Robust Face Tracking using Improved Mean Shift Algorithm

S B Setyawan¹, H Arrosida², A Sudaryanto³, R G P Yudha⁴

Department of Computer Control Engineering, Politeknik Negeri Madiun, Madiun, Indonesia

Email: sulfan@pnm.ac.id¹, hanum@pnm.ac.id², albert.success@pnm.ac.id³, gaguk@pnm.ac.id⁴

Abstract. Face tracking is important system for robotic. The problem is that it is less resistant to occlusion and noise. Robust face tracking algorithm in this paper using improved mean shift algorithm is proposed. This algorithm consist of face detection and tracking. The face identification is done by the Viola-Jones algorithm. Then, Mean Shift algorithm is done to track face as a target and improve robustness from occlusion. CBWH is added in face as target to reduce noise. The experimental results prove that it is robust from noise background and occlusion.

Keywords: Face tracking, mean shift, occlusion, background interference

1. Introduction

The human robot interface is a very popular field of knowledge today. Robots that are useful today are those that can help work for people so that the robot have to be able to interact well with users or people. For this condition, the robot needs a robust face tracking system in order to interact well with other people[1][2]. Face detection is needed in face tracking system. There were a lot of methods to detect face[3][4][5]. The common method to detect human face is haarcascade which is used to recognize faces[6].

There were many methods used for facial tracing. Optical flow is a facial tracking method. Shen Fu et al used it in facial tracking[7]. This face tracing can be done well. However, the tracking is still an error when the object is moving quickly. Tracking objects using optical flow can reduce the accuracy of face tracking[8]. One of the well-known ones is the template matching method. Bin et al also wrote a template matching method using dynamic templates[9]. In this research, object tracking can be done robustly, but it requires a lot of computation to update the template.

The other method is the mean shift tracking method[10],[11],[12]. MS tracking is a robust tracking for occlusion. However, this tracking is not robust against the same background color disturbance as the target object. Comaniciu et al improved ms method to reduce noise from the background features tracked using a bwh[13], but this method does not develop the performance of means shift algorithm because the values of the target object and candidate object are the same. Ji feng et al also developed mean shift with the CBWH method to remove background noise[14]. Sulfan et al have also developed an developed mean shift for robust object tracking against occlusion[15].

The contribution of this paper is combining face detection and CBWH Mean Shift(MS) tracking system. The aims of combining that method could make the facial tracking is resist from background features and occlusion. MS tracking algorithm is used to reduced occlusion that interference the face as target. Background noise is removed using CBWH which is inserted into target representation.
2. System Overview

In this chapter, facial detection system is described. Face detection is used to detect the face as target model representation. Then MS algorithm is presented. This method is very effective in tracking objects. This method uses a colour histogram probability in determining the face model and face candidate. CBWH is embedded in the object model to reduce background noise. With the updated target candidate and model, the location of a new face can be found using the mean shift tracking method. The block diagram of face tracking can be shown in Fig 1. Face tracking block diagram is shown in Figure 1. Green line shows face as target model which is added CBWH.

![Face tracking block diagram](image)

2.1. Facial Detection

Facial detection used is the viola jones technic. Detection is carried out by inputting RGB images taken from the camera. After that it is compared with the database that has been made by Haarcascade. So that will be marked on the detected face. Equation (1) shows the intensity for the viola jones algorithm. In this equation using adaboost to select the features used for training.

\[ i(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \]  

(1)

where \( i(x, y) \) is the intensity of image and \( i(x', y') \) is the original image.

