Stand-off Iris Recognition System

Frederick W. Wheeler, A. G. Amitha Perera, Gil Abramovich, Bing Yu, and Peter H. Tu

***Abstract*— The iris is a highly accurate biometric identifier. However widespread adoption is hindered by the difficulty of capturing high-quality iris images with minimal user cooperation. This paper describes a first-generation prototype iris identification system designed for stand-off cooperative access control. This system identifies individuals who stand in front of and face the system after 3.2 seconds on average. Subjects within a capture zone are imaged with a calibrated pair of wide-field-of-view surveillance cameras. A subject is located in three dimensions using face detection and triangulation. A zoomed near infrared iris camera on a pan-tilt platform is then targeted to the subject. The iris camera lens has its focal distance automatically adjusted based on the subject distance. Integrated with the iris camera on the pan-tilt platform is a near infrared illuminator that is composed of an array of directed LEDs. Video frames from the iris camera are processed to detect and segment the iris, generate a template and then identify the subject.**

I. INTRODUCTION

Rapid iris image collection with minimal subject cooperation or delay at biometric checkpoints will enable many new applications, and will increase adoption of iris recognition systems. Iris recognition can be much more accurate than other biometric modalities such as face, fingerprint, retina, and hand geometry [1], [2]. However, a major impediment to widespread use of the iris as a biometric identifier is the difficulty of iris image collection. Our motivation is to overcome this barrier by developing a system that captures iris images and performs identification with minimal cooperation on the part of the user. The key challenge is to quickly obtain a well-focused image of the iris.

Iris recognition systems [3] generally operate on iris images collected in the near infrared spectrum (typically in the 700–900 nm range) with 200 or more pixels across the diameter of the iris. Typically the iris is illuminated with a controlled near infrared light source and a matched near infrared passband filter is mounted on the camera lens to reduce the effect of ambient light. At these wavelengths, all individuals exhibit unique and identifiable iris patterns. However, it is challenging to capture a suitable image of the iris with these parameters.

Iris images are processed by first detecting and segmenting the iris. Almost all iris recognition systems use a Daugman-style template for matching [4]. The iris region of the image is warped (unrolled) to a normalized rectangle and a predefined set of complex Gabor wavelets is applied. The resulting Gabor coefficients are quantized in phase to produce the template. Recognition is then achieved by matching the template against a gallery of templates using normalized Hamming distance.

We have designed and built a prototype stand-off iris recognition system (Fig. 1). The system is designed for access control at moderate ranges (up to 1.5 m). We have designed the system for cooperative users of any height who approach and pause in front of the device and look toward the iris camera. We anticipate applications to gain access through locked doors.

The system uses a wide-baseline stereo pair of fixed wide-field-of-view (WFOV) surveillance cameras to locate the subject. Face detection is performed in each camera view, allowing for the detection of the presence of a subject and to determine the location of their face in 3D. This allows us to control a pan-tilt platform to direct a strobed near infrared illuminator and an iris camera toward the subject. The focal distance of the iris camera lens is also controlled based on the distance to the subject. In the iris camera video frames we perform iris detection and iris segmentation. In the prototype, identification against a gallery is achieved with the Masek implementation [5] of the Daugman algorithm [4].

This paper provides an overview of the system design and performance. Below we briefly describe several other related iris capture systems. In later sections we will provide more detail on the motivation and design of major components of our system.

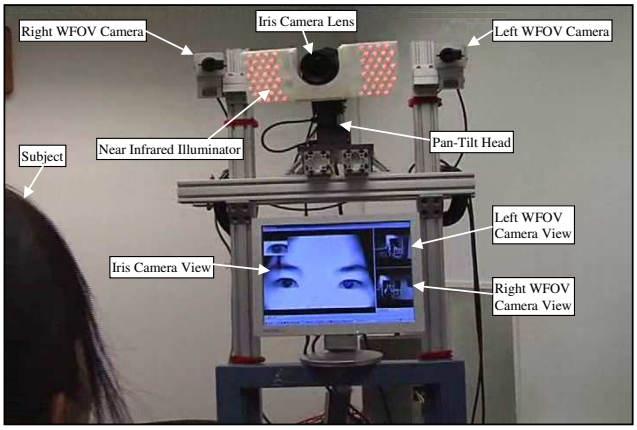

Fig. 1. Primary physical components of the iris recognition system in operation. The monitor shows system activities for development and demonstration purposes, and would not be part of a final system.

