Technium Vol. 4, No. 9 pp.47-52 (2022)
ISSN: 2668-778X
www.techniumscience.com

Real-time Face Tracking for Service-Robot

Dhuha Basheer Abdullah [1]
Prof.dhuha_basheer@uomosul.edu.iq
Mohanad Rafaa Mohammed [1]
mohanad.20csp81@student.uomosul.edu.iq
[1] Computer Science Department, University of Mosul

Abstract. Real-time human face tracking is suggested in this paper for use while interacting with robots. The system consists of two main parts, the first works to discover the human face, determine its location in relation to the original image, and find the dimensions of the face to be used in the second part of this system. The second part receives the location and dimension of the face and tracks it by controlling the movement of the camera according to the offset between the interval. In order to detect human faces for the earlier job, the Haar cascade method is used, whereas the Kanade-Lucas-Tomasi (KLT) algorithm is used for face tracking under various circumstances. As a trace output from the prior stage, the camera is offset by its offset between image frames. The findings of the experiments demonstrate that real-time tracking of human faces was successful even when the participants were donning glasses, hats, or face-side positions. At a maximum frame rate of 26 fps, experiments were conducted.
Keywords. Real-time System, Face Tracking, Haar cascade algorithm, KLT Algorithm.

1. Introduction

In the present and the future, human-robot interaction (HRI) will be a prominent area of robotics. The robot's purpose is to serve as a helpful helper who can aid humans in their tasks. The robot must be able to interact and communicate well with people. In this scenario, the tracking of human faces becomes the primary necessity for the vision system in these robots. Applications for face tracking systems dealing with forward orientations are numerous. The Kanade-Lucas-Tomasi (KLT) method may be used to resolve the issue in real time while utilizing several face modes. Based on a variety of facial traits, this system allows for precise face tracking. The extraction, selection, and tracking of the best characteristics are all steps in this algorithm's procedure for monitoring human faces. The haar cascade method is used for face detection to gather facial information.

2. Related Works

The face tracking system in the vision robot system is first supported by the face detection system [1]. The Haar cascade algorithm [4] is one of several strategies that have been used to quickly and accurately recognize human faces [3][2]. The Haar cascade method is an object identification technique that is 15 times faster than its contemporary rival and has a pretty good accuracy level of 93.7% [5].

Real-time face identification and tracking systems have been discussed, put into practice, and used in a variety of domains [6] through [10]. The face tracking objective idea can later be extended to smart equipment like robots in order to pursue moving face targets and look for lost face targets [7], [10].
Based on object movement registration techniques like optical flow, human face tracking is then enhanced by Kanade, Lucas, and Tomasi (KLT algorithm). The suggested method for monitoring human faces continuously inside a video frame employs the KLT algorithm. In comparison to previous tracking techniques, this algorithm performs better [11]–[15]. Tracking is difficult because the characteristics of the face frequently vary as a result of lighting, facial expression, orientation, and the usage of accessories.

3. System Overview

The proposed face tracking system primarily employs the Haar cascade algorithm for face recognition and the KLT algorithm for face tracking. The procedure for synchronizing the tracking system and the actuator device is also described in this section.

3.1. Face Recognition

Face recognition is a technology that is able to match the human face taken from an image or video with a group of faces stored in the database by identifying certain features in the face by analyzing the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw.
Haar cascade is an Object Detection Algorithm used to identify faces in an image or a real time video. The algorithm uses edge or line detection features . The algorithm is given a lot of positive images consisting of faces, and a lot of negative images not consisting of any face to train on them. The model created from this training is available at the OpenCV GitHub repository.
The repository has the models stored in XML files, and can be read with the OpenCV methods. These include models for face detection, eye detection, upper body and lower body detection, license plate detection etc. The subset of all 6000 features will again run on the training images to detect if there's a facial feature present or not. Now the authors have taken a standard window size of 24x24 within which the feature detection will be running. It's again a tiresome task.
To simplify this, they proposed another technique called The Attentional Cascade. The idea behind this is, not all the features need to run on each and every window. If a feature fails on a particular window, then we can say that the facial features are not present there. Hence, we can move to the next windows where there can be facial features present.
Features are applied on the images in stages. The stages in the beginning contain simpler features, in comparison to the features in a later stage which are complex enough to find the nitty gritty details on the face. If the initial stage won't detect anything on the window, then discard the window itself from the remaining process, and move on to the next window. This way a lot of processing time will be saved, as the irrelevant windows will not be processed in the majority of the stages.
The second stage processing would start, only when the features in the first stage are detected in the image. The process continues like this, if one stage passes, the window is passed onto the next stage, if it fails then the window is discarded as shown in the figure.