2.2. Object Model

In this research the object model is obtained from facial detection. The result of facial detection is the face area. From this area, the target model can be calculated using Equation (2), by taking the probability histogram.

\[ \hat{q}_u = C \sum_{i=1}^{j} k \left( \left| x_i \right|^2 \right) \delta \left( b(x_i) - u \right) \]  

(2)

\( \{x_i\}_{i=1...n} \) is pixel that have been normalized from facial location as target, and \( b(x_i) \) describes pixel RGB at location \( x_i \). \( C \) is constant of normalization in Equation (3)

\[ C = \sum_{i=1}^{n} k \left( \left| x_i \right|^2 \right)^{-1} \]  

(3)
2.3. Corrected Background Weighted Histogram

CBWH is a method used to increase the MS tracking capability in placing the object tracking location due to background noise. The background area is obtained from 2 times the length and width of the face area as the target. Coefficient $v_u$ in Equation (4) is used to define transformation between object model and object as a candidate.

$$V_u = \min\left(\frac{\hat{\delta}_u}{\delta_u}, 1\right)_{u = 1..m}$$

(4)

Then, Equation (2) can be changed to get a new object model as shown in Equation (5)

$$\hat{q}_u = C^* V_u \sum_{j=1}^{n} k\left(\left\|x_i - x_j\right\|\right)^d (b(x_i) - u)$$

(5)

2.4. Object Candidate Representation

The object candidate is obtained from the same area as the previous target model. From the same area, if the face moves, the probability value will change so that it can be compared with the previous value. So that new face locations can be found. Feature probability $u = 1 ... m$ in the object candidate can be calculated in Equation (6)

$$\hat{p}_u = C^* \sum_{i=1}^{n} k\left(\left\|x_i - x_j\right\|\right)^d (b(x_i) - u)$$

(6)

2.5. Mean Shift (MS) Tracking

The mean shift (MS) tracking is calculated after the model target that has been inserted with CWBH and the target candidate is calculated. The new location can be specified with Equation (7)

$$\hat{y}_1 = \frac{\sum_{i=1}^{n} \omega(x_i)\delta\left(\left\|y_0 - x_i\right\|\right)^2}{\sum_{i=1}^{n} \omega(x_i)\left(\left\|y_0 - x_i\right\|\right)^2}$$

(7)

Where

$$\alpha = \sum_{u=1}^{m} \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\hat{y}_0)}} \delta(b(x_i) - u)$$

(8)

3. Proposed Algorithm

In this paper, Algorithm of face detection and CWBH Mean Shift tracking is presented. The aims of this algorithm could make the face tracking is robust tracking. Face as target is detected using viola-jones algorithms. MS tracking algorithm is used to reduced occlusion and background interference is reduced using CBWH which is added in target model. That methods can be summarized as follows.

a. Detecting face using the Viola-Jones algorithm
b. If face detected go to step 3, otherwise go to step 1
c. Compute the object model, weighted background histogram
d. Compute the transformed object model
e. Let $d=0$
f. Compute object candidate in the current frame

g. Compute the weights

h. Compute the new position

i. Let \( d = \|y_1 - y_0\| \), \( y_1 = y_0 \), \( d = d + 1 \). Set the error threshold

   If \( d < \varepsilon \) or \( k > N \)

   Stop iteration

   Otherwise

   Go to step 4

j. Go to next frame, step 3

4. Results of experimental

In chapter 4, the traced face is tested. The test was carried out 3 times. The first is testing with a similar background. Then performed the test with object rotation. The last one is testing for occlusion. Figures 1 and 2 show the test for background noise. In this test, the authors compare the original means shift with the proposed algorithm. In Fig 2 shows the proposed method, in figure 1 shows the ms tracking method.

Fig. 1. Facial tracking using original ms

Fig. 2. Facial tracking using proposed method.

Figure 3 describes face tracking error. It shows that the proposed algorithm has a smaller error than the original one. The ms tracking error is 20.84 pixels and 6.9 deviation. While the proposed method has an error of 5.39, deviation of 3.59.
Second trial is used to test robustness of facial tracking. The aims of this experimental is evaluate face tracking from rotation face as target. In this experimental, face will gradually turned from right to left and otherwise. 20th, 100th and 130th frames are taken. Figures 4 and 5 show the second experimental. In this experimental haarcascade algorithm and proposed method are compared. Figure 4 shows the experimental used Viola-Jones algorithm with uses Haar-like features. Figure 5 shows proposed method. In Figure 4 shows face tracking using haarcascade algorithm cannot track well. When face turn left and right, the face tracking is missing. Then, in figure 5 shows face tracking using proposed method. The result shows that proposed method could track well even though the face turn from right to left, otherwise.

