Face Recognition Method Based on Fixed and PTZ Camera Control for Moving Humans

Seong-Hoon Kim
Department of IT Convergence Engineering
Gachon University
Seongnam-si, Republic of Korea
Email:limiteddaily@gachon.ac.kr

Jin-Woo Choi
Culture Technology Institute
Gachon University
Seongnam-si, Republic of Korea
Email:jwchoi@gachon.ac.kr

***Abstract*— Face recognition methods have been generally used for authentication purposes, in which subjects position their faces in the camera field. Lately, face recognition has begun to be used for security applications at specific locations as well as for authentication. Face recognition for security purposes involves human and face detection based on the image created by a fixed camera. Consequently, the success rate for face recognition is low for low-resolution cameras. To address this problem, this study proposes a face recognition method in a multi-camera environment that includes both a fixed and a pan-tilt-zoom (PTZ) camera. The proposed method involves detecting humans within the fixed camera's field of view while tracking the detected person by adjusting the PTZ camera's pan and tilt to keep the person in view. By zooming simultaneously with tracking to obtain a high-resolution facial image, the proposed method outperformed conventional face recognition methods that use only a fixed camera.

Keywords—Human Detection; Object Tracking; Face Detection; Face Recognition; PTZ Camera Control

I. Introduction

Face recognition technology employed for user authentication typically involves an access control system or a smartphone camera. To improve face recognition, a high-resolution facial image is required, and so subjects need to position their faces in the camera's field of view [1, 2]. Lately, face recognition has begun to be used to secure specific locations by recognizing people in the field of view. However, this method typically uses a fixed camera, so that if the person is far from the camera, a poor-resolution facial image will likely result. Moreover, since people are typically in motion, human detection, face detection, and face recognition need to be performed quickly before the person leaves the camera's field of view. Therefore, object detection typically uses a Haar-based detection algorithm, while face recognition uses algorithms such as principal component analysis (PCA) and linear discriminant analysis (LDA) [3, 4]. These algorithms run swiftly, but may yield low-accuracy outcomes, depending on image resolution. Therefore, this study proposes a face recognition method in a multi-camera environment involving both a fixed and a pan-tilt-zoom (PTZ) camera. This method involves human detection and tracking using the fixed camera; the PTZ camera is used to shoot the subject detected by the fixed camera with pan and tilt values corresponding to the person's coordinates, while tracking the person in motion and performing face recognition by zooming in to obtain a high-quality facial image.

II. Proposed Method

A. Pan and Tilt Matching between PTZ and Fixed Camera

The face recognition method proposed in this study uses the pan and tilt values obtained for the person acquired by the fixed camera to track that person using the PTZ camera, which is also used for facial detection.

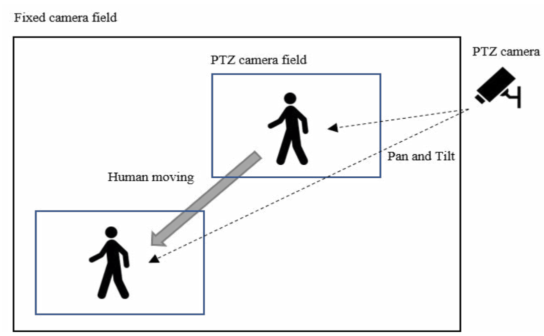

Fig. 1. Example of using a PTZ camera to track a subject detected in a fixed camera shot

Reference pan and tilt values are identified by obtaining the matching points between the fixed and PTZ cameras using nine 20×20 reference zones (Fig. 2) and locating the matching points at the center of the PTZ camera's field of view.

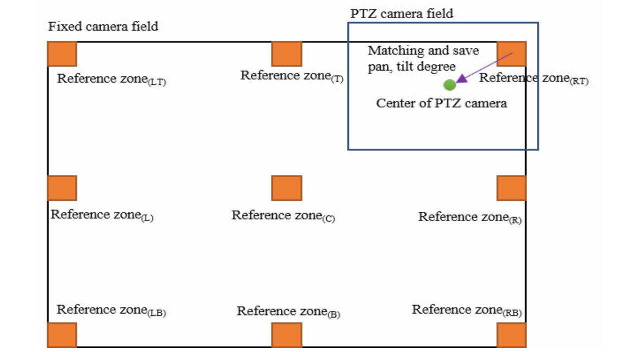

Fig. 2. Nine reference zones for calculation of the pan and tilt values identified by matching between the fixed and PTZ cameras

The nine pan and tilt reference values are calculated using four pan and tilt values proximal to the location of the person detected by the fixed camera (Fig. 3).

