International Journal of Multimedia and Ubiquitous Engineering
Vol.10, No.6 (2015), pp.341-346
http://dx.doi.org/10.14257/ijmue.2015.10.6.32

Automatic Photography Shooting Using Hand Gesture Recognition

Kwang Baek Kim[1] and Young Woon Woo[2]

[1]*Department of Computer Engineering, Silla University, Busan, 617-736, Korea*
[2]*Department of Multimedia Engineering, Dong-Eui University, Busan, 614-714, Korea*
gbkim@silla.ac.kr, ywwoo@deu.ac.kr

Abstract

As smartphones rapidly improve and as their supply increases, not only high quality images but also creative and diverse functions are expected from the cameras mounted on smartphones. In this paper we propose a method of automatic photography shooting by recognizing common hand gestures used during photography shooting. From the image acquired from a smartphone camera, skin area is extracted based on the skin color data value from YCbCr color space in which characteristics of the color of the skin is well expressed. From the extracted skin area, labeling method is applied to analyze contour data, and then skin object is extracted. By using the location data of the fingers, hand area is extracted from the skin object. Then the hand gesture can be recognized by the camera which takes the photograph automatically. We were able to confirm through experiments that recognition speed was fast even under the low-end environment and that it is more effective than smartphone camera's timers.

Keywords: *Smartphone, Skin color, Labeling, Hand gesture recognition*

1. Introduction

Supply of smartphones is increasing as the market for smartphones is rapidly growing recently. In addition, diverse functional features are added to smartphones and the quality itself is also being enhanced. This change can also be seen in cameras mounted on smartphones. Not only the high quality images but also useful and distinguishing functions of smartphone cameras are currently under active development. There are a few well known image analyzing programs such as face recognition, QR code/bar code recognition, image decoration and language translation. Due to the upgraded function of smartphones, images can be processed in real-time. Thus many real-time image processing software are under research and various image processing programs will be developed [1].

Results from real-time image processing had a limitation depending on the function of the hardware itself. However, high-end smartphones and application of open CV library enabled the smartphones to process the images in real-time with satisfying results. In this paper, we suggest a method of automatic photography shooting by applying real-time image processing to recognize generally used hand gestures during photography shootings.

2. Extraction of Skin Area

In order to recognize hand gesture, skin area extraction must be preceded. Before extracting skin area, skin color data is analyzed. Images acquired from a smartphone camera are in YUV420 color space. In YUV420 color space, Y is luminance component (brightness of the light) whereas U and V are chrominance components. This YUV420

ISSN: 1975-0080 IJMUE

color space is generally used in televisions and video cameras. When CCD or CMOS of a camcorder or camera receives light, electrical signal is generated in the RGB sensor, which is converted to digital signal by analog-digital converter (A/D convert). Then the digital signal is finally converted to YUV signal by RGB-YUV converter. In YUV420 color space, subsampling ratio for Y, U and V is 4:2:0. U and V is each subsampled horizontally and vertically. Then they are effectively set perpendicular to the center of the lines from the image. HD camcorders often use this method [2].

In order to extract skin area, YUV420 color space is converted to YCbCr color space. YCbCr color space can be integrated with YUV420 color space into YUV color space. However, integrating the two color spaces into "YUV color space" is technically incorrect. YUV420 is analog color signal in which U and V each represents blue-luminance and red-luminance. In contrast to YUV420, YCbCr is digital signal in which Cb and Cr each means the deviation of blue-yellow axis and red-cyan axis from gray [3].

The reason of extracting skin area using YCbCr color space is due to these following advantages. First, human visual system is less sensitive to color. Thus, our eyes would not be able to detect small differences in resolution between RGB images and YCbCr images. Secondly, file size of YCbCr images are smaller than that of RGB images. In addition, YCbCr images are less affected by the influence of light source and can recognize face color by just modifying Cr and Cb components. Thus, despite of environmental change during photography shooting, skin area can be easily extracted by modifying values of Cr and Cb components. Skin color data [4] of YCbCr color space is binarized to extract skin area. And then morphology operation is applied to remove noise and connect discontinuous parts in the skin area binarized image. Morphology operation is an operation that modifies the shape of a specific object in an image. This operation is used to describe the shape of the object. The resulting images of extracting skin area using the above methods and the image obtained by applying the morphology operation to binarized image is shown in [5].

After morphology operation, labeling method is applied to categorize skin objects. Then contour data of labeled objects are analyzed to receive information about each skin objects. Labeling method is a method in which adjacent pixels are labeled with the same number whereas pixels that are not connected to each other are labeled with different numbers. During labeling method, edge tracing algorithm [6], which has a short running time, is applied to analyze contour data. Edge tracing algorithm finds edges of labeled numbers. The edge tracing process is described below.

