International Journal of Engineering & Technology, 7 (2.12) (2018) 25-28

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Facial region detection robust to changing backgrounds

Seok-WooJang [1] *, SiwooByun [2]

[1] *Assistant Professor, Department of Software, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, Republic of Korea*

[2] *Professor, Department of Software, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, Republic of Korea*

**Corresponding author E-mail: swjang7285@gmail.com*

Abstract

Background/Objectives: These days, many studies have actively been conducted on intelligent robots capable of providing human friendly service. To make natural interaction between humans and robots, it is required to develop the mobile robot-based technology of detecting human facial regions robustly in dynamically changing real backgrounds.
Methods/Statistical analysis: This paper proposes a method for detecting facial regions adaptively through the mobile robot-based monitoring of backgrounds in a dynamic real environment. In the proposed method, a camera-object distance and a color change in object background are monitored, and thereby the skin color extraction algorithm most suitable for the measured distance and color is applied. In the face detection step, if the searched range is valid, the most suitable skin color detection method is selected so as to detect facial regions.
Findings: To sum up the experimental results, algorithms have a difference in performance depending on a distance and a background color. Overall, the algorithms using neural network showed stable results. The algorithm using Kismet had a good perception rate for the ground truth part of an original image, and a skin color detection rate was greatly influenced by pink and yellow background colors similar to a skin tone, and consequently an incorrect perception rate of background was considerably high. With regard to each algorithm performance depending on a distance, the closer a distance with an object was to 320cm, the more an incorrect perception rate of a background sharply increased. To analyze the performance of each skin color detection algorithm applied to face detection, we examined how much a skin color of an original image was detected by each algorithm. For a skin color detection rate, after the ground truth for the skin of an original image, the number of pixels of the skin color detected by each algorithm was calculated. In this case, the ground truth means a range of the skin color of an original image to detect.
Improvements/Applications: We expect that the proposed approach of detecting facial regionsin a dynamic real environment will be used in a variety of application areas related to computer vision and image processing.

Keywords: *Intelligent Robot; Facial Region; Dynamic Environment; Changing Background; Ground Truth; Skin Tone; Algorithm Performance.*

1. Introduction

These days, many studies have actively been conducted on intelligent robots capable of providing human friendly service in various areas, such as home service robot, intelligent vehicle, and security system [1-6]. To make natural and friendly interaction between humans and robots, it is required to develop the technology of detecting human facial regions accurately. Particularly, in dynamically changing real surroundings, it is very important for a mobile robot to detect human facial regions robustly [7-11].
Most of conventional facial region detection methods made use of already-known or learned skin color model information to find skin color regions from a captured image and judge if the extracted region is involved in face [12-15]. It is based on the premise that a human skin has a certain range of consistent colors differentiated from many other objects. Such a method is to find an already-known human skin color from a full image. It is simple but has many difficulties.
In other words, a human skin is different depending on individuals and races. Among skin regions, the face used most appears externally so that a skin color can be different depending on special makeup or coloring makeup. Also, an image can have different colors depending on the characteristics or lighting conditions of an optical device photographing an object. To overcome the different colors in artificial or natural environmental changes, new methods tried to be found variously.
In a mobile robot environment, face detection should be robust not only in a static circumstance, but in a dynamic one. For instance, if a robot comes close to or moves away from a person, or follows a person moving in a random direction, the surroundings of the person continue to change. Therefore, such a condition can negatively influence the performance of a system that detects a human face with the use of a skin color. In short, the face detection methods should have good performance in dynamic environmental situations, such as a variable distance between a robot and a human and a dynamic color change in surroundings. Unfortunately, conventional face detection methods mainly focus on static circumstances. There is not much research on dynamically changing circumstances. Therefore, this study proposes a method for detecting facial regions adaptively and robustly in a dynamically changing real environment through the mobile robot-based monitoring of surroundings.
The rest of this paper is as follows.Section 2 introduces a method of monitoring dynamically changing surroundings. Section 3 describes a method of detecting facial regions. Section 4 shows the experimental results of this study. Section 5 explains the conclusion of this study and its future research direction.

