

## Detection of Face Direction by Using Inter-Frame Difference

Bongseog Jang<sup>1</sup> and Sang-Hyun Bae<sup>2†</sup>

### Abstract

Applying image processing techniques to education, the face of the learner is photographed, and expression and movement are detected from video, and the system which estimates degree of concentration of the learner is developed. For one learner, the measuring system is designed in terms of estimating a degree of concentration from direction of line of learner's sight and condition of the eye. In case of multiple learners, it must need to measure each concentration level of all learners in the classroom. But it is inefficient because one camera per each learner is required. In this paper, position in the face region is estimated from video which photographs the learner in the class by the difference between frames within the motion direction. And the system which detects the face direction by the face part detection by template matching is proposed. From the result of the difference between frames in the first image of the video, frontal face detection by Viola-Jones method is performed. Also the direction of the motion which arose in the face region is estimated with the migration length and the face region is tracked. Then the face parts are detected to tracking. Finally, the direction of the face is estimated from the result of face tracking and face parts detection.

**Keywords:** Learner Concentration, Face Detection, Motion Direction, Inter-frame Difference

### 1. Introduction

Image processing is popularized by improving performance of algorithm and processor. In the application regards, image processing is applied for auto focus of digital camera, tracking car and tracking human. Human recognition is used for security such as personal identification and monitoring system implementation. In recent years, human recognition begins to be used for market research. Efficiency of advertisement is estimated from face data in video. This video is recorded camera installed poster. Moreover, Efficiency of this measurement is improved by identification of age group and sex. And furthermore, recommendation of merchandise is performed based on result of the identification.

Face detection on first frame and face tracking is required for estimation of face direction for multiple target. Viola-jones<sup>[1]</sup> method can detect face from image in

high speed and accuracy. However, this method cannot detect the face if its direction is different from the learned one. Furthermore, this method requires a large amount of training data. Mean-shift tracking is based on color information<sup>[2-5]</sup>. This method is robust for differences of the direction. But in case the background color is similar to the training target, there is some possibility of failures.

Meanwhile, the estimation system of concentration degree using facial expression and movement of face is researched for the teacher training<sup>[6,7]</sup>. Estimation system for a student can determine the concentration degree from gaze and eye condition. Measurement of teaching skill requires estimation of concentration degree for plural students. The number of student is limited by number of camera when this estimation system is used for measurement of teaching skill.

This paper proposes the detection method of face direction in class. This system executes face detection, face tracking and face parts detection in video recorded students in classroom. In first frame of video, frontal face region is detected by Viola-jones method. Color information and movement direction is used for face tracking and estimation of direction. The movement direction and distance is estimated by result of frame

<sup>1</sup>Mokpo National University, 61 Dorim, Moo-An, Jena-Nam, 534-729 Korea

<sup>2</sup>Chosun University, 309 Phimun-daero, Dong-gu, Gwangju, 501-759 Korea

<sup>†</sup>Corresponding author : shbae@chosun.ac.kr

(Received: April 15, 2016, Revised: June 14, 2016,

Accepted: June 25, 2016)

subtraction method. And the face direction is estimated from location of face parts and the movement direction.

## 2. Related Works

The reference<sup>[8]</sup> is the technique to detect each face part using that a brightness level of eyes is lower than other parts and that a mouth is redder than other parts. It is hardly affected by the direction of the face, but it is less accurate.

There is a method to perform template matching by light and shade image, but this method is not well worked in case of the change of directions. This method uses a template made from 4 direction surface characteristic quantities, which offer a good result in case of the facial movement and the individual difference of face direction.

When performing the extraction of the direction of face in the video object, it is necessary to track the face region. In the technique of the object tracking, there are techniques and Mean-Shift tracking<sup>[2-5]</sup> using the optical flow<sup>[9]</sup>, etc. The method using the optical flow is able to perform the tracking with a high accuracy when an object is moving in parallel. However, the estimation accuracy of tracking is decreased when the direction of target object changes.

In addition, depending on the technique to determine the optical flows, there is a problem of processing time increment. Mean-Shift tracking is a method of performing object tracking using color information, and the faster processing speed can detect the changes in the orientation of face. However if the background color to detect is similar to the target object, it can be sometimes tracking failed.

## 3. Inter-frame Difference

Inter-frame difference is a method for detecting a moving object, and is to extract the area within the moving object using three successive images. As shown in Fig. 1, the difference between 2 frames from the image of 3 images  $\{f_{c-2}, f_c, f_{c+2}\}$  is taken in each, then the threshold processing is performed. And it converts into the binary image for the result.

