SAND96-1033
Printed April 1996
UNLIMITED RELEASE

Distribution
Category UC-900

Laboratory Evaluation of the IriScan Prototype Biometric Identifier

Frank Bouchier
Janet S. Ahrens
Grant Wells
Entry Control /Systems Engineering Department
Sandia National Laboratories
Albuquerque, NM 87185

ABSTRACT

One thing that all access control applications have in common is the need to identify those individuals authorized to gain access to an area. Traditionally, the identification is based on something that person possesses, such as a key or badge, or something they know, such as a PIN or password. Biometric identifiers make their decisions based on the physiological or behavioral characteristics of individuals. The potential of biometrics devices to positively identify individuals has made them attractive for use in access control and computer security applications. However, no systems perform perfectly, so it is important to understand what a biometric device's performance is under real world conditions before deciding to implement one in an access control system.

This paper will describe the evaluation of a prototype biometric identifier provided by IriScan Incorporated. This identifier was developed to recognize individual human beings based on the distinctive visual characteristics of the irises of their eyes. The main goal of the evaluation was to determine whether the system has potential as an access control device within the Department of Energy (DOE). The primary interest was an estimate of the accuracy of the system in terms of false accept and false reject rates. Data was also collected to estimate throughput time and user acceptability.

The performance of the system during the test will be discussed. Lessons learned during the test which may aid in further testing and simplify implementation of a production system will also be discussed.

The scope of this testing and evaluation was limited to performance characterization. Vulnerability analysis to identify inherent vulnerabilities of the technology was not performed. Thus, the potential exists that serious vulnerabilities may be identified in later evaluations.

CONTENTS

Figures

Purpose

The purpose of this report is to describe the evaluation of a prototype biometric identifier provided to Sandia by IriScan, Inc. This identifier was developed to recognize individual human beings based on the distinctive characteristics of the irises of their eyes. The IriScan system was tested during FY95 because it is a new technology which appears promising and may be of interest to Sandia and the Department of Energy (DOE) for use in entry control applications.

The main goal of the evaluation was to determine whether the system has potential as an access control device within DOE. The primary interest was an estimate of the accuracy of the system in terms of false accept and false reject rates. Data was also collected to estimate throughput time and user acceptability. Since the system evaluated was a prototype, it is reasonable to expect that the performance of the production units may change somewhat. Other parameters of interest such as size, tamper resistance, form factor, and interface requirements were not examined during this evaluation. These other parameters are likely to change a great deal when the prototype technology is transferred to production units and must, therefore, be included in a full evaluation of a production unit.

In order to perform this evaluation, a test involving human volunteers was performed. This test was approved by Sandia's Human Studies Board under proposal number SNL9502. The loan of the IriScan equipment was subject to a nondisclosure agreement with IriScan, Inc., who has authorized the release of the proprietary information contained in this report (see Appendix A).

System Description

This brief system description is only provided to facilitate discussion of the testing methodology and results. It is based on the tester's understanding of the documentation and operation of the system. It is also understood that production units have changed substantially from the system tested. For a full description of the currently available systems, the reader should contact the manufacturer.

Overview

The prototype system was largely self contained. The biometric input device, processor, and memory were contained in one box. Jacks were provided that allowed connection to an external monitor, keyboard, and remote computer. The external monitor was useful to monitor user's performance and is required for enrollment and system maintenance functions. The keyboard was required for entering user data during enrollment. The remote computer connection facilitated backing up templates, updating software, and performing other system management tasks. However, none of these external devices would be required during normal operation of the device as a recognizer. The system was mounted in an equipment rack along with an external monitor and keyboard as shown in Figure 1.

