Face Tracking with Protection of the Privacy using Color Histogram and Evolutionary Video Processing

Daichi Oikawa, Junya Sato, Takuya Akashi*

Iwate University, 4-3-5 Ueda, Morioka, Iwate, Japan

*Corresponding Author: akashi@iwate-u.ac.jp

Abstract

In this paper, a face tracking method is proposed to monitor a person with protection of the privacy using a monocular camera. By applying a processing, which shuffles an interested pixel with neighbor pixel, the privacy can be protected. Although some image features, such as edge and corner are lost, color histogram feature is not lost extremely. Therefore, the proposed method uses the histogram. By combining with Genetic Algorithm (GA), the proposed method can track the face with robustness to location, scaling, and rotation changes. As a result of experiments, the average accuracy and the processing time of the proposed method are 63% and 23 fps, a better result than other methods.

Keywords: face tracking, color histogram, genetic algorithm, protection of the privacy.

1. Introduction

Protection of the privacy is a importance things for surveillance system. As one of the examples, there is a system for prevention of the drowning while bathing. Some accidents are happened by various reasons, such as disorder of a brain or a heart, rapid increase of blood pressure, and so on. These risks are especially high with older adults. If they become unconscious suddenly, they must be taken to a hospital immediately. However, early discovery of the unconscious person is difficult since there is basically only one person while bathing. Because the number of older population of living alone increases associated with aging population combined with the diminishing number of children, increase of the number of troubles while bathing can be predicted. A surveillance camera in a bathroom can prevent the accidents. However, setting the camera is difficult in terms of the privacy. For these backgrounds, a system

with protection of the privacy is necessary. In this paper, a novel method, which can track the face in a shuffled image, is proposed. All the pixels in target sequences are shuffled with the neighbor pixels. By applying this process, it is possible to protect the privacy. Some image features, such as edge and corner are lost. Nevertheless, color histogram feature is not lost extremely because it is frequency of the pixel value in a region of interest. Therefore, it is suitable for our system. Furthermore, the feature of the face does not change extremely even if the face direction is changed. It means that the color histogram is robust to the change of appearance and can be applied to a video processing.

As one of the related works of object tracking, there is CAMSHIFT proposed by Bradski⁽¹⁾. By applying Mean Shift algorithm and operating a color probability distribution image derived from color histograms of the target object, it can track an object fast with robust to scaling, angle, and noise. However, as mentioned in the paper⁽²⁾, it is difficult to track again if the object is occluded for a while and a background has the same color as the object.

There is a face detection and tracking method, which is proposed by Viola and Jones⁽³⁾. It uses boosted cascade, which is constructed from many weak classifiers based on AdaBoost using Haar-Like features. The method has low calculation cost, high accuracy, and robustness to a change of illumination. However, detection of a multiview face using only one detector is difficult. This is because the detector is leaned with many faces, whose direction and pose are the same. If we want to detect the multiview faces, some detectors, which can detect the particular face direction, must be created. Hence, a method, which does not require pre-training, is necessary since it needs to collect many images and take a long time for learning. This preparation is hard work for a user. However, our proposed method does not require these works.

Fig. 1. Flowchart.

2. Proposed Method

2.1 Outline of the Proposed Method

Flowchart of the proposed method is shown in Fig.1. Firstly, a color histogram template is created from a face detection result⁽³⁾. Secondly, a shuffled target image, which is obtained from an input video frame, is created. After that, a region, whose color histogram is similar to the color histogram template is searched in the target image by GA. The GA optimizes 5 parameters to represent a candidate region of the face. The parameters are x and y coordinate, scaling (x and y axis), and angle of rotation. Then, each candidate region is evaluated by comparing a distance between color histograms, which are created from the candidate region, and the color histogram template. The closeness of the distance represents the similarity of the face. GA can obtain an optimum solution, which is the face, by genetic operations. From the next section, the detail information is explained.

2.2 Creation of Color Histogram Template

A color histogram template is created from a face region of a subject. The face is detected using Viola-Jones method in an input video frame⁽³⁾. Next, the detected region is scaled down. This is because it includes a background. The width is 0.6 times and the height is 0.7 times of the detected rectangle. These values are determined empirically. Next, after the image is obtained, it is resized to 1/3 times using bicubic method. After that, color histograms of Cb and Cr component are created after the resized image is converted to YCbCr color space. The reason why YCbCr color space is used is often used to extract a skin color^(4,5).