Third experimental result is used to evaluate occlusion within face tracking. The aims of this experimental is evaluate robustness face tracking from occlusion. In this experimental, face will occlude with other objects. 20th and 50th frames are taken. Figures 6 and 7 show the third experimental. In this experimental haarcascade algorithm and proposed method are compared. Figure 6 shows the experimental used haarcascade algorithm. Figure 7 shows proposed method. In Figure 6 shows face tracking using haarcascade algorithm cannot track well. When face occluded, the face tracking is missing. Then, in figure 7 shows face tracking using proposed method. The result shows that proposed method could track well even though the face occluded.
5. Conclusion

Proposed face tracking algorithm is done to develop facial tracking robustness from background noise of background features and occlusion from other object. Novel tracking method has 5.39 mean error, original MS algorithm has 20.84 mean error. In the second experiment the object rotation can be resolved The third trial shows that faces with noises can still be detected. It can be concluded that the proposed surface can be traced against background disturbances and occlusion.

References

[1] Maneesha K, Shree N, Pranav D R, Sindhu S K and Gururaj C 2017 Real time face detection robot RTEICT 2017 - 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, Proceedings

[2] Gao Y, Lv X and Jia H 2016 Real-time multi-view face detection based on optical flow segmentation for guiding the robot 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2015

[3] Zhang T, Li J, Jia W, Sun J and Yang H 2018 Fast and robust occluded face detection in ATM surveillance Pattern Recognit. Lett.

[4] Machaca Arceda V E, Fernández Fabián K M, Laguna Laura P C, Rivera Tito J J and Gutiérrez Cáceres J C 2016 Fast Face Detection in Violent Video Scenes Electron. Notes Theor. Comput. Sci.

[5] Putro M D, Adji T B and Winduratna B 2016 Adult image classifiers based on face detection using Viola-Jones method Proceeding of 2015 1st International Conference on Wireless and Telematics, ICWT 2015

[6] Viola P and Jones M J 2004 Robust Real-Time Face Detection Int. J. Comput. Vis.

[7] Hsiao S F and Tsai C Y 2018 Design and implementation of low-cost LK optical flow computation for images of single and multiple levels Proceedings - 21st Euromicro Conference on Digital System Design, DSD 2018

[8] Chen Z, Cao J, Tang Y and Tang L 2011 Tracking of moving object based on optical flow detection Proceedings of 2011 International Conference on Computer Science and Network Technology, ICCSNT 2011

[9] Zheng B, Xu X, Dai Y and Lu Y 2012 Object tracking algorithm based on combination of
dynamic template matching and kalman filter Proceedings of the 2012 4th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2012

[10] Comaniciu D and Ramesh V 2000 Mean shift and optimal prediction for efficient object tracking IEEE International Conference on Image Processing

[11] Panahi R, Gholampour I and Jamzad M 2013 Real time occlusion handling using Kalman Filter and mean-shift Iranian Conference on Machine Vision and Image Processing, MVIP

[12] Zhao J, Qiao W and Men G Z 2009 An approach based on mean shift and kalman filter for target tracking under occlusion Proceedings of the 2009 International Conference on Machine Learning and Cybernetics

[13] Comaniciu D, Ramesh V and Meer P 2003 Kernel-based object tracking (Cited by: 2296) IEEE Trans. Pattern Anal. Mach. Intell.

[14] Ning J, Zhang L, Zhang D and Wu C 2012 Robust mean-shift tracking with corrected background-weighted histogram IET Comput. Vis.

[15] Setyawan S B, Purwanto D and Mardiyanto R 2016 Visual object tracking using improved Mean Shift algorithm 2015 International Conference on Information Technology Systems and Innovation, ICITSI 2015 - Proceedings