2. Background monitoring

Figure 1 illustrates the overview of the proposed face detection algorithm.

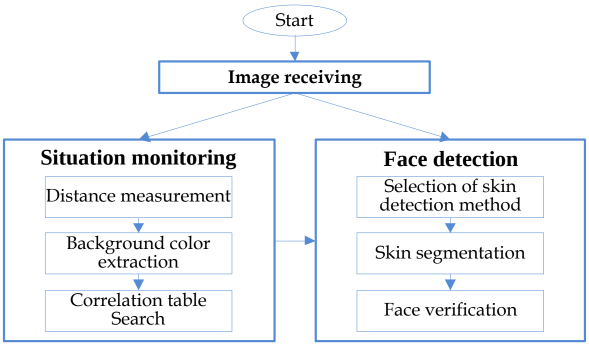

Fig. 1: Overall Flow of the Proposed Approach.

As shown in Figure 1, the proposed system has two steps: circumstance monitoring step and face detection step. In the circumstance monitoring step, a distance between a robot and an object and color information around the object are extracted, and then a relevant range is searched for in correlation table. In the face detection step, if the searched range is valid, the most suitable skin color detection method is selected so as to detect facial regions.

In the proposed method, dynamic surroundings are monitored. Generally, an intelligent robot is capable of moving in a random direction from its current position. Therefore, surroundings of the object, for example a distance between the robot and the object and the background of the object, continue to change. Accordingly, the result of monitoring dynamic surroundings is applied to the face detection step.

In order to measure a robot-object distance, point cloud data, which is 3D depth information entered from a bumblebee stereo camera, is used. In the proposed method, a point cloud sample [16-18] is extracted from the block with a certain size in the center of the input image, as shown in Figure 2. It is based on the assumption that an object is generally located in the center of an image.

$$B_{distance} = \frac{\sum_{i}^{m}\sum_{j}^{m} PC(i,j)}{m \times m} \tag{1}$$

As shown in the Formula (1), the mean of the values is calculated to be used as distance information. In the Formula (1), m means the length and width values of the block located at the center, and PC (i,j) is point cloud information extracted from the block (i,j). Additionally, a robot-object distance can be calculated with the laser scanner installed in a robot, aside from point cloud.

For the background color around an object, four sample blocks are extracted from the areas around the center of the image as shown in Figure 2, and then the mean of the extracted samples is calculated to be used as a background color. Generally, an object is located at the center of an image. Therefore, background samples are extracted from the up and down and left and right areas of the space except for the center.

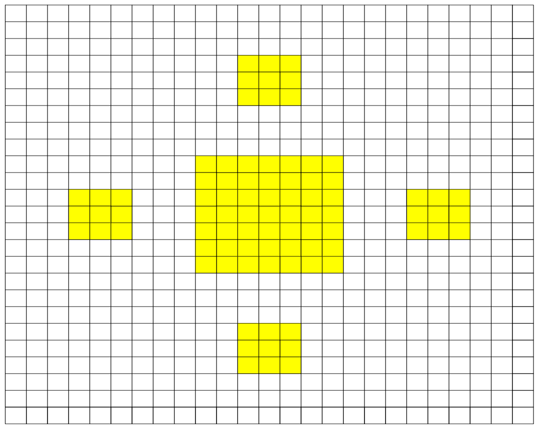

Fig. 2: Extraction of Samples.

The color mean of the extracted sample blocks can be calculated as written in the Formula (2), where n means the length and width values of a block; k is an index of representing four blocks; $H_k(i,j)$ is the hue value of the pixels located at (i,j) of the kth block. In the proposed method, the hue value is used as H, the color factor of HSV color space.

$$B_{color} = \frac{\sum_{k}^{4}\sum_{i}^{n}\sum_{j}^{n} H_k(i,j)}{4 \times n \times n} \tag{2}$$

$$H_k(i,j) = \cos^{-1}\left(\frac{[(R-G)+(R-B)]}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right)$$