Step 1: if p(x, y) ==1 then insert queue[head]

Step 2: check 5-way, queue[head]

Step 3: Go to Step 2 until no more to go. Save point(x, y) as contour sequences head++

First, if the algorithm meets a labeled pixel $P(x_i, y_i)$, it tracks down a pixel with identical labeling number. The direction of tracking starts from 90 degrees to the left and turns 45 degrees at a time until the direction is heading at 90 degrees to the right. In this way, the algorithm tracks down a pixel with the same labeling number in five directions. During the navigation, if the algorithm encounters a labeled object, that object is included in the queue. Then the algorithm repeats the 5-directional navigation, searching for another pixel with the same labeling number. At the end, if the algorithm meets the starting point, this area is recognized as a proper object whereas if the algorithm fails to meet the starting point, the area is deleted. The advantage of edge tracing algorithm is that it only labels once and yet does not require relabeling mechanism after the first labeling. Skin object is then extracted after removing noise by analyzing contour data that includes the area, circumference and width of the object obtained from edge tracing algorithm.

3. Recognition of Hand Gesture

From the extracted skin objects, hand area and other skin area is separated. The following steps are applied to separate the two areas [7]. Skin objects' centers of gravity are calculated and a circle is drawn for each skin objects around their center of gravity. Additional information about finger location is used to ultimately extract hand area. Center of gravity, *Center(x, y)* of every pixels' coordinates $P(x_i, y_i)$ within a skin object is calculated by using the equation (1).

$$x = \frac{\sum_{i=0}^{k} x_i}{k} \quad and \quad y = \frac{\sum_{i=0}^{k} y_i}{k} \tag{1}$$

In equation (1), k represents the number of pixels. A semicircle is drawn around the center of gravity, *Center(x, y)*. The radius of semicircle is *0.25 x (sum of the height and width of a skin object)*. As shown in Figure 1(b), knuckles and parts of the semicircle drawn on hand object meet at certain points. Therefore, the number of meeting points is equal to the number of unfolded fingers. Using this information of fingers' location, hand object image can be acquired.

(a) Face Object Image (b) Hand Object Image

Figure 1. Images with Semicircle Drawn

However, in Figure 2(a) not only the hand but also the wrist and the arm are extracted together. When wrist and arm areas are extracted together, center of gravity cannot be correctly calculated. Thus, hand and wrist should be distinguished. Then a new center of gravity is measured. In order to distinguish hand and wrist, the width-to-height ratio is used as described in equation (2).

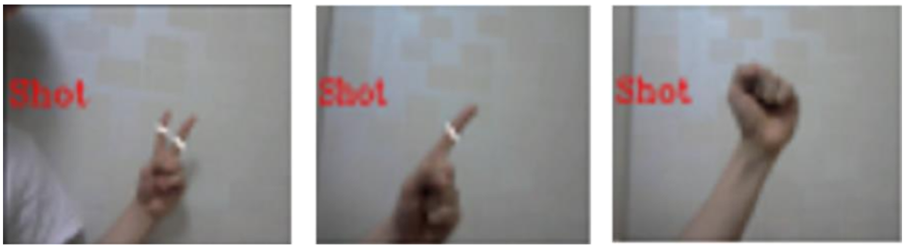
$$if(ObjectHeight > 2 \times ObjectWidth), \quad y = \frac{\sum_{i=0}^{k} y_i}{k} - \frac{ObjectHeight + ObjectWidth}{8} \tag{2}$$

If *ObjectHeight*, height of hand object, is larger than the doubled value of *ObjectWidth*, width of hand object, y value of the center of gravity, *Center(x, y)* is re-measured. The re-measured center of gravity and the semicircle drawn around the new center of gravity is depicted in Figure 2(b).

In equation (1), k represents the number of pixels. A semicircle is drawn around the center of gravity, *Center(x, y)*. The radius of semicircle is *0.25 x (sum of the height and width of a skin object)*. As shown in Figure 1(b), knuckles and parts of the semicircle drawn on hand object meet at certain points. Therefore, the number of meeting points is equal to the number of unfolded

(a) Extracted Hand Wrist Object Image (b) Enhanced Hand Object Image

Figure 2. Image of Semicircle Drawn Around the Improved Center of Gravity



(a) V Gesture (b) Counting Gesture (c) Clenching Gesture

Figure 3. Hand Gestures that are Recognized

From the extracted hand object images, certain hand gestures should be recognized. The standard for selecting hand gestures which are to be recognized is whether the hand gesture is used commonly during photography shooting. Suggested hand gestures include making a V, counting numbers and clutching a fist. These three gestures are described in Figure 3.

In the first gesture, two fingers are used to make a V as in Figure 3(a). The hand should not move for 0.5 seconds to be recognized. In the second gesture, three fingers are unfolded in the beginning. Each finger is folded one by one as in Figure 3(b). In the last gesture, the hand stays outspread for 0.5 seconds as shown in Figure 3(c) then the hand is clenched into a fist. When the suggested hand gestures are recognized, the camera automatically takes a photograph.