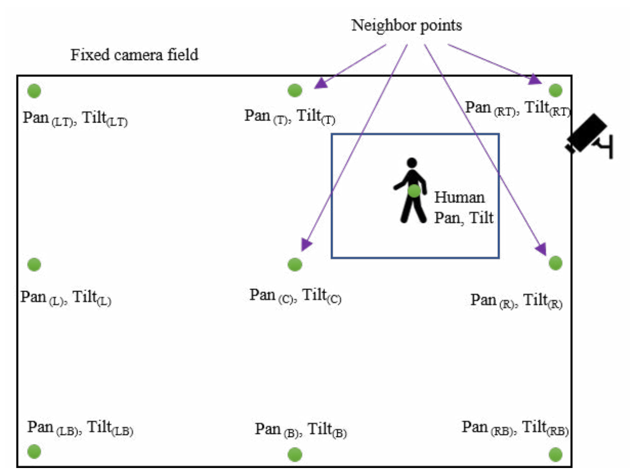

Fig. 3. Example of calculating pan and tilt reference values for a subject located in the fixed camera shot, using four proximal pan and tilt values

B. *Human Detection and Recognition*

Human detection is performed at every 10th fixed-camera frame, and a detected person is added as a new tracking target if there is currently no person being tracked. If a person is already being tracked, the similarity between the newly detected person and the person already being tracked is measured, and if the similarity measure is lower than the set threshold, the detected person is regarded as a new individual and added as a tracking target. Otherwise, the detected person is not added as a new tracking target. For human detection, a Haar cascade–based detection algorithm was used; for object tracking, we used a kernelized correlation filter (KCF) algorithm [5, 6].

C. *Face Detection and Recognition*

Face detection and recognition were performed using a PTZ camera. Face detection involved the PTZ camera simultaneously tracking and zooming in on the moving person to obtain high-quality facial images; a Haar cascade–based face detection algorithm was again used, as in the human detection aspect. For face recognition, feature extraction and recognition were performed using linear discriminant analysis (LDA) [7, 8, 9].

III. Experimental Results

A. *Experimental Environment*

An experiment to determine the accuracy of the proposed method was conducted by setting up one fixed and one PTZ camera, both manufactured by Axis Communications and having specifications as shown in Table 1.

TABLE I. FIXED AND PTZ CAMERA SPECIFICATIONS

Model	Items	Specification
AXIS-M2026LE Fixed Camera	Effective pixel	2688 × 1520
	Angle of view	Vertical: 73° Horizontal: 130°
	Encoding	H.264, Motion JPEG
AXIS-P5515 PTZ Camera	Effective pixel	1920 × 1080
	Focal length	3.8–42.9 mm, 20×
	Pan and tilt ranges	Pan: 360° Tilt: 1.8°–180°
	Angle of view	Vertical: 33.3°–3.1° Horizontal: 59.2°–5.2°
	Encoding	H.264, Motion JPEG

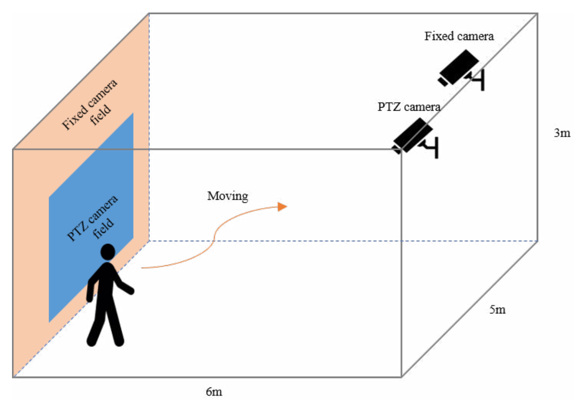

Fig. 4. Arrangement of fixed and PTZ cameras for the experiment

B. *Comparison of Recognition Rate between Fixed-Camera and Multi-Camera Environments*

The facial images acquired using the proposed and conventional methods, followed by normalization to equal size, reveal a clear difference in resolution (Fig. 5).

