [14], where data is being perceived from each image and mapped by some amount of data. The mapping from two different views is used in biological vision to perceive the depth which is inversely proportional to the disparity. Fig. 2 (a) shows the binocular vision, where f is focal length, A_l and A_r show the points on the left and right image plane, D is the distance between left and right points in the image plane and B shows the base line. Fig. 2 (b) shows the disparity and depth inverse relation where disparity is represented in gray scale to show the brightness of closer or far-

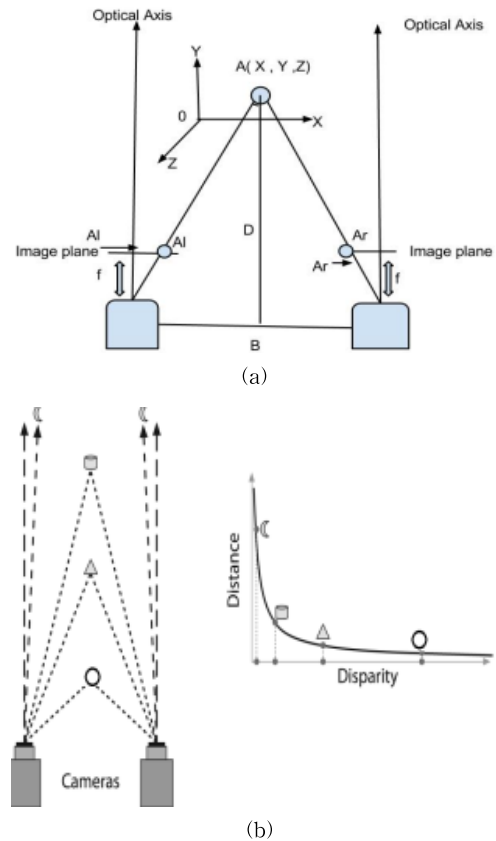


Fig. 2. Stereo vision structure. (a) Stereo vision geometry. (b) Depth vs. Disparity.

these objects, and depth represents the numerical distance of objects from camera. Therefore, the brightest object in disparity will be closer to the camera, having shortest distance and the darkest object would be far from camera by the longer distance.

In order to exclude background, we obtain left and right views by stereo vision at the same time, and the sum of absolute difference based on block matching is used to estimate the depth map. The disparity map of an input image is shown in Fig. 3 (b) that is used to extract the hand area as the nearer to the camera within disparity range 200 to 255, while background is excluded in disparity 200. After extracting the hand from disparity map see Fig. 3 (c), it is mapped to the input image to get the color information on hand as shown in Fig. 3

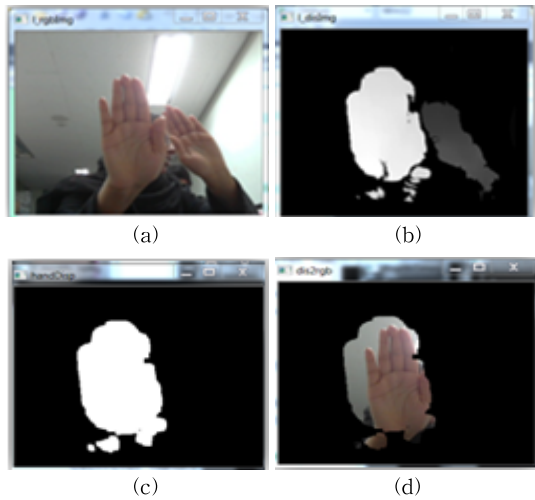


Fig. 3. Hand area extraction. (a) Input image. (b) Disparity map of input image. (c) Disparity of hand area. (d) Hand disparity mapped to input image.

(d) that is used to segment the hand more precisely. Later on, the hand is cropped based on the biggest contour presented in the image shown in Fig. 4 (a).

