March-April 2020
ISSN: 0193-4120 Page No. 3945 - 3952

Improved Viola-Jones Face Recognition UsingTracking

Mangaras Yanu Florestiyanto[1], Hidayatullah Himawan[1] and Parasian DP Silitonga[2]
[1]Informatics Engineering Department, Veterans National Development University, Indonesia
[2]Informatics Engineering Department, Santo Thomas Catholic University, Indonesia

Article Info
Volume 83
Page Number: 3945 - 3952
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 26 March 2020*

Abstract

Optimization of accuracy has been an important study in the study of face recognition for 2 decades. Varied facial features and changing poses in a short amount of time make optimizing accuracy even more complicated. This study aims to improve the accuracy of face recognition of the Viola-Jones method on moving objects by adding a tracking algorithm. The tracking algorithm used in this study is the Continuously Adaptive Mean Shift (Camshift) algorithm. This algorithm is the development of the Mean Shift algorithm which continuously adapts or adjusts to the color probability distribution that is always changing with each change of frame of a video sequence. The addition of tracking significantly increases accuracy compared without tracking by 96%.

Keywords: *Face Recognition, Tracking, Viola-Jones, Euclidean Distance*

I. Introduction

The issue of accuracy in face recognition is indeed complicated because faces are complex and complex objects. Faces have varied features and can change from time to time, even changes that can take place in a very short time [1]. The complexity of facial features and this rapid change significantly affect the accuracy of the matching process. In addition to the complexity of facial features and their changes, the accuracy of face matching is also influenced by the image quality [2] that is input because of the uncontrolled image capture environment.

A face recognition algorithm accuracy study has been conducted [3] comparing the performance of 2 (two) popular libraries in face recognition, namely Dlib and Opencv. The study results [3] state that 2 (two) libraries being studied have a good performance in detecting faces. Another study was conducted [4] by comparing several face matching methods based on Principal Component Analysis (PCA). The results of the study stated that recognition with Euclidean Distance has an average accuracy of 93.75%, Chebychev Distance has an average accuracy of 96.87% and Manhattan Distance has an average accuracy of 96.87%. While studies [2] claim that the Eigenface method has the highest accuracy rate of 96% compared to Fisherface and Local Binary Pattern Histogram (LBPH).

Study [5] evaluates the performance of face recognition methods that are commonly used, the study classifies face recognition methods into 3 (three) groups, namely 1) appearance-based, where the Kernel PCA method using SVM classifier has an accuracy level 99.05% highest

Published by: The Mattingley Publishing Co., Inc.

compared to other methods in this group; 2) feature-based, the highest level of accuracy in this group is the SIFT & MLBP method of 99.73%; and 3) soft computing-based, Gabor Jet using Borda Count Classification has the highest accuracy rate of 99.8% compared to other methods in this group.

Development Network (DN) is used [6], [7] to increase the accuracy of face recognition of an object with a complex background. The result is the accuracy of face recognition reaches more than 95%. In addition to face recognition studies of single objects, there are several studies that examine face recognition for multiple objects. Study [8] combines the Haar Cascade and Eigenface methods and is able to detect 55 faces at once in a single process with an accuracy rate of 91.67%. Study [9] claims that the Minimum Eigen Value algorithm using SVM classifier has a better level of accuracy in face recognition for plural objects which is 83%. Study [10] conducted a face recognition study for plural objects in real time using the Viola-Jones method as a face detection method, Speeded up Robust Features (SURF) to extract features and utilize the M-estimator Sample Consensus (MSAC) for face matching. The result has an accuracy of 95.9%.

This study aims to increase the level of accuracy of face recognition of the Viola-Jones method on moving objects by adding a tracking algorithm. The tracking algorithm used in this study is the Continuously Adaptive Mean Shift (Camshift) algorithm. This algorithm is the development of the Mean Shift algorithm which continuously adapts or adjusts to the color probability distribution that is always changing with each change of frame of a video sequence.

