*Proceedings of the 2000 IEEE/RSJ
International Conference on
Intelligent Robots and Systems*

Robust Face Detection Against Brightness Fluctuation and Size Variation

K. Hidai H. Mizoguchi K. Hiraoka M. Tanaka T. Shigehara T. Mishima

Department of Information and Computer Sciences (DICS),
Saitama University
255 Shimo Okubo, Urawa 338-8570, Japan
{hidai | hm | hira | sigehara | mtanaka | mishima}@me.ics.saitama-u.ac.jp

Abstract

This paper proposes a novel face detection method intended to be used for practical intelligent environment and human-interactive robot. Face detection and recognition are very crucial for such applications. However, in the real situation, it is not easy to realize the robust detecting function because position, size, and brightness of face image are much changeable. The proposed method solves these problems by combining correlation-based pattern matching, histogram equalization, skin color extraction, and multiple scale images generation. The authors have implemented a prototype system based upon the proposed method and conducted some experiments using the system. Experimental results support effectiveness of the proposed idea.

1. Introduction

Based upon increasing demand for better interface and security, significance of human face recognition by machine has raised more and more in recent years. Typical application of the face recognition is intelligent environment and human interactive robot. Towards future aging society, there are expectation and social demand for a robotic environment that is possible to interactively collaborate and support human. Research activities on such environment have increased[1]-[8]. If the intelligent environment could recognize human face, quality of its interaction with human would be improved more properly. Thus face recognition function by machine is very crucial.

Conventional face recognition technology[9][10] is possible to identify face image at accuracy over 99%, if the face is located at specified position and scaled at predefined size in a scene. That is the case in which frontal face image can be taken with simple or no background under well controlled lighting/illuminating condition and the face area occupies large amount part in the scene.

But in case of the real situation, such good condition cannot be obtained. There is a possibility that the face may not be included in the scene. Even if included, position of the face is not fixed. Brightness of the image may be so fluctuant. Size and orientation of the face may vary quite much. Thus, to use the current face recognition technology in the real situation practically, a robust method to detect and extract the face in a scene is keenly required.

Therefore this paper proposes a novel face detection method intended to be used for practical intelligent environments. The proposed method is robust against the variations of face position, size, and brightness. And the method is fast enough for robotic applications. Key idea of the method is combination of skin color extraction, correlation based matching, and multiple scale images generation. This combination realizes advanced features of the method described above. Based upon the proposed idea, the authors have implemented a prototype system. Utilizing the prototype, we have conducted some experiments to confirm feasibility of the idea. Experimental results prove the feasibility and effectiveness of the proposed idea.

Section 2 discusses problems in detail and solutions to them. Section 3 describes a technique to reduce search space utilizing skin color extraction. Section 4 describes an implementation based upon the above solutions. Experiments using the implementation are also described. Section 5 is conclusion.

2. Problems and Solutions

As described above, there are three problems. These are position variation, size variation and brightness fluctuation.

2.1. Position Variation

In the real situation, it is unknown whether a face exists or not in a scene. Even if it exists, its location is not pre-defined and also unknown. System must detect the face at anywhere in the scene.

The proposed method solves this problem by utilizing correlation-based pattern matching. The method performs the matching process for whole input image scanning from the upper left corner to the lower right as shown in fig. 1. Success rate of the detection heavily depends on the template image of the matching. In case a particular person is to be detected, the most suitable template is an averaged image of his or her faces. On the contrary, in case of an arbitrary person, an averaged image of various faces of different persons is the most suitable. And besides, if several template images of various face orientations are prepared, it will be possible to realize orientation independent face detection.