Concretely, the result of threshold processing by the difference of image  $f_{c-2}$  and  $f_c$  is  $B_{c-2}$  in equation (1). Also it carries out the threshold processing by the dif-

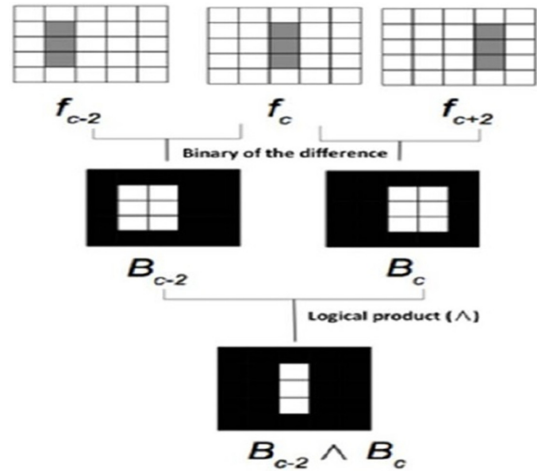


Fig. 1. Inter-frame subtraction method.

ference of image  $f_c$  and  $f_{c+2}$ . The result is  $B_c$  in equation (2).

$$B_{c-2}(x,y) = \begin{cases} 1 & |f_{c-2}(x,y) - f_c(x,y)| > \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

$$B_c(x,y) = \begin{cases} 1 & |f_c(x,y) - f_{c+2}(x,y)| > \text{threshold} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

The results obtained using logical product (AND,  $\wedge$ ) of the difference  $B_{c-2}$  and  $B_c$  is a common area of the difference  $B_{c-2}$  and  $B_c$ . And it represents the areas where the moving object exists in the image  $f_c$ . In the above equations, *threshold* is a binary value.

The movement of direction and face region of the motion are obtained from difference  $B_{c-2}$  and  $B_c$ . The following shows the processing procedure for each pixel in the face region.

① Whether it satisfies the condition  $(B_{c-2} \wedge B_c) = 1$  in coordinate  $(x, y)$  is confirmed.

② The condition is confirmed  $(B_{c-2}=1) \wedge (B_c=0)$  in the pixel which adjoins for the relative position if the condition of ① is satisfied.

③ Whether it satisfies  $(B_{c-2} \wedge B_c) = 1$  at the coordinate which adjoins with the coordinate which satisfied the condition of ② reverse-directional is confirmed.

④ Transfer to the coordinate and confirmation of the condition which adjoin for the direction equal to ③ are repeated, until  $B_{c-2} = 0$  is satisfied.

⑤ The number of the line is counted, when the coor-

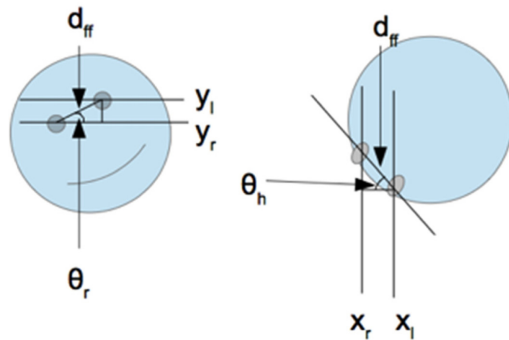


Fig. 2. Estimation of face angle by eye location.

dinate which satisfied the condition of ④ satisfies the condition  $B_C=1$ .

⑥ It moves to the coordinate as  $(B_{C-2}=0) \wedge (B_C=0)$ .

⑦ The movement from condition ⑤ to condition ⑥ is recorded, when the coordinate in the transfer tip is located on the outside in the face region. It is more renewed in the high value, when there is recorded value.

The difference between inter-frames is used to determine the movement region. The estimated region of difference  $B_{C-2}$  is the time series data which are a starting point for the estimation. In the above figure, perpendicular and horizontal line that it passes the movement region makes the region of difference  $B_C$ , which is used as an end point. And the transferring direction of face region is judged from the number of segments. That the movement is occurred is determined according to its value calculated by the moving positions. The movement as the end point in terms of the transferred by processing (6) is obtained when the end point is located at the outside of the face region. In this case, the movement is recorded by every its relative positions, and the largest value calculated from the origin is determined as the movement points.

#### 4. Detection of Face Direction

For the face direction detection, we use the coordinate of eyes. Using the number of detected eyes, the estimation of a face in previous processing is applied. After the performed face detection in the previous single face case, we make a generous decision process from the number of eye domain detected as the standard method in Table 1.