II. MATERIALS AND METHOD

The study begins with the development of a face recognition application as a tool to obtain recognition accuracy data. The common steps in face recognition are image acquisition, pre-processing, face detection, facial feature extraction, processing feature data with machine learning algorithm, matching training set with test set, and evaluating the results of matching [11]. The image acquisition in this study was taken using a CCTV camera. After getting the image from the camera, normalization process with gray scaling is done as a pre-processing step and face detection process is done with Viola-Jones, as shown in Figure 1.

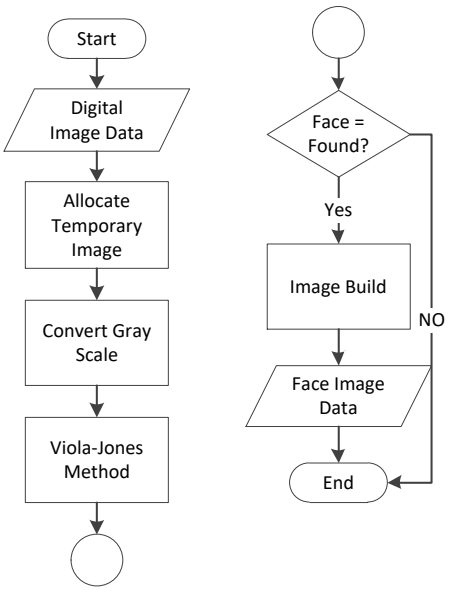

Fig. 1. Steps to Study Face Recognition Without Tracking

March-April 2020
ISSN: 0193-4120 Page No. 3945 - 3952

Furthermore, according to the purpose of the study to improve the accuracy of the face recognition steps in Figure 1, it was developed by adding a tracking algorithm. Camshift tracking algorithm is implemented in the face detection process. The tracking process will label every object detected by the camera, so no re-detection is made on the same object. The steps of the CAMSHIFT algorithm are as follows [12] [13] :

1. Determine the initial size of the window search.
2. Determine the starting location of the window search.
3. Specify the calculation area in the middle of the window search with a size larger than the window search.
4. Convert the video image frame into the HSV color system (Hue, Saturation, Value), and do a color histogram lookup in the calculation region that will produce the image color probability distribution.
5. Perform the Mean Shift algorithm as above (one or many iterations) with input in the form of the size and location of the search window as well as the color probability image distribution, save the zeroth moment.
6. Set the x, y, z, and head roll values obtained from step e.
7. Use the values x, y to determine the midpoint of the search window, (2√area) to set the size of the search window. For image color probability image distribution, the mean (centroid) area in the search window can be calculated as follows :

   Zeroth Moment :

$$M_{00} = \sum_x \sum_y I(x,y) \quad (1)$$

   First Moment for x and y :

$$M_{10} = \sum_x \sum_y x\, I(x,y)$$
$$\text{and } M_{01} = \sum_x \sum_y y\, I(x,y) \quad (2)$$

   so the mean location in the search window (Centroid) is :

$$x_c = \frac{M_{10}}{M_{00}} \qquad y_c = \frac{M_{01}}{M_{00}} \quad (3)$$

   Where (x, y) is the pixel color value at position (x, y) in the search window. Object orientation is obtained by calculating the second moments as follows [14]:

$$M_{20} = \sum_x \sum_y x^2\, I(x,y)$$
$$\text{and } M_{02} = \sum_x \sum_y y^2\, I(x,y) \quad (4)$$

   The object's orientation is :

$$\theta = \frac{arctan\left(\frac{2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right)}{\left(\frac{M_{20}}{M_{00}} - x_c^2\right) - \left(\frac{M_{02}}{M_{00}} - y_c^2\right)}\right)}{2} \quad (5)$$

