International Journal of Internet, Broadcasting and Communication Vol.10 No.1 16-22 (2018)
https://doi.org/10.7236/IJIBC.2018.10.1.3

IJIBC 18-1-3

Real-time Tracking and Identification for Multi-Camera Surveillance System

Yo-Hoon Hong[1], Seung June Song[2], Jungkyu Rho[3]*
[1]*Sechang Instruments, Korea,* [2]*Dae Han Instrument, Korea*
[3]*Dept. of Computer Science, Seokyeong University, Korea*
dcnmoon@gmail.com, junesong@dhinst.co.kr, jkrho@skuniv.ac.kr

Abstract

This paper presents a solution for personal profiling system based on user-oriented tracking. Here, we introduce a new way to identify and track humans by using two types of cameras: dome and face camera. Dome camera has a wide view angle so that it is suitable for tracking human movement in large area. However, it is difficult to identify a person only by using dome camera because it only sees the target from above. Thus, face camera is employed to obtain facial information for identifying a person. In addition, we also propose a new mechanism to locate human on targeted location by using grid-cell system. These result in a system which has the capability of maintaining human identity and tracking human activity (movement) effectively.

Keywords: *Surveillance system, heterogeneous camera, human identification and tracking, human segmentation, realtime Tracking*

1. Introduction

With the emerging digital signage technology, a smart content delivery system (SCDS) for digital advertisement is needed. The ultimate goal of SCDS is delivering digital advertisement content based on customer's personal interest. To be able to deliver content to customers based on their interest, one of the key components is a sophisticated surveillance system. This surveillance system should be able to identify humans and track their activity on targeted area, such as: indoor store or exhibition room, so that it is possible to deliver advertisement contents which match to customer's interest.

Human tracking is a task of finding location of a person and estimating its location continuously. Generally, it is done by first locating initial position of a person, and then calculating the next position by analyzing the surveillance video/camera stream frame-by-frame. The location of the tracked person is usually represented by a bounding box (2D rectangle) which encapsulate the human body.

There are several tracking algorithms which have been developed such as MIL[1], BOOST[2], TLD[3], and Cv-Track[4]. BOOST uses an on-line AdaBoost feature selection algorithm which updates object classifier while tracking the object. TLD divides the process into three tasks: tracking, learning, and detection which introduces P-N learning method to estimate tracking error. MIL uses multiple instance learning to build an adaptive appearance model for object tracking. CvTrack uses probabilistic appearance model to track objects. Since none of BOOST, TLD, and MIL have proven to successfully track multiple

Manuscript Received: Jan. 3, 2018 / Revised: Jan. 17, 2018 / Accepted: Jan. 25, 2018
Corresponding Author: jkrho@skuniv.ac.kr
Tel: +82-2-940-7753, Fax: +82-2-6292-1099
Dept. of Computer Science, Seokyeong University, Korea

targets, CvTrack is the only suitable algorithm to our system because of its ability to track multiple targets.

Human identification by using facial information is commonly known as face recognition. In general, it consists of two main process: face registration and recognition. Most of face recognition algorithms perform off-line training process which requires several face images to be registered before recognition process. However, we need a solution which is able to do on-line face training process and robust to various face poses.

In this paper, we propose a system which combines human tracking and human identification for surveillance system. We assemble several existing algorithms which can supports us in terms of tracking and identifying human. We use bird's eye view dome camera to observe targeted location and RGB-D face camera to acquire facial information in detail. In addition, we propose a novel approach for localizing human using grid-cell system.