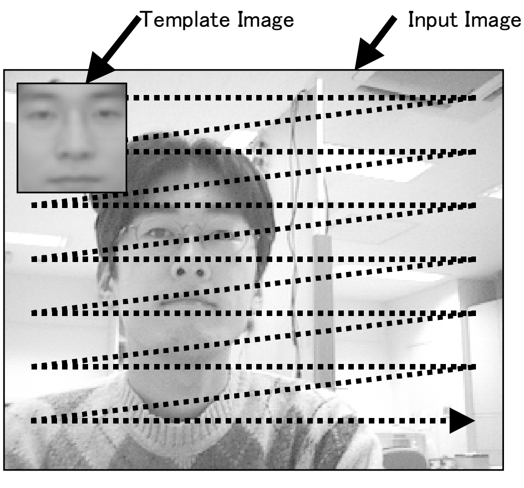

Fig. 1 Pattern Matching Process for Whole Image

In the matching process, correlation score is calculated as follows.

$$cor = \frac{A \cdot B}{\sqrt{A \cdot A} \times \sqrt{B \cdot B}}$$

Where, A and B are image vectors and cor is the correlation score between A and B.

2.2. Size Variation

As mentioned above, problem of the size variation is quite serious for automatic face detection. If the size of a face to be detected is so different from the size of the template, the pattern matching fails to detect the face. This is because the correlation based matching process calculates low similarity score and thus is impossible to detect the face candidate.

In order to overcome the problem, the proposed method generates multiple scale images of an original input image and performs the matching process against these generated images (fig. 2). If a face exists in the input image, there must be an generated image in which the face size is similar to the template size. Thus the matching process calculates high similarity score and succeeds to detect the face.

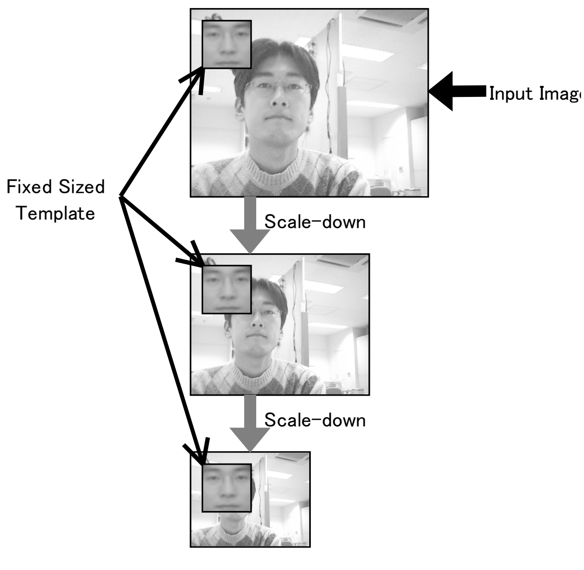

Fig. 2 Multiple Scale Images

To generate a scaled-down image, simple discrete sampling is not adequate. Average of neighbor pixels is better and adopted in this paper. Though the simple discrete sampling is fast, it spoils features of the original image. On the contrary, the averaged sampling is slow but retains features. Fig. 3 illustrates comparison of the two sampling methods.