Fig. 2 a) original image; b) shuffled image.

Because the color histogram does not include the position information, it is not extremely changed even if a head is moved. Therefore, the histogram feature is robust to change of appearance. Furthermore, it can handle the scale change by normalization if the size of search window is change. A weak point is that the histogram is not robust to illumination change even if Cb and Cr component are used.

2.3 Shuffled Target Image

In order to protect the privacy, pixels in an image are shuffled with neighbor pixel as a preprocessing. The shuffle is performed in 31×31 pixels. An interested point, which is center of the region, is shuffled with the neighbor pixel randomly. This processing is adapted in the whole image. The examples of the shuffled image are shown in Fig.2.

2.4 Evolutionary Video Processing

GA is one of the methods, which optimizes parameters based on an objective function. In this research, it is histogram distance. The parameters are position, scaling (x and y axis), and angle of rotation. Some search points, which are called as individuals, have the parameters as chromosomes. It is binary of 0 and 1. Next, the processing procedure is explained. Firstly, the chromosomes are initialized randomly. Secondly, the parameters are acquired from them. By using them, a rectangle, whose size is the same to the face detection rectangle for creation of a color histogram template, is transformed. After that, the transformed rectangle is located. It is called as a candidate region of a face. Then, after color histograms of Cb and Cr component are created from the candidate region, the histogram distance is calculated. Each individual is evaluated after a fitness value is calculated using equation (1), (2), and (3).

$$fitness = c - w \times (\rho_{Cr} + \rho_{Cb}) \quad (1)$$

$$w = \frac{2R}{r_x + r_y} \quad (2)$$

$$\rho = \sum_{i=0}^{255} \{p(i) - q(i)\}^2 \quad (3)$$

Let c is a constant number. w is a weight. It is based on the scale factor. r is scale factor, and R is maximum scale factor. $p(i)$ and $q(i)$ indicate the each frequency of index i . The calculation method of a color histogram is squared euclidean distance. This is because it is the best result in preliminary experiment. After a fitness value is acquired, the genetic operations, such as selection, crossover, and mutation are performed based on the value. These operations are iterated until a termination condition is satisfied. This iteration is called as generation iteration. This processing can create better population. In a final generation, one individual, whose fitness value is the highest, is displayed as a detection result.

In evolutionary video processing⁽⁶⁾, when a new frame is input, initialization of the population is not necessary. By inheriting the population to the first generation in the next frame, the search is effectiveness. The reason is that a subject does not change extremely between the frames.

3. Experiments

3.1 Experimental Environment

The number of target sequences is 3 (Fig.3). They were recorded using a webcam with 30 fps. Since the proposed method uses a randomseed, we used 50 randomseeds. The accuracy and the processing times were obtained by calculating the average. The number of individuals and generation iteration per image were 20 and 100. Crossover rate and mutation rate were 0.95 and 0.01. These parameters were decided based on a preliminary experiment and literature⁽⁷⁾.

3.2 Evaluation method

An answer face region for judgment is an upright rectangle, and it is manually drawn based on some conditions. The answer rectangle includes a region from a top of eyebrow to the bottom of the chin. Also, the rectangle is maximum region, which does not include the sideburn. If the eyebrow is occluded by the frontal hair, eyebrow is predicted. When the equation (4) and (5) are satisfied, the result is correct.

$$\sqrt{\{m_{gt} - m_p\}^2} < \frac{1}{2} \sqrt{\{v_{gt} - m_{gt}\}^2} \quad (4)$$

$$0.5 < \frac{a_p}{a_{gt}} < 1.5 \quad (5)$$

Let m is centroid. v is a vertex. a is a surface. gt and p

Fig. 3. Example of target sequences.

are the answer rectangle and the detection rectangle. As a comparative method, Viola-Jones method⁽³⁾ and CAM-SHIFT⁽¹⁾ were used. The computer, whose CPU is 3.3 GHz, and RAM is 8 GB, was used.

