To appear in Biometrics: Theory, Applications and Systems (BTAS10), Washington DC, September 2010

# PTZ Camera Assisted Face Acquisition, Tracking & Recognition

Hyun-Cheol Choi, Unsang Park, and Anil K. Jain

***Abstract*— Face recognition systems typically have a rather short operating distance with standoff (distance between the camera and the subject) limited to 1∼2 meters. When these systems are used to capture face images at a larger distance (5∼10 m), the resulting images contain only a small number of pixels on the face region, resulting in a degradation in face recognition performance. To address this problem, we propose a camera system consisting of one PTZ camera and two static cameras to acquire high resolution face images up to a distance of 10 meters. We propose a novel camera calibration method based on the coaxial configuration between the static and PTZ cameras. We also use a linear prediction model and camera control to mitigate delays in image processing and mechanical camera motion. The proposed system has a larger standoff in face image acquisition and effectiveness in face recognition test. Experimental results on video data collected at a distance ranging from 5 to 10 meters of 20 different subjects as probe and 10,020 subjects as gallery shows 96.4% rank-1 identification accuracy of the proposed method compared to 0.1% rank-1 accuracy of the conventional camera system using a state-of-the-art matcher.**

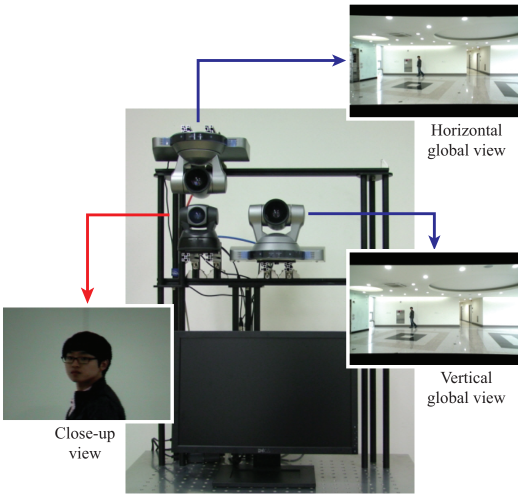


Fig. 1. Proposed camera system and images of global and close-up views.

## I. INTRODUCTION

Video surveillance has gained wide attention and interest due to the increasing concerns about security. Face recognition in surveillance environments is crucial to identify potential terrorists and criminals on a watch list. While the performance of face recognition systems has improved substantially [1], [2], the intrinsic (expression, aging, etc.) and extrinsic (pose, illumination, etc.) variations are still the major bottlenecks in face recognition. Face recognition at a distance (e.g., in surveillance scenarios) introduces another challenge, namely the low image resolution problem. Typical commercial face recognition engines require face images with at least 60 pixels between the eyes for successful recognition, which is difficult to achieve in many surveillance scenarios.