In case of the number of the eye is 0 or 1, the direc-

Table 1. Estimation of face direction by result of eye detection

Number of eye	Direction of face
2	Frontal face
1	Face suitable for either about 90 degree (right and left)
0	Face suitable for either about 90 degree (top and bottom)

tion is specified by the comparison of face conditions. The face region has the stopped location in the past and it coordinates with present face region. Specific description is as follows.

- When the states of the face of the past and present are different, it can be judged that the change of the direction occurs and the direction of the face is determined from the comparison result of the coordinates.

- When the states of the face is the same, and then the state of the face of the past and present are compared by the coordinate of the face domain. In this case, the judgments are different as follows.

- When the distance of the coordinate is more than width of the face domain, 1/2 of the height, it is judged that a face does a translation.

- When the distance of the coordinate is less than width of the face domain, 1/2 of the height, it is judged that a face does not translated.

In case of the number of the eye is 2, the rotation angle around the axis and a horizontal angle and line of sight of the face are estimated using the distance of the eyes obtained by the face parts detected from the first image.

By comparing the x-coordinate of the eyes, the larger value is becoming the left eye coordinates  $(x_l, y_l)$  and the smaller value is becoming the right eye coordinates  $(x_r, y_r)$ . With the coordinates, the rotation angle  $\theta_r$  of the face by equation (3) and the lateral angle  $\theta_h$  by equation (4) are calculated respectively. Since  $\theta_r$  is the counter-clockwise direction, if the value is positive, it is viewed from the camera.

$$\theta_r = \sin^{-1}\left(\frac{y_l - y_r}{d_{ff}}\right) \quad (3)$$

$$\theta_h = \cos^{-1}\left(\frac{x_l - x_r}{d_{ff}}\right) \quad (4)$$

Since  $\theta_n$  cannot determine the direction from the eye coordinates, that the previous direction is not changed is assumed.

## 5. Experiments and Results

### 5.1. In Case of 1 Examinee

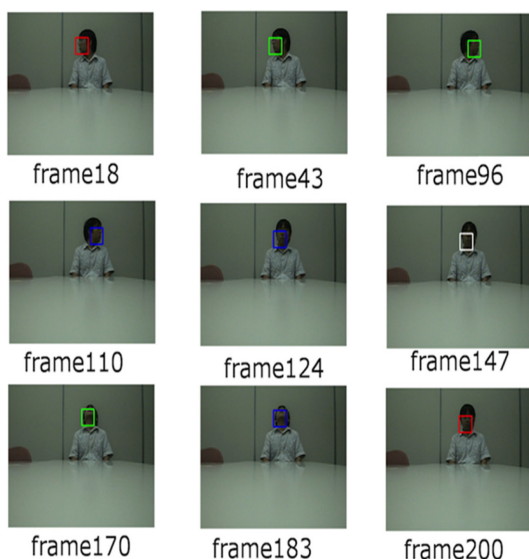
The video recording was carried out for examinee of 1 person. The video photographed by this experiment takes 640×480 pixels per frame, 30fps frame rate, approximately 6 seconds recording time, and the distance about 1m between camera and examinee.

To begin with, the result of face detection in the first image is shown in Fig. 3. Because the detection of the face fails when the examinee does not turn toward the direction of the camera, it is impossible to detect the direction of face in a continuing step.

Rectangle windows which shows the face region of the figure are shown by the different colors in terms of the direction of face. This windows are used in the pro-

**Table 2.** Relation between rectangle color and face direction

Face Direction	Rectangular Color
Up	White
Down	Red
Left	Blue
Right	Green



**Fig. 3.** Results of face detection (1 examinee).

posed algorithm. The relationship between color of the rectangle window and direction of the face is shown in Table 2. In the Fig. 3, the result of the face direction estimation is compared with actual movement. It shows a confirmation that the case of examinee is one person is possible for face direction estimation using the proposed technique in this paper.

### 5.2. In Case of 2 Examinees

Two examinees are photographed. The distance between examinee sitting at left and camera is approximately 2m, and the examinee sitting at right is approximately 1m. The examinee of the left side is numbered to examinee of 1, and the examinee of the right side is numbered to examinees of 2. Size and frame rate of the video are same to the one examinee case. The total frame number is 447 and the video playback length is about 14 seconds.