After a robot-object distance and a background color are measured, the measured distance is set to one of the four sections between 60cm and 320cm, and the measured color is set to one of five colors (black, orange, pink, yellow, and white) in the distance and color correlation table shown in Table 1. In other words, a combination of the measured distance and color is set to one of 20 split sections M_{pq}. In the next face detection step, a correlation value is used as the basic data to select the most skin color detection algorithm for a robot-object distance and a background color. In the proposed method, a distance between a camera and an object was limited to 60cm-320cm in consideration of the performance of the applied stereo camera, and the mentioned five were used as a background color, for it was hard to take into account all colors actually. How the skin color detection algorithm allocates a value to each item in the correlation table of distance and color is described in detail in Section 4. In the proposed method, there are 20 items correlation table of distance and color. The items are planned to be expanded.

Table 1: Correlation Table

	Black	Orange	Pink	Yellow	White
60-119	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}
120-179	M_{21}	M_{22}	M_{23}	M_{24}	M_{25}
180-239	M_{31}	M_{32}	M_{33}	M_{34}	M_{35}
240-320	M_{41}	M_{42}	M_{43}	M_{44}	M_{45}

3. Adaptive face detection

After a distance and a background color are extracted, a skin region is segmented with the use of the skin detection algorithm to be applied to face detection in the correlation table. The skin detection algorithms used in the proposed method are the face detection method of the emotional robot Kismet developed by MIT [20], the color clustering-based method running in the '15-second frame' product installed in the museum wall [21], the neural network-based method in HSV color space, the neural network-based method in YUV color space, and the lookup table-based method in YUV color space. The five methods are described as A_{Kismet}, A_{clust}, A_{NNHSV}, A_{NNYUV}, and A_{LutYUV}, respectively.

Among them, A_{Kismet} and A_{clust} were developed as applications, and are used in widely popular systems. A_{NNHSV} and A_{NNYUV} algorithms were developed to segment a skin color with the use of MLP (multilayer perceptron neural network) method which is highly accurate in YUV and HSV color spaces. The two methods make use of color clustering and create a color model with the use of three-layer MLP. LutYUV was also developed for efficient color clustering with the use of lookup table in the YUV color space. Color spaces used for face detection and their relevant algorithms are described in detail in the related literature [22-26].

If a distance from an object and background color information in a dynamically changing real environment are able to be extracted, it is possible to select the skin color detection algorithm suitable for the current surroundings and consequently to perform adaptive facial region detection. The proposed method makes use of the skin color detection algorithm in the relevant item M_{pq} of the correlation table of distance and color so as to separate an image into pixels with skin color and with non-skin color. Labeling [27-30] is applied to the pixels judged to be skin color, in order to segment in skin regions and examine if the segmented skin region is involved in face. To do that, the proposed method makes use of EyeMap proposed by Hsu and judges a skin region as face if the region has eyes [12].

$$EyeMap = (EyeMapC\ ) \ AND \ (EyeMapL) \tag{3}$$

$$EyeMapC = \frac{C_b^2 + (255 - C_r)^2 + (C_b / C_r)}{3},$$

$$EyeMapL = \frac{Y(x,y) \otimes g_\sigma(x,y)}{Y(x,y) \oplus g_\sigma(x,y) + 1},$$

In the Formula (3), EyeMapC is the map created for colors, and EyeMapL is the map created for brightness values. The proposed method focuses on selecting the skin color algorithm suitable for dynamically changing surroundings and detecting face adaptively. Therefore, for more details of EyeMap, see references.

4. W6ZBExperimental results

A mobile robot has the stereo Bumblebee camera of Point Gray installed. For the use of the camera, various circumstances were set up and photographed in experiment. The experimental environment has general light and indoor light with 350lux of illumination intensity, and a distance between the camera and a person changes in the limit of 60cm to 320cm. In addition, the background colors of the person change to one of five colors: black, orange, pink, yellow, and white.

