

FACE RECOGNITION VENDOR TEST 2002

Technical Appendices March 2003

**P. JONATHON PHILLIPS^{1,2}, PATRICK GROTH², ROSS J. MICHEALS²,
DUANE M. BLACKBURN³, ELHAM TABASSI², MIKE BONE⁴**

¹DARPA
3701 North Fairfax Dr.
Arlington, VA 22203

²National Institute of Standards and Technology
100 Bureau Drive, Stop 8940
Gaithersburg, MD 20899

³DoD Counterdrug Technology Development Program Office
17320 Dahlgren Rd, Code B07
Dahlgren, VA 22448

⁴NAVSEA Crane Division
300 Highway 361, Code 4041
Crane, IN 47522

Sponsors and Supporters:

Defense Advanced Research Projects Agency
Department of State
Federal Bureau of Investigation
National Institute of Justice
National Institute of Standards and Technology
Transportation Security Administration
ONDCP Counterdrug Technology Assessment Center
United States Customs Service

Department of Energy
Drug Enforcement Administration
Immigration and Naturalization Service
United States Secret Service
Technical Support Working Group
Australian Customs
Canadian Passport Office
United Kingdom Biometric Working Group

This page intentionally blank.

**Face Recognition Vendor Test 2002:
Technical Appendices**

**P. Jonathon Phillips^{1,2}, Patrick Grother², Ross Micheals², Duane M. Blackburn³, Elham Tabassi²
J. Mike Bone⁴**

¹DARPA
3701 North Fairfax Drive
Arlington, VA 22203

²National Institute of Standards and Technology
100 Bureau Drive; Stop 8940
Gaithersburg, MD 20899

³DoD Counterdrug Technology Development Program Office
17320 Dahlgren Rd, Code B07
Dahlgren, VA 22448

⁴NAVSEA Crane Division
300 Highway 361, Code 4041
Crane, IN 47522

Abstract

The *Face Recognition Vendor Test 2002: Technical Appendices* is a companion document to the *Face Recognition Vendor Test 2002: Evaluation Report*. This document provides supplemental material that is not included in the *Evaluation Report*. The *Technical Appendices* present all material provided to the Face Recognition Vendor Test (FRVT) 2002 participants, detailed descriptions of the evaluation protocols and testing procedures, additional performance results, and description of participant systems and comments on the results². Appendix E contains submissions to the frequently asked questions section of the website and the answers to the questions.

¹ Please direct correspondence to Jonathon Phillips at jphillips@darpa.mil or Jonathon@nist.gov. The *Face Recognition Vendor Test 2002: Evaluation Report* is available at <http://www.frvt.org>.

² Participant documents (system descriptions and comments on results) are the opinions of the Participants and are provided in this text for reference. Inclusion of these documents does NOT imply that FRVT 2002 Authors/Sponsors/Supporters agree with any statements or derivative results within those documents. The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology or any other FRVT 2002 Author/Sponsor/Supporter.

The *Face Recognition Vendor Test 2002: Technical Appendices* is a companion document to the *Face Recognition Vendor Test 2002: Evaluation Report*. The *Technical Appendices* provide supplemental information about the procedures used in FRVT 2002, documentation provided to participants prior to testing, additional Face Recognition Vendor Test (FRVT) 2002 results and participant-provided documentation describing their systems and their comments to the *Evaluation Report*. The appendices provide a comprehensive description of the evaluation protocol and testing procedure, and provide the necessary background for preparing and administering future technology evaluations. Table i list the title of the appendices in the *Technical Appendices*. A short description of each appendix follows.

Table i - List and Title of Appendices.

Appendix	Content
B	FRVT 2002 Announcement
C	FRVT 2002 Website
D	Forms to Participate in FRVT 2002
E	Frequently Asked Questions
F	FRVT 2002 Methods and Procedures
G	FRVT 2002 Technical Data Set
H	FRVT 2002 Test Plan
I	Image Collection Process and Image Description
J	FRVT 2002 Normalization API
K	Development Image Set
L	Detailed Results
M	Participant Product Description
N	Participant Comments on FRVT 2002 Evaluation Report
O	Additional Information on the 3D Morphable Model Used in FRVT 2002

Appendix B contains the documents that announced the FRVT 2002.

Appendix C contains a copy of the FRVT 2002 Website that was posted during the evaluation. This includes the publicly available portion of the website as well as the restricted-access participant area. The participant area contained information that was only available to participants and tentative participants during the evaluation process.

Appendix D contains copies of the two forms that FRVT 2002 participants were required to complete and sign to participate in FRVT 2002. One form explained the ground rules associate with participating in FRVT 2002. The second form explained restrictions on the use of FRVT 2002 imagery.

Appendix E contains a copy of submitted frequently asked questions and their answers. The questions and answers provide additional clarification on the FRVT 2002.

Appendix F is a description of the methods, procedures and timelines used to manage FRVT 2002.

Appendix G contains the documentation that was included in the FRVT 2002 technical datasets. The technical dataset documented the interfaces for reading target and query sets and writing the similarity files.

Appendix H contains the documentation provided to FRVT 2002 participants that describes the on-site procedures followed during the administration of the test.

Appendix I describes the imagery in the HCInt and MCInt datasets.

Appendix J describes procedures and API documentation for the normalization component of FRVT 2002.

Appendix K shows the sample images given to FRVT 2002 participants prior to testing. The sample image was designed to give participants a basic “feel” for the types of images that were included in the FRVT 2002 data sets. The sample image set was not designed for either developing or tuning of face recognition systems.

Appendix L. In order to clearly present an assessment of the state-of-the-art in face recognition, it was necessary in the *Evaluation Report* to report summary statistics of performance. Appendix L presents full cumulative match characteristics (CMC) and receiver operating characteristics (ROC) for all MCInt experiments (Section 8 and Appendix A.4 in the *Evaluation Report*). Also, this appendix presents HCInt verification error ellipses ROCs (Figure 8) for all participants in Figure 8.

Appendix M provides a description and cost of each system that was tested. The descriptions were given to the FRVT 2002 organizers by the participants prior to commencement of the FRVT 2002 evaluation.

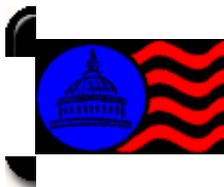
Appendix N contains the position papers written by FRVT 2002 Participants after viewing the evaluation results.

Appendix O contains additional information on the 3D Morphable Model used in FRVT 2002.

This page intentionally blank.

Appendix B

FRVT 2002 Announcements



FBO DAILY ISSUE OF APRIL 27, 2002 FBO #0146
SOURCES SOUGHT

66 -- SOURCES SOUGHT FOR FACIAL RECOGNITION SYSTEMS

Notice Date

4/25/2002

Notice Type

Sources Sought

Contracting Office

Department of Commerce, National Institute of Standards and Technology (NIST), Acquisition and Logistics Division, 100 Bureau Drive, Building 301, Room B129, Mail Stop 3571, Gaithersburg, MD, 20899-3571

ZIP Code

20899-3571

Solicitation Number

SOURCESSOUGHTFACIALRECOG

Response Due

5/10/2002

Archive Date

5/25/2002

Point of Contact

Tamara Grinspon, Contract Specialist, Phone (301) 975-4390, Fax (301) 975-8884, - Tamara Grinspon, Contract Specialist, Phone (301) 975-4390, Fax (301) 975-8884,

E-Mail Address

tamara.grinspon@nist.gov, tamara.grinspon@nist.gov

Description

The National Institute of Standards and Technology (NIST) wishes to identify potential sources to provide commercial, or soon-to-be commercial, face (facial) recognition systems. This is a market research survey being conducted for planning purposes only and is not to be construed as a commitment by the Government. THIS IS NOT A REQUEST FOR PROPOSALS. Available systems may be selected by the Government to participate in the Facial Recognition Vendor Test 2002, a government-sponsored technology evaluation of face recognition technologies. Potential sources of face recognition technology will not be funded, nor charged, to participate in FRVT

2002. A special format for gathering this information has been established in the FRVT 2002 "How to Participate" page at <http://www.frvt.org> . A deadline for submission is shown on the "Calendar" page of <http://www.frvt.org> . --

Place of Performance

Address: National Institute of Standards and Technology, Shipping and Receiving, Building 301, 100 Bureau Drive, Gaithersburg, MD 20899-0001

Record

SN00066141-W 20020427/020425213206 (fbodaily.com)

Source

[FedBizOpps.gov Link to This Notice](#)
(may not be valid after Archive Date)

| [FSG Index](#) | [This Issue's Index](#) | [Today's FBO Daily Index Page](#) |

Blackburn Duane M DLVA

From: Duane Blackburn [BlackburnDM@nswc.navy.mil]
Sent: Thursday, April 25, 2002 7:56 PM
To: BIOMETRICS@PEACH.EASE.LSOFT.COM
Subject: FRVT 2002 Announcement

In the USA Patriot Act, the National Institute of Standards (NIST) is mandated to measure the accuracy of biometric technologies. In accordance with this legislation, NIST, in cooperation with 14 other Government agencies, will conduct the Facial Recognition Vendor Test 2002. The evaluation will be held in the summer of 2002 with published results expected in the fall. The same individuals that managed the FRVT 2000 are also managing the FRVT 2002 evaluation.

Information about FRVT 2002 has been made available on the Internet at <http://www.frvt.org> including:

- Overview of FRVT 2002
- Sponsoring agencies
- Qualification requirements for participants
- Calendar of events

The FRVT 2002 is a government-sponsored technology evaluation of face recognition technology. Participation is open to anyone selling a commercial or mature prototype face recognition system, subject to the requirements and procedures outlined on the FRVT 2002 website.

The sponsors and supporters of FRVT 2002 thank you for the opportunity of providing this service for the government, public and biometrics industry.

The preceding was forwarded by the Biometric Consortium's Electronic Discussion Group. Any opinions expressed here do not necessarily reflect those of the Biometric Consortium. Further distribution is prohibited.

Problems and questions regarding this list should be sent to BIOMETRICS-request@PEACH.EASE.LSOFT.COM.

To remove your name from this list please send the command "SIGNOFF BIOMETRICS" to LISTSERV@PEACH.EASE.LSOFT.COM. Please do not send the "SIGNOFF BIOMETRICS" command to the BIOMETRICS list.

To update membership information (new e-mail address etc.), please send a message to bailey@biometrics.org providing the updated information.

From: "FRVT2002" <frvt2002@brtrc.com>
Subject: FRVT 2002 Announcement
Date: Thu, 25 Apr 2002 19:57:22 -0400
To: frvt2002@brtrc.com



In the USA Patriot Act, the National Institute of Standards (NIST) is mandated to measure the accuracy of biometric technologies. In accordance with this legislation, NIST, in cooperation with 14 other Government agencies, will conduct the Facial Recognition Vendor Test 2002. The evaluation will be held in the summer of 2002 with published results expected in the fall. The same individuals that managed the FRVT 2000 are also managing the FRVT 2002 evaluation. Information about FRVT 2002 has been made available on the

Internet at <http://www.frvt.org> including:

- Overview of FRVT 2002
- Sponsoring agencies
- Qualification requirements for participants
- Calendar of events

The FRVT 2002 is a government-sponsored technology evaluation of face recognition technology. Participation is open to anyone selling a commercial or mature prototype face recognition system, subject to the requirements and procedures outlined on the FRVT 2002 website.

The sponsors and supporters of FRVT 2002 thank you for the opportunity of providing this service for the government, public and biometrics industry.

BIOMETRICS

CATALOG

[Results List](#)[Modify Catalog Entry](#)**Modified:** 4/25/02 8:05:00 PM**Category:** Government Tests and Deployments**Biometric Type:** Facial Recognition**Vendor:** n/a**Title:** Facial Recognition Vendor Test 2002

Description: In the USA Patriot Act, the National Institute of Standards (NIST) is mandated to measure the accuracy of biometric technologies. In accordance with this legislation, NIST, in cooperation with 14 other Government agencies, will conduct the Facial Recognition Vendor Test 2002. The evaluation will be held in the summer of 2002 with published results expected in the fall. The same individuals that managed the FRVT 2000 are also managing the FRVT 2002 evaluation.

Information about FRVT 2002 has been made available on the Internet at <http://www.frvt.org> including:

- Overview of FRVT 2002
- Sponsoring agencies
- Qualification requirements for participants
- Calendar of events

The FRVT 2002 is a government-sponsored technology evaluation of face recognition technology. Participation is open to anyone selling a commercial or mature prototype face recognition system, subject to the requirements and procedures outlined on the FRVT 2002 website.

The sponsors and supporters of FRVT 2002 thank you for the opportunity of providing this service for the government, public and biometrics industry.

Additional References

Related sites: [FRVT Homepage](#)

Appendix C

FRVT 2002 Website

FACE RECOGNITION VENDOR TEST 2002



Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

Facial Recognition
Vendor Test 2000

FERET

Evaluation
Methodology

Home

Introduction

In the USA Patriot Act, the National Institute of Standards (NIST) is mandated to measure the accuracy of biometric technologies. In accordance with this legislation, NIST, in cooperation with other Government agencies, is conducting the Face Recognition Vendor Test 2002. The evaluation is being held in the summer of 2002 with results expected to be published in the fall 2002. FRVT 2002 Sponsors and Supporters are:

- Defense Advanced Research Projects Agency
- Department of State
- Federal Bureau of Investigation
- National Institute of Justice
- National Institute of Standards and Technology
- Transportation Security Administration
- ONDCP Counterdrug Technology Assessment Center
- United States Customs Service
- Department of Energy
- Drug Enforcement Administration
- Immigration and Naturalization Service
- Secret Service
- Technical Support Working Group
- Australian Customs
- Canadian Passport Office
- United Kingdom Biometric Working Group

The following individuals supervise and manage FRVT 2002:

- Dr. P. Jonathon Phillips, Test Director
- Duane Blackburn
- Mike Bone
- Patrick Grother
- Dr. Ross Michaels

One objective of the Patriot Act legislation is to "...develop and certify a technology standard, including appropriate biometric identifier standards, that can be used to verify the identity of persons applying for a US visa or such persons seeking to enter the US pursuant to a visa for the purposes of

conducting background checks, confirming identity, and ensuring that a person has not received a visa under a different name...". FRVT 2002 satisfies the US Patriot Act face recognition evaluation mandate. Results provide as input for future decisions to be made in accordance with the Patriot Act and subsequent legislation.

Participation

To assess the state-of-the-art in face recognition technology, the FRVT 2002 is conducting a technology evaluation of both mature prototype and commercially available systems face recognition systems. Core face recognition technology developers were invited to participate. The deadline for signing up for the FRVT 2002 was 10 May 2002. A list of FRVT 2002 Participants can be found on the [Participants Page](#).

FRVT 2002 Status Update - 17 August 2002

Administering the FRVT 2002 to the participants has been completed, and the FRVT 2002 organizers are in the processing of computing the performance measures and writing the report. The report is due to be released in November 2002, and will be available from this website.

To find out more about the FRVT 2002, attend the two FRVT 2002 talks at the [Biometric Consortium Conference 2002](#) 23-25 September in Crystal City, VA. The talks are "Overview of FRVT 2002" to be presented Monday morning 23 September, and "New Evaluation Statistics for Measuring the Performance of Biometric Technologies", Wednesday morning.

FRVT 2002 was administered to fourteen [participants](#) at the Navy facility in Dahlgren, Virginia, from 10 July through 9 August 2002.

In the original proposed design, the FRVT 2002 consisted of four tests. However, because of higher than expected response, the organizers redesigned the FRVT 2002 to consist of two tests. These are the High Computational Intensity (HCInt) Test and the Medium Computational Intensity (MCInt) Test. Both test required the systems to be full automatic, and manual intervention was not allowed. Participants could sign up to take either or both tests.

The High Computational Intensity (HCInt) Test was designed to test state-of-the-art systems on extremely challenging real-world images. These were full-face still frontal images. This test compares still database images against still images of an unknown person. The HCInt required participants to process a set of approximately 121,000 images, and match all

possible pairs of images from the 121,000 image set. This required performing 15 billion matches. The results from the HCIInt will measure performance of face recognition systems on large databases, examine the effect of database size on performance and estimate variability in system performance.

The Medium Computational Intensity (MCInt) Test consists of two separate parts: still and video. The Medium Computational Intensity Subtest was designed to provide an understanding of a participant's capability to perform face recognition tasks with several different formats of imagery (still and video) under varying conditions. This subtest will also help identify promising face recognition technologies that did not take the High Computational Intensity test. The still portion of the MCInt is similar to the FERET and FRVT 2000 evaluations. It compares a database of still images against still images of unknown people. The still portion of the MCInt is designed to measure performance on different categories of images. Examples of different effects that will be measured are time between images, changes in illumination, and variation in pose. The video portion is designed to provide an initial assessment of whether or not video helps increase face recognition performance. This portion uses video style imagery that is extracted from digital video sequences. The video portion compares both a database of still images against video sequences of an unknown person and a database of video sequences against video sequences of an unknown person.

FRVT 2002 Status Update - 26 November 2002

The Face Recognition Vendor Test (FRVT) 2002 conducted experiments with approximately 121,000 images in its database. This resulted in approximately 15 billion comparisons and 60 gigabytes of raw data for each participant. This represents a major advancement in face recognition evaluations.

The FRVT 2002 Evaluation Team has completed scoring on the initial set of experiments planned for FRVT 2002. In the course of scoring the initial experiments, the FRVT 2002 Evaluation Team discovered a number of new and interesting results. These results led to the creation of additional experiments that will describe strengths and weaknesses of face recognition systems that have not previously been studied. The FRVT 2002 Evaluation Team is currently in the process of scoring participants for these additional experiments.

The decision was made to include these additional experiments in the FRVT 2002 Evaluation Report,

thus causing a slight delay in its release. The FRVT 2002 Evaluation Team apologizes for this delay, but feels it is justified as the results of these new experiments have the potential to directly impact how and where face recognition systems should be fielded.

The experiments, analysis, and report generation will be completed in December 2002. Once the report is completed it will be submitted for public release approval. Once this approval is received, the FRVT 2002 Evaluation Report will be posted on the [FRVT 2002 website](#).

Communications

To maintain fairness, the Government is adhering to the following procedures for FRVT 2002. Information provided to any single participant will also be provided to all participants. The primary and preferred source of communications between Participants and FRVT 2002 government personnel is the frequently asked question (FAQ) page. A government employee has been named as the FRVT Liaison. The FRVT Liaison will be the primary point of contact and will manage communications with Participants. Questions that are outside the purview of the FAQ can be submitted to the FRVT Liaison via email at frvt2002@nist.gov. All e-mails sent to the FRVT Liaison and germane to the FRVT 2002 will be posted on the FAQ page. This policy ensures all Participants are guaranteed equal access to information concerning FRVT 2002. Contact with the FRVT Liaison other than via the FAQ areas, or contact with someone other than the non-FRVT Liaison is only allowed for extenuating circumstances. Media inquiries should be sent to Philip Bulman of the NIST Public Affairs Office at philip.bulman@nist.gov.

The FRVT 2002 Evaluation Report will be placed on the FRVT website at <http://www.frvt.org> when it has been completed and approved for public release. We will announce the release on the [Biometrics Consortium listserv](#).

FACE RECOGNITION VENDOR TEST 2002

How to Participate

Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

*Facial Recognition
Vendor Test 2000*

*Face Recognition
Vendor Test 2002*

FERET

*Evaluation
Methodology*

Home

In order to request participation in the FRVT 2002, potential participants must electronically fill out two forms:

[Application to Participate in FRVT 2002](#)

[Application for Access to the FRVT 2002 Technical Datasets](#)

Both forms are to be filled out on-line, printed, signed, and sent to the locations designated on each form. Upon receipt of both original signed forms by the Government, applicants will be classified as a Tentative Participant. The Government must receive both forms by the due date described in the [FRVT 2002 Calendar Page](#).

The Application to Participate in the FRVT 2002 explains the rules, process, and procedures governing the administration of FRVT 2002. Additional technical details will be made available at a later date in the FRVT2002 Test Plan.

The Application for Access to the FRVT 2002 Technical Datasets explains rules governing access and use of the:

FRVT 2002 evaluation data sets.

FRVT 2002 training set.

Software provided with the FRVT 2002 training set.

Conducting a technology evaluation requires that the sets of imagery provided to the Participants and results of the evaluation are in specified formats. The FRVT 2002 training set and associated software provides an example of the formats that will be used in FRVT 2002. This will provide Participants with the necessary imagery and tools to check that their system can read FRVT 2002 test image sets and that their systems output the results in the correct format. The imagery data sets and results format are in XML. The specifications for the XML format are included in the software provided with the FRVT 2002 training set. **Very important:** The baseline PCA algorithm included in the FRVT 2002 Technical Datasets requires normalized (eye-coordinates given) images. Images provided during FRVT 2002 will **NOT** be normalized. Systems evaluated will need to automatically process facial imagery without eye coordinates being provided.

The FRVT 2002 evaluation protocol is based on the evaluation protocol used in the September 1996 FERET and Face Recognition Vendor Test 2000 (FRVT 2000) evaluations. FRVT 2002 will evaluate systems that can automatically process facial imagery without eye coordinates being provided. For a subtest,

a system is given two lists of signatures; list **A** and list **B**. In the FRVT 2002, a signature consists one still image or a set of still images for the video subset. The system then has to match every signature in list **A** with every signature in list **B**. Each match yields a similarity score. The complete set of similarity scores between list A and list B is a similarity matrix. The format for the similarity matrix is contained in the FRVT 2002 Technical Datasets distribution. From the similarity matrix, performance scores and statistics are computed.

The desired method of communications between participants and the FRVT 2002 is the Participant FAQ page in the Participant Area. The Participant Area FAQ is not limited to simple questions, but will serve as the method of communications between the FRVT 2002 Liaison and Participants. All pertinent e-mail sent to the FRVT 2002 Liaison will be posted on the Participant FAQ page. This policy is to ensure that all participants are guaranteed equal access to information concerning FRVT 2002. Exceptions will only be allowed for extenuating circumstances. Two examples of the exceptions are delivery of login information for Participant Area or resolving major issues with similarity files.

FACE RECOGNITION VENDOR TEST 2002



Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

Facial Recognition
Vendor Test 2000

Face Recognition
Vendor Test 2002

FERET

Evaluation
Methodology

Home

Company: AcSys Biometrics Corp
Web Site: <http://www.acsysbiometricscorp.com>

Company: C-VIS Computer Vision und Automation GmbH
Web Site: <http://www.c-vis.com>

Company: Cognitec Systems GmbH
Web Site: <http://www.cognitec-systems.com>

Company: Dream Mirh Co., Ltd
Web Site: <http://www.dreammirh.com>

Company: Electronic System Products Inc.
Web Site: <http://www.espipd.com>

Company: Eyematic Interfaces Inc.
Web Site: <http://www.eyematic.com>

Company: Iconquest

Company: Identix
Web Site: <http://www.identix.com>

Company: Imagis Technologies Inc.
Web Site: <http://www.imagistechnologies.com>

Company: IRID Inc.
Web Site: <http://iridinc.com>

Company: Phoenix Systems Corporation
Web Site: <http://www.itsphoenix.com>

Company: VicarVision
Web Site: <http://www.vicarvision.nl>

Company: Viiasge Technology
Web Site: <http://www.viisage.com>

Company: VisionSphere Technologies Inc.
Web Site: <http://www.visionspheretech.com>

FACE RECOGNITION VENDOR TEST 2002



Sponsors

Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

Facial Recognition
Vendor Test 2000

Face Recognition
Vendor Test 2002

FERET

Evaluation
Methodology

Home

Multiple federal agencies are sponsoring and supporting the FRVT 2002. Sponsoring agencies are those providing financial, physical, and manpower support for FRVT 2002. Supporting agencies are not providing direct sponsorship, but have given their support for FRVT 2002.

Sponsors

Defense Advanced Research Projects Agency. DARPA is the central research and development organization for the Department of Defense (DoD). It manages and directs selected basic and applied research and development projects for DoD, and pursues research and technology where risk and payoff are both very high and where success may provide dramatic advances for traditional military roles and missions and dual-use applications.

U.S. Department of State. As the lead U.S. foreign affairs agency, the Department of State helps to shape a freer, more secure, and more prosperous world through formulating, representing, and implementing the President's foreign policy.

Federal Bureau of Investigation. The Mission of the FBI is to uphold the law through the investigation of violations of federal criminal law; to protect the United States from foreign intelligence and terrorist activities; to provide leadership and law enforcement assistance to federal, state, local, and international agencies; and to perform these responsibilities in a manner that is responsive to the needs of the public and is faithful to the Constitution of the United States.

National Institute of Justice. NIJ is the research agency of the U.S. Department of Justice. Created by the Omnibus Crime Control and Safe Streets Act of 1968, as amended, NIJ is authorized to support research, evaluation, and demonstration programs, development of technology, and both national and international information dissemination.

National Institute of Standards and Technology. NIST was established by Congress "to assist industry in the development of technology ... needed to improve product quality, to modernize manufacturing processes, to ensure product reliability ... and to facilitate rapid commercialization ... of products based on new scientific discoveries." An agency of the U.S. Department of Commerce's Technology Administration, NIST strengthens the U.S. economy and improves the quality of life by working with industry to develop and apply technology, measurements, and standards.

Transportation Security Administration. On November 19, 2001, the President signed into law the Aviation and Transportation Security Act, which among other things established a new Transportation Security Administration (TSA) within the Department of Transportation. TSA's

mission is to protect the Nation's transportation systems to ensure freedom of movement for people and commerce.

Supporters

[Counterdrug Technology Assessment Center](#). The Counterdrug Technology Assessment Center (CTAC) was established within the Office of National Drug Control Policy as the central counterdrug technology research and development organization of the U.S. Government. In 1998, CTAC's role was expanded to support a program to transfer successfully developed counterdrug technologies directly to state and local law enforcement agencies. The Technology Transfer Program concentrates on providing state-of-the-art, affordable, easily integrated and maintainable tools to enhance the capabilities of state and local law enforcement agencies for counterdrug missions.

[Customs Service](#). The United States Customs Service is the primary enforcement agency protecting the Nation's borders. It is the only border agency with an extensive air, land, and marine interdiction force and with an investigative component supported by its own intelligence branch.

[Department of Energy](#). The Department of Energy develops and implements energy policy and manages a vast array of technical programs. The DOE complex includes unique capabilities in science and engineering that are applied to meet the Department's goals in Energy Resources, Nuclear National Security, Environmental Quality, and Science.

[Drug Enforcement Administration](#). The mission of the Drug Enforcement Administration (DEA) is to enforce the controlled substances laws and regulations of the United States and bring to the criminal and civil justice system of the United States, or any other competent jurisdiction, those organizations and principal members of organizations, involved in the growing, manufacture, or distribution of controlled substances appearing in or destined for illicit traffic in the United States; and to recommend and support non-enforcement programs aimed at reducing the availability of illicit controlled substances on the domestic and international markets.

[Immigration and Naturalization Service](#). The INS is a Federal agency within the US Department of Justice (DOJ) that administers the nation's immigration laws. The INS conducts immigration inspections of travelers entering (or seeking entry) to the United States as they arrive at officially designated ports of entry, regulates permanent and temporary immigration to the United States, maintains control of U.S. borders, and identifies and removes people who have no lawful immigration status in the United States.

[Secret Service](#). The mission of the Secret Service is to protect the President, the Vice President, and their families, heads of state, and other designated individuals; the investigation of threats against these protectees; the protection of the White House, Vice President's Residence, Foreign Missions, and other buildings within Washington, D.C.; and security design, planning, and implementation at designated National Special Security Events. The Secret Service is also responsible for the enforcement of laws relating to counterfeiting of obligations and securities of the United States, investigation of financial crimes including, but not limited to access device fraud, financial institution fraud, identity theft,

computer fraud, telecommunications fraud, and computer based attacks on our nation's financial, banking, and telecommunications infrastructure.

[Technical Support Working Group](#). The Technical Support Working Group (TSWG) is the U.S. national forum that identifies, prioritizes, and coordinates interagency and international research and development (R&D) requirements for combating terrorism. The TSWG rapidly develops technologies and equipment to meet the high-priority needs of the combating terrorism community, and addresses joint international operational requirements through cooperative R&D with major allies.

[Australian Customs Service](#). The Australian Customs Service is responsible for managing the integrity and security of Australia's borders. It facilitates legitimate movements of people and goods into and out of Australia, while detecting and deterring unlawful movements. Australian Customs works closely with other government and international agencies. It has a fleet of ocean-going patrol vessels and contracts two major aerial surveillance providers for civil maritime surveillance and response. Interception of illegal drugs is a high priority and sophisticated techniques are used to target high risk aircraft, vessels, cargo, postal items and travellers. This includes intelligence analysis, computer-based analysis, detector dogs and various other technologies. Customs vision is to be a world leader in Customs administration delivering high quality service to the community, industry and commerce.

[Canadian Passport Office](#). The Passport Office of Canada, an Agency of the Department of Foreign Affairs and International Trade, is charged with the issuing, revoking, withholding, recovery and use of Canadian passports. It provides guidance to missions issuing passports abroad and supervises all matters relating to Canadian travel documents. The fundamental purpose of the Passport Office is to facilitate the travel of Canadians throughout the world by providing internationally respected travel documents and services. The Passport Office works in alliance with other government agencies to provide secure identification documents and services.

[United Kingdom Biometric Working Group](#). The UK Biometrics Working Group (BWG) co-ordinates the Office of the e-Envoy (OeE) Biometrics Programme, the goal of which is to enable the use of biometric authentication technology to support the OeE e-government aims and to facilitate the adoption of biometrics in support of wider government business. The BWG's mission includes providing general advice on biometric technology and procurement, addressing privacy and Human Rights issues, directing technical areas such as evaluation and testing, and offering biometric policy advice and security assurance.

FACE RECOGNITION VENDOR TEST 2002



Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

Facial Recognition
Vendor Test 2000

Face Recognition
Vendor Test 2002

FERET

Evaluation
Methodology

Home

April 25 Announcement to Public

May 10 Deadline to Request Participation

July 10 - August 9 FRVT 2002 Testing

November FRVT 2002 Evaluation Report Released

FACE RECOGNITION VENDOR TEST 2002



Participant Area

Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

Facial Recognition
Vendor Test 2000

Face Recognition
Vendor Test 2002

FERET

Evaluation
Methodology

Home

Login:

Password:

[Change Password](#)



Welcome to the Participant Area of the Face Recognition Vendor Test 2002 (FRVT 2002). This area is for participants only and contains information that is not to be divulged to anyone (even others within participant's companies) besides those directly involved with the FRVT 2002.

Technical Data Set Overview

The [Technical Data Set](#) has been made available on the participant area of the FRVT 2002 webpage. Three documents are included in this set:

1. [FRVT 2002 Implementer's Guide](#). Explains the format you will be receiving image data and the requirements for your output similarity data. This is the document that needs to be read to be able to take the FRVT 2002.
2. [Computational Requirements for FRVT 2002 High Computational Intensity Test](#). An example of how to compute the time required for participant systems to complete the subtests. Participants taking the high computation intensity test need to read this and complete a short white paper described in this document.
3. [HEF Guide](#). Sample image files, FR algorithm, and scoring code. This is a reference guide.

Schedule

June 3	Technical Data Set available on Participant Area
June 14	Deadline for High Computational Intensity Test Participants to submit their computational requirement white paper. Details for this white can be found in Computational Requirements for FRVT 2002 High Computational Intensity Test.
June 14	Last day for Participants to withdraw from FRVT 2002 without the withdrawal being mentioned in the FRVT 2002 Evaluation Report
June 14	Last day to reach agreement with FRVT Liaison for non-standard hardware requests
June 14	Names of participants announced on FRVT 2002 website

June 19	FRVT 2002 Organizers release test schedule to participants
June 28	4 pm, EDT. Deadline for participants to provide FRVT 2002 Organizers sample similarity files based on the Technical Data Set & 11 June Supplemental. Note that this is the final deadline. Participants are encouraged to submit the sample similarity files as early as possible. Providing these files early will enable participants to fix any mistakes in their files. Resubmissions of similarity files are encouraged prior to the 28 June deadline, but may not be accepted after this date.
July 9	FRVT 2002 testing begins at NSWC Dahlgren



Original Technical Data Set Release

[HEF-0.99.0.zip](#) (Windows)

[HEF-0.99.0.tar.gz](#) (Unix)

Two files above are formatted for different operating systems but contain the same information.

11 June Supplemental

In response to some participants concerns that the example Technical Datasets was misleading, we have assembled a limited set of files that are in exactly the same form as participants can expect to see during the tests.

The supplemental .zip (alternatively .tar) archives contain:

1. the XML files for the target and query sets.
2. the "images" directory containing FERET images in JPEG format.
3. the empty "output" directory where participants output files will be written.

Participants should be able to take SOLELY the contents of this .zip archive and produce valid binary similarity files. The image quality itself may or may not be representative of that included in the tests. This supplemental data, like the HEF Technical Datasets before it, is included ONLY to allow participants to see the input-output relationships.

Notes:

1. This supplemental data does not include video or still-set signature data. Participants should read the FRVT 2002 Implementer's Guide and code accordingly.
2. The code distributed with the Technical Datasets does not include a face detector; instead it uses prior manually located eye coordinates from a text file. This kind of information will NOT be made available to FRVT 2002 participants. The supplemental .zip archive is a real distribution that vendors should be able to

process as is. It does not include face or eye localization information and therefore the PCA implementation that accompanies the original Technical Datasets will not run on it.

[HEF-supplement.zip](#) (Windows)

[HEF-supplement.tar](#) (Unix)

In summary, this zip/tar file contains:

1. 512 FERET images in JPG format (256 targ + 256 query) in a directory "images"
2. An empty directory "output"
3. FERET_target.xml
4. FERET_query.xml
5. A short README.txt file

It does NOT contain:

1. Any nrm files
2. Any eye coordinates
3. Any code

FRVT 2002 Scoring Suite for Participant Review

[FRVT2002-SS-PR1.tar.gz](#) (Unix)



[FRVT 2002 Test Plan](#)

[FRVT 2002 Normalization Specifications](#)

[FRVT 2002 Visitor Information](#)

The FRVT 2002 Visitor Information document contains information regarding your upcoming visit to NSWC Dahlgren Division for the purpose of participating in the FRVT 2002. Notice that some personal information must be gathered for access to the base. This information will only be used for gaining access to the facilities. Deadlines for sending this information are very soon, so please review this attachment as soon as possible.

Appendix D

Forms to Participate in FRVT 2002

FACIAL RECOGNITION VENDOR TEST 2002


 Forms

Status Update
26 November 2002

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

*Facial Recognition
Vendor Test 2000*

*Face Recognition
Vendor Test 2002*

FERET

*Evaluation
Methodology*

Home

Application for Participating in Face Recognition Vendor Test 2002

FRVT 2002 consists of four independent subtests. Each subtest will be performed separately and will have a maximum amount of time for completion. For each subtest, Participants will provide computer(s) and a face recognition system. Participants are required to complete each subtest in the allotted time. Time cannot be saved from one subtest for use in another subtest. Results will be provided on a subtest by subtest basis (scores for participants will only be provided for subtests they fully complete). The four FRVT 2002 subtests are:

Low Computational Intensity. This subtest consists of a relatively small number of images (approximately 1,000). It is designed to test slower systems and to determine if additional allotted time yields superior performance.

Medium Computational Intensity. This subtest is designed to allow for the measurement of performance under a variety of conditions. For example, images will be included with changes in illumination and subject pose. This subtest will require the processing of approximately 10,000 images.

High Computational Intensity. This subtest evaluates a systems ability to perform face recognition tasks using very large data sets. This subtest will require the processing of approximately 120,000 images.

Video. This subtest is designed to evaluate face recognition performance on short video sequences of faces.

Participants will be able to select their own computer equipment following these guidelines:

Low Computational Intensity. One desktop computer

Medium Computational Intensity. One desktop computer

High Computational Intensity. One stand-alone computer network consisting of up to three desktop computers.

Video. One desktop computer

For commercially available systems, desktop computers and networked computers must only contain hardware and software that is commercially available. For mature prototypes, participants must provide a list of system components and identify those that are and are not commercially available. Component and cost information

about the computer(s) each participant is using must be included in their system description document.

To maintain fairness, the Government will adhere to the following procedures for FRVT 2002. Information provided to any single Participant will also be provided to all Participants. The primary and preferred source of communications between Participants and FRVT 2002 government personnel is the frequently asked question (FAQ) page. A government employee has been named as the FRVT Liaison. The FRVT Liaison will be the primary point of contact and will manage communications with Participants. Questions that are outside the purview of the FAQ can be submitted to the FRVT Liaison via email at frvt2002@nist.gov. All e-mails sent to the FRVT Liaison and germane to the FRVT 2002 will be posted on the FAQ page. This policy ensures all Participants are guaranteed equal access to information concerning FRVT 2002. Contact with the FRVT Liaison other than via the FAQ areas, or contact with someone other than the non-FRVT Liaison will only be allowed for extenuating circumstances (for example, delivery of login information for Participant Area or major issues with similarity files).

1. Evaluation Procedure

1.1 The FRVT 2002 technology evaluation commenced 25 April 2002.

1.2 In order to request participation in the FRVT 2002, potential participants must electronically fill out this form ([Application to Participate in FRVT 2002](#)) and the image data set form ([Application for Access to the FRVT 2002 Technical Datasets](#)), sign the forms and send to the locations designated on each form. Upon receipt of both original signed forms by the Government, you will be classified as a "Tentative Participant". The Government must receive both forms by the due date described in the FRVT 2002 Calendar Area.

1.3 The Government will determine which Tentative Participants will be selected as FRVT 2002 Participants. It is the Government's desire to select all Tentative Participants as Participants. However, if demand for participation exceeds the Government's ability to properly evaluate the technology, the Government will be required to select Participants from the Tentative Participant list. If this occurs, the Government will instruct all Tentative Participants to provide the Government a white paper that will be used as the basis for selection as a FRVT 2002 Participant. The Government will provide Tentative Participants with the format and requirements for this white paper, if the need exists. A list of Tentative Participants, and later, Participants will be made available on the FRVT 2002 web site.

1.4 FRVT 2002 Technical Data Sets will be made available in the Participant Area of the FRVT 2002 web site ([Application for Access to the FRVT 2002 Technical Datasets](#)). Participants will be required to perform trials using this test set, at their own facilities, and provide similarity files to the Government. The Government will check these similarity files for compliance. A deadline for delivery of these similarity files will be provided on the web site. Participants who miss this deadline may not be evaluated. Should the similarity files not

comply with standards, the Government will notify Participants and may allow re-submission of new similarity files to ensure compliance. **Very important:** The baseline PCA algorithm included in the FRVT 2002 Technical Datasets requires normalized (eye-coordinates given) images. Images provided during FRVT 2002 will **NOT** be normalized. Systems evaluated will need to automatically process facial imagery without eye coordinates being provided.

1.5 Evaluation dates and locations will be provided at a later date. The Government will attempt to accommodate each Participant on the determination of dates to the maximum extent possible.

1.6 FRVT 2002 Participants will be given a large database of images on the day they arrive for actual testing. The images will be provided on a USB Hard Drive (IDE hard drive, formatted NTFS with one partition, in an external enclosure connected to host computer via USB port). Participants will be required to partition their hard drives into two areas, one for the OS and face recognition algorithms, and the other for working space. Participants will be allowed to copy these images onto their working space hard drive partition only. The participant perform a low-level format on the working space partition of their hard drive. The Government will assure that none of the images, or data derived from the images, are still resident on the test computer. Participants will allow the Government to inspect all disks on the system to verify compliance. This inspection will involve, at a minimum, the government deleting files generated during testing and wiping free space on all disk drives. At the completion of the test the Participants will transfer all the required output files from their system onto the original hard drive via USB. The drive will then be returned to the Government.

1.7 Participants will be required to submit a five-page (maximum) system description document, in electronic form, on the first day of testing. These documents will be included in the final FRVT 2002 report that will be released to the public. Failure to provide this document, in its proper form, may result in not being evaluated in FRVT 2002. This document must adequately address the following topics:

- Overview of the evaluated system.
- Component list for the evaluated system.
- Detailed cost breakdown of the submitted system (commercial vendors only).
- Detail any modifications required to take FRVT 2002.

1.8 Precise details of the FRVT 2002 Test Plan will be made available on the FRVT 2002 web site at a later date.

1.9 FRVT 2002 Participants will not be allowed to comment on their participation in FRVT 2002 until the FRVT 2002 Evaluation Report is released.

1.10 Testing activities will be recorded using video cameras for documentation of the evaluations. Footage from this documentation will not be made available to the public without review and comment from any participant that is named in the video.

2. Release of Evaluation Results

Upon the completion of the evaluations, the Government will combine all results into a final report. The final report will contain, at a minimum, descriptive information concerning FRVT 2002, descriptions of each experiment, evaluation results, and each Participant's five-page system description document. A pre-release version of this report will be made available to Participants. Participants will be invited to provide comments which will be included as an appendix to the FRVT 2002 Evaluation Report. More specific guidance concerning the report and Participant comments will be provided at a later date. Participants are reminded that they will not comment publicly or privately on the pre-release version of the FRVT 2002 Report until it has been released to the public.

3. Additional Information

3.1 Any data obtained during these evaluations, as well as any documentation required by the Government from the participants, becomes the property of the Government. Participants will not possess a proprietary interest in the data and/or submitted documentation.

3.2 By signing and submitting the FRVT 2002 forms, Tentative Participants and Participants attest that they will not file any claim against FRVT 2002 Sponsors, Supporters, or other agency of the U.S. Government, or otherwise seek compensation for any equipment, materials, supplies, information, travel, labor and/or other participant provided services.

3.3 The Government is not bound or obligated to follow any recommendations of the Participant. The United States Government, or any individual agency, is not bound, nor is it obligated, in any way to give any special consideration to FRVT 2002 Participants on future contracts.

3.4 If a Participant decides to use results of these evaluations in any way for their own purpose, it must be accompanied by the following phrase: "Results shown from the Face Recognition Vendor Test 2002 do not constitute endorsement of any particular system by the Government." It must also be accompanied by an internet hyperlink to the final report on the FRVT 2002 web site.

3.5 With the signing of this form, Tentative Participants and Participants are authorizing their company/organization to participate in the Face Recognition Vendor Test 2002 according to the provisions in this form and later documents supplied by the Government.

3.6 With the signing of this form, Tentative Participants and Participants realize that they must send signed original copies of this form as well as "Application for Access to the FRVT 2002 Technical Datasets" to the locations documented on each form.

3.7 With the signing of this form, Tentative Participants and Participants realize that any test details and/or modifications that are provided in the Participant Area of the FRVT 2002 web site

supersede the information on this form.

3.8 With the signing of this form, Tentative Participants and Participants realize that they can withdraw from the FRVT 2002 evaluations at any time up to the day prior to the start of their testing. Tentative Participant and Participant withdrawals will be documented in the FRVT 2002 Evaluation Report.

4. Request to Participate

4.1 Participant Information

Company Name:

Responsible Party:

(Prefix/First Name -- Last Name/Suffix)

**Responsible Party
Title:**

Alternate/Witness:

(Prefix/First Name -- Last Name/Suffix)

Mailing Address:

**City/St/Postal
Code:**

Phone Number:

Fax Number:

Email Address:

Web Site Address:

FRVT 2002 POC:

(Prefix/First Name -- Last Name/Suffix)

POC Phone:

POC Email:

USA

You will receive a confirmation note that has your ID and password for access to the participant area of the web site upon receipt of both required forms.

FACIAL RECOGNITION VENDOR TEST 2002



Application for Access to the FRVT 2002 Technical Datasets

The FRVT 2002 Technical Data Sets consist of

- FRVT 2002 evaluation data sets.

- FRVT 2002 training set.

- Software provided with the FRVT 2002 training set.

FRVT 2002 evaluation data sets are used in the evaluation and will be made available at the time a system is evaluated. FRVT 2002 Training Data and Supporting Software are supplied to Participants to assist in preparing for the FRVT 2002. These two components will be made available to Participants in the FRVT 2002 Participant Area. To receive access to the FRVT 2002 Technical Data Sets, participants must sign the agreement at the bottom of this page. Failure to observe the restrictions in this document will result in access being denied for future releases of the image database and removal from FRVT 2002. Offenders may also be subject to civil damages.

1. Restrictions for the FRVT 2002 Training Data and Supporting Software

The FRVT 2002 training data and supporting software includes:

- A subset of normalized images from the FERET database

- A baseline PCA face recognition algorithm

- Scoring code to generate receiver operator characteristics (ROC) and cumulative match characteristics (CMC)

- Instructions on how to use the software tools

- Documentation on similarity file output requirements for your algorithms

- Instructions for submission of sample similarity files for Government review

Very important: The baseline PCA algorithm included in the FRVT 2002 Technical Datasets requires normalized (eye-coordinates given) images. Images provided during FRVT 2002 **NOT** be normalized. Systems evaluated will need to automatically process facial imagery without eye coordinates being provided.

The Participant agrees to the following restrictions on the FRVT 2002 Training data and supporting software:

1. The FRVT 2002 Technical Datasets will not be further distributed, published, copied, or further disseminated in any way or form whatsoever, whether for profit or not. This includes further distributing, copying or disseminating to a different facility or organizational unit in the requesting university, organization, or company.
2. All requests for copies of the FRVT 2002 Technical Datasets will be forwarded to the Agent at the National Institute of Standards and Technology (NIST).

*Status Update
26 November 2002*

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

*Facial Recognition
Vendor Test 2000*

*Face Recognition
Vendor Test 2002*

FERET

*Evaluation
Methodology*

Home

3. All publication or other release of any part of the Technical Datasets, or data derived from using the Technical Datasets, in reports, papers or other documents must first be approved in writing. Approval may be sought by writing from the NIST technical agent.

2. Restrictions for the FRVT 2002 Evaluation Data Sets

The Participant agrees to the following restrictions on the FRVT 2002 Technology Evaluation Test Set:

1. The database will not be used for any purpose except for participation in FRVT 2002 Technology Evaluations.
2. The database will only be copied to the Participant's system in the working space partition of their hard drive. Participants will perform a low-level format of this partition at completion of their testing.
3. If required, a separate laptop may be used to facilitate transfer to/from USB hard drive and face recognition system.
4. All data generated from this database during FRVT 2002 will be deleted from all Participant systems by the Government at the conclusion of the evaluation.

3. Request for Access to Databases

3.1 Participant Information

Company Name:

Responsible Party:

(Prefix/First Name -- Last Name/Suffix)

**Responsible Party
Title:**

Alternate/Witness:

(Prefix/First Name -- Last Name/Suffix)

Mailing Address:

**City/St/Postal
Code:**

Phone Number:

Fax Number:

Email Address:

Web Site Address:

****** Note: The application deadline of 10 May 2002 has past and we are no longer accepting applications for FRVT 2002.**

3.2 Signatures

With my signature, I authorize my company/organization to use the FRVT 2002 Technical Datasets in association with the Facial Recognition Vendor Test 2002, and promise to do so according to the rules and limitations listed on this form.

3.3 Application Fulfillment

To Complete the application process:

1. Print this page(s) from your browser
2. Obtain the appropriate signatures
3. Mail the completed form to:

FRVT 2002 Liaison
Code B06
1470 Room 1101
17320 Dahlgren Road
Dahlgren, VA 22448
USA

You will receive a confirmation note that has your ID and password for access to the Participant Area of the web site upon receipt of both required forms.

This page intentionally blank.

Appendix E

Frequently Asked Questions

FACE RECOGNITION VENDOR TEST 2002



*Status Update
26 November 2002*

How to Participate

Participants

Sponsors

FAQ

Calendar

Participant Area

*Facial Recognition
Vendor Test 2000*

*Face Recognition
Vendor Test 2002*

FERET

*Evaluation
Methodology*

Home

Submit an unanswered question

1. On the FRVT 2002 Application to Participate Form, what is the difference between 'Responsible Party', 'Alternate/Witness', and 'FRVT 2002 POC'?
2. For the low/medium/high computational intensity tests, can we use a desktop computer with multiple-processors and count that as one desktop computer?
3. For the Video Subtest, what is the video data format (framerate , file format)?
4. Is there a fee to participate in FRVT 2002?
5. Is there a penalty or fine for withdrawing from FRVT 2002?
6. For the first three subtests (low, medium and high computational subtests), can participants use the exact same code that was used for the Recognition Performance Test of FRVT2000?
7. The calendar lists June 21st as the date testing begins. Does that mean systems and software must be ready for all four subtests by that date? Or, will each vendor perform the first subtest in turn, then the second, and so forth?
8. For the video subtest, how many computers can be used? Can co-processing devices be used in addition to the PC(s)?
9. For the video subtest, do the video images have to be analyzed in real-time, or can we take as long as we like to analyze each frame?
10. For the video subtest, how many seconds of video are available in each video sequence?
11. For the video subtest, what is the total number of images?
12. For the video subtest, can the video watch list be hand-aligned (human processed)?
13. For the low, medium, and high computational subtests, are the numbers provided (1K, 10K and 120K) the number of images that have to be processed? Or are these numbers the total number of comparisons to be made?
14. For each of the subtests, how many images per person does the gallery include?
15. What metrics will be used to define performance for the

static and video subtests?

16. Will infrared systems be evaluated in FRVT 2002?
17. What is the purpose and requirements for the white paper?
18. Can non-US companies participate in FRVT 2002?
19. What are the time periods allowed for each test?
20. Why don't you 'level the playing field' by making all Participants perform on the same hardware configuration?
21. Is it possible to submit more than one facial recognition system for each sub test?
22. Will the actual time a participant uses for a subtest appear in the evaluation report?
23. Is it possible for the public to witness any of the face recognition testing that is taking place?
24. Could a detailed definition of the face recognition tests, possibly what is contained in the Test Plan, be sent to me?
25. How will the public be notified that the test results from FRVT 2002 are available?
26. Do you know the exact date the FRVT 2002 results will be published online?
27. Will the data released publicly from FRVT 2002 be selective due to homeland security concerns?
28. Would it be possible to obtain, purchase, or rent the FRVT 2002 dataset(s)?

1. On the FRVT 2002 Application to Participate Form, what is the difference between 'Responsible Party', 'Alternate/Witness', and 'FRVT 2002 POC'?

Responsible Party: An individual at a potential participant's company that has the authority to request participation in FRVT 2002 on behalf of that company.

Alternate/Witness: An individual who is witnessing the company application and signature of the responsible party.

FRVT 2002 POC (Point of Contact): This individual will be the lead person on-site during the evaluations and will be the company's designated individual for contact with the FRVT Evaluation Team for all aspects of the evaluation.

2. For the low/medium/high computational intensity tests, can we use a desktop computer with multiple-processors and count that as one desktop

computer?

For FRVT 2002, a "single desktop computer" is defined to be a computer with case dimensions and power requirements similar to commercially-available computers for an individual's use. Multiple processors are allowed inside that case and will be counted as a single desktop computer. Other configurations approximating a single desktop may be allowed on a case-by-case basis, but must be coordinated with the FRVT 2002 Liaison prior to participation. A detailed component list, complete with pricing information, must be included in the participant's write-up that explains the system the participant used for FRVT 2002.

3. For the Video Subtest, what is the video data format (framerate , file format)?

The video will be supplied as a sequence of separate standard JPEG files. The frame rate of the camera was 29.97 frames per second. The frames were obtained by transcoding from the original DV format.

4. Is there a fee to participate in FRVT 2002?

There is no participation fee for FRVT 2002. Participants will be required to provide hardware and software necessary for evaluating their system, as well as on-site representation during the evaluation.

5. Is there a penalty or fine for withdrawing from FRVT 2002?

Tentative Participants and Participants can withdraw from the FRVT 2002 evaluations at any time up to the day prior to the start of their testing. Tentative Participant and Participant withdrawals will be documented in the FRVT 2002 Evaluation Report. There are no other penalties or fines.

6. For the first three subtests (low, medium and high computational subtests), can participants use the exact same code that was used for the Recognition Performance Test of FRVT2000?

No. The required input and output files have different formats than those used in FRVT 2000. In the FRVT 2002, the input and output files will be in specified XML formats. The formats will be documented when the "Technical Datasets" are released to participating vendors. That distribution will include working code that demonstrates the use of XML.

7. The calendar lists June 21st as the date testing begins. Does that mean systems and software must be ready for all four subtests by that date? Or, will each vendor perform the first subtest in turn, then the second, and so forth?

A test schedule will be developed and published at the conclusion of the application process. It is our intention to have participants perform different subtests simultaneously on different machines. It may be possible, or desirable for some vendors to perform some subtests sequentially. For example, low and medium tests on the same machine.

8. For the video subtest, how many computers can be used? Can co-processing devices be used in addition to the PC(s)?

A single desktop computer is allowed for the video subtest. For FRVT 2002, a "single desktop computer" is defined to be a computer with case dimensions and power requirements similar to commercially-available computers for an individual's use. Multiple processors are allowed inside that case and will be counted as a single desktop computer. Other configurations approximating a single desktop may be allowed on a case-by-case basis, but must be coordinated with the FRVT 2002 Liaison prior to participation. A detailed component list, complete with pricing information, must be included in the participant's write-up that explains the system the participant used for FRVT 2002.

9. For the video subtest, do the video images have to be analyzed in real-time, or can we take as long as we like to analyze each frame?

Individual video sequences may be analyzed over an extended period of time. Real-time processing is not required as long as the video subtest is completed in the allocated time. This subtest is intended to determine if superior performance can be obtained from video sequences rather than still images.

10. For the video subtest, how many seconds of video are available in each video sequence?

The video sequences have a mean duration of approximately 3 seconds. They were obtained from cameras operating at approximately 30 frames per second.

11. For the video subtest, what is the total number of images?

The precise numbers have not yet been determined. There will likely be fewer than two thousand sequences each being comprised of approximately 90 frames. Note that because subjects' faces are visible in all frames a vendor may use just one frame from a sequence in order to attempt recognition.

12. For the video subtest, can the video watch list be hand-aligned (human processed)?

No. All systems for all sub-tests must be fully automated. All image enrollments and subsequent recognition computations for all the FRVT 2002 tests must be performed automatically, without human inspection or intervention.

13. For the low, medium, and high computational subtests, are the numbers provided (1K, 10K and 120K) the number of images that have to be processed? Or are these numbers the total number of comparisons to be made?

There will be approximately 1K, 10K, and 120K images that need to be processed. This implies that there could be approximately 1M, 100M and 14.4 Billion required comparisons between images. However, a subtest may be designed so that fewer comparisons are required. A complete discussion of the FRVT 2002 computational requirements will be included with the release of the Technical Datasets.

14. For each of the subtests, how many images per person does the gallery include?

The image sets contain between 1 and 20 images per person. The test is structured such that the identities of the subjects are not given to participants.

15. What metrics will be used to define performance for the static and video subtests?

The FRVT 2002 report will give a richer set of performance numbers than those quoted in FRVT 2000. The FRVT 2002 will report figures appropriate to assessing performance on tasks that include identification, verification, and watch-list operations. In addition, the effect of gallery size on performance will be included. Other analyses may be performed and reported. These include measures of statistical significance and estimates of variance of performance statistics.

16. Will infrared systems be evaluated in FRVT 2002?

Infrared imagery will not be included in the FRVT

2002.

17. What is the purpose and requirements for the white paper?

If demand for participation in FRVT 2002 exceeds the ability of the Government to properly test all participant face recognition systems, the Government may ask potential participants to submit a technical paper describing their system and its capabilities. The Government will review all submitted 'white papers' and use them as a basis for selecting final participants in FRVT 2002. Format and requirements for white papers will be provided should the need arise.

18. Can non-US companies participate in FRVT 2002?

Yes.

19. What are the time periods allowed for each test?

We have established the following tentative time limits:

Low computational: 1 Week

Medium Computational: 1 Week

High Computational: 2 Weeks

Video: 1 Week

20. Why don't you 'level the playing field' by making all Participants perform on the same hardware configuration?

FRVT 2002's objective is to "...provide independent government evaluations of commercially available and mature prototype face recognition systems." Vendor/Participants are generally free to have their systems (software and hardware) evaluated as they are sold commercially. This may include non-desktop based systems on a case-by-case basis as some vendors sell systems with specific hardware/software integrations. FRVT 2002 will publish information concerning each Participant's system in its final report to support FRVT 2002 Sponsor requirements. This policy is in the best interests of both participants and FRVT 2002 Sponsors.

21. Is it possible to submit more than one facial recognition system for each sub test?

Multiple face recognition systems from one participant may be allowed on a case-by-case basis, but must be coordinated/approved with the FRVT 2002 Liaison. The deadline for approval is 20 May 2002.

22. Will the actual time a participant uses for a subtest appear in the evaluation report?

No. The report will only document if the participant finished the subtest or not.

23. Is it possible for the public to witness any of the face recognition testing that is taking place?

No. To ensure fairness for all participants, and thus a fair evaluation, no public visits to the test facility are permitted.

24. Could a detailed definition of the face recognition tests, possibly what is contained in the Test Plan, be sent to me?

No. A thorough description of the test procedures will be provided in the FRVT 2002 Evaluation Report.

25. How will the public be notified that the test results from FRVT 2002 are available?

Results will be made available on <http://www.frvt.org>. A notification message will also be sent to the Biometrics Consortium listserv.

26. Do you know the exact date the FRVT 2002 results will be published online?

It is not possible to forecast a specific date at this time.

27. Will the data released publicly from FRVT 2002 be selective due to homeland security concerns?

FRVT 2002 is a technology evaluation, so it is not application-specific. All results will be published in the FRVT 2002 Evaluation Report

28. Would it be possible to obtain, purchase, or rent the FRVT 2002 dataset(s)?

The images contained in the FRVT 2002 datasets are for use in the FRVT 2002 evaluation and are not currently available for other purposes. At a later date, some of the images in the MCIInt may be released. Images in the HCIInt will not be released. If you desire face image data for your own development or baseline evaluation efforts, the entire FERET database is currently available. Please see

<http://www.nist.gov/humanid/feret>.



[Submit an unanswered question](#)

1. At what location will the tests be held?
2. What is the reason for not using a video file format for the video subtest?
3. What will be the range of inter-eye distances, in terms of pixels, for each subtest?
4. Will there be experiments within subtests? If so, what will they be?
5. Do any of the tests involve scenes that are cluttered, busy, or contain moving background objects?
6. Is the score format similar to FRVT2000? How about the Large DB Test for 120,000 images?
7. Will the USB protocol (used to transfer data to and from vendor machines) be 1 or 2?
8. Is JPEG only image format used for FRVT 2002? What is the percentage of color and grayscale images for each test?
9. What kind of USB drive will be used? Let us know exact specification so that we can buy one and test it.
10. Will all participants be required to arrive on July 9th, or will some participants be scheduled to arrive later to stagger the testing?
11. Are we required to send a white paper based on "Computational Requirements for FRVT 2002 High Computational Intensity Test" from the technical dataset if we are not participating in the High Computational Intensity Test?
12. The FRVT test begins with the organizers providing a dataset to the participants. Will there be any other forms that need to be signed to receive this information?
13. Where do Participants send the sample similarity files based on the sample images in the Technical Dataset?
14. Is there a typo on page 5 of "The FRVT Implementer's Guide"? We read "4N + 24 bytes". Shouldn't this be "4N + 28 bytes"?
15. Are the variations (lighting, facial expression, etc) in the Technical Data Set representative of the type of variations that the vendor software will be tested on?
16. For the Medium Computational Intensity Subtests, will the query signature set contain still set or video type images?
17. Your e-mail message dated June 6 announces the test reorganization and the fact that the new Medium test will contain 7500

signatures. Where are we supposed to get this new data/subset for the Medium test?

18. What can we expect for backgrounds in the images in FRVT 2002? (For example, complex outdoor B/D, Etc.)
19. Regarding "The FRVT 2002 Implementer's Guide": (1) It is not clear which xml is the target and which one is the query. (2) Also, the actual content of the xml files is confusing. The <file /> entries (ex.: <file name="00003fa010_940307.nrm"/>) point to inexistant files of unknown format (p. 3 of "The FRVT 2002 Implementer's Guide" states that all pictures in the FRVT2002 are JPEG).
20. What is the exact number of images for each subtest? (We need exact number in order to calculate processing time for white paper.)
21. The Document "The FRVT2002 Implementer's Guide" lists 3 kinds of signatures to be used in the test (Fig 2.1). How does each signature relate to each subtest?
22. Signature #2 ("The FRVT2002 Implementer's Guide" figure 2.1) suggests a stereo camera system. Will stereo data be used in any subtest?
23. Why are your HEF_TD library just suitable for debug version?
24. Do we need to enroll total of N images for query and target set in each test?
25. Is it possible for a participant to start the test run and then leave the testing? FRVT Organizers would need to copy results to disk and then ship the computer back.
26. Will FRVT Organizers be providing UPS for participants use during testing?
27. We downloaded everything for generating the sample similarity files. In one of your documents it says that we need to produce some XML documents from the results. However, from the stuff we downloaded it seems as if the result files are just binary files. Did we misunderstand something? Do we have to deal with XML files at all?
28. Is the query and target set be the exactly same or different?
29. Why have you provided the technical dataset in a format and structure different from the procedural requirements of the test? Why not provide a technical dataset that is identical to what we will receive? Am I really supposed to convert all of the PGM format images to JPG and modify the target and query XML files accordingly? This seems to be required if I am to create a similarity file based on the target and query XML documents provided (FERET-256-1.sigs.xml, FERET-256-2.sigs.xml)
30. Am I really supposed to convert all of the PGM format images to JPG and modify the target and query XML files accordingly? This seems to be required if I am to create a similarity file based on the target and query XML documents provided (FERET-256-1.sigs.xml, FERET-256-2.sigs.xml)
31. Can FRVT 2002 Organizers provide monitors for participants to use during the evaluations.
32. What is the minimum and the maximum relative size of a face in the image? Can the image be of factor 1000 wider than the eye distance, for example?
33. Do all images show one face only, or do you have images with several faces, e.g. in the video test?
34. What will be the total size of images (in GB) for 121,000 images?
35. How does a participant withdraw from participating in the FRVT 2002?
36. Will we need to modify the target.xml and query.xml files to get the schema resources from a local drive during FRVT 2002 tests?
37. Please elaborate on how the FRVT 2002 will evaluate 3D, infrared, or other novel face recognition systems.
38. Is the test site available 24 hours a day during the test period including weekends?
39. I am still having trouble understanding the use of XML in FRVT 2002. Can you help me?
40. How many target set and query set will be given for each subtest?

41. Is it necessary that a participant's POC be on site, or could it be a different employee of the participant or some other person engaged by the participant?
42. Each output file consists of similarity data. Must this data be sorted in a specific format?
43. Can you supply some sample data in color, or indicate where a database of color face image data which is representative of test data may be obtained?
44. What is the minimum and maximum for the width and height (in pixels) of the imagery in each test?
45. What is the scaling for similarity and distance values (0-10, 0-100....)?
46. Medium Computational Intensity Test will be administered with two targets and query sets - one for still images, one for video images. Are you going to provide us with a disk containing tree structure (Chapter 3 of Implementers guide) for each set?
47. Will the subtests be run sequentially or in parallel? Can we use the same machine for both subtests?
48. You have said that the query & target set of HCI nT are identical. Will FRVT 2002 organizers provide two XML files or only one XML file?
49. Are we allowed to bring other people, such as a key technician or software engineer, in addition to the Technical Point of Contact during the 9 to 5 period when we have access to the system?
50. FAQ # 3 states that the number of pixels between the eyes will be 30 - 500 pixels. How does this relate to FAQ # 44?
51. The technical dataset contains 3 video signatures named "video 00-02". We do not find any video sigs in the FRVT supplemental directory, so we assume that we do not have to process any video sigs for the sample similarity files we will submit for the FRVT evaluation procedure. Is this correct?
52. If the test program crashes, may our representative or representatives restart the program?
53. Can we place phone calls from the test facility to our offices? Are cell phones allowed? Are laptop PC's allowed?
54. Will the test area in Dahlgren have 20-amp AC outlets?
55. Do we need to get clearance from NSWC Dahlgren in advance for anyone that we bring, both the Technical Point of Contact and others? Are there special issues, especially long delays, for clearance for non-US citizens?
56. Can you provide an example of the 5-page document we are required to submit?
57. Are face locations nearly center in the image or could the face be located on a side?
58. What is the approximate breakdown of signatures in the Medium Computational Intensity Test?
59. How many still images are in a still set signature?
60. Is any remote access to the system such as TELNET or FTP available? Can we monitor the test remotely, e.g. over phone lines or the Internet? If the test program crashes, may we restart another debugged version of the test program?
61. Participants have been given 264 hours for testing. A participant finishes early, will they be permitted to leave early?
62. In the event that the voice-only telephone is not available and cell phone calls cannot be placed from within the building due to the metal covering mentioned in the FRVT 2002 Test Plan or other reasons, may participants leave the building to place cell phone calls from outside for technical support or other reasons?
63. May participants monitor the system at any time during the hour, or is it a specified time such as the start of the hour (9:00, 10:00, 11:00,...)? How much time do we have for the hourly monitoring, such as a maximum of five minutes?
64. May participants examine the contents of files generated by the test program such as similarity files or other intermediate status files written by the application?
65. May participants view a console output from the program, messages such as "processing probe 64,123...."?
66. What is the extension for the similarity file? .SIM? .BIN?
67. Will you supply the power and data cables for the IDE drives in USB2.0 enclosures containing your image database?

68. We will bring only Linux based systems for the test. How shall we do the virus scan?
69. How often does power go out at NSWC Dahlgren?
70. What are the deadlines for submitting a normalization function?
71. Would a windows based library be acceptable as well for the normalization function?
72. What is the number of signatures and images for the tests in FRVT 2002?
73. Are each signatures type and modality name in XML file? What about directory structure?
74. After completing the medium computational intensity test, may we add that computer to the system that is running the high computational intensity test?
75. We are planning to bring more than two people for FRVT2002 because of our functional role and contractual schedule. If you are very strict about the number of people (only two people will be allowed at the test site) for the testing, we want to be able to switch people during the test still maintain two people at the test site (including POC).
76. What is the directory structure and naming format for files we generate in the MCInt?
77. In the Outline and Overview for Use of Normalization Functions in FRVT 2002 we are told that a possible normalization can be based, in part, on similarity scores for all pairs of gallery images in G. Does this imply that G consists of a subset of the target set which is also represented as probe sets in the query set Q? Will there be a way to identify these subsets other than by brute force searches of the XML target and probe set definitions?
78. For the High Computational Intensity test, are we allowed to copy the FRVT data (pictures, xml,...) on the three computers? At the end, are we allowed to copy back to the USB drive from each of the three computers or do we need to centralize all similarity files on one computer before copying back?