Fred Wheeler (wheeler@research.ge.com), Gil Abramovich (abramovi@research.ge.com), Bing Yu (yub@research.ge.com), and Peter Tu (tu@research.ge.com) are with GE Global Research, Niskayuna, NY, USA. A. G. Amitha Perera (amitha.perera@kitware.com) is with Kitware Inc., Clifton Park, NY, USA.

Related Systems

There are a variety of commercially available iris capture devices that use a lens with fixed focal distance and that require the user to position themselves so that their eye is well focused, often with audible directions from the device. This can be a significant burden for the user. Several organizations have developed more sophisticated collection devices to overcome this requirement.

Sarnoff and Sensar have developed the IrisIdent™ iris capture system [6], [7]. This system also uses a stereo pair of WFOV cameras for 3D localization of the subject and pan-tilt and focus control for the iris camera. Though our system and the IrisIdent system have many high-level similarities, there are several important differences. The IrisIdent system finds the head using cross-correlation-based stereo and then one of the eyes of the subject using the WFOV cameras, while we take the approach of face detection in each WFOV camera view. Instead of a pan-tilt mirror we place the iris camera directly on a pan-tilt head. While the IrisIdent system uses a pair of stationary illumination sources, we have designed and constructed an LED-based illuminator that is mounted on the pan-tilt head with the iris camera. The IrisIdent system detects eyes based on specular reflections of the illumination source on the cornea. Our eye detection approach specifically looks for the characteristic shape of the pupil and iris.

The Iris on the Move™ system by Sarnoff Corp. is a high-throughput iris recognition system that identifies subjects as they walk through a portal [8]. The system includes several iris video cameras with fixed focal distances. These cameras are arranged in a stack so that as a subject walks through the portal, their irises will be imaged by at least one of the cameras and will be in focus for at least one of the frames. This system can identify up to 20 walking subjects per minute (3 seconds per subject), but requires several expensive iris cameras and lenses. The most important difference between our system and the Iris on the Move system is the application scenario. Iris on the Move is designed for walking users. Our system requires subjects to stop momentarily in order to be identified, such as at a locked door that they wish to enter.

Guo et al. [9] have described an iris capture system that uses a pan-tilt iris camera directed by face detection in a wide-field-of-view camera on the same pan-tilt platform. The distinguishing feature of our system is the use of a stereo pair of WFOV cameras with face detection so we can determine both the direction to the face and the distance. The wide-baseline stereo camera approach has the potential to more accurately determine subject distance, which is critical in active near infrared illumination systems. To allow low radiation levels on the eye we want a large aperture, which results in a small depth of focus and a stringent focal distance accuracy requirement.

Yoon et al. have described an iris capture system with similar cooperative access control operating goals [10]. This system also uses a pan-tilt iris camera. To find the position of the user's face, the system determines the standing position of a user using a light stripe laser and WFOV camera. This establishes the pan angle for the iris camera. Because the height of the subject is unknown, the system then scans the tilt angle of the iris camera until a face is detected through the iris camera. Our approach of using a stereo pair of WFOV cameras and face detection to localize the face has the advantage of eliminating mechanical scanning for the face position. Instead, after face detection, the iris camera is pointed directly at the subject's face.

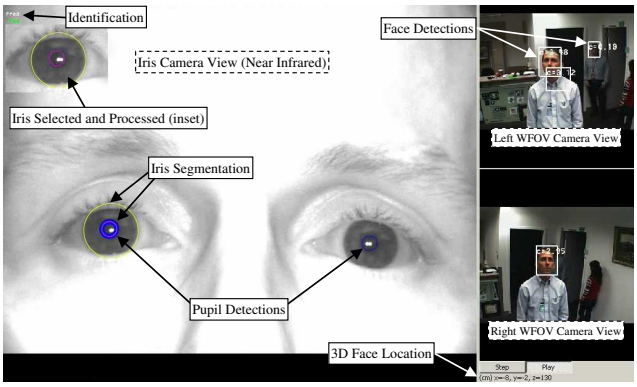

Fig. 2. The development application for the system shows video from each of the cameras along with face detections, 3D face position, pupil detections, iris segmentation, and identification results.