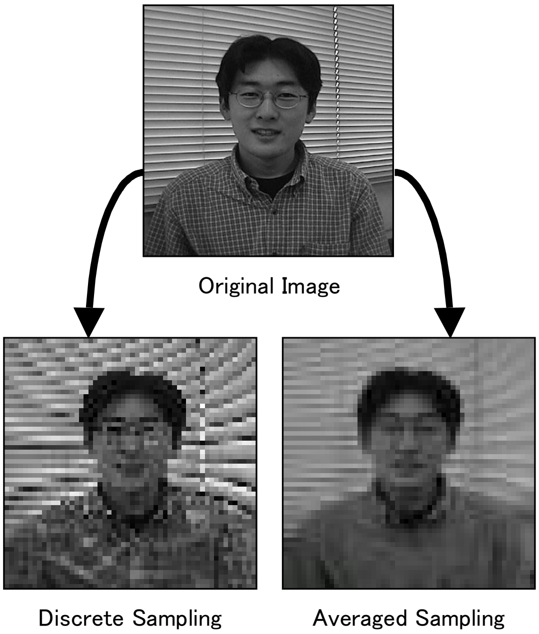

Fig. 3 Comparison between Discrete Sampling and Averaged Sampling

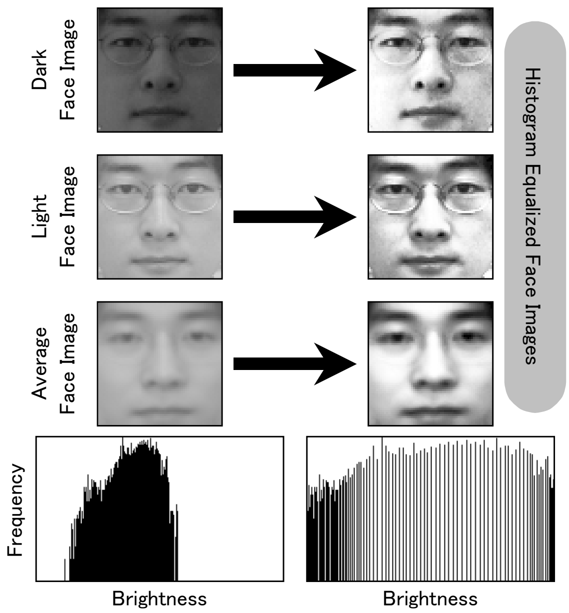

Fig. 4 Histogram Equalization

2.3. Brightness Fluctuation

When brightness of the input image is so different from the template face image, the matching process calculates low similarity score even though both images have similar features. This is serious problem in the supposed real situation because light condition is so fluctuant.

To solve the problem, the proposed method utilizes histogram equalization as a preprocess for the pattern matching. The equalization absorbs the brightness fluctuation. As shown in fig. 4, the histogram equalization expands and flattens brightness values into full range. Thus brightness is normalized and features are emphasized.

3. Reduction of Computation

In the previous section, three problems and their solutions are described. These solutions must contribute to robustness of the face detection, but they consume much time because they process whole input image. These solutions may be usable for *offline* tasks. However, for the intended applications, such as intelligent environment and human interactive robot, the face detection is a time-critical *online* task and fast execution is required. For example, if it took minute order time for an intelligent environment to detect human face, it is meaningless for interactive human support. Thus the above solutions cannot be applied simply to the time-critical *online* task.

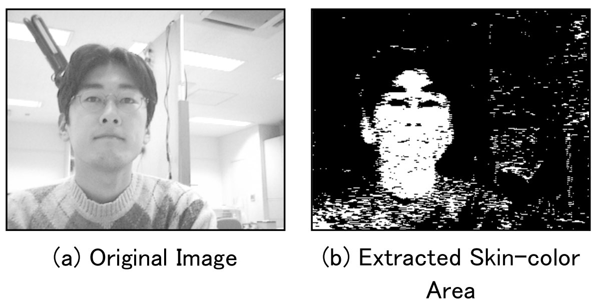

Fig. 5 Skin Color Area Extraction

To solve the problem, the proposed method utilizes skin color extraction technique to reduce search space and time for the face detection. Skin color area is a candidate of face area. Search space for the correlation-based pattern matching can be limited to the extracted skin color area. Another time consuming process, the histogram equalization, can be also performed to only the skin color area, instead of the whole image.

Color representation used for the skin color extraction is Y-Cb-Cr representation, instead of R-G-B which is commonly used for color image input/output and storage. In the Y-Cb-Cr representation, components Cb and Cr denote chromaticness and are independent from the brightness component, Y. Thus, by utilizing Cb and Cr components, brightness independent skin color extraction can be realized.