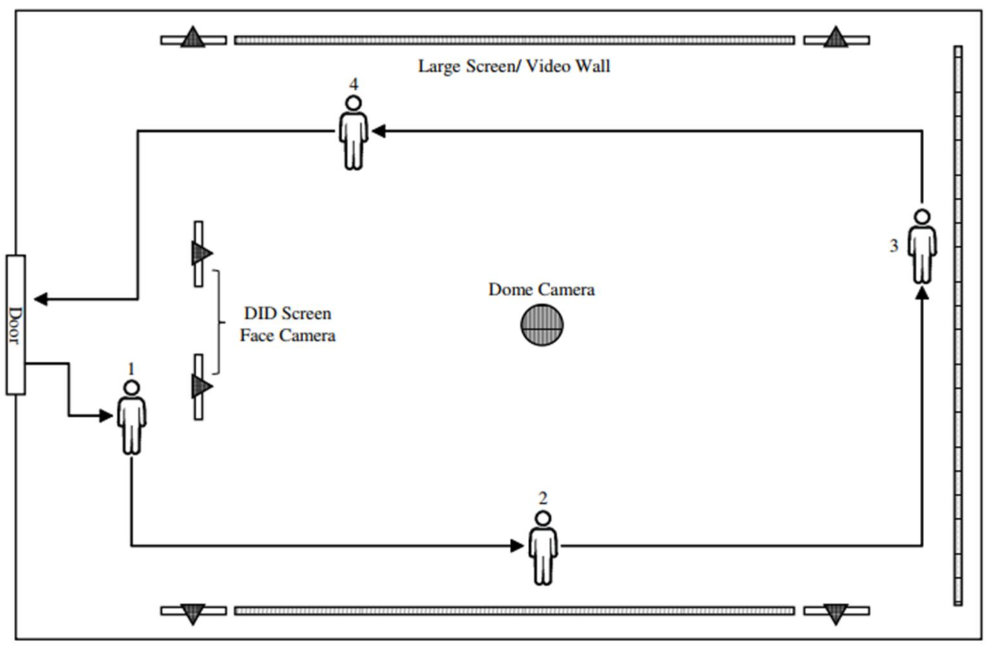

Figure 1. System Illustration

Figure 1 illustrates the system that we develop. A person walks in and interested in an advertisement content displayed on a screen. Our human identification system extracts personal identity based on facial information and our human tracking system extracts the location of that person. Whenever that person moves to another location in the screen, our human tracking system will track the movement so that additional or relevant advertisement contents can be displayed on the larger screens (video wall). Our human tracker will continuously follow the movement, until that person is going out of sight.

The rest of this paper is organized as follow. Section 2 shows the overview of system that we develop. Section 3 describes human tracking in detail. Section 4 explains about human identification using facial information. Finally, we conclude our contribution in Section 5.

2. Overview

2.1 System Architecture

In general, there are two cameras: dome and face camera, two software modules: human tracker and face recognizer, and server application. Human tracker will use dome camera observation to track human on the targeted location. If there is a new tracked human, it will send track-id, position, and time information to server. It will also send a notification message to server as long as the tracked human is moving. Face recognizer will send face-id, age, gender, and time information to sever. It will also notify the server whenever recognized face is going out-of-sight. Server application will receive all notification messages from human tracker and face recognizer. Figure 2 depicts general system architecture that we develop.

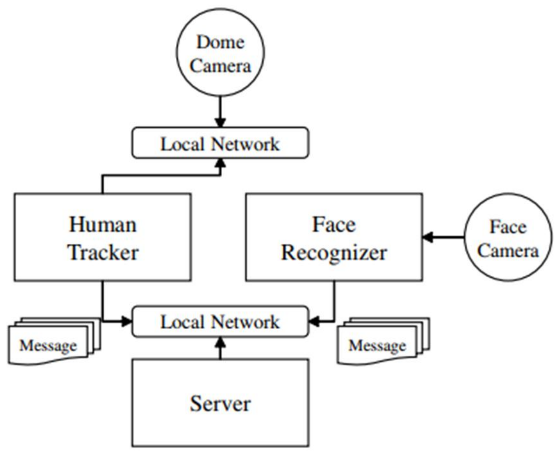

Figure 2. System Architecture

2.2 Hardware Components

Our system uses two types of cameras to observe humans in targeted location: dome and face camera. The dome camera is employed specifically for human tracking and the face camera is used mainly for human identification.

2.2.1 Dome Camera

A 360-degree IP-based (Samsung SNF-7010) [5] is used as dome camera. It will be installed on the roof facing towards the floor resulting a bird's eye view of the entire location. The motivation of using 360 degree bird-eye view point is to minimize occlusion among objects.

2.2.2 Face Camera

In addition to dome camera, our system also employs RGB-D camera (Intel RealSense F200) [6] as face camera. It captures not only image (RGB pixels) but also depth information. By using RGB-D camera, we can capture human facial feature better than RGB-only camera, so that the face recognition can perform better. It will be installed on top of the DID (Digital Information Display) screen. In addition, this face camera is also accompanied by software development kit (SDK) which includes face recognition functionality.

2.3 Software Modules

In accordance of two main tasks that will be performed by our system: human tracking and identification, there are two software modules which are built to do each task. For simplicity, we named these two modules as Human Tracker and Face Recognizer.