 Participant Area FAQ closed during evaluation period from 2 July through 13 August.

79. When is the normalization function due?
80. There seem to be time limits imposed on the normalization functions. And if a F2 normalization is taking too long, the testers will revert to using the F1 normalization function. How much time is allotted per gallery siz?. Meaning, for a gallery size of X how long should F2 or the F1 function take before it is deemed too long?
81. There are several math functions that our normalization function uses such as sqrt and exp. Is it okay to expect that whatever application is linking in our .o file will also link in the standard c/c++ code for these functions (either libm.a or libstdc++.so seems to do it fine on our machines here)?
82. Will the normalization functions be used on the high computational intensity test as well as the medium computational intensity test? If so, will there be a way for the functions being called to know in advance whether the scores being passed in are coming from the high versus the medium computational intensity test?
83. What is the minimum size of the gallery being passed to the normalization functions?
84. In FRVT 2002 Normalization Specifications: What is the meaning of "identification, verification, and watch list tasks"?
85. Is it planned to send the participants a pre-release of the FRVT 2002 Evaluation Report, allowing them to comment on it (as was done for the FRVT 2000 Evaluation Report)?

1. At what location will the tests be held?

FRVT 2002 will be administered at NSWC Dahlgren in Dahlgren, Va. Please see http://www.nswc.navy.mil/visitor_info.html for general visitor information. Specific visitor procedures and requirements will be given at a later date.

2. What is the reason for not using a video file format for the video subtest?

There are a myriad of different methods/file formats to share video data. Video data, however, is a sequence of still images. Since the rest of the subtests use JPG, the evaluation team decided to also use a sequence of JPG images for the video test. Also, this evaluation addresses face recognition, not video signal handling.

3. What will be the range of inter-eye distances, in terms of pixels, for each subtest?

The distance between the eyes in all tests will be at least 30 pixels and not more than 500.

4. Will there be experiments within subtests? If so, what will they be?

As in FRVT 2000 and the [FERET evaluations](#), virtual galleries and probe sets will be constructed from the target and query sets. See "[The FERET Evaluation Methodology for Face-recognition Algorithms](#)" for a detailed description of these sets. The virtual gallery and probe sets will be designed to analyze the effects of non-frontal images, time delay between photographs, illumination changes, outdoor lighting and other factors.

5. Do any of the tests involve scenes that are cluttered, busy, or contain moving background objects?

No.

6. Is the score format similar to FRVT2000? How about the Large DB Test for 120,000 images?

This issue is addressed in the Technical Datasets.

7. Will the USB protocol (used to transfer data to and from vendor machines) be 1 or 2?

The USB model supports both USB 1 and USB 2. We recommend using USB 2 as it is much faster.

8. Is JPEG only image format used for FRVT 2002? What is the percentage of color and grayscale images for each test?

JPEG is the only image format used for FRVT 2002. The vast majority of images will be color.

9. What kind of USB drive will be used? Let us know exact specification so that we can buy one and test it.

It's a standard IDE drive placed inside a USB-2.0 enabled enclosure. An example of such a product is http://www.adstech.com/products/USB_2_Drivekit/intro/USB2DriveKit.asp?pid=USBX-804. This may or may not be the actual device used. This commercial equipment was identified to adequately describe the subject matter above. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the equipment is necessarily the best available for the purpose.

10. Will all participants be required to arrive on July 9th, or will some participants be scheduled to arrive later to stagger the testing?

Test times will be staggered.

11. Are we required to send a white paper based on "Computational Requirements for FRVT 2002 High Computational Intensity Test" from the technical dataset if we are not participating in the High Computational Intensity Test?

No. We recommend that you still perform the assessment for the Medium Computational Intensity Subtest to ensure that you will be able to complete the test in the allotted time.

12. The FRVT test begins with the organizers providing a dataset to the participants. Will there be any other forms that need to be signed to receive this information?

No, this was taken care of by submitting the "Application for Access to the FRVT 2002 Technical Datasets" that we have already received.

13. Where do Participants send the sample similarity files based on the sample images in the Technical Dataset?

Zip them and attach it in an email to frvt2002@nist.gov

14. Is there a typo on page 5 of "The FRVT Implementer's Guide"? We read "4N + 24 bytes". Shouldn't this be "4N + 28 bytes"?

Yes. The correct phrase should be "4N + 28"

15. Are the variations (lighting, facial expression, etc) in the Technical Data Set representative of the type of variations that the vendor software will be tested on?

The FERET images in the Technical Datasets are not necessarily representative of the imagery to be used in FRVT 2002. These images were included to allow replication of the required functionality.

16. For the Medium Computational Intensity Subtests, will the query signature_set contain still_set or video type images?

In the Medium Computational Intensity Subtests, both target and query sets may contain both still and video imagery.

17. Your e-mail message dated June 6 announces the test reorganization and the fact that the new Medium test will contain 7500 signatures. Where are we supposed to get this new data/subset for the Medium test?

The Medium computational intensity test will contain approximately 7,500 signatures. The format for the signatures is contained in the documentation distributed in the Technical Data Sets.

18. What can we expect for backgrounds in the images in FRVT 2002? (For example, complex outdoor B/D, Etc.)

The key feature of the imagery is that each image consists of a single face that is the largest object in the image. Background varies from a relatively uniform background to images taken outdoors. The distance between the eyes in all tests will be at least 30 pixels and not more than 500.

19. Regarding "The FRVT 2002 Implementer's Guide": (1) It is not clear which xml is the target and which one is the query. (2) Also, the actual content of the xml files is confusing. The <file /> entries (ex.: <file

`name="00003fa010_940307.nrm"/>`) point to inexistant files of unknown format (p. 3 of "The FRVT 2002 Implementer's Guide" states that all pictures in the FRVT2002 are JPEG).

A1 - The two signature-set files in the distribution, FERET-256-1.sigs.xml and FERET-256-2.sigs.xml, can interchangeably serve as target or query sets.

A2 - The files specified in this format do not exist in the distribution as they're created as an intermediate step of our implementation. Once they are created (by csu-normalizer) the files can be used. The .nrm format is an internal binary format for our geometrically and photometrically normalized images; they're most likely not usable with any third party viewers.

FRVT 2002 participants should be able to read JPEG files specified in the general form: . The previous line should not be construed to indicate that the directory depth is 2: it may be shallower or deeper as needed.

Participants should also be able to write the required binary files in the form . They can assume that the required paths will already exist, so that any "mkdir"-like operation will not be necessary. The Technical Datasets signature-set files do not include paths, the result being that the output files would be written into the working directory.

20. What is the exact number of images for each subtest? (We need exact number in order to calculate processing time for white paper.)

The number of images in the High Computational Intensity Subtest is exactly 121590. The number of images in the Medium Computational Intensity Subtest has not been finalized.

21. The Document "The FRVT2002 Implementer's Guide" lists 3 kinds of signatures to be used in the test (Fig 2.1). How does each signature relate to each subtest?

For the High Computational Intensity Test only still signatures (and thus single still images) will be used. For the Medium Computational Intensity Test mixed still, still-set and video signatures will be used.

22. Signature #2 ("The FRVT2002 Implementer's Guide" figure 2.1) suggests a stereo camera system. Will stereo data be used in any subtest?

The still-set signature given as an example in figure 2.1 of the implementers guide lists two images. Their names should NOT be construed to indicate that they were taken simultaneously or under the same lighting conditions or even on the same day; there are just two images of one person. No stereoscopic imagery will be used in FRVT 2002.

23. Why are your HEF_TD library just suitable for debug version?

To aid development, we recommend using the "debugging" version of Xerces and the HEF libraries. Help with the "release" version of the libraries will not be supported. Note that this refers primarily to the Windows version of the libraries.

24. Do we need to enroll total of N images for query and target set in each test?

The FRVT 2002 Implementer's Guide describes the three kinds of signature that may occur in the target and query sets: a "still" signature contains only one image, but the "still-set" signature and the "video" signature will contain one or more images.

In the High Computational Intensity test every signature is a "still" signature. Therefore the number of images is identical to the number of signatures. The Medium Computational Intensity test, however, contains all types of signature, and therefore there are more images than signatures. For the MCIInt, therefore, it is not strictly necessary to enroll or process all the images, although the expectation is that better performance would be obtained by doing so. It is only necessary to report for all signatures.

The FRVT 2002 Organizers do not make any recommendations on which images should be considered, or on how they should be processed.

25. Is it possible for a participant to start the test run and then leave the testing? FRVT Organizers would need to copy results to disk and then ship the computer back.

Participants must be on site for the setup and check out procedures. The check out includes (but not limited to) transferring the similarity files from the participant system to the FRVT 2002 test agent and the inspection of the participant's hardware. Participants may leave while the test is being conducted. Participants are, however, responsible for insuring their systems complete all tests. The FRVT 2002 organizers are not responsible for informing participants that their system has crashed or ceased to work for any reason. This includes (but is not limited to) power outages.

26. Will FRVT Organizers be providing UPS for participants use during testing?

No. Participants are encouraged to bring their own UPS.

27. We downloaded everything for generating the sample similarity files. In one of your documents it says that we need to produce some XML documents from the results. However, from the stuff we downloaded it seems as if the result files are just binary files. Did we misunderstand something? Do we have to deal with XML files at all?

FRVT 2002 participants will need to be able to read signature set XML documents and output binary result files.

28. Is the query and target set be the exactly same or different?

For the High Test the target and query sets are identical.

For the Medium Computational Intensity test participants SHOULD assume the sets are DIFFERENT.

29. Why have you provided the technical dataset in a format and structure different from the procedural requirements of the test? Why not provide a technical dataset that is identical to what we will receive? Am I really supposed to convert all of the PGM format images to JPG and modify the target and query XML files accordingly? This seems to be required if I am to create a similarity file based on the target and query XML documents provided (FERET-256-1.sigs.xml, FERET-256-2.sigs.xml)

The purpose of the technical dataset is to provide an example system that can read XML signature sets and output binary similarity files. It is not the goal of the technical dataset to provide an example standalone face recognition system that is FRVT 2002 "ready." Note that successful compilation and execution of the Technical Dataset is not a requirement for participation in FRVT 2002, but was provided to enable participants to gain an end-to-end feel for how the evaluations will be run.

Other than the names, quantity, and type of files, the XML that vendors will receive during the FRVT 2002 are identical in structure to what is provided in the technical dataset (i.e., they will validate with respect to one of the provided signature set schemas).

A Technical Dataset Supplemental is now available that contains data in the same format as you will be expected to utilize for the FRVT 2002 Evaluation.

30. Am I really supposed to convert all of the PGM format images to JPG and modify the target and query XML files accordingly? This seems to be required if I am to create a similarity file based on the target and query XML documents provided (FERET-256-1.sigs.xml, FERET-256-2.sigs.xml)

During FRVT 2002, the images provided will be in "JPEG" format. If, for internal testing purposes, you wish to use the

provided imagery and XML files for your own system, then converting the images to a format understood by your recognition system may be required.

31. Can FRVT 2002 Organizers provide monitors for participants to use during the evaluations.

We will not be able to accommodate this request.

32. What is the minimum and the maximum relative size of a face in the image? Can the image be of factor 1000 wider than the eye distance, for example?

The key feature of the imagery is that each image consists of a single face that is the largest object in the image.

33. Do all images show one face only, or do you have images with several faces, e.g. in the video test?

All images contain only one face.

34. What will be the total size of images (in GB) for 121,000 images?

The total amount is less than 5GB.

35. How does a participant withdraw from participating in the FRVT 2002?

The responsible party (as designated on the applicant's "Request to Participate in FRVT 2002" form must send an e-mail that explicitly states their desire to withdraw. This e-mail must be sent to frvt2002@nist.gov.

36. Please elaborate on how the FRVT 2002 will evaluate 3D, infrared, or other novel face recognition systems.

The FRVT 2002 is designed to evaluate the performance of face recognition from still and video still imagery in the visible part of the spectrum. The FRVT 2002 will not evaluate performance of face recognition signatures gathered by other means. This includes three-dimensional signatures and infrared imagery. Since other types of facial signatures will not be provided, participants that require other types of signatures are encouraged to withdraw from the FRVT 2002 by 14 June 2002. To withdraw, the responsible party (as designated on the applicant's "Request to Participate in FRVT

2002" form) must send an e-mail that explicitly states their desire to withdraw. This e-mail must be sent to frvt2002@nist.gov.

37. Will we need to modify the target.xml and query.xml files to get the schema resources from a local drive during FRVT 2002 tests?

Vendors will not have Internet access during FRVT 2002. If validating the given signature sets is desired, it should be done from the provided schema files.

38. Is the test site available 24 hours a day during the test period including weekends?

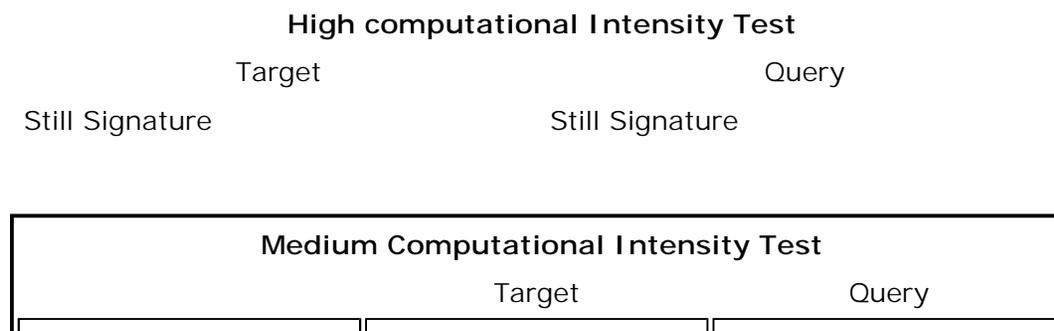
The test site will be available from 0900-1700, seven days a week.

39. I am still having trouble understanding the use of XML in FRVT 2002. Can you help me?

There have been a number of questions regarding the use of XML in FRVT 2002. Please [click here](#) to review an important document where we attempt to clarify these issues. First is a recommendation to participants. After, is a detailed discussion of how validation works with respect to the technical dataset.

40. How many target set and query set will be given for each subtest?

The High Computational Intensity Subtest has one target set and one query set. The Medium Computational Intensity Subtest has two target and query sets. See the two tables below for more information.



Still Target & Query Set	Still signature Still-set signature	Still signature Still-set signature
Video Target & Query Set	Still signature Video signature	Video signature

41. Is it necessary that a participant's POC be on site, or could it be a different employee of the participant or some other person engaged by the participant?

A Participant's POC serves a very important role during the on-site portion of the evaluation as they will serve as that participant's official representative. To change the individual designated as a participant's POC for the on-site portion of the evaluation, the responsible party (as designated in the "Request to Participate in FRVT 2002" form you submitted), must send an e-mail to frvt2002@nist.gov stating the new POC name and e-mail address. FRVT Organizers must receive this change notification two weeks prior to the participant's scheduled arrive date in order to process visit requests to the test facility.

42. Each output file consists of similarity data. Must this data be sorted in a specific format?

It is very important that data in the similarity files are sorted in a specific way. Any reordering or missing values in the similarity files will make scoring impossible. Please see page 5 of the Implementer's Guide for a complete specification of the format.

43. Can you supply some sample data in color, or indicate where a database of color face image data which is representative of test data may be obtained?

Data being used for FRVT 2002 was collected using the DARPA HumanID data collection protocols. You can look at Appendices G, K and L of the FRVT 2000 Evaluation Report for an understanding of the specific collection setup and resultant images used for FRVT 2000. While somewhat similar, there will be some deviation in the collection setup and images used for FRVT 2002.

44. What is the minimum and maximum for the width and height (in pixels) of the imagery in each test?

Min Width	Min Height	Max Width	Max Height
-----------	------------	-----------	------------

HCInt	252	300	252	300
MCInt	200	240	300	360

45. What is the scaling for similarity and distance values (0-10, 0-100....)?

The output values can be between $-\infty$ to $+\infty$. Participants are however encouraged to use a limited scale (such as 0-1, 0-10, etc). If possible, try to avoid duplicate/tied scores.

46. Medium Computational Intensity Test will be administered with two targets and query sets - one for still images, one for video images. Are you going to provide us with a disk containing tree structure (Chapter 3 of Implementers guide) for each set?

The precise directory tree structure of the MCInt test, like the HCInt test, should not be needed by a participant prior to testing. The input and output paths and filenames are embedded in the XML signature-set files.

To summarize the implementers guide:

Input: The JPEG images are named in the "name" attribute of the "file" element. This includes the path relative to the location of the XML file it is in. For example:

```
<file name="images/12345/zz98765.jpg"/>
```

Output: The output similarity files MUST be written to the file named in the "name" attribute of the query signature, again with path relative to the current XML file. For example:

```
<signature name="output/abcde/54321.sim">
```

Participants are reminded that they can also send email to frvt2002@nist.gov with technical XML-specific questions.

47. Will the subtests be run sequentially or in parallel? Can we use the same machine for both subtests?

Participants have been given a 12-day window (approximately noon on day one until noon on day twelve) for testing their face recognition systems. Also, the total time for each subtest is 264 hours and there are different computer limitations for each subtest. Therefore, most participants will be performing both sub-tests in parallel on two computer

systems. (Note: The two sets within the Medium Computational Intensity Test will be performed sequentially on the same computer.) The only way the two subtests can be performed sequentially is if a participant can complete both subtests, using the more rigid medium computational intensity test computer limitations, within the given 12-day testing window.

48. You have said that the query & target set of HCI nT are identical. Will FRVT 2002 organizers provide two XML files or only one XML file?

FRVT 2002 Organizers will provide two XML files. These files (target and query set) are identical.

frvt2002_high_intensity_target.xml
frvt2002_high_intensity_query.xml

49. Are we allowed to bring other people, such as a key technician or software engineer, in addition to the Technical Point of Contact during the 9 to 5 period when we have access to the system?

Each participant will be allowed to bring two representatives: Their FRVT 2002 POC and one other individual.

50. FAQ # 3 states that the number of pixels between the eyes will be 30 - 500 pixels. How does this relate to FAQ # 44?

FAQ #3 was answered much earlier - before FRVT 2002 Organizers had finalized the imagery. The answer to FAQ #3 was based on possible imagery. The answer to FAQ #44 (a more recent FAQ) is based on actual imagery used. Answers to both questions are technically correct.

51. The technical dataset contains 3 video signatures named "video 00-02". We do not find any video sigs in the FRVT supplemental directory, so we assume that we do not have to process any video sigs for the sample similarity files we will submit for the FRVT evaluation procedure. Is this correct?

Correct. No sample submission from video files is required.

52. If the test program crashes, may our representative or representatives restart the program?

Yes.

53. Can we place phone calls from the test facility to our offices? Are cell phones allowed? Are laptop PC's allowed?

Cell phones and laptops are allowed. We are working to get a dedicated land-line phone installed for participant's use but cannot guarantee its availability at this time. We will provide a number where participants can be reached for emergencies.

54. Will the test area in Dahlgren have 20-amp AC outlets?

Yes.

55. Do we need to get clearance from NSWC Dahlgren in advance for anyone that we bring, both the Technical Point of Contact and others? Are there special issues, especially long delays, for clearance for non-US citizens?

There are no long delays of getting individuals onto the base. The FRVT 2002 Liaison will be contacting all participants soon to gather personal information for access to the base.

56. Can you provide an example of the 5-page document we are required to submit?

See Appendix J from the FRVT 2000 Evaluation Report for examples.

57. Are face locations nearly center in the image or could the face be located on a side?

The key feature of the imagery is that each image consists of a single face that is the largest object in the image. The position of the face could be anywhere in the image.

58. What is the approximate breakdown of signatures in the Medium Computational Intensity Test?

The following numbers are high-bound estimates and could change somewhat:

Video approximately 2500
Still approximately 5000
Still-set approximately 500

59. How many still images are in a still set signature?

Between 2 and 10.

60. Is any remote access to the system such as TELNET or FTP available? Can we monitor the test remotely, e.g. over phone lines or the Internet? If the test program crashes, may we restart another debugged version of the test program?

The goal of FRVT 2002 is to test fully functioning systems. Each participant has signed up to have one system evaluated for each test. The option to have more than one system evaluated could have been negotiated up to the sign-up deadline. Participants are restricted to only changing system configuration components of their system (e.g. the USB drive was given a different drive position than you anticipated). See test plan for more details. Changing the code or otherwise modifying a system once it arrives on site is not allowed. This includes debugging, changing the code, downloading patches, or changing face recognition system parameters. External monitoring/control of the system during testing (via TELNET, FTP, PC Anywhere, dialup, Internet, wireless, etc.) is not permitted.

61. Participants have been given 264 hours for testing. A participant finishes early, will they be permitted to leave early?

Yes. Once you have completed the test(s) and successfully checked out, you will be permitted to leave.

62. In the event that the voice-only telephone is not available and cell phone calls cannot be placed from within the building due to the metal covering mentioned in the FRVT 2002 Test Plan or other reasons, may participants leave the building to place cell phone calls from outside for technical support or other reasons?

Yes. Foreign nationals will have to be escorted.

63. May participants monitor the system at any time during the hour, or is it a specified time such as the start of the hour (9:00, 10:00, 11:00,...)? How much time do we have for the hourly monitoring, such as a maximum of five

minutes?

Participants will be taken into the testing area at the top of the hour for approximately 5-10 minutes. The timing and amount of time may change according to various circumstances (power failure, computer crash, etc.)

64. May participants examine the contents of files generated by the test program such as similarity files or other intermediate status files written by the application?

Yes.

65. May participants view a console output from the program, messages such as "processing probe 64,123...."?

Yes.

66. What is the extension for the similarity file? .SIM? .BIN?

The output name is EXACTLY that already specified in the name attribute of the query's signature. It MAY include an extension, or it may not. It may also include a path. In any case participants must NOT add to, subtract from, or change the name in any way. This is discussed more in section 2.2.1 in the Implementer's Guide.

67. Will you supply the power and data cables for the IDE drives in USB2.0 enclosures containing your image database?

Yes.

68. We will bring only Linux based systems for the test. How shall we do the virus scan?

Participants using Linux will not be required to install anti-virus programs.

69. How often does power go out at NSWC Dahlgren?

Power is fairly stable at NSWC Dahlgren. Power occasionally goes out during large thunderstorms, usually for only a few minutes.

70. What are the deadlines for submitting a normalization function?

You must let the FRVT 2002 Liaison know by 5:00 PM EDT 3 July 2002 if you plan to submit a normalization function. A reasonable amount of time will be given to participants to submit their normalization functions. These may result in some participants submitting their normalization functions after they have completed taking the FRVT 2002.

71. Would a windows based library be acceptable as well for the normalization function?

In a separate document, we will release the specification for the normalization functions. The scoring package for the FRVT 2002 will run on a Linux. Hence, the normalization function will need to run on a Linux. The proposed method of implementation is: the function will be a gcc or g++ routine supplied to the FRVT 2002 organizers as a compiled object binary (.o file). Then we can link the object binary when we compile the code.

72. What is the number of signatures and images for the tests in FRVT 2002?

Approximate (non-statistically rounded) numbers are given below. Please see Notes below the chart for additional caveats for this data.

Test	Input			Output		Total
	#Signatures	#Images	#Megabytes	#Files	#Megabytes	#Megabytes
High	121589	121589	1500	121589	57000	58500
Med Video	1900	40000	1000	2000	20	1020
Med Stills	4500	4500	1200	4500	100	1300
						61000

Notes:

1. The numbers for the MCInt will change somewhat before testing begins.
2. The MCInt Video segment includes signatures of two types: video and still; it does not contain still-set signatures.

3. The MCInt Stills segment includes signatures of two types: still and still-set; it does not contain video signatures.

73. Are each signatures type and modality name in XML file? What about directory structure?

Please see the example XML file [here](#). Participants should note that the tree structure of the tests is implied by the paths given in the xml lines:

```
<file name="apath/bpath/cimage.jpg"/>
<signature name="output/dpath/">
```

Users will not have to create any of the directories in the output tree, as they'll be present with the supplied materials. No tree structure will be supplied beforehand.

74. After completing the medium computational intensity test, may we add that computer to the system that is running the high computational intensity test?

Yes, under the following conditions:

- 1) For the HCInt, time begins when the first computer starts. You will still be on the clock when you are adding the additional computer to the HCInt network.
- 2) At no time may the number of computers working on the problem exceed the limit of three workstations. For example, when the MCInt test is being taken, you will have one computer run the MCInt and two computers run the HCInt. When the MCInt is finished, you will have three computers running the HCInt.

75. We are planning to bring more than two people for FRVT2002 because of our functional role and contractual schedule. If you are very strict about the number of people (only two people will be allowed at the test site) for the testing, we want to be able to switch people during the test still maintain two people at the test site (including POC).

Only two individuals will be permitted at the test site at any given time. The individuals at the test site can change, but only two will be allowed at one time. According to the test plan, participants are also allowed to switch POCs after notifying test agents and receiving their permission.

76. What is the directory structure and naming format for files we generate in the MCInt?

The output name is EXACTLY that already specified in the name attribute of the query's signature. It MAY include an extension, or it may not. It may also include a path. In any case participants must NOT add to, subtract from, or change the path and name in any way. All the subdirectories in the output tree will already be present; users will not have to call "mkdir" or an equivalent. They should be able to just open the output file for writing.

77. In the Outline and Overview for Use of Normalization Functions in FRVT 2002 we are told that a possible normalization can be based, in part, on similarity scores for all pairs of gallery images in G. Does this imply that G consists of a subset of the target set which is also represented as probe sets in the query set Q? Will there be a way to identify these subsets other than by brute force searches of the XML target and probe set definitions?

Normalization is a post processing procedure that occurs when the FRVT 2002 organizers score the results. As such, the participants do not need to know what are the galleries at the time the test is administered. At the time scoring takes place, the FRVT 2002 organizers will give the appropriate gallery similarity scores to the normalization routine.

78. For the High Computational Intensity test, are we allowed to copy the FRVT data (pictures, xml,...) on the three computers? At the end, are we allowed to copy back to the USB drive from each of the three computers or do we need to centralize all similarity files on one computer before copying back?

You may copy all data to the three machines as needed. For the output files you may transfer them among machines as you see fit. All the output files must ultimately be placed on the USB drive.

Participant Area FAQ closed during evaluation period from 2 July through 13 August.

79. When is the normalization function due?

1700 EDT on 20 August

80. There seem to be time limits imposed on the normalization functions. And if a F2 normalization is taking too long, the testers will revert to using the F1 normalization function. How much time is allotted per gallery siz?. Meaning, for a gallery size of X how long should F2 or the F1 function take before it is deemed too long?

Participants that are concerned about the speed of their normalization functions should submit, to frvt2002@nist.gov, the following information:

- an approximation of how long the normalization routines will take on
 - a 256 x 256 score matrix
 - the same estimate for a 512 x 512 score matrix.
- the approximate machine specifications of the machine used for the timing test
- an order of speed in "big-O" notation

81. There are several math functions that our normalization function uses such as sqrt and exp. Is it okay to expect that whatever application is linking in our .o file will also link in the standard c/c++ code for these functions (either libm.a or libstdc++.so seems to do it fine on our machines here)?

It is safe to assume that the normalization code will be linked against the standard math library (libm). This includes, but it not limited to, functions such as abs(), exp(), log(), log10(), pow(), sqrt(), sin(), cos(), tan(), and so on.

82. Will the normalization functions be used on the high computational intensity test as well as the medium computational intensity test? If so, will there be a way for the functions being called to know in advance whether the scores being passed in are coming from the high versus the medium computational intensity test?

The normalization function could be used for either or both the MCInt and the HCInt. Distinguishing between the two tests within the normalization function is not permitted nor necessary. Section 3 of the FRVT 2002 Normalization Specifications clearly states the number and terms of normalization functions that can be submitted.

83. What is the minimum size of the gallery being passed to the normalization functions?

The FRVT 2002 Organizers will design experiments to adequately characterize the impact of normalization on performance. No further information regarding the experiments within FRVT 2002 will be discussed prior to release of the FRVT 2002 Evaluation Report.

84. In FRVT 2002 Normalization Specifications: What is the meaning of "identification, verification, and watch list tasks"?

Identification: Closed universe test. Ranks the gallery by similarity to a probe. A probe is correctly identified if the identity of the top rank is the same as the probe.

Verification: Open universe test: Determines if the claimed identity of a face is correct.

Watch List: Open universe test: The gallery is the watch list. Consists of a two stage process. First - determine if the person in a probe is a person in the gallery (on the watch list). Next - if the probe is someone in the gallery, determine which person is it.

85. Is it planned to send the participants a pre-release of the FRVT 2002 Evaluation Report, allowing them to comment on it (as was done for the FRVT 2000 Evaluation Report)?

Yes

Appendix F Methods and Procedures

F.1 Introduction

To offer full public disclosure of the methods and procedures used during the Face Recognition Vendor Test 2002 (FRVT 2002) evaluation, this appendix provides a description of the administration of FRVT 2002. It is the authors' goal to fully convey to the reader the amount of work that is required to properly perform a technology evaluation of biometric systems. By stimulating discussion and analysis of the methods and procedures used for FRVT 2002, future evaluations may be improved.

The remainder of section F.1 provides a description of communication methods between the FRVT 2002 evaluation team and participants and a high-level timeline of important events in FRVT 2002. Section F.2 provides information regarding the announcement and invitation to participate in FRVT 2002, section F.3 discusses the requirements to participate and section F.4 discusses the participant selection process. Section F.5 discusses pre-test activities, and F.6 discusses post-test activities, that were required to successfully administer FRVT 2002.

F.1.1 Communication Methods

To maintain fairness, the FRVT 2002 evaluation team took special care to ensure that all participants received information about the evaluation at the same time. A government employee was named the FRVT 2002 Liaison. Except for extenuating circumstances, this individual was the only link from FRVT 2002 evaluation team to the outside world, including participants.

Two methods of communication were developed for discussions between the FRVT 2002 evaluation team and participants. First, the evaluation team established a frequently asked questions (FAQ) page on the participant area of the FRVT 2002 web site. Both tentative and committed participants could submit questions to the evaluation team through the web site, or through a dedicated e-mail address. The evaluation team then posted the questions on the FAQ page along with the answers. This primary and preferred method of communication ensured that all tentative participants and participants had equal access to information concerning FRVT 2002.

F.1.2 Timeline

FRVT 2002 was performed in a much more compressed timeframe than FRVT 2000. Whereas FRVT 2000 took thirteen months to prepare, six weeks to proctor, and eight months to develop a report, FRVT 2002 took six months to prepare, five weeks to administer, and five months to develop this report. Successful completion of FRVT 2002 in this timeline was only possible because the four principles from FRVT 2000 (Duane Blackburn, Mike Bone, Patrick Grother, and Jonathon Phillips) were able to perform similar duties in FRVT 2002, and Ross J. Micheals and Elham Tabassi were able to join the team at NIST with their own unique expertise. All told, approximately twenty individuals directly assisted in FRVT 2002 in some capacity.

Preliminary discussions regarding FRVT 2002 occurred at NIST and in a telephone discussion on 2 January 2002. Less than a week later, the FRVT 2002 evaluation team had agreed upon a rough outline. A brief review of available data sources that met the evaluation ideals was followed by an intense two-month period where the team gathered, analyzed and selected images for use in FRVT 2002. This activity culminated in the first formal meeting of the evaluation team in mid March. The purpose of this meeting was to finalize the makeup of FRVT 2002, review options for test location, define individual roles on the team for the remainder of the evaluation, and review participation forms and website content.

Federal agencies were first told of the existence of FRVT 2002 on 24 January. A follow-up announcement to federal agencies was sent on 15 April to prepare them for the 25 April public announcement. Potential participants had from that date until 10 May to apply for participation. Participants were notified individually on 20 May and were provided their ID and Password for the Participant Area of the FRVT 2002 website. The Participant Area was an area of the FRVT 2002 website where additional information was provided to tentative participants and participants.

FRVT 2002 Evaluation team provided the FRVT 2002 Technical Dataset on 3 June. The technical data set contained documents that formally specified the input and output that would be used during the actual test, a sample face recognition system that operated under these protocols and an estimate of the computational power required to complete the evaluation. Participants were required to submit correct sample output files based on the dataset to FRVT 2002 Evaluation team by 28 June.

To ensure that applicants would be able to successfully complete the HCInt evaluation, the evaluation team mandated that applicants provide a resource assessment and to submit a white paper by 14 June that showed their anticipated time to complete the HCInt. The final list of FRVT 2002 participants was placed on the Participant page of the FRVT 2002 website on 18 June. Test dates were given to the participants on 13 June and the FRVT 2002 Test Plan was provided on 25 June. Testing occurred in two periods to help the government manage the evaluation process. The first period of testing went from 10 July through 21 July and the second from 24 July through 9 August.

F.2 FRVT 2002 Announcement

The FRVT 2002 Liaison announced the FRVT 2002 evaluation to the public on 25 April via five methods. First, an announcement was sent to the Biometrics Consortium listserv.¹ This free and publicly available listserv currently has over 800 members from across the world and is generally regarded as one of the primary methods of sharing information throughout the biometrics community. Second, an announcement of the FRVT 2002 was placed on *Federal Business Opportunities* (formerly Commerce Business Daily or CBD)². *Federal Business Opportunities* is the single government point-of-entry for Federal government procurement opportunities over \$25,000. Through this one portal, commercial vendors seeking federal markets for their products and services can search, monitor and retrieve opportunities solicited by the entire federal community. Third, a posting on FRVT 2002 was placed in the free-to-use Biometrics Catalog³. Fourth, the FRVT 2002 Liaison sent an individual e-mail to 88 laboratories, universities and companies that had previously claimed to the U.S. government to have commercial or mature prototype face recognition products. Finally, e-mail was sent to various integrators, resellers, industry associations, conference organizers, and biometric reference publications/websites for them to further disseminate.

On 6 May, the FRVT 2002 Liaison sent a second series of e-mails announcing the FRVT 2002 to the Biometrics Consortium Listserv, the group of 88 agencies, universities and companies, and the group of other entities in the biometrics field.

F.3 Qualifications to Participate

To assess the state-of-the-art in face recognition technology, developers of the technology were invited to participate in FRVT 2002. The evaluation was limited to commercially available systems and mature prototypes. Additionally, only developers of core technology were permitted to participate, rather than developers licensing software from other core technology developers. In order to request participation, potential participants were required to electronically fill out the two forms shown in Appendix D, print and sign them, and send them back to the FRVT 2002 evaluation team. Those whose forms were received by 10 May were classified as tentative participants. Tentative participants were then reclassified as participants on 18 June after successful completion of several preliminary activities.

¹ <http://www.biometrics.org/html/listserv.html>

² <http://fedbizopps.gov/>

³ <http://www.biometriccatalog.org/>

F.4 Participation Overview and Test Schedule

F.4.1 Participation Overview

Everyone that asked to participate in FRVT 2002, and successfully completed all pre-evaluation processes, was allowed to participate. Twenty-two tentative participants signed up to take FRVT 2002 and received the technical datasets. Of these, 14 later committed to participating and were listed on the FRVT 2002 website as “participants.” Ten actually attempted the evaluation. All ten participants completed the evaluation.

F.4.1 Developing the Test Schedule

When creating the test schedule, the FRVT 2002 evaluation team planned on having no more than twelve participants on site at any time. Another constraint was to have no more than three participants start on the same day. The start dates were staggered to reduce the human traffic caused by multiple participants moving in equipment on the same day. Since there were more than twelve tentative participants at the time the schedule was originally created, it was decided to run the evaluation in two sessions with the second session starting after all participants in the first session had completed testing. Each participant was scheduled to have twelve days for testing, with actual test time of 264 continuous hours for each test.

On 13 June, there were 20 tentative participants. Each was randomly assigned one of the possible starting test dates and notified of this date via e-mail. Several tentative participants withdrew from participation after this date. Some participants also requested that their assigned dates be changed. The FRVT 2002 evaluation team granted these requests by reassigning them to empty slots in the schedule. Table 1 shows the start dates and latest allowed end dates used for FRVT 2002 testing.

Table 1 - Scheduled start and end dates for each FRVT 2002 Participant. Note that not all participants required their full allotment of test dates.

Participant	Start Date	End Date
Eyematic Interfaces	10 July	21 July
Viisage	24 July	4 August
AcSys Biometrics	25 July	5 August
Cognitec Systems GmbH	25 July	5 August
Imagis Technologies	25 July	5 August
Dream Mirh	26 July	6 August
VisionSphere Technologies	26 July	6 August
C-Vis	26 July	6 August
Iconquest	29 July	9 August
Identix	29 July	9 August

F.5 Pre-test Activities

Before taking part in FRVT 2002, tentative participants and the FRVT 2002 evaluation team were required to complete several preliminary steps to prepare for the evaluation. These steps were designed to help guide them through some of the technical issues of the evaluation so that problems could be resolved before the actual testing began. The following subsections describe those activities.

F.5.1 Technical Datasets

A technical data set (Appendices G) was made available on the participant area of the web site on 3 June. It provided documentation on the format of the image data that would be distributed to participants during the evaluation and the required output format that participants would need to use for submitting results. It also contained a set of sample images, the source code for a sample face recognition algorithm, source code for software used to generate scores from matching results, and results based on the images and a sample face recognition system. Tentative participants were encouraged to use the source code as a guide for modifying their own software to run the evaluation. It also provided a means for them to test their own software and compare the resulting scores with those obtained using the sample face recognition algorithm. The technical data set also contained an example of the required HCInt Computational Intensity White Paper.

F.5.2 HCInt Computational Intensity White Paper

The FRVT 2002 evaluation team anticipated that many applicants might have underestimated the difficulty of the High Computational Intensity Test (HCInt). As part of the Technical Data Set, provided on 3 June, participants were given an example of how to derive their anticipated time and computational requirements to participate in that test. Applicants were mandated to perform a similar computation for their system and to submit a white paper by 14 June that shows their anticipated time to complete the HCInt. This allowed HCInt participants to prove that it was theoretically possible for their systems to completely process the large number of images included in the test. Failure to submit this paper resulted in their removal from participating in the HCInt. Applicants were also strongly encouraged to perform a similar assessment for the Medium Computational Intensity Test, but were not required to do so. All participants that started the HCInt and MCInt completed within the allotted times.

F.5.3 Sample Similarity Files

Without proper similarity files, the FRVT 2002 evaluation team could not produce the results shown in this report. The evaluation team and tentative participants worked together to ensure the accuracy of participant's similarity file development process prior to actual testing.

The technical data set contained a set of sample image files. Tentative participants were required to process these images using their software and produce similarity files in the proper format. They were required to submit these files to the FRVT 2002 evaluation team for analysis. If any problems were encountered with the submitted files, tentative participants were allowed to make corrections and resubmit as many times as necessary until the deadline of 28 June.

F.5.4 Normalization Protocol

The FRVT 2002 uses the Sep96 FERET evaluation protocol. The key innovation in this protocol is the usage of virtual galleries and probe sets. Virtual galleries and probe sets allow for the measurement of performance for different categories of images from one set of similarity files. This approach did not allow the use of normalization that some participants regularly use in their product. Similarity scores are classically computed by comparing two images and reporting a measure of similarity between the two faces in those images. Normalization adjusts these scores based on all the images in a gallery.

On 6 August, the FRVT 2002 Liaison announced an amendment to the test protocol to permit the usage of a participant's normalization routine if they desired. For FRVT 2002 purposes, normalization was a post-test administration procedure that was executed by NIST after similarity files had been submitted after testing. Participants who elected to utilize the FRVT 2002 normalization option were required to provide NIST, by 20 August, an object file that contains the result of the compilation of C code of specific functions. This file was then called at an appropriate time by the NIST scoring code during the scoring process. For intellectual property reasons participants were not allowed to submit source code or details of their normalization techniques to the FRVT 2002 evaluation team. The Normalization API is available in Appendix J of this document.

F.5.5 Dry Run

The FRVT 2002 evaluation team performed a dry run at the test facilities on 9 July 2002. NIST provided a baseline face recognition system and acted as a participant as the FRVT 2002 test agents stepped through the entire test plan twice. Although no major issues were found, the dry run was very beneficial as the majority of the contracted test agents had not been previously involved in the planning and had several questions concerning what they were and were not allowed to do while proctoring the tests.

F.6 Post-test Activities

Participants were also required to perform several activities in support of the FRVT 2002 Evaluation Team after physical testing had been completed. These are discussed in the sections below.

F.6.1 Scoring Code Inspection

The scoring suite used for the Face Recognition Vendor Test 2002 was an updated version of those used for the FERET and FRVT 2000 evaluations. The FRVT 2002 Evaluation Team implemented a multi-step process to ensure that the newer version of the scoring suite is accurate. Several of these steps were performed by NIST to ensure that results using the FRVT 2002 scoring code would be identical to results using either the FRVT 2000 scoring code or the FERET scoring code.

The last step in this process was to provide the FRVT 2002 participants a copy of the scoring suite and the similarity files from a baseline system so that they could study the suite and report any implementation errors. The FRVT 2002 Liaison provided participants a set of similarity files, the scoring software and the results from that scoring software on 30 October 2002. Participants were asked to study the validity of the scoring code and provide feedback to the evaluation sponsors if they found any software implementation errors by 7 November 2002. No errors were reported.

F.6.2 Participant Review of FRVT 2002 Evaluation Report

The FRVT 2002 Evaluation Team provided a copy of this report to the participants for their review seven days prior to public release. Participants were given the option of preparing and submitting a position paper on the evaluation results within these seven days. These position papers are included in Appendix N of this report without modification.

This page intentionally blank.

Appendix G
Technical Data Set

The FRVT 2002 Implementer's Guide

Version 0.99.0

Initial FRVT-2002 Technical Dataset Release

Patrick Grother Ross J. Micheals P. Jonathon Phillips

National Institute of Standards and Technology
Information Access Division, Image Group
Gaithersburg, MD 20899

May 31, 2002

Chapter 1

Introduction

This paper details what Face Recognition Vendor Test 2002 (FRVT 2002) participants need to be able to do to complete the FRVT 2002 tests. The document is split into four chapters. This introduction gives an overview of the face recognition evaluation protocol in use. Chapter 2 describes the format of the files given to, and required from, FRVT 2002 participants. Chapter 3 is a procedural description of what participants should do during the test. Chapter 4 details a PCA based face recognition algorithm that is capable of reading and writing the required FRVT 2002 files.

Participants are also directed to the following related documents:

1. *The FRVT 2002 Technical Datasets*. This collection contains a face recognition system capable of reading and writing the files used in FRVT 2002. It also includes a small extract of the FERET database (256 images) used for demonstration purposes.
2. The short paper *Computational Requirements of the FRVT 2002* is an analysis of the time required to complete the FRVT 2002. It is vital that participants read this document, consider its implications for their systems, complete the attached form, and return it to the FRVT 2002 organizers.
3. The logistical aspects of the FRVT 2002, for example times, dates, locations are formally described in the *Face Recognition Vendor Test 2002 Test Plan*.
4. FRVT 2002 is an implementation of *The HumanID Evaluation Framework* that defines a general mechanism for the evaluation of arbitrary multimodal biometric systems. The document is not required reading, but participants may refer to it for a more complete reference on the use of XML in FRVT 2002.

Readers of this document are directed to specific files in the *Technical Datasets* archive that accompany this document. Such pointers are intended to make concrete the concepts explained here.

1.1 Overview

Image recognition systems and biometric based identification systems usually operate by comparing new unknown images with a set of known images, often referred to as the gallery. The unknown, or probe images, are usually obtained some time after those in the enrolled gallery. For example, in a verification setting the new image is obtained from a person claiming to already be in the gallery. In an identification setting the new image may belong to a person who is, or is not, already present in the gallery. The performance of recognition systems on these various tasks can be estimated off-line in a so-called technology evaluation by requiring the system to report numerical comparisons between pairs of images which act respectively as probe and gallery images. The FERET protocol ¹ formalized this

¹See Phillips et al. *The FERET Evaluation Methodology For Face Recognition Algorithms*, IEEE PAMI 22:1090-1104, 2000.

procedure, and the FRVT 2002, like the FRVT 2000² before it, uses the FERET methodology. One of the FERET protocol's features is that images of subjects taken under many different conditions are present in the testing sets, and consequently the performance of a system may be estimated for any image or subject condition represented by a subset of the image set. For example, if verification performance is required on a population of persons who have aged by six years and who have grown beards, then that number can be computed from *just* that subset of the reported recognition scores that correspond to those specific images. Many such recognition tasks can be embedded in the one test.

The FRVT 2002 uses an extended version of the FERET protocol known as the *Human ID Evaluation Framework*³ (HEF) that allows for the quantitative testing of recognition systems on heterogeneous sets of multimodal biometric signals. The multiple face-face comparisons of FERET are replaced by the abstract notion of comparing a *signature* with a *signature*. In this context a *signature* could be a single face image as before, or a set of faces, or a video of a person, or a set of ten fingerprint images, a speech recording, a composite fingerprint and face record, or, indeed, any other biometric information in any data format.

Thus we define the terms of FRVT 2002. Two *tests* are planned, referred to as the Medium Computational Intensity and High Computational Intensity tests. Each contain imagery representative of several imaging and/or subject circumstances. The imagery is specified by two *signature sets* called the *target* and *query* sets. The related terms *gallery* and *probe* refer, respectively, to subsets of *target* and *query*, and it is these subsets that specify the targeted application-specific experiments embedded in a *test*.

The *test* proceeds by a recognition system reporting real scalar *similarity* or *distance* values for all pairs of *signatures* in the *target* and *query* sets. The resulting *similarity matrix* forms the basis for many statistical analyses that result in various performance values.

1.2 Contact Information

The primary and preferred source of communications between participants and FRVT 2002 government personnel is the FRVT 2002 Frequently Asked Question (FAQ) page, which is located off of the FRVT 2002 website, <http://www.frvt.org>. A government employee has been named as the FRVT Liason. The FRVT Liason will be the primary point of contact and will manage communications with participants. Questions that are outside the purview of the FAQ can be submitted to the FRVT Liason via email at frvt2002@nist.gov. All e-mails sent to the FRVT Liason and germane to the FRVT 2002 will be posted on the FAQ page. This policy ensures all participants are guaranteed equal access to information concerning FRVT 2002. Contact with the FRVT 2002 Liason other than via the FAQ areas, or contact with someone other than the non-FRVT Liason will only be allowed under extenuating circumstances.

²See Blackburn et al. *Facial Recognition Vendor Test 2000*, February 2001.

³See Grother et al. *The Human ID Evaluation Framework* a copy of which accompanies the *FRVT 2002 Technical Datasets*.

Chapter 2

Input and Output File Specifications

The primary requirement for FRVT 2002 is that for each *test* that participants undertake they must have the ability to read two files, one each for the *target* and *query* sets, compare the named *signatures* therein, and write binary files that contain measures of the likeness of the individual appearing in those *signatures*.

The formats of the input and output files are detailed in the following sections.

2.1 Signatures

The input *signature set* files are in an XML format¹. XML is distinguished by its flexible and extensible usability in document mark up, and its machine readability. The ability of vendors to read XML is a fundamental requirement of FRVT 2002.

The *target* and *query* sets are *signature sets* containing many *signatures*. A *signature* is a named collection of arbitrary biometric information from one *individual*. FRVT 2002 is structured around the ability of a Face Recognition algorithm to compare any two *signatures* and produce a scalar real number, either a *distance* or a *similarity* value. Specifically, participants are required to compute and write a scalar for each *signature* pair occurring in the *target* and *query* sets. Further there are three kinds of *signature* in FRVT 2002, each of which contains only face imagery. Their definitions follow, and an example of each kind is given in figure 2.1.

1. A *still* is a single still image of one person's face,
2. A *still-set* is a set of still images of one subject, possibly taken at different times or under different conditions
3. A *video* is a sequence of frames of an individual taken at regular intervals over some short period of time.

Note the *file* elements at the deepest level of XML nesting are the most concrete. They give the physical location of the images, including a path relative to the location of the XML *signature set* file it is in. In FRVT 2002, all images specified in the *name* attribute are standard JPEG files, as indicated in the *type="jpeg"* attribute.

When a FRVT 2002 participant's algorithm encounters a pair of single-file still *signatures* it is clear that the system should do the conventional matching and report a real scalar number. For multi-file *signatures* (i.e. still set or video) the same is true, a real scalar must be reported. How participants effectively use all *n* files of *signature* A and *k* files of *signature* B is up to the participant's application,

¹See <http://www.w3c.org/xml> for example

```

1:  <signature name="2010123/1394">
2:    <sigmember xsi:type="simple-face-still-type" modality="simple face still">
3:      <dataset media="digital still" type="jpeg">
4:        <file name="solo_pic.jpg"/>
5:      </dataset>
6:    </sigmember>
7:  </signature>

8:  <signature name="2010122/5079">
9:    <sigmember xsi:type="simple-face-still-type" modality="simple face still">
10:     <dataset media="digital still" type="jpeg">
11:       <file name="left_camera_pic.jpg"/>
12:     </dataset>
13:   </sigmember>
14:   <sigmember xsi:type="simple-face-still-type" modality="simple face still">
15:     <dataset media="digital still" type="jpeg">
16:       <file name="right_camera_pic.jpg"/>
17:     </dataset>
18:   </sigmember>
19: </signature>

20: <signature name="50002/02463579">
21:   <sigmember xsi:type="simple-multifile-face-video-type"
22:     modality="simple multifile face video">
23:     <dataset>
24:       <file name="0002-01-frame00.jpg"/>
25:       <file name="0002-01-frame01.jpg"/>
26:       <file name="0002-01-frame02.jpg"/>
27:       <file name="0002-01-frame03.jpg"/>
28:       <file name="0002-01-frame04.jpg"/>
29:     </dataset>
30:   </sigmember>
31: </signature>

```

Figure 2.1: The three kinds of *signatures* that will be used present in FRVT 2002.

Examples of XML *target* and *query* sets are *FERET/FERET-256-1.sigs.xml* and *FERET/FERET-256-2.sigs.xml* in the *Technical Datasets* distribution. Code for reading such *signature sets*, and constructing a usable internal representation thereof, is contained in the file *software/src/SigSet_XML.cpp* as called from the application *software/src/apps/HEF-Validator-app.cpp* on the *Technical Datasets* distribution.

2.2 Similarity Files

While section 2.1 was concerned primarily with marking up standardized *input* for evaluations, this section defines the binary format that face recognition systems must *output*.

The overall requirement for a specific *test* is that if there are N *signatures* in the *target* set and M *signatures* in the *query* set, then a participant's code must produce M output files each containing N *similarities* or *distances*. The matrix of NM values is referred to as a *similarity matrix* although is not usually stored or manipulated as such.

2.2.1 File Names

The output files take their names from the *signature name* attribute. This will generally include a path so that an application that has parsed the *query signature set* can write the similarity scores to a file whose name is exactly that in *query signature*. The application may assume that any directories present in the *name's* path will already be present in the provided output directory tree.

2.2.2 File Format

For all tests these output files are in a binary, unformatted I/O, format. The XML format described in *The HumanID Evaluation Framework* document is disallowed in FRVT 2002 because, for large *signature sets*, the total storage requirement for such XML similarity files could be several hundred gigabytes.

The file consists of a 20 byte header, followed by as many *similarity* or *distance* values as there are *signatures* in the *target* set, and a final closing 8 byte footer. The header consists of four items: a “magic number”², a value for testing the byte order (endian-ness) of the subsequent values, a count of the number of *similarity* or *distance* values, and a value for *polarity* (see section 2.2.4). The bulk of the data, N binary floats, follows. It is **imperative** that the i -th element of the similarity data corresponds to the i -th *target* signature. Any reordering or missing values will make scoring impossible. The file is terminated with a repeat of the magic number.

The similarities of the N *target signatures* to each *query signature* are stored in a single file whose name is taken from the *name=*”” attribute of the *query*. For a *target* set of size N , the file **must** contain precisely $4N + 24$ bytes. The 4 arises because the required number of bytes used to represent one scalar real number is the size of a single precision float on most contemporary architectures. Participants must not output higher precision values, and they should ensure that their values are representable in the range of a float. Either Big or Little Endian byte ordering must be used consistently throughout. Sample binary files must be submitted to the FRVT 2002 organizers before the **tests** are commenced. Formally, the numbers must be reported in IEEE 754 format that is used on Intel based PCs, and many UNIX platforms. Participants with any concerns as to their floating point formats should collaborate with the FRVT 2002 organizers proactively.

In summary, the contents of each binary file is, in order,

Description	Value	Data type	Number of items	Total Bytes
Magic number	FRVT2002	char	8	8
Endian Indicator	0x12345678	unsigned int	1	4
Number of Distances	N	unsigned int	1	4
Polarity	0 or 1	unsigned int	1	4
Similarity	3.14159	float	1	4
...	...	float	$N - 2$	$4(N - 2)$
Similarity	8.31441	float	1	4
Magic number	FRVT2002	char	8	8

Strict adherence to this format is mandatory because recovery from erroneously generated files may not be possible.

2.2.3 Endian Indicator

The endian indicator is used to ensure that when similarity files are written and read across architectures, different byte orderings³ can be corrected for. A mechanism for detecting the byte order of the machine that *wrote* the similarity scores is as follows. First, take an integer of value, and store it in memory (line 34), and write it to a file *one byte at a time* (line 35). Later, when this file is read in the same fashion, the “natural” byte ordering imposed by the machine architecture is reflected in the input.

2.2.4 Polarity

For each *signature* in the *query* set, the FRVT 2002 participant must produce a *similarity file*. All such files may contain one of two categories of value, termed *similarity* or *distance*. A large value for a *similarity* implies sameness of identity of the subject appearing in the *signature*. A large value for *distance*, on the other hand, implies difference of identity. This concept is termed *polarity* and must be recorded in all *similarity files* as discussed above. The permissible values are 0 for *similarity* values and 1 for *distances*.

²A simple heuristic means for distinguishing this kind of file from all others.

³Byte order is also known as *endianness* or *byte sex*.

```

1: #ifndef FRVT2002_simwriter
2: #define FRVT2002_simwriter

3: #include <iostream>           // for fatal errors to cerr
4: #include <stdio.h>           // for file output
5: #include <string>            // internal rep. of signature names
6: #include <vector>            // for the actual data
7: using namespace std;

8: class simfileWriter
9: {
10:     private:
11:         bool isdistance;      // are the values distances of similarity

12:         string queryname;     // will create a file of this name, that
13:         typedef pair<string, float> PSV; // will contain real-numbered similarity
14:         vector<PSV> results;  // values in either a binary or XML format
15:         simfileWriter() {}    // this is not allowed, so hide it.

16:     public :
17:         // construct with two arguments:
18:         // 1. a string containing the output file name. This is the name
19:         //    of a query signature, as described in the FRVT 2002 guide.
20:         // 2. a bool that is true if the real numbered data will be
21:         //    distances, or false if similarities. This is the polarity
22:         //    described in the FRVT 2002 Implementers Guide.
23:         simfileWriter(const string &queryname, const bool values_are_distances);

24:         // add a name, value pair to the container.
25:         void push_back(const string &targetname, const float simvalue)
26:             { results.push_back( PSV(targetname, simvalue) ); }

27:         void write() const;    // write the data to file
28:         // void write_xml_version() const; // unused in FRVT 2002 but in HEF chap 3
29:     };

30: #endif

```

Figure 2.2: A C++ class specifying a container for holding similarity data during processing.

2.2.5 Sample Code

Figure 2.4 gives C++ code for writing similarity files. It shows the output method of the *simfileWriter* container class shown in Figure 2.2. The pseudo code in figure 2.3 shows how the class could be used in a real FRVT 2002 system. Alternative working C++ writing code is included with the accompanying *Technical Datasets* distribution; specifically the file *software/src/CSU/csuSubspaceTest.cpp* contains a function called *writeDistances()* that outputs binary data to file.

```

1: // FRVT 2002 pseudo code for one test

2: // this code is illustrative only, it is included only in the context of the
3: // simfileWriter class. Readers are warned that the code will not compile as
4: // as it stands, and the calculations that it attempts to depict are not
5: // necessarily recommended, worthwhile, bug free, realizable, or in any way
6: // appropriate for a FRVT 2002 participant to use or copy.

7: // bring in the results container class
8: #include <simfileWriter.hpp>

9: // querynames and targetnames are strings
10: for ( unsigned int q = 0 ; q < querynames.size() ; q++ )
11: {
12:     // each query of the query set results in an output file
13:     // so construct a container for the output data
14:     simfileWriter result(querynames[q]);

15:     // retrieve the previously computed low-dimensionality feature vectors
16:     // of the image(s) that comprise this query signature
17:     const feature_vecs qrep(querynames[q]);

18:     // compare each query with all members of the target set
19:     for ( unsigned int t = 0 ; t < targetnames.size() ; t++ )
20:     {
21:         // retrieve the previously computed low-dimensionality feature vectors
22:         // of the image(s) that comprise this target signature
23:         const feature_vecs trep(targetnames[t]);

24:         // compute some measure of the distance or similarity
25:         const float v = compute_facial_similarity_metric(qrep, trep);

26:         // store the similarity/distance of the query sig to this target sig
27:         result.push_back(targetnames[t], v);
28:     }

29:     result.write(); // write the required output file
30: }

```

Figure 2.3: Pseudo code representing how the *simfileWriter* class could be used in a working FRVT 2002 implementation.

```

1: #include <simfileWriter.hpp> // for fatal errors to cerr

2: // !! warning this code is intended only to demonstrate the output format
3: // !! required of FRVT 2002. It is not suggested nor recommended that this
4: // !! is code be used without first checking that it conforms with the
5: // !! formal binary specification in the main text.

6: // member:   simfileWriter::write
7: // purpose:  write an unformatted (binary) similarity/distance file
8: // args:     none (they come from the members of the class)
9: // fails:    code will cause program to write an error message and exit if
10: //          1. the output file specified on construction is not openable
11: //          2. if any of the required about data cannot be written to file
12: //          3. the output file cannot be closed successfully
13: // limitations: code should throw exceptions instead of calling exit

14: static void bail(const string &msg, const string &fn, const int errorcode)
15: {
16:     const string routine = "simfileWriter::write(): failed to ";
17:     cerr << routine << msg << endl; // may write errno too
18:     if (errorcode) exit(errorcode);
19: }

20: void simfileWriter::write() const
21: {
22:     const size_t one = 1;

23:     // open the output file for binary writing
24:     FILE *fp = fopen(queryname.c_str(), "wb");
25:     if (!fp)
26:         bail("fopen for writing binary file: ", queryname, 40);

27:     // write a simple header containing two unsigned ints:
28:     // 1. the number of similarity/distance entries to follow
29:     // 2. the polarity: a value of 1 indicates a "distance" measure
30:     //    a value of 0 indicates a "similarity" measure
31:     const char *magic = "FRVT2002";
32:     if (strlen(magic) != fwrite(magic, sizeof(char), strlen(magic), fp))
33:         bail("fwrite first magic number into file: ", queryname, 41);
34:     const unsigned int endian_tester = 0x12345678;
35:     if (4 != fwrite((void*) &endian_tester, sizeof(char), 4, fp))
36:         bail("fwrite an endianness confirmation into file: ", queryname, 42);
37:     const unsigned int nvalues = results.size();
38:     if (one != fwrite(&nvalues, sizeof(unsigned int), one, fp))
39:         bail("fwrite an unsigned int into file: ", queryname, 43);
40:     const unsigned int polarity = isdistance ? 1 : 0;
41:     if (one != fwrite(&polarity, sizeof(unsigned int), one, fp))
42:         bail("fwrite an unsigned int into file: ", queryname, 44);

43:     // write the real numbered similarities/distances one at a time
44:     // ignoring the targetname string, the name is omitted for economy (previously
45:     // the FERET and FRVT 2000 tests wrote text files that included the name)
46:     for ( vector<PSV>::const_iterator b = results.begin() ; b != results.end() ; b++ )
47:         if (one != fwrite(&b->second, sizeof(float), one, fp))
48:             bail("fwrite a float into file: ", queryname, 45);

49:     // write the magic number again, and close
50:     if (strlen(magic) != fwrite(magic, sizeof(char), strlen(magic), fp))
51:         bail("fwrite last magic number into file: ", queryname, 46);
52:     if (0 != fclose(fp))
53:         bail("fclose file: ", queryname, 47);
54: }

```

Figure 2.4: The write method code for outputting FRVT 2002 similarity files. The functionality of this routine **must** be emulated by participants.

Chapter 3

Procedural Requirements

For each test, participants will be given a disk containing tree structures of the form in figure 3.1. The two XML files at the top level will be named after the test, one for each of the *target* and *query* sets as shown. Those files will contain *file* elements whose name attributes will include paths, having the general form *images/somemajor/someminor/someimage.jpg*. Participants are required to write output files specified by the name attribute of the *query signatures*, **not** the *file*'s name, that also will generally include a pathname into the output directory tree.

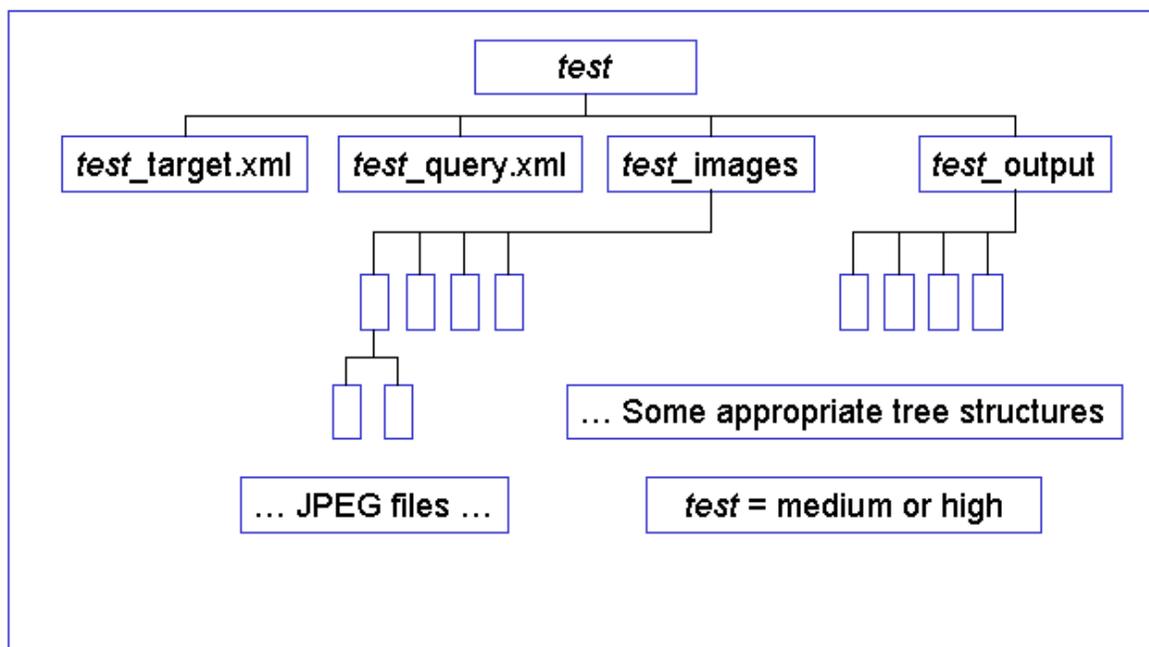


Figure 3.1: The tree structure given to participants for each *test* .

Thus a procedural description of what vendors must do is given below, but the document *FRVT 2002 Test Plan* should be consulted also.

1. Mount the drive supplied by the FRVT 2002 Organizers. This may simply consist of plugging in a USB drive. Alternative media such as CD or DVD may be available, by prior request.
2. Recursively copy the whole tree to a working partition. Note that the output directory tree should

also be copied because it contains subdirectories that the application will populate as it proceeds.

3. Change directory into the top-level directory. This implies that the participants application has this as its working directory.
4. Execute the test. Participants must prepare to do this successfully before attending the formal FRVT 2002 test. In this respect they should see the document *Computational Requirements of the FRVT 2002* that addresses this issue more extensively.
5. Ensure the correctness of the output files. This should be done by participants own software.
6. Recursively copy the output tree back to the supplied media (the USB drive), overwriting the (empty) one that was there.
7. The FRVT 2002 Test Agents will then delete all data on the working partition. There are strict requirements that all the material supplied to participants must be purged from the host machines. The FRVT 2002 proctors will further specify this process in the *FRVT 2002 Test Plan*.

Chapter 4

FRVT 2002 Technical Datasets Reference Implementation

The *HumanID Evaluation Framework* (HEF), is a mechanism for the quantitative testing of recognition systems. The *FRVT 2002 Technical Datasets Reference Implementation* is a software package that implements a variety of functionality described by the HEF, with respect to the components of the HEF that will be used in FRVT 2002. Throughout this document, the term HEF-TD will be used to refer to this, the HumanID Evaluation Framework Technical Dataset software.

4.1 Caveats

Using this code in your own software for FRVT 2002 is neither required, nor necessarily recommended. It is provided only as an example implementation of a system that can read XML and output binary files of the same structure to be used during the actual tests. The provided code has been tested, but should not be considered to be “bug-free.”

4.2 Directory Structure

The HEF-TD has the following directory structure. The top directory has three main subdirectories, *documents/*, *FRVT-2002/* and *software/*. The contents of the *documents/* and *software/* directories are self-explanatory. Files similar in structure and nature to those that will be used in FRVT-2002 are located in the *FRVT-2002/* directory.

The *software/* directory has the most complex structure. Unless noted otherwise, there is a single copy of each source code file (shared between both Unix and Windows) in *software/src/* or *software/include/*. Under Unix-like operating systems, object files are stored on the *objs/* directory. Under both Unix and Windows, library files and executable binaries and are stored in *software/lib/* and *software/bin/* respectively. The files required to build HEF-TD under windows are in the *Projects* directory. Each supported compiler has its own directory within *Projects/VC6/*. At this time, because **Microsoft Visual Studio 6** is the only supported Windows compiler, there is a single subdirectory, *Projects/VC6/* in this folder. The *software/import/* directory is reserved for external packages (such as Xerces). This directory is further subdivided by operating system i.e., *software/import/unix/* and *software/import/Win32/*.