3.3 Experimental Results

The experiment results are shown in Table.1 and 2. The accuracy of the proposed method is 63% while CAM-SHIFT and Viola-Jones method are 0%. Although CAM-SHIFT can track a face, the result includes a neck region. This is the reason of the low accuracy. In our system, it is a purpose to track an only face region. Therefore, CAM-SHIFT is not suitable for our system. Viola-Jones method uses difference of intensity between face parts, such as palpebral fissures, nares, and a mouth and skin region as a Haar-like feature. If a shuffled image is used as a target image, that information is lost. From this reason, the accuracy is 0%. The fastest processing speed is CAMSHIFT (328 fps). The proposed method is 23 fps, and Viola-Jones method is 9 fps. Each tracking results are shown in Fig.4.

The accuracy of the proposed method in Seq. 1 is higher than the other sequences, although the subject has short hair and there is no face occlusion. Most face directions which fail to track correct face region is up. This is because appearance change is caused by a face direction, color by vary of lighting condition, and a weight increases fitness value of individuals whose chromosome of scale is large. Sequence. 2 was the worst accuracy in the three sequences. It is considered to be the cause of occlusion by a hair when the face direction is down and right. The accuracy of Seq. 3 decreases by 20% compared with Seq. 1 although the subject has short hair. The reason of this is contemplated that the light in Seq.3 is stronger, and the amount of the change of face direction is larger than Seq. 1.

3.4 Another evaluation method

As mentioned in section 3.3, the accuracy of the proposed method is the best. Next, the tracking performance in shuffled image is evaluated. When a center of a detection rectangle is in an answer rectangle, the result is correct answer. Table 3 shows the accuracy. The average accuracy is 95%. This result is enough for any applications.

Table 1. Tracking accuracy (%).

	CAMSHIFT	Viola-Jones	Proposed method
Seq. 1	0	0	79
Seq. 2	0	0	55
Seq. 3	0	0	57
Average	0	0	63

Table 2. Processing time (fps).

CAMSHIFT	Viola-Jones	Proposed method
328	9	23

Table 3. Tracking result of proposed method (%).

	1	2	3	Average
Accuracy	99	90	94	95

4. Conclusions

In this paper, a novel face tracking method with protection of the privacy is proposed. By shuffling pixels with neighbor pixel, the privacy can be created. Some image features, such as edge and corner are lost. However, a color histogram feature is not lost extremely. By combining with genetic algorithm with evolutionary video processing, the proposed method can track the face with a high accuracy and fast processing speed. In experiment, we compared with CAMSHIFT and Viola-Jones method. The effectiveness of the proposed method is the best. The accuracy of the proposed method may decrease under an incandescent lamp because it uses the color feature. We would like to experiment in a real environment and improve the method.

References

- (1) Gary R. Bradski : "Computer vision face tracking for use in a perceptual user interface", Fourth IEEE Workshop on Applications of Computer Vision, 1998
- (2) Patrick Pérez, Carine Hue, Jaco Vermaak, and Michel Gangnet : "Color-based probabilistic tracking", Springer Berlin Heidelberg, pp. 661-675, 2002
- (3) Viola Paul and Michael J. Jones : "Rapid object detection using a boosted cascade of simple features", IEEE Computer Society Conference on, Vol. 1, 2001
- (4) Jorge Alberto Marcial Basilio, Gualberto Aguilar Torres, Gabriel Sánchez Pérez, and L. Karina Toscano Medina : "Explicit image detection using YCbCr space color model as skin detection", Applications of Mathematics and Computer Engineering, pp. 123-128, 2011
- (5) Douglas Chai and King N. Ngan : "Face segmentation

(a)

(b)

(c)

Fig. 4. Experiment results. The green and red rectangle represent correct answer and incorrect answer. a) CAMSHIFT; b) proposed method; c) Viola-Jones.

- using skin-color map in videophone applications", IEEE Transactions on Circuits and Systems for Video Technology Vol. 9 Issue 4, pp. 551-564. 1999
- (6) Takuya Akashi, Yuji Wakasa, Kanya Tanaka, Stephen Karungaru, and Minoru Fukumi : "Downsized Evolutionary Video Processing for Lips Tracking and Data Acquisition", Journal of Advanced Computational Intelligence, Vol. 11.8, 2007
- (7) Sadiq M. Sait and Habib Youssef : "Iterative Computer Algorithms with Applications in Engineering: Solving Combinatorial Optimization Problems", Wiley-IEEE Computer Society Press, 2000