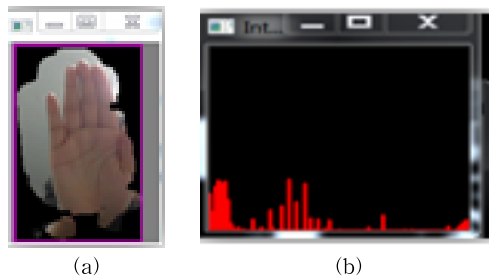


Fig. 4. Hue histogram. (a) Hand cropped image. (b) Hue histogram of figure (a).

3. ADAPTIVE THRESHOLD

The adaptive threshold of hand area is required to work for the hand variations in color, scale, rotation, and in different lightning condition. By the hue histogram analysis we find the peaks and valleys in the hue histogram. And we segment the hue histogram to regions by analyzing these peaks and valleys. We perform color clustering through corresponding colors of the well distributed points on

the hand area with the segmented regions. Then, we get the region with maximum color cluster as threshold to segment hand adaptively.

We use the following steps for determining the adaptive threshold for hand color segmentation as in Procedure 1, and the detail explanation of each step given in following sections.

Procedure 1: Adaptive threshold for hand color segmentation.

1. Drawing hue histogram of cropped hand.
2. Segmenting the color regions by histogram analysis.
3. Matching the color regions with color of well distributed points on the cropped hand.
4. Identifying the color region matching with the maximum color cluster of well distributed points.
5. Deciding of adaptive threshold by taking the start and the end point of the identified color region.

3.1 Hue Histogram Analysis

After the cropping of hand area (see again Fig. 4 (a)), it is used to perform the hue histogram analysis that is the graphical representation of hue presented by counting the hue value in distinct category (known as bins) of hue ranges. The hue histogram of cropped hand image is shown in Fig. 4 (b).

The histogram has variations of colors on hand area, which is smoothed to get the valuable valley and peak for the region segmentation. We perform the histogram smoothing using equation 1, where $n-1$ is the total number of hue, $h(i)$ is the height of histogram on bin i , and $H(i)$ is the height of histogram on bin i after smoothing. The smoothing operation in result controls the quantization level in histogram as shown in Fig. 5.

$$H(i) = (h(i-1) + h(i) + h(i+1)) / 3, i = 1 \text{ to } n-1 \quad (1)$$

After the histogram smoothing, the histogram is separated into regions by finding the peaks and

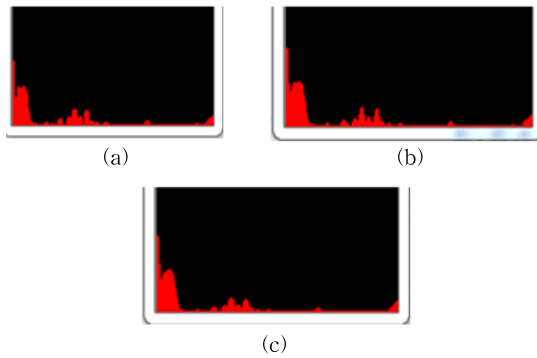


Fig. 5. Histogram smoothing. (a) First time smoothing, (b) Second time smoothing, (c) Third time smoothing.

valleys according to the variation of slopes of the histogram, as shown in Fig. 6. The segmented regions are expressed in the form of vector using equation 2. Where, n is the number of regions, and $\{s_{-p_i}, e_{-p_i}\}$ is the start and end point of the i^{th} region.

$$\begin{cases} \text{Regions} = \{R_1, R_2, \dots, R_i, \dots, R_n\} \\ R_i = \{s_{-p_i}, e_{-p_i}\} \end{cases} \quad (2)$$

Fig. 6 shows the region separation using slope up and slope down approach, where p and v represent the peak and valley respectively, and the valley is the end point of a region. Fig. 7 (b) shows the region segmentation on Smoothed Histogram (see Fig. 7 (a)), where $R_0, R_1, R_2, R_3, R_4, R_5, R_6$ and R_7 represent the number of the segmented regions and s_{-p_i} and e_{-p_i} represent the start and end point of each region.