II. SYSTEM

Fig. 1 shows the primary visible components of the system. The two WFOV cameras and the iris camera are arranged horizontally and the illuminator is long and narrow to facilitate installation above a door. This baseline positioning of the cameras can easily be changed for other situations. Fig. 2 is a screen-shot of the development and demonstration application that shows the annotated video feeds. Fig. 3 is a block diagram of the main hardware and processing components of the system. A face detector is applied to video frames from a calibrated pair of WFOV cameras and the subject's face is located in 3D through triangulation. The iris camera and strobed illuminator are then directed to point at the subject's eyes and the iris camera is focused. Frames from the iris camera are processed to detect, segment and match irises against a gallery. In the following sections we will describe major system components in further detail.

A. Wide-Field-of-View Cameras

The WFOV stereo cameras are ordinary 30 Hz 640 by 480 pixel color surveillance cameras. The cameras are positioned horizontally about 0.5 m apart. Their lenses are selected and adjusted for close-range focus and each camera views the entire capture zone. It would be a natural extension to improve the resolution of one or both of these cameras so that face recognition could be performed at the same time as iris recognition. The cameras are connected to a frame grabber on the system computer.

The two WFOV cameras are geometrically calibrated using the procedure described by Zhang [11]. Each WFOV camera simultaneously collects video of a checkerboard

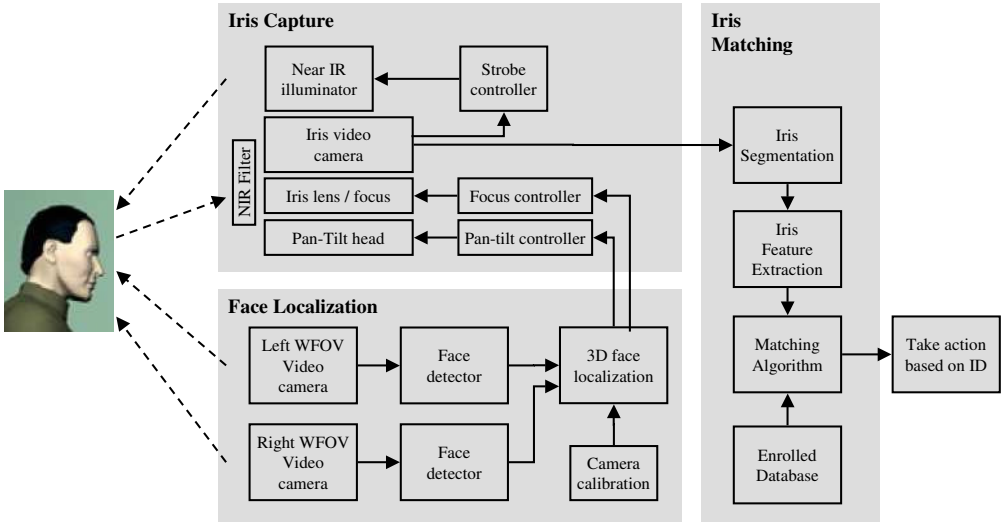

Fig. 3. Primary components of the system.

calibration pattern (Fig. 4). The calibration pattern is held in the capture zone of the system, in view of both WFOV cameras, and is slowly rotated in multiple directions. The calibration procedure extracts corner locations on the calibration pattern, and determines the internal parameters of the individual cameras and also their relative positions. The calibration of these two cameras establishes a metric world-based coordinate system [12]. The WFOV cameras must be calibrated whenever they are moved or the zoom setting of a lens is changed. In production, calibration would be a one-time operation.

In all subsequent processes, the coordinates of all 3D points are specified in this world-based coordinate system and are found through stereo triangulation. The calibration procedure yields a camera projection matrix for each WFOV camera that maps 3D points to their image location. Using the pseudo-inverse of this matrix, one can find the 3D line containing all points that map to a given image location. When an object is located in each WFOV image, the triangulation process finds the 3D point that is closest to the two associated 3D lines [12].