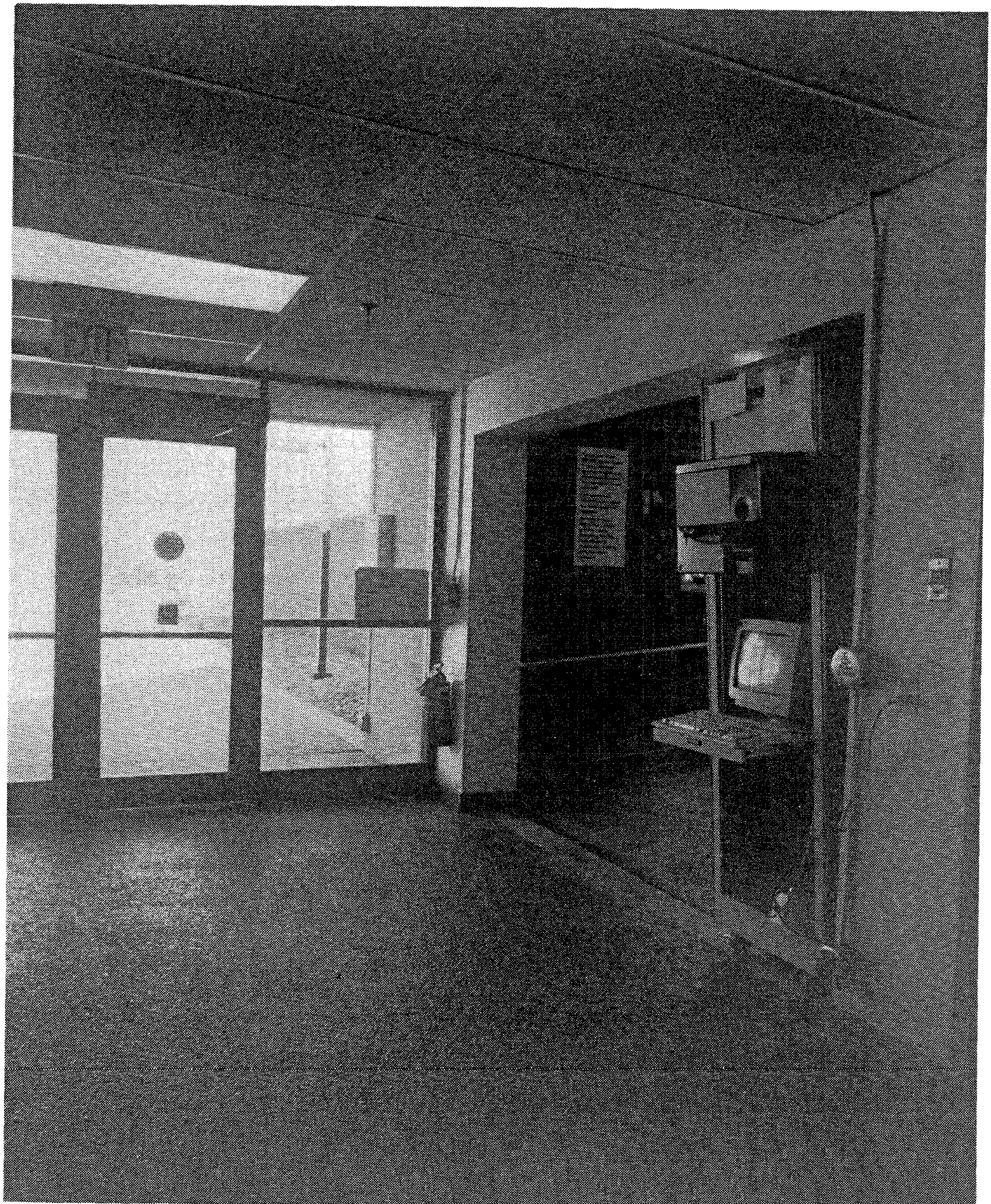
Figure 1 - Test System Configuration

User Interface

The users interface with the system by looking at a backlit LCD monitor that is mounted in a rocking subchassis on the right side of the system chassis. This monitor is located behind a partially reflecting mirror. The reflected image is directed at a video camera that acquires the incoming image and provides the input to the LCD monitor and a frame grabber located in the processor chassis. In this manner, the user can align and focus the image of his own eye by the combination of moving his head and tilting the input subchassis until he sees a centered, focused image of his own eye in the monitor. A small halogen light, located about 2 inches below the imager, provides illumination through a red filter.

Enrollment Process

First, the user aligns his eye as described above. The operator monitors the user's alignment on the external monitor that displays the frame grabber's output. When the user is properly aligned, the operator initiates the enrollment process. The system then acquires three successive images from the frame grabber and generates an iriscode for each image. If the average hamming distance[1] among the three iriscodes is sufficiently low, the operator is then given the option of enrolling that eye. If the operator chooses "yes," he is prompted by the system to provide the name of the user and which eye was imaged. Finally, the operator is offered the option of enrolling the other eye of the same user.