A full listing of the contents of the technical dataset are as follows:

File or Directory	Description
documents/	Documentation directory
documents/implementers_guide/	Implementer's Guide directory
documents/implementers_guide/FRVT2002-implementers_guide.ps	Implementer's Guide (Postscript version)
documents/implementers_guide/FRVT2002-implementers_guide.pdf	Implementer's Guide (PDF version)
documents/manual/	HEF Manual directory
documents/manual/HEF-FRVT2002-TDS.ps	HEF Manual (Postscript version)
documents/manual/HEF-FRVT2002-TDS.pdf	HEF Manual (PDF version)
documents/xml/0.99.0	Version 0.99.0 XML Schemas
documents/xml/0.99.0/sigset-schemas	Signature Set Schemas
documents/xml/0.99.0/sigset-schemas/multiple-simple-face-stills-set.xsd	HEF Multiple Simple Face Stills Set schema
documents/xml/0.99.0/sigset-schemas/sigset-general.xsd	General (abstract) signature set schemas
documents/xml/0.99.0/sigset-schemas/sigset-unrestricted.xsd	HEF Unrestricted Signature Set schema
documents/xml/0.99.0/sigset-schemas/simple-face-still-type.xsd	HEF Simple Face Still Set schema
documents/xml/0.99.0/sigset-schemas/simple-face-still-type.xsd	Simple Face Still schema type definition
documents/xml/0.99.0/sigset-schemas/simple-face-video-set.xsd	HEF Simple Face Video Set Schema
documents/xml/0.99.0/sigset-schemas/simple-face-video-type.xsd	Simple Face Video schema type definition
documents/xml/0.99.0/sigset-schemas/simple-multifile-face-video-set.xsd	HEF Multifile Face Video Set schema
documents/xml/0.99.0/sigset-schemas/simple-multifile-face-video-type.xsd	Simple Multifile Face Video type definition
documents/xml/0.99.0/sigset-schemas/single-sigmember.xsd	HEF Single Sigmember schema
FRVT-2002/	FRVT 2002 data and examples directory
FRVT-2002/FERET/	FERET data directory
FRVT-2002/FERET/coords.512	FERET imagery ground truth data
FRVT-2002/FERET/FERET-256-1.sigs.xml	Example signature set
FRVT-2002/FERET/FERET-256-2.sigs.xml	Example signature set
FRVT-2002/FERET/m150x130.dat	Image mask used during normalization
FRVT-2002/FERET/raw/	Directory containing raw image data
FRVT-2002/FERET/raw/*.pgm	FERET imagery
FRVT-2002/FERET/norms/	Target directory for image norms
FRVT-2002/video/	Video data directory
FRVT-2002/video/raw/	Directory containing raw video frames
FRVT-2002/video/raw/*.pgm	Example (short) video sequences
FRVT-2002/video/video-coords	Ground truth data for video clips (for normalizer)
FRVT-2002/video/video.nrm.ss.xml	Video data signature set
README	Main README file (points to this document)
software/	HEF Technical Dataset software directory
software/bin/	General binaries and executables directory
software/bin/demos/	Target directory for demonstration binaries
software/bin/apps/	Target directory for application binaries
software/config.guess	Required component of <i>configure</i> script
software/config.sub	Required component of <i>configure</i> script
software/configure	Unix configuration script
software/configure.in	Autoconf script (generates <i>configure</i>)
software/import/	General target directory for external (imported) software
software/import/unix/	Target directory for external (imported) Unix software
software/import/Win32/	Target directory for external (imported) Win32 software
software/include/	General include file directory
software/include/CSU/	Directory for CSU include files
software/include/CSU/	Directory for CSU include files
software/include/CSU/csuDebuglevel.hpp	CSU debugging macros
software/include/CSU/csuFileCommon.hpp	CSU common function declarations
software/include/CSU/csuGlobal.hpp	CSU globals
software/include/CSU/csuMatrix.hpp	CSU matrix declarations
software/include/getopt.hpp	Helps provide <i>getopt()</i> capabilities under Win32
software/include/HEF/	Directory for HEF include files
software/include/HEF/config.hpp.in	Used by <i>configure</i> script to generate <i>config.hpp</i>
software/include/HEF/dummy.h	Dummy header file for <i>configure</i> script
software/include/HEF/Signature.hpp	Signature class declarations
software/include/HEF/SigSet.hpp	Abstract Signature Set class declaration
software/include/HEF/SigSet_XML.hpp	XML-based Signature Set class declaration
software/include/HEF/std_DefaultHandler.hpp	std_DefaultHandler class declaration
software/include/HEF/string-utils.hpp	String utility function declarations
software/install-sh	Binary installation script
software/lib/	Target library directory
software/Makefile	Default Makefile
software/Makefile.in	With <i>configure</i> , used to generate Makefile
:	:
:	:

File or Directory	Description
⋮	⋮
<code>software/objs/</code>	Target directory for object files
<code>software/Projects/</code>	General IDE directory
<code>software/Projects/VC6/</code>	Directory for Visual C++ 6 files
<code>software/Projects/VC6/HEF/</code>	Directory for HEF Visual C++ 6 files (redundant)
<code>software/Projects/VC6/HEF/CSU Normalizer/</code>	VC6 directory for CSU Normalizer
<code>software/Projects/VC6/HEF/CSU Normalizer/CSU Normalizer.dsp</code>	VC6 CSU Normalizer project
<code>software/Projects/VC6/HEF/CSU Subspace Engine/</code>	VC6 directory for CSU Subspace Engine
<code>software/Projects/VC6/HEF/CSU Subspace Engine/CSU Subspace Engine.dsp</code>	VC6 CSU Subspace Engine project
<code>software/Projects/VC6/HEF/CSU Subspace Train/</code>	VC6 directory for CSU Subspace Engine
<code>software/Projects/VC6/HEF/CSU Subspace Train/CSU Subspace Train.dsp</code>	VC6 CSU Subspace Train project
<code>software/Projects/VC6/HEF/HEF.dsw</code>	VC6 HEF workspace
<code>software/Projects/VC6/HEF/LibHEF/</code>	VC6 LibHEF directory
<code>software/Projects/VC6/HEF/LibHEF/LibHEF.dsp</code>	VC6 LibHEF project
<code>software/src/</code>	General source directory
<code>software/src/CSU/</code>	CSU source files directory
<code>software/src/CSU/csu1face2norm.cpp</code>	CSU/NIST normalization code
<code>software/src/CSU/csuEigenTrain.cpp</code>	CSU Eigenface training code
<code>software/src/CSU/csuFileCommon.cpp</code>	Common file manipulation code
<code>software/src/CSU/csuFisherTrain.cpp</code>	CSU Fisherface training code
<code>software/src/CSU/csuMatrix.cpp</code>	CSU Matrix code
<code>software/src/CSU/csuSubspaceTest.cpp</code>	CSU subspace training code
<code>software/src/CSU/csuSubspaceTrain.cpp</code>	CSU subspace engine code
<code>software/src/CSU/cveigen.cpp</code>	CSU eigenvector code
<code>software/src/CSU/getopt.cpp</code>	Provides <code>getopt()</code> under Win32
<code>software/src/demos/</code>	General demonstration code directory
<code>software/src/HEF-config.in</code>	With <code>configure</code> generates <code>HEF-config</code>
<code>software/src/Signature.cpp</code>	Signature class code
<code>software/src/SigSet.cpp</code>	SigSet class code
<code>software/src/SigSet_XML.cpp</code>	XML SigSet class code
<code>software/src/std_DefaultHandler.cpp</code>	std_DefaultHandler class code
<code>software/src/string-utils.cpp</code>	string utility functions code

4.3 Required Software

HEF-TD has been designed to be platform agnostic — currently, the supported platforms are Linux and Microsoft Windows. In order to build HEF-TD on either system, you will need the following software.

4.3.1 HumanID Evaluation Framework FRVT-2002 Reference Implementation

Naturally, to build the HEF-TD, you will need the latest version of the HEF-TD source code. If you do not have it already (i.e., you are reading a standalone version of this document) HEF-TD can be downloaded from the participant area of the FRVT-2002 website <http://www.frvt.org>.

4.3.2 C++ Compiler Suite

HEF was designed to work with **GCC** (the GNU compiler collection) and **Microsoft Visual C++**. Although **GCC** is supported, HEF-TD currently cannot be compiled under the Cygwin environment, since Xerces-C++ does not yet support Cygwin.

To build HEF-TD under Linux, you'll need **GCC 2.96** or better. Because **HEF-TD** makes heavy use of the C++ Standard Library (i.e., C++ strings, and templates), you'll need a version of **GCC** that has fairly mature template support. As of early 2002, most Linux distributions come with a sufficient compiler. If necessary, you can download and build your own `gcc` from <http://gcc.gnu.org>.

To build HEF-TD under Microsoft Windows, you'll need **Microsoft Visual Studio 6** with the components necessary for compiling C++ programs already installed. If you have not done so already, apply **Microsoft Visual Studio 6.0 Service Pack 5**. The code should also compile under **Microsoft Visual Studio .NET**, although .NET is not yet an “officially” supported platform.

4.4 Xerces-C++ 1.7.0

The HEF-TD makes heavy use of the Apache Group's free XML parser, **Xerces-C++**. To build under Linux, you will need the **Xerces-C++ 1.7.0 Stable Source Package for Unix's**, or *xerces-c_1_7_0.tar.gz*. This can be downloaded for free from the Apache Group at <http://xml.apache.org>. For Windows, you will want to download the **Xerces-C++ 1.7.0 Stable Source Package for Windows**, or *xerces-c_1_7_0.zip*. To minimize the potential for link-related problems, it is recommended that you use the same compiler for both HEF-TD and Xerces-C++.

4.5 Suggested Software

Although the following are not required to compile and run the HEF-TD, they are recommended if you want to see the original code used for the PCA implementation, and how it was adapted (a.k.a. "hacked") to accept XML input and output.

4.5.1 CSU Face Recognition Algorithm Source Code

Although HumanID Evaluation Framework does not include any face recognition algorithms, so that developers can see an example of a working system, in this reference implementation, we have included a version of the Colorado State University PCA implementation which has been modified to work with HEF. The original CSU code is available from the website <http://www.cs.colostate.edu/evalfacrec/>.

4.5.2 The FERET Database

To run the code with the same datasets used in developing and testing this package, you will want a copy of the **FERET Database: March 2001 Release**. For more information about obtaining your own copy, see the NIST FERET website at <http://www.nist.gov/humanid/feret>.

4.6 Build Instructions

If you are compiling the HEF-TD for Windows, you can skip directly to the next section (4.6.2). Otherwise, continue on to section (4.6.1). We recommend that you carefully read the installation instructions *completely* before you begin.

4.6.1 Linux Build Instructions

To build HEF-TD under Linux using the GCC compiler, follow these instructions:

1. On your development machine, untar the file *HEF-0.99.0.tar.gz* in a directory of your choosing. This can be accomplished with the command¹

```
% tar xvfz HEF-0.99.0.tar.gz
```

The package will untar into its own directory.

2. Next, you will need to compile Xerces. If you already have Xerces on your system, set the *XERCE-SCROOT* environment variable and continue to step 3.

¹The leading "%" represents the system prompt (which on your system, will most likely look different than this) and should not be typed.

- (a) Copy the *xerces-c-src1_7.0.tar.gz* into the *HEF-0.99.0/software/import/unix* and untar it. (You will need to adjust these commands for your system — below it is assumed that the Xerces tarball is located in your current directory, which may not be the case for your system).

```
% cp xerces-c-src1_7.0.tar.gz HEF-0.99.0/software/import/unix
% cd HEF-0.99.0/software/import/unix
% tar xvfz xerces-c-src1_7.0.tar.gz
```

- (b) Set your *XERCESCROOT* environment variable to the **Xerces** root directory. (Make sure you do this with *export* (bash) or *setenv* (csh), so that the *configure* scripts can see it).
- (c) Out of the box, **Xerces** needs some modification. From the *xerces-c-src1_7.0/* subdirectory, change into the *src/xerces/*. And run:

```
% chmod 755 ./runConfigure
% chmod 755 ./configure
```

to make *runConfigure* and *configure* executable.

- (d) Next, run the *runConfigure* command with arguments appropriate to your system (run it without arguments to see the complete usage information). On a “typical” Linux system the command

```
% ./runConfigure -p linux -c gcc -x g++
```

should suffice.

- (e) Finally, from the *HEF-0.99.0/software/import/unix/xerces-c-src1_7.0/src/xerces/* directory, compile **Xerces** with

```
% gmake
```

or

```
% make
```

Xerces suggests using the command “gmake” in case you have a version of make on your system other than GNU make. This will take some time. If Make returns with no errors, then **Xerces** was probably built correctly. Note that a “make install” is *not* required.

3. From the *HEF-0.99.0/software/* directory, run the *configure* command.

```
% ./configure
```

The leading “./” is necessary if “.” (the current directory) is not part of your *PATH*.²

4. Finally, to make the example PCA implementation, run

```
% make csu-pca
```

This will make the *libHEF.a* library (saved in *lib/*) and variety of executables in the *bin/* directory.

4.6.2 Windows Build Instructions

If you are compiling the HEF-TD under Linux, you can skip this section. Otherwise, to build HEF-TD under Windows using Microsoft Visual Studio, follow these instructions:³

1. On your development machine, unzip the file *HEF-0.99.0.zip* into a directory of your choosing. You will be accessing this directory many times, so it is not recommended that you nest it more than one or two levels deep from a drive’s root.

²This Unix feature that is seemingly an annoyance is actually for security. If the current directory is in your path, you might unknowingly be running someone else’s malicious executable if you happen to be in their directory.

³Note that we use the forward-slash (/) in these instructions, instead of the backslash, (\) the “correct,” but not strictly enforced, directory delimiter in DOS/Windows.

- Unzip the file *xerces-c-1.7.0.zip* into the directory *HEF-0.99.0/software/import/Win32/*. Afterward, you should have a directory called *HEF-0.99.0/software/import/Win32/xerces-c-src1.7.0/* which contains the Xerces-C++ 1.7.0 source.
- Start up **Microsoft Visual Studio 6.0**. Open the Xerces workspace (*xerces-all.dsw*), located in *HEF-0.99.0/software/import/Win32/xerces-c-src1.7.0/Projects/Win32/VC6/xerces-all/*.
- Within this workspace, you will see a large number of projects. To the *XercesLib* project, within the workspace panel, right-click on *XercesLib* and select “Set as Active Project”. Build the project by selecting *Build xerces-c-1.7.0D.dll* from the *Build* menu. Depending on your system, building the library will take some time. Upon successful compilation, you should see the message

```
xerces-c-1.7.0D.dll - 0 error(s), 0 warning(s)
```

under the *Build* tab of the output panel. This is a good time to build and play with other projects in the workspace. However, for HEF, you will only need to build *XercesLib*.

- Close the *xerces-all* workspace.
- From Visual Studio, open the HEF workspace (*HEF.dsw*), which is located in the directory *HEF-0.99.0/software/Projects/VC6/HEF/*.
- Build HEF-D.lib, the HEF static library. From within the workspace panel, right-click on the LibHEF project and select “Set as Active Project”. Build the project by selecting *Build LibHEF-D.lib* from the *Build* menu.
- Build the rest of the software. From the *Build* menu, select “Batch Build...”. Unselect all of the “Release” versions of the projects — there is no real need to build these. Although it is not necessary, you may also want to unselect LibHEF, since it was built already. Make sure “Selection only” is marked (i.e., has a check in it), and select the “Build” button. As part of the build process, the static library file *LibHEF-D.lib* will be copied to the *software/lib/* directory, and a variety of executables will be copied into *software/bin/*.
- Close the workspace and quit **Microsoft Visual Studio 6.0**.

4.7 Run Instructions

In this section, a full walk-through of a face recognition experiment is given. Before beginning, make sure the that your *PATH* environment variable includes the *HEF-0.99.0/software/bin/* directory. Additionally, under Unix, you’ll want your *LD_LIBRARY_PATH* environment variable to include the location of the *xerces-c-1.7.0D.so* or *xerces-c-1.7.0.so* library (the latter, if you chose to compile the release version of the library). Under Windows, there is no separation of executable path, and dynamic library path. Your Windows *PATH* must include a path to the directory containing *xerces-c-1.7.0D.dll* (or *xerces-c-1.7.0.dll*). You might just want to copy into *C:/Windows/System*. In either case, make sure you point to the same library that you linked against during the compilation of HEF.

Here is a command-by-command example run of a “typical” experiment, as run on a Linux system, under the (GNU) bash shell. Under Windows, you can use the *cmd* shell (with slightly modified commands), or use the **bash** shell that comes packaged with the **Cygwin** environments (<http://www.cygwin.com>).

- Convert the Imagery** (if necessary). If you are working with the original FERET distribution (and not the PGMs included with HEF), then you will need to convert the TIF images to PGMs, so that the normalization code can operate on the images. This can be done with the “convert” command provided by the **ImageMagick** package. To convert the TIFF *file.tiff* to a *pgm* with the name *file.pgm* use the command:

```
% convert file.tiff pgm
```

To convert all of the files in a directory, use the bash "for" loop construction:

```
% for file in *.tif; do convert $file pgm; done
```

2. **Normalize the Imagery.** From this step on, we assume that our current directory is this package's *FRVT-2002/FERET/* directory, and that we have all of our images, as PGMs, in *FRVT-2002/FERET/raw/*. In this directory, you should have files *FERET-256-1.sigs.xml*, *FERET-256-2.sigs.xml*, *coords.512*, and *m150x130.dat*. These are included in the distribution.

To normalize the images, we use the CSU normalization code, *which does not yet support XML*. We run the command

```
% csu-normalizer coords.512 raw norms -m m150x130.dat
```

where

- *coords.512* is the ground truth file provided with FERET,
- *raw* is the directory containing the PGM imagery,
- *norms* is the directory to put the normalized images, and
- *-m m150x130.dat* indicates an image mask file, (also from the FERET distribution).

This generates normalized images, and puts them in the *norms/* directory. Notice that *csu-normalizer* tries to normalize every file listed in the ground truth file, and may complain if it does not find a corresponding PGM.

3. **Make a copy or symbolic link to the schemas.** When using a validating XML parser, the application will need access to the schema documents themselves. If you are running under Linux, then, from the *FRVT-2002/FERET/* directory, run the command

```
% ln -s ../../documents/xml/0.99.0 .
```

which will create a symbolic link to the XML schema directories. Under Windows, you will need to copy the *0.99.0/* directory tree, from *documents/xml/0.99.0/* to *FRVT-2002/FERET/*. This can be accomplished via the command line, or Windows explorer. If the schemas are not copied or symlinked correctly, then the example implementations will fail to run correctly, since the applications will not be able to properly locate the HEF schema files.

4. **Train the algorithm.** Next, we need to train the PCA algorithm (i.e., generate the feature space upon which we will project the normalized images). Do this with the command:

```
% csu-subspace-train -imDir norms/ FERET-256-1.sigs.xml FERET-256-1.fs
```

where

- *-imDir norms/* is a directory (note the trailing forward-slash) where the normalized images can be found,
- *FERET-256-1.sigs.xml* is an XML document listing 32 normalized images. (This file actually contains more than 256 norms, but 224 of them have been commented out).
- *FERET-256-1.fs* is the name of the output file in which the generated feature space will be saved.

The feature space file is used in the next step, where we actually run the algorithm. **Generating the feature space can take some time**, especially with a large number of images.

5. **Run the algorithm.** Finally, we can run the PCA algorithm itself. First a "results" directory should be present to store the similarity data. If not:

```
% mkdir results
```

Then, run PCA with the following command (shown on multiple lines for clarity)

```
% csu-subspace-engine -imDir norms/ FERET-256-1.fs
                        FERET-256-1.sigs.xml FERET-256-2.sigs.xml
                        results/ Euclidean
```

where

- *-imDir norms/* is the path to the normalized images
- *FERET-256-1.fs* is the feature space (a.k.a. the output from the training step)
- *FERET-256-1.sigs.xml* is the target file
- *FERET-256-2.sigs.xml* is the query file
- *results/* is a directory where the similarity scores will be written
- *Euclidean* is the distance metric to use (run *csu-subspace-engine* without arguments to see more metrics).

Note that in this run, we used the same file for the training, gallery, and *query* sets, but in a real experiment, this would likely not be the case. The outputs of the experiment are saved in the *results/* directory.

4.7.1 About the Code

There are two major components that were added for accepting XML input. First, is the class *std_DefaultHandler* (defined in *src/std_DefaultHandler.cc* and *include/HEF/std_DefaultHandler.h*). These files are wrappers for the Xerces-C++ *DefaultHandler* class, but replace *signatures* that use the Xerces structures (such as *XMLCh* and *Attributes*) with the C++ standard library classes *string* and *map*. The class *SigSet_XML*, which inherits from the class *std_DefaultHandler* and the purely virtual class *SigSet*, can be used to take an XML-based *signature set* file, and parse it into a collection of data structures. These data structures, defined in *Signature.h*, are based on the HEF XML specification.

If you are integrating HEF into your own code (or vice versa) here are a few helpful tips.

- Under both Linux and Windows, it's easy to get paths, binaries, and libraries inconsistent with each other. Most of the time, such errors manifest themselves in the form of either static or dynamic linking problems. Whenever you encounter such a problem, look carefully at your paths, libraries, and versions first.
- Under both Linux and Windows, the order in which files are linked is critical. When linking your executables, make sure that your own object files are first, with the libraries last.

4.7.2 Known Bugs

There is a bug in **Xerces-C++ 1.7.0** which causes *uniqueness* constraints in schemas to not always be applied properly. To patch this bug “by hand,” then you will need to apply the following diff to the file *QName.cpp*.

```
417,419c417,421
< return (XMLString::compareString(fPrefix, qname.getPrefix())==0) &&
< (XMLString::compareString(fLocalPart, qname.getLocalPart())==0) &&
< (fURIId == qname.getURI());
---
> if (fURIId == 0) // null URI
> return (XMLString::compareString(getRawName(), qname.getRawName())==0);
>
> return ((fURIId == qname.getURI()) &&
> (XMLString::compareString(fLocalPart, qname.getLocalPart())==0));
```

This is not necessary to ensure proper operation of the software, but is recommended if you want to check yourself for duplicate *signature* name and ids.

Computational Requirements for FRVT 2002

1 Introduction

This document addresses the significant computational requirements of the FRVT 2002. It is imperative that participants understand the implications of this material as it pertains to their own systems. Participants are directed to the form included on the last page: It **must** be completed by participants and returned to the organizers at *frvt2002@nist.gov* by 17:00 on June 14 2002.

This document's scope is limited. The correct references for implementers are:

1. *The FRVT 2002 Technical Datasets*. This collection contains a face recognition system capable of reading and writing the files used in FRVT 2002. It also includes a small extract of the FERET database (256 images) used for demonstration purposes.
2. The logistical aspects of the FRVT 2002, for example times, dates, locations are formally described in the *Face Recognition Vendor Test 2002 Test Plan*.

The FRVT 2002 tests are extensions of the FERET protocol [2] as used previously in FRVT 2000 [1]. Because the number of images used in the High Computational Intensity test, ≈ 121000 , is an order of magnitude larger than that used in FRVT 2000, the computational loads are **significantly** larger. Indeed, the amount of computation required grows superlinearly with the number of images, as discussed below. Also the storage requirements are large also.

2 Timing Model

The FERET evaluation requires that N pieces of imagery in the target set \mathcal{T} are compared with M pieces in the query set \mathcal{Q} . These two sets are not necessarily disjoint i.e. $|\mathcal{T} \cap \mathcal{Q}| \geq 0$

The discussion that follows assumes the use of a modular recognition system comprised grossly of two parts: the first performs image analysis and dimensionality reduction on the two input sets; the second computes pairwise comparisons between those processed features. FRVT 2002 participants should augment this structure as necessary.

Thus the total processing time for any of the FRVT 2002 tests can be modelled as the sum of the times for input and image processing (enrollment), feature comparison, and output.

$$T(N) = \frac{N}{R_E} + \frac{M}{R_T} + \frac{NM}{R_C} + \frac{NM}{R_W} \quad (1)$$

where

Symbol	Name of Operation	NIST's Empirical Rate (s^{-1})
R_E	target image analysis and enrollment	1
R_T	query image analysis and enrollment	1
R_C	Pairwise comparisons of k -dimensional feature vectors	$1.6 \cdot 10^8 / k$
R_W	Four byte binary floats output to file	$2.5 \cdot 10^6$

Note the dependence of the comparison rate on feature space dimensionality, k . Note the quadratic nature of equation 1. Given the very large rates for the quadratic terms it is clear that for small N and M the expensive enrollment process dominates the total time. But as N increases the quadratic terms eventually dominate. The graph in 1 plots the linear enrollment times, the quadratic components, and

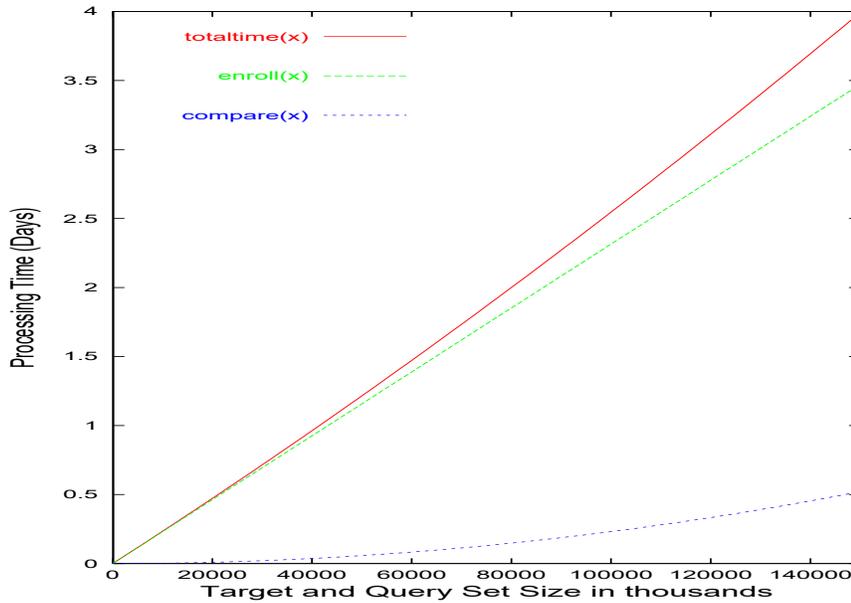


Figure 1: Processing times as a function of the *target* set size.

the total time. With our rate parameters the quadratic matching parts would exceed the enrollment times with $N \approx 1000000$.

The two enrollment figures in the final column were based on estimates from commercial products, and eye detection software running in-house. The computation rates were obtained using a short C++ program, compiled with g++ -O3, running under linux 2.4.7 on a PC equipped with a 1.5GHz Xeon Pentium III and 1GB memory. The feature comparison numbers were obtained using a plain $L2$ distance metric on feature vectors of dimension $k = 256$.

Example: For the High Intensity *test* in which $N = M = 125000$, $k = 256$ and the rates are those shown above, the resulting total time is

$$\begin{aligned}
 T &= 2 \frac{125000}{1} + 125000^2 \left(\frac{256}{1.6 \cdot 10^8} + \frac{1}{2.5 \cdot 10^6} \right) \\
 &= 2.5 \cdot 10^5 + 3.13 \cdot 10^4 \text{ s} \approx 3.26 \text{ days} \quad (2)
 \end{aligned}$$

This figure is well within the 11 days allowed for the *tests*.

Note in particular that the output time is considerable and the code is not simply CPU bound. Note also that the leading factor of 2 can be ignored if the *target* and *query* sets are identical.

3 Storage Model

The storage needed for FRVT 2002 is broken down into three parts. Storage in this context refers to either disk space or memory depending on the architecture of the system.

1. **Input:** The raw imagery supplied to vendors will not exceed 5 GB.
2. **Working Space:** If a system typically generates templates that persist then the storage requirement is likely linear in the size of the *target* set, $|T|$

Example: If the enrolled template from \mathcal{T} is a $k = 256$ dimensional vector then for the High Computational Intensity test with $N \approx 121000$ the target features would require 236 MB of main memory, assuming an eight byte double precision representation.

3. **Output:** The protocol requires that a vendor produce similarity files that contain the NM real-valued distance measures. This information must be written in a basic binary format. It is not clear that compression of this output will yield any significant reduction in storage. Thus any use of *zip* (or *tar -z*) is useful only as a bundling program.

Example: Given $N = M \approx 121000$ the the complete result will occupy about 58.6 gigabytes.

4 Caveats

1. **System Architecture:** The above models will not apply if the architecture of a system does not decompose the problem into separate enrollment and feature comparison parts. For example, a particularly naive system making NM calls of an SDK function $f(image1, image2)$ would embed NM enrollments and be very expensive; using example numbers, $T(125000) = 500$ years!
2. **Operating System Dependencies:** The empirical rate estimates quoted above are subject to measurement error. Particularly the disk IO is difficult to measure because the the operating system buffers and caches certain operations, and because of the underlying hardware (the RAID-1 striped partition used in the example above is substantially faster than a standalone SCSI disk).
3. **Restart:** If for any reason a recognition system ceases processing this could effect seriously effect times. Because our computations above only simulate an FRVT 2002 *test* without actually doing it, we cannot assert that uninterrupted processing is possible. Accordingly a mechanism for restarting an incomplete *test* might be needed.
4. **Video:** Our simulated times above assume the existence of single-file *signatures*, and so does not include any overhead associated with the processing of the multi-file *still-set* and *video signatures* as defined in the *FRVT 2002 Implementer's Guide*. This time could be considerable. See the FRVT 2002 online FAQ for approximate numbers of images in multi-file *signatures*.

5 Conclusion

The above models may be inappropriate for various reasons including and in addition to those given in 4 Accordingly vendors should assess the computational requirements of their own systems and bring sufficient CPUs, machines, and memory so as to complete the test in the allotted time. Certain limits on hardware are described in the *FRVT 2002 Test Plan*.

References

- [1] D. Blackburn, M. Bone, and P. J. Phillips. Facial recognition vendor test 2000. Technical report, 2001.
- [2] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face recognition algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22:1090–1104, 2000.

6 *FRVT 2002 : Statement of Competence*

We have received a large number of applications for participation in FRVT 2002, most of whom will take all *tests*. The FRVT 2002 Evaluation Team requests you review the data in the document above, *Computational Requirements of the FRVT 2002*, to gain an understanding of the processing and storage resource requirements associated with each FRVT 2002 *test*. Based upon this information, you may decide to modify your participation request. The deadline for withdrawal from FRVT 2002 without it being noted in the final report is June 14 2002. Our goal is to ensure that you are able to complete each *test* that you have applied for.

Participants must submit this page, duly completed with information requested, to the Evaluation Team at frvt2002@nist.gov by 17:00 EDT on June 14 2002.

Value	Name or Value
Operating System + Version	
Number and Speeds of Processors	
Machine Memory (MB)	
USB Connectivity (Type 1, 2)	
Disk type and capacity (GB) (IDE, SCSI, other)	
Implmentation Name (Marketing or or Algorithm name)	
Software Language (C, C++, ADA etc.)	
Feature Space Dimension (if appropriate)	
Time to enroll an image (in seconds)	
Time to compare two enrolled feature sets (in microseconds)	
Time to write 500 KB binary file	
Name of person filing form	
Date	

The HumanID Evaluation Framework

Version 0.99.0

Initial FRVT-2002 Technical Dataset Release

Patrick Grother Ross J. Micheals P. Jonathon Phillips

National Institute of Standards and Technology
Information Access Division, Image Group
Gaithersburg, MD 20899

May 30, 2002

Chapter 1

Introduction

The HumanID Evaluation Framework, or HEF, is a mechanism for the quantitative testing of biometric recognition systems. It is an extension of the FERET protocol, which defined a framework for the evaluation of face recognition technologies. The HEF, however, is general enough to apply to virtually any recognition task, and can be applied to arbitrary, heterogeneous mixtures of biometric systems.

The goal of the HEF effort is to design, implement, and deploy standards for the robust and complete documentation of the biometric system evaluation process. A mandate of the scientific method is experimental repeatability — achievement of this goal in the computer science and electrical engineering field is often thwarted by, among other things, inadequate documentation of experimental procedures.

This initial release of HEF is coupled with the 2002 Face Recognition Vendor Test, or FRVT 2002. Although FRVT 2002 has played a key role in shaping the current form of the HEF, it should be noted that they are separate, but related efforts. FRVT 2002 provides a unique opportunity for the development and implementation of an “initial” version of HEF. The lessons learned here will help determine the future shape of the HEF.

The primary focus of HEF is the documentation of the input test suites and output recognition hypotheses, and *not* the algorithms embedded within a particular recognition system. In various domains (character recognition, fingerprint matching, automatic target recognition), much effort has been put into the gathering of standard training and testing suites that facilitate algorithm development and subsequent evaluation. The use of test data for empirical testing is widespread in the scientific literature. The HEF is an attempt to leverage contemporary technologies for the formal description of such tests. Accordingly, the HEF defines a suite of XML-based markups for the inputs to, and outputs of, recognition systems.

The HEF is designed to facilitate black-box empirical testing. A recognition engine takes two sets of biometric signatures, the enrolled and the “unknown” test samples, and produces some form of identification data. Currently, the HEF assumes that this output data is a collection of scores, with each score indicating the similarity between a pair of signatures.

1.1 About this Document

This document is structured as follows. Chapter 2 discusses the mark up of signature sets — or the inputs to an evaluation. Chapter 3 discusses the mark up of similarity scores, or the outputs of an evaluation. Finally, in Chapter 4 instructions on how to build and run the example implementation are given.

This document assumes familiarity with software development, biometric systems, XML, and XML Schemas. An excellent introduction to XML Schemas is the XML Schema Primer, available from <http://www.w3.org/TR/xmlschema-0/>.

Throughout this document, we may use the following shorthand phrases and terms:

- The term *schema* refers to an XML Schema document. The term *instance document* refers to an XML document that conforms to the constraints defined by some fixed schema. An instance document is “valid”, or will “validate successfully” if it meets the constraints set out by a particular schema. Because an instance document might be valid under multiple schemas, or may validate with different success under different schemas, validation will often be referred to as “with respect to” or “under” a particular schema. Schemas themselves may be referred to by schema name (e.g. **HEF Similarity Set Schema**) or file name (e.g. “*similarity-set.xsd*”).
- The phrase “a valid $\langle x \rangle$ (element) contains n valid $\langle y \rangle$ (elements)” should be considered equivalent to, “a instance document that contains an $\langle x \rangle$ element and does not have n children tags of type $\langle y \rangle$ will not successfully validate.” With XML Schema, having the proper number of child elements is always a necessary, but not always sufficient, condition for successful validation.
- When referring to attributes, the phrase “attribute a is a b ” should be considered equivalent to “an instance document that contains a a attribute, and is not of type b , will not successfully validate.” Unless specified otherwise, types (such as “boolean,” “token,” or “double”) are as defined by the XML Schema specification, located at <http://www.w3.org/XML/Schema>.

It is important to note that schema validation enforces only the constraints encoded in the schema. An application may have its own set of constraints that are not reflected in the schemas. For example, a schema might enforce that a “sorted” attribute take on a “descending,” “ascending,” or “unsorted” value, but an *application* might enforce that they are actually sorted.

Participants are also directed to the following related documents:

Versions

This document describes the 0.99.0 release of the HEF FRVT-2002 Technical Dataset Release. For major and minor version numbers, the following version convention is proposed

Major	Minor	
0	98	FRVT-2002 Internal Release
0	99	Initial FRVT-2002 Release with Request for Comments
1	0	Final FRVT-2002 Release
1	1	Initial Public Release

For example, the first internal release will be version 0.98.0, the third revision of the Final FRVT-2002 Release will be 1.0.2, and so on.

Contact Information

Since this initial release of HEF is in conjunction with FRVT 2002, there are some unique rules that govern the contact information regarding this version of the HEF and the FRVT 2002 Technical Datasets. The primary and preferred source of communications between participants and FRVT 2002 government personnel is the FRVT 2002 Frequently Asked Question (FAQ) page, which is located off of the FRVT 2002 website, <http://www.frvt.org>. A government employee has been named as the FRVT Liason. The FRVT Liason will be the primary point of contact and will manage communications with participants. Questions that are outside the purview of the FAQ can be submitted to the FRVT Liason via email at frvt2002@nist.gov. All e-mails sent to the FRVT Liason and germane to the FRVT 2002 will be posted on the FAQ page. This policy ensures all participants are guaranteed equal access to information concerning FRVT 2002. Contact with the FRVT 2002 Liason other than via the FAQ areas, or contact with someone other than the non-FRVT Liason will only be allowed under extenuating circumstances.

Chapter 2

Signature Sets

A primary goal of the HEF is to have a well-defined means of marking up sets of biometric data. This section describes the markup of “signature sets,” which list the biometric data from a group of individuals. For greater flexibility, HEF provides not a single schema, but a family of schemas that may be used to validating different kinds of signature sets. Using a set of schemas allows applications to validate signature sets using domain-specific criteria, while still maintaining a high-level, or “global” consistency.

2.1 Terminology

In the HEF model, the following terminology describes biometric information at different grouping levels. Each human subject of interest is an *individual*. A collection of biometric data for a single individual makes up a *signature*. A collection of signatures constitutes a *signature set*.

For a given individual, a particular biometric recording event corresponds to a *sigmember*. An event in this context is typically a time-localized period during which the subject is imaged. Since a subject may have many biometric recordings, a single *signature* can contain one or more *sigmembers*. For example, in the FERET database, there are images of some subjects taken several months apart. In this case, each image is a different *sigmember*. In the general case a *signature* will be comprised of *sigmembers* that contain heterogenous biometric data; for example, a fingerprint and a mugshot.

A *sigmember*, or recording event, could also contain one or more data components. For example, a stereoscopic video might consist of two (simultaneously captured) video sequences. A *dataset* corresponds to a logical component of a biometric recording. It follows that a *sigmember* may contain one or more *datasets*. The precise definition of a dataset is expected to change according to the mode of biometric. For most biometrics, however, a single dataset is often sufficient.

Under the HEF, it is assumed that for each *dataset*, there exists a set of one or more files containing the raw biometric data of interest. Therefore, each *dataset* may contain one or more *files*. Each *file* corresponds to a data file that contains biometric data. Note that the HEF does not attempt to restrict the permissible file formats (JPEG, PNG, MPEG, AVI, VRML, MP3 etc) in any way.

2.2 Detailed Example

To illustrate the above terminology, consider the following examples. Suppose we have biometric data on three different subjects, and we wish to create a structured document that describes this data.

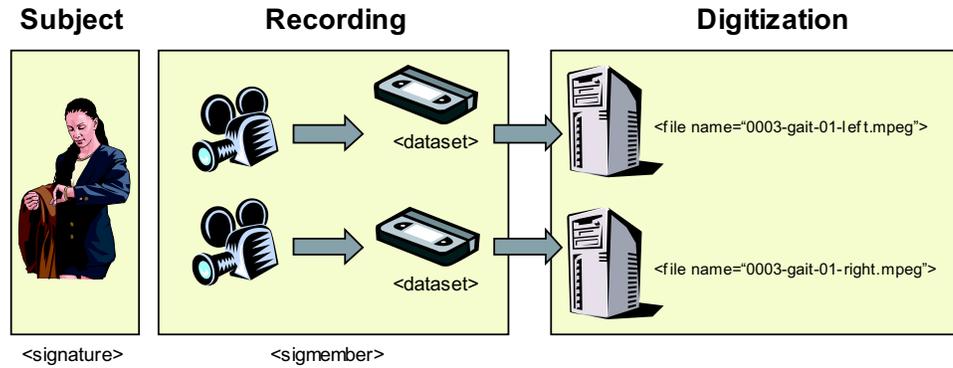


Figure 2.1: An illustration of the differences between subject 0003's *signature*, *sigmember*, *dataset*, and *file* elements.

- **Subject 0001.** Suppose the data to be used for subject 0001 is a **single facial image**. The signature for subject 0001 has a single sigmember, with a single dataset, with a single file that contains an image of subject 0001's face.
- **Subject 0002.** Suppose the data to be used for subject 0002 is a **short video clip**. The video, however, is not stored in a single file, but as a collection of five individual frames or images. The signature for subject 0002 has a single sigmember, with a single dataset, with five files that each contain a different frame. For this subject, there is only one sigmember since the video clip is from a single recording event. There is also only one dataset, since the individual frames, are part of the a larger logical component — the "video."
- **Subject 0003.** Suppose the data to be used for subject 0003 includes an **iris scan**, **three facial images** each taken on different days, and a **stereoscopic gait video**. The signature for subject 0003 therefore contains five sigmembers (one for the iris scan, three for each facial image, and one for the gait video). For the first, iris scan, sigmember, there is a single dataset with a single file that contains the iris data. Three sigmembers, for the facial imagery, each have a single dataset, each with a single file that contains facial images. The fifth sigmember (the gait video) has two datasets — one for each video. The datasets would each have a single file if the data was encoded in a single video (such as an MPEG), or would have a collection of files, where each file corresponded to a particular frame. See Figure 2.1 for an illustration.

In Section 2.2.1, we describe an example document marking up the biometric information for these three synthetic signatures. After the overview, we give a detailed treatment of each element and attribute.

2.2.1 Document Overview

Figures 2.2–2.3 shows an example of how one would mark up the above information using the HEF signature set schemas. The non-whitespace lines have been numbered so that we can easily reference the different parts of the document — a real signature set document would not have them.

The first line of the document is a standard XML header, used to indicate that the document is indeed a (version 1.0) XML document. If there is any whitespace before this header, then the instance document would not validate successfully.

Lines 2 through 5 and line 60 compose the opening and closing root elements of the document. The `<signature-set>` element also contains several attributes used for XML namespace "bookkeeping". We will briefly describe them, but for most signature sets, copying these lines verbatim will most likely be sufficient — it is not necessary to completely understand them. On line 2, the `xmlns` attribute defines the "target" namespace of the document. This associates the elements of the document with the string, or

“namespace” *http://www.nist.gov/humanid/hef/xml/0.99.0*. On line 3, the *xmlns:xsi* attribute associates the prefix *xsi* with the standard name *http://www.w3.org/2001/XMLSchema-instance*. This is a special string “understood” by XML parsers which allow the prefix *xsi* to be used to access XML Schema elements and attributes. Lines 4 and 5 indicate the location of the signature set schema. It is a pair that consists of a namespace (which matches the *xmlns* value) and the schema filename itself. For most XML parsers (like Xerces), if the schema filename is indicated with a relative path (such as line 5 in our example), then the base path is the directory in which the instance document is located, *not* the working directory of the application. For example, if our example instance document was located in */usr/local/HEF/example*, then the XML parser would expect to find the schema in */usr/local/HEF/example/0.99.0/sigset-schemas/sigset-unrestricted.xsd*.

Lines 7 through 13 describe the biometric information for the first subject — subject 0001. This signature contains a single sigmember (lines 8 and 12), a single dataset (lines 9 and 11), and a single file (line 10). As indicated by lines 8 and 9, the data is a JPEG facial still.

Lines 14 through 25 describe the biometric information for the second individual. Subject 0002’s signature consists of a single video recording, their signature contains a single sigmember (line 15, 16, and 24), and a single dataset (lines 17 and 23). Each frame of the video clip, however, is stored in its own file, and therefore, the dataset (line 17) contains multiple file elements (lines 18–22).

Lines 26 through 59 describe the biometric information for the third subject — subject 0003. Recall that for subject 0003 there are five separate recordings, and therefore, five separate sigmembers. Lines 27–31 are a sigmember for the iris recording information. It contains a single dataset (lines 28 and 30) and a single file (line 29). Lines 32 through 47 mark up the next three sigmembers. Since they are face stills, they are similar in structure to subject 0001’s sigmember (lines 8–12). We reiterate that because they are stills taken at different times — i.e., they are different recordings — each face still is a different sigmember (as opposed to different datasets or different files). Finally, lines 48–57 describe the stereoscopic gait video. There is only one sigmember (lines 48–57) since both videos were taken during a single biometric recording event. Each video, however, is a logical component of this recording, and therefore each corresponds to a dataset element (lines 49–52 and lines 53–56). The files themselves (one for each video) are marked up in lines 50–51 and lines 54–55).

2.3 Document Structure

Because of the wide variety of biometric systems, and the varying nature of the constraints that a recognition system may want to apply on a set of signatures, the HEF includes a family of related schemas that can be used “as is” for face recognition systems, or easily extended to accommodate new ones. HEF uses *derived types* (see Section 2.3.3) as its main vehicle for accomplishing this kind of flexibility.

Although the HEF schemas will ultimately be available online, they have been designed primarily for standalone operation (no network or web client/server access). The *sigset-general.xsd* schema is the main schema which defines the base and abstract types used in different signature sets. Because it contains abstract types it is not meant to be used directly by an instance document. In this section, unless otherwise specified we are referring to the schema defined by *sigset-unrestricted.xsd*, which is a relatively unrestricted, non-abstract redefinition of *sigset-general.xsd*. This schema is also referred to in this document, and in the schema files themselves as the **HEF Unrestricted Sigset Schema**.

The structure of an HEF signature set document follows directly from the above grouping terminology. Each term corresponds to a tag, and the containment relationship is represented by the nesting of elements. For example, since a signature set contains multiple signatures, the *<signature-set>* may have one or more *<signature>* elements as child elements — a *<signature>* element may have one or more *<sigmember>* elements as children — and so on.

In the remainder of this section we detail the role of each element in a signature set document. To aid understanding, we will incrementally recreate the example signature set, one element hierarchy at a time.

```

1:  <?xml version="1.0" encoding="UTF-8"?>
2:  <signature-set xmlns="http://www.nist.gov/humanid/hef/xml/0.99.0"
3:      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4:      xsi:schemaLocation="http://www.nist.gov/humanid/hef/xml/0.99.0
5:                          0.99.0/sigset-schemas/sigset-unrestricted.xsd"
6:      name="example set">
7:    <signature name="signature 00" subject_id="0001">
8:      <sigmember xsi:type="simple-face-still-type" modality="simple face still">
9:        <dataset media="digital still" type="jpeg">
10:         <file name="0001-01.jpg"/>
11:        </dataset>
12:      </sigmember>
13:    </signature>
14:    <signature name="signature 01" subject_id="0002">
15:      <sigmember xsi:type="simple-multifile-face-video-type"
16:        modality="simple multifile face video">
17:        <dataset>
18:         <file name="0002-01-frame00.jpg"/>
19:         <file name="0002-01-frame01.jpg"/>
20:         <file name="0002-01-frame02.jpg"/>
21:         <file name="0002-01-frame03.jpg"/>
22:         <file name="0002-01-frame04.jpg"/>
23:        </dataset>
24:      </sigmember>
25:    </signature>
26:    <signature name="signature 02" subject_id="0003">
27:      <sigmember modality="iris scan">
28:        <dataset media="digital still">
29:         <file name="0003-iris.dat"/>
30:        </dataset>
31:      </sigmember>
32:      <sigmember xsi:type="simple-face-still-type" modality="simple face still">
33:        <dataset media="digital still" type="jpeg">
34:         <file name="0003-01.jpg"/>
35:        </dataset>
36:      </sigmember>
37:      <sigmember xsi:type="simple-face-still-type" modality="simple face still">
38:        <dataset media="digital still" type="jpeg">
39:         <file name="0003-02.jpg"/>
40:        </dataset>
41:      </sigmember>
42:      <sigmember xsi:type="simple-face-still-type" modality="simple face still">
43:        <dataset media="digital still" type="jpeg">
44:         <file xsi:type="spatial-file-type" name="0003-03.jpg"
45:           roi="(10,30) (125,110)"/>
46:        </dataset>
47:      </sigmember>
48:    :
49:    :

```

Figure 2.2: The completed XML markup of the example data (continued in Figure 2.3).

```

:
:
48:   <sigmember modality="stereo gait video">
49:     <dataset media="digital video" type="mpeg">
50:       <file xsi:type="temporal-file-type" name="0003-gait-01-left.mpeg"
51:         start="10" stop="230" unit="frame">
52:     </dataset>
53:     <dataset media="digital video" type="mpeg">
54:       <file xsi:type="temporal-file-type" name="0003-gait-01-right.mpeg"
55:         start="10" stop="230" unit="frame">
56:     </dataset>
57:   </sigmember>
58: </signature>
60: </signature-set>

```

Figure 2.3: The completed XML markup of the example data (continued from Figure 2.2).

2.3.1 Signature Set Elements

The root tag of every valid signature set document must be `<signature-set>`. The optional *name* attribute is a token that may be used to name the signature set, so that an application can refer to a particular signature set by using its name, as opposed to its filename, or other characteristic. A valid `<signature-set>` must have one or more valid child `<signature>` elements. If a `<signature-set>` does not have at least one `<signature>` child element, then the document will not validate successfully.

2.3.2 Signature Elements

Each `<signature>` corresponds to a collection of biometric information for a single individual. Each signature must have a unique *name* which is a token that can be used to refer to a particular signature. An optional attribute *subject_id* is a token associated with a particular *subject*, as opposed to a *signature*. It is important to understand the difference between *name* and *subject_id*. Given the nature of an evaluation, the *name* may or may not contain information about the subjects true identity. In an external evaluation, there may be a need to hide a subjects identity from a recognition system, and it would be expected that the *subject_id* attribute not be provided. However, there is still a need to provide handles for specific signatures — this is what the *name* attribute is for.

To ensure that both signatures and individuals can be referred to without ambiguity, within a single signature set file, no pair of `<signature>` elements may share the same *name* value. A valid `<signature>` contains one or more valid `<sigmember>` elements. The enforcement of uniqueness constraints across different files is an option of an application.

Figure 2.4 is an example of a valid use of `<signature-set>` and `<signature>` tags based on the example signatures discussed in Section (2.1). In these examples, an ellipsis (...) represents information to be added later — they are not actually contained in the instance document itself. Since our example uses three different subjects, the root `signature-set` element has three `signature` children elements. Notice how each subject has its own set of signature tags, and each signature has its own unique *name* and *subject_id*.

2.3.3 XML Schema Types

This subsection provides some background material on XML Schemas that will aid understanding of the nature of certain HEF XML constructions. If you have experience using and writing XML Schemas, you may want to skip to the next section.

A feature of XML Schemas is that each element and attribute is associated with a particular *type*. This allows the validation process to enforce *type constraints* on instance documents, much as a compiler enforces type constraints on source code. For example, in an XML Schema specifies that all elements

```

<signature-set name="example set">
  <signature name="signature 00" subject_id="0001">
    :
  </signature>
  <signature name="signature 01" subject_id="0002">
    :
  </signature>
  <signature name="signature 02" subject_id="0003">
    :
  </signature>
</signature-set>

```

Figure 2.4: A valid nesting of *<signature-set>* and *<signatures>*.

must be of type *decimal*, and an instance document contains the fragment “*<temperature>*Herr Dr. Merkwuerdigichliebe *</temperature>*” will clearly not validate successfully, since “Herr Dr. Merkwuerdigichliebe” is not a decimal value, such as “37.0”.

Types in XML Schemas are not limited to such simple schemes, however. Users can generate their own complex types, create type hierarchies, and redefine new types as extensions or restricted versions of other types — much like the object oriented concept of inheritance. In XML Schemas, the “parent” or “superclass” is referred to as the *base* type, and the “derived”, “child”, or “subclass” is called the *derived* type.

Consider the following example (adapted from the XML Schema Primer). We start with the compound *address-type* type:

```

<complexType name="address-type">
  <sequence>
    <element name="name" type="string"/>
    <element name="street" type="string"/>
    <element name="city" type="string"/>
  </sequence>
</complexType>

```

We can then use *address-type* as a base class for a new derived type *us-address-type*, which extends *address-type* by including a “state” and “zip” field.

```

<complexType name="us-address-type">
  <complexContent>
    <extension base="address-type">
      <sequence>
        <element name="state" type="string"/>
        <element name="zip" type="positiveInteger"/>
      </sequence>
    </extension>
  </complexContent>
</complexType>

```

The type *us-address-type* is an “address derived type” of base type *address-type*. It has “name,” “street,” and “city,” elements, just as *address-type*, but “state” and “zip” elements as well.

With XML Schemas, the association of element names and types does not occur automatically. For example, to have an element *<bill-to>* of type *address-type*, within a schema, we must explicitly tie them together. Because of this lack of implicit coupling with XML Schemas, in some cases, it is possible to associate multiple types with the same element, where the nature of the enforced type changes according to some explicit instruction in an instance document, or because of the context of its occurrence. We will not show how to accomplish this here. Instead, we will show how to indicate which type an element

should be validated against, in the case there an element could be associated with potentially multiple, and therefore ambiguous, types.

To explicitly indicate that an element should be validated with respect to a specific type, one simply needs to set the *xsi:type* attribute to the name of desired type. The *xsi:type* is a special attribute shared by *all* XML elements, and it not parsed in the same way user-defined attributes are — it is used by the parser to indicate that a different type should be used for validating a particular element.

For example, suppose we had an XML document where the `<bill-to>` element is of type *address-type*. The following fragment *would* be valid XML

```
<bill-to>
  <name>Lionel Mandrake</name>
  <street>Burpleson Airforce Base</street>
  <city>Burpleson</city>
</bill-to>
```

However, the fragment

```
<bill-to>
  <name>Lionel Mandrake</name>
  <street>Burpleson Airforce Base</street>
  <city>Burpleson</city>
  <state>NV</state>
  <zip>12345</zip>
</bill-to>
```

would *not* validate successfully, since `<state>` and `<zip>` are only appropriate elements for *us-address-type*, and not *address-type*. However, by using the *xsi:type* attribute, we can indicate to the parser that the *us-address-type* should be used, instead of the *address-type*. This is valid, since *us-address-type* is of base type *address-type*, which satisfies the original schema requirement that `<bill-to>` be of type *address-type*. The fragment

```
<bill-to xsi:type="us-address-type">
  <name>Lionel Mandrake</name>
  <street>Burpleson Airforce Base</street>
  <city>Burpleson</city>
  <state>NV</state>
  <zip>12345</zip>
</bill-to>
```

would now validate successfully. Again, note that in the *xsi:type* is defined in neither *address-type* nor *us-address-type* — it is a “global” attribute common to all XML schema elements.

2.3.4 Sigmember Elements

Each `<sigmember>` corresponds to a particular biometric recording event — the process of collecting new biometric information about a subject. There are two kinds of `<sigmember>` types. First, a `<sigmember>` element may be of the base type *sigmember-type* — a generic type that can be used to describe a recording of arbitrary mode. The other types are defined as restricted derivations of *sigmember-type*. Currently, there are three such types, *simple-face-still-type*, *simple-face-video-type* and *simple-multifile-face-video-type*.

- The *simple-face-still-type* is designed to be the preferred type for describing facial images stored in a single data file. Data that might be described with this type include digital stills, or scanned photographs.
- The *simple-face-video-type* is designed to be the preferred type for describing facial video that is stored in a single data file. An MPEG video clip might be described with this type.

- The *simple-multifile-face-video-type* is designed to be the preferred type for describing facial video stored in multiple data files, typically one image or datafile per frame of video.

All of these types are accessed through the use of the `<sigmember>` tag, but require the use of the special attribute `xsi:type` to indicate, to the XML parser, that they are not of the base type `<sigmember-type>`, but of a specified derived type. Typically, the derived sigmember types versions of the base type `sigmember-type`, but with additional constraints. If the derived types are insufficient for describing your own biometric data, either the base type `sigmember-type` can be used, or a new derived type could be written.

The required *modality* attribute describes the modality of the recording. Do not confuse this with the `xsi:type` attribute, which indicates to the XML parser, the proper derived type to use when validating the document. Typically, the restricted derivations of the `sigmember-type` (e.g., *simple-face-still-type*) will require that the *modality* attribute take on a specific value. Note that in an XML file, there is no implicit typecasting when setting the *modality* attribute to the value required in the derived sigmember. To validate according to the derived type, both `xsi:type` and the *modality* attribute are required. It is *highly* recommended that whenever possible, you use reserved/predefined type values only when using derived sigmember types.

All `<sigmember>` elements may also use the optional *metadata* attribute. This attribute is a token that an application can use to reference recording meta-information, such as sensor information, persons involved in the data collection, contact information, and so on. Although it is recommended that the value for the *metadata* attribute be used for filenames, there is no requirement to do so.

Returning to our common example, Figure 2.5 shows the preferred way of describing our example biometric data with `<sigmember>` tags. For subject 0001, we choose to use the *simple-face-still-type*. Notice how both that `xsi:type` and *modality* attributes are included. The `xsi:type` indicates to the parser (in the same way that a pragma indicates to a compiler) that this particular sigmember is of derived type *simple-face-still-type*, and the *modality* indicates to our application that the sigmember is a simple face still. The *simple-face-still-type* is defined in such a way that the document will not validate successfully if the *modality* attribute does *not* have the value “simple face still.” Due to limitations in the current version of XML schemas, setting the *modality* to “simple face still” cannot lead to coercion of the element from the base type `sigmember-type` to the derived type *simple-face-still-type*. Therefore, as mentioned previously, to avoid confusion, setting the value of *modality* to “simple face still” when using the default type `sigmember-type` is not recommended.

2.3.5 Dataset Elements

Every valid `<sigmember>` must contain at least one `<dataset>`, where each `<dataset>` corresponds to some logical component of a biometric recording. Naturally, all `<dataset>`s with a common parent should correspond to the same biometric recording. The precise definition of dataset is expected to change according to the mode of the biometric. For most biometrics, such as a facial image still, it is expected that a `<sigmember>` has a single `<dataset>`. It is anticipated that more complex recordings would contain more than one `<dataset>` (for example, one dataset per camera in the case of stereoscopic video). Until more derived sigmembers are added to HEF, it is expected that HEF users will define their own conventions for determining the nature and quantities of datasets for their own biometric recordings.

There are two optional attributes that may be used with the `<dataset>` element. The *media* attribute is a token which describes the media of the original recording — “digital still”, “35mm film”, or “Hi-8 video” for example. The *datatype* attribute is a token which should be used to describe the format of the recorded data — “JPEG”, “MPEG”, “TIFF”, and so on. Like most of the optional attributes, there is no strict convention for the values of *media* and *datatype*.

Figure 2.6 shows the proper nesting of `<dataset>` tags for our common example. Notice that except for the stereoscopic gait video, all `<sigmember>`s have a single `<dataset>`.

```

<signature-set name="example set">
  <signature name="signature 00" subject_id="0001">
    <sigmember xsi:type="simple-face-still-type"
      modality="simple face still">
      :
    </sigmember>
  </signature>

  <signature name="signature 01" subject_id="0002">
    <sigmember xsi:type="simple-multifile-face-video-type"
      modality="simple multifile face video">
      :
    </sigmember>
  </signature>

  <signature name="signature 02" subject_id="0003">
    <sigmember modality="iris scan">
      :
    </sigmember>
    <sigmember xsi:type="simple-face-still-type" type="simple face still">
      :
    </sigmember>
    <sigmember xsi:type="simple-face-still-type" type="simple face still">
      :
    </sigmember>
    <sigmember xsi:type="simple-face-still-type" type="simple face still">
      :
    </sigmember>
    <sigmember type="stereo gait video">
      :
    </sigmember>
  </signature>
</signature-set>

```

Figure 2.5: A valid nesting of <signature-set>, <signature> and <sigmember> tags.

```

<signature-set name="example set">
  <signature name="signature 00" subject_id="0001">
    <sigmember xsi:type="simple-face-still-type" modality="simple face still">
      <dataset media="digital still" type="jpeg">
        :
      </dataset>
    </sigmember>
  </signature>
  <signature name="signature 01" subject_id="0002">
    <sigmember xsi:type="simple-multifile-face-video-type
      modality="simple multifile face video">
      <dataset>
        :
      </dataset>
    </sigmember>
  </signature>
  <signature name="signature 02" subject_id="0003">
    <sigmember modality="iris scan">
      <dataset>
        :
      </dataset>
    </sigmember>
    <sigmember xsi:type="simple-face-still-type" modality="simple face still">
      :
    </dataset>
  </sigmember>
  <sigmember xsi:type="simple-face-still-type" modality="simple face still">
    <dataset media="digital still" type="jpeg">
      :
    </dataset>
  </sigmember>
  <sigmember xsi:type="simple-face-still-type" modality="simple face still">
    :
  </sigmember>
  <sigmember modality="stereo gait video">
    <dataset media="digital video" type="mpeg">
      :
    </dataset>
    <dataset media="digital video" type="mpeg">
      :
    </dataset>
  </sigmember>
</signature>
</signature-set>

```

Figure 2.6: A valid nesting of `<signature-set>`, `<signature>`, `<sigmember>`, and `<dataset>` tags.

2.3.6 File Elements

Finally, we reach the bottommost tag, the `<file>` element. Each `<file>` element is terminal — if it contains any child elements, then the document will not validate successfully. Each file element is used to refer to a particular datafile, which, depending on the context of the parent tags, may contain complete or partial data of a biometric recording.

Like `<sigmember>`, the `<file>` has a base and derived types. The base type, `file-type` is designed for general use, and has a single required attribute `name` which is a token that indicates the biometric data's file name. The two derived types provide mechanisms for indicating to an application more information about data of interest. For example, there may be several faces in a digital still, or several people captured in a gait video, but only one spatiotemporal region of interest.

The `spatial-file-type` extends the base type by offering an optional `roi` attribute, which is a token that can be used to specify a region of interest within the data file. The `temporal-file-type` is designed to be the preferred type for describing recordings with temporal information. This type requires the `start` and `stop` attributes, which are integer indexes that indicate to an application the logical beginning and ending of the biometric data. The optional `unit` attribute is a token to indicate the unit of measure to associate with the start and stop indexes. For example, in a video sequence, the unit value may be "frame". A type of `temporal-file-type` may also use the `roi` attribute.

Figure (2.2) shows a complete description of our example data, using all of the required tags, and the proper XML headers. There is only a single `<file>` tag for subject 0001, since there is only one facial image available for this subject. Subject 0002 has five `<file>` tags, all part of a single `<dataset>`, each corresponding to a single frame of the (short) video clip. The use of `<file>` tags for subject 0003 is more complex. The iris scan, and the first two facial images use a plain `<file>`. In the third face still, instead of using the default base type `file-type`, the derived type `spatial-file-type` is used, as well as a region of interest, or `roi` attribute. Without indicating the use of a derived type through the `xsi:type` tag, using an `roi` attribute would result in failure to validate correctly. Similarly, the derived type `temporal-file-type` is used to mark up the two gait videos.

2.4 Supported Schemas

As mentioned in Section (2), the HEF is not limited to a single schema, but a family of schemas. So far, we have limited our attention to the schema defined in `sigset-unrestricted.xsd`, since it is the most general schema in the family. However, there are several other schemas available for situations in which stricter constraints are desired. These schemas are briefly described here. For complete information, see the extensively commented schema files themselves.

- **HEF Unrestricted Sigset Schema.** The schema defined in `sigset-unrestricted.xsd`, and discussed in the previous section, defines the most general form of signature set. With this schema, signature sets may contain any (non-zero) number of signatures of any sigmember type. Unless otherwise specified within the document, each sigmember type may have any (non-zero) number and manner of datasets, each of which may have any (non-zero) number and manner of files.
- **HEF Single Sigmember Schema.** The schema defined in `sigset-single-sigmember.xsd` will validate documents in which every signature has one (and only one) sigmember. All documents that validate under this schema, should also validate under the HEF Unrestricted Sigset Schema.
- **HEF Simple Face Still Set.** The schema defined in `simple-face-still-set.xsd` will validate documents in which every signature has a single sigmember, and each sigmember is of derived type `simple-face-still-type`. This schema is designed to be used for evaluating most facial recognition systems that use still imagery. All documents that validate under this schema, should also validate under the HEF Unrestricted Sigset Schema.
- **HEF Multiple Simple Face Stills Set.** The schema defined in `multiple-simple-face-stills-set.xsd` will validate document in which every signature has one or more sigmembers where each sigmember

is of derived type *simple-face-still-type*. This schema is designed to be used for evaluating most facial recognition systems capable of using multiple stills of a single subject. Note that the multiplicity is at the *sigmember* level, indicating that the multiple images of a subject correspond to *different* recording events. All documents that validate under this schema, should also validate under the HEF Unrestricted Sigset Schema.

- **HEF Simple Multifile Face Video Set.** The schema defined in *simple-multifile-face-video-set.xsd* will validate documents in which every *sigmember* is of derived type *simple-multifile-face-video-type*. This schema is intended to be used for evaluation of facial recognition systems that can operate on a video sequence where each frame has been stored in a separate file. All documents that validate under this schema, should also validate under the HEF Unrestricted Sigset Schema.

2.5 Built-In Derived Sigmember Types

This section provides information about the built-in derived sigmember types *simple-face-still-type*, *simple-face-video-type*, and *simple-multifile-face-video-type*.

2.5.1 HEF Simple Face Still Type

The **HEF Simple Face Still Type**, is defined in the file *simple-face-still-type.xsd* and, in instance documents, should be referred with the token "*simple-face-still-type*". This derived type is designed to be the preferred type for representing facial stills. A *<sigmember>* of this type must have the attribute *xsi:type* set to the value "*simple-face-still-type*" and attribute *modality* set to "*simple face still*". Valid *<sigmember>* elements of type *simple-face-still-type* must only contain one *<dataset>* element, which must contain a single *<file>* of type *spatial-file-type*. Because this is specified in the schema, there is no need to explicitly set the *<file>*'s *xsi:type* attribute in the instance document.

2.5.2 HEF Simple Face Video Type

The **HEF Simple Face Video Type**, is defined in the file *simple-face-video-type.xsd* and, in instance documents, should be referred to with the token "*simple-face-video-type*". This derived type is designed to be the preferred type for representing facial video stored in a single file. A *<sigmember>* of this type must have the attribute *xsi:type* set to the value "*simple-face-video-type*" and attribute *modality* set to "*simple face video*". Valid *<sigmember>* elements of type *simple-face-video-type* must contain one *<dataset>* element, which must contain a single *<file>* of type *temporal-file-type*. Because this is specified in the schema, there is no need to explicitly set the *<file>*'s *xsi:type* attribute in the instance document.