B. Face Detection and Localization

The system uses the WFOV cameras to localize the face of the user in 3D. Pittsburgh Pattern Recognition's Software Development Kit (FT-SDK) performs face detection [13], [14] on each view independently (Fig. 2). As a person enters the capture zone, their face is detected in each view, and is then localized in 3D.

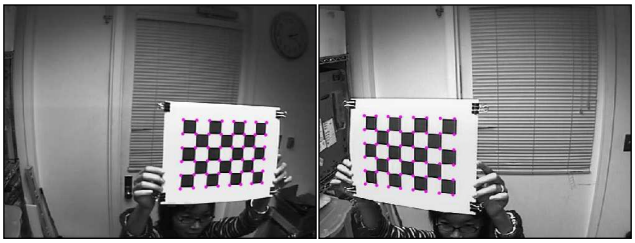

Fig. 4. Left and right WFOV camera views of the calibration target showing detected corners (magenta dots).

The face detector represents detections as a rectangular box in the image. Since the midpoint between the eyes is approximately centered on the face, we use the center of a detection box as the location of the eyes. After triangulation we have the approximate location of the point between the eyes in 3D.

Multiple face detections may occur in each WFOV video frame (left and right), typically due to the presence of multiple individuals. Since the detection in the views is independent, each pairing of faces corresponds to a 3D location and the system must determine which pair corresponds to the user's location. The 3D location that is closest to the center of the capture zone is selected as the target 3D location.

C. Iris Camera

The iris camera is the monochrome Cooke PixelFly QE, which is equipped with a 2/3"-format Sony ICX 285AL CCD sensor, and has boosted sensitivity in the NIR spectrum. High NIR sensitivity is desirable to reduce the required level of illumination. The sensor resolution is 1394 by 1024 pixels, with a frame rate of 12 Hz. The sensor elements of this camera are relatively large at 6.45 by 6.45 mm, limiting the adverse effects of diffraction and thus allowing for high image contrast.

The iris camera lens is the potentiometer-equipped Fujinon D16x73A-Y41. This lens provides programmable motorized focus and zoom (focal length 10–160 mm). A study of the trade-off between diffraction effects and depth of field resulted in an aperture setting of f/9.5, where the system is (just) diffraction limited. The depth of field is about 30 mm, making clear the need for accurate focal distance control. The lens is equipped with a 715 nm long-pass filter that blocks most of the ambient light and allows most of the LED

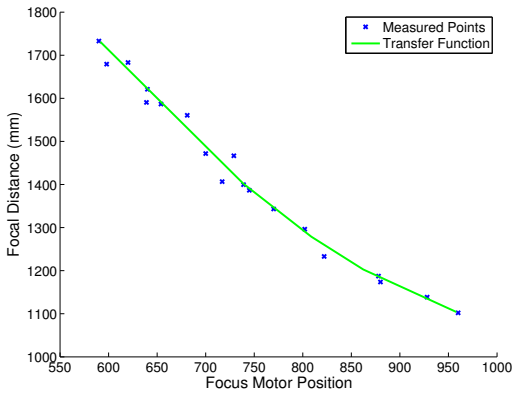

Fig. 5. Measured focal distance vs. focus motor position.

NIR illumination. The iris camera and lens are configured to capture iris images at a resolution of approximately 200 pixels across the iris, with a field of view of 76 mm at a subject distance of 1.5 m.

D. Near Infrared Illumination

The system illuminator uses 64 LEDs, each with an 810 nm center wavelength and a lens providing $\pm 6^\circ$ viewing half-angle (Epitex L810-06AU). Each LED provides a nominal radiant intensity of 110 mW/sr at 50 mA. The LEDs are mounted on a custom plastic mount that holds them such that they are all on a small region of the surface of a sphere with a 1.5 m radius and point toward the center of that sphere. The spherical mounting surface provides sufficient illumination intensity at 1.5 m while the distribution of LEDs is designed to prevent unsafe radiation at closer range. The LED mount is attached to the pan-tilt head so it and the iris camera always point in the same direction (seen in Fig. 1).