   If :

$$a = \left(\frac{M_{20}}{M_{00}} - x_c^2\right) \qquad b = 2\left(\frac{M_{11}}{M_{00}} - x_c y_c\right) \qquad c = \left(\frac{M_{02}}{M_{00}} - y_c^2\right)$$

   then the length l and w width of the centroid distribution are [15]:

$$l = \frac{\sqrt{(a+c) + \sqrt{b^2 + (a-c)^2}}}{2} \quad (7)$$

$$w = \frac{\sqrt{(a+c) - \sqrt{b^2 + (a-c)^2}}}{2} \quad (8)$$

   The equation produces x, y values, object rotation, length and width (area or z value)
8. Repeat step c for each change of the video image frame

March-April 2020
ISSN: 0193-4120 Page No. 3945 - 3952

The whole step of the face recognition process by adding tracking can be seen in Figure 2 and Figure 3.

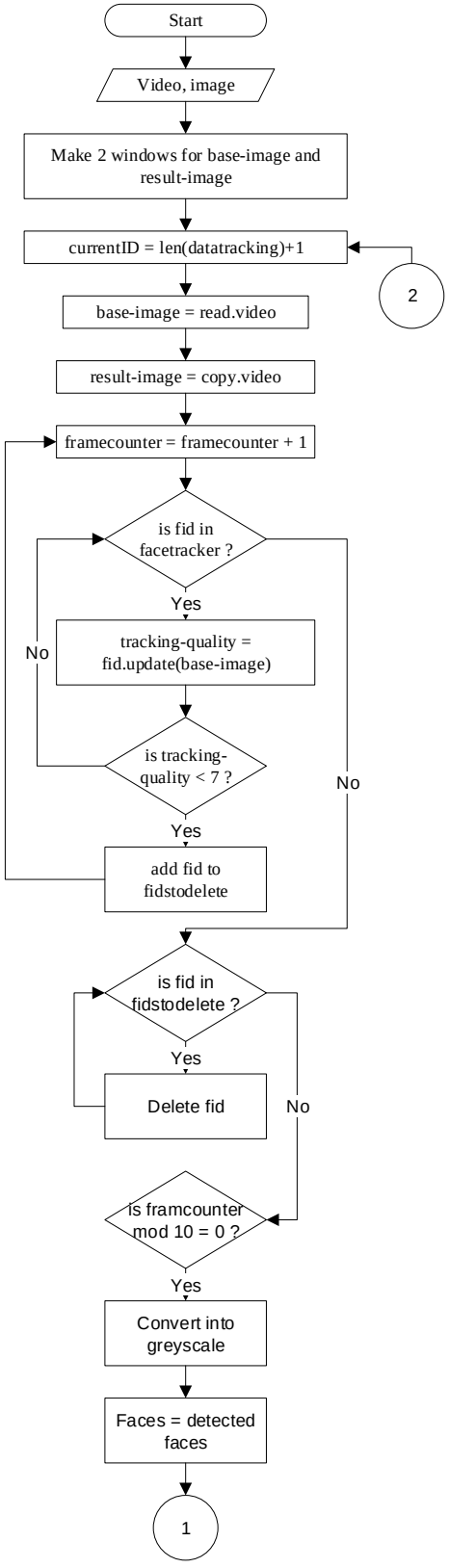

Fig. 2. Step by Step Face Rrecognition Study Tracking

Published by: The Mattingley Publishing Co., Inc.