2.5.3 HEF Simple Multifile Face Video Type

The **HEF Simple Multifile Face Video Type**, is defined in the file *simple-multifile-face-video-type.xsd* and, in instance documents, should be referred to with the token "*simple-multifile-face-video-type*". This derived type is designed to be the preferred type for representing facial video where each frame is stored in a separate data file. A *<sigmember>* of this type must have the attribute *xsi:type* set to the value "*simple-face-video-type*" and attribute *modality* set to "*simple face video*". Valid *<sigmember>* elements of type *simple-face-video-type* must contain one *<dataset>* element, which must contain a one or more *<file>*, each of type *spatial-file-type*. Because this is specified in the schema, there is no need to explicitly set the *<file>*'s *xsi:type* attribute in the instance document.

2.6 Custom Derived Sigmember Types

In the future this section will provide instructions on how to create your own derived sigmember types, and package them so that they can be provided in the next formal release of HEF. It will also discuss the schema files themselves, including the overall design of the schema family, the files structure of the signature set schema family, and the purpose of each such file.

Chapter 3

Similarity Scores

While the previous chapter was concerned primarily with marking up standardized *input* for evaluations, this brief chapter defines the XML format for describing a regular *output* of a face recognition system.

In the context of the HEF, a system consists of a black box recognition engine that takes two XML signature sets, the enrolled and the “unknown” test samples, and produces a set of pairwise identification scores. More formally, the protocol requires that an algorithm compare all elements of a *target set*, \mathcal{T} , with those of a *query set*, \mathcal{Q} , and report a similarity matrix, S , whose elements, s_{ij} , state some measure of the sameness of the identity of the individual in the i -th element of \mathcal{T} and the j -th element of \mathcal{Q} . Because subsets of \mathcal{T} and \mathcal{Q} have different subject- or acquisition-specific properties, certain elements of S , corresponding to *gallery* (i.e. enrolled), $\mathcal{T}_{\text{gallery}}$, and *probe* (i.e. unknown) sets, $\mathcal{P} \subseteq \mathcal{Q}$, may be used retroactively to make performance estimates on targeted recognition tasks (e.g. in face recognition, say, image acquisition issues like resolution or subject specific changes in, say, eye-wear or facial hair).

3.1 Document Structure

In this section, we will describe the structure of a similarity score file valid under the “HEF Similarity Set Schema”. Fortunately, the structure of a file containing similarity scores is much simpler than for signature sets. Instead of a family of schemas, all valid similarity sets use the single *similarity-set.xsd*.

Because of the potential for running large experiments, and therefore, obtaining millions of similarity scores, similarity data from a given test may be stored in multiple files. In a *standalone* similarity set, all similarity data is stored in one large, monolithic file. In a *multifile* similarity set, data is stored within a collection of files, where each file contains the similarity scores for one a particular query signature. It is also possible, but not recommended, to have a hybrid of both types.

3.1.1 Detailed Example

An example will best illustrate the structure of a similarity set document. Suppose our target set consists of our three example signatures — “signature 00”, “signature 01”, and “signature 02” — and our query sets contains three unknown signatures of the same subjects — “signature 00a”, “signature 01a”, and “signature 02a”. Our similarity set file will have nine similarity scores (three targets, three queries). Figure 3.1 shows an example of how these scores would be marked up under the HEF XML formats.

Like the document from section (2.2.1) the first line of the document is the standard XML header, which indicates that the document conforms to the XML (version 1.0) standard. Leading whitespace or blank lines would render the instance document invalid.

Lines 2 through 6 and line 27 compose the opening and closing root elements of the similarity set document. The `<similarity-set>` contains the same XML namespace “bookkeeping” documents as discussed in section (2.2.1), and, as indicated by line 5, uses the schema defined in the file `0.99.0/similarity-schemas/similarity-set.xsd`. Line 6 indicates that this similarity set file does not depend on the presence of any other similarity file — more on this later.

In a similarity set document, the similarity scores are grouped by *query signature*. In other words, a particular query signature is “fixed”, and the scores from all pairwise comparisons involving that query signature are indicated in the enclosing children tags. For example, since there are three signatures in the query signature set, there are three `<similarity>` elements in the document — the scores where the query signature is “signature 00” are marked up on lines 7 through 14, the “signature 01” query signature is marked up on lines 14 through 20, and the remaining “signature 02” query signature on lines 21 through 27.

Each `<similarity>` element has a *query* attribute that indicates which query signature is fixed. Line 7 “fixes” the query signature to “signature 00”, line 14 to “signature 01” and line 21 to “signature 02”. Each `<similarity>` element also has a `<values>` element which gives information about the nature of the similarity scores. For a set of “distance” scores, such as lines 9–11 and lines 23–25, lower numbers indicate stronger evidence that two signatures are from the *same* subject. For a set of “similarity” scores, like those in lines 16–18, lower values indicate that a two signatures are from two *different* subjects. The similarity or distance values themselves are marked up using the `<s>` element (lines 9–11, 16–18, 23–25h). For example, line 9 indicates that the “distance” between target signature “signature 00a” and the query signature “signature 00” is 0.02342, line 18 indicates that “similarity” between target signature “signature 02a” and query signature “signature 01” is 2.87322, and so on.

In Figure 3.1 all of the similarity scores were stored in a single file. When an evaluation consists of hundreds or thousands of signatures, storing all the results in a single file, may become unfeasible. Therefore, there is a mechanism that allows the similarity set to be broken into a collection of files, where each file contains the values for a single query signature. The structure of such a multifile similarity set is straightforward. Essentially, it involves replacing all the `<values>` elements with `<file>` elements. Each `<file>` element points to a separate XML file where the root element is a `<similarity>` tag, instead of a `<similarity-set>` tag.

Figures 3.2–3.5 show a multifile version of our example similarity set document. The similarity set document of Figure 3.2 is the same as Figure 3.1 except that the `<values>` elements (lines 8–12, 15–19, 22–26 in Figure 3.1) are replaced with `<file>` elements (lines 8, 11, and 14). These `<file>` elements point to separate XML files, *signature_00.xml* (Figure 3.3), *signature_01.xml* (Figure 3.4), and *signature_02.xml* (Figure 3.5). Note that Figures 3.3–3.5 are using the *similarity.xsd* schema, and not the *similarity-set.xsd* schema — each of these files have a single `<similarity>` element at their root.

3.1.2 Similarity Set Elements

The root of a similarity set document must be a single `<similarity-set>` element. A valid `<similarity-set>` element must have one or more valid `<similarity>` as child elements. The *standalone* attribute is a boolean used as a “hint” to indicate if, for a given test, the file contains all available similarity data. In the current similarity set schema, the value of this attribute does not effect the behavior of the validator (there is nothing encoded in the schema that may enforce a standalone constraint).

3.1.3 Similarity Elements

A similarity set document consists of a collection of `<similarity>` elements. Each `<similarity>` element corresponds to a collection of pairwise comparisons, where a particular query signature is fixed, and compared against all elements in the target set. For convenience, `<similarity-set>` elements may contain an optional *name* attribute which is a token that may be used to name the similarity set. A `<similarity-set>` element may also contain the *standalone* attribute, which when set to *true* indicates that the file contains all available similarity data.

All `<similarity>` elements must contain a `query` attribute, which is a token that indicates the *name* of the query signature that is being pairwise compared over all the target (a.k.a. gallery, or enrolled) signatures. Note that the signature *name* should be used here, *not* the `subject_id` since, by definition, within a signature set, all signatures are assigned a unique name, but may not even contain a value for the `subject_id`.

A `<similarity>` element may also contain the optional `meta` attribute, which is a token that may be used to point to a file that contains meta-information about this particular set of comparisons (such as algorithm parameters, training information, etc.).

3.1.4 Values and File Elements

Each `<similarity>` element must have either a single valid `<values>` element, or, a single valid `<file>` element. The `<values>` element is used to indicate to the parser, that the raw similarity scores are to follow. Using a `<file>` element, however, indicates to the parser that this is *not* a standalone document, and that for the query signature indicated by the parent element, the similarity scores are stored in a separate file. It is expected, that this separate file will be a valid **HEF Similarity** (not **HEF Similarity Set**) document. A **HEF Similarity** document is a subset of the **HEF Similarity Set** schema. It is identical to a **HEF Similarity Set** document, except the root element must be a `<signature>` element, and not a `<signature-set>` element. For multifile similarity sets, the naming convention for `<file>` elements and file names is application dependent.

For both standalone and multifile similarity sets, the raw similarity scores corresponding to a particular query signature are found in the child elements of a `<values>` element. Each `<values>` tag must have `polarity` and `sort` attributes. The `polarity` attribute must take on either the value “distance” or “similarity” depending on the nature of the comparison score returned by a particular recognition engine. If a large positive value indicates that two signatures are *different*, then `<polarity>` should be set to “distance”. If a large positive value indicates that two signatures are *similar*, then `<polarity>` should be set to “similarity”. The tag `sort` is used to indicate the order of the values contained in the child elements. Permitted values for `sort` are “ascending”, “descending” and “unsorted”. In the current version of HEF, the `sort` attribute is more of a “hint” rather than an enforcement. If the order of the values in the child elements does not match that indicated by the value of the `sort` attribute, by the current implementation, the document may still validate successfully.

3.1.5 “s” Elements

A `<values>` element contains one or more `<s>` element, where each `<s>` element corresponds to a pairwise comparison of two signatures. Every `<s>` element must have a `n`, for name, attribute and a `v`, for value, attribute. Each pairwise comparison has a query signature and a target signature. The query signature is indicated by the value of the `query` attribute in the enclosing `similarity` element. The target signature is indicated by the value of the `<s>` element’s `n` attribute. Like the `query` attribute of the `<similarity>` element, the `n` attribute is a token that should refer to the *name* of a signature. The `v` attribute is used for the raw similarity or difference score itself.

```

1: <?xml version="1.0"?>
2: <similarity-set xmlns="http://www.nist.gov/humanid/hef/xml/0.99.0"
3:                 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4:                 xsi:schemaLocation="http://www.nist.gov/humanid/hef/xml/0.99.0
5:                                     0.99.0/similarity-schemas/similarity-set.xsd"
6:                 standalone="true">
7:   <similarity query="signature 00">
8:     <values polarity="distance" sort="ascending">
9:       <s n="signature 00a" v="0.02342"/>
10:      <s n="signature 01a" v="8.44832"/>
11:      <s n="signature 02a" v="9.23437"/>
12:     </values>
13:   </similarity>
14:   <similarity query="signature 01">
15:     <values polarity="similarity" sort="descending">
16:       <s n="signature 01a" v="9.98432"/>
17:       <s n="signature 00a" v="3.34231"/>
18:       <s n="signature 02a" v="2.87322"/>
19:     </values>
20:   </similarity>
21:   <similarity query="signature 02">
22:     <values polarity="distance" sort="unsorted">
23:       <s n="signature 00a" v="5.43432"/>
24:       <s n="signature 01a" v="7.43645"/>
25:       <s n="signature 02a" v="0.44834"/>
26:     </values>
27:   </similarity>
28: </similarity-set>

```

Figure 3.1: A standalone similarity set document for our example data.

```

1: <?xml version="1.0"?>
2: <similarity-set xmlns="http://www.nist.gov/humanid/hef/xml/0.99.0"
3:                 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
4:                 xsi:schemaLocation="http://www.nist.gov/humanid/hef/xml/0.99.0
5:                                     0.99.0/similarity-schemas/similarity-set.xsd"
6:                 standalone="false">
7:   <similarity query="signature 00">
8:     <file name="signature_00.xml"/>
9:   </similarity>
10:  <similarity query="signature 01">
11:    <file name="signature_01.xml"/>
12:  </similarity>
13:  <similarity query="signature 02">
14:    <file name="signature_02.xml"/>
15:  </similarity>
16: </similarity-set>

```

Figure 3.2: A multifile version of the similarity set document for our example data. When an evaluation has a large number of signatures, it is often useful to break the similarity data into separate files.

```

<?xml version="1.0"?>
<similarity xmlns="http://www.nist.gov/humanid/hef/xml/0.99.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.nist.gov/humanid/hef/xml/0.99.0
    0.99.0/similarity-schemas/similarity.xsd"
  query="signature 00">
  <values polarity="distance" sort="ascending">
    <s n="signature 00a" v="0.02342"/>
    <s n="signature 01a" v="8.44832"/>
    <s n="signature 02a" v="9.23437"/>
  </values>
</similarity>

```

Figure 3.3: The contents of file *signature_00.xml* — a component of the non-standalone version of the similarity set document and the similarity scores for the query signature named “signature 00”.

```

<?xml version="1.0"?>
<similarity xmlns="http://www.nist.gov/humanid/hef/xml/0.99.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.nist.gov/humanid/hef/xml/0.99.0
    0.99.0/similarity-schemas/similarity.xsd"
  query="signature 01">
  <values polarity="similarity" sort="descending">
    <s n="signature 01a" v="9.98432"/>
    <s n="signature 00a" v="3.34231"/>
    <s n="signature 02a" v="2.87322"/>
  </values>
</similarity>

```

Figure 3.4: The contents of file *signature_01.xml* — a component of the non-standalone version of the similarity set document and the similarity scores for the query signature named “signature 01”.

```

<?xml version="1.0"?>
<similarity xmlns="http://www.nist.gov/humanid/hef/xml/0.99.0"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://www.nist.gov/humanid/hef/xml/0.99.0
    0.99.0/similarity-schemas/similarity.xsd"
  query="signature 02">
  <values polarity="distance" sort="unsorted">
    <s n="signature 00a" v="5.43432"/>
    <s n="signature 01a" v="7.43645"/>
    <s n="signature 02a" v="0.44834"/>
  </values>
</similarity>

```

Figure 3.5: The contents of file *signature_02.xml* — a component of the non-standalone version of the similarity set document and the similarity scores for the query signature named “signature 02”.

Appendix H

FRVT 2002 Test Plan

Face Recognition Vendor Test 2002

Test Plan

June 25, 2002

1 Introduction

This document describes the procedures that will be followed during the administration of Face Recognition Vendor Test 2002 (FRVT 2002). As described in the FRVT 2002 Implementer's Guide, participants will be given two sets of images to compare, from which they will produce a set of similarity files giving the matching scores for query images compared to target images. Participants will also supply a document describing their tested system(s) and associated costs. The results will be presented in the reports including a FRVT 2002 Final Report.

2 Test Description

2.1 Overview

The FRVT 2002 consists of two parts: the medium and high computational intensity tests¹.

Medium Computational Intensity: This test will require the processing of approximately 7500² signatures. It is divided into two separate parts. The first part measures performance on video style imagery. The second part resembles the FRVT 2000 evaluations using still images taken under different circumstances.

High Computational Intensity: This test will evaluate performance of a very large data set of approximately 121,000 images.

Each test will be performed separately and will have a maximum amount of time for completion. For each test, participants will provide a face recognition system complete with all necessary hardware. Time cannot be saved from one test for use in another test. Tests may be run concurrently. The FRVT 2002 will report results for complete and correctly formatted test submissions. Additionally, results for incomplete submissions may be reported where appropriate. Failure to complete a test will be noted in FRVT 2002 report(s). Results will be provided on a test-by-test basis. Prior to arriving at the test

¹ For a precise description of the medium and high computational intensity tests see the appropriate documentation available at <http://www.frvt.org>.

² The precise number, at the time of writing, remains to be determined.

site, participants will have selected the tests they will take (check the website for the deadline for making such arrangements).

2.2 Personnel

Personnel involved with FRVT 2002 will be test agents and participants. Test agents are the government representatives and contractor assistants proctoring the evaluation. Participants are representatives of the organization whose system is being evaluated. Each organization will be allowed to send up to two people.

Each participant will assign one point of contact (POC) to interface with test agents when receiving images, submitting results, and signing the event checklist shown at the end of this document. A different POC may be assigned after first notifying the test agents. Participants are required to have a POC on site during setup and closeout procedures. They are not required to have a person on site between setup and closeout procedures. However, the FRVT 2002 evaluation team will not inform participants if their machine crashes or power goes out. If the POC is not on site when the time limit is reached to stop their program from running, then the participant will be disqualified for exceeding the time limit. There will be no penalty for finishing early and participants are free to leave the test area at any time after similarity files have been submitted and closeout procedures have been completed.

2.3 Facilities

Testing will take place in Building 194 at NAVSEA Dahlgren. Directions and area information has been sent to participants and is available on the “Test Plan” page of the FRVT 2002 Participant Area web site. Several participants may undergo testing at one time, possibly in the same room.

A voice-only telephone may be available for use by participants. First priority will be given to participants setting up equipment on their first day. Second priority will go to other participants with system problems. Due to the metal covering on building 194, cell phones may not work indoors.

NAVSEA Dahlgren is located approximately one hour from the nearest shopping district, so it is advisable to anticipate any spare parts that may be needed and bring them to the evaluation. FRVT 2002 Organizers recommend bringing your equipment with you. If necessary, parts may be shipped to the following address.

Duane Blackburn
DoD Counterdrug Technology Development Program Office
NSWCDD Code B07
1470 Room 1101
17320 Dahlgren Road
Dahlgren, VA 22448

2.4 Computer Requirements

The following table shows the maximum number of desktop computers allowed for each test.

Test	Desktop Computers Allowed
Medium Computational Intensity	1
High Computational Intensity	3

For the High Computational Intensity Test, the three (maximum) computers may be networked using a participant-supplied local area network (LAN) to perform operations in parallel. Wireless networking will not be allowed.

A desktop computer is defined to be a computer with case dimensions and power requirements similar to commercially available computers for an individual's use. Multiple processors are allowed inside that case and will be counted as a single desktop computer. Other configurations approximating a single desktop may be allowed on a case-by-case basis, but must be coordinated with the FRVT 2002 Liaison prior to participant arriving at the test site (check the website for the deadline for making such arrangements).

Participants must supply power strips to accommodate all hardware associated with submitted systems as well as uninterruptible power supplies, as needed. Systems should be setup and tested in advance so that no design, coding, or debugging is necessary at the evaluation site. Only minor configuration changes such as mapping of USB hard drive letters will be allowed.

Participants must install virus-scanning software on all submitted systems and scan with the latest virus definitions on the first day of their evaluation in the presence of a test agent.

2.5 System Description Documents

Participants will be required to submit a five-page (maximum) document, in electronic form, on the first day of their testing that describes the submitted system(s). This document will be a section in the final report that will be released to the public. This is the participant's opportunity to describe their system(s) and costs to those that may read the FRVT 2002 report. This document must be submitted on the first day of testing. Target and query sets will not be distributed to the participants unless this document has been received. Failure to meet this requirement will effectively disqualify a participant. This document must, at a minimum, adequately address the following:

- Provide an overview of the submitted system(s)
- Provide a component list for the submitted system(s)
- Provide a detailed cost breakdown of the submitted system(s)
- Detail any modifications specifically for the purpose of FRVT 2002

Additionally, participants must use the following guidelines for submitting system description documents.

1. Use Microsoft Applications: All participants should ensure their reports are submitted in a Microsoft Office application format. For instance, word processing documents should be submitted in Word, charts and spreadsheets should be submitted in Excel, etc. Using Microsoft applications will ensure uniformity and facilitate compilation of the final report.
2. Send All Original Documentation: Send original documents for anything that is to be inserted into your report. For instance, send original Excel spreadsheets, PowerPoint slides, and image files (such as PS, EPS, JPG, TIF, etc.). Once the final report is compiled, it will be converted to PDF format for distribution to the community. Sending original documentation ensures uniformity and accuracy when the report is compiled and converted to PDF format.
3. Use Standard Fonts: Use either “Arial” or “Times” fonts (NOT Times Roman) in your reports (except for equations).
4. Embed All Fonts: Not all fonts are available on all machines. If you must use a font that is not standard with Microsoft (for instance, with equations, etc.), you should embed the font in the document so it will not change when the document is converted to PDF format.
5. PDF or Hard Copy: Include either a PDF file or hard copy of your document so your desired layout may be preserved.

2.6 Time Limits

Each test selected by a participant must be completed within a designated time limit. The following table shows the time limit for each test.

Test	Time Limit
Medium Computational Intensity	264 hours
High Computational Intensity	264 hours

The time allotted for each test includes the time necessary to copy the target and query sets from the supplied USB hard drive(s) to the participant hard drive(s) and to copy the resulting output back to the USB hard drive(s). Participants will setup their system(s) when arriving on their start date then process a small set of sample images and return the results on the supplied USB hard drive(s). Test agents will then verify that the results are in the proper format then distribute the main body of images. The evaluation period will start at the time the main body of images is released and end 264 hours later. The evaluation period must begin by 1700 EDT on the participant’s start date.

The FRVT 2002 will report results for complete and correctly formatted test submissions. Additionally, results for incomplete submissions may be reported if appropriate. Systems may run continuously during the allotted time, but access to the testing area will be limited to the hours of 0900 through 1700 EDT each test day, including weekends. During that time, participants may monitor their system(s) once each hour, including weekends, with the exception of the participant’s start date, during which full time access

will be allowed from 0900 through 1700 EDT. In the event of an overnight system crash, participants will not be allowed to restart their system(s) until 0900 EDT the following day. Participants are encouraged to implement their system(s) in a manner that allows restarting from the point where a crash occurred rather than restarting from the beginning.

In the event of a power outage, time will be added to the evaluation period of all affected participants. If the outage occurs between 0900 and 1700 EDT, the amount of time added will be the amount of time that the power was out. If a power outage occurs outside these hours, the amount of time added will cover the period from the start of the outage until 0900 EDT the next day.

2.7 Images

FRVT 2002 participants will be given a large database of images on the day they arrive for testing. The images will be provided via IDE hard drives in USB 2.0 enclosures. The drives will contain a single 80GB partition containing a NTFS file system. Any participant requiring a different configuration must contact the FRVT 2002 Liaison prior to arrival at the test site (check the website for the deadline for making such arrangements). A separate hard drive will be used for each test. If required, a separate laptop computer may be used to facilitate transfer to/from the USB hard drive and face recognition system(s).

Participants will be required to use two separate partitions or hard drives, one for the OS and face recognition algorithms, and the other for working space. Participants will be allowed to copy the images onto their working space partition or drive only. Any intermediate files derived from the images must also be stored on the working space partition or drive.

2.8 Similarity Files

Participants must produce matching results in the form of similarity files that will be submitted to the test agents at the end of the evaluation. Participants will be allowed to store similarity files on their working space partition or drive only. At test completion, the similarity files must be transferred to the same USB hard drive(s) used for distributing the image files. If required, a separate laptop computer may be used to facilitate transfer to/from the USB hard drive(s) and face recognition system(s).

Participants must return the media to the government at test completion and perform a low-level format on their working partitions and drives. A test agent will assure that none of the images, or data derived from the images, are still resident on the test computer. The inspection will involve, at a minimum, the government deleting files generated during testing and wiping free space on all disk drives.

2.9 Video Recording

Testing activities will be recorded using video cameras for documentation of the evaluations. Footage from this documentation will not be made available to the public without review and comment from any participant that is named in the video.

2.10 Schedule

Participants will begin and end their evaluation period on staggered dates in order to prevent traffic jams while equipment is being moved into the test area. Participants will report to the test area at 0900 EDT on their scheduled start date. The USB hard drive(s) containing the images to be processed will be distributed the same day, after the participant has completed system setup and sample output has been verified. Participants may return the complete results on the hard drives to the test agents at any time within 264 hours after the images have been released. Participants will be allowed to monitor their systems once each hour from 0900 through 1700 EDT while the test is in progress, including weekends, with the exception of the participant's start date, during which full time access will be allowed from 0900 through 1700 EDT.

3 Test Procedures

The evaluation will proceed according to the procedures detailed in the following subsections. Certain events will be recorded using the event checklist shown at the end of this document. A test agent and participant POC will initial the checklist as events occur. Each participant will assign one spokesperson to perform this function.

3.1 Setup

1. Test area opens at 0900 EDT.
2. Participant provides system description document(s) to test agent.
3. Participant and test agent initial checklist indicating that a system description document was submitted, in its proper form, for each system.
4. Participant sets up computer system(s) in test area.
5. Test agent records available hard drive space on checklist.
6. Participant runs participant-supplied virus scanning software on each system.
7. Participant and test agent initial checklist acknowledging available hard drive space and the fact that a virus scan has been successfully completed.
8. Test agent releases USB hard drive(s) containing sample images to participant when system setup is complete.
9. Participant copies the set of sample images from USB hard drive to working space of face recognition system, performs matching to produce similarity files, then submits results to test agent using same USB hard drive.
10. Test agent verifies that similarity files are in proper format.
11. Participant and test agent initial checklist indicating that sample similarity files are in the proper format.
12. Test agent releases USB hard drive containing main body of images.
13. Participant and test agent initial checklist indicating that the main body of images has been released to participant. At this point, the time is recorded and the evaluation period begins. The evaluation period must begin on or before 1700 EDT on the participant's start date.

3.2 Evaluation Period

1. Participant copies images for medium and/or high computational intensity tests from USB hard drive(s) to their system(s).

2. Participant begins matching query images against target images.
3. Participant may monitor system(s) once each hour between the hours of 0900 and 1700 EDT each testing day to handle system crashes, with the exception of the participant's start date, during which full time access will be allowed from 0900 through 1700 EDT.
4. Participant copies resulting similarity files to USB hard drive(s).

3.3 Completion

1. Participant returns USB hard drive(s) containing images and similarity files to test agent.
2. Test agent verifies a random sampling of similarity files to confirm that they are in the proper format.
3. If the similarity files are not in the proper format, participant may make corrections and resubmit as many times as needed until the evaluation period ends.
4. When similarity files have been successfully verified, participant and test agent initial checklist indicating that USB hard drive(s) have been returned and that a random sampling of similarity files has been verified.
5. Test agent copies all similarity files from the USB hard drive to a RAID server then makes additional copies using other media.
6. Test agent inspects all hard disks of submitted system(s), deletes images, derived files, and similarity files then wipes drive free space.
7. Participant and test agent initial checklist indicating that hard disks have been inspected and free space has been wiped.
8. Participant and test agent sign checklist.

FRVT 2002 Event Checklist

Organization Name: _____

Event	Test Agent Initials	Participant Initials	Date / Time
System description documents submitted			
System 1	_____	_____	_____
System 2	_____	_____	_____
System hard drive free space recorded and virus scan successfully completed			
Computer 1 (bytes): _____	_____	_____	_____
Computer 2 (bytes): _____	_____	_____	_____
Computer 3 (bytes): _____	_____	_____	_____
Computer 4 (bytes): _____	_____	_____	_____
Sample similarity files verified to be in proper format			
Medium Computational Intensity	_____	_____	_____
High Computational Intensity	_____	_____	_____
USB hard drive(s) released to participant (timer starts)			
Medium Computational Intensity	_____	_____	_____
High Computational Intensity	_____	_____	_____
Random sampling of similarity files verified to be in proper format.			
Medium Computational Intensity	_____	_____	_____
High Computational Intensity	_____	_____	_____
USB hard drive(s) returned to test agent			
Medium Computational Intensity	_____	_____	_____
High Computational Intensity	_____	_____	_____
System hard drives inspected, free space recorded and wiped			
Computer 1 (bytes): _____	_____	_____	_____
Computer 2 (bytes): _____	_____	_____	_____
Computer 3 (bytes): _____	_____	_____	_____
Computer 4 (bytes): _____	_____	_____	_____

Test Agent Signature / Date _____

Participant Signature / Date _____

Appendix I : FRVT 2002 Imagery

I.1 Introduction

This appendix describes the imagery used in the Face Recognition Vendor Test 2002 (FRVT 2002). It discusses the image sets used in both the High and Medium Computational tests, and includes examples of the images, both stills and video sequences, that were used in the various studies described in the main body of the report. This appendix gives brief descriptions of the acquisition, processing, properties and use of the images in each subset. Some common aspects of all images of the FRVT 2002 corpus are that they contain the face of exactly one individual, and that the face (and neck and part of the chest) is the only foreground object, and that the image was acquired in the visible spectrum in color. All the images provided to FRVT 2002 participants are in standard JPEG format, regardless of whether they were acquired that way.

I.2 High Computational Intensity Test Database

Imagery used for the High Computational Intensity Test is a subset of a much larger collection furnished to the FRVT2002 Evaluation Team by the Visa Division of the Department of State. That collection consisted of approximately 6.8 million images from about 6.1 million visa applicants who received various non-immigrant visas in Mexico over the period 1996 to 2002. The data supplied to NIST is itself a random subset of all images gathered in Mexico during that period. The imagery was obtained in the field and in large quantities such that its utility in quantifying face recognition performance is for the most part unreachably. As discussed below the images are of good quality, gathered in a consistent manner. This homogeneity leaves natural anthropometric population variation as the leading cause of error in the systems tested. This is in contrast to much of the academic literature in which vision systems are designed to handle particular difficulties in images such as non-frontal views, or non-uniform or low contrast illumination.

The vast majority of the images were collected at United States consular offices using standard issue digital imaging apparatus whose specification remained fixed over the collection period. Consular officials who collected the images instructed the applicant to face the camera. The result is a set of well-posed (i.e. frontal to within 10 degrees) images of cooperative subjects usually with a mouth-closed neutral facial expression. The subject usually occupies a quarter of the image area. The top of the subject's shoulders is almost always visible. The background is universally uniform, with a white background in most cases. White balance artifacts occur less than 5% of the time. Figure I-1 shows images of the authors at NIST, included here in place of actual State images because privacy issues preclude their inclusion.



Figure I-1. Images of NIST personnel included here as reasonable representations of those used in the FRVT 2002 High Computational Intensity test.

Metadata The images, as supplied to the FRVT 2002 Evaluation Team, were accompanied by image and subject specific information: subject id, age, country of birth and nationality, image acquisition date and consular office code. Because sex information was not available NIST inspected one image of each of the 37,440 individuals used in HCInt, and keyed either male, female, or indeterminate. This latter category was used in 89 cases where the there individual's sex could not readily be determined.

I.2.1 The FRVT2002 High Computational Intensity Subset

We applied a sequence of image and population specific selection criteria to the supplied State Corpus to arrive at the HCInt subset. The original State set contains 6.8 million images of 6.1 million subjects. The HCInt subset contains 121,589 images from 37,440 persons which we obtained as follows. We retained only color images, discarding greyscale images for which the precision was one byte per pixel, or for which $R = G = B$ for all pixels; second we retained only those images of height and width 300 and 252 pixels; and third we kept only those images that did not produce even very minor JPEG file format irregularities as reported by our Independent JPEG Group March 1998 implementation. These three steps reduced the number of candidate images by about 8.2% of the original total.

We then invoked two population specific criteria. The first is to consider only those images of person who were 18 years of age or older at the time they first applied for a visa. This obviates legal complications involved with the use of minors, and also removes effects arising from the rapid change of appearance associated with childhood and adolescent growth. Secondly, and most significantly, we retained only those persons with three or more images. This multiple images per person criterion is essential to a recognition evaluation. We chose to use three or more images as this led to a final database of size that is tractible in the FRVT 2002 test.

The result is a collection of color images, all of size height 300 and width 252 pixels, stored as standard JPEG images with a mean file size 9.5 kB. The mean interocular distance, as reported by a known effective eye location algorithm, is 75.2 pixels. The images given to participants were byte-for-byte the same as those provided by the Department of State; they were not recompressed, resized or in any way processed by NIST.

Selection Biases The criteria described above impart selection biases on the population. For example, the demographic structure of the whole State corpus is plotted with that of the partial set in the figure I-2. It

shows that persons born through the 1960s are under-represented by our study, but those born after 1967 are over-represented. However our study reports performance on the population who applied for visas on three or more occasions. In that sense the appropriate population density we should compare against is that from which persons apply on TWO or more occasions.

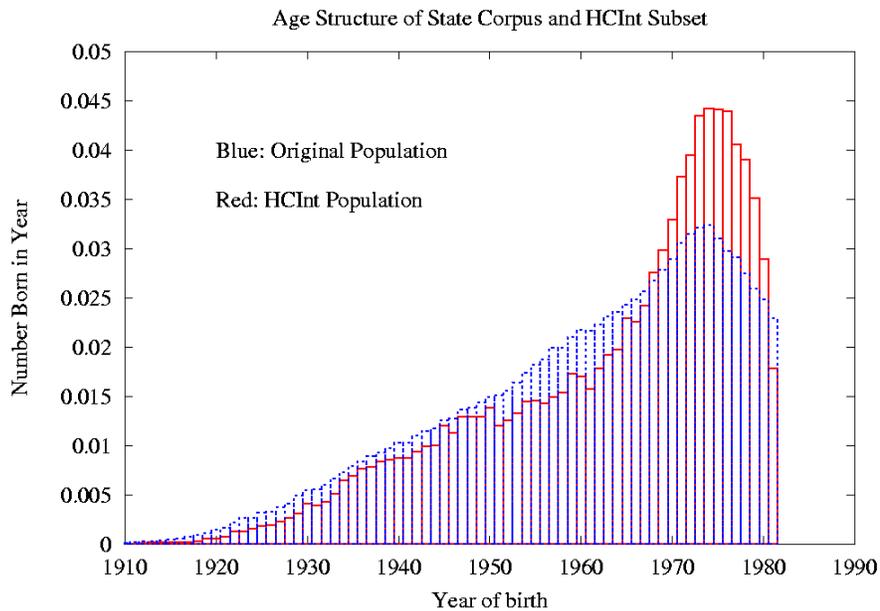


Figure I-2. Age structures of the entire State corpus and the HCInt subset

If recognition performance is affected by age of the subject then an under- or over-representation in our sample would change the results somewhat. It should be noted, however, that this effect depends on the application; if the recognition task is a new one-to-many search on an entire corpus, then the number of images per person is irrelevant.

Other distributions were also affected by our selection criteria. For instance 62% of the population used in HCInt were issued visas at just three of the consular offices, whereas in the State corpus as a whole that figure is just 37% for the same three offices. This concentration occurred because repeat applicants preferentially attend a select few consular offices. Given the uniformity of the imaging apparatus across offices, any changes in the recognition performance estimates is likely to be small. This latter assertion can be tested using the framework we have.

Metadata information is vital to recognition evaluations as it allows recognition performance to be broken out by category, for example by age, or by time between image acquisition dates.

Subject ID Given the vital nature of subject identity in recognition performance testing, this section discusses the derivation and integrity of the HCInt ID metadata field. The ID field itself is a 32 character hexadecimal string (i.e. 128 bits) from the MD5 hash function applied to the concatenation of four metadata strings, namely the subjects name, date and place of birth, and nationality.

There are two types of error that generally affect identity information in biometric databases. First, is that several different persons have the same ID, and, second, that a single person has more than one ID. In the case of HCInt the process by which two or more persons have the same ID is that the four metadata fields

are identical. This is clearly an unlikely coincidence in a general population and can be expected to introduce errors much less than the recognition variation observed over disjoint populations in the FRVT 2002 results. Another mechanism for this class of error is that the hash function produced the same output string, a hash-collision, for different inputs. This is an even more unlikely scenario given the quality of the MD5 cryptographic algorithm and the number of trials.

The second class of error is, at first glance, more problematic. If a typographical error occurred during the data-entry phase of the subject enrollment then two IDs would be obtained from the same person. However for HCInt this event would, via MD5, yield a "new" person with only one image, and these individuals will be excluded by the criterion that we only use persons with three or more images. Likewise if an individual within in the State corpus has applied under two different names, then this will not be revealed in FRVT 2002 unless it has occurred three or more times. We have validated this finding by inspection of the most similar ostensibly non-matching images, as reported by the leading FRVT 2002 systems, and found no errors of this kind.

Nationality As we stated above the State imagery has unreprouachable size and real-world provenance. The notable exception to this is that the images are from one nation, and although the population is large enough to span anthropological face variation in Mexicans we cannot strictly extend the conclusions to, say, a Scandinavian population. Although the entire corpus contains images of persons of 185 nationalities and Mexicans comprise 98.76% of them, the HCInt subset contains images of persons of 66 nationalities of which 96.59% or 36165 are Mexican. The decrease in the proportion of Mexican nationality indicates that the non-Mexican population in Mexico applies more frequently for visas. The next most represented nationality within the HCInt subset is Chinese at 1.68% or 628 persons. This fraction was considered insufficient to evaluate the effect of nationality (and ethnicity) on recognition performance.

Footnote to the Results Many of the FRVT 2002 results are based on a gallery of 37437 individuals. This is three fewer than the number of persons present in the HCInt subset described above. This difference arose because one FRVT 2002 participant produced incomplete similarity files for a single image of those three persons. The FRVT 2002 committee elected to remove those three subjects from all scoring computations.

1.3 Medium Computational Intensity Test Database

In contrast to the HCInt image set, the Medium Computational Intensity image set is composed of a heterogeneous set of still images and video sequences of subjects in a variety of, poses, activities and illumination conditions. The images originate from two distinct sources. The first is the joint National Institute of Standards and Technology (NIST), Naval Surface Warfare Center (NSWC, Dahlgren), University of South Florida (USF) still face image corpus gathered between 1999 and 2002. The second, from University of Texas, Dallas, consists of video and stills, taken in 2001. Both sets of images are partitions of the hBase database being assembled as part of DARPA's Human Identification at a Distance Program¹. Each piece of imagery there is accompanied by ground truth information describing its acquisition and properties. The images of the MCInt were supplied to participants in two disjoint partitions. The first contained 7722 still images from 1477 individuals; the second contained 1477 video sequences and stills, from 63 subjects.

The two partitions share in common that the subjects appearing in the images are all unpaid, informed, and consenting volunteers. For privacy reasons the data was gathered anonymously; a volunteer's name was never recorded and is instead replaced by a randomized integer identifier that is used to label all the imagery ever taken of that individual. IDs are maintained over time by having subjects self-identify from a previous image.

¹ See P. Jonathon Phillips, "Human Identification Technical Challenges", in Image Processing for Biometrics in the International Conference on Image Processing (ICIP), 2002, Rochester, New York.

I.3.1 The NIST-NSWC-USF Collection.

This collection is comprised of images taken indoors and outdoors. Although the images were taken over more than three years at three sites they were acquired in almost identical circumstances such that their geometric and photometric characteristics are approximately the same. This was achieved by using consumer grade autofocus digital CCD cameras (Sony Mavica) in to photograph individuals against a background of photographic supply 18% gray paper, illuminated alternately under a two photoflood arrangement, and an ambient overhead fluorescent setting.

Although the imagery itself has consistent properties the subject populations of the three sites differ. The USF segment consists of subject population of diverse ethnicity that is distinct from the older predominantly caucasian mix at NIST and NSWC. The collection is therefore not free from unavoidable selection biases.

Four images of each subject are used in the MCInt study. The first two, under the two photofloods have the subject with a neutral then an alternate facial expression. The latter two are a corresponding pair obtained under ambient lighting. In most cases the alternative expression is classified as a smile, but in a significant fraction other expressions are present.

The images of figure I-3 form the gallery and probe sets used for testing changes in expression, overhead lighting and outdoor lighting, and elapsed time. The leftmost image is drawn from the gallery of the earliest examples of neutral expression two-light frontal stills from all the persons in the NIST-NSWC-USF collection. The second image is a corresponding image from a later collection. The third shows a subject with an alternative facial expression taken on the same day as his gallery image. Finally the last image shows fluorescent room lighting, without changes in the camera's white balance setting or facial expression. The images are either 240x320 or 256x341 in size with an average compressed JPEG file size of approximately 18 kB. The median interocular distance is 51.2 pixels.



Figure I-3. Four sample indoor images from the NIST-NSWC-USF Collection used in the MCInt test.

Outdoor Still Images The outdoor stills are characterized by changing background and directional sunlight illumination. A small fraction of the images were taken with cloud cover. Each subject stood facing the camera at a distance of 2 meters from a zoomed Sony Mavica digital camera in auto-focus mode. In all cases subjects were asked to provide a neutral facial expression. Figure I-4 shows two images taken on the same-day as those in Figure I-3, then two more recent samples for studying indoor outdoor and time delay effects. All images have size 240x320 pixels and compressed JPG file size of about 26 kB. The median interocular distance is 45 pixels.



Figure I-4. Four sample outdoor images from the NIST-NSWC-USF Collection used in the MCInt test.

Frontal Reconstructions The MCInt image set included partitions included to test the efficacy of using 3D morphable models as a preprocessing step in the recognition of non-frontal images. The concept is that a non-frontal image can be fit to a 3D model of human heads, and then a frontal view can be generated. Specifically by starting from an example set of 3D face images, a morphable face model is derived by transforming the shape and texture of the examples into a vector space representation. New faces and expressions can be modeled by forming linear combinations of the prototypes. Shape and texture constraints derived from the statistics of example faces can be used to guide manual modeling or automated matching algorithms. This method² has been advanced by Thomas Vetter and his group at the University of Freiburg, and funded recently by DARPA's Human Identification at a Distance program.

To investigate the utility of the Vetter method, NIST departed from its standard two-light source illumination protocol and instead used a single photoflood source situated behind and above the camera on the optical axis. This was done to satisfy a single light source assumption of the Freiburg model. NIST consulted with Volker Blanz in Freiburg, supplied him with nearly 800 of the new NIST images and incorporated the original and frontal reconstructions into the MCInt set. Examples of the original images are shown in Figure I-5. Reconstructions are shown in Figure I-6. The images used for the test were reduced from their original size to have width 256 and height 341 pixels. The average file size is 14.7 kB.



Figure I-5. Frontal images taken from persons at NIST using a single light source (and ambient overhead lighting). The use of a single light source is assumed by the 3D morphable model used to remove pose variation from non-frontal images.

² See Volker Blanz and Thomas Vetter, *A Morphable Model For The Synthesis Of 3D Faces*, Siggraph 1999, Computer Graphics Proceedings, Pages 187-194, ed. Alyn Rockwood, Addison Wesley Longman, Los Angeles.



Figure I-6. The top row shows original images as matched against those in the gallery of Figure I-5. The subject, illuminated by one light source, is looking left and right at about 45 degrees, and up and down at about 30 degrees. The second row shows the corresponding frontal reconstructions from the 3D morphable model.

I.3.2 Imagery from the University of Texas at Dallas

A database of face and person imagery was created at The University of Texas at Dallas for use in human memory experiments that test the importance of natural face and body animation, as well as posture. Close-range video clips and static images were taken of more than a hundred individuals on at least two different occasions to make duplicate sets separated by one week to six months. We used a 63-person subset who appeared on two occasions. The image file size is 720 by 480 pixels, with file size 22 kB.

Facial Speech Videos Two five-second video clips were taken of each individual speaking, first in a neutral way, then in an animated way. Figure I-7 shows two examples; all images have width 520 pixels and height 480 pixels, and the filesize is 16 kB.



Figure I-7. The rows show selected frames from examples of the UT Dallas "facial speech" videos lasting 150 frames. The two rows show the subject gathered on different occasions.

This page intentionally blank.

Appendix J

Normalization in FRVT 2002

FRVT 2002 Normalization Specifications

1. Introduction

As stated on the Face Recognition Vendor Test (FRVT) 2002 website (www.frvt.org), the FRVT 2002 uses the Sep96 FERET evaluation protocol (Phillips et al. [1]). The key innovation in the Sep96 FERET evaluation was the introduction of virtual galleries and probe sets. Virtual galleries and probe sets allow for the measurement of performance for different categories of image from one set of similarity files.

Using normalized similarity scores is an idea that has emerged in the biometrics field. Normalization is most commonly used in fingerprint matching. Classically, similarity scores are computed by comparing two images and reporting a measure of similarity between the two faces in those images. Normalization adjusts the scores based on all the images in a gallery. An overview of normalization is presented in section 2, which precedes a detailed description of how FRVT 2002 participants may proceed with such an option if they elect to do so.

2. Normalization

To make the idea of normalization more concrete, consider the following notation. For details on the notation and terms, see Phillips et al. (2000). Let G be a gallery that is a subset of a target set T , and p a probe from a probe set P that is a subset of a query set Q . Let S be the matrix of non-normalized similarity scores between target set T and query set Q . Let S_p^G be the set of similarity scores between probe and all signatures in G , and $|G|$ be the number of signatures in gallery G . The set S_p^G consists of non-normalized scores. One possible normalization function, f_1 , takes S_p^G as input, and gives as output, a set of normalized scores N_p^G - i.e. $N_p^G = f_1(S_p^G)$ where N_p^G is a new set of $|G|$ normalized similarity scores between probe p and all images in G . The set of normalized score N_p^G is computed separately for each p in P , and the performance scores are computed from $|P|$ sets of normalized scores.

A second possible normalization function takes into account the similarity scores between all pairs of gallery images. Let S^{GG} be a $|G|$ by $|G|$ similarity matrix between all pairs of gallery images in G . A possible normalization function, f_2 , takes input S^{GG} and S_p^G and outputs N_p^G i.e. $N_p^G = f_2(S^{GG}, S_p^G)$.

3. Implementation

Specific guidelines for the use of normalization in FRVT 2002 are given below. Normalization is a post-test administration procedure that will be executed by NIST after similarity files have been submitted. Participants who elect to utilize the FRVT 2002 normalization option will need to provide NIST with an object file that contains the result of the compilation of C code of specific functions. For intellectual property reasons **vendors must not, at any time, submit source code, nor any details of any algorithm** to the FRVT 2002 committee.

1. **Tasks:** Participants may submit up to six functions: zero, one, or two for each of the identification, verification, and watch list tasks. To ensure proper interoperability, the calling semantics *cannot* be varied from those described below.

2. **Object Files:** The functions will be supplied as a single compiled object (i.e. .o) file that can be run on the FRVT 2002 analysts' computers performing the scoring. All code should be written in ISO C. Other languages can only be used if the call semantics are *identical* to those for C. Particularly C++ must *not* be used as is because of non-standardization of function name mangling.
3. **Linux:** The object files will be linked into programs that run on Intel-based machines running the Linux operating system¹. To minimize interoperability problems, we recommend using a recent Linux distribution such as RedHat 7.x, Mandrake 8.x, SuSE 8.x, or Gentoo 1.x.
4. **Report:** The FRVT 2002 will report both normalized and non-normalized performance scores.
5. **Responsibility:** Participants are responsible for providing a normalization function that works with the FRVT 2002 scoring system.
6. **Resources:** The normalization functions need to work within a reasonable time, and within reasonable memory. The normalization functions will need to be called *each time* a probe is processed. This means that the function is embedded in the heart of the scoring code. Use of normalization cannot cause a significant delay in the scoring process. Note, to process a gallery of 10,000 images and probe set of 10,000 images will require calling the normalization routine 10,000, and each call will contain 10,000 similarity scores.
7. **Side-Effects:** The normalization functions must have no unspecified inside effects such as sending data to standard output or standard error, writing files to disk, etc.
8. **Usage:** The FRVT 2002 organizers will not tailor galleries to fit specific properties of a normalization function. For example, if a normalization function assumes that a gallery is homogeneous in terms of *image quality* or properties of the subjects appearing therein, the FRVT 2002 organizers do not guarantee that all galleries will be homogeneous.

3.1 Function Specifications

Participants may submit an object file containing *up to six* different normalization functions --- zero, one, or two functions for each of the *identification*, *verification*, and *watch list* tasks. Each task may use a function with the either of two prototypes, *F1* and *F2*, as detailed below. The function name itself is discussed in section 3.2.

```
void F1(const unsigned int g_in, const float* sPG_in, float* sPG_out);
```

where

g_in is the number of elements in the gallery;
sP_in is a pointer to the similarity scores for a given probe or gallery (i.e., a one-dimensional array);
sPG_out is where the normalized similarity scores are to be written.

Or, with the *F2* call signature:

```
void F2(const unsigned int g_in, const float* sGG_in, const float* sPG_in, float* sPG_out);
```

¹ Linux is a freely-downloadable, Unix-like operating system, available inexpensively online and in bookstores. See <http://www.tldp.org/> for more information on how to obtain, install, and use Linux.

where *sGG in* is a pointer to a matrix of similarity values between all pairs of *gallery* elements. Because this matrix may not be symmetric it contains all elements represented as an array of floats. The (i,j)-th element is accessed using `sGG in[j * g_in + i]` and corresponds to the i-th element from the j-th similarity file; this implies column major ordering of the data, with each column of the matrix coming from a single similarity file.

Participants may assume that the proper storage for `sPG_out, g_in * sizeof(float)`, has already been allocated. Storage for these variables will be allocated outside of the functions.

Within time and memory limitations, performance scores will be generated for normalization functions provided. Because of potential time and memory limitations associated with F2 class functions, it is recommended that vendors provide an F1 function for each F2 function.

3.2 Naming Conventions

The name of the functions that a participant may submit in their object file is determined by:

1. the name of the participant
2. the application (i.e., identification, verification, watch list)
3. the signature's structure (i.e., *F1()* or *F2()*)

Each participant and application has its own abbreviation. Participants will be e-mailed their own particular abbreviation by the FRVT 2002 liaison. The application abbreviations are:

Identification	<i>ident</i>
Verification	<i>verif</i>
Watch List	<i>watch</i>

Each call signature will take the following general form.

```
void [participant code] [application] [function type](const int g, ...)
```

where *[participant code]* is the participant's three-letter code, *[application]* is the target application, or context within which the normalization code should be used, and *[function type]* is the structure of the function (i.e., *F1()* or *F2()*).

The name of the submitted object file must be *[participant code]-norm.o*. The participant may submit code to be used during scoring on identification, verification and watch list testing. For example, consider that a generic participant, whose three-letter code is NST, provides a *F1()* structured function for identification and a *F2()* function for the watch list task. Their full declarations would be

```
void NST_ident_F1(const unsigned int, const float *, float *);
```

and

```
void NST_watch_F2(const unsigned int, const float *, const float *, float *);
```

The NST code would be written into a single source file named *NST-norm.c*. The following command would be used to compile the normalization code:

```
helix% gcc -O3 -I./ -c NST-norm.c -o NST-norm.o
```

The flag `-c` suppresses linking, `-I./` indicates that the code can include header files from the current directory, `-O3` enables optimization, and `-o` directs the output to the specified file. Note the use of **gcc not g++** here. Also note that **gcc** running under the Cygwin package on non-Linux platforms does not interoperate with the linux runtime environment. The code must *not* be

compiled with the -g flag. This prevents meaningful symbols being visible to the FRVT 2002 analysts. The code should be as self-contained as possible. Any non-standard library code should be included, and to avoid link problems, private functions and variables should be declared static in a participant's code.

Sample code is shown in the box below; note that it does not necessarily perform a useful normalization - **it is included for illustration of data access and interoperability only.**

```
void NST_ident_F1(const unsigned int g,
                 const float* sPG_in, float* sPG_out)
{
    unsigned int i = 0;
    float max = sPG_in[0];
    float min = sPG_in[0];

    for (i = 1; i < g; i++)
    {
        if (sPG_in[i] > max) max = sPG_in[i];
        if (sPG_in[i] < min) min = sPG_in[i];
    }

    for (i = 0; i < g; i++)
        sPG_out[i] = (sPG_in[i] - min) / (max - min);
}

void NST_watch_F2(const unsigned int g, const float* sGG_in,
                 const float* sPG_in, float* sPG_out)
{
    unsigned int i = 0, j = 0;
    float max = sGG_in[0];
    float min = sGG_in[0];

    for (i = 0; i < g; i++)
        for (j = 0; j < g; j++)
        {
            if (sGG_in[i*g + j] > max) max = sGG_in[i*g + j];
            if (sGG_in[i*g + j] < min) min = sGG_in[i*g + j];
        }

    for (i = 0; i < g; i++)
        sPG_out[i] = (sPG_in[i] - min) / (max - min);
}
```

The resulting object file should be sent to frvt2002@nist.gov according to a timetable that the FRVT 2002 administrator will send with the participants' three-letter participant abbreviation. The object file submission will initiate a correspondence with an FRVT 2002 analyst to verify the behavior of the normalization routines.

References

[1] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The FERET Evaluation Methodology for Face Recognition Algorithms. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22: pp. 1090-1104, 2000.

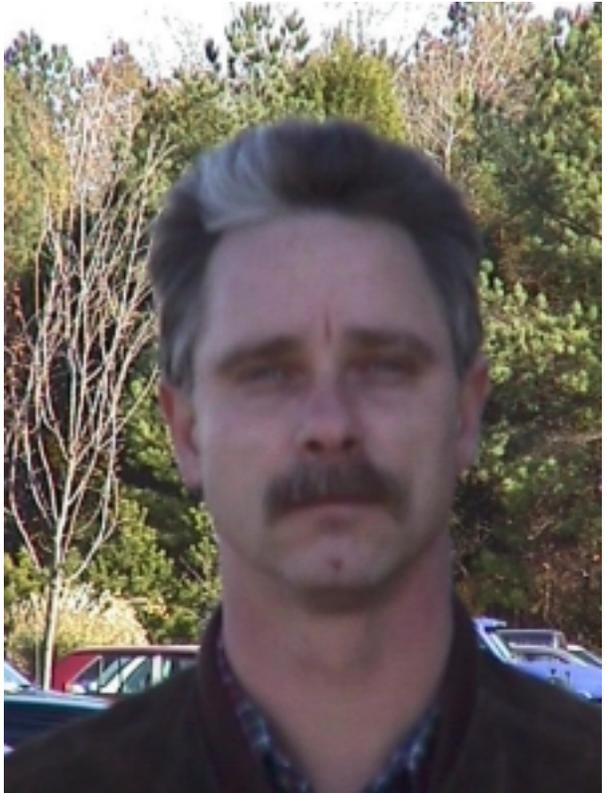
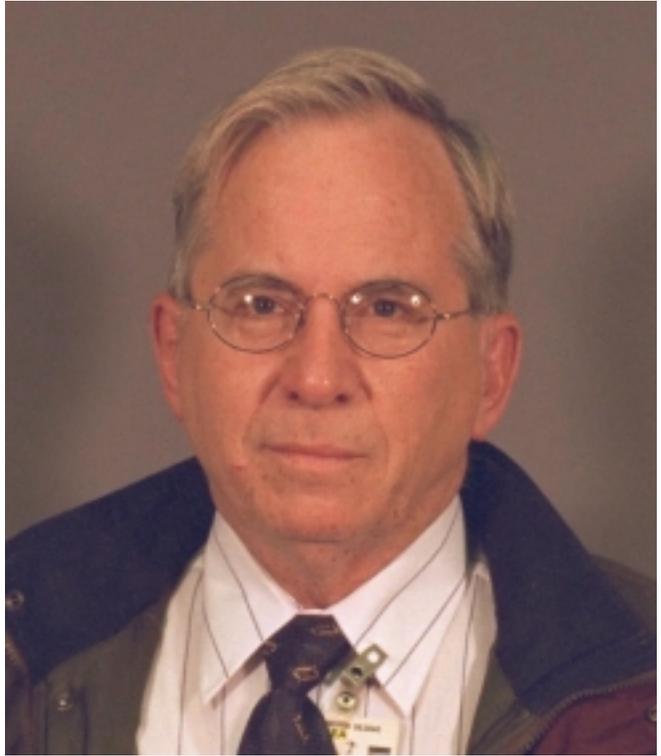
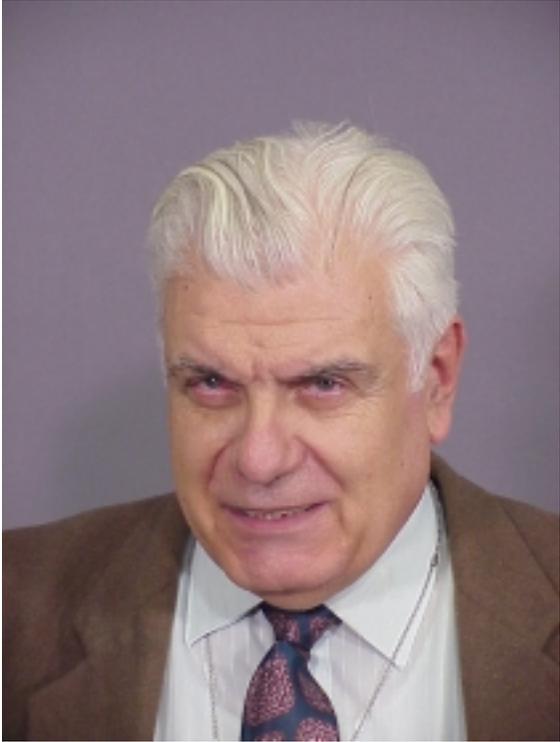
This page intentionally blank.

Appendix K

Development Image Set

The following images were provided to FRVT 2002 Participants as examples of images they would use during the FRVT 2002. A larger database of FERET images was also provided for participant's use during the pre-evaluation check of their APIs to ensure compatibility with the FRVT 2002 scoring code. Images have been non-uniformly resized for printing purposes.









This page intentionally blank.

Appendix L

In order to clearly present an assessment of the state-of-the-art in face recognition, it was necessary in the *Evaluation Report* to report summary statistics of performance. Appendix L presents full cumulative match characteristics (CMC) and receiver operating characteristics (ROC) for all MCInt experiments (Section 8 and Appendix A.4 in the *Evaluation Report*). Also, this appendix presents HCInt verification error ellipses ROCs (Figure 8) for all participants.

L.1 HCInt Verification Error Ellipses Experiment

Figure 8 in the *FRVT 2002: Evaluation Report* showed standard error ellipses for verification performance for Cognitec, Eyematic and Identix. The standard error was computed from twelve HCInt small galleries of size 3,000. The centerline is the ROC performance for the aggregate for all twelve galleries. The ellipses are two times the standard deviation at selected performance points, and the points clustered around the ellipses are the performance of one of the twelve galleries at the selected performance point. Figure L.1 through Figure L.8 show error ellipses for each of the HCInt participants separately, and Figure L.9 show error ellipses for all participants on one graph.