The LED circuit arrangement uses eight branches of eight serially connected LEDs. This arrangement is dictated by the voltage drop through each LED (1.5–2.5 V, depending on the current draw and the strobe mode) and the strobe controller characteristics. The strobe controller is triggered at the frame rate by the camera controller and has an illumination duration of 20 ms, matched to the iris camera exposure period to minimize motion blur and the effect of ambient light. The average irradiance at the desired maximum range of 1.5 m is 650 μW/cm^2. The illumination system was thoroughly analyzed to operate well within eye safety standards and recommendations [15], [16].

E. Pan-Tilt Calibration and Control

After a face is detected and its 3D coordinates are determined, the pan-tilt head is controlled to direct the iris camera toward the subject's face. The Directed Perception PTU-46-70 pan-tilt head used on the prototype has an interface that allows for setting the pan and tilt angles relative to its home position. To control the pan-tilt head we need to know the position of the iris camera in the world coordinate system, established by the WFOV camera calibration process, and the direction in which it will point given particular pan and tilt settings. As long as the pan-tilt unit is not physically moved this is a one-time procedure.

The calibration procedure for the pan-tilt head involves collecting three pairs of 3D world coordinate points. First, the pan-tilt head is placed in its home position and two points along the iris camera boresight (center pixel) are determined, one close to the camera (about 0.5 m), and another further away (about 1.5 m). Second, the pan-tilt head is panned approximately 10 degrees, and again, two boresight points are measured. Finally, the pan-tilt head is returned to its home position, and then tilted by 10 degrees, and two boresight points again are measured.

Each of the six calibration points is determined by placing a calibration target in front of the iris camera and carefully positioning it so that it is along the boresight of the iris camera. Then, the row/column position of the calibration target is measured in each of the WFOV cameras. Since the WFOV cameras are calibrated, this establishes the 3D location of the calibration target.

The six calibration points enable the specification of a coordinate system for the iris camera pan-tilt head within the world coordinate system. Each pair of points defines a line in the 3D world coordinate system. The intersection of these lines is the origin of the pan-tilt coordinate system. Due to measurement noise, the intersection is, of course, approximated.

A straightforward orthogonalization procedure allows us to determine the axes of the pan-tilt coordinate system. The result yields the pan-tilt coordinate system axes, as unit vectors in the world coordinate system. By projecting a world coordinate face detection location into the pan-tilt coordinate system we can readily determine the pan and tilt angles to point the iris camera boresight at the subject.

F. Focus Control

The iris camera focal distance is controlled by setting the lens focus motor position through a motor controller. The relationship between the motor position and focal distance is determined using another manual calibration procedure.

We collect a set of measured distances and the motor positions that focus the iris camera at that distance. A calibration target with fine lines is placed in several positions over the operating range of distances in front of the

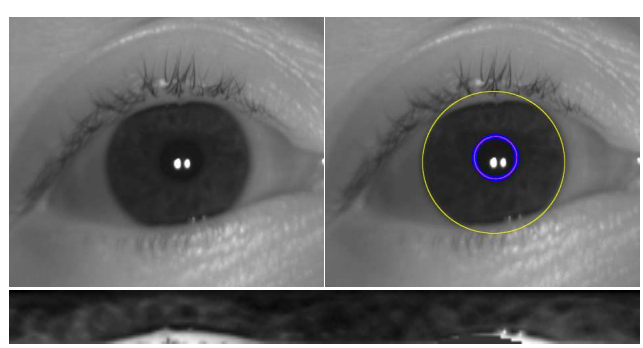

Fig. 6. An iris image from the prototype (left), also with the segmentation results (right) and unwrapped (bottom).