Recognition Process

The user simply pushes a button on the front panel of the main chassis and aligns his eye in the monitor. The system then will make multiple attempts to recognize the individual's eye. The default number of attempts is nine, but this can be changed by the operator. For each attempt, the system acquires images until one of sufficient contrast to be judged in reasonable focus is obtained. It then generates an iriscode for the image and attempts to match the iriscode to the iriscodes in its database. If the hamming distance between the new iriscode and one of the enrolled iriscodes is sufficiently low, the system decides it has found a match, sounds an acceptance beep, and discontinues acquiring images. At this time, it reports the name associated with the enrolled iriscode, which eye it is (left or right), and the hamming distance on the external monitor (if connected), and optionally to a log file. If no match is found after nine attempts, a rejection tone is sounded.

Test Description

The system was set up in the building 822 breezeway. This location was chosen due to the high volume of pedestrian traffic that passes through it every day. This location also had the advantage of simulating an entrance portal, as this breezeway is an entrance to Sandia's Technical Area 1. The disadvantage of this location was the large windows that allowed sunlight in from many directions, which caused reflections in the iris images. Some of these windows can be seen in Figure 1.

---

[1] Hamming distance refers to the number of digit positions in which two binary numerals, characters, or words of the same length are different.

Volunteers were solicited by posters placed in buildings around the test location. Approximately 400 eyes were previously enrolled in the system by the manufacturer, several of them were Sandian's who had been enrolled at trade shows. An additional 122 individuals were enrolled as part of this test. The original intention had been to enroll both eyes of each volunteer. During testing, however, it was observed that some individuals had trouble aligning one of their eyes, making it difficult to enroll that eye. At that point, it was decided to enroll only one eye if the volunteer had difficulty with the second eye. As a result, we ended up with 199 eyes, from 122 individuals, enrolled during the test.

All of the volunteers were given a three-page informed consent form and a brief overview of the system before proceeding with the test. If they chose to proceed with the test, they were enrolled in the system as described above and asked to use the system as often as possible over the 8 days that the system was in place. Each recognition attempt was observed by an operator. The results of these attempts were recorded manually, along with the operator's observations. Each rejection of a valid user was called a false reject (FR). We will refer to all of these transactions as FR attempts.

After eight days of operation, the Sandia enrollments were deleted and the database was restored to the original 403 enrollees as supplied by the manufacturer. The volunteers were then requested to attempt to be recognized one more time. These results were also recorded manually. Each recognition of an individual was to be called a false accept (FA). We will refer to all of these transactions as FA attempts. Test subjects who had been enrolled previously by the manufacturer were excluded from this phase of testing since their enrollment data was not deleted.

Test Results

Enrollment

Enrolling new users was a tricky process. It took some time for most individuals to become oriented with the imaging system. Demonstration of the system by the operator while allowing the new users to watch the external monitor helped this situation and became part of our enrollment process. While the new users attempted to align their eyes, the operators attempted to coach them by monitoring their progress on the external monitor. However, the slow update rate of the external monitor made this difficult. A faster processor, which was provided after our testing was completed, improved that update rate significantly. Once users were oriented, most were enrolled quickly.

Most people naturally enrolled the eye they felt most comfortable using first. This eye was usually what is known as their dominant eye. When we attempted to enroll their other eye, many had difficulty keeping that eye aligned. Some had to go as far as covering their dominant eye in order to enroll the other eye. This difficulty, combined with the discomfort some users expressed at having the system's light shining in their eye while they made repeated attempts to enroll, led us to drop the enrollment of the second eye when the users had difficulty.

One user could not be enrolled on the system despite repeated attempts with both eyes over a period of several days. Detailed photographs of this user's eye were acquired and analyzed. No unique characteristics which would prevent enrollment were identified.

A sampling of enrollment times was recorded. The average enrollment time was 2 minutes and 15 seconds. This time was measured from when we began our overview of the system to when the first eye was completely enrolled. The time to read and sign the consent form was not included.

False Reject Testing

During the first phase of the test, 895 FR transactions were recorded with 106 actual rejections for a raw FR rate of 11.8%. Next, the comments were analyzed and various reasons for false rejects were tabulated. This categorization is subjective, but it yields other ways to look at the data.