March-April 2020
ISSN: 0193-4120 Page No. 3945 - 3952

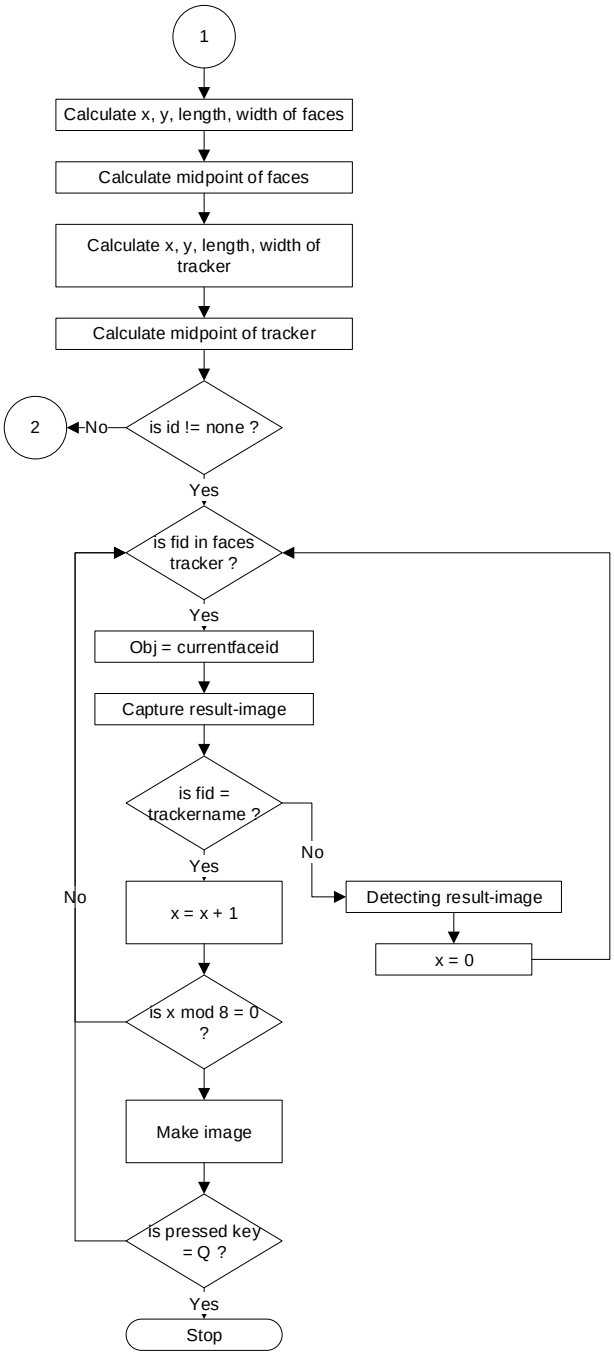

Fig. 3. Advanced Steps for Face Recognition Study With Tracking

III. RESULT

The process begins with taking values from the training data, starting with face detection on the results of digital images recorded by the camera. The detected face image will be normalized and extracted to get the minimum value used for the matching process. If the face value of the test object is smaller than the threshold value then it will be recognized as a member of the training data.

The face recognition application is tested with 5 (five) objects that have been trained and 5 (five) objects without training. The average level of accuracy of face recognition with the addition of tracking algorithm and without tracking is 96%

Published by: The Mattingley Publishing Co., Inc.

and 74.92%. Comparison of the level of accuracy can be seen in Table 1. An accuracy comparison chart is presented in Fig. 4. and Fig. 5.

Table 1. Comparison of The Level of Accuracy

Object	*Tracking*		*Non-Tracking*	
	Detection Amount	**Accuracy (%)**	**Detection Amount**	**Accuracy (%)**
Object1	5	100	69	62.37
Object2	5	100	93	71.49
Object3	5	80	78	60.57
Object4	5	100	63	66.85
Object5	5	100	62	74.26
Object6	5	100	78	85.39
Object7	5	100	78	81.13
Object8	5	100	70	86.33
Object9	5	80	76	78.06
Object10	5	100	65	82.74