In the figure, the frames where two examinees are sit are shown. In the presented experiment, frame number 71, 132, 168, 229, 292, and 336 are an arrested state. The result of face tracking by the proposed technique including the installation alignment is shown in the figure. The case has failed to the installation alignment such as frame number 336, rectangle window of examinees of 2 is perfectly deviated from the face. If there is no performance in the error arisen from frame number 380, it is continued the processing. The position in face region of examinee 1 as shown in frame number 380 is shown to Table 3, and the position in face region examinee 2 is shown in Fig. 4.

In the position, 0 shows the distance from the position at the initial stage by XY axes. Large slippage is not generated in the face region of examinee 1. However the slippage increases in the face region of examinee 2.

**Table 3.** Relation between rectangle color and face direction

Frame Interval	Face Direction	Detection Result
1 ~ 18	Left	Left
19 ~ 42	Stop	Stop
43 ~ 96	Right	Right
97 ~ 109	Stop	Stop
110 ~ 124	Left	Left
125 ~ 146	Stop	Stop
147 ~ 170	Up	Up
171 ~ 182	Stop	Stop
183 ~ 200	Down	Down



Fig. 4. Results of face detection (2 examinee).

The slippage begins to greatly deviate from frame number 300 toward downward direction, and it is increasing along with the time passing. It seems to be considered that this is due to the size of detected face region. Based on the video for examinee 1, from comparing Fig. 3 and 4, the rectangle which shows the face region is smaller than the actual face region. Therefore it is proven that the rectangle window gets larger than the face region when each examinee is moving.

The installation alignment of the face region is carried out in order to contain the flesh color region on the boundary line over the fixed proportion. In the video for examinee 1, it is able to confirm that the direction of the face could be appropriately detected by this proposed algorithm. However it can be guessed that the alignment of face region becomes unstable by the color which is similar to the background color. Finally it has perfectly come off in the video for examinee 2.

## 6. Conclusion

In this paper, the directional estimation of the face was carried out using the face part detection, using the template of face tracking, and using the difference between frames and feature quantity of 4 directions and luminance value. The frontal face is detected by the Viola-Jones method from the first frame of video, and flesh color reference value is obtained from the color histogram. The position of both eyes is detected by template matching, and the distance between both eyes is obtained.

The face tracking is based on information of the difference between frames. Installation alignment and face part detection of the face region are carried out when the slippage of the face region is not generated. The picture element number which could judge that it is the color of the skin can make the face region move. If pixel value on boundary line in the face region is compared with flesh color reference value, and if there is boundary line under  $1/4$ , then it can be concluded until picture element number of flesh color consisted over  $1/4$ .

Face parts detection is carried out after the installation alignment. Template matching is carried out using feature quantity of 4 directions and template of the luminance value, as it done even in the first frame, and the eye was detected with the judgment of whether it is an eye from the value of the resemblance. For 1 examinee video and 2 examinee video, the proposed technique is applied, and the effectiveness is verified.

## Acknowledgement

This work was supported by research fund from Chosun University 2014 Research Program.

## References

- [1] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features", *Computer Vision and Pattern Recognition*, Vol. 1, pp. 1511-1518, 2001.
- [2] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", *IEEE T. Pattern Anal.*, Vol. 24, pp. 603-619, 2002.
- [3] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking", *IEEE T. Pattern Anal.*, Vol. 25, pp. 564-577, 2003.
- [4] Y. Tsuduki, H. Fujiyoshi, and T. Kanade, "Mean shift-based point feature tracking using SIFT", *J. Information Processing Society*, Vol. 49, pp. 35-45, 2008.
- [5] N. Oshima, T. Saitoh, and R. Konishi, "Real time mean shift tracking using optical flow distribution", *SCIE-ICASE*, pp. 4316-4320, 2006.
- [6] E. Watanabe, T. Ozeki, and T. Kohama, "Analysis of behaviors by lecturer and students in lecture based on piecewise auto-regressive modeling", *Intelligent Computer Communication and Processing*, pp. 385-390, 2011.

- [7] K. Kunihiro, K. Matsuura, and Y. Yano, "Cultivating the performance of presentation through monitoring presenter's action", Proceedings of the 18<sup>th</sup> International Conference on Computers in Education ICCE 2010, Putrajaya, Malaysia, November 29-December 03, 2010.
- [8] Y. Araki, N. Shimada, and Y. Shirai, "Detection of faces of various directions and estimation of face direction in complex backgrounds", PRMU, Vol. 217, pp.87-94, 2001.
- [9] B. K. P. Horn and B. G. Schunck, "Determining optical flow", Artif. Intell., Vol. 17, pp. 185-203, 1981.