Face tracking with camshift algorithm for detecting student movement in a class

M Harahap, A Manurung, Priya, A Prakoso, M F Tambunan
Faculty of Technology and Computer Science, Universitas Prima Indonesia, Indonesia
Email: mawaddah@unprimdn.ac.id

Abstract. Face detection (face tracking) has been widely applied for various purposes, including in the fields of entertainment, education and security. Face detection can certainly be done with the camera in real time. For example on a camera on a laptop or camera in a room in real time will detect facial movements. Face detection is implemented using the camshift algorithm. The camshift algorithm works on a search window that can find facial movements in each frame. The camshift algorithm that has been applied can calculate the size and location of the search window that will be used for the next frame. The camshift algorithm can be used for detection such as face detection. The distribution used is hue in the HSV color dimension (Hue, Saturation, Value). The use of this hue distribution is done to overcome differences in human skin color and the background used when taking frames.

1. Introduction
In information technology, biometrics usually refers to technology to measure and analyze the characteristics of the human body such as fingerprints, eyes, sound patterns and facial patterns that were first used for the authentication process[1]. Face Tracking is a way to detect changes in image changes from one frame to another to find its location[2][4][24]. The use of face tracking plays an important role in computer vision and can be widely applied in a variety of applications such as automatic surveillance, traffic monitoring and robot vision[2][21]. Face tracking moves are very complex due to flexible movement changes, changes in light and changes in point of view[3][19]. The use of automatic surveillance is used to be able to find objects in the form of faces on someone in a room. Certainly, the
camera certainly only takes all body objects to the person who enters the room, so it is necessary to look for faces on a person's body object to be able to recognize the person[4][17]. There are several factors that affect the results of face tracking, namely the presence of other objects that cover part of the object, the background of objects that have the same characteristics and colors[5][19][22]. So we need a method that has a good accuracy of resistance to changes that occur in the object[6].

Viola Jones is a method that is quite widely used by other researchers and is quite good in handling face tracking problems in real time, but this method requires a long time to handle, with the position of the front head suitable, and cannot detect the head with face black. The Local Binary Pattern (LBP) method is very effective for describing image features. LBP has advantages such as high speed in an invariant calculation and rotation, facilitating extensive use in the field of shooting, texture checking, facial recognition, image segmentation. However, this method is very difficult to detect faces if there is a small change in the face, and is not accurate enough because it is only used in binary and gray images[7][13][27].

AdaBoost is a method used in enhancements in the field of machine learning. The AdaBoost method has very accurate predictions by combining several relatively weak and wrong rules. This method is easy to apply, and this method is quite good in handling many faces in one image. However, even though this method produces results that are very suitable for the object being searched for, the calculation of this method also requires considerable time[8][15].

The SMQT Features and SNOW Classifier methods are methods that have 2 steps of completion. The first step is face lighting, where there will be a search for pixel information on the image to detect the face. The second step is classifying objects with the aim of getting results from face detection. However, face lighting in the first step can result in the determination of skin color, for example in an image with an area of gray may be detected as a face[9][21][26].

From some research results, the camshift method is very well used for object tracking. Camshift algorithms are used for security, vehicle navigation, surveillance cameras, car driver assistants, biometrics, video games, and industrial automation[10][11][15][20]. The Camshift (Continuously Adaptive Mean-Shift) algorithm known as an improvement to the mean-shift method states that this algorithm has quite good tracking quality on various objects based on the base color of the object [9][23][27]. The CamShift method is also very good if you want to be combined with several other methods in various studies. So this study will design an application that is able to detect faces using the CamShift algorithm.

2. Related research

Zhang NaNa, and Zhang Jin researched about optimizing the improvement of face tracking by using camshift algorithms, this study refers to tracking faces where first looking for a moving object whether the object is someone or another moving object, with these two conditions the result of the object traceable can be determined, if you find a face on the object then it can be determined that the object is someone who was caught on camera, if there is no face on the object it is determined as an ordinary object that moves on the camera[12][23].

Muhammad Haris Khan, John Mc Donagh examines face tracking with various expressions obtained in an image, it aims to be able to capture the best possible face to eliminate a little error when tracking faces using the camshift algorithm, this study has a slight increase because it not only captures faces on the object, but the expression on the face is also determined to get better results, and of course it can still be developed to determine the feeling that is being experienced by the face that is traced to someone[14][28].