3.2. Face Tracking

The KLT algorithm handles face tracking, following any previously identified moving facial characteristics. The KLT chooses the best traits from the forehead, brows, eyes, nose, cheeks, lips, and chin [14], [17], which are where the majority of the features are concentrated. The selected characteristics are then tracked by optical flow in the following phase. The following model serves as a representation of the optical flow calculations:

$$(x,y,t)=(x+u,y+v,t+1) \quad (1)$$

Where the intensity of a certain time and position of image $(x,,)$ and its neighbour position of $I(x+u,y+v,t+1)$ on the next time frame is same [14].

3.3. Interaction with Device

A camera is a component of the proposed real-time system that feeds the processing portion of the system, as illustrated in Figure 1. Tracking the location of the human face in the actual environment is the primary objective of the hardware system. So, the x and y coordinates of the bounding box acquired by face tracking in the picture plane are sent as the tracking algorithm's output to Arduino. A pair of DC servo motors that are each assigned to rotate the webcam according to an imagined x-axis and y-axis carry out the rotation (two degrees of freedom).

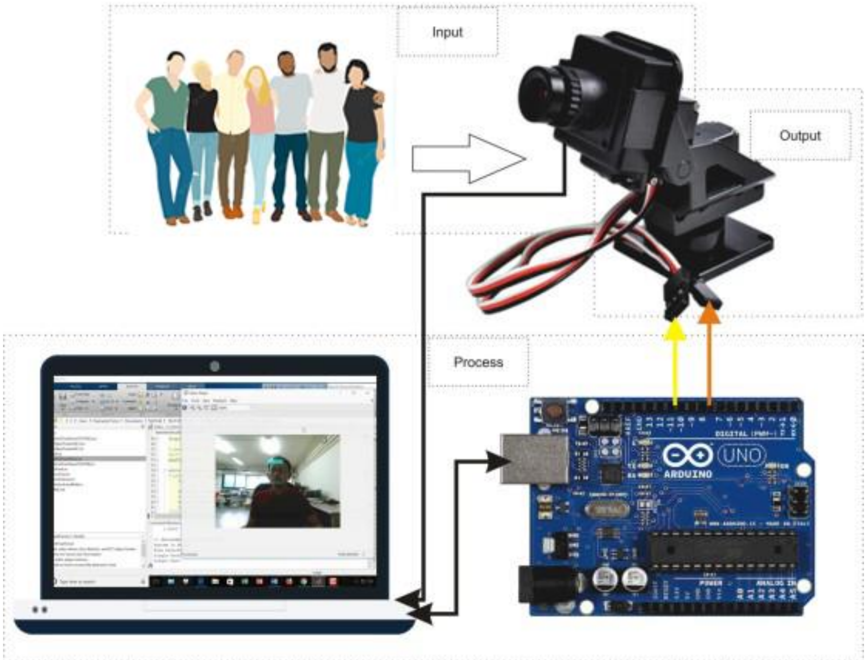


Figure 1. Interface between Devices

Based on the value of the error, the actuator will move. If the face is not in the centre of the image, the system will display an error value. Vertical error for the y-axis and horizontal error for the x-axis make up the error value. Fig. visually depicts the error position for the identification of a single face. Due to the tiny change in position detection that frequently takes place while the face is already at the centre point, the yellow region is a tolerance area as in figure 2. close to the centre point that helps to stabilize the movement of the servo actuator. The resolution of the live frame that was taken is indicated by the x-size and y-size. In the experiments, the resolution was 640 x 480.