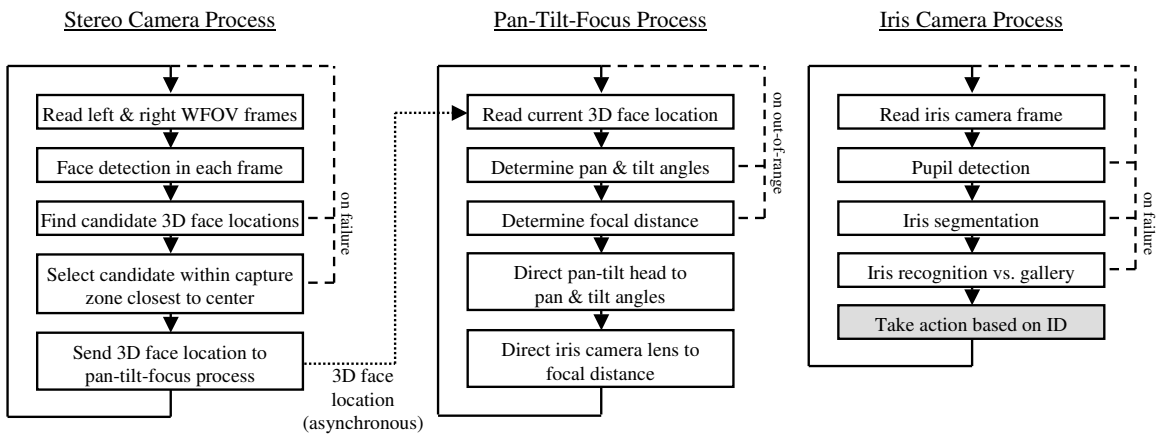

Fig. 7. The three primary processes (threads) of the system. The stereo camera process communicates 3D face locations to the pan-tilt-focus process, but the iris camera process is logically independent.

system. First, at each position, the location of the calibration target is manually selected in the views of each of the WFOV cameras. Through triangulation we then determine the distance from the iris camera to the target. Second, we manually adjust the focus motor until the calibration target is well focused by the iris camera. This establishes the motor position required to focus at the measured distance.

Fig. 5 shows these measured data points. Also shown is a smooth transfer function fitted to these points. This transfer function is used to map desired focal distance to motor position during system operation.

Note that the iris camera zoom is not adjusted. The lens we are using supports electronically adjusted zoom. However, we anticipate future systems using less expensive lenses with fixed zoom. The zoom setting is set so that iris images have about 200 pixels across the diameter when the subject is at the maximum distance of 1.5 m away. Iris resolution increases when the subject is closer.

G. Pupil/Eye Detection

Once a subject has been targeted, the iris camera views one or both eyes and the surrounding region of the face, as seen in Fig. 2. We apply a pupil detector in each frame to localize an eye in order to crop out a smaller region of the image to pass to the iris segmentation and recognition component, and to initialize that procedure.

The pupil detector relies on the fact that the pupil is round and is typically much darker than its surroundings. Straightforward image magnitude thresholding, morphological closing with a small disk and connected component analysis then produces a number of pupil candidates. The characteristic specular reflection of the illuminator is not explicitly used by this method of pupil detection. Candidates that have a radius within an acceptable range and that have a radius to area ratio that is consistent with being circular are declared to be pupils. Figs. 2 and 6 show pupil detections.

H. Iris Recognition

Iris segmentation and recognition is performed using the Masek implementation [5], [17] of the Daugman algorithm [4]. A Daugman-style iris-matching algorithm matches iris images by computing templates for each iris image. Identification is performed by computing the distance between the template for a probe iris image and the templates stored in a database, or gallery. Templates are computed in several steps. First, the iris is detected and segmented in the portion of the image around the detected pupil. Segmentation finds the inner and outer boundaries of the iris and may also find a mask indicating which portions of the image are obscured by eyelids, lashes, or specular reflections. Then, the iris portion of the image is unwrapped by warping with a modified polar to rectangular coordinate system transformation that accounts for the pupil diameter. Fig. 6 shows an iris image, segmentation and the unwrapped iris image.

Next, a predefined set of complex Gabor wavelets is applied to the now rectangular iris image. The complex Gabor coefficients are quantized in phase to produce two template bits per coefficient. The collection of these bits forms the template. These templates are matched very quickly by simply computing a normalized Hamming distance between probe and gallery templates.

I. System Computer

The prototype system in implemented on a single general purpose PC with a 3.20 GHz single-core CPU. There are three primary independent and asynchronous threads of control (Fig. 7). One thread is responsible for continuously detecting faces, determining the 3D face location, and the second for pointing the pan-tilt head toward the face and setting the focal distance. Independently, a third thread controls iris recognition. This thread continuously reads frames from the iris camera, detects pupils, crops eye regions, and performs iris segmentation and recognition. Not all iris camera frames are processed. Currently, the prototype can process about 5 iris camera frames per second, and 15 WFOV frames per second.