Cihan H. Dagli developed research on tracking many faces in one image, where the K-Means method became a calculation that determines how many faces there are in one image, this study aims to collect the number of faces and trace one by one face starting from one face to the face others in the image, use the color of facial skin as a reference and ignore other colors to accelerate proper facial determination[13][29].
3. Methods

3.1. Camshift Algorithm

CamShift stands for Continuously Adaptive Mean Shift, which is the development of the Mean Shift algorithm that is done continuously (repetitively) to adapt or adjust to the ever-changing probability distribution of colors each frame change from the video sequence [18] [23]. While the steps of the CamShift algorithm are as follows: For the image of the probability of skin color distribution, the mean area (centroid) in the search window can be searched by the equation:

Search zeroth moment:

\[
M_{00} = \sum_x \sum_y I(x, y) \quad M_{00} = \sum_x \sum_y I(x, y)
\]  

Searching first moment for x and y:

\[
M_{01} = \sum_x \sum_y x I(x, y) \\
M_{01} = \sum_x \sum_y y I(x, y)
\]

And then location of mean on search window (centroid) is a:

\[
X_c = \frac{M_{10}}{M_{00}} \quad Y_c = \frac{M_{10}}{M_{00}}
\]

where \(I(x, y)\) is the pixel color value in the position \((x, y)\) of the image and \(x, y\) is in the search window.

2-dimensional (2D) orientation of face objects is obtained by performing a second moments calculation with the equation:

\[
M_{20} = \sum_x \sum_y x^2 I(x, y) \\
M_{01} = \sum_x \sum_y y^2 I(x, y)
\]

where the object's orientation is

\[
\theta = \tan^{-1} \left( \frac{2(M_{11} - X_c Y_c)}{(M_{20} - X_c^2) - (M_{02} - Y_c^2)} \right)
\]

If

\[
\alpha = \frac{M_{20}}{M_{00}} \quad b = 2 \left( \frac{M_{11}}{M_{00}} - X_c Y_c \right) \quad c = \left( \frac{M_{11}}{M_{00}} - Y_c^2 \right)
\]

then the length \(l\) and width \(w\) of the centroid distribution are:

\[
l = \sqrt{\frac{(a+c)+\sqrt{b^2+(a-c)^2}}{2}}
\]
The use of these equations in the face tracking system will result in $x, y$, face rotation, length and width (area or $z$ value).

3.2. The Proposed Model
Starting with displaying the form from the application, after the camera is ready to record, then proceed with the process of shooting, it will be continued by counting conversion of RGB to HSV which mean color panels, RGB stands for red, green, blue. The colors formed by the color model are the result of a mixture of primary colors red, green, and blue based on certain compositions. Converting the HSV color panel to Grayscale then to binary do the camshift algorithm with input in the form of the size and location of the search window and the image of the color probability distribution, and save the zeroth moment. Repeat the calculation region in the middle of the search window with a size larger than the search window for each change in the video image frame. When the camshift calculation results are successful, the next step is to create a detection box on the face that is obtained and display the results of the application that has been tested. To explain the processes that occur in the face tracking application using the camshift algorithm, the author uses a flowchart. The shape of the flowchart of the design process that occurs as shown in Figure 1.

![Figure 1. Application Process Flowchart](image-url)
4. The Result and Discussion

Research requires a surveillance camera which is placed in the classroom at an angle that allows it to get some of the faces in the room. This study uses a surveillance camera with a resolution of 1MegaPixel. Security cameras are mounted as high as 2 meters, 2 meters high due to the high right on the camera allows the camera to capture your face and prevent just got a head if it is too high. His discussion uses an example image made and taken from the Faculty of Education of the English Literature Study Program at the Universitas of Prima Indonesia as an example image that will be used in this study. This study uses images as research material to look for faces in several photos. The photo is the result of recording surveillance cameras installed in class. In the recording of the surveillance camera that has been connected to the application, 21 frames of images will be taken from 21 seconds during video recording, as a test material to display the face results of recording several frames, Figure 2 is an example of a picture in the first second. In the captured image from the video camera, then the size of the search window and the initial location of the search window are first determined, then the calculation region in the middle of the search window with a size larger than the search window.

