Environment Independent Hybrid
Face Recognition System
Using a Fixed Camera and a PTZ Camera

Euisun Kim\textsuperscript{1}, Green Bang\textsuperscript{1}, Daniel Chung\textsuperscript{2}, Ilju Ko\textsuperscript{1}

\textsuperscript{1} Department of Media, Soongsil University, 369 Sangdoro, Dongjak-gu, Seoul, 156-743, Republic of Korea
\{kes33, banglgreen, andy\}@ssu.ac.kr
\textsuperscript{2} Department of Sports Information Technology, Soongsil University, 369 Sangdoro, Dongjak-gu, Seoul, 156-743, Republic of Korea
danielec@ssu.ac.kr

Abstract. With the growth of the high resolution CCTV camera market, face recognition from farther distances is required in addition to short distances. Thanks to advanced face recognition technology, faces can be recognized without any concern for the condition of the area around the face. In this paper, we used two cameras, the fixed format camera and the PTZ camera, a fixed format camera and a PTZ camera, to separate face detection and face recognition functionalities for load distribution of the face recognition module. This face recognition method can be integrated with global video monitoring, unmanned technology, etc. We suggest that an environment independent hybrid face recognition system enables the face detection at far distances with multi-channel CCTV video input processing. We can expect a new intelligent video security monitoring market that uses the new face recognition system.

Keywords: object detection, environment independent face recognition, distributed processing

1 Introduction

Due to the demand for crime prevention in modern societies, high resolution CCTV cameras are used widely. These cameras have replaced the existing low resolution CCTV cameras. This trend shows that face recognition technology using high resolution CCTV cameras is important. As high resolution CCTV cameras provide high quality images, face detection is easier and the detected face can be tracked continuously [1]. Thanks to advanced face recognition technology, faces can be recognized without any concern for the condition of the area around the face. Therefore, we can implement an environment independent face recognition system. To implement this system, we need a new face recognition method that uses video monitoring schemes with a combination of short and far distances. This face recognition method can be
integrated with global video monitoring, unmanned technology, etc. We can expect a new intelligent video security monitoring market that uses the new face recognition system. In Japan, companies such as NEC, and OMRON are interested in face recognition. NEC developed an application with face detection and recognition, as well as gender detection [2]. In the United States, companies such as Microsoft, Apple, and Google are interested in face recognition technology and have invested in accordingly. Face recognition technology for short distances is sufficiently mature, but face recognition system for far distances is still behind [3].

In this research, we used two cameras, a wide-area fixed format camera and a Pan-Tilt-Zoom (PTZ) camera, instead of only one camera. By using these two cameras, we separated the Object Detection (OD) and Face Detection (FD) functionalities, whereas objects are the body area of a person. In this way, the load on the face recognition system can be reduced, and the face recognition module works with multiple channel images from different sites simultaneously. In this reason, the face detection performance in the far distance is improved, and the faces can be recognized from the multiple channel input videos simultaneously [4] [5]. Figure 1 shows the overall architecture of this face recognition system.

![Fig. 1. Overall architecture of the hybrid face recognition system](image)

### 2 Face detection using a fixed format and a PTZ camera

In this section, we propose a face detection method using a fixed format camera and a PTZ camera. First, the background image from a far distance is captured by the fixed format camera. Second, using the PTZ camera, we identify and extract the differing image from the input background image. Third, we find and separate the object from the differing image. Last, we control the PTZ camera to zoom in on the object we found.
Figure 2 shows the procedure of image processing using a fixed format camera. We make the background model using the background image without objects from the input image of the wide area monitoring camera. The fixed format camera tracks objects in the background considering motion prediction and the direction of motion using the background model. The face detection method works using the masked image of the separated object. The PTZ camera tracks the object using the information the PTZ camera control module received: the masked image and the location of the object.

![Flow of applying the background model](image)

**Fig. 2.** Flow of applying the background model

In Figure 3, we select the Region of Effect (ROE) from the four quadrants of the PTZ camera input image. We can use this method thanks to the high image quality of the high resolution PTZ camera. Then, we select a specific ROE and scan each different ROE for efficient image processing. This method is efficient because of the reduction of the scanning range and time of the PTZ camera and the load reduction of the face detection system caused by filtering outside the ROE which does not apply to image processing. Since the input image quality of the fixed format camera is quite high at 720p, we can detect multiple faces with the advantage of high resolution input. We use presets for short and far distance objects to focus. Therefore, we can detect objects automatically, and actively monitor. The presets of the PTZ camera provide unmanned monitoring and efficient camera management functions. Using these functions, we can move the camera and monitor suspicious objects in any specific area automatically.

![Screenshot of applying the background model](image)

**Fig. 3.** Screenshot of applying the background model
Figure 4 shows the window of object detection with four quadrants of input images. This window also shows the specific quadrant containing the detected face itself and the masking and tracking area of the detected face. Figure 5 shows the result of the configuration of the monitoring area corresponding to the environment of the installed cameras. We configure the monitoring area by using the face detection area selection of the PTZ camera.

![Screenshot of the partitioned search area of the input video](image1)

**Fig. 4.** Screenshot of the partitioned search area of the input video

![Example of selecting the effective area for face detection](image2)

**Fig. 5.** Example of selecting the effective area for face detection: (a) selecting the effective detection area, (b) selecting the distance of the target

Recently developed CCTV cameras that support high resolution videos can detect multiple faces automatically. In addition, the camera can detect an object without using the PTZ method by magnifying the specific area containing the object [6]. In this paper, we use the automatic face detection and tracking method for multiple faces from the input image of the PTZ camera if the camera supports high resolution. Figure 6 explains the method of ROE selection in an image from a PTZ camera.
3 Experiments and conclusion

In this paper, we proposed a hybrid face recognition system with multiple face detection that can process input from multiple CCTV cameras. This system can recognize up to 15 faces and process up to 200 simultaneous CCTV videos with the face recognition module using a distributed thread method. Figure 7 shows the processing procedure of the face recognition module. Once a face is detected, the detected face information is clipped and that clipped face information is transferred to the face recognition module. This module can process in real-time, and we can develop the module further considering the processing power of individual systems.
Figure 8 shows the face detection and recognition results. In (a), the face detection module searches and detects multiple faces simultaneously and indicates the detected faces in rectangles. In (b), the face recognition module thread recognizes the detected faces using a distributed process with a transferred input order. This shows that face detection and object detection can be executed in a local site, and that face recognition can be executed in a remote site with expandable functionalities.

![Figure 8](image)

**Fig. 8.** Face detection and recognition results: (a) Simultaneous multiple face recognition result, (b) Face recognition results of the transferred face image

In our system, the face detection module can be separated from the other parts. The face detection module reads the video input from the PTZ camera in a local site and detects faces from the input. Then, the face detection module transfers the detected image to the face recognition server. In this way, our system can be flexible with regard to how it is configured. As the face recognition server can be separated from the other parts of the system, the monitoring coverage can be extended with additions and upgrades to the servers. We expect the smart video monitoring market to be formed as security monitoring is required for detecting and recognizing people who enter and leave public areas such as boarding platforms in a subway station [6, 7].

**Acknowledgment** This work was supported by Institute for Information & Communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) (No. B01901520330004003, Non environment sensitive Face Detection Core Technology Development)

**References**