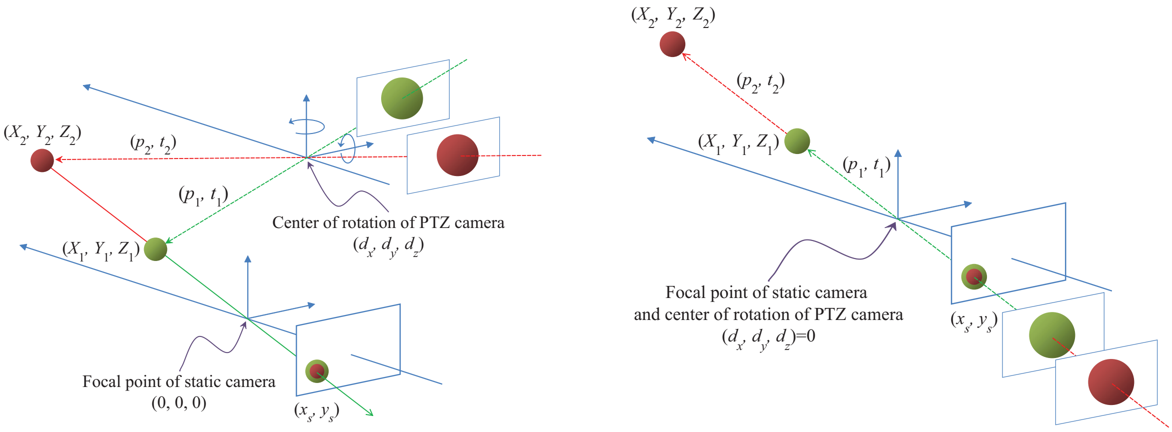
There have been only a few studies reported on face recognition at a distance. These approaches can be essentially categorized into two groups: (i) generating a super resolution face image from the given low resolution image and (ii) acquiring high resolution face image using a special camera system (e.g., a high resolution camera or a PTZ camera). Dedeoglu et al. [3] recognized faces in low resolution images using the super-resolution method. Park et al. [4] proposed a stepwise reconstruction of a high-resolution facial image based on the Extended Morphable Face Model. The performance of their system is highly dependent on the training data and the recognition accuracy rapidly drops when the image resolution drops below 16×16. Yao et al. [5] used a high magnification static camera to capture face images at a very long distance (50∼300 m). However, the camera does not provide pan and tilt motion, resulting in a very small field of view. Bernardin et al. [6] proposed an automatic system for the monitoring of indoor environments using a single PTZ camera. However, their system requires frontal face in every frame to properly control the PTZ camera and they have to zoom out when it fails to detect the face. Marchesotti et al. [7] used a pair of static and PTZ cameras to capture high resolution face images. Hampapur et al. [8] used multiple static cameras and a PTZ camera to accurately estimate the 3D world coordinates of a subjects face and then zoom into the face to capture a high resolution image. Stillman [9] used multiple static cameras to estimate the location of a person in a calibrated scene and the PTZ camera tracks the detected face. Most of these systems rely on the reconstruction of 3D world coordinates or poorly defined relationship calibration between static and PTZ cameras. The 3D world coordinate estimation is computationally expensive and not suitable for real time applications. Tistarelli et al. [10] summarized the problems in face recognition at distance as low image resolution, out of focus, interlace in video images, and motion blur.

In this paper, we propose a camera system with dual static cameras and a PTZ camera to capture face images

H.-C. Choi is with the Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea hcchoi@korea.ac.kr

U. Park is with the Department of Computer Science and Engineering, Michigan State University, E. Lansing, MI 48824, USA parkunsa@cse.msu.edu

A. K. Jain is with the Department of Computer Science and Engineering, Michigan State University, E. Lansing, MI 48824, USA and with the Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea jain@cse.msu.edu



(a) Non coaxial configuration

(b) Coaxial configuration

Fig. 2. Schematics of non coaxial and coaxial camera systems. The two targets projected on the same point in the image plane correspond to the same (different) pan and title angle of PTZ camera in (non) coaxial camera system.

at a distance of up to 10 meters. We propose (i) a novel calibration method to coordinate the static and PTZ views and (ii) PTZ camera control framework in terms of the speed of pan and tilt motion to smoothly track and capture the target face. A face recognition test with 20 probe subjects captured at a distance of between 5 to 10 m and 10,020 gallery subjects shows a rank-1 identification accuracy of 96.4%.

## II. Camera Calibration

### A. Problem Formulation

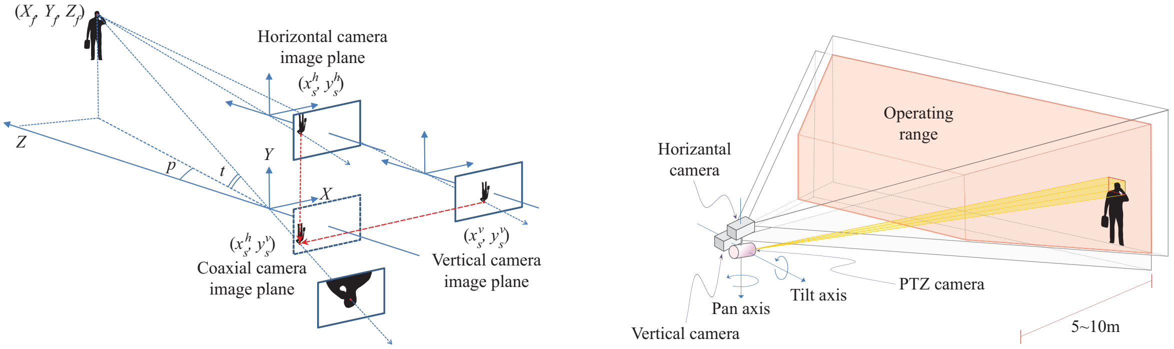
We first define the variables used to describe the proposed camera system.