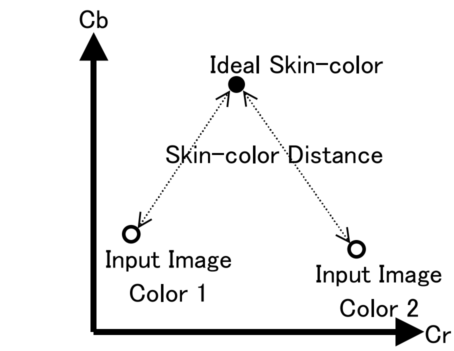

Fig. 6 Skin Color Distance on Cb-Cr plane

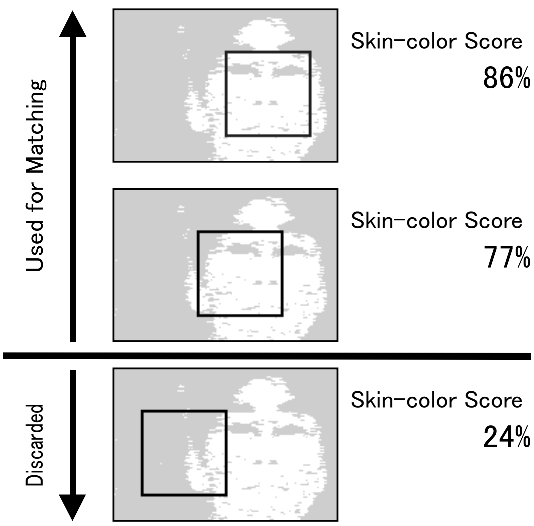

Fig. 7 Skin Color Score (Threshold is 50%)

The process of the skin color extraction is as follows. Firstly, we obtain an ideal skin color by averaging pre-captured face images. The ideal skin color is represented as a point on Cr-Cb plane. Secondly, at each pixel of an input image, we calculate *skin color distance*, a distance from the ideal point to an input image color on the plane as shown in fig. 6. Lastly, by thresholding the skin color distance at each pixel, we extract a potential skin color area.

To limit the search space for the matching within the extracted skin color area, here we introduce *skin color score*. It is a percentage of the number of pixels included in the skin color area to the number of whole pixels of the search area as shown in fig. 7. The pattern matching is only performed if the score is higher than the pre-defined threshold.

4. Implementation and Experiments

To confirm possibility to realize the proposed method, the authors have implemented a prototype system based upon the method. Fig. 8 shows examples of face detection done by the prototype. Fig. 9 illustrates overview of the process performed by the system.

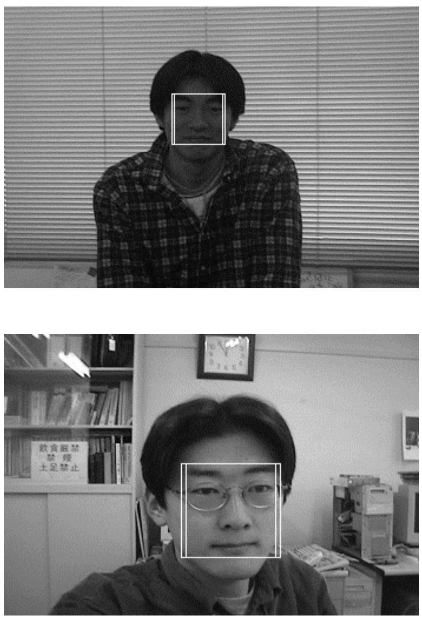

Fig. 8 Examples of Face Detection

Utilizing the prototype, some experiments are conducted to confirm feasibility and effectiveness of the proposed method. For this purpose, success rate and execution time of the detection are measured. Effect of the histogram equalization is also measured.