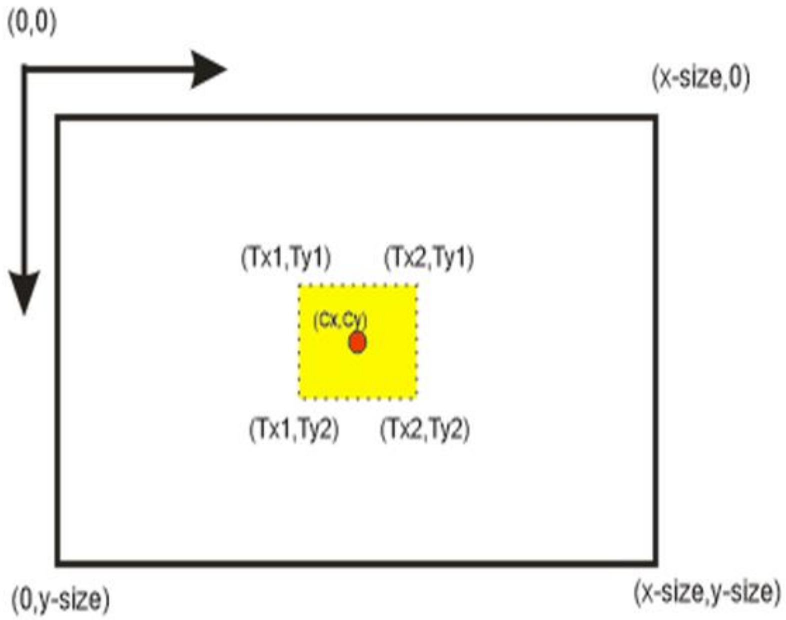


Figure 2. The tolerance area of the centre image

The yellow portion of the image is a threshold region that surrounds the center of the frame and measures 60 by 60 pixels (Cx,Cy). The lowest boundary area (Tx1 for the x-axis and Ty1 for the y-axis) and maximum boundary area represent the threshold area (Tx2 for x-axis and Ty2 for y-axis). By manipulating the actuator that moves the camera, the system will maintain the face in the yellow box region. In figure 3, the tracking system algorithm is displayed. The target face's centre area is determined by the threshold value, which is produced by the Bbox point centre extraction to generate the centre point bounding box (Xc, Yc). The extracted centre point is then compared to the area's threshold value, which determines how quickly the servo motor acts.

