MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

A System for Non-intrusive Human Iris Acquisition and Identification *

K. Hanna, R. Mandelbaum, D. Mishra, V. Paragano, L. Wixson
David Sarnoff Research Center & Sensar Inc.
{khanna, rmandelbaum, lwixson, dmishra, vparagano} @sarnoff.com

Abstract

An automated system for the non-intrusive acquisition of images of human irises for the purpose of identity verification is described. This system uses active machine vision techniques, and does not require the user to make any physical contact with the system, or to assume any particular pose except that he stand with his head within a designated calibrated volume.

The system integrates a broad range of vision and control algorithms, and hence represents a major achievement in system integration. Currently, the system has been developed to the level of a robust prototype. It has been tested on over 600 people, and the results are impressive. A proposed use of the system is for identity verification at automatic teller machines to reduce the frequency of fraud.

In this paper, the overall control structure of the system is delineated, and each of the algorithmic modules is described. Preliminary test results are reported and discussed.

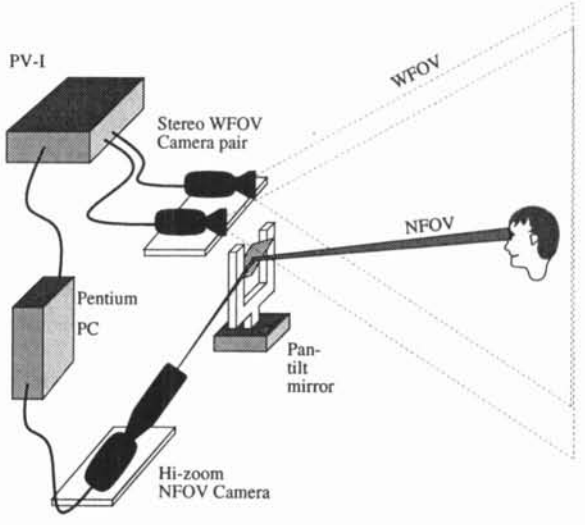

Figure 1: Schematic of the components of the system for the acquisition of high-resolution images of human irises for identity verification.

1 Introduction

We describe an application of machine vision for the non-intrusive active *acquisition* and *identification* of human irises. Daugman [4] has shown that the pattern of every individual's iris is highly distinctive, and may therefore be used as a distinguishing biometric. The IRISCAN™ system that uses this algorithm is currently in use for identification of authorized personnel in restricted areas. It is a passive system, performing no image acquisition on its own. The user is required to place his eye at a certain position relative to an imaging camera, and self-position until the image is focussed. Another system for iris-identification requiring similar levels of user cooperation is described in [7].

In this paper, we describe an *active* system for image acquisition using machine vision techniques. The user simply stands in front of the system, an image of their iris is acquired, and their identity is verified or refuted. This process currently takes approximately 5 seconds to perform, including the time for the user to comply with a request to look up. The system consists of a stereo pair of wide field-of-view (WFOV) cameras, a narrow field-of-view (NFOV) camera, a pan-tilt mirror allowing the NFOV to be moved relative to the WFOV, a PV-I™ real-time vision computer, and a PENTIUM™ front-end computer (see Figure 1). The system actively finds the position of the user's eye and acquires a high-resolution image to be passed to Daugman's system. No particular pose is required of the user except that he stand such that his head is within the WFOV and within a volume for which the system has been calibrated.

Figure 2 delineates the algorithm control structure and dataflow of the system. In particular, the system performs the following steps:

1. **Head and depth finding:** The presence of a person within the WFOV camera image is detected and the nearest person in the scene is distinguished from background users. A region of interest (ROI) containing the head of the person within the WFOV image is located, and the distance of the person's face from the WFOV camera is estimated.

*This research is funded by **Sensar, Inc.**, 121 Whittendale Drive, Moorestown, New Jersey 08057. The intellectual content of in this paper is the property of **David Sarnoff Research Center**, Princeton, NJ, **Sensar, Inc.**, Moorestown, NJ, or **IriScan, Inc.**, and is protected by several existing and pending patents.