The groups we came up with included the following:

1. User or environmental error. This included a number of errors such as a user presenting an eye that was not enrolled or not keeping their eye open until the machine executed its nine attempts. Reflections from external light sources that obscured the iris' image were also counted in this group.
2. Reflections from glasses. All users who normally wore glasses were asked to use the system wearing their glasses. Reflections of the system's light off the user's glasses caused glare to appear in the digitized image. Most users were able to orient themselves so that the glare did not distort the iris in the digitized image. Some users, particularly those with very thick glasses or dirty glasses had difficulty doing this.
3. User Difficulty. User difficulty included hair obscuring the image, difficulty focusing, more difficulty using the system with one eye than the other, etc.

Next an attempt was made at analyzing the error rate without the errors that could be attributed to extreme environmental conditions or deliberate misuse. When the 17 transactions that were flagged as group 1 were removed, we were left with 878 "good" transactions (with 89 FRs) and a FR rate of 10.2%. The other errors would be typical of an actual system in use at an entry control point. Since several of the rejects fell into more than one group, it is difficult to determine what the error rate would be if multiple groups were deleted. Each record needs to be examined in context to see if it needs to be included in a particular scenario.

Many of the volunteers who were rejected were willing to try again. Frequently, they passed on the second attempt. This led us to try to determine the 2-try error rate. For this analysis, the transactions were grouped into events. A 2-try event consisted of an attempt to be recognized, and the repeated attempt with the same eye if rejected. Of the 47 rejections where the users were willing to make a second attempt, 31 passed on the second attempt. Of the 878 "good" transactions, 47 were retries, leaving 831 2-try events. Examining the 47 retries, it was found that 31 passed and 16 were rejected.

This leaves us with 73 rejections on the first attempt, and 42 still rejected after being given the opportunity of a second attempt. This leads to a 1-try FR rate of 8.8% and a 2-try FR rate of 5.1%.

Another way to look at this data is to try to determine the FR rate if only one of the two eyes needed to match the database. This would be a realistic scenario for an entry control point and could be thought of as the 2-eye error rate. Our testing was treating each eye as equally likely to find a match while practice was showing a marked difference between the performance of each eye for some users. With this in mind, we tried to sort through the data to determine for each rejection, if the user had been recognized using the other eye. We found that for 42 of the rejects, the users had made an attempt with the other eye. Of these 42 attempts, the users were recognized 39 times. By a process similar to that used for the 2-try FR rate, we came up with 836 2-eye events. There were 86 rejects after the first attempt and 47 of those were still rejected after being given the opportunity to retry with the other eye. This results in a 1-eye FR rate of 10% and a 2-eye FR rate of 5.6%.

It should be noted that only about half of those rejected after the first attempt tried again either with the same eye, or the other eye. If all the volunteers had tried again, we might have achieved a lower error rate. However, we were limited to the number of transactions that the volunteers felt comfortable providing.

We have already discussed the dominant eye's surprising effect on the system's performance. Several other unanticipated factors were noted that may have significant effects on performance. These factors were noted by reviewing the operator's comments. The factors were not part of the test design, so we do not have consistent records that would allow correlations to be determined. The comments indicate that user's height, thickness of glasses, and weight may affect their false reject rates.

Extremely tall and short people seemed to have more difficulty because the system is designed for someone of average height to look directly into it. Very tall and short people had to hold their heads at an angle to keep their eye parallel with the imager. The combination of awkward bends made it difficult for them to achieve clear images. We had one 5'1" volunteer who needed to use a stepstool to use the system. Other tall users were looking down, sometimes to the point where the lower rim of their glasses interfered with the image. As mentioned earlier, wearers of thick glasses had trouble providing glare-free iris images. Some overweight people seemed to have more of their iris covered by their eyelid, causing higher reject rates. Any follow up study should look into how these factors affect performance

The average transaction time of a sampling of transactions was 14 seconds. The minimum time recorded was 6 seconds. The maximum time was 23 seconds. This time was measured from when the user first pushed the start button to when the system beeped indicating recognition or rejection.

False Accept Testing

Many biometric devices are verifiers. The IriScan is a recognizer. The difference is that verifiers require the user to claim an identity first and then attempts to see if the biometric sample is a close enough match to the enrolled template that matches that identity. When doing false accept testing, the users are typically asked to try to match a set of other volunteers' identities, leading to a large number of transactions. When testing a recognizer, we only require one transaction of each person that is not in the database. The recognizer tries to match against each member of the database.