Participant	Figure
Cognitec	Figure L.1
C-Vis	Figure L.2
Dream-Mirh	Figure L.3
Eyematic	Figure L.4
Identix	Figure L.5
Imagis	Figure L.6
Viisage	Figure L.7
Visionsphere	Figure L.8
All Participants	Figure L.9

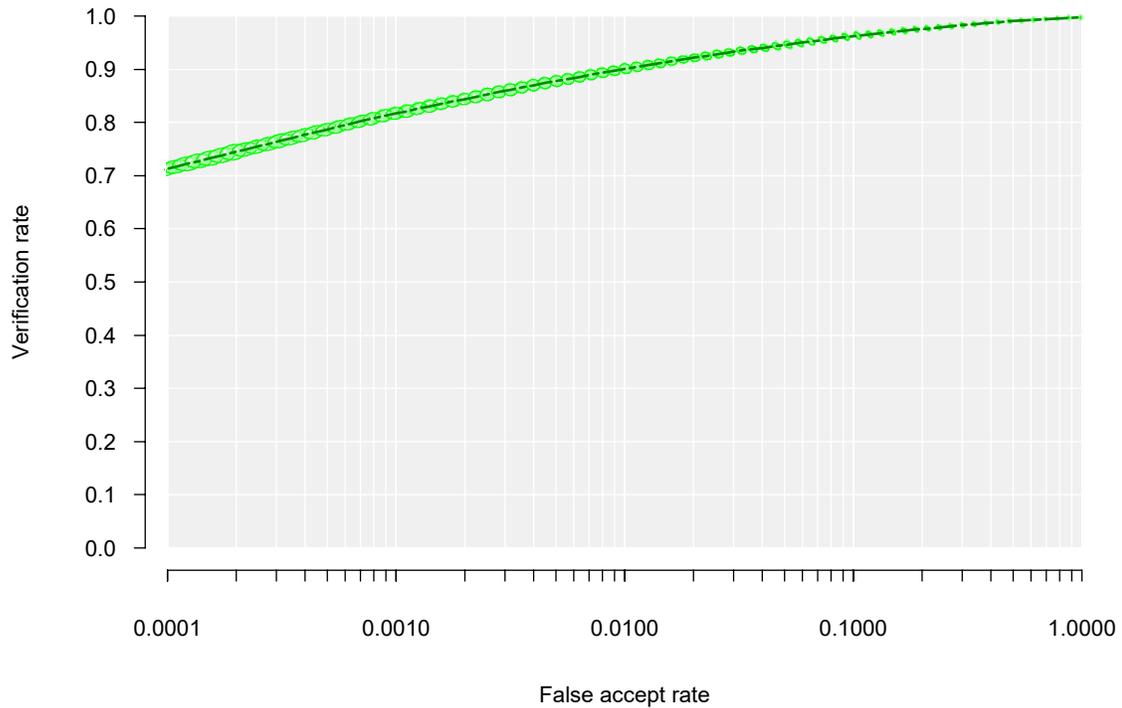


Figure L.1 - Cognitec Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

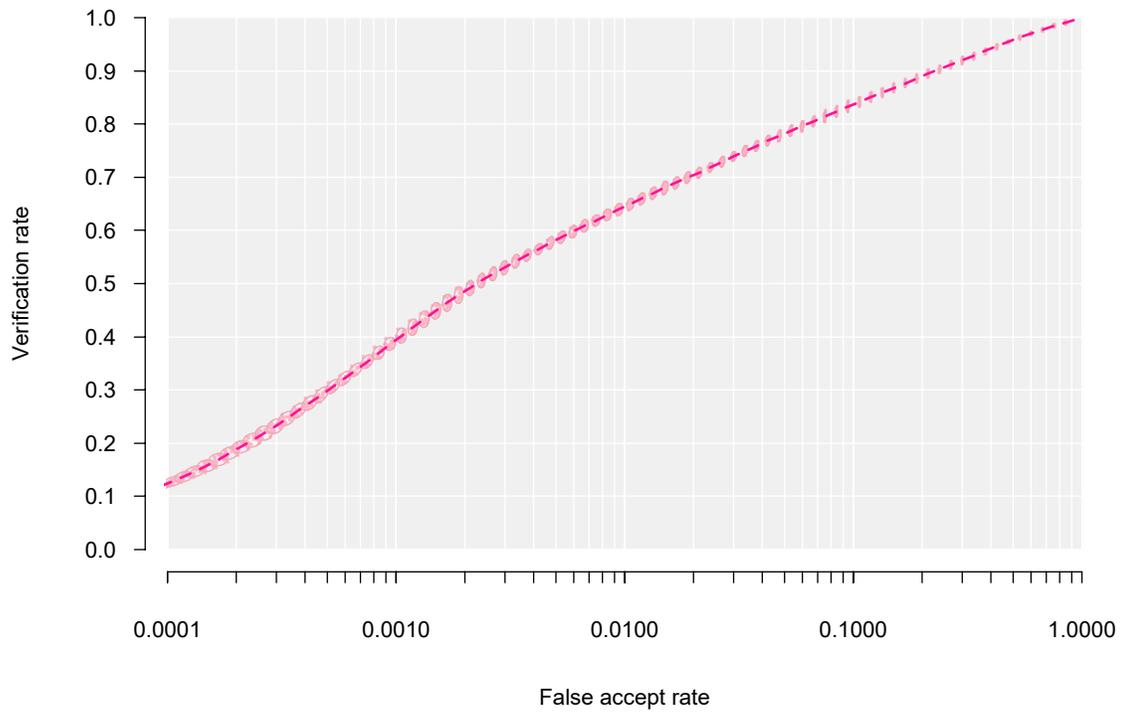


Figure L.2 - C-VIS Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

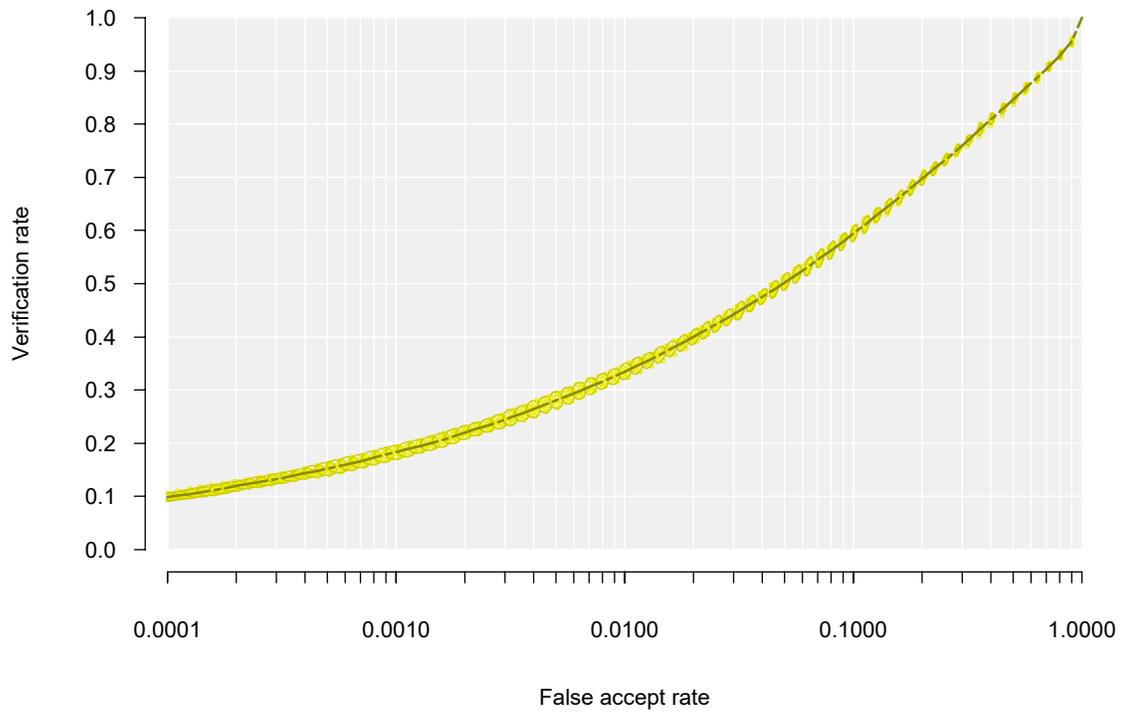


Figure L.3 - Dream Mirh Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

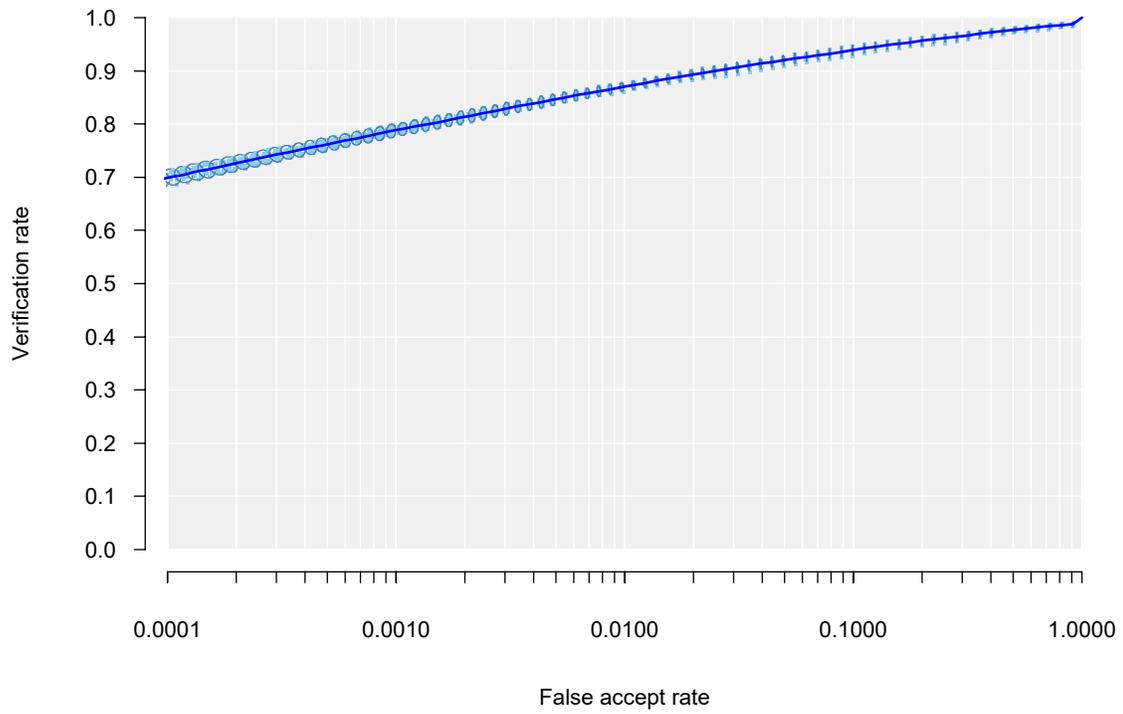


Figure L.4 - Eyematic Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

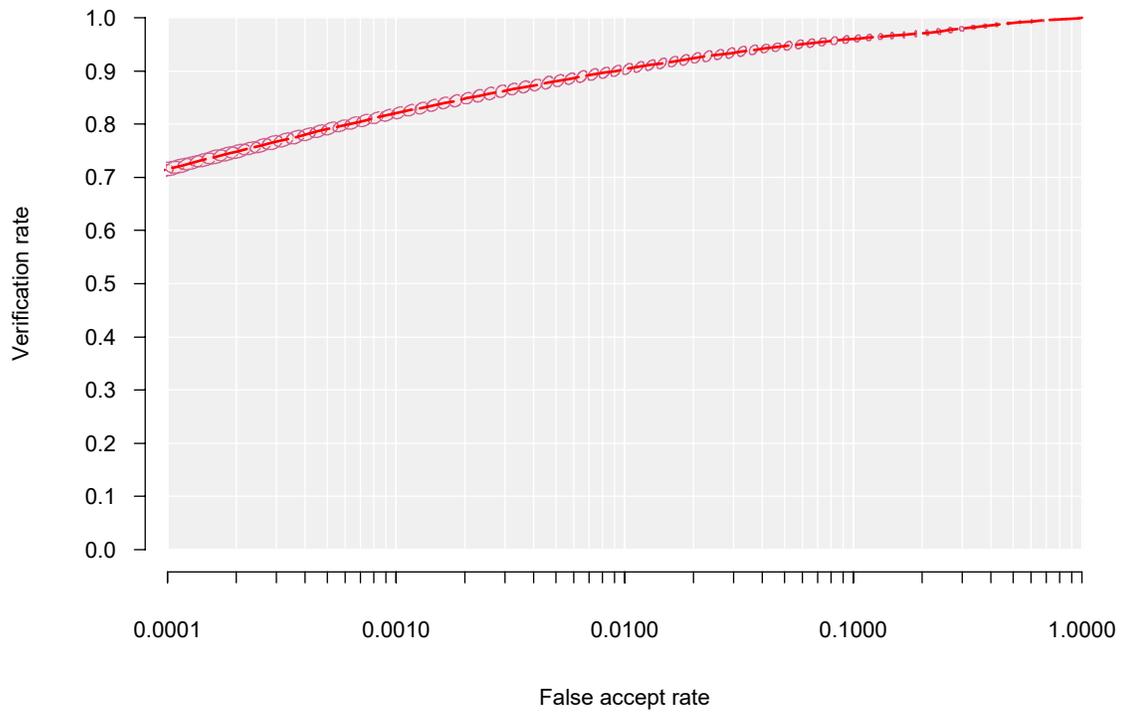


Figure L.5 - Identix Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

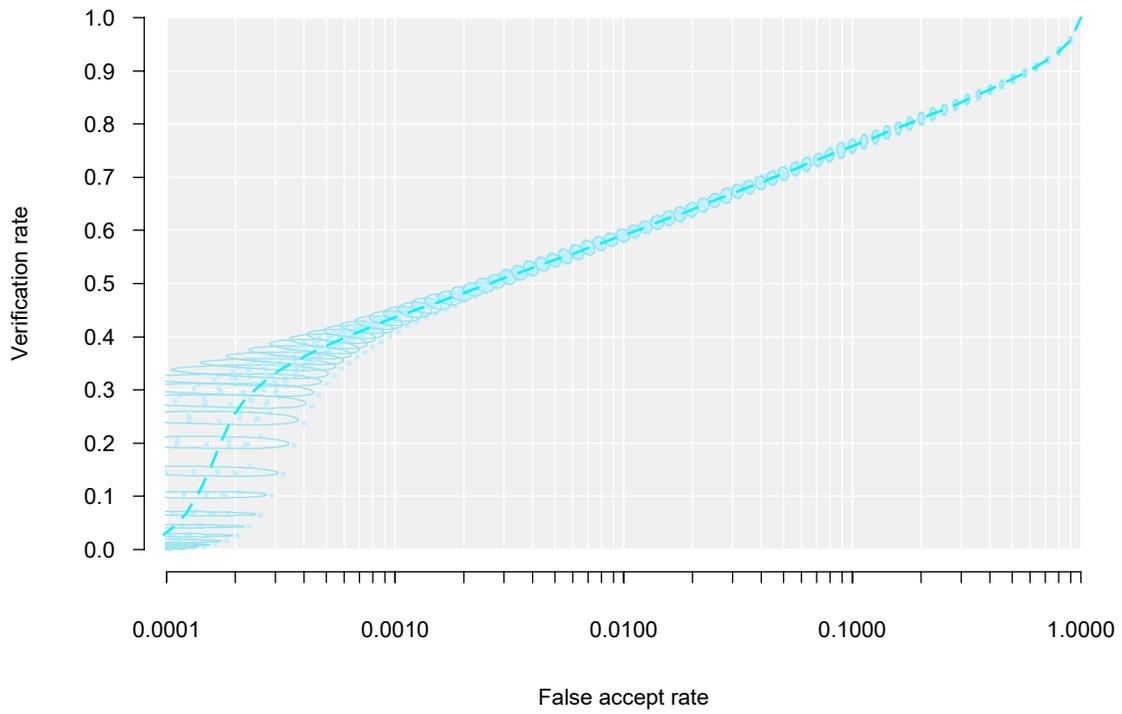


Figure L.6 - Imagis Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

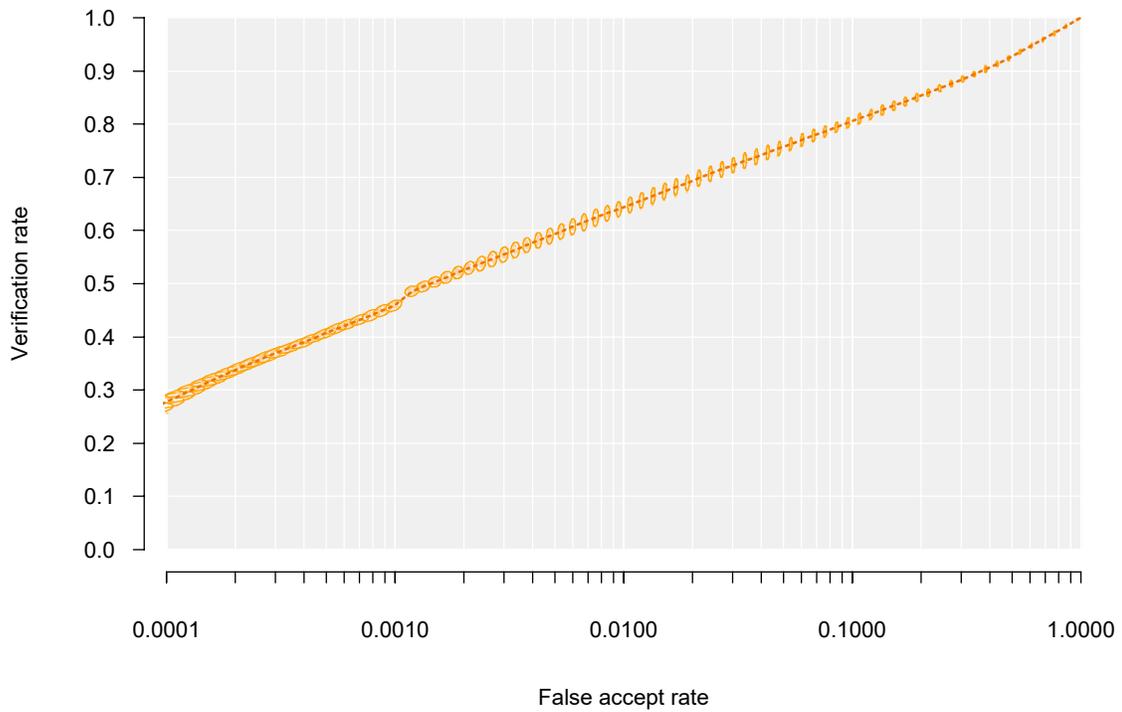


Figure L.7 - Viisage Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

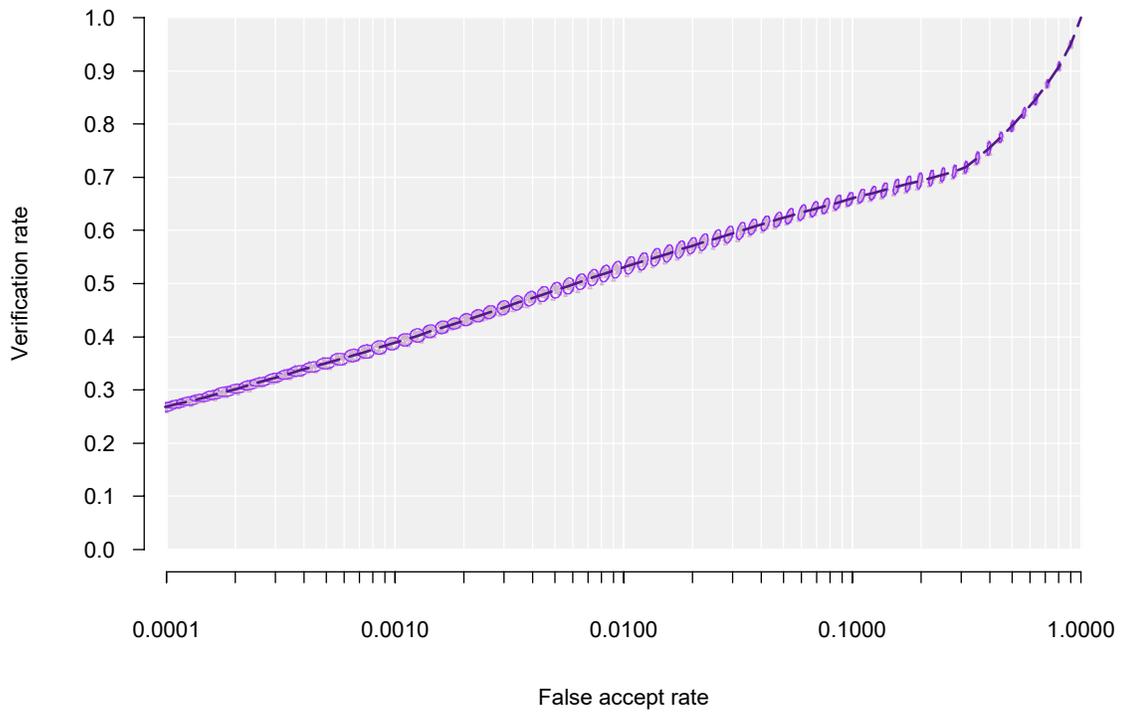


Figure L.8 - VisionSphere Verification Error Ellipses. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

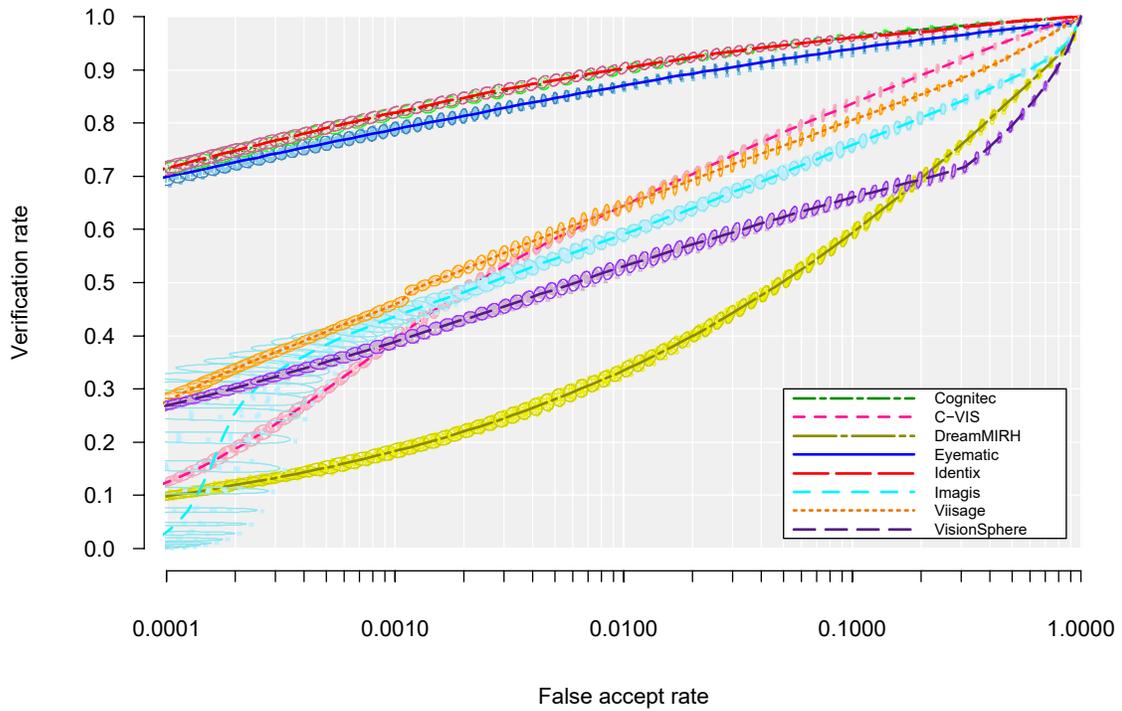


Figure L.9 - Verification Error Ellipses for all Participants. This figure corresponds to Figure 8 in the *FRVT 2002: Evaluation Report*.

L.2 Normalization Experiment

Figure 9 in the *FRVT 2002: Evaluation Report* highlighted changes in performance between normalized and non-normalized verification performance. Figure L.10 through Figure L. 13 provide full normalized and non-normalized ROC curves from this experiment.

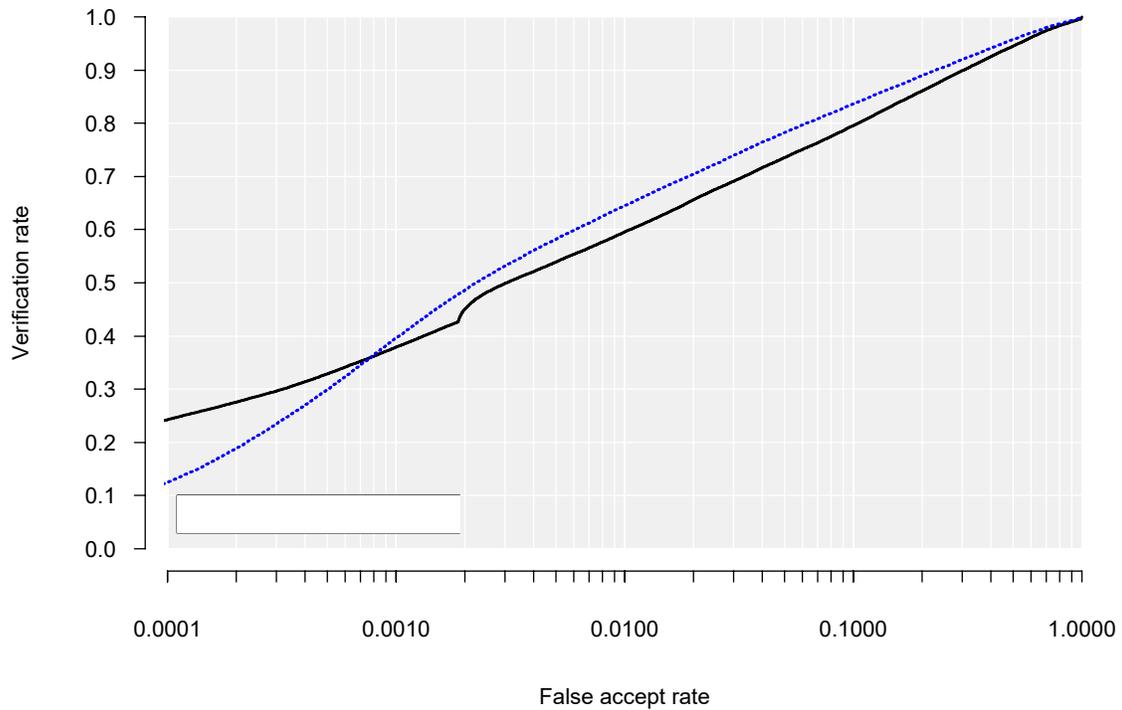


Figure L.10 – Normalized and non-normalized ROC for C-Vis. This figure corresponds to Figure 9 in the *FRVT 2002: Evaluation Report*.

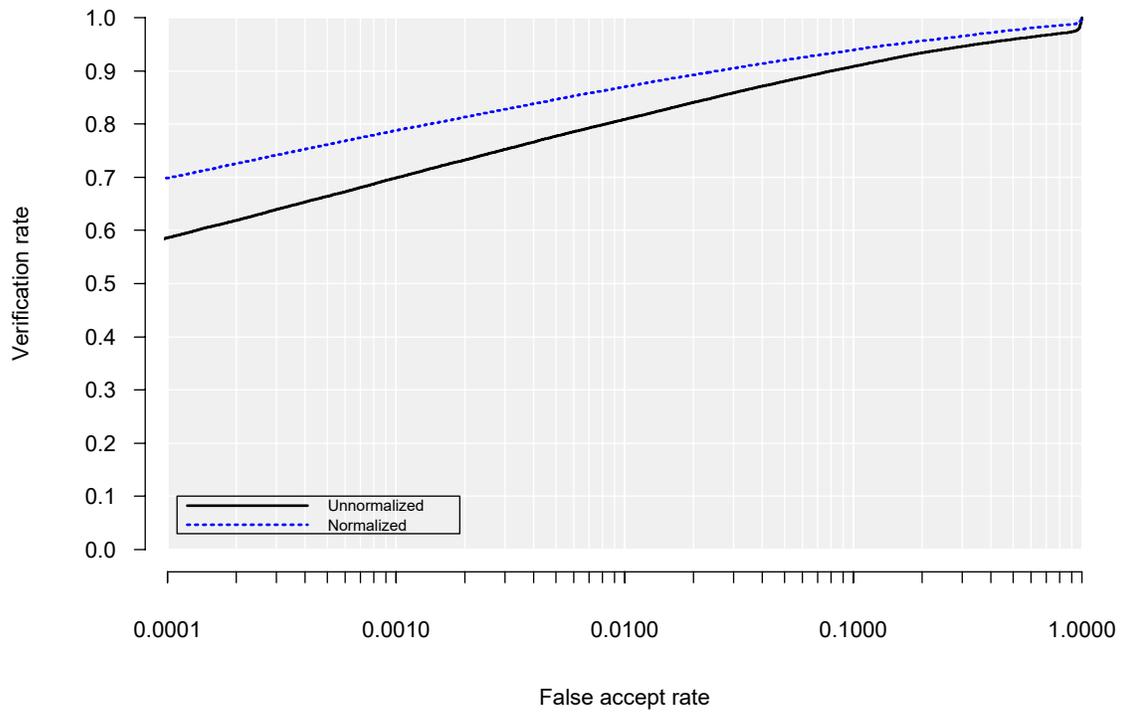


Figure L.11- Normalized and non-normalized ROC for Eyematic. This figure corresponds to Figure 9 in the FRVT 2002: Evaluation Report.

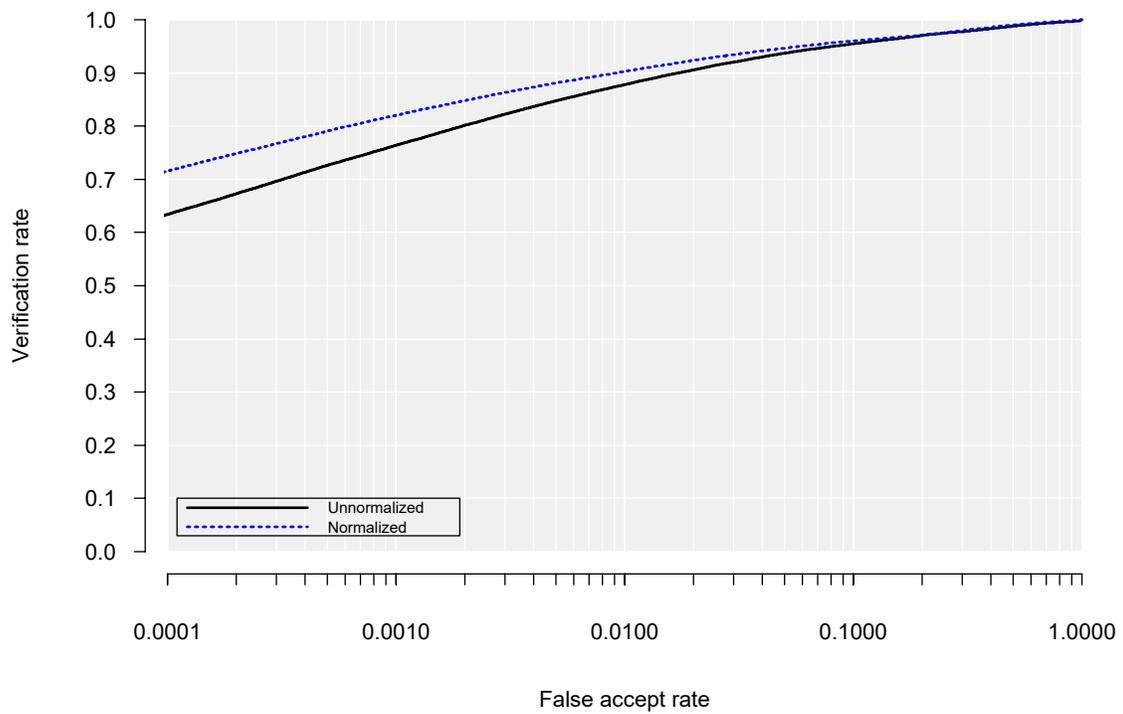


Figure L. 12 - Normalized and non-normalized ROC for Identix. This figure corresponds to Figure 9 in the FRVT 2002: Evaluation Report.

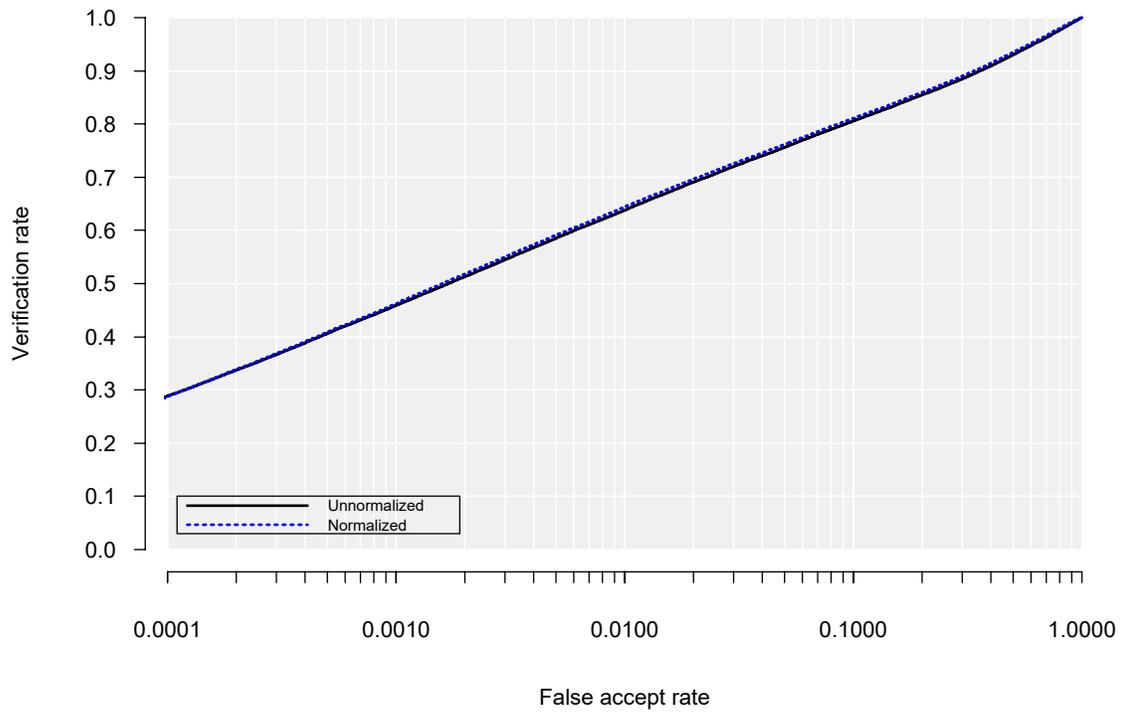


Figure L. 13 - Normalized and non-normalized ROC for Viisage. This figure corresponds to Figure 9 in the FRVT 2002: Evaluation Report.

L.3 MCInt Frontal Face Experiment

Figure 22 in the *FRVT 2002: Evaluation Report* was a comparison of system performance for the verification task using various frontal face probes. Figure 28 showed identification results for the same experiments. Figure L.14 through Figure L.23 provide full ROC and CMC charts for each lighting condition, as outlined in the following table.

Lighting Condition	Performance Curve	Figure Number
Indoor (same day)	ROC	Figure L.14
	CMC	Figure L.15
Indoor (same day, overhead)	ROC	Figure L.16
	CMC	Figure L.17
Indoor (different day)	ROC	Figure L.18
	CMC	Figure L.19
Outdoor (same day)	ROC	Figure L.20
	CMC	Figure L.21
Outdoor (different day)	ROC	Figure L.22
	CMC	Figure L.23

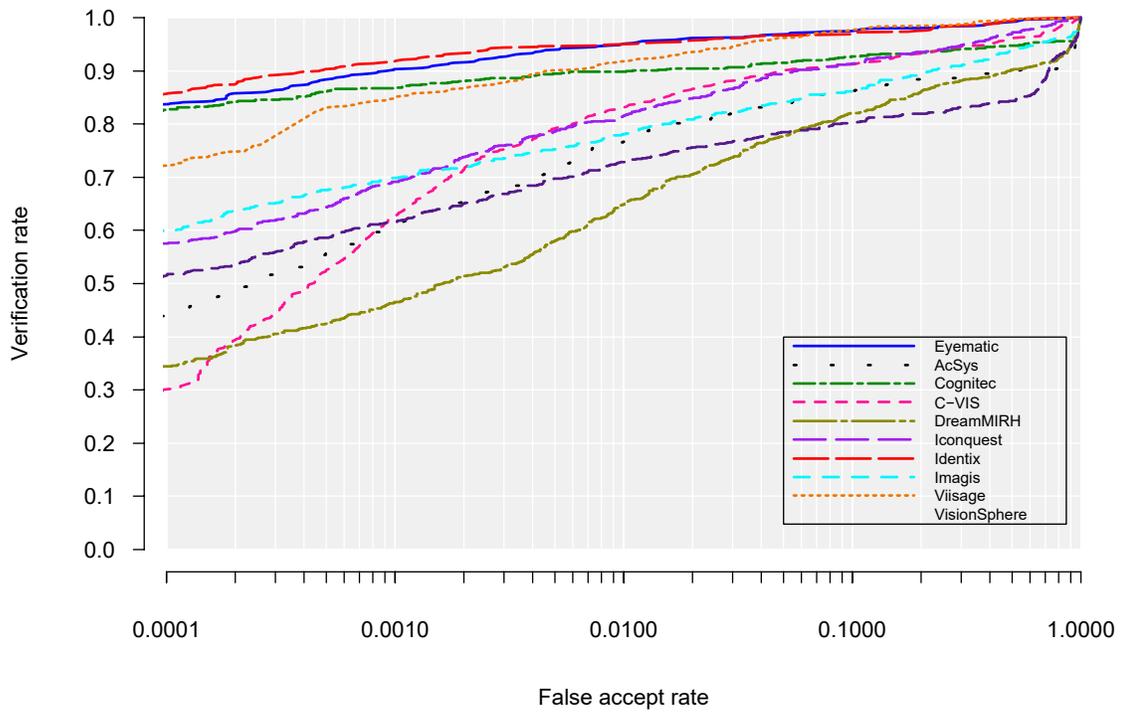


Figure L.14 - Verification Performance for Indoor (same day) Probes. This figure corresponds to Figure 22 in the *FRVT 2002: Evaluation Report*.

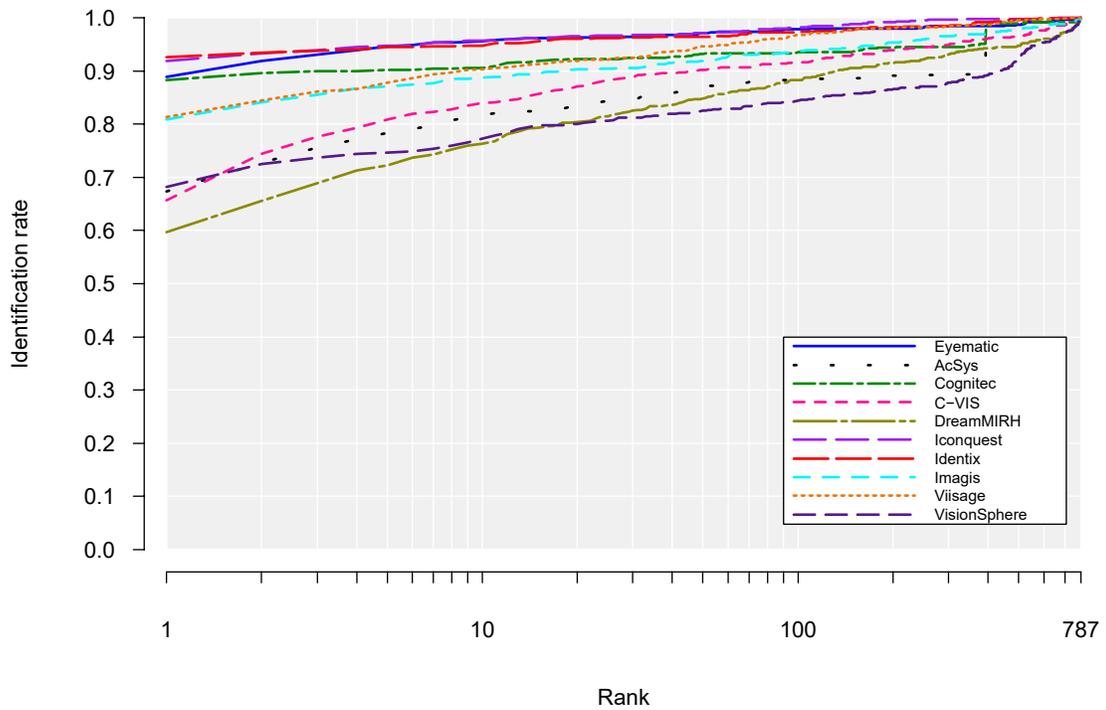


Figure L.15 - Identification Performance for Indoor (same day) Probes. This figure corresponds to Figure 28 in the *FRVT 2002: Evaluation Report*.

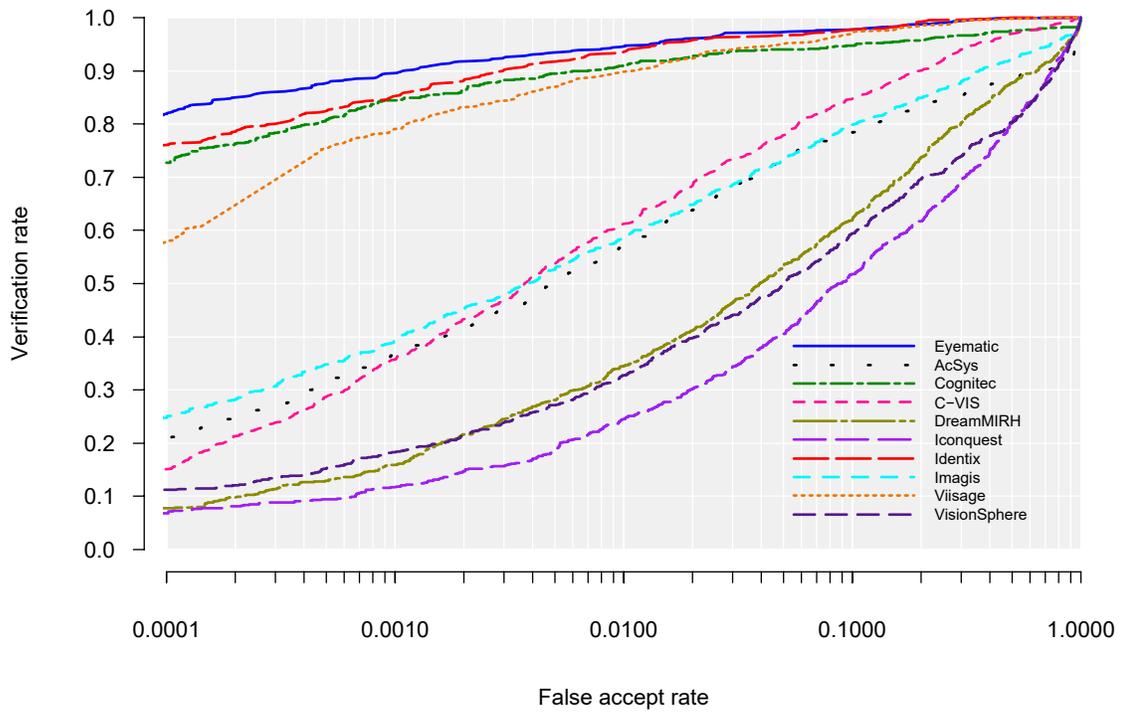


Figure L.16 - Verification Performance for Indoor (same day, overhead) Probes. This figure corresponds to Figure 22 in the *FRVT 2002: Evaluation Report*.

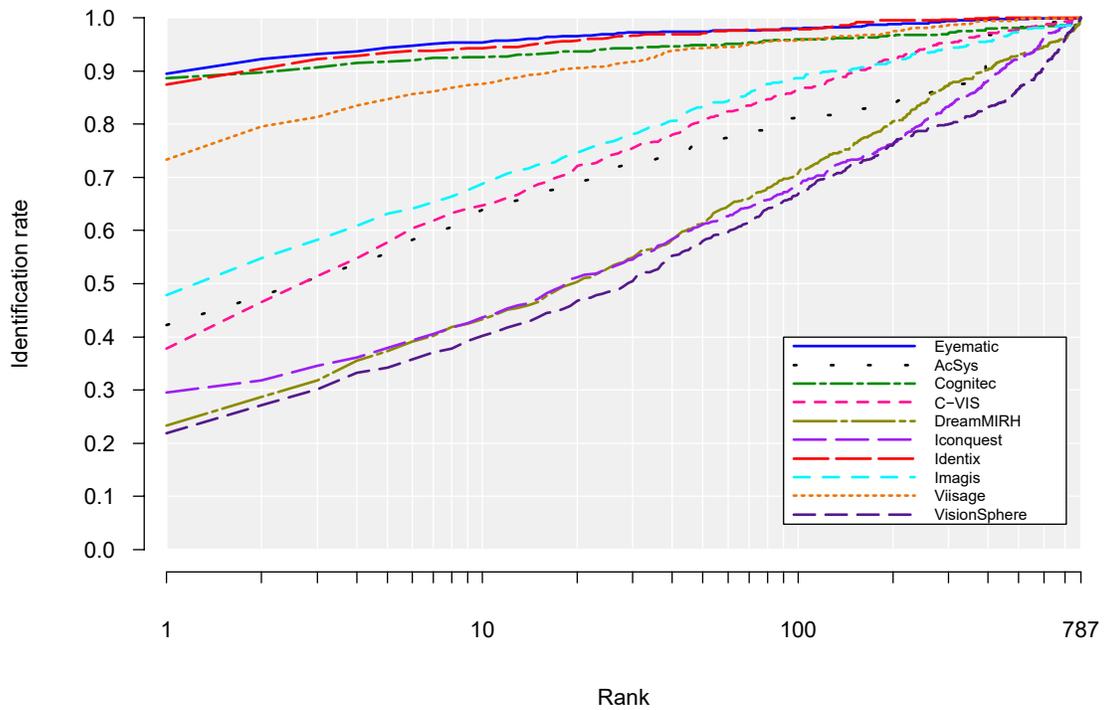


Figure L.17 - Identification Performance for Indoor (same day, overhead) Probes. This figure corresponds to Figure 28 in the *FRVT 2002: Evaluation Report*.

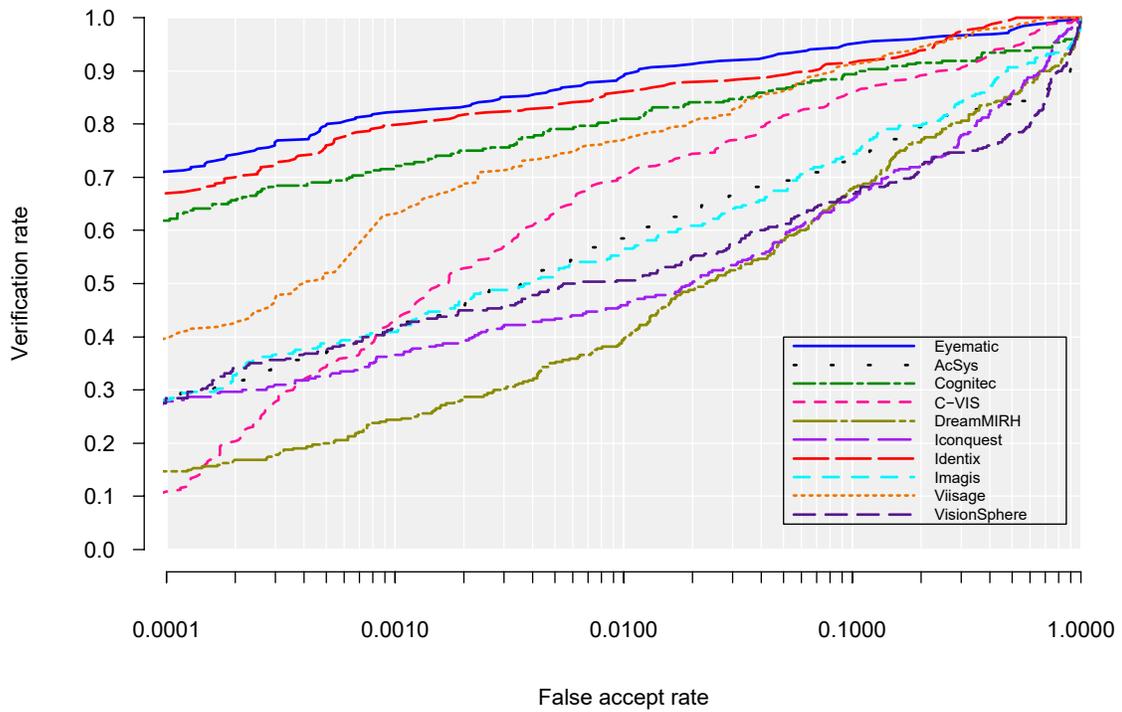


Figure L.18 - Verification Performance for Indoor (different day) Probes. This figure corresponds to Figure 22 in the *FRVT 2002: Evaluation Report*.

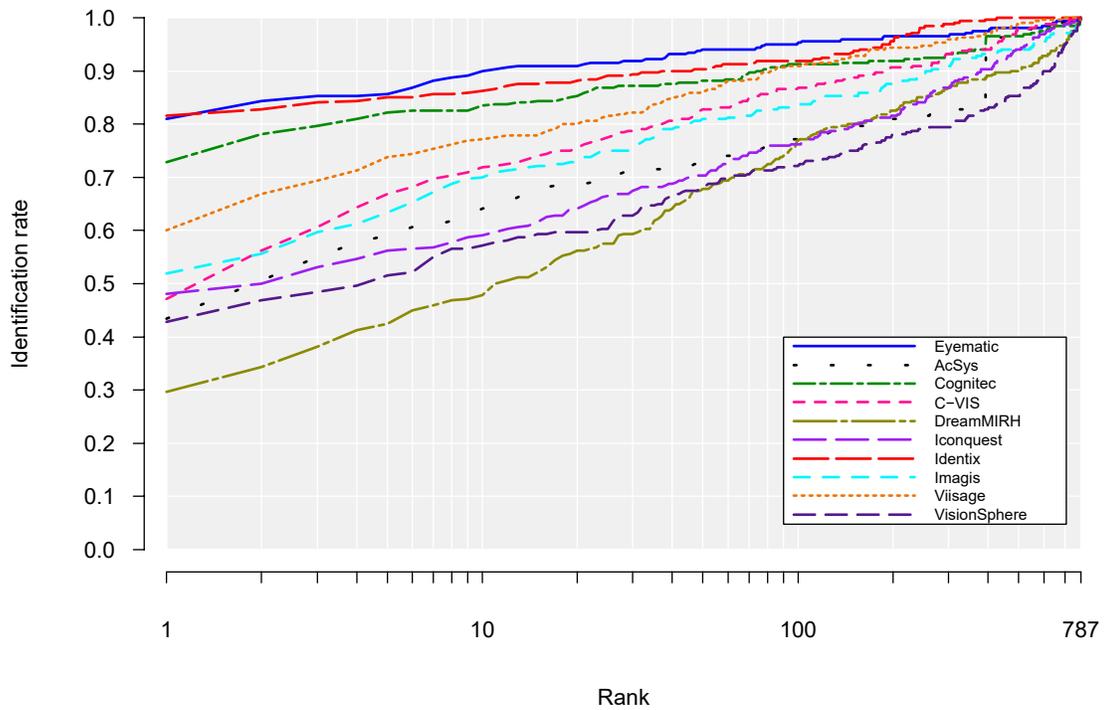


Figure L.19 - Identification Performance for Indoor (different day) Probes. This figure corresponds to Figure 28 in the *FRVT 2002: Evaluation Report*.

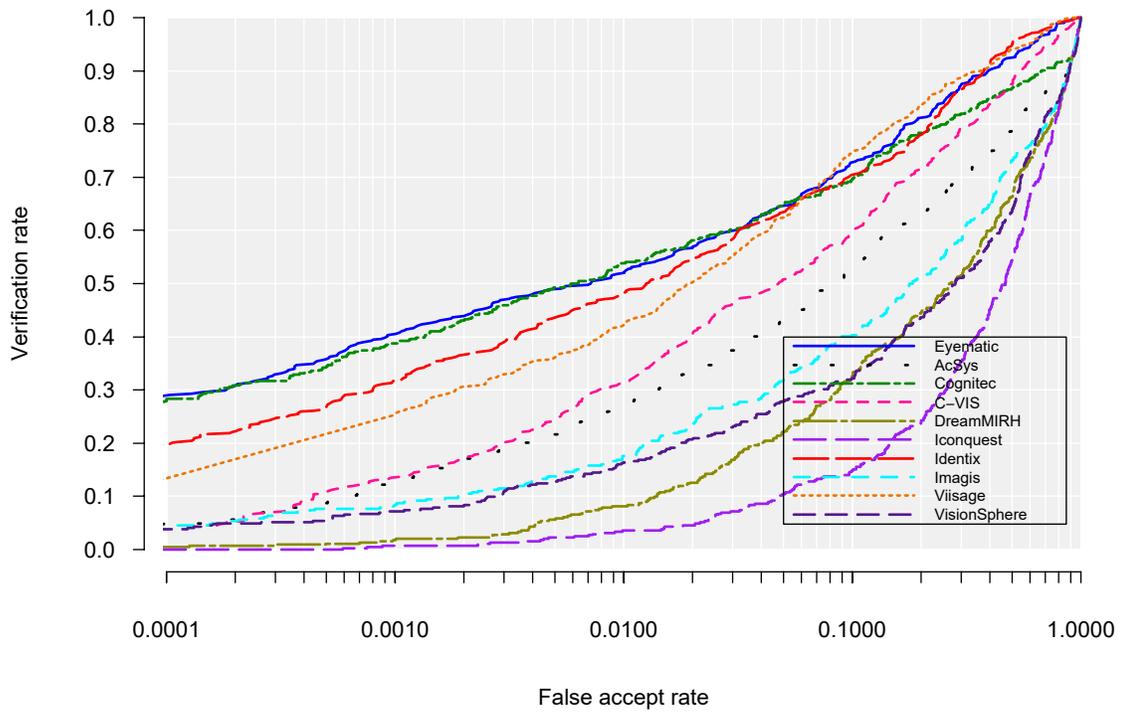


Figure L.20 - Verification Performance for Outdoor (same day) Probes. This figure corresponds to Figure 22 in the *FRVT 2002: Evaluation Report*.

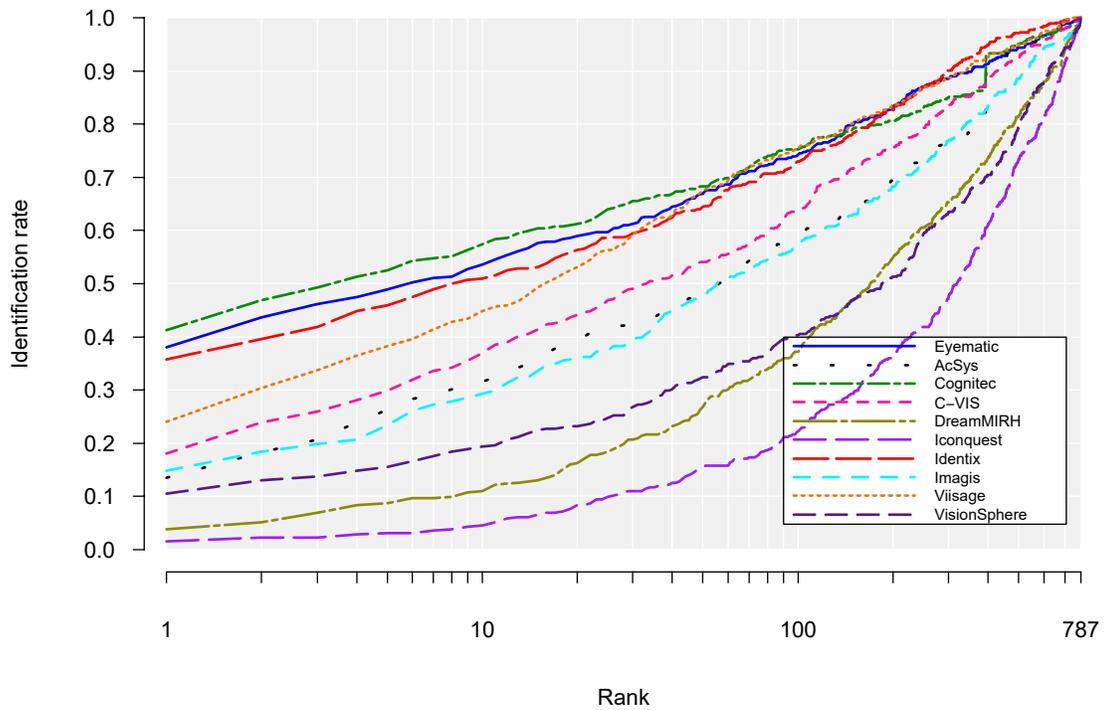


Figure L.21 - Identification Performance for Outdoor (same day probes). This figure corresponds to Figure 28 in the *FRVT 2002: Evaluation Report*.

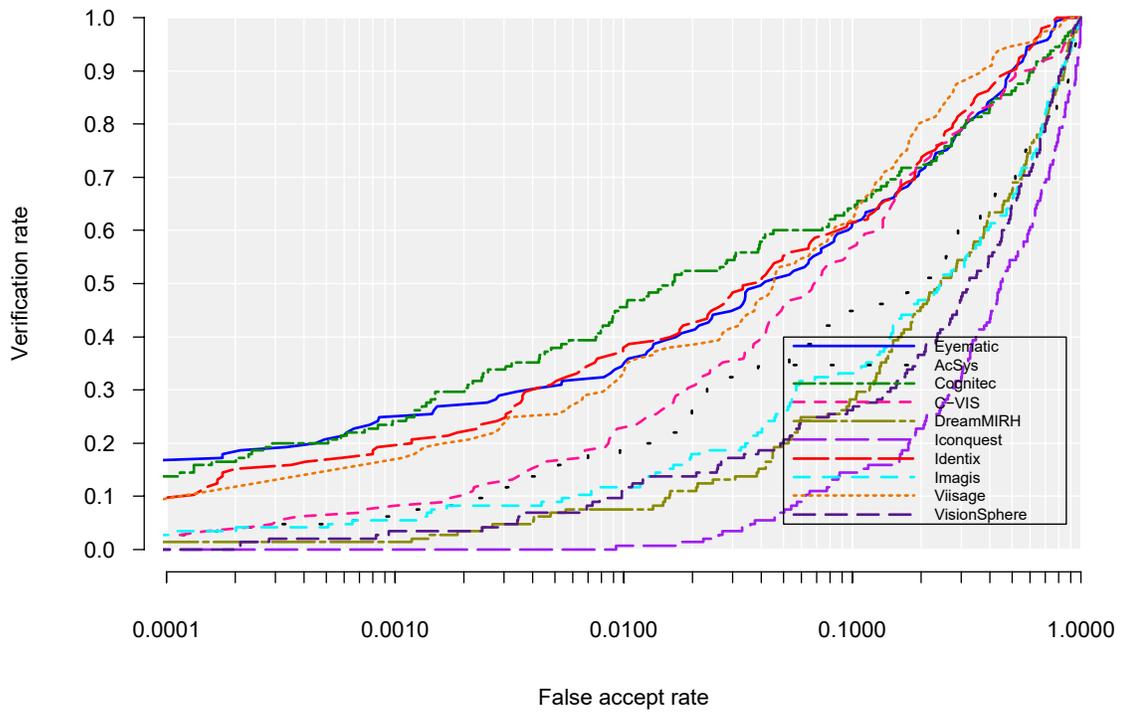


Figure L.22 - Verification Performance on Outdoor (different day) Probes. This figure corresponds to Figure 22 in the *FRVT 2002: Evaluation Report*.

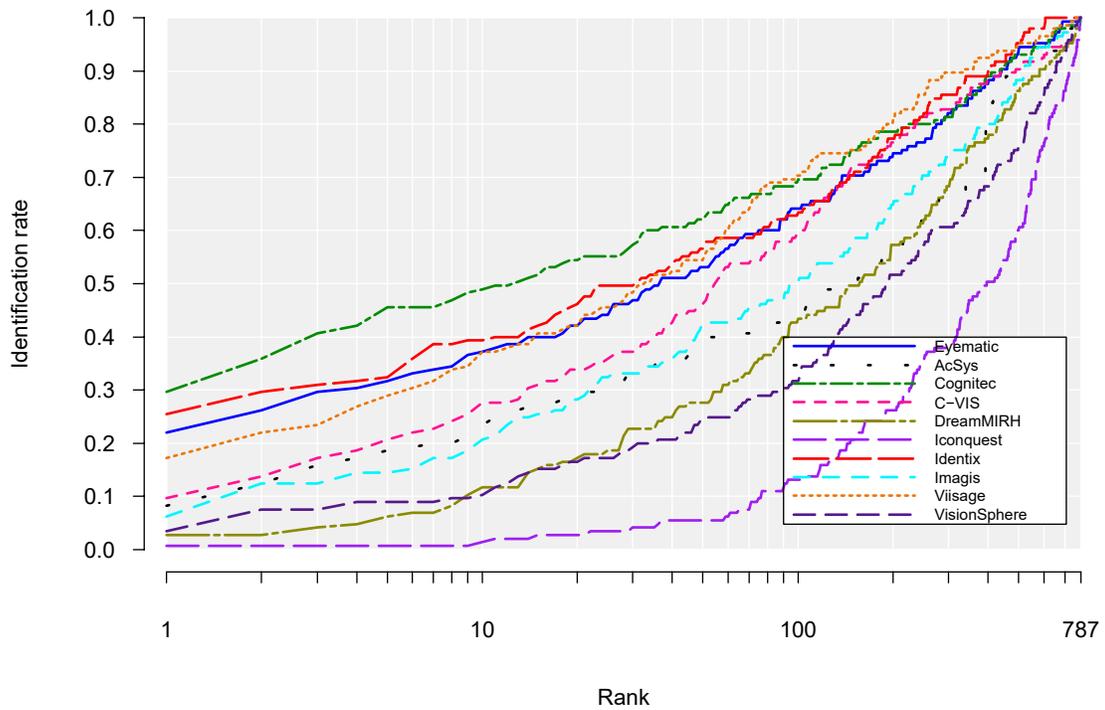


Figure L.23 - Identification Performance on Outdoor (different day) Probes. This figure corresponds to Figure 28 in the *FRVT 2002: Evaluation Report*.

L.4 MCInt Morphable Model Experiment

Figure 23 in the *FRVT 2002: Evaluation Report* was a comparison of verification system performance using probes of non-frontal poses as well as 3D Morphable reconstructions. Figure 29 showed identification performance for the same experiments. Figure L.24 through Figure L.41 provide full ROC and CMC charts for each pose and reconstruction, as outlined in the following table.

Pose	Performance Curve	Figure Number
Front (Morph)	ROC	Figure L.24
	CMC	Figure L.25
45L	ROC	Figure L.26
	CMC	Figure L.27
45L (Morph)	ROC	Figure L.28
	CMC	Figure L.29
45R	ROC	Figure L.30
	CMC	Figure L.31
45R (Morph)	ROC	Figure L.32
	CMC	Figure L.33
30U	ROC	Figure L.34
	CMC	Figure L.35
30U (Morph)	ROC	Figure L.36
	CMC	Figure L.37
30D	ROC	Figure L.38
	CMC	Figure L.39
30D (Morph)	ROC	Figure L.40
	CMC	Figure L.41

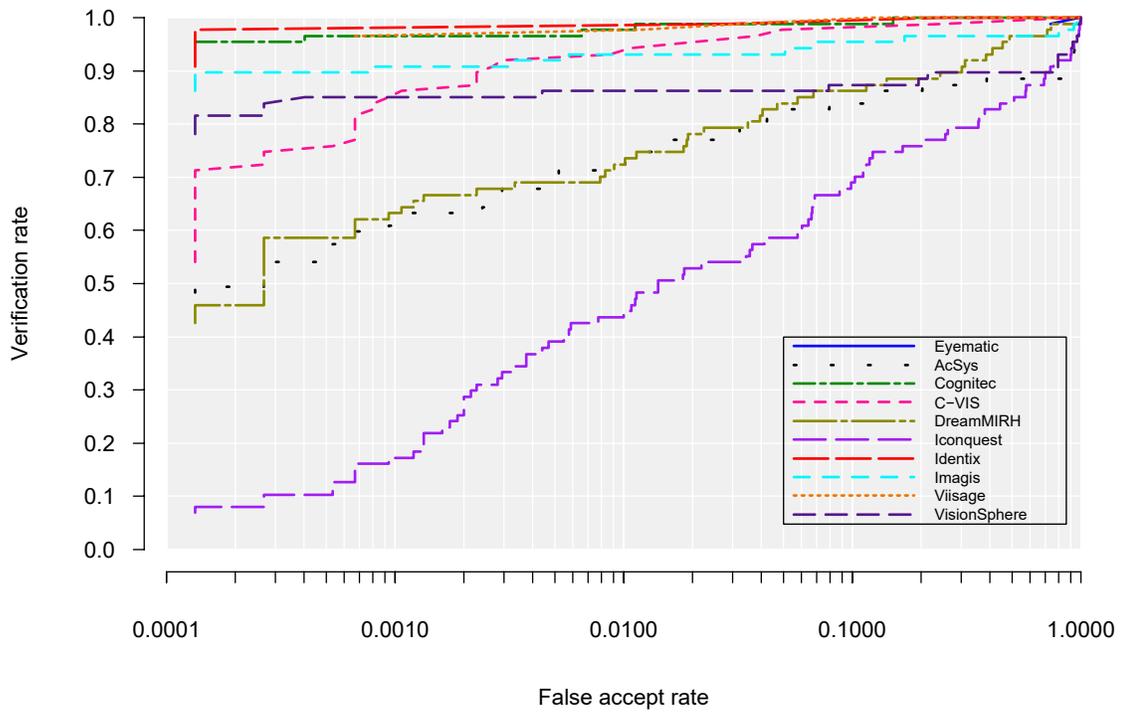


Figure L.24 - Verification Performance on Front (Morph) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

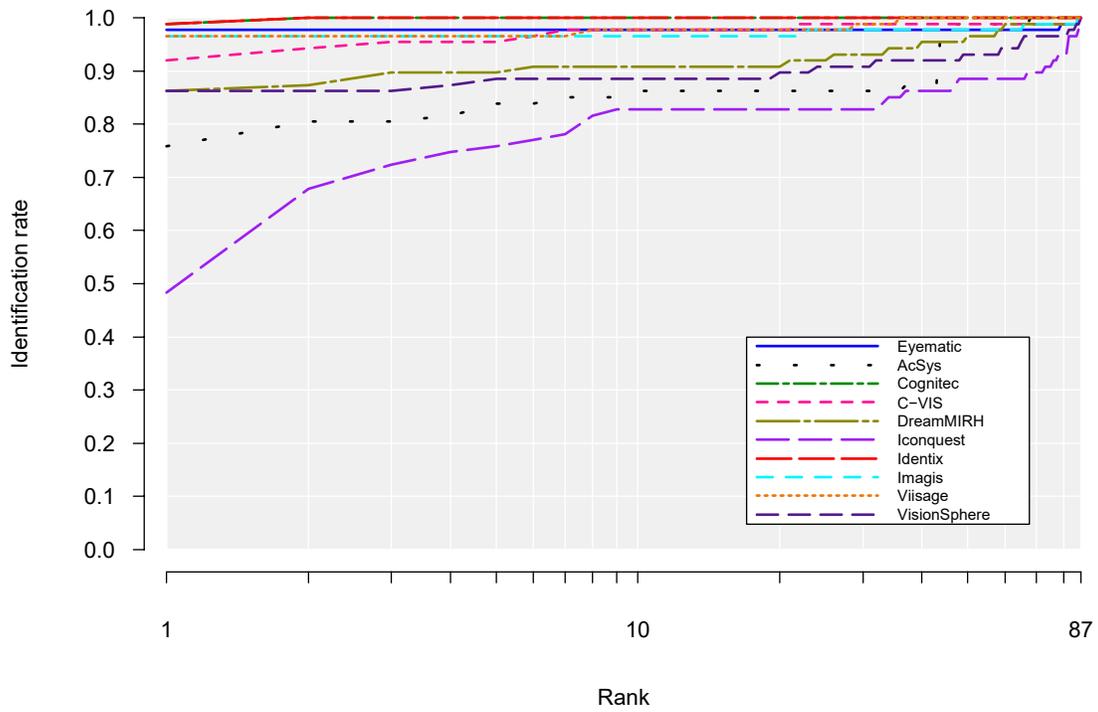


Figure L.25 - Identification Performance on Front (Morph) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

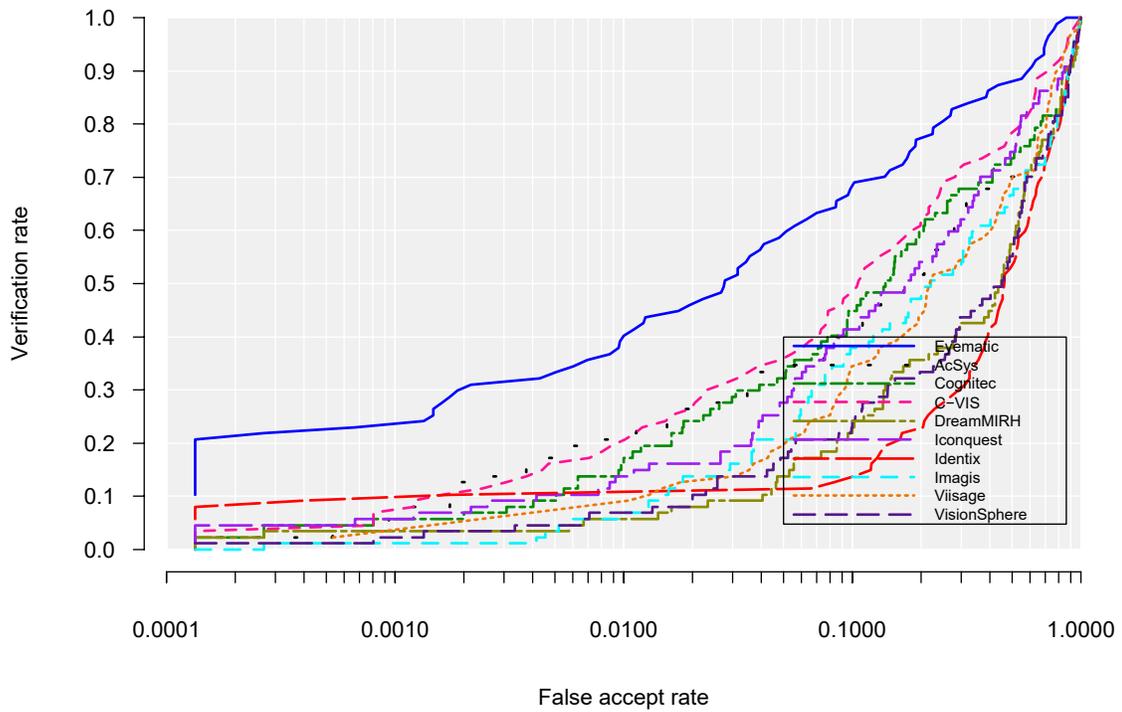


Figure L.26 - Verification Performance on 45 Degree (Left) Pose. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

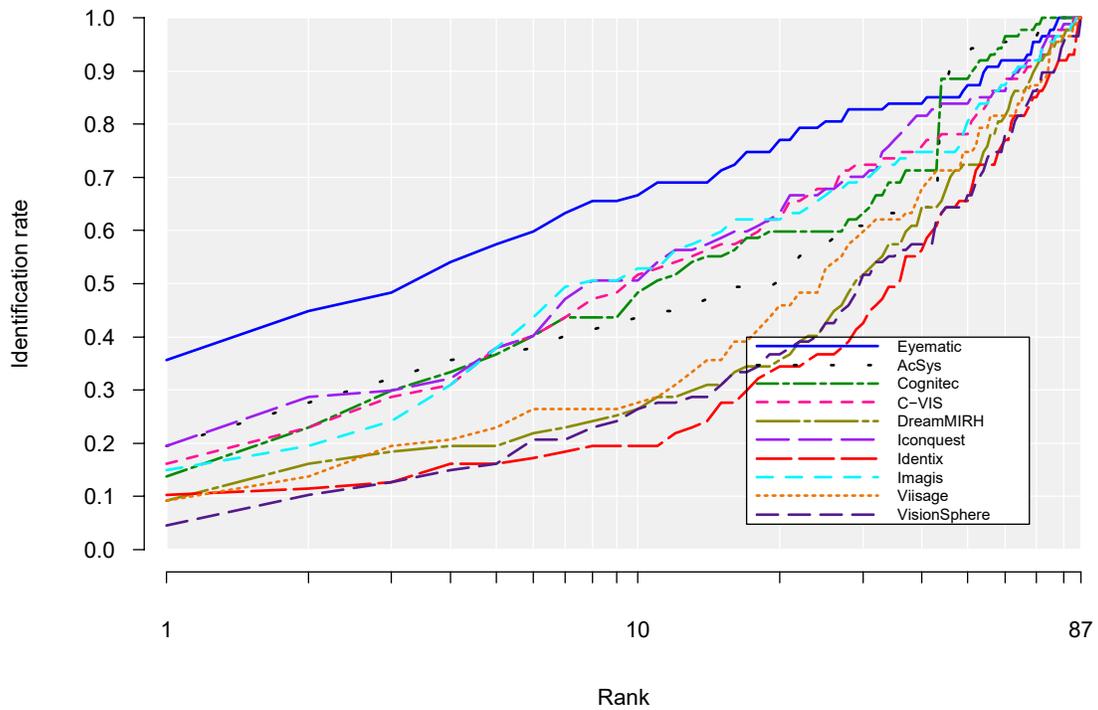


Figure L.27 - Identification Performance on 45 Degree (Left) Pose. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