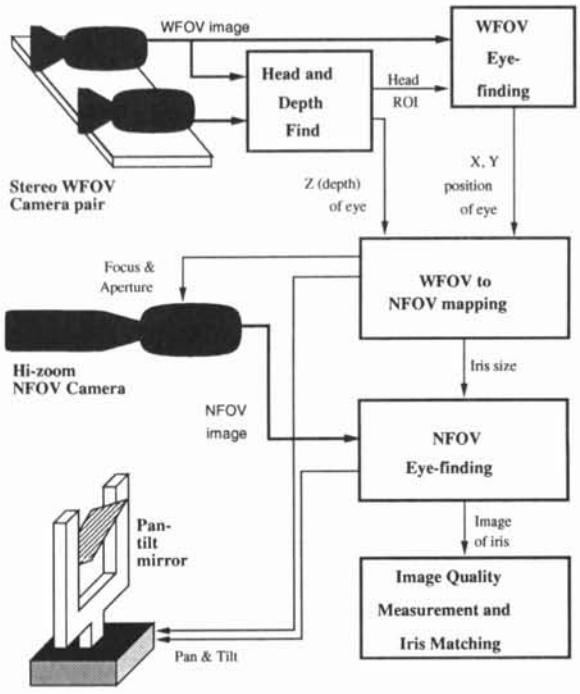

Figure 2: Algorithm control structure and dataflow diagram of the iris identification system, showing the five major components.

2. **WFOV eye-finding:** Template methods are employed within the ROI described in step 1 to locate precisely the right eye of the person's face in WFOV image coordinates.

3. **WFOV to NFOV mapping:** The depth and WFOV image coordinates of the eye are used to estimate the pan, tilt and focus parameters of the mirror and the NFOV camera necessary to align the NFOV camera with the position of the eye.

4. **NFOV eye-finding:** Real-time closed-loop tracking of the eye in the NFOV (with a FOV of only about 3 degrees) is performed, in order both to center the eye and reduce motion-blur.

5. **Image quality measurement and iris matching:** The high-resolution (about 200x200 pixels) image of the iris is checked for image quality. Daugman's algorithm is then invoked to validate or refute the identity of the user.

Currently, the allowable volume accommodates anyone between 4 feet 8 inches (1.43 meters) and 6 feet 3 inches (1.90 meters) in height, standing between 1 foot (30 centimeters) and 3 feet (1 meter) from the rig and laterally displaced by up to about 1 foot 6 inches (0.5 meters) to either side of the centerline of the rig. This volume corresponds to the region in which most people's heads appear relative to an Automatic Teller Machine. A proposed use of the system is for identity verification at such money-vending machines to reduce the frequency of fraud.

We have tested the system on over 600 people of various races, heights, and eye-colors. Performance is very impressive, both in terms of speed of acquisition and discrimination power during verification. The system has been shown to be robust against facial hair, face-masks, eye-patches (on one eye), prescription glasses and even many types of sunglasses.

The image of the eye is currently acquired and verified with a mean time of 5.5 seconds, though significant speed-ups are envisioned in the near future. This rapidity of acquisition is achieved mainly through use of hierarchical methods such as those described in [1, 2] implemented on the PV-I™ real-time vision computer. This computer is capable of constructing Gaussian and Laplacian pyramids, correlating images for stereo and performing template-matching almost at frame-rate. The probability of erroneously verifying an eye has been analyzed in [4].

In section 2 we describe the five major modules. Section 3 delineates our test results, while section 4 describes areas of future research and development.

2 System modules

2.1 Head and depth finding

The head and depth finding module takes as input a pair of stereo WFOV images, an example of which is shown in Figure 3. The purpose of the module is to select the **nearest** user to the system (in the case of multiple users), to find the **position** of the user's head in the image, and to estimate the **depth** of the user's eye from the system.

The module implements a cross-correlation-based stereo algorithm to build a disparity map of the WFOV scene. Details of the pyramid implementation of this algorithm are described in [3]. The disparity map is then analyzed and the closest region of approximately the size and shape corresponding to that of a human head is extracted. The disparity corresponding to the user's face is then taken to be the mean disparity of this region. As is well known, the 3-dimensional depth of the face is proportional to the inverse of this disparity [6, 5]. The depth of the face is used in subsequent modules and is especially useful for attaining correct focus of the iris image. The inset in the upper right corner of Figure 3 shows an example, at reduced resolution, of the area selected by the head-finding stereo algorithm.