Wehave experimented face detection from the images photographed with the changes in a camera-person distance and a backgroundcolor. The first experiment is in the condition where a distance between the camera and a person is 60cm and a background color is black. The second experiment is in the condition where the distance is 70cm and the color is orange. The third experiment is with 80cm in distance and pink in color. The fourth experimentis with 160cm in distance and yellow in color. The last experiment is with 180cm in distance and white in color.

To sum up the experimental results, algorithms have a difference in performance depending on a distance and a background color. Overall, A_{NNHSV} and A_{NNYUV} algorithms using neural network, and A_{LutYUV} algorithm showed stable results. A_{Kismet} algorithm had a good perception rate for the ground truth part of an original image, and a skin color detection rate was greatly influenced by pink and yellow background colors similar to a skin tone, and consequently an incorrect perception rate of background was considerably high. In addition, A_{clust} algorithm was influenced a lot by a background. Just as A_{Kismet} algorithm, this algorithm perceived a background in pink, yellow, and orange as a skin, and therefore its incorrect perception rate was high. With regard to each algorithm performance depending on a distance, the closer a distance with an object was to 320cm, the more an incorrect perception rate of a background sharply increased.

In order to analyze the performance of skin color detection algorithm applied to face detection, we examined how much a skin color of an original image was detected by each algorithm. For a skin color detection rate, after the ground truth for the skin of an original image, the number of pixels of the skin color detected by each algorithm was calculated. In this case, the ground truth means a range of the skin color of an original image to detect. With Photoshop, data were obtained manually. The skin color detection rate R of each algorithm is calculated as written in the Formula (4).

$$R = \frac{number\ of\ \ detected\ \ pixels}{number\ of\ \ ground\ truth\ \ pixels} \times 100 \tag{4}$$

For example, Figure3illustrates the skin color detection rate graphs of five algorithms in the condition where a distance with an object is 130cm and a background color is black.

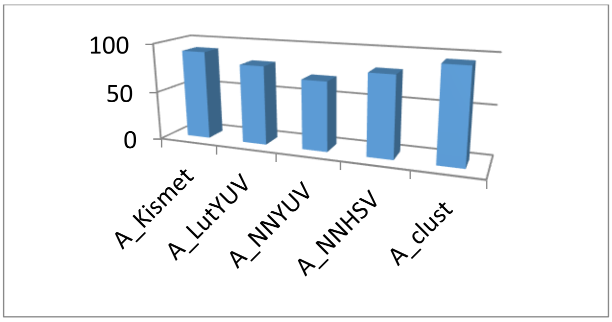

Fig. 3: Comparison of Skin Detection Rate.

According to the experiment, the most excellent skin color detection algorithms depending on a camera-person distance and a background color of a person are summarized in Table 2.

Table 3 shows the skin color detection algorithms selected by the proposed circumstance monitoring-based face detection method depending on a camera-person distance and a background color.

Table 2: Comparison of Face Detection Performance Depending on Circumstantial Changes

	Black	Orange	Pink	Yellow	White
60-119	A_{Kismet}	A_{NNHSV}	A_{NNYUV}	A_{NNHSV}	A_{Kismet}
120-179	A_{clust}	A_{NNHSV}	A_{LutYUV}	A_{NNHSV}	A_{clust}
180-239	A_{clust}	A_{NNHSV}	A_{LutYUV}	A_{LutYUV}	A_{NNHSV}
240-320	A_{NNHSV}	A_{NNYUV}	A_{LutYUV}	A_{NNYUV}	A_{LutYUV}

When the results of Table 2 are considered to be the ground truth data for the skin color detection algorithm selection depending on a circumstance, the results of Table 3 are found to be satisfactory, though there are some errors.