3.2 Decision of Threshold by color clustering

For determining the threshold adaptively, we get

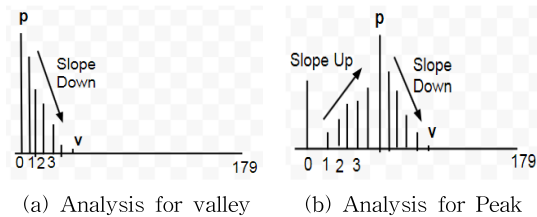


Fig. 6. Peak and valley for region segmentation.

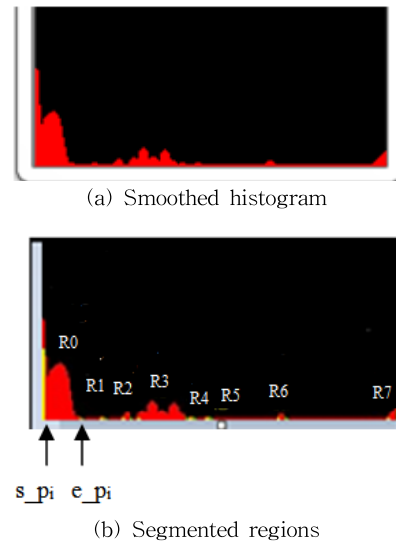


Fig. 7. Hue region segmentation.

nine hue values from well distributed points on the center area of the hand in the input image. In addition to that, four more points are taken around center point of the hand. These points on the hand area are shown in Fig. 8. The adaptive threshold of the hand is estimated by taking the maximum color cluster on the hand area. The clusters are the frequency of colors in regions that are determined by matching the color of the well distributed points on the hand area with each segmented region in the histogram. Fig. 9 shows the color matching of the well distributed points corresponding with the regions in the histogram, where the numbers show the degree of color clustering in hue regions as shown in Fig. 9 (c).



Fig. 8. Well distributed Points.

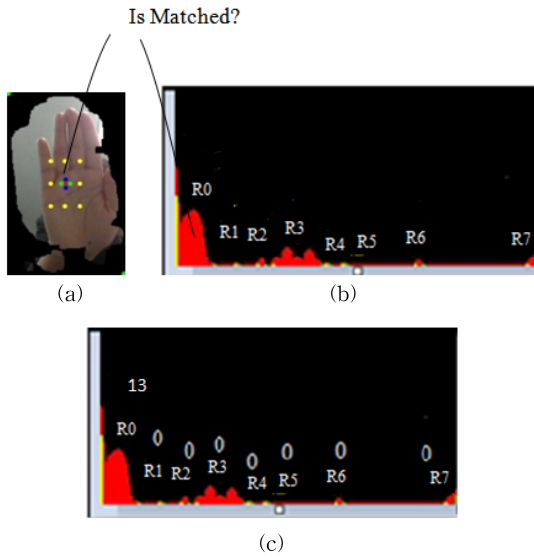


Fig. 9. Color Clustering. (a) Well distributed points on the hand area. (b) Color clustering on each region. (c) Result of color Clustering.

The next step is to select the threshold point that is taken by getting the maximum color cluster in hue histogram, which belongs to one of the segmented regions in the hue histogram. The maximum color cluster is “region 0” in Fig. 9, where the hue value of the start and end point is the color threshold of hand. In other words, the hue value of the start and end point denote the range of colors on the adaptive threshold. After the decision of threshold, hand segmentation is done adaptively from the hand area by the procedure 2.

Procedure 2: Hand segmentation (sp, ep)

```

/* sp, ep : start and end point of the adaptive threshold
/* f(x, y) and g(x, y) are the input image and the binary image

Begin
  if ( (sp == hueStart) and (endPoint_lastRegion == hueEnd))
    if ( sp < f(x, y) < ep) and
      (startPoint_lastRegion < f(x, y) < endPoint_lastRegion))
      then g(x, y) = object; else g(x, y) = background;
  else if ( ep == hueEnd) and (startPoint_firstRegion == hueStart))
    if (sp < f(x, y) < ep) and
      (startPoint_firstRegion < f(x, y) < endPoint_firstRegion))
      then g(x, y) = object; else g(x, y) = background;
  else if (sp < f(x, y) < ep)
    then g(x, y) = object; else g(x, y) = background;
End
    
```

4. EXPERIMENTS AND DISCUSSION

Segmentation experiments have been done to show the usefulness of our proposed method. This algorithm was implemented on Intel PC-i5 in C++ and OpenCV libraries [17] with two Logitech HD 310 Webcams.