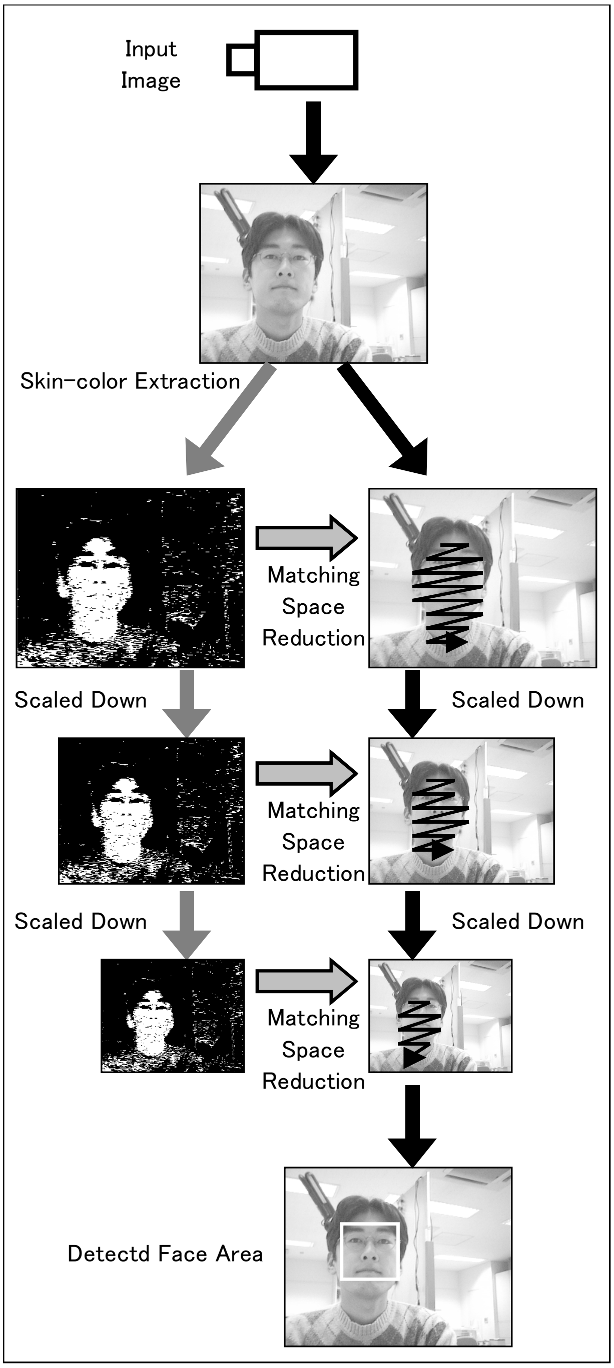

Fig. 9 System Overview

Table 1 describes experimental condition. Input images contain face of various positions, sizes and brightness. Template image for the matching is the averaged face images of various people. If right area is detected, that is an area similar to the template image is detected, the detection is judged as "success". Success rate and execution time are shown in table 2 and table 3 respectively.

Table 1 Experimental Condition

Input Image Size	160×120 pixels
Multiple Scale Images	10 scales
Skin Color Score	50 %
Number of Input Images	48
OS	Linux 2.2
CPU	Pentium II 333Mhz

Table 2 Experimental Result (Success Rate)

Template Image Size (Pixels)	Success Rate of Detection (%)	
	Without H.E..	With H.E.
16×16	21.1	68.2
24×24	57.9	95.8
32×32	84.2	97.9
48×48	78.9	93.7

(cf. H.E.: Histogram Equalization)

Table 3 Experimental Result (Execution time)

Template Image Size (Pixels)	Processing Time (seconds)
16×16	0.762
24×24	0.844
32×32	0.851
48×48	0.853

Table 2 and fig. 10 show that the histogram equalization contributes to higher success rate. This result supports that the histogram equalization is effective against the brightness fluctuation. The table and figure also show that the larger template results in the higher success rate. However the larger template consumes much execution time. And too large sized template, for example 48×48 in this case, decreases the success rate because smaller sized face image cannot be detected.

Table 3 shows that the execution time is sub-second order for each sized template. This performance is sufficient for some applications but insufficient for other applications. Especially for the real-time task, ten to hundred times speedup is required. Since there must be some parallel nature in the multiple scale images generation, the histogram equalization, the skin color extraction and the correlation based pattern matching, speed of the proposed method can be improved by introducing parallel processing.