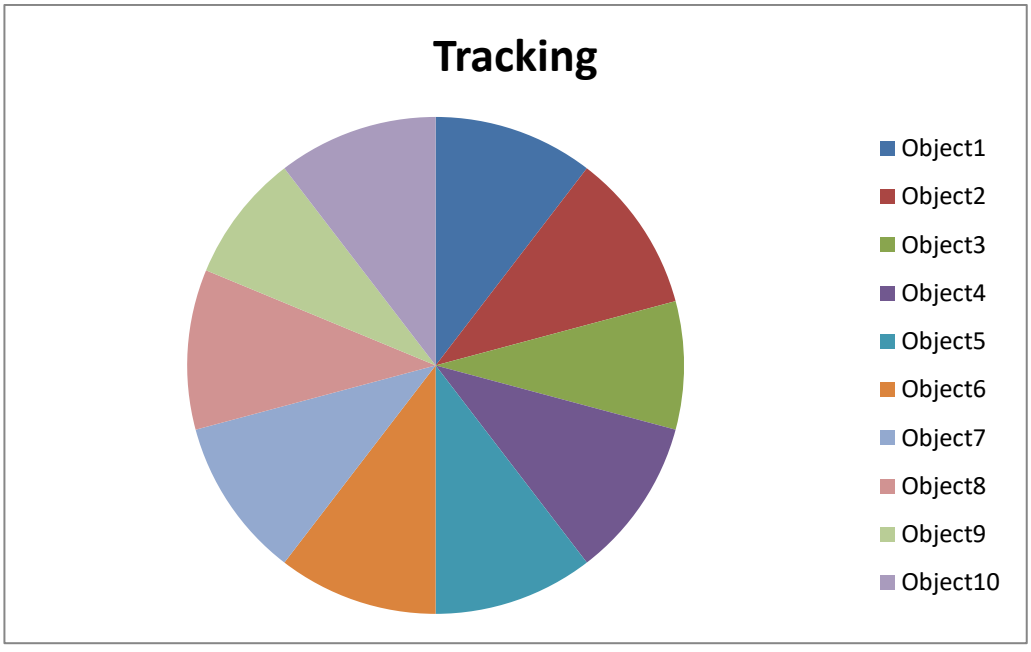

Fig. 4. Accuracy Detection With Tracking

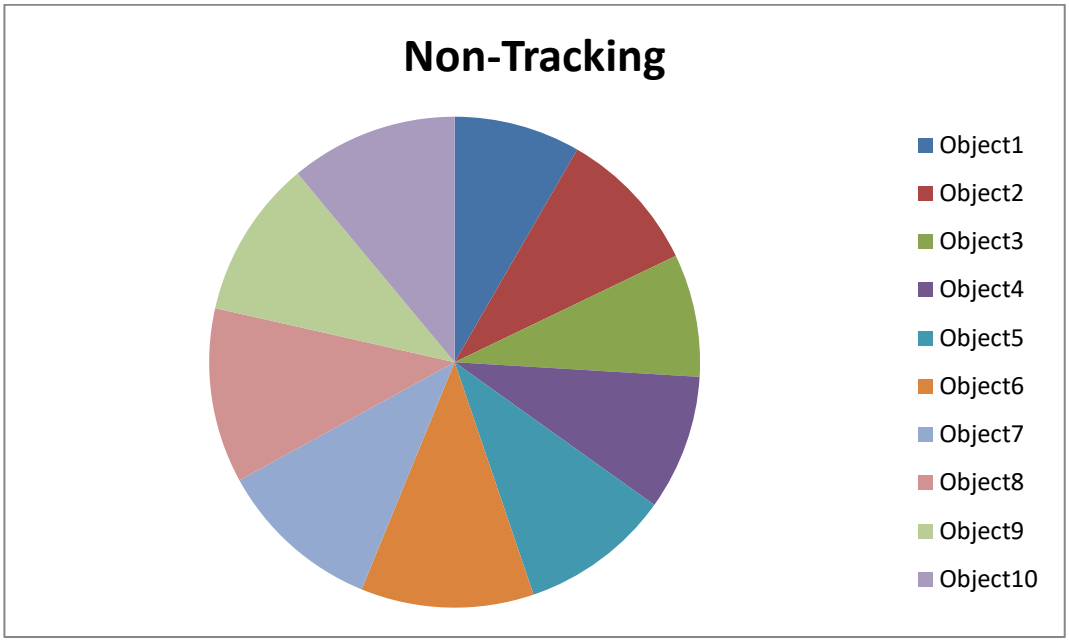

Fig. 5. Accuracy Detection With Non-Tracking

IV. CONCLUSION

This study aims to improve the accuracy of face recognition of the Viola-Jones method on moving objects by adding the Camshift tracking algorithm. The study results show that the implementation of the tracking algorithm significantly increases the accuracy of face recognition when compared to the results without the tracking algorithm. Average recognition accuracy of test results with the addition of tracking by 96%.

V. ACKNOWLEDGEMENTS

The author thanks the Institute of Research and Community Service at Yogyakarta National Veterans Development University for financial support in this research. The author hopes that the results of this study can be useful for other researchers.

REFERENCES

[1] S. Jaiswal, "Evaluation of face recognition methods," *J. Glob. Res. Comput. Sci.*, vol. 2, no. 7, 2011.

[2] S. Zarei, "Face recognition methods analysis," *Int. J. Artif. Intell. Informatics*, vol. 1, no. 1, p. 01, 2018.

[3] N. Boyko, O. Basystiuk, and N. Shakhovska, "Performance Evaluation and Comparison of Software for Face Recognition, Based on Dlib and Opencv Library," *Proc. 2018 IEEE 2nd Int. Conf. Data Stream Min. Process. DSMP 2018*, pp. 478–482, 2018.

[4] R. Saha, D. Bhattacharjee, and S. Barman, "Comparison of Different Face Recognition Method Based On PCA," *Int. J. Manag. Inf. Technol.*, vol. 10, no. 4, pp. 2016–2022, 2014.

[5] M. Roomi and D. M. P. Beham, "A Review Of Face Recognition Methods," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 27, 2013.