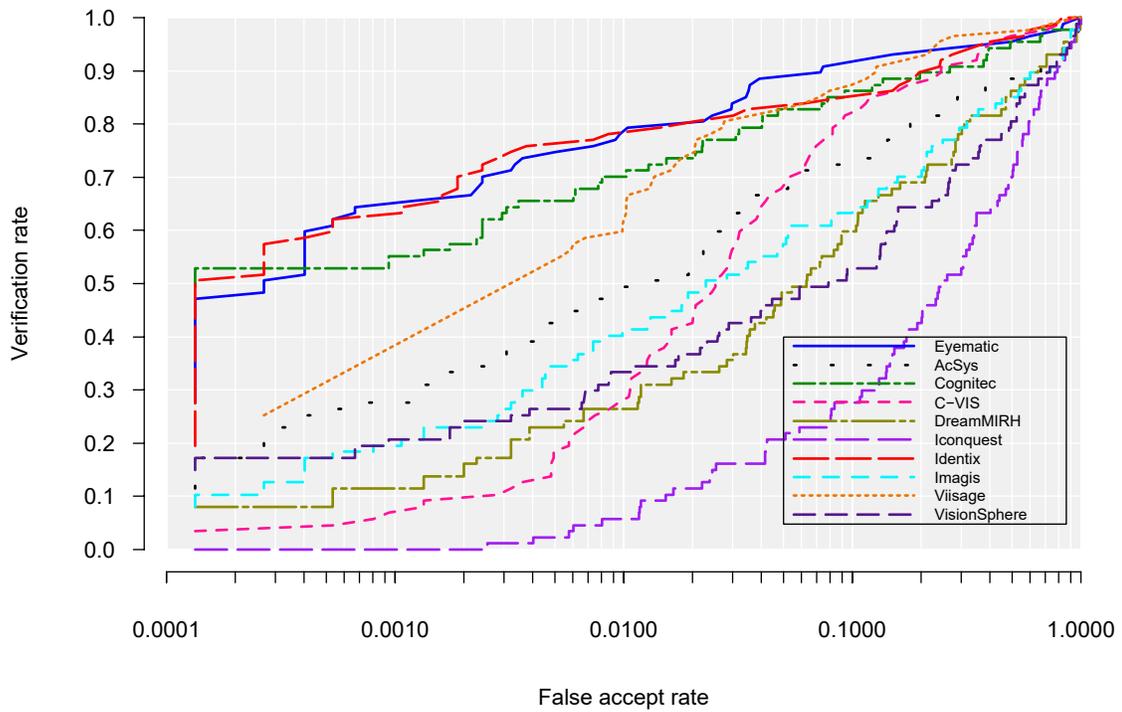


Figure L.28 - Verification Performance on 45 Degree (Left, Morphed) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

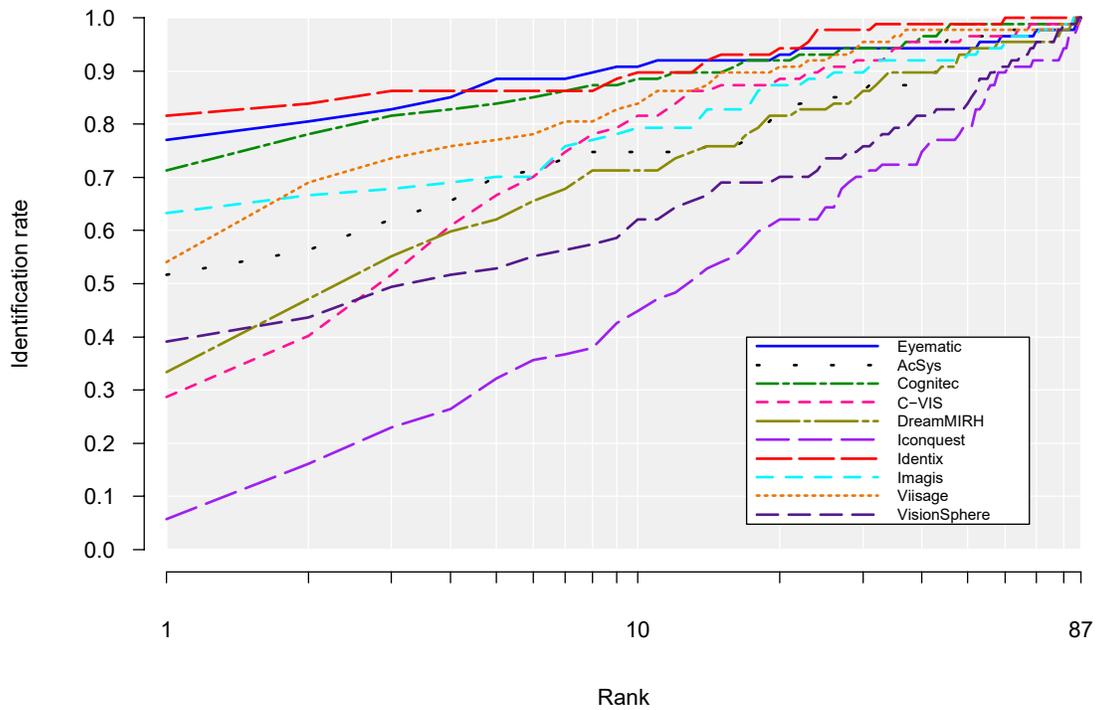


Figure L.29 - Identification Performance on 45 Degree (Left, Morphed) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

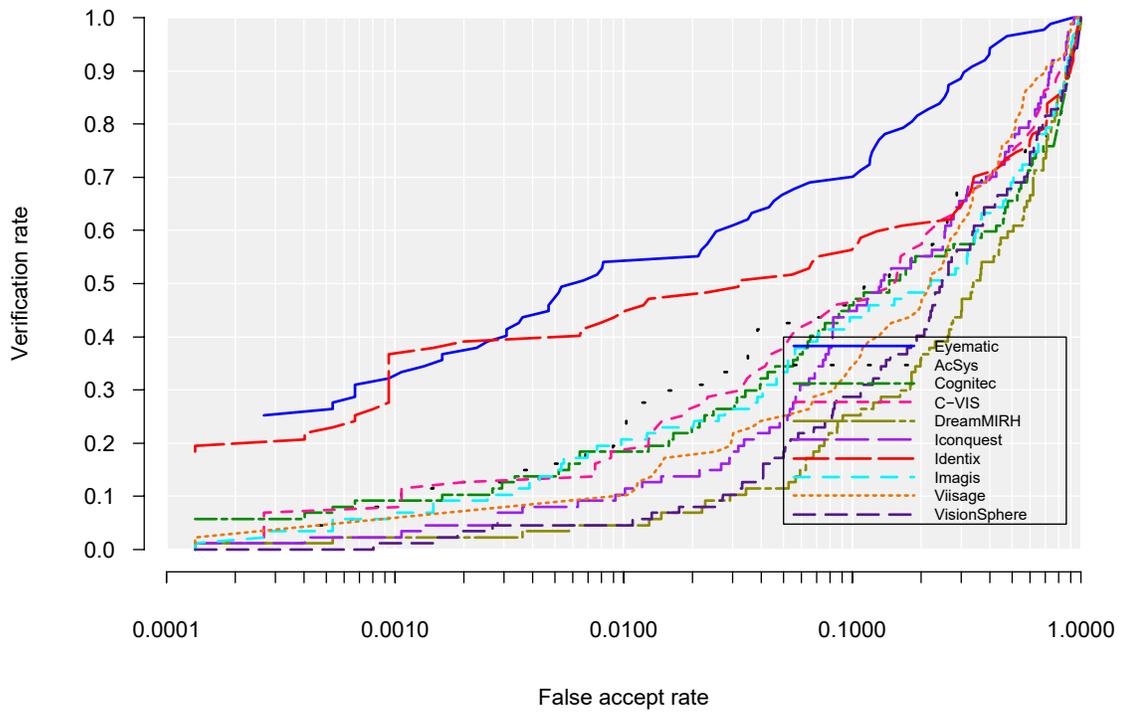


Figure L.30 - Verification Performance on 45 Degree (Right) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

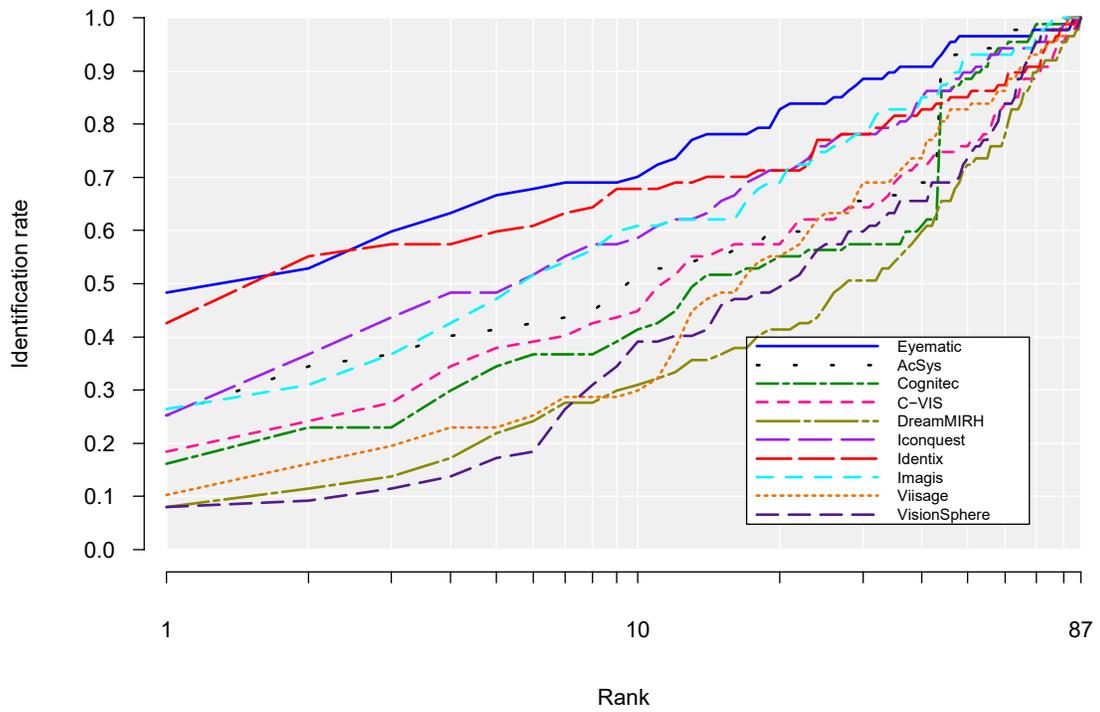


Figure L.31 - Identification Performance on 45 Degree (Right) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

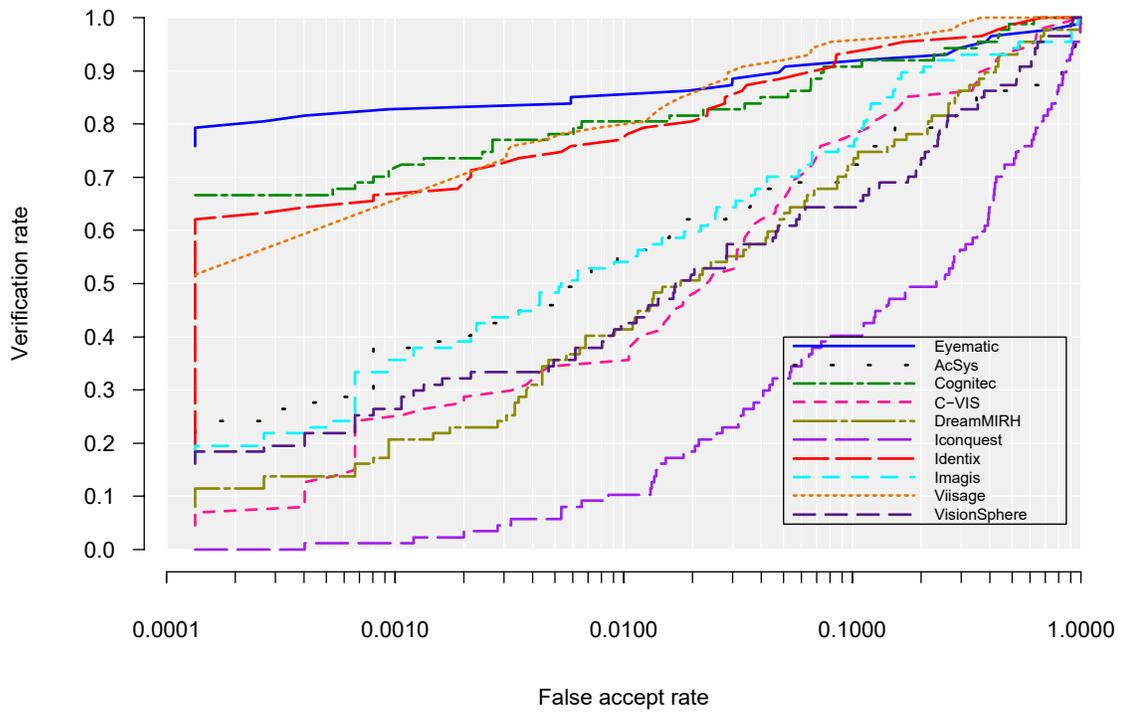


Figure L.32 - Verification Performance of 45 Degree (Right, Morphed) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

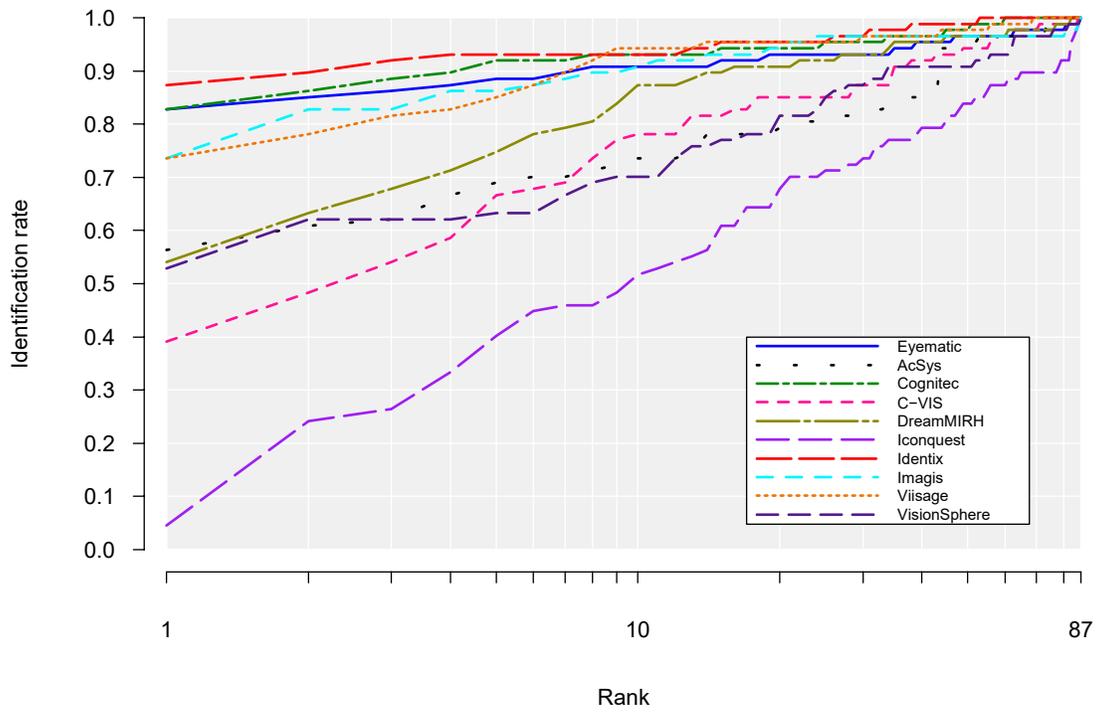


Figure L.33 - Identification Performance of 45 Degree (Right, Morphed) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

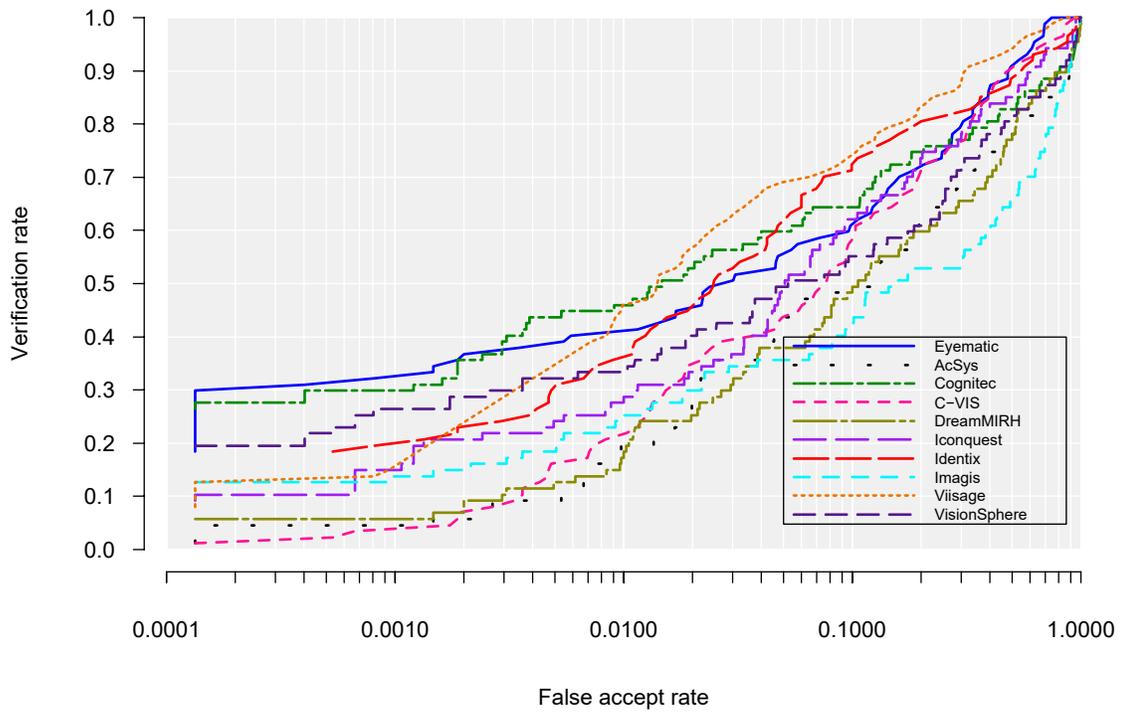


Figure L.34- Verification Performance of 30 Degree (Up) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

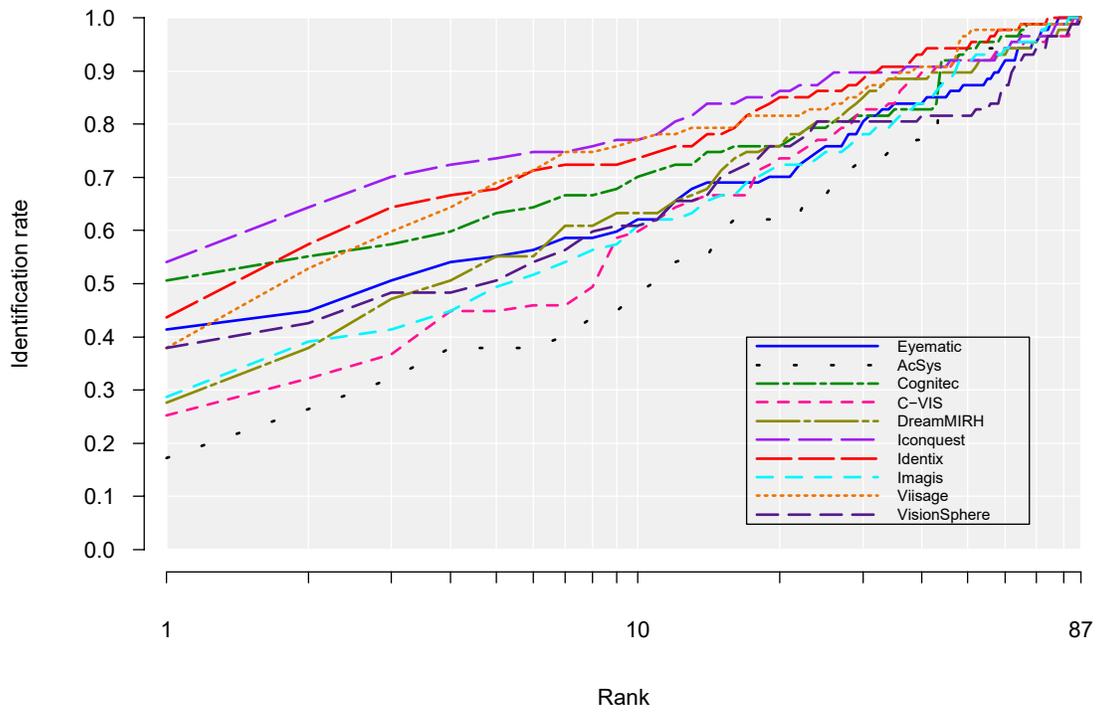


Figure L.35 - Identification Performance of 30 Degree (Up) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

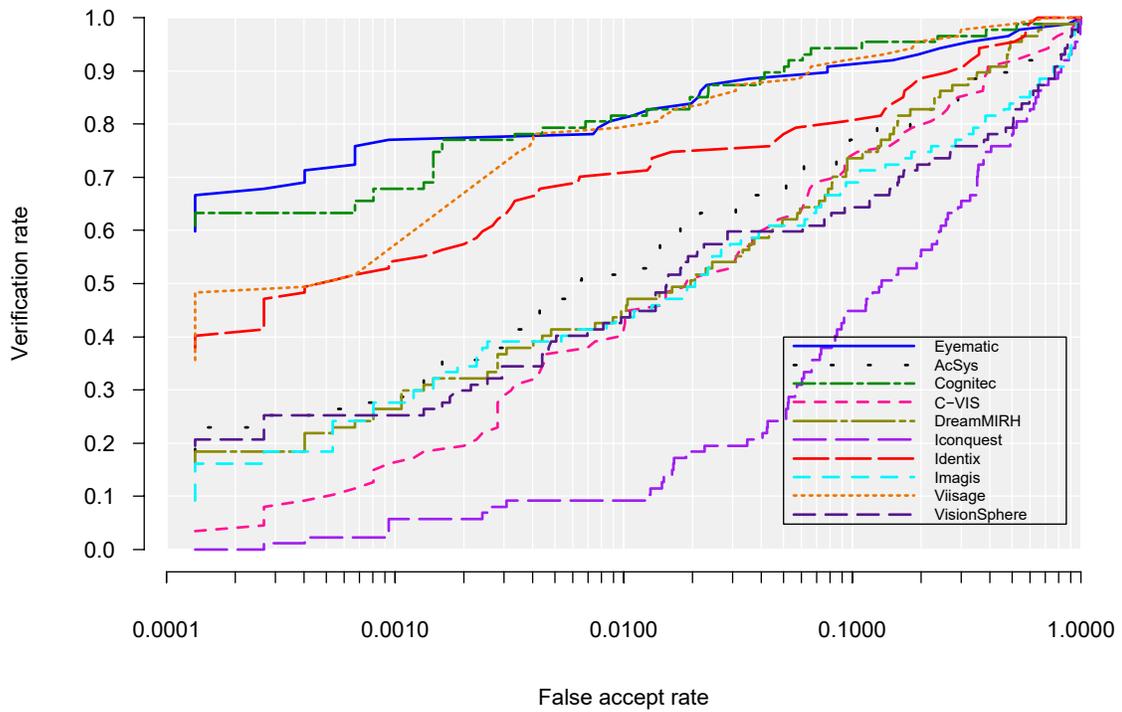


Figure L.36 - Verification Performance of 30 Degree (Up, Morphed) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

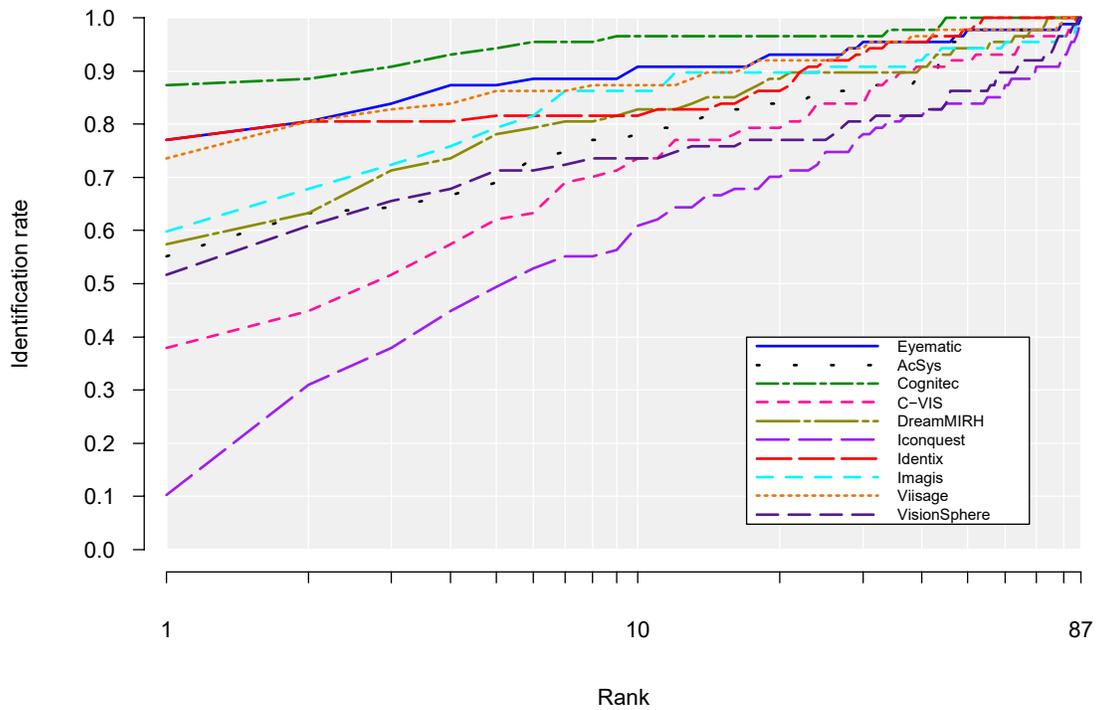


Figure L.37 - Identification Performance of 30 Degree (Up, Morphed) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

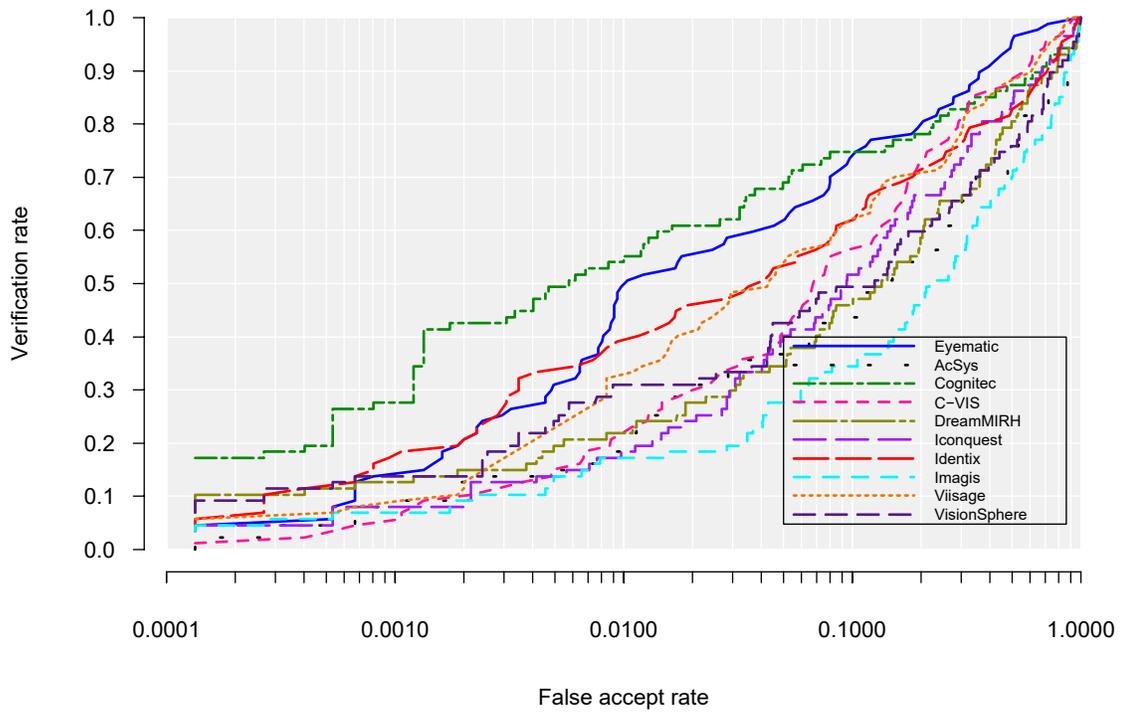


Figure L.38 - Verification Performance of 30 Degree (Down) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

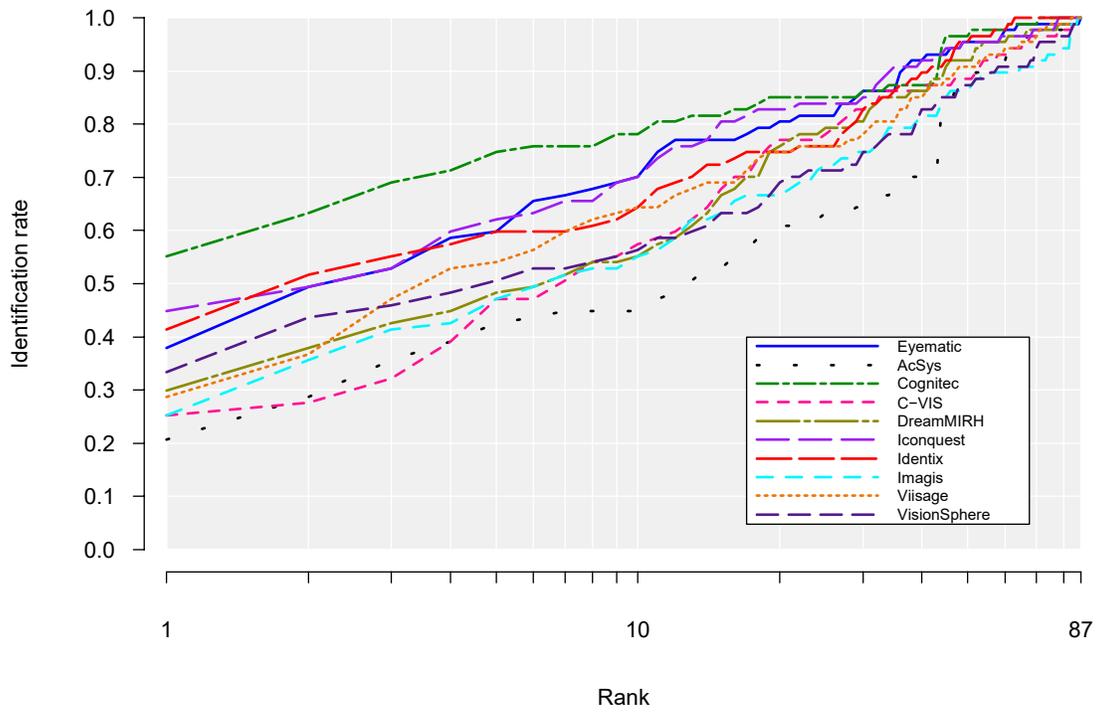


Figure L.39 - Identification Performance of 30 Degree (Down) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

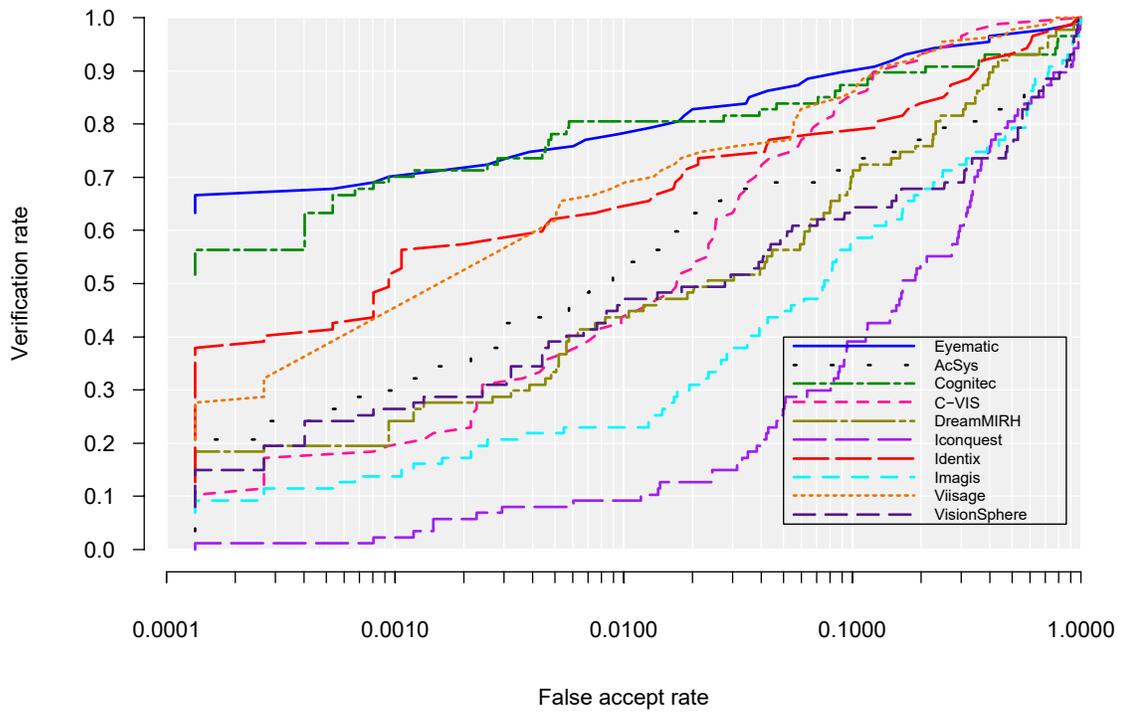


Figure L.40 - Verification Performance for 30 Degree (Down, Morphed) Probes. This figure corresponds to Figure 23 in the *FRVT 2002: Evaluation Report*.

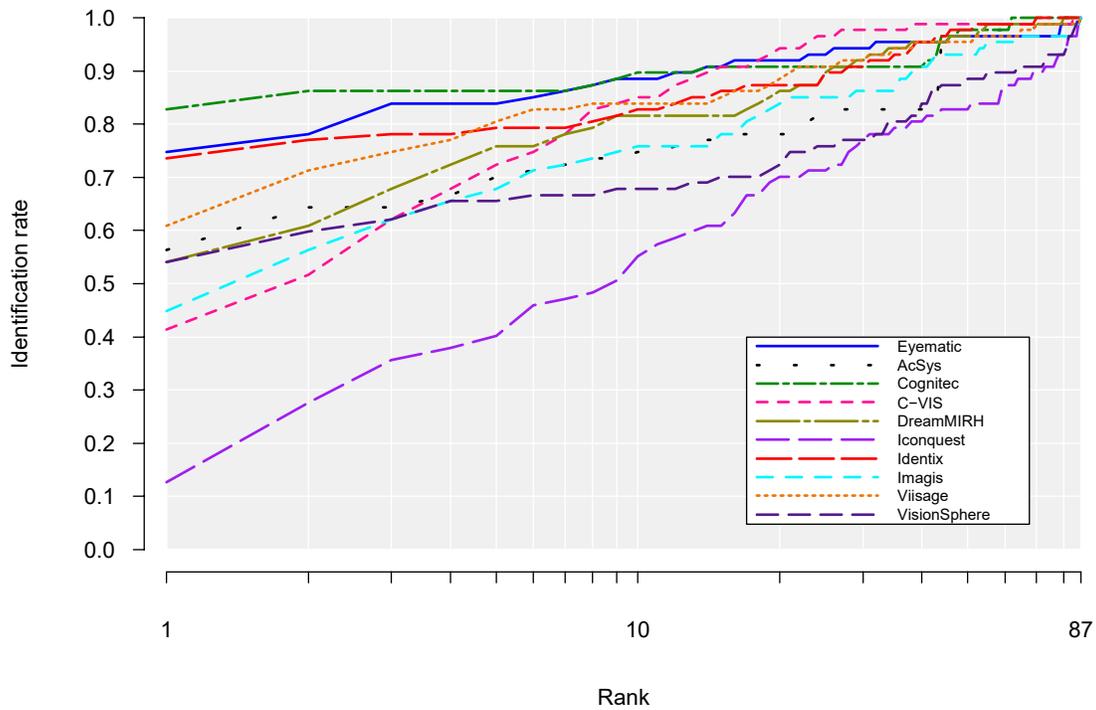


Figure L.41 - Identification Performance for 30 Degree (Down, Morphed) Probes. This figure corresponds to Figure 29 in the *FRVT 2002: Evaluation Report*.

L.5 MCInt Still-Video Experiment

Figure 24 in the *FRVT 2002: Evaluation Report* was a comparison of system performance using still and video probes. Figure 30 showed identification performance for the same experiments. Figure L.42 through Figure L.45 provide full ROC and CMC charts for each.

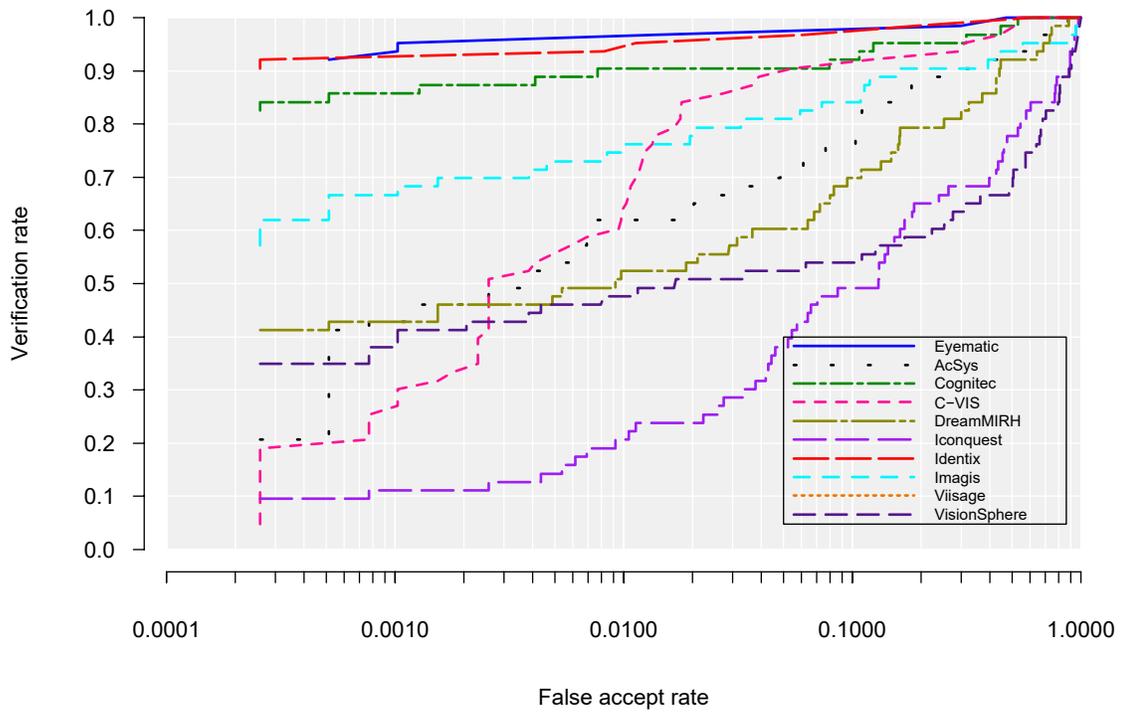


Figure L.42 - Verification Performance for Frontal (Still, Different Day) Probes. This figure corresponds to Figure 24 in the *FRVT 2002: Evaluation Report*.

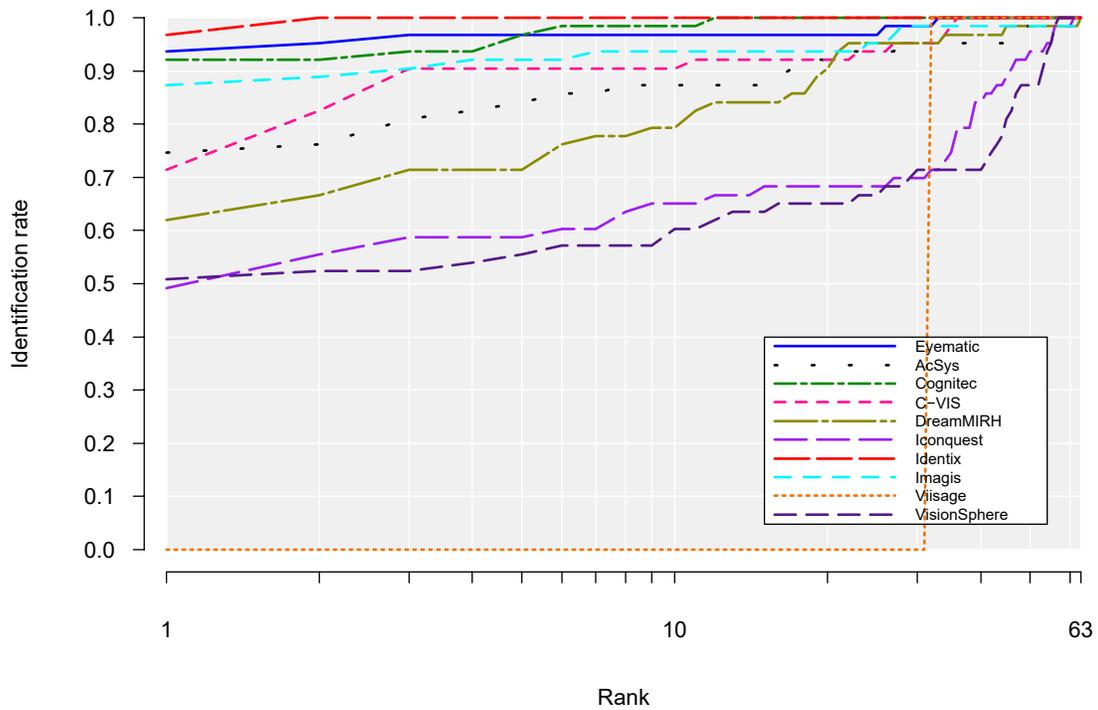


Figure L.43 - Identification Performance for Frontal (Still, Different Day) Probes. This figure corresponds to Figure 30 in the *FRVT 2002: Evaluation Report*.

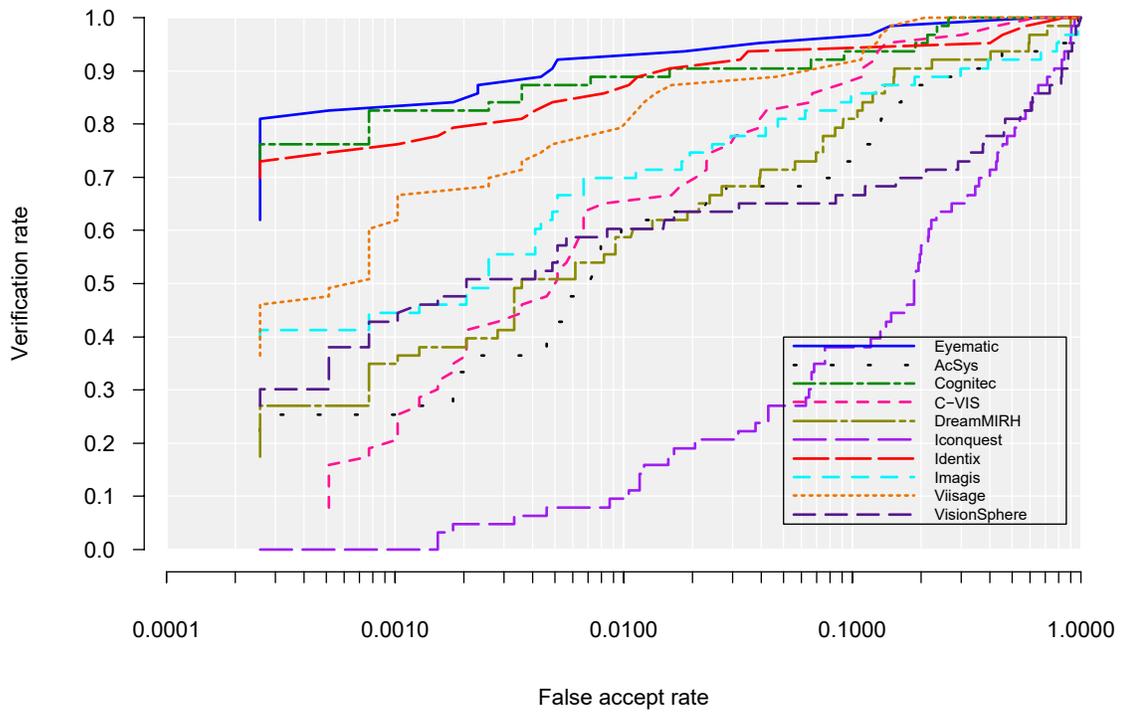


Figure L.44 - Verification Performance for Frontal (Speech "Video", Different Day) Probes. This figure corresponds to Figure 24 in the *FRVT 2002: Evaluation Report*.

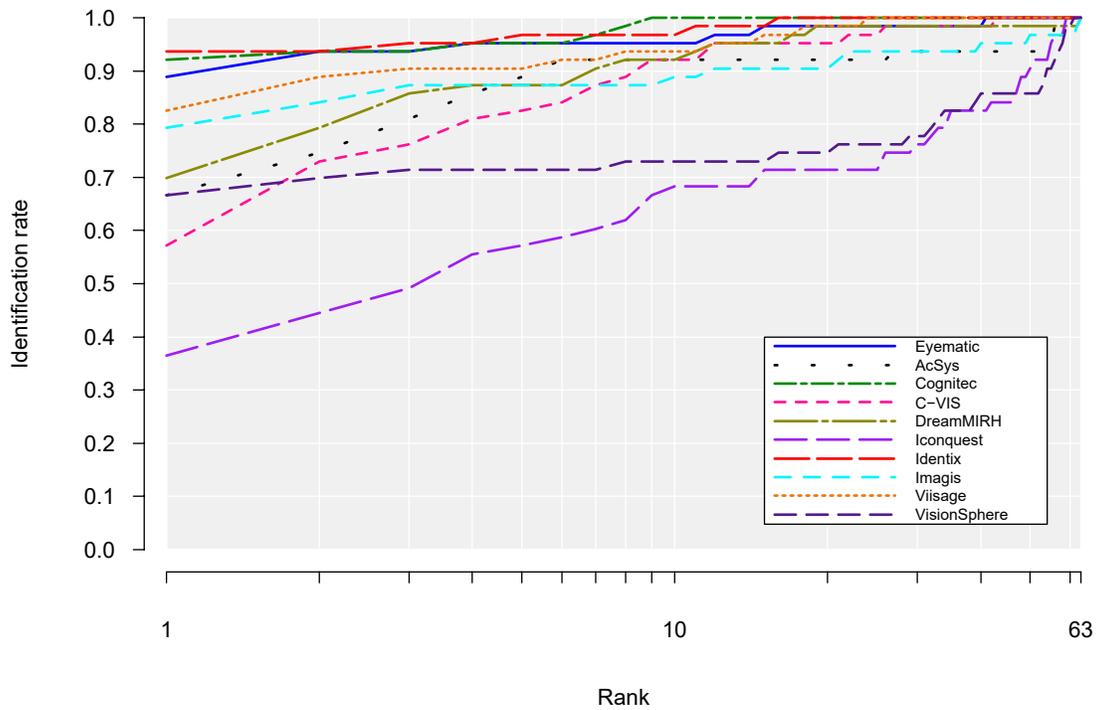


Figure L.45 - Identification Performance for Frontal (Speech "Video", Different Day) Probes. This figure corresponds to Figure 30 in the *FRVT 2002: Evaluation Report*.

L.6 HCInt Images in the MCInt Subtest.

FRVT 2002 Organizers placed a small number of images from the corpus that comprised the HCInt sub-test into the MCInt sub-test. The experiment in section L.6 has:

512 images in the gallery

512 images in the probe

1024 images total.

These are images of 512 randomly selected individuals from the HCInt corpus. The image in the gallery is the earliest image of that person, while the probe entry is the most recent image of that person.

This experiment was developed to see if any participants that only took the MCInt experiment could potentially do well on HCInt imagery and to see if there are any major deviations between scores for those participants that took both the MCInt and HCInt tests.

Verification and Identification performance for these images in the MCInt sub-test are provided in Figure L.46 and Figure L.47.

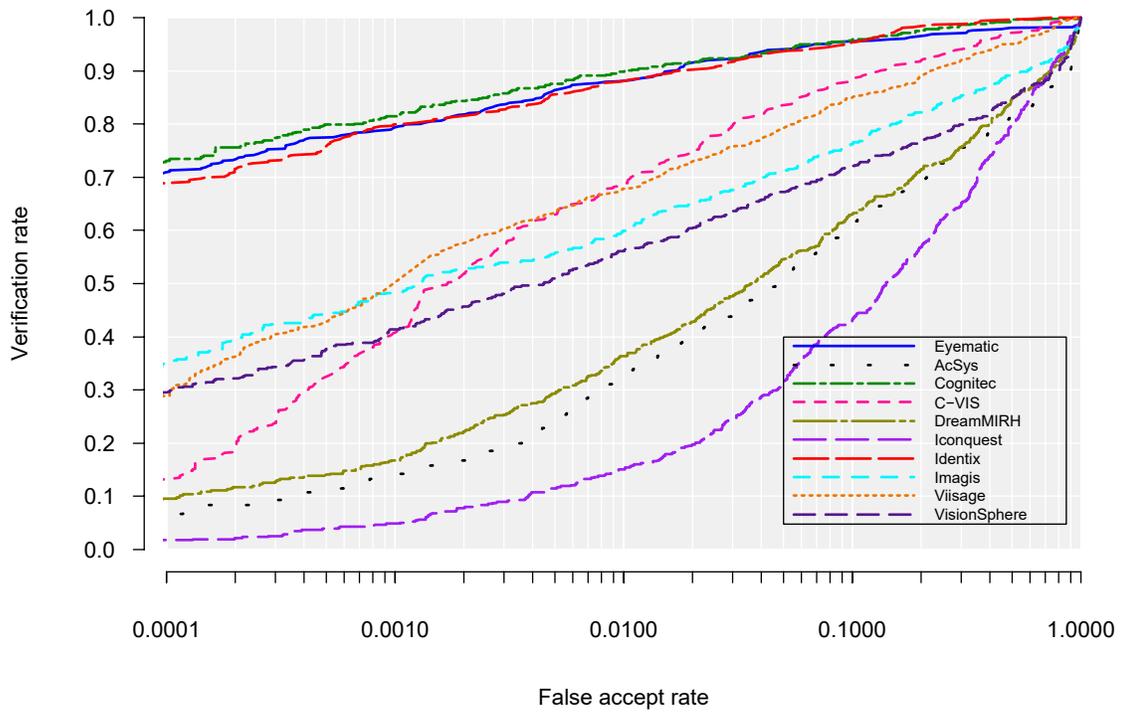


Figure L.46 - Verification Performance of Images from the HCInt Corpus in the MCInt sub-test.

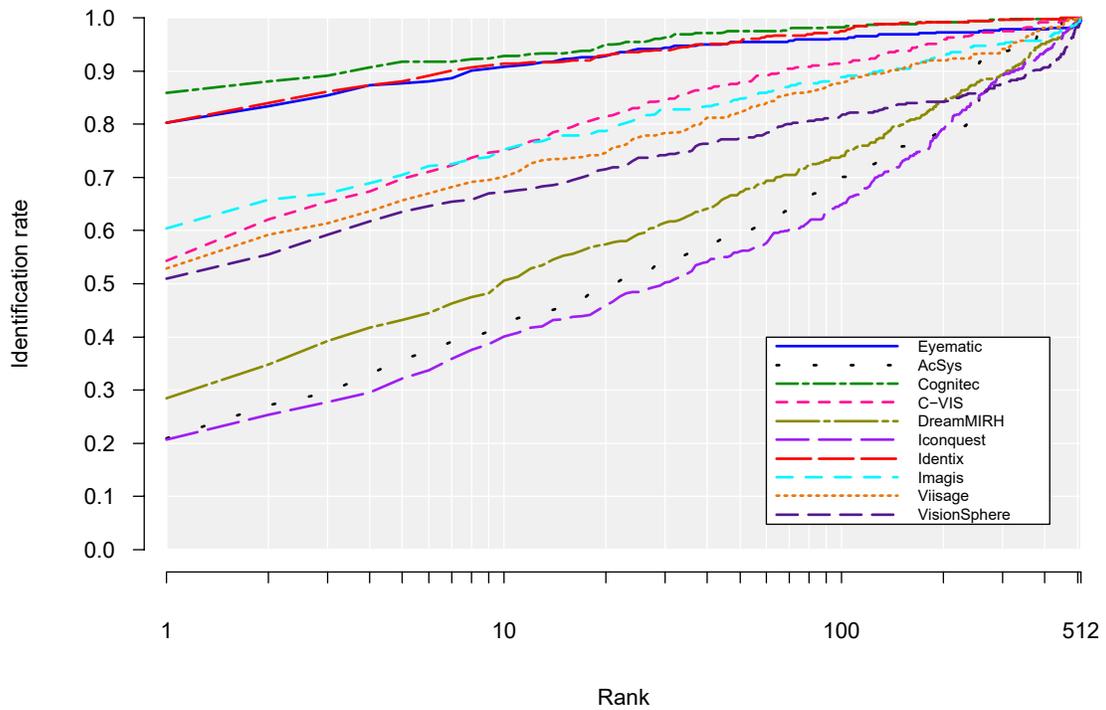


Figure L.47 - Identification Performance of Images from the HCInt Corpus in the MCInt sub-test

Appendix M

Participant Product Descriptions

FRVT 2002 Participants were required to submit a paper describing their face recognition approach, modifications made to current systems for FRVT 2002 and an estimated cost of the system(s) used in FRVT 2002. These documents were turned in to FRVT 2002 organizers prior to beginning the FRVT 2002 on-site evaluations.

Participant documents are the opinions of the Participants and are provided in this text for reference. Inclusion of these documents does **NOT** imply that FRVT 2002 Authors/Sponsors/Supporters agree with any statements or derivative results within those documents. The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology or any other FRVT 2002 Author/Sponsor/Supporter.



AcSys Biometrics Corporation
399 Pearl Street
Burlington, ON L7R 2M8 - Canada
Phone: (905) 634-4111
Fax: (905) 634-1101
E-mail: info@nxsgrp.com
Web Site: www.acsysbiometricscorp.com

AcSys Watchlist Application

Overview

AcSys Biometrics Corp. is pleased to participate in the 2002 Facial Recognition Vendor Test. AcSys Watchlist, the prototype we are introducing for FRVT 2002, applies the speed, accuracy, and intelligence of AcSys face recognition technology to a 2-D watch list application. AcSys Watchlist is our prototype for a new product line that represents an entirely new direction for AcSys face recognition technology.

AcSys Watchlist versus AcSys FRS

Our mature product line, AcSys FRS, is not compatible with the test procedures proposed by FRVT 2002. The focus of the AcSys FRS product line is on the reliable authentication of identity. It accomplishes this authentication through the comparison of a number of images (representing a three-dimensional view of the individual) against a biometric template that is itself a product of a substantial number of images. This technique allows AcSys FRS to verify faces presented at a variety of angles.

This three-dimensional approach to face recognition makes AcSys FRS an ideal technology in access control situations and for other applications where individuals wish to authenticate a claimed identity, or are otherwise attempting to be recognized in a cooperative manner within a live environment. The AcSys FRS product line is designed to provide unparalleled speed and accuracy in such situations using an approach to authentication that is radically different from that of other face recognition technologies.

AcSys FRS is designed to continue to learn an individual's face over time and provide quick, accurate, and reliable authentication of the individual time after time. In contrast, AcSys Watchlist looks for degrees of resemblance between a single facial image (or small set of images) and a database of biometric templates.

AcSys Watchlist Functionality

AcSys Watchlist is a tool for performing two-way analysis of similarities between facial images (on a query list) and biometric templates (on a target list). Whereas our established AcSys FRS line of products works on a 3-D principle, examining facial images within live video to determine or authenticate identity, AcSys Watchlist:

- Compares a 2-D still image and/or live video to biometric templates
- Compares a biometric templates against 2-D images
- Returns a list of top matches, along with similarity values (scores), and displays associated images
- Creates biometric templates from sets of image files
- Updates biometric templates with new images without increasing the size of the template

The system's ability to build biometric templates from a large number of images and update the templates with additional images without increasing template size makes it a scalable solution

for all watch list applications. The system's ability to recognize faces will improve with the addition of more images to templates, but the size of the template database will not grow.

AcSys Watchlist is designed to assist officials in the performance of background checks by allowing them to compare images captured for visa and license applications against databases of existing visa and license applications and watch list databases.

AcSys Watchlist includes the same set of utilities that have made the AcSys FRS line of products easy to use and administer:

- **FaceDatabase** – Allows you to administer target information
- **FaceTrainer** – Creates and updates templates from facial images and live video placed in a training queue

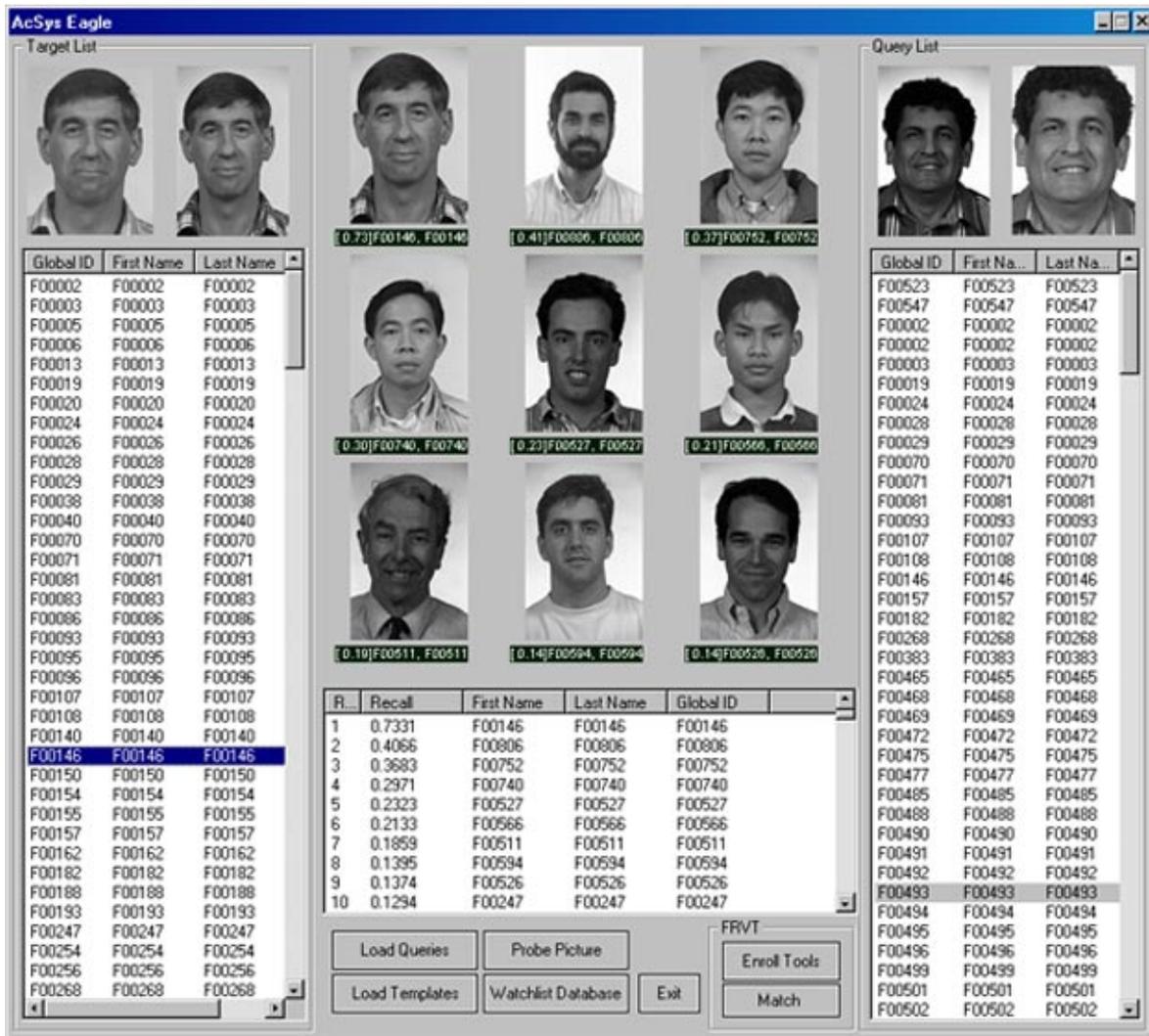


Figure 1: The AcSys Watchlist Main Interface

Component List

System includes a video capture board, even though it is not required for this test, as image capture is out of scope.

Hardware

- Intel Pentium III 1 GHz CPU x 2 installed
- ASUS CUV4X-D Motherboard
- 512 MB SDRAM @ 133 MHz
- Winnov PCI Videum Capture Board
- SoundBlaster 128 PCI
- ATI Rage XL Pro 2x AGP
- 3com 3C905C-TX NIC
- Maxtor 5T040H4 Hard Drive
- CD ROM
- 3.5 Floppy Drive

Software

- | | |
|---|--|
| Operating System: | <ul style="list-style-type: none">• Windows XP Pro |
| AcSys Face Recognition Software: | <ul style="list-style-type: none">• AcSys Watchlist• AcSys FaceTrainer• AcSys FaceDatabase |

Cost Breakdown

When the system is released to general availability, software costs will start at approximately US\$25,000, with peripheral modules (e.g., database, administrative tools, capture and control units, training stations, enrolment stations) to be priced separately.

Modifications for FRVT 2002

1. The **Target List** and **Query List** were renamed for clarity.
2. Two options were added to the application specifically for FRVT 2002:
 - **Enroll Tools:** This option allows you create biometric templates from images and add the templates to the Target List.
 - **Match:** This option runs the FRVT 2002 test process. Specifically, it performs a one-to-many comparison of each template in the Target List against each signature in the Query List and generates a similarity file for each signature.

The Systems Used by Cognitec in the FRVT 2002

Cognitec Systems GmbH
An der Flutrinne 12
D-01139 Dresden
Germany

1 Overview of the systems

Cognitec participated in the Face Recognition Vendor Test 2002 with two different systems, one for the High Computational Intensity (HCInt) test, the other for the Medium Computational Intensity (MCInt) test. The systems differed only in the computer hardware: a PC with a fast processor and 1 GB of main memory for the HCInt test, a notebook computer for the MCInt test. The software was the same in both systems and mainly consisted of two command line tools running under Windows 2000 and written specially for the Vendor Test using Cognitec's FaceVACS® Software Development Kit (FaceVACS® SDK).

1.1 FaceVACS® SDK

The FaceVACS® SDK provides functions for using all parts of Cognitec's face recognition engine such as face finding, eyes finding, creation of a biometric template from one or more images (enrollment), 1:1 comparison (verification) and 1:n comparison (identification) of an image against templates. The SDK thus gives integrators the opportunity to use high quality face recognition technology in a broad range of applications.

The main features of FaceVACS® SDK are:

- State of the art face recognition API for easy integration of biometric technology into a broad range of applications
- Support for biometric operations: enrollment, verification and identification from image streams (live video) and image collection (image databases)
- C++ interface to support modern paradigms of software development:
 - Object oriented design
 - Extensibility and modularity
- Abstract interfaces support easy adaption to various video streaming hardware and databases
- Detailed documentation including API reference, user guide containing installation and redistribution manuals
- Fully documented examples illustrating the main use cases and variation points

FaceVACS® SDK is available for Linux (glibc 2.x) and various Windows platforms (95/98/Me/NT4.0/2000/XP)¹.

1.2 Face Recognition Technology

The FaceVACS® face recognition engine processes each image (8 bit greyscale) as follows (see also the two block diagrams below):

- **Face localization:** The image is analyzed to determine the position and size of one or more faces. (In all of the following steps, we assume that only a single face is found.)
- **Eye localization:** The positions of the eye centers in the face are determined.
- **Image Quality Check:** The quality of the face image is checked to see whether it is sufficient for the steps that follow.
- **Normalization:** The face is extracted from the image and is scaled and rotated such that an image of fixed size results, with the eye centers at fixed positions in that image.
- **Preprocessing:** The normalized image is preprocessed with standard techniques such as histogram equalization, intensity normalization, and others.
- **Feature extraction:** In the preprocessed image, features are extracted that are relevant for distinguishing one person from another.
- **Construction of the reference set:** During enrollment the facial features of (usually) several images of a person are extracted and combined into a reference set, also called the "biometric template".
- **Comparison:** For verification, the set of extracted features is compared with the reference set of the person who the person in the image just processed is claimed to be; for identification, the feature set is compared to all stored reference sets, and the person with the largest comparison value is selected; in both cases the recognition is considered successful if the (largest) comparison value - which is interpreted as a similarity value - exceeds a certain threshold value.

¹ Windows 95 is only supported as a target platform for applications developed with the SDK, not as a platform for the development itself.