III. OPERATION AND PERFORMANCE

Users are enrolled and recognized by the same stand-off system. In enrollment mode, the system repeatedly collects

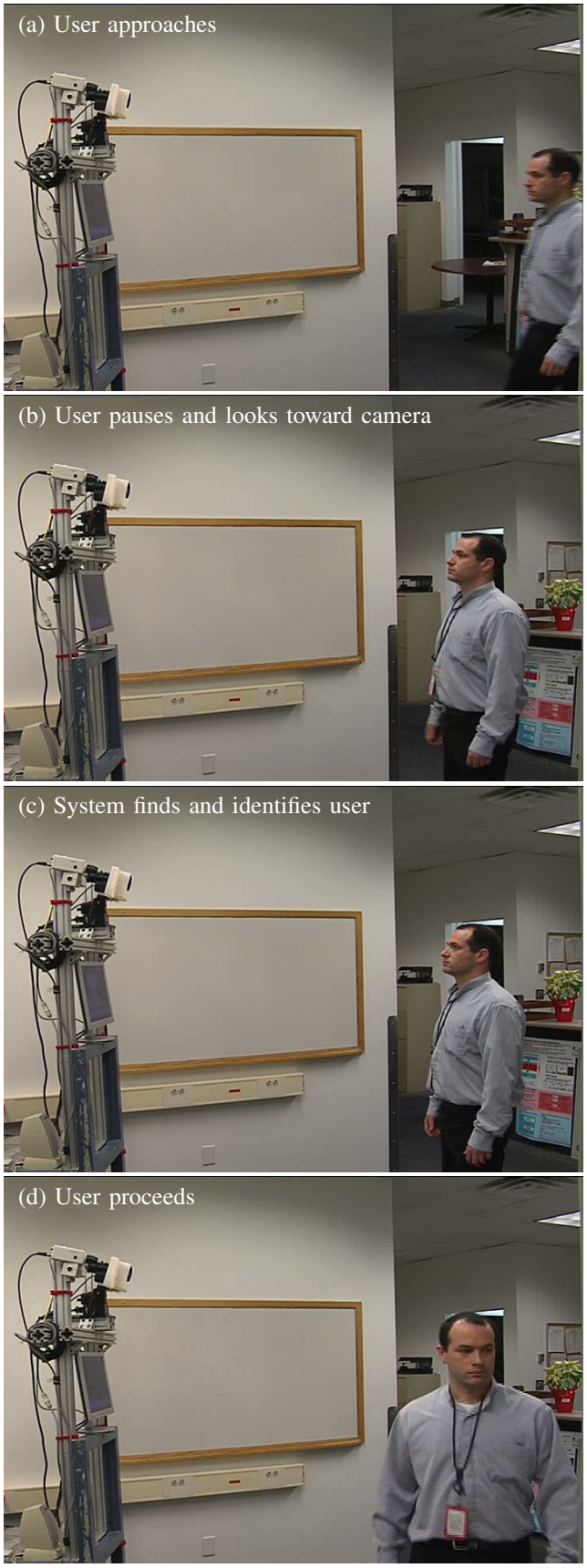

Fig. 8. Typical walk-up operation of the system. The time between pausing (b) and identification (c) currently averages 3.2 sec.

iris images and templates and maintains a queue of the 10 most recent templates. When two templates in the queue match with a normalized Hamming distance (NHD) less than 0.30, one of the templates from the best matching pair is placed in the gallery. In practice we have enrolled both irises and declare a recognition if either is recognized.

Instead of capturing one, or a fixed number of iris images and attempting recognition, our approach for evaluating this prototype is to continuously capture iris images and attempt recognition until it succeeds. We use an NHD recognition threshold of 0.30, which corresponds to a *verification* false recognition rate of $1 : 10^6$. A more refined design will include safeguards against attempts to achieve a false recognition by simply standing at the system indefinitely. However, at present we are more concerned with how fast we can achieve true recognition.