- $(X, Y, Z)$ : 3D real world coordinate
- $(X_f, Y_f, Z_f)$ : face location in the real world coordinate system
- $(x_s^i, y_s^i)$ : image coordinate of the $i^{th}$ static camera
- $(x_{s,f}^i, y_{s,f}^i)$ : face location in the image coordinate of the $i^{th}$ static camera
- $(x_{ptz}, y_{ptz})$ : image coordinate of the PTZ camera
- $(p, t, z)$ : pan, tilt, and zoom parameters to control the PTZ camera

Our objective is to drive the PTZ camera via the $(p, t, z)$ parameters toward the face location $(X_f, Y_f, Z_f)$ to capture a high resolution face image (at least 60 pixels of interpupillary distance). To determine the desired $(p, t, z)$ values, we can either try to directly estimate $(X_f, Y_f, Z_f)$ or use the relationship between $(x_s^i, y_s^i)$ and $(p, t, z)$. A number of different camera systems which use various combinations of multiple PTZ and static cameras can be considered. We will review some of the representative camera systems and introduce our proposed approach.

- Single PTZ camera [6]: face location is first estimated in the zoomed-out view and the camera is controlled to acquire a high resolution face image. Camera needs to be continuously zoomed in and out because it is very easy to lose track of moving subjects.
- Single PTZ and single static camera [7], [11]: face location is estimated in the static view and the PTZ camera is controlled to capture a high resolution face image. Due to the lack of depth information ($Z$ coordinate), it is difficult to accurately obtain the $(p, t)$ values.
- Single PTZ and dual (multiple) static cameras [8], [9]: multiple static views allow stereographic reconstruction to estimate $(X, Y, Z)$. However, the stereographic reconstruction is computationally expensive and has a limited operational distance. Multiple static cameras are sometimes considered only to increase the surveillance coverage. Our proposed method belongs to this category. However, we use multiple static cameras to achieve the coaxial camera configuration that enables a larger standoff.
- Single static high resolution camera [5]: by using a telescope attached to the camera, face image can be acquired at a large distance, but the field of view is severely limited. By using a high definition video camera, the field of view is increased, but the operational distance is much limited compared to the system using PTZ cameras.

Various camera calibration processes are required in the aforementioned camera systems to correlate the world coordinates, image coordinates of static cameras and parameters to control the PTZ cameras. We propose to use one PTZ camera and two static cameras with relative camera calibration scheme between the image coordinate of static camera, $(x_s^i, y_s^i)$, and PTZ camera parameters, $(p, t, z)$. Compared to other camera systems proposed in the literature, our approach has the following advantages: (i) calibration process does not involve the world coordinates, (ii) only one relative calibration process is required and the calibrated system can be easily deployed in other locations with no addition calibration, (iii) with the novel coaxial calibration scheme, face images can be captured irrespective of the distance between the camera and subject, and (iv) by using subjects location predictor and staged camera speed control, we



(a) Schematic of the proposed co-axial camera system

(b) Schematic operational range of the co-axial camera system

Fig. 3. Schematic of the proposed coaxial camera system: Two static cameras are placed above and beside the PTZ camera to generate the virtual camera in a coaxial position w.r.t. the PTZ camera.

obtain a smooth PTZ camera control capability.