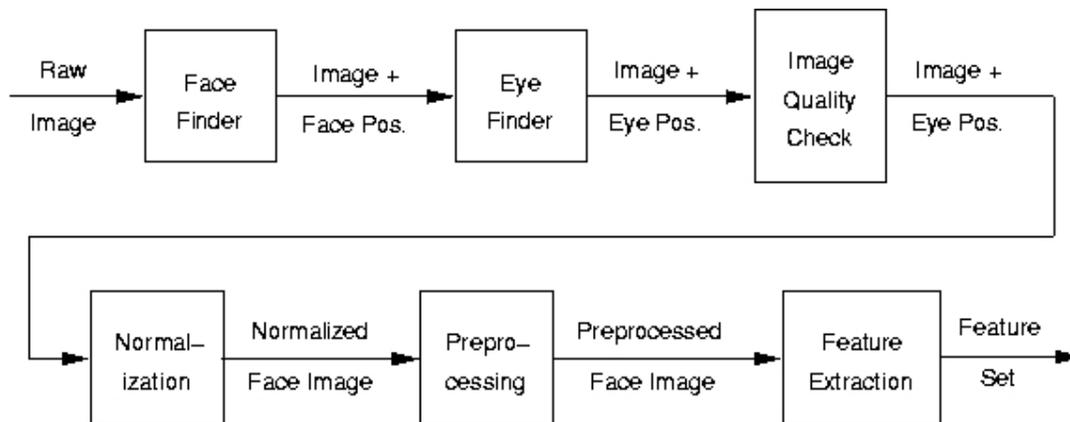


Figure 1: FaceVACS architecture: Feature set creation

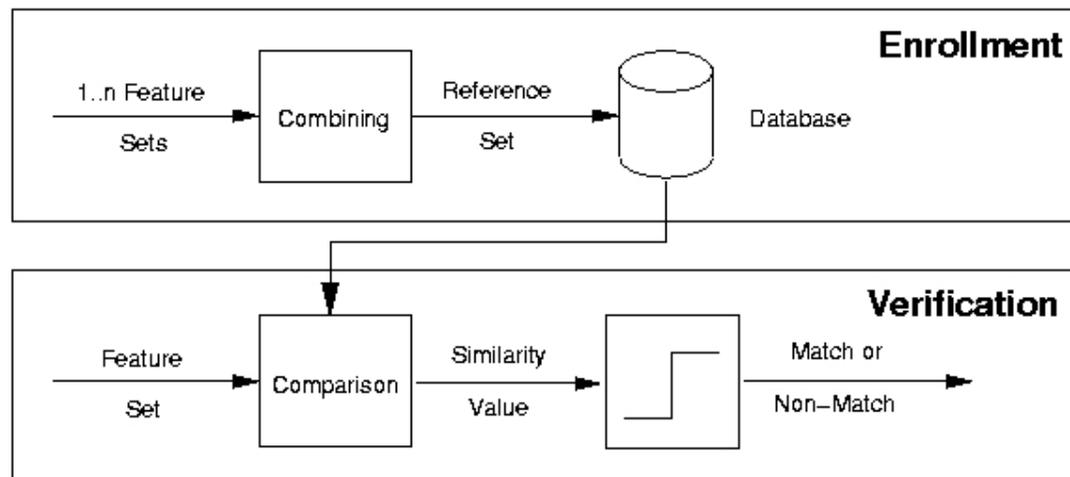


Figure 2: FaceVACS architecture: enrollment and verification

In addition, FaceVACS has a "live check" to ensure that the face in front of the camera is a real one and not just a photograph. To this end, the changes in appearance occurring during movement of the face (rotations around the vertical axis in particular) are exploited. Due to the special 3D structure of a real face, those changes are quite different for a real face than for a photo. So when the user wants to pass the live check, he or she should briefly rotate his or her head back and forth.

For more details see the technology description on Cognitec's website (<http://www.cognitec-systems.com>).

2 Component list and cost breakdown

The submitted systems had in common the software, consisting mainly of two simple command line tools (about 500 lines of C++ code) that can be quickly written on the basis of the FaceVACS® SDK. The FaceVACS® SDK is currently priced at \$ 3,800².

2.1 System for HCInt test

The following table shows the components and the cost breakdown of the PC used for the HCInt test (all prices are average mail order prices in Germany without tax as of July 18th 2002):

Item	Vendor	Price
Case Power Supply 300W	SilentMax Fortron	\$ 175 (case + power supply)
Mainboard MSI K7N 420Pro	MSI	\$ 120
CPU Athlon XP1700+	AMD	\$ 95
CPU Cooler	AMD	\$ 25
RAM 2x512MB PC266 CL2 DDR	Infineon	\$ 310
HDD 80 GB IBM IC35L080VA	IBM	\$ 86
DVD/CDROM		\$ 65
USB 2.0 Controller		\$ 35
		\$ 911

The total system price (SDK + PC) thus amounts to **\$ 4,711**.

2.2 System for MCInt test

The notebook computer used for the MCInt test is a Fujitsu-Siemens Lifebook C 6555 equipped with the following hardware components:

- CPU Pentium III 650 MHz
- HDD 10 GB
- SDRAM 128 MB 100 MHz
- FDD

² In addition, if an SDK customer plans to sell applications developed with the SDK, runtime royalties have to be paid to Cognitec for each sold copy of the application software; those royalties are fixed individually in a contract between the customer and Cognitec.

- CDROM drive
- 4 MB Video RAM

In February 2001, the price for this notebook was DM 4,985 (German Marks) without VAT. Nowadays (July 2002), an equivalent or even somewhat more powerful notebook should be available for less than \$ 1,800, making the total system price (SDK + Notebook) be less than **\$ 5,600**.

3 Modifications for the purpose of FRVT 2002

The only modification, that was done for efficiency reasons, was the extension of the FaceVACS® SDK interface to allow the comparison of two biometric templates (in the current release, FaceVACS® SDK 1.2, only comparison between a template and a set of images is possible). This allowed to exploit the fact that in the HCInt test, query and target set were identical; so, the whole processing chain consisting of, among other steps, face finding, eyes finding, and feature extraction had to be executed only once for each image. With the interface of the current release, each image would have had to be processed twice, once as target image and once as query image. The comparison results, however, would have been exactly the same.

FaceSnap RECORDER – General Overview

The FaceSnap RECORDER digitizes standard video sources and performs real-time pattern recognition on the captured images. The FaceSnap RECORDER detects human faces in the video stream and stores a user-adjustable image region around each face. A powerful image database facilitates easy navigation and search on large numbers of stored images. For analyzing the face image database, interactive and automatic search functions are provided. As an option, the FaceSnap RECORDER can perform real-time facial identification. The FaceSnap RECORDER is a novel solution for person-oriented video surveillance, access monitoring and other applications requiring face image recording and face recognition.

The FaceSnap RECORDER targets the market for intelligent video surveillance. It captures and stores face images in near video real-time. That is, any standard video source (cameras, VCRs ...) can be analyzed for one or more persons of whom the face is visible in the image. The face images are stored in a real-time data base. The user can select images as reference faces and then search all the recorded data for images that may match the reference image(s).

The most important applications for the FaceSnap RECORDER are in law enforcement (e.g. police observations), in casinos, airports, border control, and banks.

The FaceSnap RECORDER is a complete ready-to-use system comprising a powerful PC platform and a factory installed software. It is currently shipped in three versions: Standard, "law enforcement" (LEF), and LEF Portable.

The FaceSnap RECORDER also is designed for operating in a network. That is, several FaceSnap RECORDER may be accessed from a central server and other systems can easily be interfaced to a FaceSnap RECORDER using standard internet protocols.

A powerful real-time facial identification system can be setup by connecting a FaceCheck SERVER to a FaceSnap RECORDER over a network link.

There are several modes of operation:

- Standard recording mode
The FaceSnap RECORDER screen shows the live camera image and a number of most recently recorded persons. The user can switch between a number of different video inputs. The captured face images are displayed in a normalized format and can be stored in real-time on the local disk.
- Replay mode
In replay mode, the user can easily navigate through a large amount of face image data. Images are time stamped and can be directly accessed through their time keys or the graphical event histogram that provides an overview of the temporal distribution of recorded images.

- Search mode
In search mode, the user can define image groups as a general purpose tool for data management. In addition, image groups are used to define training sets for face identification. Automatic functions for image pattern recognition are provided to search the face image database and to support content-based data management.
- Training mode
The FaceSnap RECORDER can be trained to recognize faces based on one or more sample images collected by the user. In training mode, an image group can be imported to establish or extend a training set for a particular person. There are two hyper classes of persons: a) the class of all persons known to the FaceSnap RECORDER and b) the class of all persons the FaceSnap RECORDER is supposed to look for.
- Selective recording mode (FaceCheck SERVER)
Face images captured by a FaceSnap RECORDER operating in recording mode can be sent to a second FaceSnap RECORDER over a high-speed 100BaseTX network connection in real-time. The second system, which has to be put in the selective recording mode, receives the stream of face images and tries to identify all persons belonging to the hyper class of "active/interesting" people (watch list). Only those images for which the similarity value exceeds a user-defined threshold are displayed and optionally stored in the database.

Technology

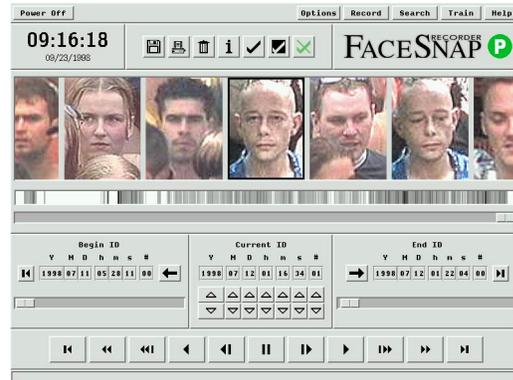
The most important technology built into the FaceSnap RECORDER is the automatic selection of human faces in an image. The underlying technical problem is to detect and localize face-like patterns in a digitized image of an arbitrary scene. Moreover, the computer has to complete this search within fractions of a second since the person may not look straight toward the camera much longer than that. An additional requirement FaceSnap was designed to meet is the independence of object cues like motion or color. FaceSnap looks for face patterns in an image, not for regions showing a certain color range or being in motion relative to the image background. FaceSnap solves these problems by using a fuzzy abstract face model in combination with a neural net classifier. The outcome of the face finding stage are face image windows which are ranked according to a faceness value indicating how much the selected image portion resembles a human face

The face detection (face spotting) of the FaceSnap RECORDER uses the FaceSnap technology as described above. The model-based pattern recognition has a processing rate of up to 25 images per second, depending on CPU type and size of the source image. Frontal views of human faces are preferred with a rotational tolerance of ca. +/- 15 degrees. Tolerance to deviations of the frontal face pose can be adjusted.

The face identification of the FaceSnap RECORDER uses the FaceCheck technology. The pattern recognition for face identification is based on user-provided training sets. The facial feature extraction is designed for frontal face views. Compensation of image rotation works up to +/- 15 degrees. Based on the location of prominent facial features, the face images are first geometrically normalized and then normalized with respect to the gray level

distribution. From the normalized images, feature vectors are produced and then used for classification. The classification algorithm is based on a combination of the elastic net matching technique and specialized neural net classifiers.

The face recognition technology used for the product FaceSnap RECORDER has been developed by C-VIS and partly by the University of Bochum. All intellectual property rights are owned by C-VIS.



Screenshots of standard recording mode (left) and replay mode (right).



Screenshots of search mode (left) and training mode (right)



Screenshot of the selective recording mode (FaceCheck SERVER)

Systems submitted to the FRVT2002

The facial recognition vendor test 2002 does not include any test with real-time video input. Therefore the FaceCheck SERVER option of the FaceSnap RECORDER was not part of the systems that were submitted by C-VIS.

The systems submitted by C-VIS for the FRVT2002 are all of the product model “FaceSnap RECORDER law enforcement”. C-VIS used a total of three systems for the test.

The component lists of the submitted systems are given below. The table also states the price of the products.

Product Model	System Price [Euro]	Enclosure	Processors / RAM	Other Components	Test used for
FaceSnap RECORDER LEF portable	10450	ATX-Portable with 12,1” TFT	Dual Pentium III, 1.4 GHz, 512 MB	CDROM, LS120, 60 GB HDD, PCI video frame grabber, 10/100 Ethernet Adapter	HCI
FaceSnap RECORDER LEF	8800	ATX-Desktop	Dual Pentium III, 1.1 GHz, 512 MB	80 GB HDD, 10/100 Ethernet Adapter	HCI
FaceSnap RECORDER LEF	8800	ATX-Mini-Tower	Single AMD Athlon 1900+, 256 MB	CDROM, FDD, 80 GB HDD, PCI video frame grabber, 10/100 Ethernet Adapter	MCI

The standard products uses for the FRVT2002 had to be modified to meet the special requirements of the Test Plan and the “FRVT 2002 Implementer’s Guide”.

The significant modifications are:

- User interface option for choosing “FRVT2002 operating modes”.
- Special “Import Mode” that uses XML documents for controlling the image access.
- Special mode for generating scoring results in the FRVT2002 file format.
- Normalization of the scoring results with an external function.
- Special matching mode with reduced accuracy for the High Computational Intensity Test.



Summary Description Documents

Dream Mirh Co., Ltd.

Overview of System

The Dream Mirh FRVT2002 system consists of hardware and software. The hardware is a generic standalone style workstation. One workstation is used for the high computational intensity test and another for the medium computational intensity test. The software is used for all tests using the MIRH Eye Face Recognition Engine.

Hardware

The structure of FRVT 2002 appears on <Table 1>. <Table 1> compares the MIRH Eye Engine with the commercial package, MIRH Eye Surveillance system.

	FRVT 2002	Commercial (Surveillance)
Computer	HP x-4000 Pentium IV 2.2GHz (Single)	Pentium IV 2.0GHz (Single)
Main Memory (RAM)	4GB	512MB
OS	Windows 2000 Server	Windows 2000 Server
Input Condition	JPEG format image	Any CCD Camera or Photo

Table 1: Hardware Configuration

FRVT 2002 uses the widely used HP Workstation x-4000 model system. The breakdown in price is as follows:

HP x-4000 Workstation

Pentium IV 2.2GHz Dual CPU

4GB RDRAM Memory

Wildcat 6110 64M

Price: \$ 4,000

17" TFT LCD Monitor

Price: \$ 700



Windows 2000 Server

Price: \$ 1,000

For the windows series or Red Hat Linux operation systems and hardware, basic configurations can be expanded upon the user's request.

Software

The software used in FRVT 2002 was developed for users wishing to use face-recognition technology. Its functions are based on MIRH Eye SDK (Software Development Kit) 2.0. MIRH Eye SDK takes image input from a camera or any other image, finds the facial area, compares that with the data stored in the database, discriminates and verifies the data. This SDK can be used with the Windows series and Visual C++ as it has optimal codes for better compatibility program development, thus ensuring a faster and more powerful operation. The structure and hardware for the MIRH Eye SDK is as follows:

MIRH Eye SDK has the following 3 components.

Detection: Finds the face in the entire image

Verification: 1:1 method. An ID is entered with the facial image

Identification: 1:N method. Only the facial image is entered so that the system needs to search the database for the match.

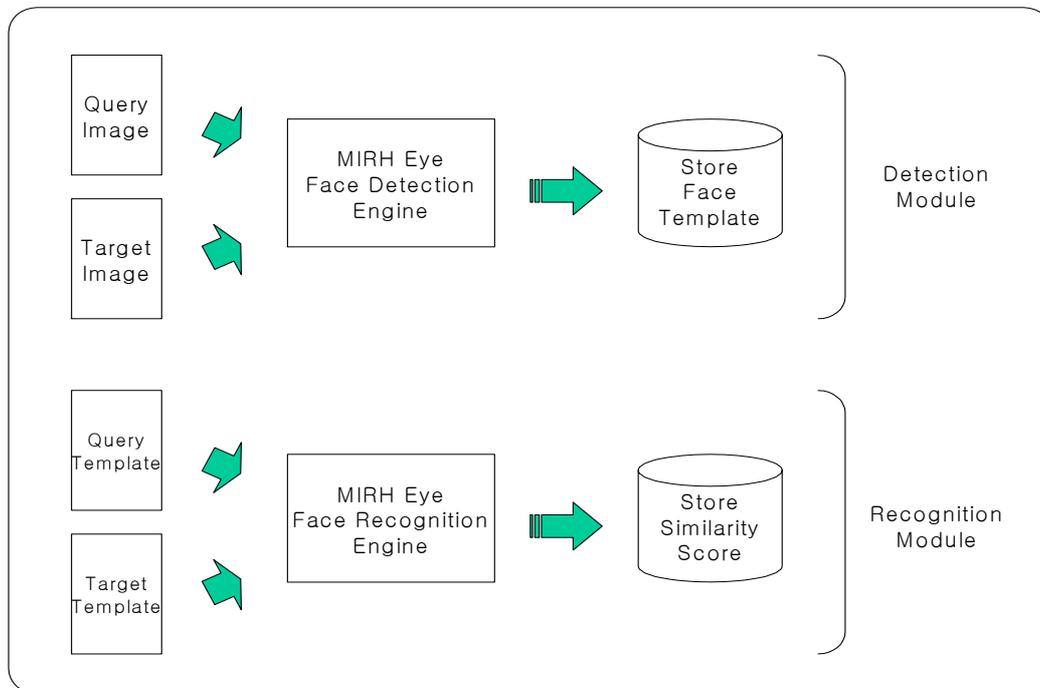
MIRH EYE SDK use the following hardware

PCs with Windows 9x, 2000, NT, XP operating systems

Intel Pentium III 500MHz or higher

64 RAM or higher

The software engine used for FRVT 2002 is the basic MIRH Eye SDK Engine that has been modified for the test environment. The structure and process for the FRVT 2002 is show in the flow chart. <Picture 1>



<Picture 1>: MIRH Eye Engine flow chart

The MIRH Eye Engine that was designed for the FRVT 2002 has two main processes Detection and Identification. In the detection process, the images on the query/target list are uploaded, the facial area is found, and a template is made for the images to be stored in xml form. After all the images are loaded in the database, the Identification process automatically begins. **The detector converts all images into template. Recognizer takes one template from Query, and stores the degree of resemblance based on the comparison with all templates in the target in the similarity file.** After all Query images are running the test ends.

For further information, please contact:

Donghoon Jeon

Dream Mirh Co., Ltd.

R & D Center

Senior Engineer

Email : smallove@dreammirh.com <http://www.dreammirh.com/>



Face Recognition System Description Eyematic Interfaces Inc.

Face Recognition Vendor Test, July 2002

Eyematic's visual sensing technology is made available through a Software Development Toolkit (SDK). The toolkit bases on Eyematic's visual sensing software, called EPL (Eyematic's Primary Library). Specific capabilities are Face detection in images, Head or Local Feature tracking in sequences, Facial Feature Analysis and Human Identification through Face Recognition. The SDK's core component is a C++ based API consisting of image processing modules.

The API was developed to allow easy Integration of Eyematic's technology in applications such as Access Control Systems, Credit Card Authorization Terminals, Point of Sales Systems, Baggage/ Passenger Matching and surveillance Systems.

OMRON Corporation (Japan) is Eyematic's largest partner integrating face recognition applications based on the SDK.

1 EPL-API SDK Overview

1.1 Components

The API architecture was designed with emphasis on modularity, adaptability and convenience for the application developer. It provides a generic system of data processing components (modules) that can be linked together to a pipeline. Each module is optimized for a specific image-processing task. Sub sequential modules process information provided by preceding modules. The data set being passed through consists of an image and image related information. Still images as well as video stream images flow through all processing stages and are gradually enriched by face related information (face position, head pose, bounding rectangle, local feature positions, etc) as the data passes each module (Fig. 1,3). Processing modules can be individually combined to serve specific requirements such as fast processing on video streams or extensive search on still images. The pipeline concept allows optimizing CPU usage on multi processor platforms as well as using multiple computers by running each module in a separate thread or process.

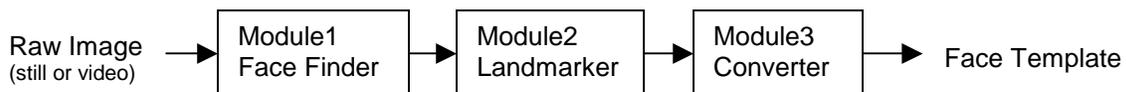


Figure 1: Typical set of SDK Modules used for face recognition.

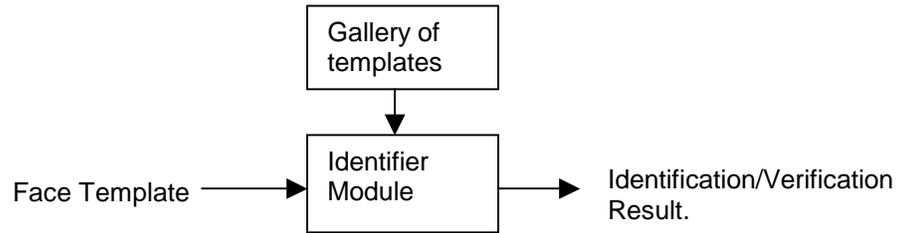


Figure 2: Typical set of SDK Modules used for face recognition.

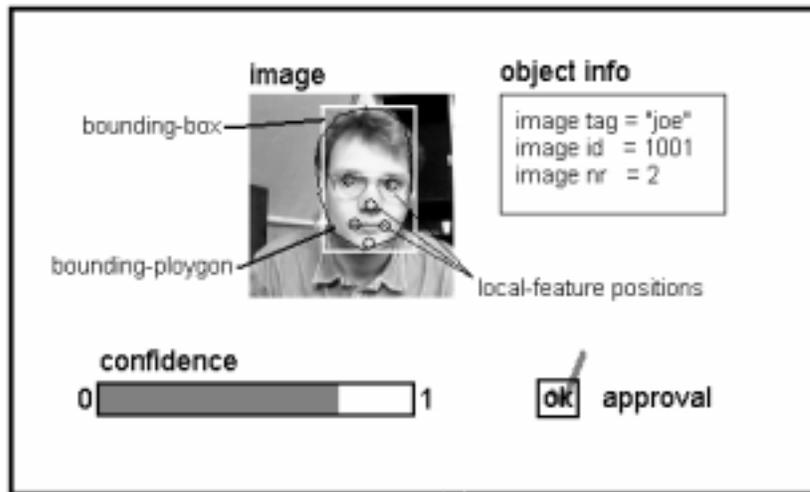


Figure 3: Data set being passed through modules.

In order to build a face recognition system, the system integrator would combine modules of 4 categories:

Module Category	Description
Face Finder	Locates the face on the image
Landmarker	Finds local feature positions
Converter	Extracts a face template from the image
Identifier	Recognizes/Verifies a person by comparing a face template to a gallery of templates.

A combination of Face Finder, Landmarker and Converter specifies a face enrollment system (Fig. 1). The face template obtained would either be stored in a gallery or identified/verified by matching it to entries of a gallery (Fig. 2). Face Finder and Landmarker modules produce a confidence value that provides a measure of quality for face detection. This is particularly important when processing video stream data. Here this measure is used as filter to select images that do display a face in sufficient quality.

1.2 Underlying Technology

The face recognition system uses Gabor Wavelet Image Transformation and Neural Networks. Face Detection Modules employ a general face model to localize a face in the image. A 3D representation of the head copes with pose variations. The general face model is learned from a large database of sample images. The face models supplied with our standard modules cover a wide variety of environmental conditions (illumination, poses, expressions, etc). Sometimes an adaptation to specific conditions is desired. For this purpose a special training tool can be provided.

1.3 Technical Specifications

For each category we have developed a variety of modules, which represent different accuracy and speed levels. The processing speed per image typically varies between 200 ms and 1000 ms depending on configuration used¹. Modules of the same category can be exchanged without the need to adapt source code. The size of a face template can range from 1 Kbytes to 6 Kbytes depending on the type of Converter used. The typical Identification/Verification speed is about 90 μ s (microseconds) for a single template-to-template match² based on a template size of 4 Kbytes. Usage of lower template sizes will yield significantly higher processing speeds at reduced accuracy.

1.4 Target Platforms

The EPL software has been developed with high emphasis on portability. The underling code was written in strict accordance with standards that ensure portability to multiple platforms. Time critical routines are optimized in using platform specific resources to ensure maximum performance. The API itself complies with the ANSI C, C++ Standard. For high performance and desktop applications conventional desktop PCs can be used. For low cost and low size embedded systems, we have developed an embedded EPL-API based on Texas Instruments Digital Signal Processors.

Full functionality is available on MS Windows based Platforms using Intel Pentium (or similar) Processors.

Operating System: Windows 95, 98, ME, 2000, XP

Recommended Development Environment: MS Visual C++ 6.0

Slightly reduced functionality is available for Linux/Unix based Platforms.

Recommended Compiler: g++ or compatible

A special embedded version of the SDK is available for the DSP Families TMS320C6200, C6400, and C6700.

Recommended Development Environment: TI Code Composer Studio

1.5 System Recommendation

The system requirements depend on desired processing speed and gallery size. We recommend using a MS Windows Intel Platform:

CPU: Pentium 3 at 800 MHz or better

RAM: 256 Mbytes or higher

¹ Speed value refers to Face Finder, Landmarker and Converter pipeline on a Single CPU Platform: Intel P4, 2GHz (or similar).

² Measured on single CPU Platform: Intel P4, 2GHz (or similar).

2 FRVT2002 Test Application

The FRVT2002 Application uses the EPL Module concept described above for still image processing. We selected a template size of about 4 Kbytes/image for both subtests (HCINT and MCINT). We bypassed the Identifier module and compared face templates directly to avoid some unneeded processing, which the Identifier Module would normally do automatically (sorting of similarities, normalization, etc). For video data we used the filtering method mentioned in the paragraph above.

For both subtests we chose face detection module configurations that take between 0.5s and 0.8s processing time for each image on a single computer. (S. Paragraph 3 for Hardware Specs.)

Part of training data of face detection module used for FRVT2002 was taken from the *FERET* database of facial images collected under the *FERET* program.

Ref.: P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The *FERET* database and evaluation procedure for face recognition algorithms," *Image and Vision Computing J*, Vol. 16, No. 5, pp 295-306, 1998.

3 FRVT2002 Component list and cost breakdown

3.1 Computers

Three desktop computers were prepared to run the test. Two computers were dedicated to HCINT. One computer was prepared for MCINT and HCINT. The computers used for HCINT were connected via 100MB/s Ethernet Network during the time HCINT was running. We used Windows 2000 as operating system on all machines.

Computer Specification (Excl. Monitor)	Quantity	Unit Price (Excl. Tax)	Total Price
DELL Dimension 8200 (Single Processor) Intel Pentium 4 Processor at 2.53 GHz 1024 MB PC800 RDRAM 80 GB 7200 RPM Ultra ATA Hard Drive	2	\$2,456.00	\$4,912.00
Compaq EVO Intel Pentium 4 Processor at 2 GHz Southland Compaq 1024 MB RAM 40 GB Hard Drive	1	\$1,482.98	\$1,482.98

3.2 Accessories

The following accessory items were prepared:

Accessory Specification / Usage	Quantity	Unit Price (Excl. Tax)	Total Price
Netgear Ethernet Hub 100MB/s <i>To connect computers</i>	1	\$169.95	\$169.95
Belkin KVM Switch/Cable Omnvm <i>To switch single monitor, keyboard, mouse between computers.</i>	1	\$226.32	\$226.32
LACIE 80GB USB 2.0 Hard Drive <i>Extra Data Storage Device (if needed)</i>	1	\$203.68	\$203.68

3.3 Face Recognition Software

Our Business Model for the API SDK is to license it to System Integrators. The price consists of an upfront license fee and royalties depending on (but not limited to) field of view, geographical region and exclusivity.

We offer evaluation licenses to customers who might be interested in testing our system. The fee for an evaluation license of the API SDK Software and Modules to replicate the FRVT2002 test setup would be approximately \$10,000. The actual amount can fluctuate depending on market conditions, special requests, etc.

We gladly refer inquiries about applications using Eyematic's visual sensing technology to OMRON Corporation (Japan). URL: <http://www.society.omron.com/faceid/index.html>

Contact:

Eyematic Interfaces Inc.
9800 S. La Cienega Blvd., 5th Fl.
Inglewood, CA 90301
USA

www.eyematic.com

Iconquest
3525 Piedmont Road
Building 7, Suite 600
Atlanta, GA 30305

System Description
for
FRVT 2002
Face Recognition Vendor Test
July 24, 2002

Iconquest's submission to FRVT 2002 consists of proprietary software, IQ ImageMatcherTM, hosted on a Dell Precision 340 desktop computer.

Description of IQ ImageMatcherTM Software

IQ ImageMatcherTM is software for personal computers that combines a fractal image comparison algorithm with both a graphical user interface and a script driven interface.

The image comparison algorithm returns a number $M(Q,T)$ for each pair of images Q and T .

The methodology is to partition Q into relatively small pixel blocks and to determine whether each such block can be better represented as a contractive transformation of a somewhat larger block chosen from Q or chosen from T . The transformations are chosen from among compositions of spatial rotation and spatio-spectral scaling and translation. In the simplest case, $M(Q,T)$ can be chosen as the ratio of number of partition blocks better represented from T to total number of partition blocks. Typically though, we scale up T somewhat and normalize to insure that this process results in $M(Q,T) = 1$ when $T = Q$.

This approach is an example of unconstrained image comparison where no assumption of structure is imposed on the underlying imagery. In contrast are constrained image comparison applications such as optical character recognition, fingerprint identification, biometric face recognition, robotic vision for quality control and automatic target recognition.

One expects unconstrained systems to be applied at least in cases where the imagery is of too poor quality for structure differentiation necessary in a constrained system, where there is not sufficient imagery available for development of constrained systems or when there is not sufficient time for development of a constrained system in response to some immediate application need.

The IQ ImageMatcherTM approach has several advantages in unconstrained image matching. First, $M(Q,T)$ exhibits continuity relative to spatial and spectral

Phone : 404-264-8060

Email: asloan@mediabin.com

Fax: 404-264-8313

transformations built from rotation, scaling and translation, i.e., if S is such a transformation that doesn't rotate, scale or translate much, then $M(Q,S(T))$ will tend to be close to $M(Q,T)$. In addition, if a portion of Q is similar to a portion of T, then $M(Q,T)$ can be relatively large in comparison to $M(Q',T)$, when Q' is unlike T.

M can be extended to sets of imagery in a number of useful ways. For example, if $T = \{T_1, T_2, \dots, T_N\}$ where each T_i is an similar view of one object, define

$$(1) \text{ ANY}(Q,T) = \text{maximum}\{M(Q,T_i): I=1, 2, \dots, N\}$$

$$(2) \text{ ALL}(Q,T) = \text{minimum}\{M(Q,T_i): I=1, 2, \dots, N\}$$

Increasing the number of images in T representing very different views of an object increases the likelihood that ANY(Q,T) will be high when Q is any view of the object.

Increasing the number of images in T representing very similar views of an object decreases the likelihood that ALL(Q,T) will be high when Q is not a similar view of the same object.

Using $\text{WHOLE}(Q,T) = \text{minimum}\{M(Q,T), M(T,Q)\}$ instead of $M(Q,T)$ makes it harder to get a high score when Q contains only a portion of the object completely represented in an image T.

The IQ ImageMatcherTM software implements these and other related functions.

One application of IQ ImageMatcherTM is to pre-screening. In a typical pre-screening application, there is a Watch List of important objects and classes of objects, whose presence or absence in Query imagery is important, in varying degrees, to determine. Pre-screening is used when the amount of unknown Query imagery exceeds the capacity of available image analysts or security personnel to review. For example, if analysts can only review 10% of available imagery in a given time frame, then how should those 10% be selected? Pre-screening selects the imagery most likely to be most important. The continuity mentioned above can be used to direct to collection of a database of Target imagery and the construction of generalizations of M similar to the examples given above to define a procedure which selects imagery more likely to be representative of important objects and classes of objects.

In applications where a database of Target imagery is available prior to use of software, the database can be analyzed and a more optimal measure can be computed by using maximum and minimum values of M across various subsets of images in the database representing similar instances of a person and different instances of a person.

As a simulation of matching with less than ideal imagery, a FERET thumbnail database of 7249 images was segmented for faces and each segmented face was scaled to 96 pixels in area, while maintaining the aspect ratio of the segmented image. The first image of

each of the 695 individuals was used to create 695 Query Sets, each consisting of 1 image. The remaining 6554 images constituted a fixed Target set. ImageMatcher™ was run 695 times, once for each Query set and the same Target set. We set Success(N) to be the fraction of the 695 Query images for which the set {those Target images having the N highest scores} contained at least one Target image representing the Query image individual and found

<u>Image Size</u>	<u>Success(1)</u>	<u>Success(20)</u>	<u>Success(50)</u>	<u>Success(100)</u>
“8x12”	.65	.87	.92	.95

Modifications for the Face Recognition Test

The ability to handle image signatures in an XML format was added to comply with Test requirements.

An image segmentation pre-processing module was added to search for the largest object and enclose in a bounding rectangle since the FRVT 2002 test protocol guarantees that the face is the largest object in each image.

IQ Imagematcher™ has a parameter which determines the output size of an image scaling pre-preprocessing module. This output size can dramatically affect the computation time. We used the each of 256 images in the FERET_query data set to create 256 Query sets and all the 256 images in the FERET_target data set to create a Target set. The FERET_query and FERET_target data sets were supplied in the June HEF supplement. We ran IQ Imagematcher™ to test the effect of the rescaling the original imagery by segmenting each image for a face and rescaling the segmented images to an area equivalent to the following Image Sizes while attempting to maintain the aspect ratio and found

<u>Image Size</u>	<u>Success(1)</u>	<u>Success(20)</u>	<u>Success(50)</u>
6 x 9	.25	.56	.71
8 x 12	.36	.70	.78
12 x 18	.41	.69	.78
16 x 24	.50	.75	.82
24 x 36	.59	.77	.85
32 x 48	.63	.77	.86

Here Success(N) is the fraction of the 256 Query images for which the system selected the correct Target image among the set of 256 Target images having the N highest scores.

Typically, performance would improve with the use of higher resolution images and the improvement changes are smoother as the resolution increases. The 1536 pixel image area, corresponding to 32 x 48, was selected to insure that the Test be completed in the specified time limit. The scaling module was invoked to resize each segmented Target

and Query image to have 1536 pixels, while maintaining the aspect ratio of the segmented imagery.

The higher resolution original image data provided in the Test was only used to convert to greyscale and to segment. The image matching was done only with the “32 x 48” imagery.

In order to reduce computation, rotations were not included in the collection of contractive transformations considered by the software.

Color information was only used to create grayscale imagery and in the segmentation step.

In cases where the signature was a still set or a video only 2 image representatives were used. This number was chosen to control the computation load, so that the Test could be completed within the constraints provided. The video was turned into a still set by selecting every tenth image. The image representatives for each still set were selected to maximize the value of $M(,)$ when applied to different images in the still set.

With $T = \{T_1, T_2, \dots, T_N\}$ and $Q = \{Q_1, Q_2, \dots, Q_P\}$ being the selected images, the reported similarity scores are computed as:

Maximum $\{M(Q_i, T_j)\}$ for $i = 1, 2, \dots, P$ and $j = 1, 2, \dots, N\}$
and for the test N and M were either 1 or 2.

System Description

Hardware

Dell Precision 340 Desktop
2.4 GHz/533 MHz Front Side Bus, P4 Processor, 512 Full Speed Cache
1 GB PC800 @ 400MHz ECC, RAMBUS memory
nVidia Quadro 2 EX, 32 MB Graphics Card
Two 36 GB SCSI U160/M Hard Drives
SCSI Ultra 160/M Controller
3.5” 1.44 Floppy Drive
1394 Controller Card
16x DVD and 24xCDRW
Windows 2000 Pro
Cost \$4,313

Software

IQ ImageMatcher – development version, not separately priced.



Identix Corporation
One Exchange Place, Suite 800
Jersey City, New Jersey 07302
Phone: 201-332-9213
Fax: 201-332-9313
<http://www.identix.com>
E-mail: info@identix.com

1.0 Overview of Submitted Systems

Facelt[®] Technology used in Facial Recognition Vendor Test 2002

All Identix Corporation face recognition products and solutions are derived from the algorithms in the Facelt[®] Identification Software Developers Kit (SDK).

THE FACEIT[®] ENGINE HAS A NUMBER OF CAPABILITIES

Face Detection Detects faces in static images or video streams, even in complex scenes or backgrounds.

Segmentation Crops, rotates, and re-scales faces out of an image.

Image Quality Evaluates the quality of a captured facial image for recognition.

Faceprint Creation Generates templates as small as 84 bytes.

Face Recognition One-to-one (verification) or one-to-many (identification).

These capabilities are packaged as

1. C++ classes and COM objects, included in the Software Developers Kits (SDKs), and Facelt[®] add-on software modules.
2. ABIS (Automated Biometric Identification System), a scalable one-to-many search engine and application framework.
3. Argus, a scalable real-time video watchlist system with capabilities to integrate into video surveillance infrastructures.
4. IBM/Informix Datablade.
5. Third-party applications such as the IBM screen saver, the Sony/Keyware screensaver, etc.



FACIAL RECOGNITION TECHNOLOGY BENEFITS

- Accurate
- Passive
- Convenient
- Human Readable Audit Trail
- Non-Invasive
- Uses Existing Photographs and Databases

APPLICATIONS

Authentication

Computer and Network Security, Token Replacement
Access Control, Banking, Border Control

Criminal Justice Systems

Mugshot and Booking Systems

Human ID at a Distance

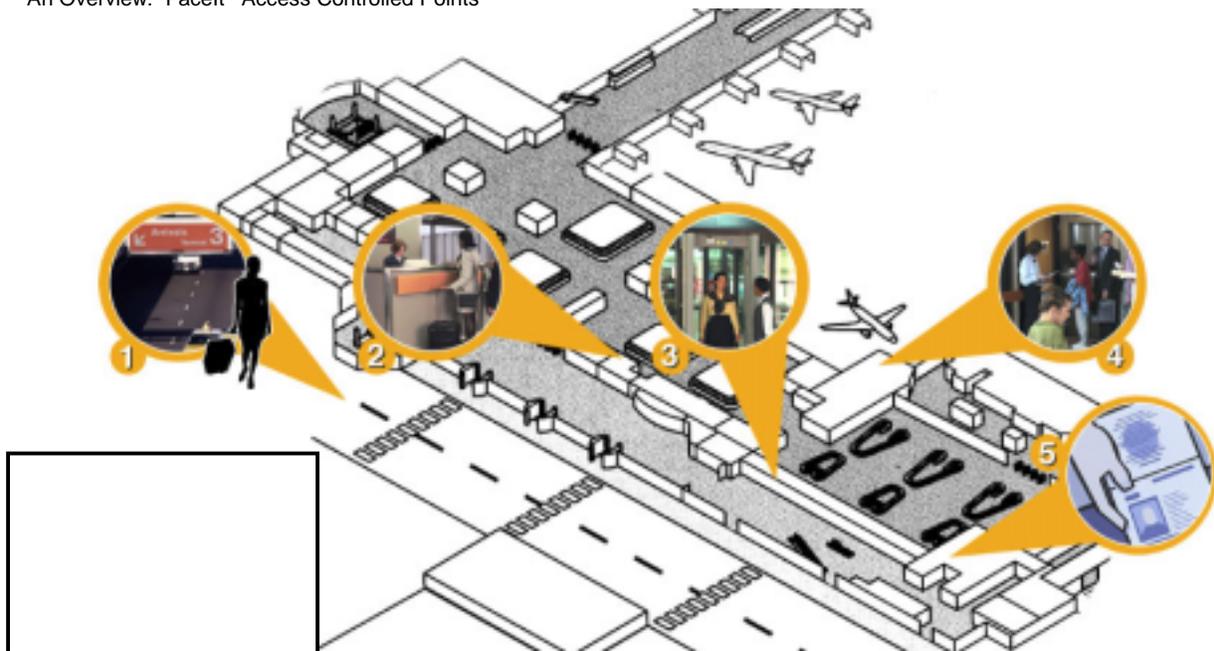
Surveillance, Smart CCTV, Human Traffic
Control, Image Indexing and
Post-Event Analysis

Identity Fraud

Voter Registration, National IDs, Passports,
Drivers' Licenses, Employee IDs

Facelt® Face Recognition Technology Airport Application

An Overview: Facelt® Access Controlled Points





WORKING WITH FACEIT[®]

Technology and Pricing

Facelt[®] technology is initially provided in the form of software developers kits and modules. Once the solution is completed, a license agreement is then structured in accordance with the application or market segment. *Evaluation opportunities are available to pre-qualified entities.*

What is Facelt[®]?

An award-winning enabling technology for detecting and recognizing human faces.

Facelt[®] is used by Identix's business partners to build a variety of finished products and solutions such as mugshot and booking systems, self-service financial kiosks, identity fraud database systems, access control solutions, network security applications and many others.

How It Works

The Core of the Facelt[®] Engine: The Local Feature Analysis (LFA) Algorithm

Fundamental to any facial recognition system is the way in which faces are coded. Initial studies into this field resulted in a global approach, known as *eigenfaces*. This theory suggests that faces can be synthesized from a set of master global facial templates, called *eigenfaces*. Combining co-efficients from the various *eigenfaces* yields a unique identity for every given face, similar to the composition of colors from the primary colors, red, blue and yellow. However, due to the fact that the *eigenfaces* are global in nature, they do not compensate for real-world variability with respect to lighting, pose, expression, facial hair, etc. Moreover, different sets of global templates apply for different ethnic groups and races, making the results significantly reliant on knowing which set of templates to apply and highly dependent on color and shading.

On the other hand, the approach used by Identix, called *Local Feature Analysis (LFA)* is based on local features that make up part of the face that differentiate between individuals. These features combine the texture information of a face image with information of the underlying anatomical bone structure. Identity is determined by combining information of these features in a statistically meaningful way.



Because its makeup is based on local features, the LFA *faceprint* is very resistant to changes in lighting, skin tone, eyeglasses, aging, facial expression, hairstyle, pose, etc.

There are roughly 80 degrees of freedom on the face that differentiate between humans. These are encoded in both the Identix “vector” *faceprint*, which is 84 bytes in size, and the Identix “Intensive” *faceprint*, with is 3.5K in size.

Application and Installations

CCTV Town Centers -- Taking a high-tech approach to fighting crime, London’s borough of Newham has tied over 300 cameras into a central control room where the Facelt[®] system identifies known criminals.

Information Security -- IBM provides Facelt[®] screensaver with all shipments of its snap-on Ultraport camera, designed for the A, T and X series Thinkpad laptops.

Eliminating Voter Fraud -- Facelt[®] is an integral component of the Mexican Federal Election Institute’s on-going program to eliminate duplicate voter registrations.

FACEIT[®] AT WORK: SELECT PARTNERS

See how other partners are implementing Facelt[®] technology at:
<http://www.identix.com/partners>

Our professional services group is also available to assist in developing enhanced Facelt[®] software modules to meet specific project requirements.

2.0 Component Lists and Cost Breakdown for Systems

Medium Computational Intensity Test

The system configuration used for the Medium Computational Intensity Test (MCINT) consisted of:

Component	Quantity	Cost (U.S. Dollars)
Dell Precision 530 Workstation Dual 2.4GHz	1	4258
Dell Model M782 Monitor	1	229
Identix Facelt [®]	1	10,000



Identification SDK		
Total Cost		14,487

High Computational Intensity Test

The system configuration used for the High Computational Intensity Test (HCINT) consisted of:

Component	Quantity	Cost (U.S. Dollars)
IBM xSeries 360 Server Quad 1.6GHz	3	111,870
APC Smart-UPS 2200VA	1	916
USB 80GB Hard drive	1	200
Ethernet 10/100 Switch	2	84
Cables for USB Drive	2	50
USB PCI Card	1	25
Dell P793 Monitor	1	259
KVM Switch 4-way	1	75
Identix Facelt [®] Identification SDK	1	10,000
Total Cost		123,479

Note that we do not recommend using Quad servers because of the cost. The same system performance can be achieved using 12 inexpensive rack-mounted single CPU machines at \$1,200 each. This reduces the cost to roughly \$27,000.

3.0 Modifications Specifically for the FRVT 2002 Test

A test application was written to perform both the medium and high intensity tests. This test application invoked the Facelt[®] Identification SDK. The test application included the following functionality specifically for the FRVT 2002 test:

- Reading and parsing the FRVT 2002 target and query XML files.
- Crash recovery to recover from power outages or other problems during the test.
- Support for running the face matching on three computers in parallel.
- Writing binary similarity files conforming to the FRVT 2002 Test specifications.
- Displaying status information on the Microsoft Windows console and writing status information periodically to log files.

ID-2000 System Description - Imagis Technologies Inc.

1. System Overview

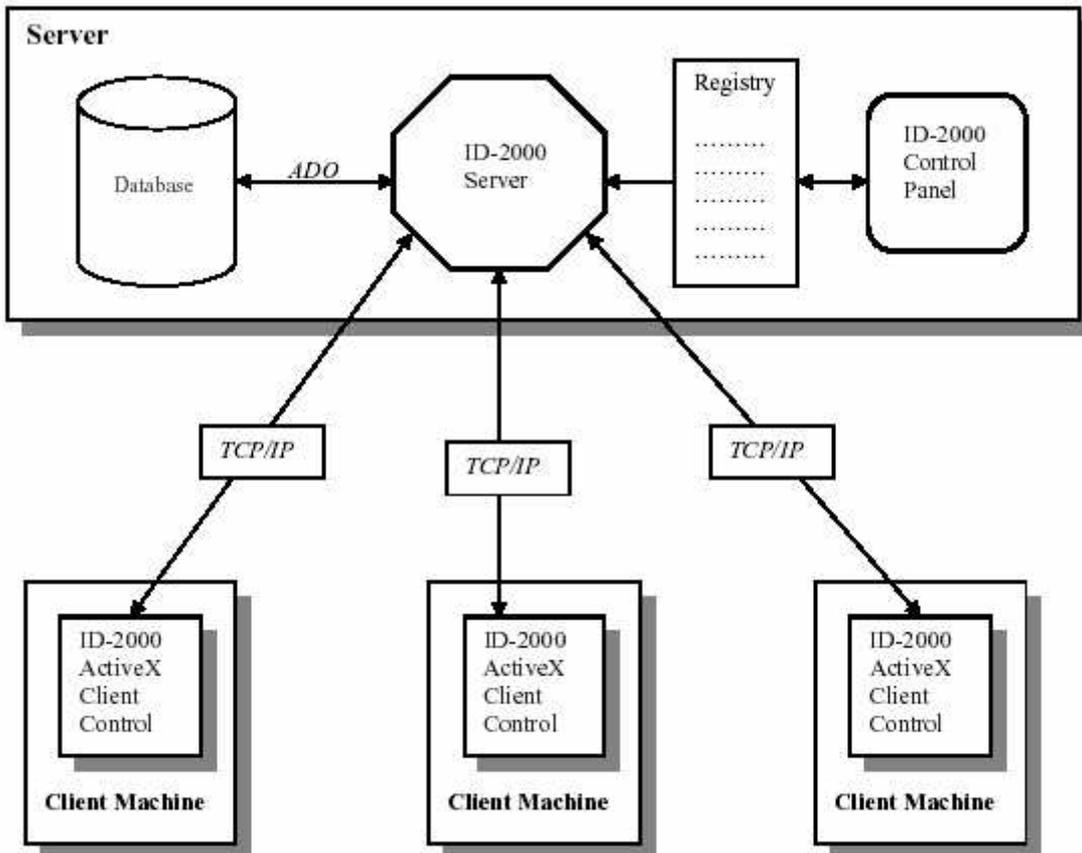
ID-2000 uses a deformable surface model to locate a face, normalize it, render it using various light source positions and find values for the pitch, yaw and roll of the head. The modeling allows the face detection to work accurately with an infinite number of face shapes (not just the "universal facial shapes" that are pre-determined by Eigenface-based approaches). Spectral analysis algorithms are then used to transform the detected face object into an encode array. The encode array can then be matched against other encode arrays in a database.

Platforms Supported	Windows 98 / NT / 2000 / ME/ XP
Architectures Supported	Client Server, Application Server, Stand-alone
Databases Supported	ODBC/ADO compliant databases such as Oracle, SQL Server, MSAccess.
Database Size	Unlimited. Note: some file based databases such as Microsoft Access have an inherent upper limit. Access has a size limit of 1GB
Input	Accepts any source of visual signals including photographs, live or recorded video. Also works with artist rendered images.
Faceprint Size	278 bytes. (269 MB of RAM per million records).
Speed	Calculations based on a single 1.7 GHz Intel Pentium 4 processor with 512 MB of RAM (1 Million records requires 269 MB of available RAM). Face finding and encoding using a 200 x 250 image: 500-800 milliseconds. One-to-many matching: 60 million per minute. The ID-2000 Server is multi-threaded so that it can respond to an unlimited number of simultaneous requests and so that it can fully utilize multiple processor machines.
Live Video Capture	Video capture is implemented through Microsoft Video for Windows (VFW). Any VFW compatible image capture system is all that is required for live video capture. The ImgWatch control provides a complete live video face recognition solution. Separate Authentication and Identification modes (as well as motion detection and scene verification modes). Images (or sequences of images) can be captured using the VFW capture window provided with the client ActiveX control.
Pose	ID-2000 finds faces where both eyes are visible, typically up to 25 degrees from frontal. Forthcoming releases of ID-2000 will address greater angles with yaw, tilt and roll.
Eyeglasses	Can work with eyeglasses provided the eyes are visible and not occluded by dark tints or glare. Thick frames may adversely affect search results.
Lighting	Does not require special lighting or background. Optimal performance is achieved in diffuse ambient lighting. In addition, performance is best if the subject is not back-lit, although this can be compensated for with gain control on the video camera. ID-2000.

Background	Finds faces in any background, plain or cluttered. Recognition is independent of background. Face recognition from live video is independent of background motion and / or camera motion
Image Color, Depth and Resolution	Functions with equal performance on color or gray scale images. Requires minimum 8-bit image depth.
Tracking	Tracks a face over time while encoding and querying the database for matches
Segmentation	Crops faces from background
Surveillance	Watch Mode enables a face in a video stream to be captured, encoded, and matched without human intervention

ID-2000 is Imagis' proprietary face recognition software technology. ID-2000 is a full ActiveX control and Server Service Executable developed with Microsoft Visual C++ 6.0 and ATL 3.0. The control is insertable and is functional with the majority of clients, including (but not limited to) VBA, VB, C++, Delphi, Internet Explorer and Netscape web browsers. The Server is an NT Service (or Tray Application on Win9X) that runs as a background process on the same machine as the Server Database is located.

The Client / Server architecture is shown in the diagram below:



The ID-2000 SDK is modeled after the client-server architecture. An ID-2000 client connects to an ID-2000 server and asks the server to do something such as find a person or encode a face. The server fulfills the request and sends the results back to the client. Under the client-server architecture, the server, in response to client requests performs the majority of processing. The client-server architecture provides the following benefits:

- An ID-2000 administrator is free to place ID-2000 servers wherever he or she wishes, with little regard to the location of client. The server can be located in the same room as a client or in another room on an entirely different continent. The location of the server is normally a well-kept secret.
- Because the server performs the majority of work in a client-server application, to increase client's responsiveness, it is often sufficient to increase the speed of the computer hosting the ID-2000 server.
- Client-server applications are ideal for use in low-bandwidth environments such as wireless mobile applications. It requires only a few hundred bytes of information exchanged to ask an ID-2000 server to find an individual. Client applications never talk to the ID-2000 server directly using TCP/IP sockets or other low-level communication primitives. Instead, they call methods and examine properties on the ID-2000 client control.

The ID-2000 server can use data in MS-Access, SQL Server 7+, Oracle 8+ and most ODBC compatible databases. The only requirements are:

- Records in the database be identified by a numeric primary key
- That the images managed by the database be in JPEG format
- That the server have write access to a single BLOB field to store the image encoding.

Pictures can be stored externally to the database in separate files, or stored in the database as binary large objects (BLOBs). Encodings must be stored in BLOB fields and be one of the following data types:

- Database Encode Array Data Type
- Microsoft Access OLE Object
- SQL Server 7+ Image
- Oracle 8+ BLOB

The ID-2000 server is implemented as a Windows NT service that uses encrypted TCP/IP as its transport. ID-2000 clients connect to the ID-2000 server on an administrator-defined protocol port number. The protocol port number defines the ID-2000 database available to the ID-2000 client. It is possible to set up different ID-2000 Servers on different protocol ports. For example, arrest and booking system may wish to keep its adult database separate from its juvenile database. To accomplish this, an ID-2000 administrator would set up two databases, each on a separate protocol port. This way, there is no chance that a juvenile would be returned inadvertently on a search of an adult database.

An ID-2000 server may be deployed where access to the database must be controlled. For this reason, a number of security features are available. The ID-2000 server administrator can restrict access to the ID-2000 server by username and password, by Internet address or both. For example, it is possible limit access to the ID-2000 server to John, Mary, Bob, James, and Darryl from computers whose Internet addresses are in the range of (142.104.6.5 .. 142.104.6.15).

Unauthorized attempts to access the server are logged in the ID-2000 event log that the ID-2000 server administrator can examine periodically.

2. Component List

a) Medium Computational Intensity Test

Hardware

- DELL Optiplex (Single Processor)
- Intel Pentium 4 Processor at 1.7 GHz
- 512 MB SDRAM
- 20GB 7200 RPM Hard Drive
- Cost - \$1600 US

Software

- ID-2000 Server Service \$25,000 US
- ID-2000 ActiveX Client Control (for static images)
\$500-\$3500 US per client depending on number of concurrent connections required. Note: Only one concurrent connection was used in the Medium Computational Intensity Test
- ImgWatch ActiveX Client Control (for live video)
\$500-\$3500 US per user depending on number of concurrent connections required. Note: Video data for the test was provided using previously captured frame clips, therefore this control was not used.

b) High Computational Intensity Test

Hardware

- DELL Optiplex (Single Processor)
- Intel Pentium 4 Processor at 2.2 GHz
- 1024 MB SDRAM
- 80GB 7200 RPM Hard Drive
- Cost - \$2200 US

Software

- ID-2000 Server Service \$25,000 US
- ID-2000 ActiveX Client Control (for static images)
\$500-\$3500 US per client depending on number of concurrent connections required. Note: Only one concurrent connection was used in the High Computational Intensity Test
- ImgWatch ActiveX Client Control (for live video)
\$500-\$3500 US per user depending on number of concurrent connections required. Note: Video data for the test was provided using previously captured frame clips, therefore this control was not used.

3. Modifications for FRVT 2002

The system used in FRVT 2002 contains some improvements to the face fitting routines.

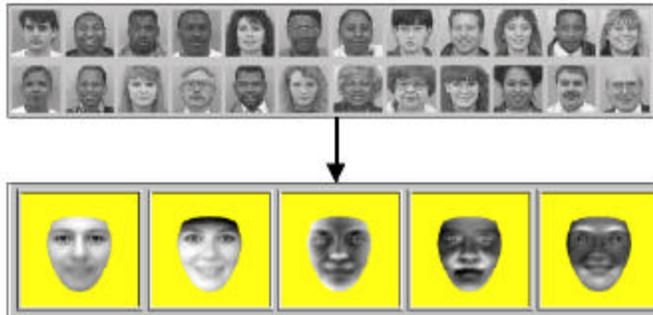


1.0 Viisage Technology Core Facial Recognition Technology

The facial recognition technology developed at the Massachusetts Institute of Technology (MIT) and exclusively used by Viisage Technology employs "Eigen faces", which are characteristics of a person's face, and maps the facial image into a multi-dimensional face space. Using special techniques developed by Viisage the Eigen faces are used to provide high speed facial matching to one or many candidate faces in a database.

Figure 1.1: Eigen Faces

The multidimensional Eigen space itself is determined through a separate process at Viisage Technology that is typically done only once. This process begins with a large diverse population of thousands of facial images. For each of these images, the head and eyes are located; the image is standardized, and then masked. The resulting ensemble of localized, standardized and masked images

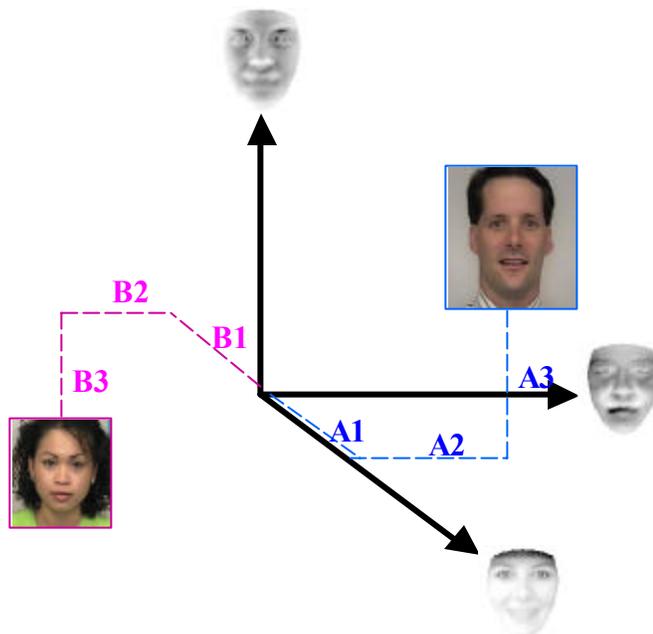


are then processed with a mathematical technique called Principal Components Analysis. The result of Principal Components Analysis a set of face-like images called Eigen faces. Each Eigen face is mathematically independent/orthogonal to all others, and is an independent degree of freedom for describing faces. In other words, these face-like images are the most efficient set of building blocks needed to build any face. Figure 1.1 shows the first five Eigen faces that result when Principal Components Analysis is performed on a large sample of face images.

Figure 1.2: Face Space

After standardization and masking, the image is projected into the multidimensional Eigen space of facial recognition as shown in Figure 1.2. The result of this projection of the facial image onto the Eigen face templates is a set of Eigen coefficients, which together form an Eigen vector. The multidimensional Eigen space is constructed of 128 mathematically orthogonal coordinates and each coordinate is representative of a single characteristic Eigen face.

The first coefficient of the image being enrolled is calculated through the projection of that image onto the primary Eigen space coordinate, which is also referred to as the average Eigen face. Once determined, the first coordinate projection is subtracted from the original image in order to produce a residual image. This residual image is then projected onto the second designated

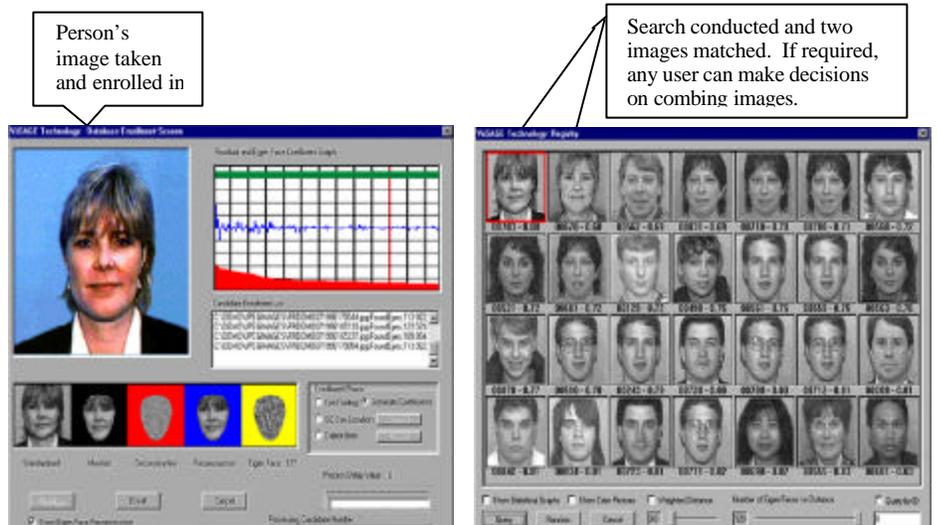




Eigen space coordinate and thus the second coefficient is obtained. Then the second projection is subtracted from the previous residual image in order to produce a new residual image. Successively, each new residual image is projected onto the next coordinate. Each projection subtracted produces a further deconstructed residual image.

This process of projecting the resulting residual image onto each coordinate produces a total of 128 characteristic Eigen coefficients. Together this set of characteristic coefficients represents a complete vector projection in the facial recognition Eigen space. The combination of these 128 coefficients with respect to their corresponding Eigen face images produces a reconstructed masked image that can be viewed and verified as being visually very similar to the original masked image.

The captured or scanned digital image is converted to Eigen coefficients (simply described as facial features) and submitted to the Facial Recognition Search Engine, which returns results in real-time to the requesting client. These results comprise the closest matches found for the individual. They are ranked in order with the closest match shown first.



2.0 The Viisage Technology Facial Recognition Software Development Kit

Viisage Technology has developed its facial recognition technology into a SDK that provides the fundamental facial recognition algorithms for development of applications that perform facial recognition. This FRSDK is used internally by Viisage Technology to produce surveillance, access control, and facial database search applications for its customers. The FRSDK itself is a product that empowers its users to provide state-of-the-art facial recognition in their own applications.

3.0 The Viisage Technology FRVT2002 Software

The facial recognition applications used by Viisage Technology in the FRVT2002 test are based on our FaceEXPLORER™. These applications consist of an automatic enrollment application and an automatic matching application. The enrollment application is customized to process the FRVT2002 XML scripts to enroll the images into an Oracle based facial recognition database. The matching application is customized to perform the comparisons specified in XML, using the data in the Oracle database, and write the results in files as specified in the XML scripts.

4.0 Component List

4.1 Medium Computational Intensity

Dell WorkStation 340



2.53GHz Pentium 4
1 GB ECC memory
80 GB ATA-100 IDE Hard drive
Microsoft Windows 2000 Professional operating system
CD-ROM, and CD-RW drives
Price: \$2507

4.2 High Computational Intensity

3 Dell WorkStation 530
Xeon 2.4 GHz processor
1 GB ECC Memory
80 GB ATA-100 IDE Hard drive
Microsoft Windows 2000 Professional operating system
CD-ROM drive
Price: \$4615 dual CPU.
1 3Com 8 port switch
Price: \$125

4.3 Software

FaceEXPLORER™ : typically Viisage Technology provides a custom quote. Pricing is complex and requires knowledge of the legacy operating systems in which one plans to convert prior database images as well as the manner in which the facial recognition technology may be used.

Our FaceEXPLORER™ system normally is a networked solution and not sold on a stand-alone basis. The product is shipped and licenses provided either per computer or per image it is application dependent. A baseline system would require, at a minimum, one software key, which would be priced \$10,000. FaceEXPLORER™ typically runs on Windows 2000, NT or XP platform. Viisage Technology would prefer to suggest the recommended components that would be integrated into a system such as cameras, lighting, backdrops, framegrabbers, databases, signature capture devices, printers and consumables.



Viisage can search a database containing millions of images and find a match in seconds.

FaceEXPLORER™ • Identity Fraud Prevention

Viisage's *FaceEXPLORER™* applications provide a real answer to the exploding problem of identity theft and fraud. Identity theft is one of the fastest growing crimes in America, claiming 1.7 million victims in 2001. Viisage has deployed the world's largest face recognition application, supporting a database containing ten million images, growing by 15,000 images a day, where State law enforcement officials maintain the integrity of their drivers' license database by decreasing the ability for individuals to fraudulently obtain multiple credentials under different names. Applications also include law enforcement booking and investigative systems, voter registration, social services benefits, weapons permits and immigration programs all of which reduce identity fraud. The technology searches millions of images in seconds – no competitor can match Viisage's performance in size, scope and functionality.





“Viisage’s system has met and exceeded our expectations, resulting in the need for us to expand our current licenses. The level of support we have received from Viisage confirms that we have partnered with the right company.”

A. Muwonge

*Uganda Secretary of the
Electoral Commission*





VisionSphere Technologies Inc.
Building M-50, National Research Council
1200 Montreal Road, Ottawa,
Ontario, Canada, K1A 0R6
Tel.: (613) 740-0245
Fax: (613) 993-8054

Overview of VisionSphere's UnMask Technology

How does *UnMask* work?

VisionSphere's products use algorithms and techniques based and developed on more than 30 years of research experience by Dr. Martin Levine of McGill University, the Chief Scientist of VisionSphere Technologies Inc.



Every face verification or recognition system must rely on the same three steps:

- Face detection and location of key features
- Extraction of facial descriptors
- Comparison of extracted information to a database

The first component of the system, which locates the face and its key features, determines the robustness of the system. The more advanced the face detection algorithms, the more invariant the system is to changes in head rotation, colour, brightness, hairstyle and facial expression. As a result, the face location and feature-finder components effectively determine how useful the face recognition system will be in the final application.

With the goal of constructing the most reliable and robust face recognition system possible, VisionSphere's UnMask takes the simplest and most effective approach to face processing. Initially, it locates the face and the eyes in the image automatically, using proprietary search algorithms. Next, the software normalizes and crops the image to provide excellent invariance to variations in lighting, head rotation, facial expression, and hairstyle. The image has now been made comparable to other face images and is ready to be analysed.

The second component of the system, which extracts the required information needed for comparing face images, determines the ultimate accuracy of the face recognition system. This stage computes a unique facial code for each individual.

For the second stage, VisionSphere's software uses a unique Holistic Feature Code (HFC). This provides a strong discriminating capability for comparing faces with very high confidence rates at fast processing speeds.

The third component of the system compares the coded information with data stored in a database. The percentage of false positives and false negatives determines the performance of this component of the system. Databases of face images that are used for law enforcement typically contain large numbers of images. For this reason, the speed and accuracy of the comparison system is critical to the usefulness of any face recognition system.

For the third stage, VisionSphere's software compares the two faces using a proprietary distance function. This function stresses the significant differences between faces and is another important factor in the system's high confidence rate and excellent speed.

Our product line is currently composed of three separate but related products:

UnMask Plus is a software AI (artificial intelligence) system for identification and removal of duplicate or multiple images from large photo-ID databases.



It's Me -- an advanced automatic computer logon authentication system. The system hardware and software components cannot only be used to verify the identity of a network or workstation user at logon time, they can also be used as part of e-Business verification, video conferencing, etc.



FaceCam -- VisionSphere's layered, biometric, user-verification terminal, designed and built specially for secure access applications, such as physical access control, time and attendance and registration systems



Face Recognition Vendor Test 2002

For the FRVT2002, two custom console-based test application written with our face recognition SDK were used for the Medium and the High Computational Intensity Tests. In general, the test application first locates the eye locations of all the images in the gallery set and the test set. Then, one face template is generated from each image in both sets. Next, it performs the matches between the face templates from the gallery set and the test set. Finally, the matching results were written in the form of similarity files according to the FRVT2002 specifications.

Component list

- Applications built using Visionsphere Technologies face recognition Software Developers Kit.
- One (1) computer for the Medium Computation Intensity Test (a single Pentium 4 1.7GHz processor with 512Mb PC133 SDRam with a 60GB Western Digital Hard Disk)
- Three (3) computers for the High Computation Intensity Test (Dual Pentium Xeon 2.0GHz processors with 1GB DDRam). Both used a 80GB Ultra ATA-100 (7200 RPM) hard drive and ran Windows 2000 Professional with Service Pack 2. Three high-end computers were used for the High Computational Intensity Test because of the nature of this test (finding the precise location of faces in 121589 images and performing nearly 15,000,000,000 comparisons in 264 hours).