2.2 WFOV eye-finding

The goal of the WFOV eye finder is to locate the right eye of the subject within the ROI computed by the head-finding module described in section 2.1. The right eye was chosen arbitrarily. Eye finding is performed by ana-

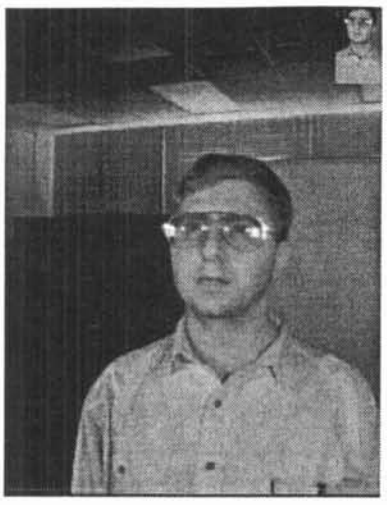

Figure 3: Typical output of the WFOV camera. The white spots on the glasses are specularities from the illumination. The inset in the upper right corner shows, at reduced resolution, the area selected by the head-finding stereo algorithm. Template methods are used on this ROI to find the eye.

lyzing and combining the results of three eye-find methods:

1. A **template**-based method which locates the face by searching for characteristic arrangements of features in the face. Two main features of the input face are used: (a) a bandpass-filtered version of the image, and (b) the orientation of particular features at a coarse resolution, for example the mouth. These features are compared to a face template comprising an a priori estimate of the expected spatial arrangement of these features. A face is detected when a set of tests using these features are successfully passed.

2. A **template**-based method which uses similar features as in (1) but locates the eyes by identifying the specularities that are visible off the surface of spectacles, if present.

3. A **specularity**-based method which locates reflections of the illuminators that are visible on the cornea.

The information from each eye-finding method is combined to determine whether an eye has been detected, and if so, its location in the image.

2.3 WFOV to NFOV mapping

A calibration look-up table (LUT) is used to map the information recovered from the WFOV processing onto the NFOV imaging parameters required to align the WFOV and NFOV images. In particular, the input to the LUT is the (x, y) image coordinates of the eye, and the z location of the head. The output includes the **pan** and **tilt** angles of the pan/tilt mirror, the **focus** of the NFOV camera, and the expected **diameter** of the iris. The calibration LUT accounts for factors that include the *baseline separations* between the WFOV cameras and the NFOV camera, *lens distortion* in the WFOV cameras, and *vergence* of the WFOV cameras.

Calibration (i.e. construction of the LUT) is performed only once at the manufacture of the system. An object of known size is placed at an (x, y) location in the image. The depth z in that region is computed by the stereo module. The pan/tilt mirror is manually slewed so that the object is centered in the NFOV image. The image is then focussed, and the diameter in pixels of the object is measured. Hence, for each point, the set of corresponding values $\{x, y, z, pan, tilt, focus, iris\ diameter\}$ is recorded. This is repeated for up to 25 points per depth plane, and at up to 4 depths inside the working volume of the system. Next, for each small set of neighboring points within the master list of up to 100 points, a vector of linear functions

$$(f_{\text{pan}}, f_{\text{tilt}}, f_{\text{focus}}, f_{\text{diam}}) : X \times Y \times Z \rightarrow P \times T \times F \times D$$

is fit to the data, where $x \in X$, $y \in Y$, $z \in Z$, $pan \in P$, $tilt \in T$, $focus \in F$, and $iris\ diameter \in D$. The result is a set of vectors of functions that define the mapping from $X \times Y \times Z$ to each of the described parameters throughout the volume. During operation, the LUT maps the user's (x, y, z) location to the linearly interpolated NFOV camera parameters $(f_{\text{pan}}, f_{\text{tilt}}, f_{\text{focus}}, f_{\text{diam}})\,(x, y, z)$.

2.4 NFOV eye-finding

Eye finding in the narrow field of view image is based on detecting a set of features visible in the image only if the eye is in the field of view. In the current system, NFOV eye-finding relies on two incandescent light sources, one on either side of the cameras. Light from these sources is reflected from the cornea of the eye, and appears as two bright spots on the iris. See Figure 4 for a typical view of these specularities. The specularities are used both to confirm the presence of the eye in the NFOV image and subsequently to determine the location of the eye.

The algorithm to center the iris consists of detecting the pair of bright spots. The separation of these specularities is estimated from the depth information obtained by the head-finding module of section 2.1. Since the specularities are approximately symmetrically located on either side of the eye center, their positions are used to estimate the coordinates of the center of the iris. Once the presence of the eye has been reliably determined, closed-loop NFOV tracking is invoked without using further WFOV information. However, in the event of large motion, the NFOV module may lose track and be forced to revert to the WFOV modules to re-acquire the eye.