### B. Coaxial Camera Calibration

Conventional camera calibration process typically refers to calculating the relationship between the world coordinate and static image coordinate systems, which has been well studied in the literature [12], [13]. The calibration process in PTZ camera systems for the high resolution face image acquisition involves calculating the relationship between the world coordinate and $(p, t, z)$ parameters via the image coordinates of static camera, where the calibration between the world coordinate and static image coordinate is not needed. Therefore, the calibration process involves calculating the mapping function from $(x_s, y_s)$ to $(p, t)$. The zoom $(z)$ parameter is obtained based on the estimated object size (see Sec. II-C). The mapping function $F$ can be calculated by a linear equation using a set of corresponding ground truth values of $(p, t)$ and $(x_s, y_s)$ as:

$$\begin{bmatrix} p \\ t \end{bmatrix} = F \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} = \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ 1 \end{bmatrix} \quad (1)$$

We find a set of corresponding $(x_s, y_s)$ and $(p, t)$ point pairs by manually driving the PTZ camera to a number of different positions (15 in this case) in the static view. Even though a non-linear mapping function showed slightly better result in terms of the residual error in our experiments, we chose to use the linear method for computational efficiency.

An image coordinate in the static view $(x_s, y_s)$ could have originated from two different world coordinates $(X_1, Y_1, Z_1)$ and $(X_2, Y_2, Z_2)$. As a result, the desired $(p, t)$ values can be different depending on these world coordinates. However, the pan and tilt angle values of PTZ camera can be mapped on to the pixel position $(x_s, y_s)$ in the image scene regardless of the distance $Z$ where static and PTZ cameras are arranged in a coaxial configuration, as shown in Fig. 2(b). A coaxial configuration of static and PTZ cameras has the following properties: (i) focal point of static camera and center of rotation of PTZ camera coincide $(0 = (d_x, d_y, d_z))$ and (ii) the camera axes of all the cameras are parallel. Due to the difficulty in designing such a hardware system, we propose a camera system that effectively satisfies the requirements of coaxial camera configuration as shown in Fig. 3(a). We configure two static cameras, one above (horizontal camera), and one beside (vertical camera) the PTZ camera, so that the $X$ coordinate ($Y$ coordinate) of the horizontal (vertical) camera's focal point coincides with the $X$ coordinate ($Y$ coordinate) of the PTZ cameras's center of rotation. All cameras are also configured to have parallel camera axes. The mapping function from the static image coordinate to the pan-tilt parameters can thus be estimated as $(p, t) = F(x_s^h, y_s^v, 1)$. Fig. 3(b) shows the schematic of the large operating distance of the proposed coaxial camera system. Fig. 4 shows the superiority of the coaxial system over the non-coaxial system. The mapping function $F$ is calculated at $Z$=5 m and the face images are captured in the five to ten meter distance range in both the systems. The coaxial system captures the face in the center of the image at all distances, while the non-coaxial system misses the face as the distance increases. The coaxial camera system can also be operated at a distance of less than 5 m or larger than 10 m. However, we determined that face detection in the static view is not reliable (too small) when the distance is larger than 10 m. The operating distance of our system can be increased by using higher definition static cameras or multiple PTZ cameras to zoom into the subjects face at multiple levels. This is the topic of our ongoing work.

### C. Face Localization and Zoom Control

We use a background subtraction method [14] followed by the morphological opening operation to obtain the silhouette of the subject in the field of view (Fig. 5(c)). After object detection, we compute the vertical and horizontal projection histograms of the silhouettes in each frame. The projection histograms are thresholded and a rectangular area is extracted to localize the region of interest that contains the subject. The location of the head is estimated using the ratio of the height of the head and silhouette (Fig. 5(d)). The head location is estimated from the horizontal and vertical static cameras as $(x_s^h, y_s^h)$ and $(x_s^v, y_s^v)$. Then $(x_s^h, y_s^v)$ is used to estimate the desired pan and tilt parameters $(p, t)$.

We use the height of detected object in static camera images for zoom control. We manually measure ten magnification factors of the PTZ camera that ensures that the