4. Experiment and Results

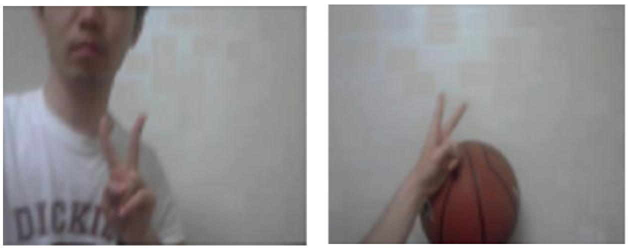
In order to test the method suggested in this paper for automatic photography shooting by hand gesture recognition, smartphone Galaxy S II(Android Platform 2.3(Gingerbread)) mounted with Exynos 4210 1.2GHz Dual-Core and 1GB RAM was used. The program was implemented into the smartphone using eclipse-Android and OpenCV library. The real-time imaging was applied to 8 million pixels HD 1080p camera of Galaxy SII. Figure 4 shows successful recognition of hand gesture and automatic photography shooting. An alarm sign was set on the upper part of the image to notify that the hand gesture was recognized.

(a) Original Images (b) Hand Gesture Recognize Image

Figure 4. Hand Gesture Recognition Results

Author names and affiliations are to be centered beneath the title and printed in Times New Roman 12-point, non-boldface type. Multiple authors may be shown in a two or three-column format, with their affiliations below their respective names. Affiliations are centered below each author name, italicized, not bold. Include e-mail addresses if possible. Follow the author information by two blank lines before main text.

Figure 5 shows incorrect cases of hand gesture recognition. In Figure 5(a), the skin area for the hand overlapped with the other skin area. In the case of Figure 5(b), an object that has a similar value to skin color was in the image.



(a) Skin Area Overlapped (b) Object with Similar Color to Skin Color

Figure 5. Incorrect Hand Gesture Recognition

5. Conclusions

In this paper, as part of developing useful and ingenious functions of smartphone cameras, we suggested a method of automatic photography shooting by hand gesture recognition. In the suggested method, images acquired by smartphone camera were first converted to YCbCr color space. Then skin area was extracted based on the skin color data value. From the extracted skin areas, hand area was extracted using locational information of the fingers. Finally, commonly used hand gestures during photography shooting were recognized to automatically take a picture. Proposed algorithm excluded unnecessary algorithm, enabling real-time processing. In addition, during the experiment, this method was confirmed to be more effective than timer function of smartphones. Further research will be focused on two points; distinguishing skin area from objects with similar color value to skin area and exactly extracting hand area despite overlapping skin objects.

References

[1] J. Y. Kim and H. S. Jun, "Implementation of Real-time Image Processing and Augmented Reality for Smart Mobile Phone," Proceedings of Korea Information Processing Society Conference, Korea, **(2011)**.

[2] Y. H. Choi, "Study on the Extraction Method Camera Motion Parameter," Ph.D. Thesis, Major in Information Technology, Graduate School of Education, Seowon University, **(2004)**.

[3] C. Poynton, "Digital Video and HDTV Algorithms and Interfaces," San Francisco, Morgan Kaufman,**(2003)**.

[4] D. Chai, "Face Segmentation Using Skin-Color Map in Videophone Applications," Proceedings of IEEE Trans. on Systems for Video Technology, vol. 9 no. 4, **(1999)**, pp.551-564.

[5] K. B. Kim and Y. W. Woo, "Image Shooting Using Gesture Recognition," Proceedings of Asia Workshop on Convergence Information Technology of KSCI, Jeju, Korea, vol. 1, December **(2014),** pp.19-22.

[6] OpenCV KOREA, "Available at: http://cafe.naver.com/opencv/."

[7] A. Malima, E. Ozgur and M. Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control," Signal Processing and Communications Applications, vol. 14, **(2006)**, pp.17-19.

Authors

Kwang Baek Kim, he received his M.S. and the Ph.D. degrees in Department of Computer Science from Pusan National University, Busan, Korea, in 1993 and 1999, respectively. From 1997 to present, he is a professor, Department of Computer Engineering, and Silla University in Korea. He is currently an associate editor for Journal of The Korea Society of Computer and Information, and The Open Artificial Intelligence Journal (USA). His research interests include fuzzy neural network and application, bioinformatics and image processing.

Young Woon Woo, he received his B.S. degree, M.S. degree and Ph.D. degree in electronic engineering from Yonsei University, Seoul, Korea in 1989, 1991 and 1997, respectively. Since 1997, he has been a professor in Department of Multimedia Eng., Dong-Eui University, Busan, Korea. His research interests are in the area of artificial intelligence, image processing, pattern recognition and medical information.