![Sample Photos](image)

*Figure 2. Sample Photos*
In the test photo in the form of photos of 21 frames of video recording taken, because 21 frames are enough to show different movements. Every face that is obtained will be a sample of face scavenging. A total of 21 images were taken and sorted according to the pose of the face contained in the 21 photos as in figure 3. Where there were 21 pieces of photos from the face to be used as data so that the photos could be known to the students in the class.

![Image of Dataset](image)

**Figure 3. Image of Dataset**

From the results obtained, in figure 4 there are 8 students who are in the photo. But in the picture only 6 students were detected. This is caused by two things, namely the first because the student's face pose is not visible in the photo, second because the face pose from the camera distance of more than three meters, because based on the application testing results, students who are more than 3 meters away will not be detected by the camera.
The camera cannot capture the exact position of the face when the face is sideways position, because the lateral position, does not indicate the position of the front face. The position on the front face caught on camera is the face that will be processed by the application to determine who the people on each face are caught. These results must be obtained in accordance with the research that has been done. At a distance of 3 meters, the face position may be caught or could not be caught. Because of the long distance will cause the camera to capture a face increasingly difficult, with the capture of the face tagging faces can not be done. At a distance of 3 meters less, appropriate facial position is at the forefront of the face will be marked and the application will mark the name of the face, the results can be seen in Table 1.

### Table 1 Face Coverage Results Based Distance

| Distance face to the camera | Face Detection | Face Recognition |
|----------------------------|----------------|-----------------|
| ≥ 1 meters                 | ✓              | ✓               |
| ≤ 3 meters                 | ✓              | ✓               |
| > 3 meters                 | ×              | ×               |

Face tracking can be said to be successful if the position of the face that are in the front position. For marking the face name that is a predictive calculation algorithm camshift. Tracking and tagging faces in the application can not tag faces is the man himself, the point here is if the camera captures a facial application will predict the face is A, that when face A camera caught, face signifies the absolute application is A.

### 5. Conclusion

The conclusions that researchers obtained based on the implementation results of this application are face tracking used in this study in the form of a system process determines the time to start and stop recording the entry of students into the classroom. Data synchronization is only done when the program looks for student names with faces in the database and those in the room. The results of the implementation of the camshift algorithm for face tracking in real time using this camera will be fully successful if done in an open room or enough light and the position of the face more perpendicular to the camera, and vice versa the face is often undetectable in closed rooms, poor lighting, and distance objects with a camera are not ideal (too close or too far away). This can still be developed with other cases such as in offices or activities that require face tracking. That way, the Camshift Algorithm can be used to create face tracking applications to detect the faces of students in a room by matching program data in the database and in the room.
References

[1] Sultoni. *Pengenalan Wajah Secara Real Time Menggunakan Metode Camshift, Lapalcial of Gaussian dan Discrete Cisine Transform Two Dimensional (LoGDCT2D).* Universitas Brawijaya. 2016.

[2] Joe Y.M dan Andria W. *Robot Perekam Objek Berbasis Face Tracking.* Universitas Klabat. 2017, ISSN : 2541

[3] Ronsen P dan Irpan, A.P. *Moving Object Tracking Using CHAMSHIFT and SURF Algorithm.* STMIK Mikroskill. April 2015, ISSN. 1412-0100

[4] Ahmad U, 2005, *Pengolahan Citra Digital.* Graha Ilmu. Yogyakarta.

[5] Muhammad R.C. *Rancang Bangun Alat Pengatur Arah Kamera Berdasarkan Pergerakan Kepala Berbasis Sensor Gyroscope.* Universitas Telkom. 2 Agustus 2016, ISSN : 2355-9365

[6] 2221Kadir dan Susanto, 2013, *Teori dan Aplikasi Pengolahan Citra.* Andi. Semarang.

[7] Li J 2015 *Research On Camera-Based Human Body Tracking Using Improved Cam-Shift Algorithm,* International Journal On Smart Sensing And Intelligent Systems Volume(8): No(2).

[8] Joshan, A J and Suresh P. 2013 *Implementation and Comparison of Kernel and Silhouette zased Object Tracking,* International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume(2): Issue(3),

[9] Gi-Woo K, and Dae-Seong K 2015 *Modified Camshift Algorithm Based on HSV Color Model For Tracking Objects,* International Journal of Software Engineering and Its Applications Volume(9), Number(7).