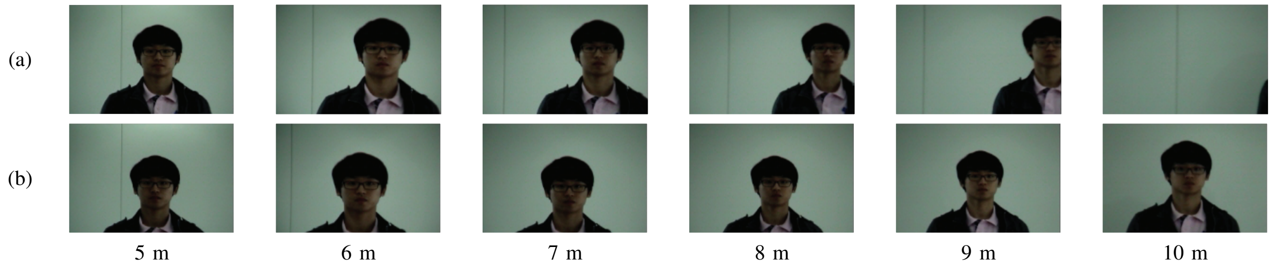


Fig. 4. Facial images at a distance of 5 to 10 m: PTZ camera was controlled using (a) a single static camera and (b) dual (vertical and horizontal) cameras in coaxial configuration after the calibration between $(x_s, y_s)$ and $(p, t)$ at a distance of 5 m.

distance between two eyes is at least 60 pixels and their corresponding silhouette heights from a set of training data. A quadratic mapping function between the height $(h)$ of silhouette and zoom values $(z)$ of the PTZ camera is obtained from

$$z = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \begin{bmatrix} h^2 & h & 1 \end{bmatrix}^T. \qquad (2)$$

### D. System Configuration

The system consists of three cameras: two Sony EVI-HD1 cameras are used as static cameras to obtain the vertical and horizontal global views and one Sony EVI-D100 camera is used as a PTZ camera to track and acquire high resolution face image at a distance. The image resolutions are 720×360 and 720×486 for the static and PTZ views, respectively. All image acquisition and processing modules are based on OpenCV [15]. The PTZ camera is controlled using the standard RS-232 serial port. The tracking and camera control components run in real time (8 fps) on a quad core Intel 2.8 GHz machine. Fig. 1 shows the proposed camera system. The static camera on top of the PTZ camera is configured upside down to minimize $d_y$. However, this is not a required set up in the coaxial camera system.

The system is decomposed into static camera processing and dynamic camera control modules (Fig. 6). The former includes image capture, background subtraction, and object and head detection using vertical and horizontal static cameras and coaxial face location estimation. The latter performs face location prediction and camera control (i.e., pan-tilt and speed control). The static processing module sends target locations of face in each frame to dynamic camera control module. The PTZ camera control module adjusts pan-tilt angles to observe the target in the field of view.

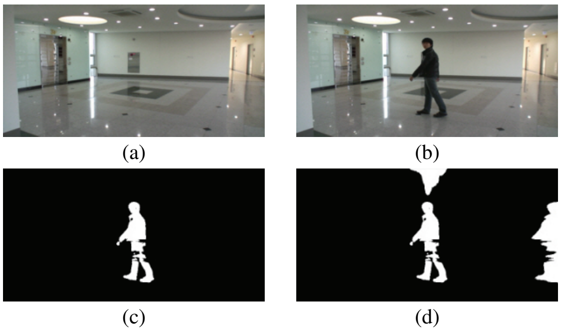


Fig. 5. Head localization: (a) background image, (b) input image, (c) silhouette, and (d) localized head region.

## III. Camera Control for Smooth Tracking

There are two components in this module: the pan and tilt parameter controller (PTC) and the motion velocity controller (MVC). The PTC predicts the next head location given the previous head trajectory. The estimated head location is converted to pan and tilt values. Given a set of pan and tilt values, the MVC controls the velocity of pan and tilt motion. While there have been a few previous studies on the static image processing part [16], no systematic study has been reported on the dynamic camera control part.

### A. Pan and Tilt Controller

The objective of the camera control is to keep the subject being tracked in the center of the PTZ view. By setting the head location to the center of PTZ view, the possibility of losing track of the face in the next frame is minimized.