For this testing, the database was limited to the original 403 templates that the manufacturer provided. Only experienced users who had participated in the false reject portion of the test were allowed to participate. Each volunteer was asked to attempt to be recognized one time. During this test, 96 FA transactions were recorded, with no actual false accepts occurring. This is analogous to performing 96 times 403 or 38,688 transactions on a verifier.

System Comments

The system that was tested included the ability to record all recognitions to a log file. It is desirable that such a system log all transactions (including enrollments), attempts to access, whether they were successful or not, and date and time. From a tester's perspective, it would be nice if the log included the hamming distances for all the enrolled templates for each transaction (or at least the several nearest templates). This feature would allow verification of the hamming distance distributions that the developer has published and would eliminate the need to do false accept testing.

One further item that might be of interest to some users would be the ability to store an image of the eye of persons who attempt to be recognized but failed. This could amount to an instant enrollment and could alarm security personnel when that same eye appears again. It could also be used to help prosecute intruders.

Conclusions

The system performed extremely well in difficult conditions. While every effort was made to closely monitor each transaction for anomalies, it is possible that reflections from the high ambient light of the test location could have led to some of the false rejects. This would indicate that if the system was to be used outdoors or near windows, an enclosure or at least side shades would be required. Still the system had a good false reject rate and no false accepts were observed indicating potential for high security applications.

The faster processor might even help reduce the false reject rate. Since the faster processor allows the system to complete each search more quickly, the user is less likely to get impatient or tired of looking into the monitor. In addition, this might allow users to simply stand with their feet against a prepositioned footstop and rock slowly back and forth in front of the imager. The system then acquires images until it gets one in good focus which has a match in its database. This process was suggested by IriScan engineers and alleviates the problem of users who had difficulty achieving sharp focus. Unfortunately, the processor was not available until our testing was complete, but a few test runs among the experimenters showed significantly lower acceptance times.

This test has shown that the system has a great deal of potential for use as a high-confidence biometric identifier. If there is sufficient interest in the DOE community, it should be a followed up by a more complete study of a production system. Such a study should take into account the performance affecting variables that were found during this test. A follow-up study should also include a vulnerability analysis. Such a complete study would determine the suitability of the IriScan system for high security applications.

APPENDIX A

IriScan

Setting the standard in biometric identification™

Company Addendum to the Sandia Test Report on Laboratory Evaluation of the Prototype IriScan Biometric Identifier

IriScan appreciates the early prototype test effort performed by Sandia National Laboratories, which was accomplished with the objective professionalism that the government and industry have come to expect of SNL.

When IriScan and SNL arranged the early test of the first prototype, there was mutual agreement that distribution of the test results would not be publicly disclosed in consideration of the fact that an early non-production system could not be expected to fairly represent the ultimate performance characteristics of production models. Indeed, the test results partially confirmed that reasoning with respect to false reject results, which are significantly different than those obtained with advanced prototypes in another government test, as well as with the System 2000EAC, now in production. At the same time, zero false accepts were expected and confirmed.

IriScan and SNL have subsequently decided to release the results of the early test. This decision was based on a compelling interest in IriScan technology by users and the overall excellent performance of the prototype under "very difficult" conditions.

Readers of this early test report should understand that a number of subjective factors can significantly impact false reject performance in a biometric system that may be unrelated to the technical capabilities of the system. In the IriScan prototype, these included: siting in an uncontrolled ambient light environment; human factors design in switch placement; height adjustment mechanism design; illuminator type and pattern; slower processor (33 MHz); and an earlier version of the software that performed image processing at less than optimum levels. Many of these factors were altered in the advanced prototypes for DoD and industry, and as a result, crossover error rates at less than 1% were consistently achieved in DoD testing as well as continuing internal evaluations. All factors that impacted rejection of authentics were also addressed in the design of the production System 2000EAC.

IriScan is eager to meet the challenge of performance testing of production systems in any appropriate operating environment.

John E. Siedlarz
President and CEO
IriScan, Inc.

133-Q Gaither Drive, Mt. Laurel, NJ 08054-1701
(609) 234-7977 • Fax (609) 234-4768 • Email: iriscan@aol.com

Distribution:

1	MS9018	Central Technical Files, 8523-2
5	0899	Technical Library, 4414
1	0619	Print Media, 12615
2	0100	Document Processing, 7613-2 For DOE/OSTI
1	0769	D. S. Miyoshi, 5800
3	0762	Safeguards and Security Library, 5807
24	0781	Entry Control/Systems Engineering, 5848