Table 3: Proposed Face Detection

	Black	Orange	Pink	Yellow	White
60-119	A_{Kismet}	A_{NNHSV}	A_{LutYUV}	A_{Kismet}	A_{Kismet}
120-179	A_{Kismet}	A_{NNHSV}	A_{LutYUV}	A_{NNHSV}	A_{clust}
180-239	A_{clust}	A_{NNHSV}	A_{LutYUV}	A_{LutYUV}	A_{Kismet}
240-320	A_{NNHSV}	A_{NNYUV}	A_{LutYUV}	A_{NNYUV}	A_{LutYUV}

5. Conclusion

This study proposed the method of detecting facial regions robustly and adaptively in dynamically changing real surroundings through the intelligent mobile robot-based monitoring of surroundings. In the circumstance monitoring step of the proposed method, a distance between a robot and an object, and information on surrounding colors are extracted, and then a relevant range is searched for in a correlation table. In the face detection step of the method, if the searched range is valid, the most suitable skin color detection method is selected to segment skin regions and thereby, a human facial region is detected robustly.

It will be planned to evaluate the performance of the proposed facial region detection system in comparison with the use of the images photographed in various circumstances. In order to create a more reliable integrated scale automatically on the basis of the monitoring information on surroundings, it will additionally be planned to apply artificial intelligence inference engines like fuzzy [31-34].

6. Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1D1A1A09917838).