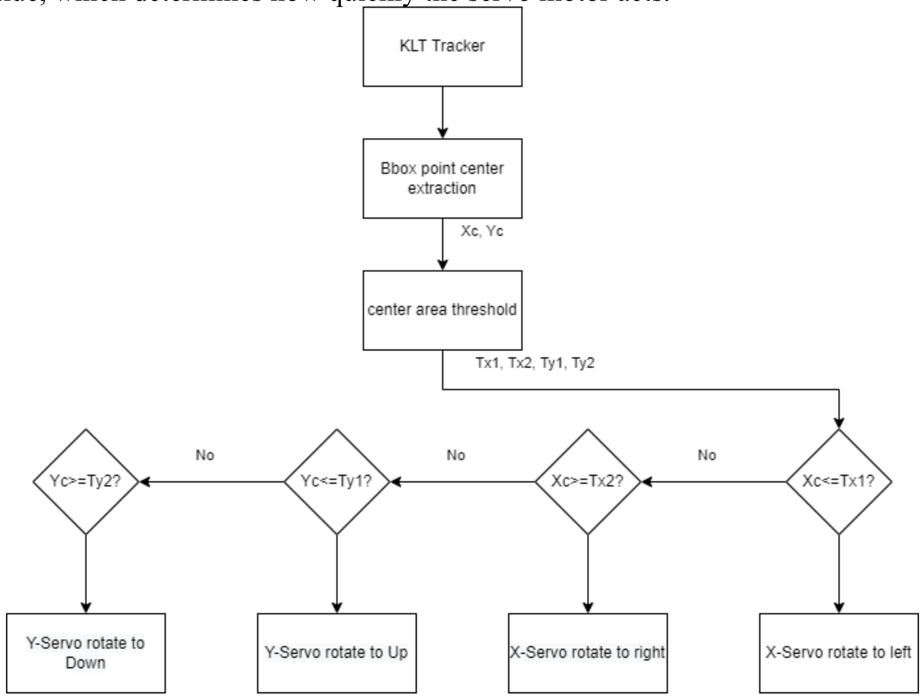


Figure 3. Flowchart of tracking system

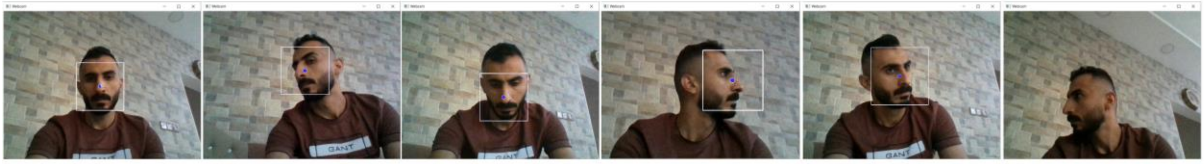
Technium Vol. 4, No. 9 pp.47-52 (2022)
ISSN: 2668-778X
www.techniumscience.com

4. Experiment Results

Experiments were conducted on the following hardware: Intel(R) Core(TM) i5-7y-57 CPU @ 1.20GHz 1.61 GHz, Ram 8.0 GB, Window 10 PRO 64-bit operating system, x64-based processor. In the evaluation of the proposed system performance, the successful system was considered when a human face was correctly detected and tracked. For all tests lighting conditions were stable.

4.1. Face Detection Result

Face detection must take place before tracking. The real-time performance is possible using the Haar cascade approach. Numerous studies have demonstrated that face identification is capable of coping with situations including various head orientations. Figure 5 displays the detection findings. If all of the eyes and brows are not visible, or when the face is in the extreme positions depicted like in figure 5(f), the proposed system fails. One of the causes of failure process that occurs in face detection is no detection of complete eye features.

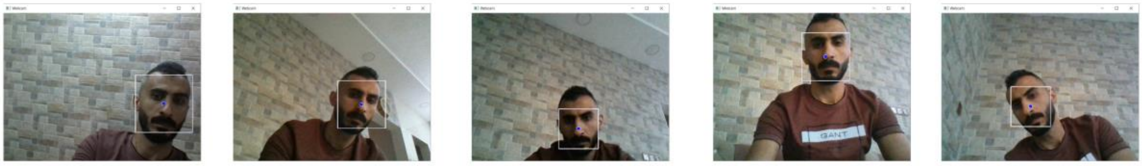


(a) (b) (c) (d) (e) (f)

Figure 5. Experiments Results of face detection

4.2. Face Tracking Result

It is decided which aspect of the markings on the face is the best. The KLT approach maintains feature detection while still giving a bounding box, identifying features that move in accordance with visual flow movements. The complete completion of the tracing process is depicted in Figure 6. Only after the face is completely covered does the tracing stop operation take place. Even though just a few tiny characteristics were discovered, the tracking system was nevertheless able to perform tracking in extreme situations. Real-time video is converted into 1,000 continuous photos to evaluate the resolution system. However, when the subject wore glasses and a mask, the results were subpar. Since no facial characteristics were found, the face tracking in this instance is unsuccessful.



(a) (b) (c) (d) (e)

عناوين الفجر

Figure 6. Experiments Results of face tracking

5. Conclusion

In order to facilitate the process of robot engagement with humans while carrying out its activities, this study introduces a real-time facial tracking system. The suggested approach can retain a face's position in the middle of the frame area while following a single face. Even while using accessories and under harsh situations, this system will always track the face and maintain it in the centre of the picture. Although the face detection procedure must be completed before face tracking, good results were obtained when conducting the tracking, but the results of face detection under harsh conditions were not excellent. The system will be improved in future work to track several faces using a more reliable face detection method.