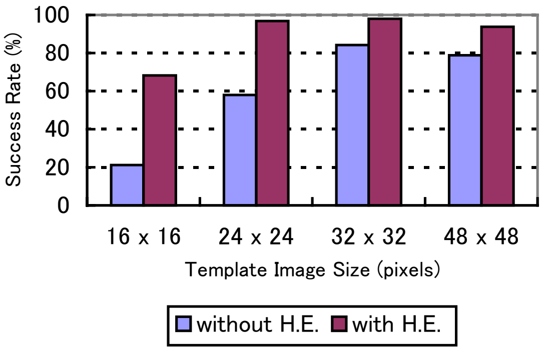

Fig. 10 Effect of Histogram Equalization(H.E.)

5. Concluding Remarkss

In this paper, a novel face detection method is proposed and presented. Since the method is to be used for intelligent environments and/or practical human-robot interaction, it is designed to be robust against brightness fluctuation of face image and variations of the face position and size. The method also has an advantage of fast execution speed. These advantages are realized by combining skin color extraction, correlation based matching, and multiple scale images generation.

A prototype implementation is presented. It confirms possibility to realize the proposed idea. The authors conduct some experiments utilizing the implemented prototype. Experimental results show over 90% success rate of detection and sub-second order execution time. They prove feasibility and effectiveness of the proposed method. Connecting to face recognition program and constructing a face detection and recognition system usable in the real situation are future works. Realizing a human-interactive intelligent environment with the proposed face detection capability is also a future work.

Acknowledgment

This work has been partly supported by CREST of JST (Japan Science and Technology) 279102. The work also has been supported in part by Grant-in-Aid for Scientific Research (B) of JSPS.

References

[1] Joo-Ho Lee, Guido Appenzeller, Hideki Hashimoto, "Physical Agent for Sensored, Networked and Thinking Space", Proc. of ICRA'98, pp.838-843, 1998.

[2] A.Pentland, "Looking at People: Sensing for Ubiquitous and Wearable Computing", Trans. on PAMI, Vol.22, No.1, pp. 107-119, 2000.

[3] T. Sato, et al., "Robotic room: Symbiosis with human through behavior media", Robotics and Autonomous Systems, No.18, pp.185-194, 1996.

[4] H. Mizoguchi, et al.,"Robotic office room to support office work by human behavior understanding function with networked machines", IEEE/ASME Trans. on Mechatronics, Vol.1, No.3, pp.237-244, 1996.

[5] M.C.Torrance, "Advances in Human-Computer Interaction: The Intelligent Room", Working Notes of the CHI95 Research Symposium, 1995.

[6] A. Pentland, "Smart Rooms", Scientific American, pp.54-62, 1996.

[7] H. Asada et al., "Total Home Automation and Health Care/Elder Care", Tech. Rep., Dept. of Mech. Eng., MIT, 1996.

[8] R. P. Picard, Affective Computing, MIT Press, 1997.

[9] R. Chellappa, C. Wilson, and S. Sirohev, "Human and Machine Recognition of Faces: A Survey", Proc.IEEE, Vol. 83, No. 5, pp.705-740, 1995.

[10] T. Kurita, K. Hotta, and T. Mishima, "Scale and Rotation Invariant Recognition Method Using Higher-Order Local Autocorrelation Features of Log-Polar Image", Proc. of ACCV'98, Vol. II, pp.89-96, 1998.

[11] H. A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection", IEEE Trans. on PAMI, Vol. 20, No.1, pp.23-38, 1998.

[12] Q. B. Sun, W. M. Huang, and J. K. Wu, "Face Detection Based on Color and Local Symmetry Information", IEEE Proc. of FG'98, pp.130-135, 1998.

[13] K. Hotta, T. Kurita, and T. Mishima, "Scale Invariant Face Detection Method using Higher-Order Local Auto-correlation Features extracted from Log-Polar Image", Proc. of FG'98, pp.70-75, 1998.

[14] T. Mori et al.,"Action Recognition System based on Human Finder and Human Tracker", Proc. of IROS'97, pp.1334-1341, 1997.