References

[1] Li B, ZhangX, FangY, Shi W. Visual servo regulation of wheeled mobile robots with simultaneous depth identification. IEEE Transactions on Industrial Electronics. 2018 January, 65 (1), pp. 460-469.
[2] DuH, WenG,ChengY,HeY, Jia R. Distributed finite-time cooperative control of multiple high-order nonholonomic mobile robots. IEEE Transactions on Neural Networks and Learning Systems. 2017, 28 (12), pp. 2998-3006.
[3] Gui K, LiuH, Zhang D. Toward multimodal human-robot interaction to enhance active participation of users in gait rehabilitation. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2017 November, 25 (11), pp. 2054-2066.
[4] NamHS,KohS, KimYJ,Beom J, LeeWH, LeeSU, Kim S. Biomechanical reactions of exoskeleton neurorehabilitationrobots in spastic elbows and wrists. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2017 November, 25 (11), pp. 2196-2203.
[5] Lotfavar A,HasanzadehS,Janabi-Sharifi F. Cooperative continuum robots: concept, modeling, and workspace analysis. IEEE Robotics and Automation Letters. 2018 January, 3 (1), pp. 426-433.
[6] Dutta A,Dasgupta P. Ensemble learning with weak classifiers for fast and reliable unknown terrain classification using mobile robots.IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2017 November, 47 (11), pp. 2933-2944.
[7] MinaryP, PichonF, MercierD, LefevreE, DroitB. Face pixel detection using evidential calibration and fusion. International Journal of Approximate Reasoning. 91,2017December, pp. 202-215.
[8] ChaudhryS, ChandraR. Face detection and recognition in an unconstrained environment for mobile visual assistive system. Applied Soft Computing. 2017April,53, pp. 168-180.
[9] WuS, KanM, HeZ, ShanS, ChenX. Funnel-structured cascade for multi-view face detection with alignment-awareness. Neurocomputing. 2017January, 221, pp. 138-145.
[10] GunasekarS, GhoshJ,Bovik AC. Face detection on distorted images augmented by perceptual quality-aware features. IEEE Transactions on Information Forensics and Security. 2014, 9 (12), pp. 2119-2131.
[11] Menotti D, Chiachia G, Pinto A, Schwartz WR, Pedrini H, Falcao AX, Rocha A. Deep representations for iris, face, and fingerprint spoofing detection. IEEE Transactions on Information Forensics and Security. 2015, 10 (4), pp. 864-879.
[12] Hsu RL, Abdel-MottalebM, Jain AK.Face detection in color images.IEEE Transactions on Pattern Analysis and Machine Intelligence.2002May, 24 (5), pp. 696-706.
[13] FangJ, QiuG.A color histogram-based approach to human face detection.Proceedings of the International Conference on Visual Information Engineering, 2003July, pp. 133-136.
[14] LeeKM, Component-based face detection and verification.Pattern Recognition Letters. 2008February, 29 (3), pp. 200-214.
[15] BiancoS, GaspariniF, Schettini R. Adaptive skin classification using face and body detection. IEEE Transactions on Image Processing. 2015, 24 (12), pp. 4756-4765.
[16] RafaelMS, AguirreE, MiguelGS.People detection and tracking using stereo vision and color.Image and Vision Computing. 2007June, 25 (6), pp. 995-1007.
[17] HuangP, ChengM, ChenY, LuoH, WangC, Li J. Traffic sign occlusion detection using mobile laser scanning point clouds. IEEE Transactions on Intelligent Transportation Systems. 2017 September, 18 (9), pp. 2364-2376.
[18] Lei H, JiangG, Quan L. Fast descriptors and correspondence propagation for robust global point cloud registration. IEEE Transactions on Image Processing. 2017 August, 26 (8), pp. 3614-3623.
[19] Wang Y, Zhu XX. Automatic feature-based geometric fusion of multiviewTomoSARpoint clouds in urban area. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2015 March, 8 (3), pp. 953-965.
[20] BreazealC,EdsingerA, FitzpatrickP, ScassellatiB.Active vision for sociable robots.IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans. 2001September, 31 (5), pp. 443-453.
[21] Kovac J, PeerP, SolinaF.Human skin color clustering for face detection.Proceedings of theInternational Conference on Computer as a Tool (EUROCON). 2003April, 2, pp. 144-1448.
[22] KakumanuP, MakrogiannisS, BourbakisN.A survey of skin-color modeling and detection methods. Pattern Recognition. 2007March, 30 (3), pp. 1106-1122.
[23] VezhnevetsV, SazonovV, AndreevaA. A survey on pixel-based skin color detection techniques.Proceedings of the International Conference on Computer Graphics and Vision (Graphicon), 2003October, pp. 85-92.
[24] HanZ, TianJ, QuL, Tang Y. A new intrinsic-lighting color space for daytime outdoor images. IEEE Transactions on Image Processing. 2017 February, 26 (2), pp. 1031-1039.
[25] LeeH, KimHS, Kim JI. Background subtraction using background sets with image- and color-space reduction. IEEE Transactions on Multimedia. 2016 October, 18 (10), pp. 2093-2103.
[26] CernadasE, Fernandez-DelgadoM, Gonzalez-Rufino E, CarrionP.Influence of normalization and color space to color texture classification. Pattern Recognition. 2017 January, 61, pp. 120-138.
[27] MaggioriE, TarabalkaY, CharpiatG, Alliez P.High-resolution aerial image labeling with convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing. 2017 December, 55 (12), pp. 7092-7103.
[28] KumarB, Dikshit O. Spectral contextual classification of hyperspectralimagery with probabilistic relaxation labeling. IEEE Transactions on Cybernetics. 2017 December, 47 (12), pp. 4380-4391.
[29] HirakawaT, TamakiT, KuritaT, RaytchevB, KanedaK, WangC, Najman L.Tree-wise discriminative subtreeselection for texture image labeling. IEEE Access. 2017 July, 5, pp. 13617-13634.
[30] PaisitkriangkraiS, SherrahJ, JanneyP, Hengel A.Semantic labeling of aerial and satellite imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2016 July, 9 (7), pp. 2868-2881.
[31] ZhouT, IshibuchiH, Wang S. Stacked-structure-based hierarchical takagi-sugeno-kangfuzzy classification through feature augmentation. IEEE Transactions on Emerging Topics in Computational Intelligence. 2017 December, 1 (6), pp. 421-436.
[32] Nnolim UA. Improved partial differential equation-based enhancement for underwater images using local-global contrast operators and fuzzy homomorphic processes. IET Image Processing. 2017 November,11 (11), pp. 1059-1067.
[33] ChenX, HuJ, WuM, Cao W. T–S fuzzy logic based modeling and robust control for burning-through point in sintering process. IEEE Transactions on Industrial Electronics. 2017 December, 64 (12), pp. 9378-9388.
[34] Chen SM, Chen ZJ. Weighted fuzzy interpolative reasoning for sparse fuzzy rule-based systems based on piecewise fuzzy entropies of fuzzy sets. Information Sciences. 2016February, 329, pp. 503-523.