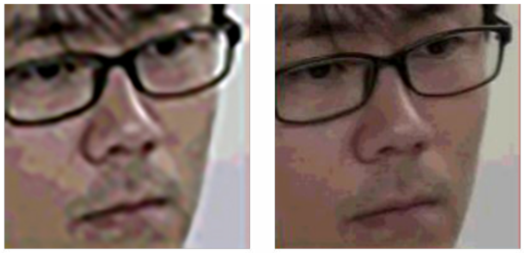

Fig. 5. Compared resolution of facial images obtained using conventional and proposed methods

A recognition-rate experiment was conducted by obtaining face recognition rates at varying distances (Table 1). The results suggest that the improvement in facial-image resolution of the proposed method over the conventional method is greater when the subject is farther away. The improvement in recognition rates was approximately 6% on average.

TABLE II. FIXED AND PTZ CAMERA SPECIFICATIONS

Method	**Distance (m)**	**Recognition rate (%)**	**Average of recognition rate (%)**
Conventional	2 m	91.7%	85.5%
	4 m	84.2%	
	6 m	80.5%	
Proposed	2 m	92.1%	91.1%
	4 m	91.5%	
	6 m	89.8%	

IV. CONCLUSIONS AND FUTURE WORK

This study proposed a face recognition method for a person in motion in a multi-camera environment utilizing one fixed camera and one PTZ camera. In the conventional fixed-camera face recognition method, face-recognition rates drop off as the subject's distance from the camera increases, due to the low resolution of the facial image extracted. In the proposed method, the moving person is detected and tracked first by the fixed camera and then by the PTZ camera. The PTZ camera acquires the facial image by both tracking and zooming in on the person and performs face recognition using an LDA algorithm. The proposed method yielded a recognition rate that was approximately 6% higher than the conventional fixed-camera method. However, when the person moved very quickly, face recognition failed due to difficulty in tracking. This issue suggests the need to shorten the time required involving pan-tilt-zoom, which in turn requires both algorithm acceleration research and an improvement in PTZ camera hardware performance.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2017R1A6A3A11034944).

REFERENCES

[1] P. M. Kumar, U. Gandhi, R. Varatharajan, G. Manogaran, R. Jidhesh and T. Vadivel, "Intelligent Face Recognition and Navigation System using Neural Learning for Smart Security in Internet of Things", Cluster Computing, The Journal of Networks, Software Tools and Applications, 2017.

[2] H. Ai and X. Cheng, "Research on Embedded Access Control Security System and Face Recognition System", Measurement, vol. 123, pp.309-322, 2018.

[3] S. D. Kakade, "A Review Paper on Face Recognition Techniques", International Journal for Research in Engineering Applications & Management, vol. 2, pp.1-4, 2016..

[4] M. J. Sawarkar, "A Review on Different Face Recognition Techniques", Internaltional Journal of Scientific Research in Computer Science, Engineering and Information Technology, vol. 5, pp.207-213, 2019.

[5] J.F. Henriques, R. Caseiro, P. Martins and J.Batista, "High-speed Tracking with Kernelized Correlation Filters", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015..

[6] X. Xie, F. Wu and Q. Liu, "Object Tracking based on KCF and Sparse Prototypes", 2018 International Conference on Control and Computer Vision, pp.69-74, 2018.

[7] L. Elrefaei, A. Alharthi, H. Alamoudi and S. Almutairi, "Real-time Face Detection and Tracking on Mobile Phones for Criminal Detection", 2017 2nd International Conference on Anti-Cyber Crimes, 2017.

[8] R. Vedantham, L. Settipalli and E.S. Reddy, "Real Time Facial Expression Reconition in Video Nearest Neighbor Classfier", International Journal of Pure and Applied Mathematics, vol.118, pp.849-854, 2018.

[9] K. Shrivastava, S. Manda, P.S. Chavan, T.B. Patil and S.T. Sawant-Patil, "Conceptual Model for Proficient Automated Attendance System based on Face Recognition and Gender Classification using Haar-Cascade, LBPH Algorithm along with LDA Model", International Journal of Applied Engineering Research, vol.13, pp.8075-8080, 2018.