[10] Jog A and Halbe S, 2012 *Object Tracking Using Camshift Algorithm in Open C,* International Journal Of Scientific Research (IJSR), Volume(1): Issue(6).

[11] Bhavish S A, R Desai J, and Saha S, 2012 *A Fast Haar Classifier based Gesture Recognition using Camshaft Algorithm and Curve Fitting Method.*

[12] Joshi S, Gujarathi S and Mirge A, 2014 *Moving Object Tracking Method Using Improved Camshift With Surf Algorith,* International Journal of Advances in Science Engineering and Technology, Volume(2): Issue(2), pp 14-18.

[13] Emami,E, and Fathy M 2012 *A Visual Tracking Algorithm Based on Camshift and Motion Segmentation,* International Journal of Information and Communication Technology Research Volume(4): Number(3), pp 99-106.

[14] Ms. Bedekar K. A and Dr. Mrs. Patil V.V, *Particle Filter and CAMShift Approach for Motion Detection: A Comparative Study,* International Journal of Latest Trends in Engineering and Technology Vol.(7) Issue(3), pp. 364-367.

[15] Salhi A And Ameni Y J, 2012, *Object tracking system using Camshift, Meanshift and Kalman filter,* International Journal of Electronics and Communication Engineering Vol:6, No:4.

[16] Shao-Fan Lien, Kuo-Hsien Hsia and Juhng-Perng Su, 2015, *Moving Target Tracking based on CamShift Approach and Kalman Filte,* Appl. Math. Inf. Sci. 9, No. 1, pp 395-401

[17] Ferdinan and Suryana Y, *Camshift Improvement With Mean-Shift Segmentation, Region Growing, And Surf Method,* International Journal of Communication & Information Technology (CommIT) ,Vol. 7 No. 2 Oktober 2013, pp. 53-64.

[18] Wang A, Li J, Lu Z, *Improved Camshift with adaptive searching window,* International Journal of Soft Computing And Software Engineering (JSCSE), ISSN: 2251-7545, Vol.2, No.3, 2012.

[19] Emami,E, and Fathy M 2012, *Object Tracking Using Improved CAMShift Algorithm Combined with Motion Segmentation.*

[20] P. Hidayatullah, H. Konik, *CAMSHIFT Improvement on Multi-Hue and Multi-Object Tracking,* International Conference on Electrical Engineering and Informatics pp 17-19 July 2011, Bandung.

[21] Smolyanskiy S , Huitema C, Liang L and Anderson E S, *Real-time 3D face tracking based on active appearance model constrained by depth data,* Image and Vision Computing 32 (2014) pp 860–869

[22] Setyatia E, Alexandre D and Widjaja, *Face Tracking Implementation with Pose Estimation*
Algorithm in Augmented Reality Technology, Social and Behavioral Sciences 57 ( 2012 ) pp 215 – 222

[23] N Zhang and J Zhang, Optimization of Face Tracking Based on KCF and Camshift, 8th International Congress of Information and Communication Technology (ICICT-2018), Procedia Computer Science 131 (2018) pp 158–166

[24] Francis F and Franck L, Theory of evidence for face detection and tracking, International Journal of Approximate Reasoning 53 (2012 ) pp 728–746

[25] Musab C, Sencer Ü, Implementation of Tracking of a Moving Object Based on Camshift Approach with a UAV, Procedia Technology, Volume 22, 2016, pp 556-561.

[26] Yunji Zhao Hailong Pei, An Improved Vision-Based Algorithm for Unmanned Aerial Vehicles Autonomous Landing, Physics Procedia 33 (2012 ) pp 935 – 941

[27] John G. Allen, Richard Y. D. Xu, Jesse S. Jin, Object Tracking Using CamShift Algorithm and Multiple Quantized Feature Spaces. Madsen Building F09, University of Sydney, NSW 2006

[28] H Khan, J McDonagh, Georgios Tzimiropoulos, Synergy between face alignment and tracking via Discriminative Global Consensus Optimization.

[29] Cihan H. Dagli, Multi-Pose Face Recognition And Tracking System, Procedia Computer Science 6 (2011) pp 381–386