[6] D. Wang, G. Zheng, and L. Liu, "Developmental network and its application to face recognition," 2015, pp. 6310–6315.

[7] D. Wang and L. Liu, "Face Recognition in Complex Background: Developmental Network and Synapse Maintenance," *Int. J. Smart Home*, vol. 9, pp. 47–62, 2015.

[8] T. Mantoro, M. A. Ayu, and Suhendi,

"Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods," *Int. Conf. Multimed. Comput. Syst. -Proceedings*, vol. 2018-May, pp. 1–5, 2018.

[9] S. Chinchu, A. Mohammed, and B. S. Mahesh, "A novel method for real time face spoof recognition for single and multiple user authentication," *2017 Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2017*, vol. 2018-Janua, pp. 376–380, 2018.

[10] M. S. I. Sameem, T. Qasim, and K. Bakhat, "Real time recognition of human faces," 2016, pp. 62–65.

[11] R. Sharma and M. S. Patterh, "A new pose invariant face recognition system using PCA and ANFIS," *Optik (Stuttg).*, vol. 126, no. 23, pp. 3483–3487, 2015.

[12] D. Sohsten and S. Murilo, "Multiple face recognition in real-time using cloud computing, Emgu CV and Windows Azure," in *International Conference on Intelligent Systems Design and Applications, ISDA*, 2013, pp. 137–140.

[13] S. R. Dubey and S. Mukherjee, "A Multi-Face Challenging Dataset for Robust Face Recognition," 2018, pp. 168–173.

[14] L. An, M. Kafai, and B. Bhanu, "Face recognition in multi-camera surveillance videos using Dynamic Bayesian Network," in *2012 6th International Conference on Distributed Smart Cameras, ICDSC 2012*, 2012, pp. 1–6.

[15] C. Huang and N. Wei, "Comparison and prospect study on multiple face detection and recognition methods," *Proc. - 2012 Int. Conf. Comput. Sci. Serv. Syst. CSSS 2012*, pp. 1619–1622, 2012.