2.3.1 Human Tracker

Human tracker is used for detecting and tracking human movement. It can track several humans which are visible to the dome camera. It uses images captured by dome camera as input, and generates a track ID for each person. Figure 3 shows the block diagram of this module.

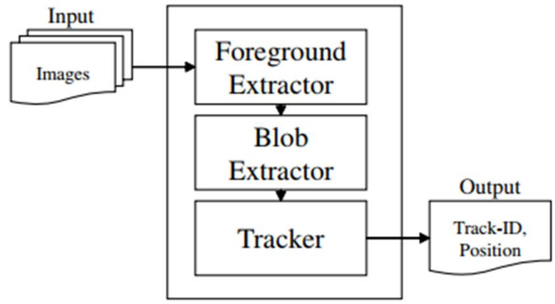

Figure 3. Human Tracker Block Diagram

There are several components, such as: foreground extractor, blob extractor, and tracker, included in this module. Foreground extractor analyzes every image captured from dome camera and extracts foreground image mask which represents moving person. Blob extractor encapsulates the body moving person. Tracker is used for following the movement of that person. As the result, this module will emit the person's track-id and position.

2.3.2 Face Recognizer

This module is mainly used for human identification process. It uses images and face information captured by RGB-D camera as input, and generates several information such as: face-id, age, and gender. Figure 4 shows the block diagram of this module.

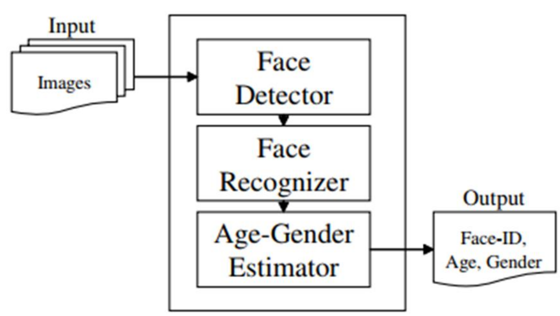

Figure 4. Face Recognizer Block Diagram

There are several components which are required to build this module: face detector, face recognizer, and age-gender estimator. Face detector checks whether there is one or more faces on image captured by face camera. If there is a face object, then face recognizer tries to recognize the person by using the facial information. Whenever a face is not recognized, then face recognizer will register the new face, extract its profile: age-gender, and store it in the face database. Finally, this module will emit the personal identity of a person: face-id, age, and gender, as output.

3. Human Tracking Implementation

In this section, we explain about the implementation of human tracker module. We describe the processing flow, data structure, server message, and user interface design of this module.

3.1 Processing Flow

Initially, this module will stream images from dome camera. And then captured images will be analyzed one-by-one to obtain the foreground mask which represents the moving object image mask. Next, the blob extractor will encapsulate human body by creating bounding box. After extracting blobs of human body, we will use CvTrack to track human movement. As long as the tracked human is visible by the dome camera, there will be messages sent to the server notifying the current location of the tracked human.

There are several properties of CvTrack which can be adjusted such as: minimum-maximum blob size, minimum active frames, maximum inactive frames, and maximum track distance. Minimum and maximum blob size can be adjusted to eliminate irrelevant blobs. For example, we can eliminate other moving objects which are irrelevant by defining the average size of human body. Minimum active frames can be used to make sure that the moving object is a walking human. Maximum inactive frames can be used to eliminate human which are having passive activity on the targeted location. And maximum track distance can be used to determine whether a track and a blob match. We can configure these properties based on various circumstances, such as: the height of dome camera position and the size of observed room.

3.2 Data Structure

To support the operation of human tracker, there are several data types which must be store by the system. There are 3 data types which are going to be used to develop this module: dome camera configuration data, tracked human information, and message which sent to server.

3.3 Server Message Specification

After all of the processing steps have been done, Human Tracker module will send a message to server. The message contains Track-ID, Position, Direction, and Current Date-time.

3.4 User Interface Design

To help user while interacting with our system, there are several UI components that should be presented on the screen such as: Menu strip contains menu access for changing dome camera capture settings and server address, Dome Stream Picture Viewer shows images captured by dome camera, Server Message Text Area shows messages which are sent to the server.

4. Human Identification

In this section, we explain about human identification process using face recognizer module in more detail. We describe the processing flow, data structure, server message, and user interface design of this module.