The experimental data consisted of 189 sets of extracted hand using depth information. Samples of experiment data are shown in Fig. 10. We used these images to evaluate our proposed hand segmentation method using adaptive threshold, which is explained in section 3. Fig. 11 illustrates the process of hand segmentation on sample image (see again Fig. 3 (a)) that produce the adaptive threshold in two hue regions due to the hue circularity.

The results of experiments are shown in table 1. The examples of good segmentation in Fig. 12 show that our method can segment hand adaptively for all human races in any lightning condition and colored background. And we compared the proposed method with our researches [3,9] and Avinash [11]. These methods provide 81.1% in [3], 85% in [9], and 90.47% in [11] respectively.

The reason for obtaining high rate of good seg-

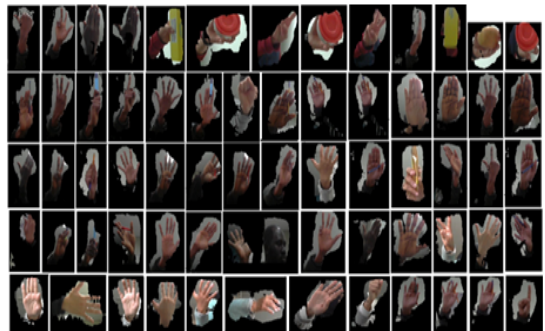


Fig. 10. Examples of experiment data.

Table 1. Experiment result

	Good	False
Number of Images	180	9
Segmentation Rate	0.952	0.047

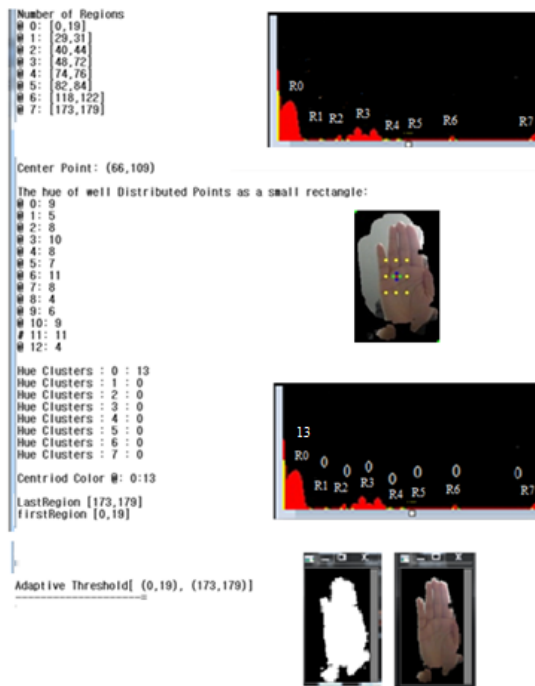


Fig. 11. Illustration of the process of the proposed method.

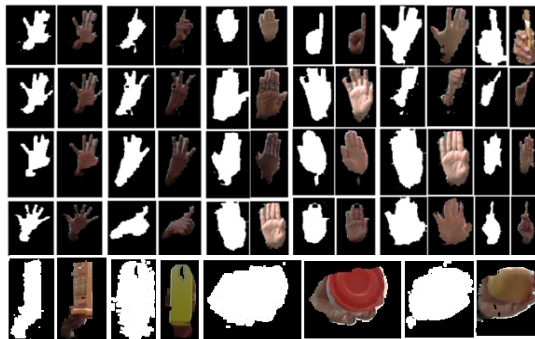


Fig. 12. Examples of good segmentation.

mentation is that the proposed method is based on depth information and adaptive threshold by hue histogram analysis with color clustering. The depth information by stereo vision is used to exclude any colored background within disparity less than 200, and extract the hand area as the nearer object to the camera than background. The appropriate disparity threshold for extracting hand can be adjusted in different application conditions. Based on depth information, we grasp that the hand

would be easy to extract from any background.