To test the throughput of the system we line up test subjects who approach the system one at a time. Each subject walks into the capture zone, and stands still while looking toward the iris camera. The system captures the iris image and identifies either iris of the subject. Once the person is identified by the system he moves along and the next person approaches. In this mode the current prototype identifies a person about every 7.1 seconds. The average time between a subject pausing and facing the iris camera and that subject being identified by the system is **3.2 seconds**. We emphasize that these throughput and identification times are from small-scale tests of our current *prototype* system. We expect that further development will improve performance.

IV. FUTURE WORK

A critical performance parameter is the time between a subject pausing at the iris identification system and the subject being identified. From timers built into the system software we know this time is currently dominated by two factors: the time to move the pan-tilt head, and the time to locate and segment the iris. We are encouraged that each of these factors has a lot of potential for improvement. The time to point the iris camera can be improved with faster hardware, or by a strategy of tracking people as they approach the system, so the iris camera can be already pointed close to their position when they pause and look at the camera. Further, we expect iris location and segmentation time to be significantly faster when using commercial iris recognition systems, relative to the Masek implementation used presently [5].

V. CONCLUSIONS

In this paper we have presented an overview of an iris capture and recognition system that uses a stereo pair of cameras and face detection for 3D face localization, and a pan-tilt head to direct an illuminator and iris camera at the subject. Iris recognition has a great deal of potential in many identification applications. Currently the primary impediment to the use of iris recognition is that capture devices are difficult to use or expensive. We believe that capture systems

that work with minimal user requirements or familiarity are the key to widespread adoption of iris recognition.

REFERENCES

[1] J. Daugman, "Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons," *Proceedings of the IEEE*, vol. 94, pp. 1927–1935, November 2006.

[2] R. Wildes, "Iris recognition: An emerging biometric technology," *Proceedings of the IEEE*, vol. 85, pp. 1348–1363, September 1997.

[3] K. Bowyer, K. Hollingsworth, and P. Flynn, "Image understanding for iris biometrics: a survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008.

[4] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.

[5] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, The School of Computer Science and Software Engineering, The University of Western Australia, 2003.

[6] K. Hanna, R. Mandelbaum, D. Mishra, V. Paragano, and L. Wixson, "A system for non-intrusive human iris acquisition and identification," in *IAPR Workshop on Machine Vision Applications*, November 1996.

[7] M. DellaVecchia, T. Chmielewski, T. Camus, M. Salganicoff, and M. Negin, "Methodology and apparatus for using the human iris as a robust biometric," in *SPIE Proceedings on Ophthalmic Technologies VIII*, vol. 3246, (San Jose, CA), January 1998.

[8] J. R. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. J. LoIacono, S. Mangru, M. Tinker, T. M. Zappia, and W. Y. Zhao, "Iris on the move: Acquisition of images for iris recognition in less constrained environments," *Proceedings of the IEEE*, vol. 94, pp. 1936–1947, November 2006.

[9] G. Guo, M. Jones, and P. Beardsley, "A system for automatic iris capturing," Technical Report TR2005-044, Mitsubishi Electric Research Laboratories (MERL), December 2005.

[10] S. Yoon, H. G. Jung, J. K. Suhr, and J. Kim, "Non-intrusive iris image capturing system using light stripe projection and pan-tilt-zoom camera," in *Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2007.

[11] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, pp. 666–673, 1999.

[12] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[13] H. Schneiderman, "Learning a restricted Bayesian network for object detection," in *Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2004.

[14] H. Schneiderman, "Feature-centric evaluation for efficient cascaded object detection," in *Proc. of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2004.

[15] Illuminated Engineering Society of North America, *Recommended Practice for Photobiological Safety for Lamps and Lamp Systems - General Requirements, ANSI/IESNA RP-27.1-05*, 2005.

[16] CIE Central Bureau, Vienna, *International Commission on Illumination, Photobiological Safety of Lamps and Lamp Systems, CIE S 009/E:2002*, 2002.

[17] L. Masek and P. Kovesi, "MATLAB source code for a biometric identification system based on iris patterns," 2003. The School of Computer Science and Software Engineering, The University of Western Australia.