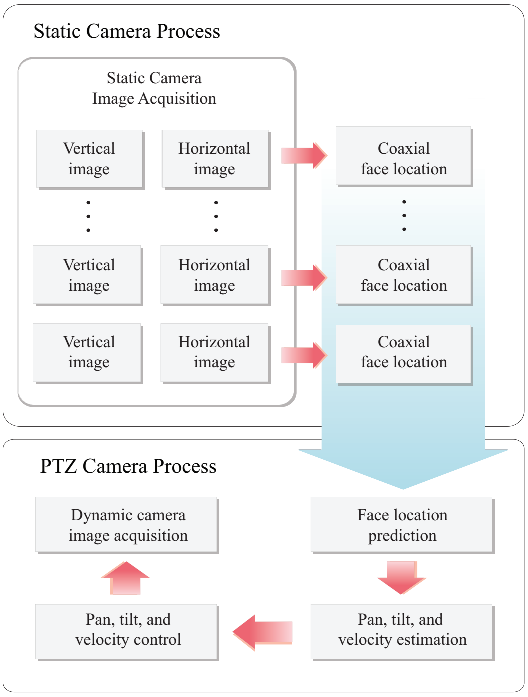


Fig. 6. Schematic of the proposed camera system.

Controlling the camera with current location of the head and its corresponding pan and tilt values does not provide robust tracking capability due to delays in image processing and mechanical camera motion.

To solve the above problem, we use a linear prediction model. Let $p_n$ and $t_n$ denote the $n^{th}$ pan and tilt control values of PTZ camera and $k_n$ be the time (ms) at the estimation of $p_n$ and $t_n$. Then, the next pan and tilt value $(p_\alpha, t_\alpha)$, after time $\alpha$ (ms) can be computed from a set of previously estimated values (=$C$) using the following two-step recursion of update

$$M_n = \begin{bmatrix} b_1 & b_2 \\ b_3 & b_4 \end{bmatrix} = D(n)K(n)^T(K(n)K(n)^T)^{-1} \tag{3}$$

$$where\ D(n) = \begin{bmatrix} p_n & p_{n-1} & \dots & p_{n-C-1} \\ t_n & t_{n-1} & \dots & t_{n-C-1} \end{bmatrix} \tag{4}$$

$$and\ K(n) = \begin{bmatrix} k_n & k_{n-1} & \dots & k_{n-C-1} \\ 1 & 1 & \dots & 1 \end{bmatrix}, \tag{5}$$

and prediction as

$$\begin{bmatrix} p_\alpha \\ t_\alpha \end{bmatrix} = M_n \begin{bmatrix} k_n + \alpha \\ 1 \end{bmatrix}. \tag{6}$$

### B. Motion Velocity Controller

The PTZ camera in our system provides 24 levels of pan speeds from 2 to 300 degrees/sec and 20 levels of tilt speeds from 2 to 125 degrees/sec. In typical PTZ camera systems, a fixed speed, usually a maximum velocity is used at each camera control command. However, the fixed speed strategy can cause discontinuous control of the camera, resulting in a higher probability of losing the subject or making the captured image blurred. In our system, the PTZ camera speeds are calculated based on the current and the next predicted head location. We control the PTZ camera at multiple stages for smoother motion.

## IV. APPLICATION TO FACE RECOGNITION

In order to verify the face recognition capability of the proposed system in surveillance applications, we conducted face recognition test at a distance of up to 10 meters. We compared the face identification accuracies using both the conventional static camera and the proposed camera system to show the effectiveness of the proposed system.

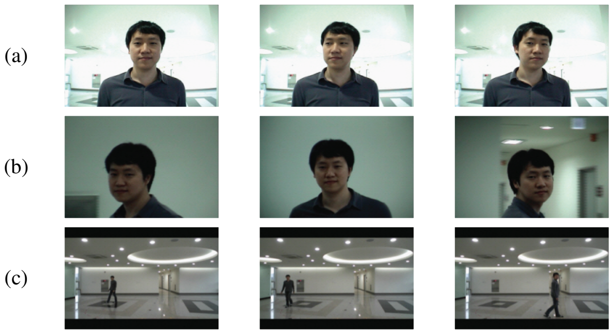


Fig. 7. Gallery and probe images: (a) front, left and right facial images for gallery and example probe images captured from (b) PTZ camera and (c) static camera.

### A. Experimental Data

We captured videos of 20 subjects at a distance of 5 to 10 m using both static (Fig. 7(c)) and PTZ cameras (Fig. 7(b)). All video data were collected indoors at Korea University campus from the university students. Each subject was asked to walk starting at about 10 m from the camera up to about 5 m distance by making an S-shaped path to evaluate the tracking capability of the proposed system in a surveillance scenario. The average duration of each video is about 25 seconds at 30 fps. The gallery data consists of three images per subject captured at about 1 m distance from the camera at three different poses (Fig. 7(a)). Additional 10,000 images of 10,000 subjects from the MORPH database [17] are also added to the gallery to increase the complexity of face identification by having a large gallery size. Even though the images in MORPH are different from the probe videos in terms of pose, overall face size, ethnicity, etc., it is the only large public domain face image database available.