4.1 Processing Flow

Initially, this module will acquire an image from face camera. After that, it will try to detect the existence of face on the captured image. If there is one or more faces within the image, then it will try to recognize the face. An unrecognized face will be registered as a new face by assigning a new face ID. After that, the newly registered face will be analyzed further to estimate age and gender. At the end, it will be stored to the face database. Meanwhile, if the detected face is recognized, then the system will get the face ID as well as its age and gender information. Finally, all new face and already recognized face information will be sent to server. Figure 5 shows the processing flow of this module.

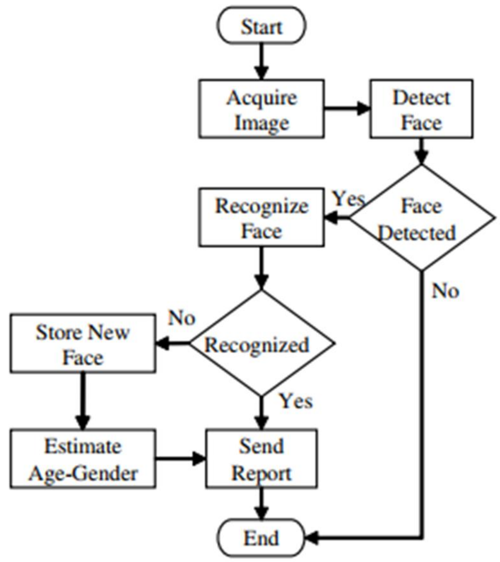

Figure 5. Human Identification Processing Flow

4.2 Data Structure

There are several entities which will be used to represent data during face recognition process: Face Capture represents capture device data, Recognized Face represents face information, Server Report represents message sent to server.

4.3 Server Message Specification

After all of the processing steps have been done, Face Recognizer module will send a message to server. The message contains Face-ID, Age, Gender, Position, and Current Date-Time.

4.4 User Interface Design

To help user while interacting with Face Recognizer, there are several UI components that should be presented on the screen such as: Menu strip contains menu access for changing face camera settings and server address, Camera Stream Picture Viewer shows images captured by face camera, Server Message Text Area shows messages which are sent to the server.

5. Conclusion

In this paper, we introduced a surveillance system which can identify and track human using heterogeneous camera. To track human, we use a dome camera and develop human tracker module which analyzes and follows human movements. And to identify human, we employs face camera to obtain facial information and develop face recognizer module to recognize human. A grid-cell system is used for locating human location on the targeted location. We combine the location information reported by human tracker and face recognizer, so that human position can be located precisely.

Acknowledgement

This work was supported by Convergence Technology Development Project, Ministry of SMEs and Startups [S2274464, Development of multilateral multi-touch device and platform based on multi-touch and image recognition technologies].

References

[1] B. Babenko, M. Yang, and S. Belongie, “Visual Tracking with Online Multiple Instance Learning,” 2009.

[2] H. Grabner, M.Grabner, and H. Bischof, “Real-Time Tracking via On-line Boosting,” in *Proc. 17th British Machine Vision Conf.*, pp. 47-56, 2006.

[3] Z. Kalal, K. Mikolajczyk, and J. Matas, “Tracking Learning Detection,” in *IEEE Trans. Pattern Anal. Mach. Intell.*, (Washington, DC, USA), July 2012.

[4] A. Senior, A. Hampapur, Y. Tian, L. Brown, S. Pankanti, and R. Bolle, “Appearance Models for Occlusion Handling,” in *Image and Vision Computing*, vol. 24, 2006.

[5] Samsung,“Samsung CCTV SNF 7010,” Retrieved from *http://www.samsungsecurity.com/product/product_view.asp?idx6930*

[6] Intel, “Intel RealSense 3D Camera (F200),” Retrieved from *https://software.intel.com/en-us/realsense/f200camera*

[7] J.G. Song, Y.S. Lee, W.T. Jang, H.J. Lee, T.Y. Kim, “Face Recognition Authentication Scheme for Mobile Banking System” Journal of Internet, Broadcasting and Communication (IJIBC), Vol.8 No.2, pp.38-42, 2016

[8] S.M. Woo, M.L. Lee, “Sensors Network and Security and Multimedia Enhancement” Journal of Internet, Broadcasting and Communication (IJIBC), Vol.8 No.1 pp.64-68, 2016