To segment hand adaptively, the color clusters are determined by matching the color of the well distributed points on the hand area of the input image, with the regions in histogram. The regions are divided using the approach of slope up and down by the shape analysis of hue histogram on hand area, which is estimated by depth information. The hand threshold is selected by taking the maximum color cluster which is belonged with one of the regions in hue histogram. This way, we can segment all colored types of hand such as black, white, red and brown of all human races in any brightness and colored background. Furthermore, this method is also useful to segment the plain colored objects.

The significant factor in false segmentation is that the hand cannot be distinguished from hand area because the hue values of overall hand area are in the adaptive threshold. The examples of false segmentation are shown in Fig. 13. This problem can be solved by using saturation factor in HSV color model which would distinguish the hand and part of background around hand.

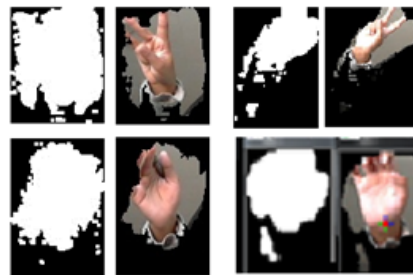


Fig. 13. Examples of false segmentation.

5. CONCLUSION

This paper describes a method for adaptive hand segmentation of various Human races in different lighting conditions, including any colored background. Segmentation of hand having colored background is difficult, because the predefined threshold is dependent on background environment and lightning condition. Moreover, skin values are

not same for all types of hands.

To work with any colored background, we use two input image planes to extract the hand area as the nearer object to the camera than the background. In order to segment hand of all human races in any lightning condition, we determine the adaptive threshold of hand color, that is taken by getting the maximum color cluster which is belonged with one of the regions in hue histogram; the clusters are determined by matching the color of well distributed points on hand area with the regions in hue histogram. These regions are deduced by the shape analysis of hue histogram on hand area, which is estimated by depth information.

The result of the experiments showed the usefulness of the proposed method in segmentation of hand as well as the plain colored objects. Future work includes solving the problem of distinguishing between the hue similarity of hand and the outside of hand area. We also plan to extend the system for hand motion tracking and gesture recognition.