References

[1] J. We, Wang. W, "Face Detection And Recognition For Home Service Robots With End-To-End Deep Neural Networks", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), New Orleans, LA, USA, 5-9 March 2017.

[2] Z. Tao, L. Jingjing, J. Wenjia, S. Jun, Y. Hiuhua, "Fast and robust occluded face detection in ATM surveillance", Pattern Recognition Letters (2017)

[3] V.E. Machaca Arceda et al. "Fast Face Detection in Violent Video Scenes", Electronic Notes in Theoretical Computer Science 329 (2016) 5–26.

[4] V. Paul, J. Michael, "Robust Real-time Face Detection", International Journal of Computer Vision 57(2), 137–154, 2004.

[5] P. Aditya, P. Erick, S. Deni, "Face Detection with Various Angular Positions on a Bunch of People by Comparing Viola-Jones and Kanade-Lucas-Tomasi Methods" , National Journal of Informatics Engineering Education (JANAPATI), Volume 5, Number 3, December 2016.

[6] C. Anitha, M.K. Venkatesha, B. Suryanarayana Adiga, "Real Time Detection and Tracking of Mouth Region of Single Human Face", 3rd International Conference on Artificial Intelligence, Modelling and Simulation (AIMS),Kota Kinabalu, Malaysia, 2-4 Dec. 2015.

[7] K. Maneesha, N. Shree, D. R. Pranav, S. K. Sindhu, C. Gururaj, "Real time face detection robot", 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 19-20 May 2017.

[8] Z. Ji, C. Zhiqian, "Real Time Face Detection System Using Adaboost and Haar-like Features", 2nd International Conference on Information Science and Control Engineering, Shanghai, China, 24-26 April 2015.

[9] L. Jianchao, Z. "Real-time Face Detection During the Night", International Conference on Systems and Informatics (ICSAI), Hangzhou, China, 11-13 Nov. 2017.

[10] G. Yutong, L. Xuewei, J. Hongyan, "Real-time Multi-view Face Detection based on Optical Flow Segmentation for Guiding the Robot", 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2015.

[11] B. Ritesh, P. M. Jasmine, "Face Detection And Tracking Using Klt And Viola Jones", ARPN Journal of Eng.and Applied Sciences, Vol. 11(23), ISSN 1819-6608, December 2016,

[12] C. Debmalya, C. Saravanan, "Comparative Study of Camshift and KLT algorithms for Real Time Face Detection and Tracking Applications", Second International Conference on Research in Computational Intelligence (ICRCICN), India, 23-25 Sept. 2016.

[13] P. Heri, "Control of 3D Object-Based Augmented Reality Movement Feature Feature with POSIT", JNTETI, Vol. 4, No.1, February 2015.

[14] B. Vinodh, T. Sunitha, M. Snehasis," Multiple object tracking by improved KLT tracker over SURF features", National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), India, 16-19 Dec. 2015.

[15] N. Binu, F. Jacob, T. Richard, D. Yakov, A. Theus, A. Vijayan, "Multi-Pose Face Recognition And Tracking System", Procedia Computer Science, Vol. 6, 2011, Pg 381-386, open access.

[16] M. D. Putro, T. B. Adji, B. Winduratna,"Adult image classifiers based on face detection using Viola-Jones method", 2015 1st International Conference on Wireless and Telematics (ICWT), Manado, Indonesia, 17-18 Nov. 2015

[17] Shi, Tomasi, "Good Features to Track", Conference on Computer Vision and Pattern Recognition (CVPR94), June 1994.

[18] Mendoza Torres, G., & Ródriguez Gómez, G. (2021). A Fuzzy Control MIMO for Stabilization and attitude of a Cube Sat. Technium: Romanian Journal of Applied Sciences and Technology, 3(1), 12–22. Retrieved from https://techniumscience.com/index.php/technium/article/view/2621

G. B. N. (2020). IRAD: The use of pesticides in maize culture: Health and environmental impacts on human beings. Technium: Romanian Journal of Applied Sciences and Technology, 2(4), 61–67. https://doi.org/10.47577/technium.v2i4.880