### B. Results and Analysis

We performed a face recognition experiment using all the frames in the collected video data set as probe (17,140 images of 20 subjects) and 10,060 images of 10,020 subjects as gallery. A commercial face recognition engine, FaceVACS [18], was used for face detection and recognition. We rejected probe images with matching scores less than 0.31 (24% rejection, scores lower than this are all 0) and 0.45 (40% rejection) in the PTZ view. The range of matching scores provided by FaceVACS is [0,1]. The probe images from static views show complete failure of face recognition and the rejection scheme did not help in improving the identification accuracy. Table I shows the Rank-1 face identification accuracies obtained from the static and PTZ views. The threshold scores for rejection is represented by $t_r$. While the identification accuracy of the PTZ view is 48.8%, that of the static view is no better than random guess. Frame level fusions using score-sum method with contiguous 2, 5, and 10 frames after rejection scheme ($t_r$=0.45) shows further improvement in the identification accuracy. For example, in the fusion with 5 frames, the matching scores of the probe

TABLE I

FACE RECOGNITION ACCURACY OF CONVENTIONAL STATIC AND PROPOSED PTZ CAMERA SYSTEM

| Methods of identification | Rank-1 identification accuracy(%) |
|---|---|
| Static view<br>(conventional surveillance system) | 0.1 |
| PTZ view, 1 frame,<br>(coaxial camera system) | 48.8 |
| PTZ view, 1 frame, $t_r$=0.31 | 64.5 |
| PTZ view, 1 frame, $t_r$=0.45 | 78.4 |
| PTZ view, fusion of 2 frames, $t_r$=0.45 | 87.7 |
| PTZ view, fusion of 5 frames, $t_r$=0.45 | 93.9 |
| PTZ view, fusion of 10 frames, $t_r$=0.45 | 96.4 |

image at time $t$ to all gallery images are summed with those of probe images at time $t-1$, . . ., $t-4$. Then, the identity is decided based on the summed scores. Fig. 8 and Fig. 9 show example probe images that were not successfully matched and those successfully matched at rank-1. Major reasons of the failures are (i) inability to track a face, (ii) off-frontal facial pose, (iii) motion blur, and (iv) non-neutral facial expression.

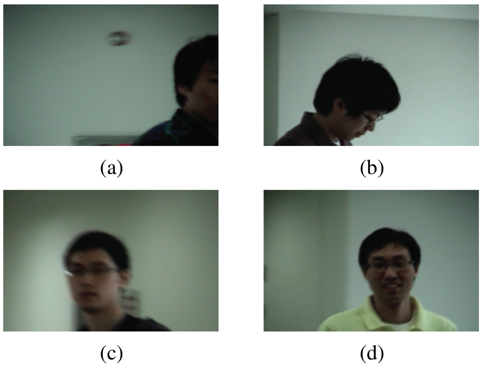
## V. CONCLUSIONS AND FUTURE WORK

We have proposed a novel coaxial camera system that can capture high resolution face images (with inter-pupillary distance of ∼60 pixels) at a distance of 5 to 10 meters for face recognition. The coaxial camera configuration and calibration method provide a large operating distance with improved face recognition accuracy. We also introduced a linear prediction model and a pan and tilt motion velocity control method for robust tracking. The face recognition test shows the effectiveness of the proposed system for subject identification at a distance.

Currently, the proposed camera system can track and identify only one person in the field of view and the operating distance is limited to ∼10 m. Also, the system can recognize a face only when it is close to frontal pose, which is the inherent limitation of the commercial face matcher used in this work as well as other available face matchers. We plan to extend the proposed system to identify multiple persons in the field of view by tracking multiple subjects in the static view and then identifying each subject sequentially using the PTZ camera. We also plan to extend the operating distance beyond 10 meters by using high definition static cameras.