REFERENCES

- [1] D.L. Baggio, S. Emami, D.M. Escriva, K. Ievgen, N. Mahmood, J. Saragih, et al., *Mastering OpenCV with Practical Computer Vision Projects*, Packt Publishing, Birmingham, 2012.
- [2] H. Park, *A Method for Controlling Mouse Movement using a Real Time Camera*, Master's Thesis of Brown University, 2010.
- [3] R. Fayyaz and E.J. Rhee, "Controlling Slides using Hand Tracking and Gesture Tracking," *Proceeding of the 37th Conference of Korean Information Processing Society*, Vol. 19, No. 2, pp. 436-439, 2012.
- [4] S. Alvarez, D.F. Llorca, G. Lacey, and S. Ameling, *Spatial Hand Segmentation using Skin Color and Background Subtraction*, Technical Report, School of Computer Science and Statistics of Trinity College Dublin, TCD-CS-2010-34, 2010.
- [5] L. Chen and C. Grecos, "A Fast Skin Region Detector for Colour Images," *IEEE Conference Publications*, Vol. 2005, No. CP509, pp. 195-201, 2005.
- [6] V. Powar, A. Jahagirdar, and S. Sirsikar, "Skin Detection in YCbCr Color Space," *Proceeding on International Conference in Computational Intelligence*. 2011. <http://research.ijcaonline.org/iccia/number5/iccia1037.pdf> (accessed Oct., 10, 2013)
- [7] S.I. Kang, A. Roh, and H. Hong, "Using Depth and Skin Color for Hand Gesture Classification," *IEEE International Conference on Consumer Electronics*, pp. 155-156, 2011.
- [8] P.W. Hawkes. *Vol. 151 of Advances in Imaging and Electron Physics*, Elsevier, San Diego, Cali., 2008.
- [9] R. Fayyaz and E.J. Rhee, "Adaptive Hand Segmentation using Depth Information and Hue Histogram Analysis," *Proceeding of The Autumn Conference of the Korea Society of Information Technology Applications*, pp. 193-201, 2012.
- [10] R.Y. Tara, P.I. Santosa, and T.B. Adji, "Hand Segmentation from Depth Image using Anthropometric Approach in Natural Interface Development," *International Journal of Scientific & Engineering Research*, Vol. 3, No. 5, pp. 605-609, 2012.
- [11] B.D. Avinash, D.K. Ghosh, and S. Ari, "Color Hand Gesture Segmentation for Images with Complex Background," *International Conference on Circuits, Power and Computing Technologies*, pp. 1127-1131, 2013.
- [12] D. Xu, Y.L. Chen, X. Wu, Y. Ou, and Y. Xu, "Integrated Approach of Skin-color Detection and Depth Information for Hand and Face Localization," *IEEE International Conference on Robotics and Biomimetics*, pp. 952-956, 2011.

- [13] T.C. Pham, X.D. Pham, D.D. Nguyen, S.H. Jin, and, J.W. Jeon, "Dual Hand Extraction using Skin Color and Stereo Information," *IEEE International Conference on Robotics and Biomimetics*, pp. 330–335, 2009.
- [14] N.J. Short, *3-D Point Cloud Generation from Rigid and Flexible Stereo Vision Systems*, Master's Thesis of Virginia Polytechnic Institute and State University, 2009.
- [15] G. Chang, J. Park, C. Oh, and C. Lee, "A Decision Tree based Real-time Hand Gesture Recognition Method using Kinect," *Journal of Korea Multimedia Society*, Vol. 16, No. 12, pp. 1393–1402, 2013.
- [16] B.S. Lee and E.J. Rhee, "Pixel-based Skin Color Detection using the Ratio of H to R in Color Image," *Journal of Information of technology Applications & Management*, Vol. 12, No. 1, pp. 231–239, 2005.
- [17] G.R. Bradski and V. Pisarevsky, "Intel's Computer Vision Library: Applications in Calibration, Stereo Segmentation, Tracking, Gesture, Face and Object Recognition," *Proceeding of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 796–797, 2000.



Rabia Fayyaz

was born in Pakistan, on August 31, 1987. She received her B.S. degree in Double Mathematics and Computer Science, and M.S degree in Geographical Information System (GIS) from University of the Punjab, Lahore, Pakistan in 2008, and 2010 respectively. In 2013, she received her M.S. degree in Computer Engineering from Hanbat National University, Daejeon, Korea. She is currently the PhD candidate in Artificial Intelligence and Computer Lab in the Graduate School of Information and Communications, Hanbat National University, Daejeon, Korea. She is conducting research on Object Detection, tracking and recognition using stereo vision. Her research interests are in image processing, computer vision, pattern recognition and artificial intelligence.



Eun Joo Rhee

was born in Korea, on December 11, 1954. He received his B.S., M.S. and Ph.D. degrees in Electronics Engineering from Chungnam National University, Daejeon, Korea, in 1980, 1984 and 1989 respectively. He was a post-doctoral fellow in Graduate School of Imaging Science and Technology of Tokyo Institute of Technology in Japan from 1994 to 1995, and a visiting professor in Oregon Graduate Institute of Science and Technology in America from 1998 to 1999. Since 1989, he has been with Hanbat National University, Daejeon, Korea, where he is currently a professor in the Department of Computer Engineering. His research interests are in image processing, pattern recognition, computer vision and artificial intelligence.