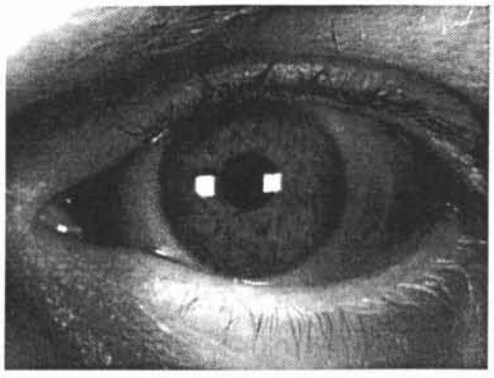

Figure 4: Typical output of the NFOV camera. The white spots are specularities from the illumination. The pattern of the iris is used for identity verification using Daugman's algorithm.

2.5 Image quality measurement and iris matching

After mechanical centering of the eye in the NFOV, electronic centering using the mean position of the detected specularities is performed to bring the eye precisely to the image center. This image is passed for matching to the algorithm developed by Daugman in [4]. This algorithm localizes the iris by locating the pupil/iris and iris/sclera boundaries, transforms the iris into a dimensionless polar coordinate system to account for iris dilation, and characterizes the iris texture by quantizing the phase of the output of a set of Gabor filters at different spatial scales, resulting in a 256 byte binary code.

3 Test results

The system is currently used in 2 modes : **enroll** and **verify**. In **enroll** mode, a single image of the eye is captured and used to produce a reference iris code. In **verify** mode, an image is captured and the resulting iris code is compared to the reference code to verify the identity of the user. The subject types in his/her account number and the system then requests the subject to look up. The system then verifies the user's identity, or declares either that the image acquisition was unsuccessful or that an impostor was detected.

The system has been tested on over 600 people to date. Results are shown in the following table:

	Number	%
Enrolled subjects	618	100.0
Successful verification in < 3 seconds	61	9.9
Successful verification in < 5 seconds	306	49.5
Successful verification in < 9 seconds	550	89.0
Successful verification (total)	611	98.9

3.1 Discussion

The time for the system to verify the user is measured from the instant the user is requested to look up to the time that the user is verified. Therefore, the timings include the time for the user to comply with the request, the time for the mirror and camera to respond mechanically, as well as processing time.

The mean time to verify a subject was 5.5 seconds. Reasons for non-acquisition include excessive height of the subject (placing the eye outside of the allowable volume), and spurious specularities detected near, but not on, the eye. In particular, specularities off the frames of spectacles occasionally caused errors.

4 Future research and development

The NFOV Eye find module currently detects the presence of an eye by locating specularities off the cornea. This can result in false eye detection, particularly from specularities off the frames of glasses. In future, this module will be supplemented by exploiting constraints that identify the eye more uniquely, such as shape constraints. The speed of operation will be increased by optimizing the movement of the mirror and the camera, as well as optimizing the algorithms. For example, the mirror can be pointed to the head region and the camera can be brought into approximate focus well before the eye is located precisely. This reduces the length of the trajectory that the mechanical components are required to traverse, hence reducing acquisition time.

References

[1] JR. Bergen, P. Anandan, K. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. *European Conference on Computer Vision*, pages 5–10, 1991.

[2] PJ. Burt and E. Adelson. The laplacian pyramid as a compact image code. *IEEE Trans Commun*, 31:532–540, 1983.

[3] PJ. Burt, L. Wixson, and G. Salgian. Electronically directed 'focal' stereo. *Proceedings of the Fifth International Conference on Computer Vision*, pages 94–101, June 1995.

[4] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(11):1148–1160, 1993.

[5] U. R. Dhond and J. K. Aggarwal. Structure from stereo — a review. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(16), November/December 1989.

[6] K. Hanna. Direct multi-resolution estimation of ego-motion and structure from motion. *Proceedings of the IEEE Workshop on Visual Motion*, pages 156–162, 1991.

[7] R. Wildes, J. Asmuth, G. Green, S. Hsu, R. Kolczynski, J. Matey, and S. McBride. A machine-vision system for iris recognition. *Machine Vision and Applications*, 9, 1996.