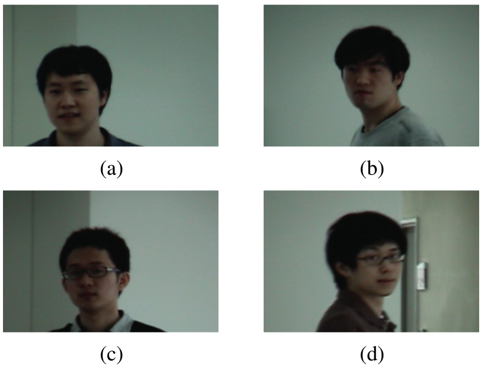
## VI. ACKNOWLEDGMENTS

Anil K. Jain's research was partially supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0).



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

Fig. 8. Example probe images that were not successfully matched at rank-1 due to (a) tracking failure, (b) off-frontal pose, (c) motion blur, and (d) non-neutral expression.



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

Fig. 9. Example probe images successfully matched at rank-1.

## REFERENCES

[1] A. Jain and S. Li, *Handbook of face recognition*. Springer, 2005.

[2] P. Phillips, W. Scruggs, A. O'Toole, P. Flynn, K. Bowyer, C. Schott, and M. Sharpe, "Frvt 2006 and ice 2006 large-scale results," *Technical Report NISTIR 7408, Natl Inst. of Standards and Technology*, Mar. 2007.

[3] G. Dedeoglu, T. Kanade, and J. August, "High-zoom video hallucination by exploiting spatio-temporal regularities," vol. 2. IEEE Computer Society, 2004, pp. 151–158.

[4] J. Park and S. Lee, "Stepwise reconstruction of high-resolution facial image based on interpolated morphable face model," in *Proc. Int'l Conf. Audio-and Video-based Biometric Person Authentication*, 2005, pp. 102–111.

[5] Y. Yao, B. Abidi, N. Kalka, N. Schmid, and M. Abidi, "Improving long range and high magnification face recognition: Database acquisition, evaluation, and enhancement," *Computer Vision and Image Understanding*, vol. 111, no. 2, pp. 111–125, 2008.

[6] R. S. K. Bernardin, F. v. d. Camp, "Automatic person detection and tracking using fuzzy controlled active cameras," in *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, 2007, pp. 1–8.

[7] L. Marchesotti, S. Piva, A. Turolla, D. Minetti, and C. Regazzoni, "Cooperative multisensor system for real-time face detection and tracking in uncontrolled conditions," in *SPIE Int'l Conf. Image and Video Communications and Processing*, 2005.

[8] A. Hampapur, S. Pankanti, A. Senior, Y.-L. Tian, L. Brown, and R. Bolle, "Face cataloger: multi-scale imaging for relating identity to location," in *Proc. IEEE Con. Advanced Video and Signal Based Surveillance*, 2003, pp. 13 – 20.

[9] S. Stillman, R. Tanawongsuwan, and I. Essa, "Tracking multiple people with multiple cameras," in *Proc. Int'l Conf. Audio-and Video-based Biometric Person Authentication*, 1999.

[10] M. Tistarelli, S. Li, and R. Chellappa, *Handbook of Remote Biometrics: for Surveillance and Security*. Springer, 2009.

[11] S. Prince, J. Elder, Y. Hou, M. Sizinstev, and E. Olevsky, "Towards face recognition at a distance," in *Proc. Crime and Security*, 2006, pp. 570–575.

[12] R. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Trans. Robotics and Automation*, vol. 3, no. 4, pp. 323–344, 1987.

[13] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. IEEE CS Int'l Conf. computer vision*. Published by the IEEE Computer Society, 1999, p. 666.

[14] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1337–1342, 2003.

[15] Open Computer Vision Library, http://www.intel.com/research/mrl/research/opencv/.

[16] R. Liu, X. Gao, R. C. X. Zhu, and S. Z. Li, "Tracking and recognition of multiple faces at distances," *Advances in Biometrics, LNCS*, vol. 4642, pp. 513–522, 2007.

[17] K. R. Jr. and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, 2006, pp. 341–345.

[18] FaceVACS Software Developer Kit, http://www.cognitec-systems.de/.