Cost of the systems used in FRVT 2002

FRVT Medium Computational Intensity Test

One computer with a single Intel Pentium 4 1.7GHz processor and 512Mb Ram	Can \$1430
Unmask Identification SDK	Call Sales

FRVT High Computational Intensity Test

Three computers with dual Xeon 2.0GHz processors and 1024Mb Ram	Can \$3,700 each
Unmask Identification SDK	Call Sales

This page intentionally blank.

Appendix N

Participant Comments on FRVT 2002: Evaluation Report

FRVT 2002 Participants were provided a pre-release copy of the *FRVT 2002: Evaluation Report* and *FRVT 2002: Technical Appendices* on 1 March 2003. They were invited to submit a 6-page (maximum) document that serves as their "position paper" on the evaluation and the results. The deadline for submission of these papers was noon EST on 10 March 2003.

Participant documents are the opinions of the Participants and are provided in this text for reference. Inclusion of these documents does **NOT** imply that FRVT 2002 Authors/Sponsors/Supporters agree with any statements or derivative results within those documents. The identification of any commercial product or trade name does not imply endorsement or recommendation by the National Institute of Standards and Technology or any other FRVT 2002 Author/Sponsor/Supporter.

AcSys Biometrics Corp

AcSys Biometrics Corp.
399 Pearl Street
Burlington, Ontario
CANADA L7R 2M8
www.acsysbiometricscorp.com

Phone 905 634-4111
Fax 905 634-4111

FRVT 2002 -AcSys Vendor Comments

AcSys Biometrics Corp

AcSys Biometrics Corp

Executive Summary

AcSys Biometrics Corp would like to thank the members of the FRVT 2002 testing group for allowing AcSys to participate in this review and evaluation of facial recognition technologies.

AcSys Biometrics participated in this evaluation utilizing a pure proof of concept application based on of our Facial Recognition Software (FRS) engine. This engine was developed for physical and logical access control using live video during facial identification and enrollment. Our participation in this test allowed us to gather information about testing procedures for static image (2D) comparison, gain access to large 2D databases for the concept application and gather performance information that would be used in developing a commercial application for the 2D image market segment.

The performance results received from this test matched our expectations considering that the AcSys application, while still in its infancy, has been designed from the perspective of operation from continuous video feed, rather than static 2D images. With the information gathered from this evaluation and the recent FRS engine enhancements, the new AcSys 2D product that is being produced will we expect to substantially exceed the performance results of this test.

On the following pages we will address these issues;

- 1)** Review of testing procedures
 - 2)** Product Status at time of evaluation
 - 3)** Obstacles for the AcSys software within the evaluation
 - 4)** Reasons for entering only the Medium Computational test
 - 5)** AcSys Face Recognition engine improvements since test
 - 6)** Information on How AcSys FRS works
 - 7)** Summary
-

1) Review of Testing Procedures

The testing procedure was setup to enroll multiple images of the same person using different identity tags. The same dataset is used as both target and query set. This is an understandable way of doing the test because several statistics can be calculated from the same enrollment. This does give the AcSys 1.6 engine a disadvantage because the system as configured expects enrolled images of a person to have the same identity tag.

The AcSys FRS system is based upon machine learning, which operates by counter-training images of the individual against images of different but similar looking individuals. The result of having different ID tags for the same individual inhibits the operation of our system as it results in images of the same individual being used in both the training and counter training sets during generation of the biometric template.

For the MCInt - Still Images dataset, there are 9 images of each person containing different identity tags.

2) Product status at time of Evaluation

When AcSys was asked to join the FRVT 2002 testing the current Face Recognition engine was AcSys FRS version 1.6. This engine was developed and used for physical or logical access control using live video feed. Enrollment was done by capturing 30-100 pictures in a fully automated enrollment process while moving towards the camera over several feet, in order to capture significant variations in head pose and ambient lighting.

The application of matching static picture (2D) facial images (versus continuous video feed) had not been attempted by us prior to this test. At the time of the test, the AcSys system analyzed gray scale facial images extracted at a 32 x 32 pixel resolution. The extracted facial image is processed through a mapping function and does not measure specific features. The product employed in this test was designed for low pixel resolution video feed with the user typically at a distance of from 5 to 20 feet from the camera (~ 20 pixels between the eyes).

3) Obstacles for the AcSys software within FRVT

FRVT 2002 Evaluation Protocol – Rule 2

“To be able to compute performance for multiple galleries and probe sets requires that multiple signatures of a person are placed in both the target and query sets. This leads to the second rule: Each signature in the target and query sets is considered to contain a unique face. In practice, this rule is enforced by giving every signature in the target and query sets a unique random identifier.”

*“Each signature in the target and query sets is considered to contain a unique face.”
The faces are not actually unique, but they are considered unique.*

This way of testing creates a problem for the AcSys 1.6 engine in the way that facial biometric templates are generated.

The AcSys 1.6 engine is based upon machine learning, and as such counter-trains images of an individual against images of other similar individuals within the database, iteratively refining this separation of faces. In the FRVT 2002 test there existed several pictures of the same individual, containing different identities. This causes the AcSys 1.6 engine to counter-train the same individual against images of that same individual, thus inhibiting system performance.

4) Reasons for entering only Medium Computational

Our business at that time was focused solely on physical and logical access based upon continuous video feed. No development effort had been expended on analysis and use of high-resolution static images. Current developments are underway oriented towards this application area. AcSys decided on entering the MCIInt test to focus on gathering information about testing procedures that would allow us to improve the engine accuracy for single picture enrollment. Our next step will be to deploy a FRS solution the High Computational tests utilizing some of the information gathering in the round.

5) AcSys Face Recognition engine improvements since test

Changes to the AcSys FRS 2D and multi-dimensional image software were already underway at the time this test took place and more will be added based on some of the information gathered in FRVT 2002.

We are integrating the following improvements into the AcSys FRS engine.

- Automated 3D morphing of static 2D images
- Increased resolution of pixel array for facial image analysis
- Application of N-tuple logic to analyze and classify distinct facial regions. The approach employed at the time of this test analyzed a single region of the face.
- Optimization of FRS algorithm for high-resolution static picture (2D) facial image recognition.
- Optimized processing of assembly templates to achieve pattern matching rates of ~500,000 images/sec.

6) Overview of How the AcSys Face software works.

The AcSys Facial Recognition system is based upon a proprietary machine learning approach that uses non-linear complex domain analysis. Logical units referred to as "cell assemblies" are applied to separate tasks such as facial tracking, determination of head orientation and facial identification.

The face-tracking cell is trained using a database of up to several thousand facial images

and the neural network component applies a synaptic pruning and regrowth operation to optimize the selection of inputs and adjust complex "weights". Inputs to the cell assemblies are simply the raw image or frequency domain coefficients of the raw image, with prior normalization applied. The learning process used in this test required a minimum of 100 synaptic pruning and regrowth cycles, or approximately 10 seconds of processing.

The process is generic in that, same cell assembly structures are used to perform the various recognition tasks i.e. facial tracking, classification, etc.

7) Summary

AcSys Biometrics core facial engine is based on an advanced learning solution that can be modified or can train and learn independently to compensate for new variables submitted to the system. Having said that, the AcSys FRS system performed well considering no new training solutions or modifications were made to the beta test system. If AcSys would have been allowed to submit the images with the appropriate ID tags, and had the systems run the appropriate training routine on those images, the AcSys FRS solution would have produced substantially higher results.

The FRVT 2002 was a very comprehensive evaluation of facial recognition technologies and AcSys was pleased to take part, and AcSys will also offer its new products for further evaluation for the FRVT group due to the unique nature of our solution. This will allow the people who test face applications to gain a better understanding of the true strengths of the AcSys FRS engine, and allow AcSys to continue to enhance its security solutions.

AcSys, like most technology companies, will continually upgrade and modify our solutions to meet the ever-changing needs of the environments they are being deployed in. This test has allowed AcSys to understand another type of testing routine that is being developed to evaluate facial recognition software and we will continue to train our product to achieve the highest results in all of the possible testing environments currently being developed around the world.

Comments on the FRVT 2002 Report

Cognitec Systems GmbH
An der Flutrinne 12
D-01139 Dresden
Germany

1 Summary

Cognitec was invited to participate in the FRVT 2002 in April 2002. We took this opportunity and provided tools based on the then latest state of our FaceVACS technology. We are very pleased with Cognitec's performance compared to the other participants. These excellent results confirm that our technology is second to none.

The FRVT 2002 report presents many important results relevant to the practical use of face recognition systems. One issue, however, could have been treated in a better way for a meaningful comparison of technologies, and this would have demonstrated Cognitec's superb comparative performance even more clearly. While the reported identification performance allows the direct comparability between the participants, this is not the case for verification and watch list experiments due to the fact that the non-normalized performances are not compared directly in most of the figures. The performance scores for the participants who provided normalization functions are 'optimized' for the gallery, while the scores for the participants that did not were achieved making no assumptions about the data. We strongly believe that an evaluation based primarily on the non-normalized scores is more appropriate for real-world problems.

Cognitec did not submit normalization functions to FRVT 2002 even knowing that normalization does improve our performance under certain conditions. Based on the organizers' announcement that "FRVT 2002 will report both normalized and non-normalized performance scores" (see Appendix J, section 3), we assumed that the non-normalized scores would also be presented for all participants. Under this assumption we saw no need whatsoever to submit normalization functions for two reasons. Firstly, we think that normalization should not be regarded as being part of core face recognition technology. In our opinion, it is just a way of optimizing the results for a specific experiment but does not tackle the classical pattern recognition problem that the learned parameters of some function do not 'generalize' properly. This becomes more evident on noticing that it is possible to define effective normalization functions independent of the classification task. Secondly, the applicability of normalization functions in real-world scenarios is limited.

We indeed used normalization in the past. With normalized scores, Cognitec's performance in the verification and watch list experiments would have been significantly higher. In the following section we show the results we have obtained on the public FERET database as an example.

Aside from our critical remarks, we strongly believe that FRVT 2002 is a major milestone in the field of biometric evaluation technology. The FRVT 2002 report shows the improvements achieved in the last years as well as the weaknesses and the directions where future research is needed.

In the next section, we comment on Cognitec's performance results. After a deeper discussion of the normalization issue, we go into the results of the HCInt and MCInt tests in more detail. We also provide data on the hardware used and on the time spent for those tests. Section 3 lists some improvements of Cognitec's FaceVACS technology that we achieved since the time of the FRVT 2002, both in recognition performance and in processing speed. We close with some concluding remarks in section 4.

8 Comments on the Performance Results

The testing protocol implemented by FRVT 2002 addressed the evaluation of state-of-the-art facial recognition technology for different use cases and applications. The high computational intensity test was the most important test of FRVT 2002 as it measured the performance on 121,589 images of more than 37,000 persons.

Some of the tests aimed at the impact of three techniques for improving the performance: three-dimensional morphable models, normalization of similarity scores, and the use of video sequences for face recognition.

2.5 Normalization vs. Non-normalization

At first, the organizers considered normalization a post-processing technique for improving performance. Unfortunately, the presentation of some performance results mixes the normalized scores of some participants with the non-normalized scores of others.

Cognitec was among the participants that did not submit normalization functions even knowing that they improve the performance under very specific conditions. This is confirmed by several experiments we conducted. The impact of our normalization function for our algorithm is shown in figure 1.

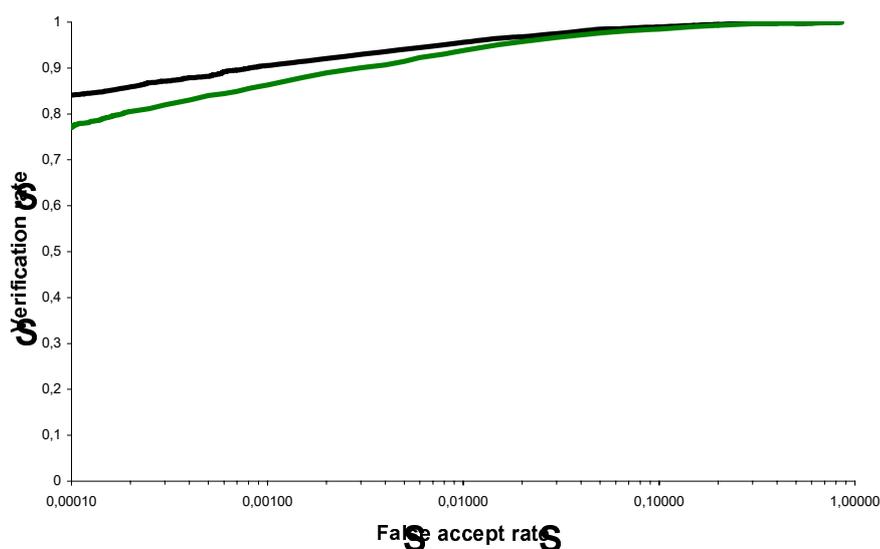


Figure 1. Normalized versus non-normalized results for Cognitec's algorithm on images from the FERET database (gallery: fa subset; probes: subsets fb (different facial expression), rb and rc (different head pose), and bk (different illumination); for additional information on these subsets, see http://www.itl.nist.gov/iaid/humanid/feret/feret_master.html)

Based on the organizers' announcement that "FRVT 2002 will report both normalized and non-normalized performance scores" (see Appendix J, section 3), we assumed that also diagrams would be included showing only non-normalized performance for the different technologies. This certainly is a fairer and more direct way of comparing the performance of the participants. Said assumption was the main reason why, finally, we decided not to provide score normalization.

Another reason is the limited applicability of normalization. In general, we think it is desirable that the result of a comparison between a probe and a gallery image does not depend on what other images are in the gallery. In the last paragraph on page 9 of the FRVT 2002 report, the authors also address this issue and ask: "When normalization is applied, is

verification still '1 to 1' matching?" Certainly it is not! Normalization requires that a probe be compared with a gallery and therefore access to the gallery has to be provided for the comparison of just two templates. Normalization not only demands more computing power for verification (as many comparisons have to be performed as images are in the gallery), but also the assessment of the performance of a system becomes a very tough practical problem when the gallery at hand changes over time.

There are applications in which the gallery is not available and thus are not amenable to normalization. An example is an access control system comprising smartcards containing the biometric template and terminals that take a live image of the cardholder and perform a verification using that live image and the template on the card. In such a system, the gallery is distributed over the smartcards and cannot be used at verification time for normalization. Access to the gallery can also be restricted due to privacy policies.

In our opinion, score normalization lies outside the face recognition process. It can be advantageous in some situations while in others simply it is not or might even cause problems. For example, in identification, normalization is irrelevant¹.

It should be possible to use a generic method for generating normalization functions based on the estimation of the similarity score distributions for a range of galleries and probe sets. This is exemplified by the normalization function we have implemented, which works as follows:

A normal distribution function is fitted to the set of scores. From the parameters of that distribution function, the parameters of the uniquely determined linear transform are computed that maps the fitted normal distribution to the standard normal distribution (i.e. zero mean, unit variance). That linear transform is then applied to all scores in the set.

Obviously, that normalization function can be used for any non-normalized scores whose distribution is near to a normal distribution, independent of the specific classification task. Thus for an objective comparison of core face recognition technologies, only the non-normalized results should be used. Figure 2 shows the non-normalized verification performance ROC on the HCInt large gallery and probe set for Cognitec and the second and third best performers.

While the non-normalized scores for verification in the HCInt test are reported for all participants (for the participants who submitted normalization functions these results are shown in Appendix L, figures L.10 through L.13), there are some results for which no corresponding results based on non-normalized scores are provided. Therefore we consider the reported data incomplete; the missing data would have been very valuable for a meaningful comparison of the participants.

¹This holds if the order of the similarity scores between faces is not changed. However, we do not see how changing the order could make sense, and it seems indeed that the submitted normalization functions do not change it.

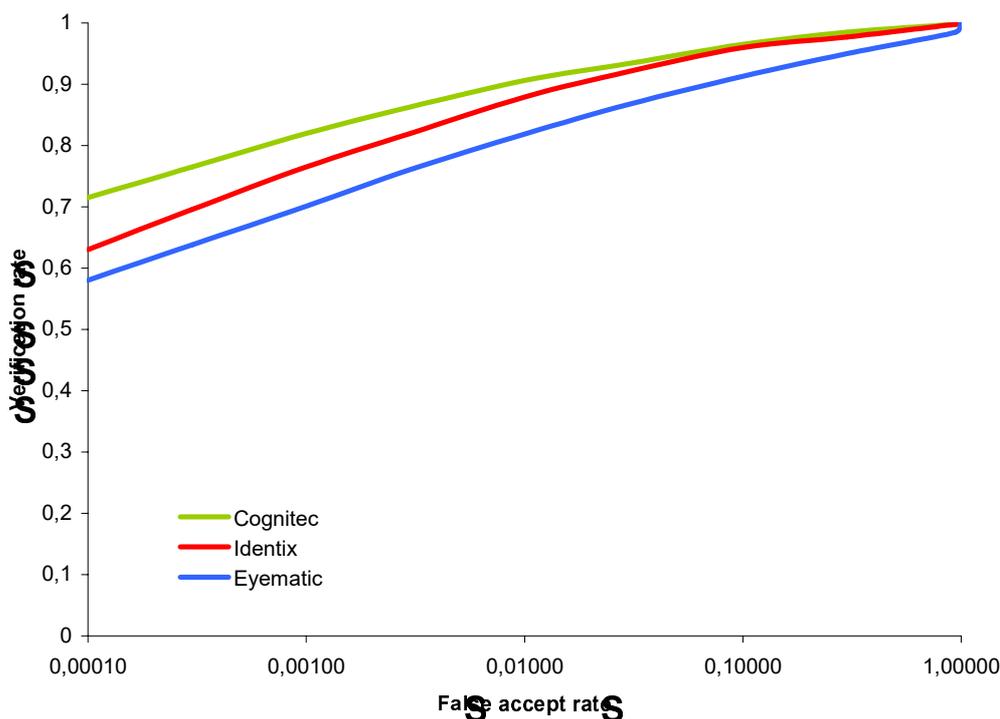


Figure 2. Non-normalized verification performance ROC on the HCInt large gallery and probe set for Cognitec and the second and third best performers.

2.2 High Computational Intensity Test

The high computational intensity test (HCInt) was the most important test in FRVT 2002. As the report says, it aimed at the evaluation of the performance of state-of-the-art systems on extremely challenging real-world problems. The excellent performance of our technology in these experiments confirms the top position of our products to satisfy best the demands of the market.

We found interesting that our algorithms discern males better than females. In our opinion, this is not a property of our algorithms. We have observed that women's facial appearance is more homogeneous than men's.

2.3 Medium Computational Intensity Test

The medium computational intensity (MCInt) test assessed the impact of further techniques for improving the performance of current technologies: three-dimensional morphable models, and the use of video sequences for face recognition.

Unfortunately, the comparison of non-normalized results for those experiments is not possible as the non-normalized data is missing in the report for those participants that used normalization.

Three-dimensional morphable models

Cognitec's algorithms were not trained specifically to deal with faces rotated by more than 15 degrees. The devised experiment shows that a large improvement can be achieved by morphing non-frontal faces to their frontal view. Although these techniques still require considerable computing power, they could become feasible in the near future.

Care has to be taken when interpreting these results because the measured performance is not only determined by the facial recognition system but also by the quality of the morphing system.

Video images

Cognitec's FaceVACS technology does not exploit any properties related to video imagery. Therefore, it is not surprising that these experiments did not show an improvement of Cognitec's performance.

2.4 Processing Speed and Hardware

The FRVT 2002 did not directly address the evaluation of computing requirement and processing speed. In real-world applications though, this is a very important aspect. In table 1 we give the hardware used for the tests and the figures for the time spent on each of the tasks. This shows that our technology is also very competitive regarding computing requirements and processing speed. Further details of the hardware can be found in Cognitec's system description in Appendix M.

Table 1. Times spent on the HCInt and MCInt tests²

Test	# of signatures	# of comparisons	Time to enroll	Time to compare	Comparisons per second	Hardware
HCInt	121,589	~15 billions	12 hours	104 hours	~40,000	Athlon XP @ 1.7 GHz 1 Gbyte RAM
MCInt	7,722	~60 millions	6.5 hours	1 hour	~16,600	Mobile Pentium III @ 650 Mhz 128 Mbytes RAM

3 Improvements of the FaceVACS technology since the test

The FRVT 2002 took place eight months ago. For facial recognition, one of the most challenging biometrics around, this is a long time, as the level of development efforts remains very high. Cognitec keeps its focus on the improvement of FaceVACS and its underlying algorithms and techniques.

3.1 Increased Performance

We increased the detection rate and the accuracy of our face and eye finder by employing better feature detectors and by employing better methods for learning the parameters of the classifiers. In addition, we improved the recognition algorithm, leading to lower overall error rates and less performance differences between faces of different ethnicities.

3.2 Current Matching Speed

We could improve the speed of the comparison function from former 40,000 comparisons per second achieved during FRVT 2002 up to 500,000 comparisons per second on a single Pentium 4 @ 2.4 GHz. This was achieved through code optimizations and through using optimized functions that take advantage of the latest SSE2 instruction sets on modern processors for the most time consuming operations.

² The given times include the time spent on file I/O. In particular, the enrollment time includes the time for reading the images and writing the templates, and the comparison time includes the time for reading the templates and writing the similarity scores. The enrollment time given for the MCInt test actually refers to enrolling the signature set twice, once as target set and once as query set.

4 Conclusions

Cognitec is very pleased that the high quality of our face recognition technology FaceVACS was confirmed by the FRVT 2002. We have been developing this technology since 1995, and will continue to improve its recognition performance and to extend its range of applications. To an increasing degree, this will encompass application areas for which other biometrics are currently preferred.

Both the biometrics industry and potential customers in government and industry will be very grateful for the results of FRVT 2002. So far it has been difficult for them to assess the different face recognition technologies on a common and neutral basis; this is what FRVT 2002 has provided with very demanding and real-world data. We think that this type of evaluation should be repeated on a regular basis.

Cognitec thanks the organizers and sponsors of the FRVT 2002 for the opportunity to participate and for the excellent and independent work they did.

Remarks on this commentary are appreciated and should preferably be sent as e-mail to info@cognitec.com.

Comments from C-VIS on the Facial Recognition Vendor Test 2002

The FRVT 2002 is an important technology evaluation. Its results provide valuable insights into the state of the art in facial recognition. However, the test will only partially give a potential user of automated facial recognition an answer on whether this technology fulfills his expectations.

The results of the High Computational Intensity test (HCI) are valid only for images of an ID photo type. Among the HCI images that we saw, no one showed a face in front of a structured non-uniform background and no images came from typical video sources, for example. As a consequence, a HCI test on a realistic mix of image types may have very different results both in terms of the general capabilities of the technology and for the individual vendors.

The main product of C-VIS, the FaceSnap RECORDER, has been designed for real-time surveillance applications. The kind of image data we are confronted with in typical applications is often quite different to most of the images used in the FRVT 2002. In the previous FRVT in the year 2000 there was an attempt to include in the test a realistic application scenario. In our opinion, the short video sequences from files that were included in the FRVT2002 are not an equal substitute to a real video surveillance test. As a suggestion for the future, a recognition experiment from a digital video footage of at least a few hours length should become part of the FRVT.

There are a number of important real world applications for which customers seek a complete and viable solution. In a real application scenario, the success or failure of a particular system may not depend primarily on the performance of a single "core technology". Instead, the quality of interaction of all system components and the suitability of the system design for a particular purpose becomes crucial. A commercial system should be tested against its claims with respect to particular applications. For future tests, it would be desirable to exactly identify the most important applications of facial recognition and to tailor test procedures to them.

In the FRVT2002, C-VIS has confirmed its position as one of the leading players in this market. The priorities of our technology developments are driven by actual customer needs. A test like the FRVT2002 can of course not reflect the whole spectrum of challenges encountered in practical applications. However, we hope that the FRVT team continues its excellent work, making the FRVT a regular benchmarking service for the commercial face recognition community.

Bochum, March 10, 2003

Prof. Dr. Thomas Zielke
President
C-VIS GmbH



Comment on FRVT2002 Evaluation Report

Eyematic Interfaces Inc.

March 2003

After reviewing the FRVT 2002 evaluation report and considering our on-site experience, Eyematic is very pleased with the overall results. The test was conducted in a well-organized manner, and special care was given to ensure equal opportunities to all vendors, which gives high credibility to the test results. We appreciate the straightforward way in which the test data was described and presented. Assistance from the FRVT Team was pleasant and helpful to complete all test runs without any problems for which we needed only half of the allotted time.

In this statement we would like to discuss our motivation for our system's configuration and consequences for the test result of HCINT. We will then consider aspects of MCINT, add a few remarks about technical advances we have made since July 2002 and finally give a few suggestions for future face recognition tests.

Importance of Adaptive Systems

Considering the diversity of applications made possible through face recognition technology, a reader of the evaluation results must understand that the variety of possible conditions under which images can be acquired are practically endless and a public test can not yet be seen as comprehensive for all aspects of use. Specific face recognition applications always impose specific requirements to the underlying technology. Image data that was acquired under controlled conditions like plain background and well-distributed illumination clearly represents the ideal scenario, which is optimal to investigate effects related to the variability of faces. But the value of a system performing well under these conditions still remains questionable with respect to applications like "Access Control" or "Surveillance" where we are confronted with unstable environmental conditions like illumination-variation, pose, etc. These conditions have a great influence on the system's performance; despite efforts to control them through hardware (e.g. artificial light). Secondly, even in so-called "Cooperative Scenes" we must make allowances for the quality level of cooperation in the absence of a human operator. This is particularly important if the application is built for convenience: A 'fast lane' for frequent flyers; access control replacing the use of ID cards or typing an ID number; quick validation of credit card or passport; etc. To properly face this challenge, the fundamental technology must become adaptable.

We understand ourselves as a supplier of the general core technology. Our customers integrate it into their applications. This requires us to allow our customers freedom in choosing the environmental conditions and still provide them with software solutions that can cope with such conditions. For years Eyematic has undertaken significant R&D efforts to solve these problems. We have developed methods that allow a customer to adapt our system to any condition, even those unknown to us, without imposing additional restrictions in the generality of usage or raising the complexity of our API. One of our most successful solutions is an automatic training tool that can optimize the entire face recognition system to specific scene conditions through sample images with some ground truth position information. Recent developments allow such training even without ground truth data.

Details about our system configuration for FRVT2002

The FRVT2002 did not allow any sort of adaptation or training during the test. Since no specific information about the image conditions was disclosed, we chose to configure our system for high invariance to pose, size, background texture and some illumination as it occurs for access control systems. Our customers frequently face these conditions. To limit the influence of background texture we suppressed hair and head-shape cues for recognition. However, such cues can be beneficial when the faces are presented before untextured background (as it was apparently the case in the HCINT)¹. After reviewing the evaluation report, we came to the conclusion that our configuration was good for most data used in MCINT but it was sub-optimal for HCINT. Additionally, since our largest customer is located in Asia, our face databases have a large portion of Asian faces. This might have given our system a slight bias.

MCINT Test Results

This test mimics better than the HCINT test the conditions present in typical applications such as access control or passport authentication.

Indoor-Outdoor Test

Our scores in all indoor conditions reflect our optimization for indoor access control systems. We are pleased to have reached highest rank here. For outdoor conditions it is possible to raise the performance through training, however, we admit that reaching an acceptable level is not at all easy under the influence of direct sunlight.

Head Pose Test

We are pleased to have reached highest or second highest scores in all pose (morph) experiments (Evaluation Report Fig. 23). However, 45 degrees head rotation is rather extreme for systems designed for frontal faces and for most practical purposes we found that tolerance to pose variations of +/- 20 degrees is sufficient. We would like to point out that our SDK can be adapted to any dominant pose as long as enough recognizable facial features remain visible (We do not require visibility of both eyes). This is necessary for applications where, due to camera mounting limitations, faces can only be viewed under an angle and consequently seldom appear frontal.

Still-Video Test

Our SDK includes specific features that were designed for optimized performance on video or still-set data. Unfortunately we introduced a bug in the FRVT2002 test setup that was discovered a few weeks after the test was finished. It affected the performance of the Still-Video Test. Despite this error we are pleased to learn that we still reached the best score (Evaluation Report Fig. 24).

Technical Advances since July 2002.

About ¾ year has elapsed since we prepared the product description for FRVT2002 as presented in Appendix M of the Evaluation Report. In the meantime we have reduced the size of a face template from about 4 Kbytes (this was used in the test) to less than 1.5 Kbytes without any loss of accuracy. With small loss of accuracy we now offer template sizes below 1 Kbytes. This has

¹ This might explain why among the three best scoring vendors our system showed the least sensitivity towards gender (Evaluation Report Figures 17 through 19).

drastically increased the template matching speed and its memory economy for large galleries. We also increased the overall processing speed by over 50% without any loss of accuracy.

Suggestions for future face recognition tests

More emphasis on real-world scenarios

While the design of HCINT was excellent to answer questions regarding temporal, gender and age effects or the use of large databases, it did not address typical challenges of real-world applications, which are: Illumination variation, pose variation, occlusion, noise, lens distortion and other. MCINT covers some of these issues but not sufficiently due to its size limitation. We also suggest (re-) introducing fine-grained tests for pose variation (10, 20, 30, ... degrees rotation) as well as variable head size.

Evaluation of methods for Adaptable Systems

Unless our computers become several orders of magnitude faster and the software much smarter, there will always be a need for adaptability to specific scenes. A system that achieves peak accuracy on any possible image containing a face under any possible condition without adaptation does not seem realistic for the near future. Consequently vendors that have a flexible system produce scores that depend on its configuration rather than its principal capability. We suggest introducing a session in which a percentage of image data is extracted from the test data set and given to the vendor for a limited time before the actual generation of similarities begins. The vendor should be allowed to use the image data in order to train his system in order to properly configure it for the subsequent test. This simply mimics the situation that is prevalent in most real world applications.

Different presentation of results

For tests where systems reach a very high score (e.g. Evaluation Report Fig. 23 (Front)), results like 99% and 98% verification rate are hardly distinguishable in the chart although they would represent significantly different performances. We suggest plotting against "Failure Rate" with using a logarithmic scale.

We would like to give thanks to all FRVT sponsors and organizers for making this test possible.

Contact:

Eyematic Interfaces Inc.
9800 S. La Cienega Blvd., 5th Fl.
Inglewood, CA 90301
USA
www.eyematic.com

IconQuest

www.iconquesttech.com

Comments on the Facial Recognition Vendor Test 2002

I. OVERVIEW

IconQuest has developed an image recognition system, IQ MatchMaker™, which is focused on solving recognition problems with poor quality images. For example, such images may be very small, blurred, partially obscured, or photographed through dirty windows. Poor quality images arise routinely in the real world.

IQ MatchMaker™ is not based on an underlying structure characteristic of faces, but applies to general imagery. For example, by simply changing the Probe and Gallery images to photos of planes and reconnaissance photos of airfields the same software used in FRVT 2002 can locate a variety of aircraft in the reconnaissance photo.

In order to apply the general image recognition engine in IQ MatchMaker™ to face images, a segmentation algorithm was incorporated as an option into the software.

In FRVT 2002, IconQuest made only minimal use of the full size imagery provided by the Sponsor. Specifically, the full size imagery was only used for conversion from color to gray scale, histogram equalization and segmentation. Once these tasks were accomplished, the images were scaled down using pixel averaging to smaller images, each containing no more than 1536 pixels. In most cases, the scaled-down imagery contained substantially less than 1536 pixels. For comparison, note that 1536 pixels could come from an image of resolution 32 x 48, while the MCInt images provided by the Sponsor varied in resolution between 240 x 320 and 720 x 480. IconQuest made no further use of the full size imagery. Only the scaled down images were used in the image comparison and analysis phases of recognition. Therefore, all results relating to IconQuest in the FRVT 2002 Evaluation Report are for segmented images having an area containing no more than 1536 pixels.

II. PROGRESS SINCE THE FRVT 2002

(A) Identification

IconQuest tested its IQ MatchMaker™ technology to see the dependence of image size and Identification performance. 695 Probes consisting of the first image of an individual in the FERET database of thumbnails and a Gallery of another 6554 images in the FERET thumbnail database were used. Pixel averaging was used to scale down the FERET thumbnail images. The results for Rank 1 are:

<u>Image Size (pixel area)</u>	<u>Identification Rate</u>
8 x 8	.72
11 x 13	.72
21 x 22	.90
37 x 41	.97

These results represent a significant improvement over the performance of the system as it existed in July of 2002, reflecting substantial modifications to the software since FRVT 2002. Additional results on small, blurred imagery are contained in the March 2003 Advanced Imaging article *Through a Glass, Darkly : Image Recognition with Poor Quality Imagery*.

(B) Verification

IconQuest had done no Verification performance testing prior to the FRVT 2002. Since that time, an algorithm for re-normalizing similarity scores has been added which substantially improves verification performance. A number of specific tests were conducted using segmented FERET thumbnails which were then scaled down using pixel averaging. An example that illustrates the improvement in Verification performance follows. A Gallery consisting of the first image of each of the first 100 individuals and Probes consisting of the second image of each of the first 608 individuals was created by scaling the segmented FERET thumbnails down to 21 x 22 pixels. An operating threshold was set to reach a false accept rate of .01. Without re-normalization, the system produced a false alarm rate of 46%. With re-normalization, the false alarm rate was reduced to 26%.

III. COMMENTS/ANALYSIS OF FRVT 2002 RESULTS

(1) Identification

IQ MatchMaker™ performed well in Identification Performance for Indoor (same day) Probes; see Figure L.15 where IconQuest results varied between best, second best and third best as rank varied.

(2) Verification

- (a) IQ MatchMaker™ performed relatively better in the Identification tests than in the Verification tests. This is not surprising since, as noted in (B) above, IconQuest conducted no verification testing prior to the FRVT 2002. Testing was critical in optimizing the software to perform better in identification tasks.
- (b) As noted in the FRVT 2002 Evaluation Report, (page 19), normalization of similarity scores increases verification rates. Of the five participants who scored highest in verification tests (see Figure 22 on page 33 of the Evaluation report), four submitted normalization functions which were used in computing verification results. As noted in (B) above, IconQuest developed a normalization function only after the FRVT 2002 time limit to submit such functions expired. Based on results typical of those noted in (B) above, substantial improvements related to verification performance are expected to result from the use of IconQuest's current normalization function.
- (c) Despite the lack of optimization based on verification testing prior to FRVT 2002 and the lack of a normalization function, IQ MatchMaker™ tended to perform mid-range in verification tests, when it performed well in identification tests; see, for example, L.14 & L.15 and L.34 & L.35

(3) Pose

- (a) IQ MatchMaker™ performed relatively well in the identification tests of pose variation. For example in Rank 1 Identification performance, IconQuest was:

- (i) best in 30 Degree Up; (ii) second best in 30 Degree Down; (iii) tied for second best in 45 Degree (Left); and (iv) tied for fourth best in 45 Degree (Right).
- (b) As noted in the Overview, all of IconQuest's results are based on similarity scores obtained from images which were scaled down to contain no more than 1536 pixels. The pose results are remarkable since they represent comparisons of images resulting from the composition of pose distortions and significant scaling transformations.

(4) Lighting

- (a) Changes in lighting (e.g., Indoor (same day) Probes Vs. Indoor (same day, overhead) Probes) challenged the system. Since we are dealing with a general image recognition engine, as opposed to a face recognition system, an ideal solution to this problem would allow recognition despite changes in illumination in general imagery. This represents an area of on-going interest.
- (b) As noted in the overview, all of IconQuest's results are based on similarity scores obtained from images which were scaled down to contain no more than 1536 pixels. Therefore the image variation resulting from the lighting changes is really the result of compositions of the lighting distortions and significant scaling transformations.

(5) Background

Samples of Outdoor imagery contained in the Evaluation Report display a much greater background complexity than those samples of the Indoor imagery. IconQuest selected an image segmentation algorithm on the basis of its speed. This non-proprietary algorithm was selected after testing on a large number of FERET images. The FERET images tend to have backgrounds closer to these Indoor imagery samples than to the Outdoor imagery samples. The segmentation algorithm, which we selected, is fast, but doesn't work as well for images with complex backgrounds. We expect that this was not an appropriate choice of segmentation algorithm for a number of FRVT 2002 images with complex backgrounds and provides at least a partial explanation of the variation in performance between Indoor and Outdoor imagery.

(6) Temporal Variations

Analysis of the variation between same-day and different-day results isn't really possible without access to the relevant image data. In the case of Outdoor images, we expect that changes in natural illumination from day to day might offer a partial explanation.

(7) Video

In the Abstract of the Face Recognition Vendor Test 2002 Evaluation Report, the Sponsor observed that "face recognition from video sequences offers only a limited increase in performance over still images." We feel that the results supporting this conclusion, at least from the vantage of IQ MatchMakerTM,

may well reflect the lack of a video database similar in variation to the FERET database for use in optimizing the software, and that rapid and substantial performance improvement would result from acquisition of such a video database.

(8) Morphable Models

- (a) The Morphable Models (“MM”) approach is an example of a more general approach in which images undergo pre-processing transformations. Despite the lack of utility of employing the specific morphable model used in FRVT 2002 together with IQ MatchMaker™, we believe other transformational approaches can have great value.
- (b) The MM tests of MCInt seem to be of a different nature than the rest of the FRVT 2002 tests. FRVT 2002 was a time-limited test in that each participant had to complete production of similarity scores in no more than 264 hours. Except for the MM tests, the Probe and Gallery images represent real world images and participants are free to spend as much time as they like pre-processing images, but their pre-processing time is counted toward the test time. However, the MM tests reflect mandating a specific pre-processing step (performed by the Sponsor), but without counting the time it takes to carry out that step towards the overall test time. Without counting this time, it is not possible to support the conclusion of the Evaluation Report that “Results show that three-dimensional morphable models. . . increase performance” especially in the context of a time-limited test. In addition, results show that the use of MM decreases performance in at least one important case: frontal imagery. In more detail:
 - (i) Utility of MM and Computation Time: The computation time, MMTime, expended by the Sponsor in the MM pre-processing should be given. Since Participants were not given external pose information concerning Gallery images, they would have had to apply MM pre-processing to all images. To the extent MMTime is negligible, then the objection raised here is minor. If MMTime is significant, then the objection raised here is material. For example, if a Participant added MMTime to the 264 hour time limit and its performance increased more than the increase observed by using MM in the 264 hour test period, then the proper conclusion would be that “Morphable Models decrease performance”, at least in the context of a time-limited applications represented by FRVT 2002.
 - (j) Utility of MM and Pose Statistics: FRVT 2002 test results (see Fig. L-25 and Fig. L-24) show that Morphable Models decrease identification and verification performance on frontal imagery. The extent of decrease varies considerably among Participants. Whether or not MM pre-processing is practically useful then clearly depends on the statistics of pose variations in an image population.
 - (k) Utility of MM and Test Structure: Except for MM subtests, FRVT 2002 tests the performance of face recognition alternatives in a situation where a user has no advance knowledge of Gallery images.

In many real world situations, the user has a database of known imagery and is attempting to perform Image Recognition (“IR”) functions such as identification, verification or watch functions on unknown Probe images. In such situations, Gallery images are known in advance and can be subject to substantial pre-processing and any time limit would apply only to performing the IR functions on Probes. Then it would be reasonable not to count MM pre-preprocessing time as part of the time limit. In such a test, Participants would be given Gallery images in advance of the test and should be able to choose how to pre-process the Gallery imagery. Then one could evaluate the state-of-the art in pre-processing and decide whether MM increases performance more than other pre-preprocessing techniques.



Comments on FRVT 2002

Identix FRVT2002 Team

March 10, 2003

INTRODUCTION	2
PERFORMANCE STRATIFICATION	2
FACIAL RECOGNITION VALIDATED FOR NATIONAL ID SYSTEMS	2
FURTHER RESULT HIGHLIGHTS	3
VARIATIONS BETWEEN TEST RESULTS AND TRUE PERFORMANCE	3
HARDWARE AND COST VARIATIONS BETWEEN TEST SYSTEM AND REAL SYSTEMS	4
WHY CHOOSE IDENTIX?	4
FINAL COMMENTS	5

Introduction

The results of the FRVT 2002 tests demonstrate Identix Facelt technology to be a leading performer in all categories. Through the various tests, facial recognition has been validated for large one-to-many applications (i.e., national ID) and verification applications. The tests have also shown the significant value that facial recognition can deliver in watch list applications. Most significantly for the field of biometrics, the tests show that facial recognition biometrics do work, and they can substantially augment fingerprint applications when fused together.

While there are a few variations between the test results and actual performance which is discussed briefly below, the evaluation overall paints a very strong picture for Facelt technology and Identix as a whole, which stands alone among the top performers in being a large, stable, US based publicly traded company. Identix is a multi-biometric company with over 20 years of real world experience, established partnerships with a large number of systems integrators, card manufacturers and other biometric solution providers.. We have worldwide facial recognition and fingerprint biometric deployments for national ID, drivers' license, voter registration, background checking, physical and logical access control, social security payments and airport security.

Performance Stratification

The results of the FRVT 2002 tests show that there are at present three top tier technologies: Identix, Cognitec and Eyematic. This top tier membership was observed consistently for Identification, Verification, and Watch list tests. The best of breed technologies outperform all others regardless of task.

The close clustering of results for these three companies indicates a convergence of performance and thus, suggests the maturation of the field of automated facial recognition. Such clustering has not been observed in previous tests.

The convergence of results does not mean that further improvements in performance should not be expected. As pointed out in the FRVT2002 report, performance gains have been significant over every previous test period. Furthermore, internal testing of current (March 2003) Identix algorithms demonstrates performance gains over the (June 2002) algorithm used in FRVT2002.

Facial Recognition Validated for National ID Systems

For the first time, the US Government has published the performance of face recognition for large databases (>100,000 images). This report should be considered a highly reliable validation for those interested in implementation of face for applications such as identity card programs and border control. While the expected performance of the system will vary with the quality of the database, these tests shows that for real-world face images and real-world size databases, *face recognition works*.

The conclusion that the performance of one-to-many identification lookup decreases only linearly with the logarithm of database size (p. 44) is very good news. This important result has been extended internally at Identix to database sizes of up to 10 million. These results indicate that face recognition is a viable option for state and even national ID systems.

Note that performance will be further enhanced when demographics are used to segment the database into bins, thus reducing search size before performing the biometric search.

Further Result Highlights

Face Recognition Works. In general, the results of the various sub-tests show that face recognition technology:

- Works well for one-to-many identification with real-world databases.
- Works very well for indoor verification.
- Provides significant value for watch list (surveillance) applications.

Improved Accuracy. Identix is particularly pleased about the improvement in accuracy that our technology has demonstrated since the last FRVT in 2000.

Minimal Sensitivity to Aging. Contrary to popular belief, Figure 16 shows that face recognition performance is only slightly sensitive to temporal variations (changes over time). Thus, we would argue that template aging is not a significant problem for adults.

Unaffected by Indoor Lighting. The indoor lighting variation problem has been solved. Facelt is unaffected by changes in lighting for images taken indoors (see leftmost two data points in Fig. 22).

Suitable for Fusion with Fingerprint. As discussed by the authors (see Fig. 25) facial recognition provides greater accuracy than single fingerprint (circa 1998) for verification applications requiring relatively low security (FAR of 0.01 or lower). This cross-over result, together with the finding that face has reached a level of maturity sufficient for large-scale one-to-many applications, suggests that fusion of biometrics is an idea whose time has come.

Identix offers both facial and fingerprint solutions, and has the expertise to develop and deploy dual and fused biometric systems where these provide the best solution.

Variations Between Test Results and True Performance

While overall we are very impressed with the scientific validity of FRVT2002, we briefly wish to list slight variations between true Identix performance and that measured by the test.

Current Technology Delivers Improved Performance. Identix prides itself on active research and development, so that already our performance is better than what is shown in this report.

Use of Video Can Benefit Facial Performance. Identix challenges the assertion that video is not an advantage for facial recognition. The source of confusion on this point is that the video used in the still versus video test was unrealistic, containing either pure pose change or pure opened/closed mouth (talking) change. In order to properly test the impact of video, sequences would contain natural variations in resolution, lighting, pose and expression associated with humans moving through space.

Video also offers many opportunities to perform real-time registration of facial features. Registration errors are a primary source of recognition failure.

For surveillance applications, Identix offers a finished product, Facelt ARGUS. We expect that Facelt ARGUS performance exceeds what is shown in Figure 12, because the

technology used in Facelt ARGUS has been optimized to take advantage of the data that real-world video applications provide.

Tuning to Different Populations is Not Required. We disagree with the notion that algorithms should be tuned to data sets, as discussed on page 40. Our core Facelt matching algorithm does not contain any information about race, age or gender, thus the technology used in the large identification test has no parameters to tune. Optimization is performed dynamically in the context of normalization (discussed in detail in the Appendix A of the main report).

Normalization Does Improve Performance for One-to-Many Applications. The benefit of normalization for identification was not fully tested in the present evaluation because the type of normalization that was allowed in this test (see Appendix A of main report) was restricted to routines that do not change the rank order of the one-to-many search result. The FRVT 2002 test givers did not allow a more appropriate normalization for identification.

Hardware and Cost Variations Between Test System and Real Systems

Use of Identix facial technology does not require the high cost hardware that was used in this test.

Our one-to-many product, Facelt ABIS (Automated Biometric Information System), utilizes a set of separate, low cost units to perform face finding, template creation, load balancing and matching. For the present tests, more than the maximum allowed number of these mini-computers would have been required. Note that Facelt ABIS is a modern three-tier http/XML system that scales with database size, number of searches required per unit time, distance between hardware components and number of users. Again, Facelt ABIS relies on the low-cost single CPU components mentioned above, which we refer to as Biometric Appliances.

For large-scale identification systems, the total cost of ownership includes many business factors other than hardware cost. The costs underlying an AFIS (fingerprint) system are an example of this fact.

Why Choose Identix?

Other than the obvious requirement for superior technology, those contemplating a choice between vendors should also consider the following important factors:

Business factors – Identix is US-based; it is a large, stable, publicly traded organization, arguably the only facial company in that class. Identix is a multi-biometric company with over 20 years of real world experience, established partnerships with a large number of systems integrators, card manufacturers and other biometric solution providers.. We have worldwide facial recognition and fingerprint biometric deployments for national ID, drivers' license, voter registration, background checking, physical and logical access control, social security payments and airport security.

Technical factors – Identix is technically active: we have a strong R&D group, an active presence in standards organizations, and we are the inventors of algorithms and thus dedicated to their continued improvement, as opposed to having licensed technology which is static and now out of date.

Identix offers core technology and products for both facial and fingerprint matching technologies, of key importance for those considering a solution that will rely on fusion of the two.

IDENTIX INCORPORATED

Product factors – Identix provides modern implementations that take advantage of the latest technologies: mature finished off-the-shelf products and software developer kits. Also, Identix has a Professional Services organization, which can be of tremendous value in performing tasks leading up to real world implementation, including development of customized software applications or finished systems, provision of benchmarking services and acceptance testing, and data analysis services.

Identix offers facial solutions for verification, one-to-many and surveillance in the form of both finished products and SDKs.

For facial one-to-many, we offer:

- Facelt ABIS, a cutting-edge scalable facial image search system akin to an AFIS that delivers speed, accuracy and throughput based on customer requirements.
- IBIS, a wireless, mobile solution for finger and face ID in the field; ideal for law enforcement and border patrol.
- Facelt Identification SDK – COM-based Win32 SDK for building one-to-many search applications.

For facial surveillance (watch list applications), we offer:

- Facelt ARGUS, an off-the-shelf facial surveillance and identification system that provides watch list and alarm capabilities for enhancing new and existing CCTV systems. Facelt ARGUS scales to any number of cameras and can be set up remote from the location under surveillance.
- Facelt MultiFace SDK – add-on to Identification SDK for building watch list applications: enables capture of faces from live video to input to search engine.

For facial one-to-one, we offer:

- Facelt Verification SDK – Win32 SDK for building access control applications; uses COM and ActiveX

Final Comments

Identix is pleased with the results of FRVT2002. In particular, we are happy with the progress that we have made since the last evaluation and with the validation of facial for large-scale one-to-many applications.

Identix applauds the momentous effort by the evaluating body. We feel that the tests were well organized and that the results are expertly presented. We value our continued strong relationship with the agencies and personnel involved in carrying out this evaluation.

3/9/2003 2:50 PM

Imagis welcomed the opportunity to participate in the Face Recognition Vendor Test 2002 and appreciates the efforts made by the administrators to provide a uniform academic laboratory test environment for face recognition accuracy assessment.

Last summer's test -- Imagis' first experience with the FRVT -- provided a snapshot of the state of the industry at that time. The test was particularly useful to Imagis in that it suggested to us a number of ways in which we could achieve significant improvements to our then-current FR algorithms. In addition, it provided us with valuable insights into the problem of balancing the need to achieve acceptable levels of accuracy while at the same time preserving our technology's intrinsic ability to search large databases quickly -- a factor that is often cited by our customers as a critical requirement in real world implementations of FR applications.

The release of ID-2000 Version 9.0 in the fall of 2002 reflected these improvements. This version not only achieved a substantial increase in accuracy by correcting certain limitations identified during the internal benchmark procedures that preceded the FRVT test itself, it also proved that significant accuracy advancements are achievable on an ongoing basis as a result of the inherently extensible nature of our core imaging algorithms. These improvements have been, and will continue to be, implemented in upcoming software releases. For example, major users around the world have already acclaimed the level of real-world performance achieved by the post-FRVT release of the Imagis' ID-2000 Version 9.0. Imagis has also recently demonstrated a 64-bit version of ID-2000, further enhancing the technology's ability to search large databases (i.e. more than two million images) without impeding the speed requirements demanded by our customers. Imagis is the only face recognition vendor that supports a 64-bit environment.

For FRVT 2002, Imagis used customer-ready software with an encode array size of 1,400 bytes, operating on a single processor. This still achieved overall performance levels that compared favorably with other vendors, some of whom required multi-processor systems and encode arrays exceeding 4,000 bytes. A forthcoming release of the software will allow the user to choose larger encode sizes suitable for environments in which encode time and match speed are less important than accuracy. A true symmetric multiple processor (SMP) version of the Imagis' ID-2000 application will be available by mid-year 2003.

We invite all interested parties to evaluate our Version 9.0 release. While laboratory testing provides a good basis for discussion, nothing can substitute for a real-world test deployment using one's own data and imagery. In such an environment, users can measure the impact that speed and database size have on the operational processes of their own organization. To discuss a pilot deployment, please visit www.imagistechnologies.com or call 1.604.684.2449.

Viisage Technology Position Paper on the FRVT 2002 Report

The FRVT 2002 represents a significant advance in the testing of facial recognition technology. As such, the FRVT 2002 team is to be commended for how they conducted the test, the in-depth report they generated, and the useful conclusions and trends they identified.

We participated in the FRVT 2002 using a commercial version of our FaceEXPLORER application. It was only customized to handle the inputs and outputs as dictated by the FRVT test specifications. No technology was customized for this test.

After reviewing our performance results, we were greatly surprised by the relative performance of Viisage. We do not believe that these test results represent the performance of our applications. The results our many customers obtain do not match with what the curves would otherwise suggest. Without a more detailed analysis of our particular results, we find it hard to reconcile these differences. However, in light of some of the figures in the report, there is troubling evidence that something was wrong with our test setup, either due to human error or a problem with the software.

We will continue to investigate what went wrong with our test software and hope to be able to offer a more informative appraisal when completed.

The primary evidence is Figures 24 and 30 in the *“FRVT 2002: Evaluation Report”* and the supporting Figures L.42 and L.43 from Section L.5, MCInt Still-Video Experiment, in the accompanying *“FRVT 2002: Technical Appendices.”* These figures compare the performance between still imagery and video for “different day” signatures with a gallery of 63 signatures, one per person, under both verification and identification scenarios. In Figure L.42, there is no verification curve for Viisage. In Figure L.43, the identification curve is a strange step-shaped function.

According to the FRVT evaluation team, these anomalies were both caused by our software generating the same score for all matches in this particular test. Clearly, this indicates a problem with our system. These results also surprised the FRVT evaluation team as the *“FRVT 2002: Evaluation Report”* says they manually checked the Viisage similarity files for confirmation. We would like to thank the FRVT evaluation team for including this manual check.

Unfortunately, at this time, we are unable to determine the source of this problem and whether or not it could have affected the results from the other tests. Given that the images are presented to the system in a presumably randomized order, we must assume that it could have affected the results of the other tests in some unknown way.

We therefore must conclude that all the results for Viisage are in question and do not represent the performance of our software.



Suite 231, Building M-50 IPF, 1200 Montreal Road, Ottawa, ON, Canada – K1A 0R6

FRVT 2002

The FRVT 2002 was managed very professionally from beginning to end by a dedicated team of people who showed competence, flexibility, and fairness in all their interactions with the participants. It was certainly a very important milestone in the field of face recognition for the raw size of the test database used, the number of participants, and the careful selection of test sets to allow for the evaluation of the sensitivity of the algorithms to various parameters such as aging and gender.

Indirectly, FRVT also tested the speed of the systems: for the high computational test, everyone had to process about 15 billion face comparisons within 264 hours, not including the time to detect the faces and code them. This represents more than 15,000 comparisons per second.

In order to complete the FRVT in the allotted time, we had to use a custom solution where we used our best comparison algorithm only a certain percentage of the time, using another, less efficient but faster, algorithm to pre-sort the pictures. The results obtained do not completely represent the full power of this algorithm when used in verification applications where more time can be used for the match.

Since FRVT, many improvements have been implemented to our face recognition algorithms to be able to deal with a larger volume of data. The comparison speed has been increased to be about 15 times faster than what was used in the competition. The raw recognition performance has also been improved significantly in the process. A recent test ran on an in-house face database of 5,800 pictures shows an increase of the recognition performance (FAR=0.1%), 81% over the software used in FRVT. We are confident this new technology would now perform very well under strict time constraints as imposed by the competition.

Static Images & Live Image Verification

Our system was originally designed to perform a one-to-one verification in which verification speed is not crucial. A verification application that takes up to one second for the face- verification part of the verification process is usually acceptable to a user. Once scrubbed, a database of facial images that does not contain multiples of duplicate images is used in numerous applications such as passport or border control to authenticate the user standing before a facial authentication device. However, there are several factors that adversely affect the performance of a verification system that are not applicable to a static image test such as the FRVT. One of those factors is the effect of ambient illumination found on a live human face. In a static image test such as that of the FRVT, illumination stayed constant in the pairs of images tested. However, this is not the case in real life situations. To cope with this problem, VisionSphere has developed and integrated into UnMask, controls that adapt the Company's camera's sensors, to varied ambient illumination found on a face. The camera sensors are embedded into the Company's several facial authentication devices.

This technology permits our cameras, embedded in our facial authentication devices, FaceCam, and our USB desktop camera, to automatically adjust their sensor settings to obtain a good image of the face, independently of the background. This process significantly improves the recognition performance in certain lighting conditions.

In a pilot project conducted on behalf of Transport Canada, in the summer of 2002, at the Ottawa International Airport, our software integrated with, and used with a FaceCam, was deployed to control door access to a restricted area for 170 airport employees over a 7-week period. Access was granted successfully by our integrated solution 96.4% of the times and not a single instance of false acceptance was reported in approximately 2,700 authentications over a seven-week period. A liveliness test (eye blink detection) was also used to confirm that the employee granted access was a human and not a photograph of that employee.



Conclusion

In spite of the excellent organization of the tests, FRVT did not evaluate the ability of the systems to deal with live verification under varied lighting conditions to verify an average cooperative user under constraints other than raw number of matches per second. The FRVT test compares static images to static images. The images are compared to locate duplicate or multiple images, which for certain applications, results in the removal or location of duplicate or multiple images of the same person. After the removal of duplicate or multiple images, the facial-image database can be considered clean.

In a real world situation, the clean database would be used to authenticate/verify a live person in varied ambient illumination, against his/her facial image template stored in the clean facial-image template database. The FRVT did not, however, take that into consideration, i.e., it did not authenticate/verify users in real-world conditions (varied ambient illumination).

When face-recognition software is rated by tests such as FRVT, consideration should be given to authenticating/verifying a live person against his/her facial image templates stored in the clean database. A static image "ONLY" test such as the FRVT does not truly reflect how face recognition technology will perform in real world situations.

March 10, 2003

Appendix O

Additional Information on the 3D Morphable Model Used in FRVT 2002

Generating Frontal Views from Single, Non-Frontal Images

Volker Blanz¹, Thomas Vetter²

¹ Max-Planck-Institut für Informatik,
Stuhlsatzenhausweg 85, 66123 Saarbrücken, Germany

² Universität Basel, Departement Informatik,
Bernoullistr. 16, 4056 Basel, Switzerland

Abstract

This paper describes a method for generating frontal views of faces, given single images at arbitrary poses and illuminations. The system is based on a Morphable Model of 3D faces, which represents face-specific information extracted from a dataset of 3D scans. It estimates the 3D shape of novel faces from the input images, and completes surface regions that are not visible in the image, due to self occlusion. The reconstructed 3D surface covers the face, the ears, and the upper half of the neck.

To generate the frontal views of the reconstructed faces at a standard position, pose, and size, the system uses 3D computer graphics. In the frontal image, the algorithm uses standard lighting conditions, after having compensated the effects of illumination in the original image. The novel drawing parameters are extracted from an example image. This image also provides standard hair, shoulders, and background, which are added to all transformed frontal views.

1. Introduction

Many face recognition algorithms perform best if the gallery and probe images are frontal views. In some applications, however, image data are taken at oblique viewing directions, such as images from cameras on ATMs. Often, authorities have not more than one random snapshot of a suspect.

To achieve robust recognition across a wider range of viewpoints, we propose a system that computes frontal views from arbitrary images. The result of this non-trivial preprocessing step can then be used by state-of-the-art face recognition systems.

Clearly, the transformed frontal views can only be estimates of the true appearance of the person: The shading information in an image, in general, is not sufficient to recover 3D depth in a unique way. Moreover, shading effects in images of faces are difficult to distinguish from variations in skin complexion, for example in the region around the eyes. Finally, parts of the face may be invisible in the input image, due to self-occlusion.

Our strategy to estimate the faces' frontal appearance, therefore, relies heavily on class-specific knowledge about human faces. This information is derived in an automated way from a dataset of textured 3D scans of faces. 3D shape of faces is

reconstructed by finding the combination of example faces that is as similar as possible to the input image. We can then redraw the reconstructed shape in a frontal view, using 3D computer graphics.

In the following sections, we briefly summarize the Morphable Model and the algorithm for 3D shape reconstruction. In Section 4, we describe the procedure that was applied to the test images, and show a set of examples.

2. Morphable Model of 3D Faces

The Morphable Model of 3D faces [Vetter and Poggio, 1997, Blanz and Vetter, 1999] is a vector space of 3D shapes and textures (surface colors), spanned by a set of examples, and capturing the variations found within this set. It is derived from a dataset of 200 textured *Cyberware* (TM) laser scans. The set is balanced with respect to male and female faces. The persons are aged between 18 and 45 years, and except for one Asian female, all are Caucasian. Previous work [Blanz et al., 2002] indicates that the model may well be applied to reconstruct 3D shape from images of a wider ethnic variety.

The shape and texture vectors are defined such that any linear combination of examples

$$\mathbf{S} = \sum_{i=1}^m a_i \mathbf{S}_i, \quad \mathbf{T} = \sum_{i=1}^m b_i \mathbf{T}_i. \quad (1)$$

is a realistic face, given that \mathbf{S} , \mathbf{T} are within a few standard deviations from their averages. Each vector \mathbf{S}_i is the 3D shape, stored in terms of x, y, z -coordinates of all vertices $k \in \{1, \dots, n\}$ of a high-resolution 3D mesh:

$$\mathbf{S}_i = (x_1, y_1, z_1, x_2, \dots, x_n, y_n, z_n)^T. \quad (2)$$

In the same way, we form texture vectors from the red, green, and blue values of all vertices' surface colors:

$$\mathbf{T}_i = (R_1, G_1, B_1, R_2, \dots, R_n, G_n, B_n)^T. \quad (3)$$

In the conversion of the laser scans into shape and texture vectors \mathbf{S}_i , \mathbf{T}_i , it is essential to establish dense point-to-point correspondence of all scans with a reference scan, to make sure that vector dimensions in \mathbf{S} , \mathbf{T} describe the same point, such as the tip of the nose, in all faces. Dense correspondence is computed automatically with an algorithm derived from optical flow (for details, see [Blanz and Vetter, 1999]). Finally, we perform a Principal Component Analysis (see [Duda et al., 2001]) to estimate the probability distribution of faces around their average, and we replace the basis vectors \mathbf{S}_i , \mathbf{T}_i in Equation (1) by an orthogonal set of eigenvectors.

2.1 Fitting the Morphable Model to Images

In order to reconstruct a 3D face from an image, we fit the Morphable Model to the input image in an analysis-by-synthesis loop: At each iteration, the current model parameters, together with 3D computer graphics techniques, define a model image. Let $I_{r,model}(x, y)$, $I_{g,model}(x, y)$, $I_{b,model}(x, y)$ be the red, green, and blue channels of this image. The goal of the fitting algorithm is to make it as similar as possible to the input image: In a gradient descent, the algorithm minimizes the image difference

$$E_I = \sum_x \sum_y \sum_{c \in \{r,g,b\}} (I_{c,input}(x, y) - I_{c,model}(x, y))^2 \quad (4)$$

with respect to the following model parameters: shape coefficients, texture coefficients, 3D position, 3D angle, focal length, red, green and blue components of ambient and parallel light, direction of that parallel light, and color offsets, gains, and contrast. For details, see [Blanz and Vetter, 1999].

In the current version, the system is initialized by providing 2D image positions of some feature points, such as the tip of the nose, or the corners of the eyes. The algorithm then converges automatically, and produces the following results: shape and texture coefficients, 3D shape and texture \mathbf{S} , \mathbf{T} , pose parameters, such as head position, orientation, and illumination conditions.

The fact that parameters for pose and illumination are recovered from the image, and that they are represented separately and explicitly just as in computer graphics, implies that we can modify any of them independently, and draw the face from any new angle and under any new illumination.

Since the linear combination of textures \mathbf{T}_i cannot reproduce all local characteristics of the novel face, such as moles or scars, which in fact may be highly relevant for recognition, we extract the person’s true texture from the image wherever it is visible. This is done by an illumination-corrected texture extraction algorithm [Blanz and Vetter, 1999]. The boundary between the extracted texture and the predicted regions (where the true texture is occluded in the input image) was still visible in some of the images used in this test. We have meanwhile improved our algorithm to achieve a smooth transition, based on a reliability criterion for texture extraction.

None of the 3D faces that form the Morphable Model has facial hair. Still, texture extraction captures the beard in a texture map, and reproduces it on the frontal view, even if the beard’s thickness is neglected in this transformation. In the same way, eyeglasses are mapped on the facial surface by the algorithm. Recently, we have extended the model to handle open-mouth facial expressions, but this was not yet applied to the FRVT02 data.

4. Standard Hair Style and Shoulders

Since our model covers neither hair, nor shoulders, nor the scene behind the head, we cannot produce an estimate of these image regions in the transformed frontal view. As a result, we could leave the image blank around the face, ears, and top of the neck. However, these images would be quite different from the type of images that most face recognition systems are designed for, so the programs might at least require readjustment. In particular, this would be true if transformed and real frontal view images occurred in the same test.

Therefore, we decided to render the reconstructed faces into a standard frontal view of one person, so all transformed images have the same hair-style, shoulders, and background (Figure 1).

As an alternative, we could render the faces into their individual frontal view images, which are part of the database. However, this would not reflect the realistic face recognition scenario, where nothing more than the oblique view is available. In fact, if we rendered the faces into the frontal views stored in the gallery, a simple comparison of some pixels around the faces would give away the identities in a trivial way.

As a standard frontal view, we selected an image with short, dark hair, that looks realistic on most faces. With additional gender recognition, which may be based on the 3D face reconstruction, it would be possible to select different standard images for different genders (or ethnic groups).



Figure 1: *From the original images (top row), we recover 3D shape (second row), by finding an optimal linear combination of example faces. Mapping the texture of visible face regions on the surface and rendering it into a standard background, which is a face image that we selected, produces transformed front views (bottom row). Note that the frontal-to-frontal mapping, which served as a baseline test, involves hairstyle replacement (bottom row, center).*

In order to determine head size, position, orientation and illumination of the standard image, we fitted the Morphable Model to the image, as described in the previous section. To be able to draw hair in front of the forehead, we manually defined a transparency map that is opaque on the hair, and transparent everywhere else, for the standard image.

Given a novel input image, we perform the following steps:

1. Manually define feature points, such as the tip of the nose, the corners of the eyes, or any points along the occluding contours on the cheeks. For FRVT02, we defined an average number of 11 feature points to ensure optimum quality. However, 6 points are often sufficient. In future systems, these points can be found by an automated feature detection algorithm.
2. Run the optimization (4 minutes on a Pentium 4 at 2GHz).
3. Render the 3D face in front of the standard image, using the parameters for position, orientation, size, and illumination of the standard image.
4. Draw the hair in front of the forehead, using the transparency map.

5. Conclusions

We have presented a method to handle images taken at arbitrary viewpoints with state-of-the-art face recognition systems. The combination of our reconstruction algorithm with other recognition systems may reduce their sensitivity to pose angle considerably. In the current state, our method requires manual labelling of a small number of feature points in the image. We anticipate that this labelling is no major problem in many applications, for example if a relatively small number of suspects' images have to be added to a gallery that is used for search. Combined with a feature detection algorithm, our system will be fully automated.

Acknowledgements Philipp Georg has labelled the feature points in the input images. The 3D face scans were scanned by Niko Troje at Max-Planck-Institut für Biologische Kybernetik, Tübingen, Germany, in the department of Heinrich Bülhoff.

References

- [Banz et al., 2002] Blanz, V., Romdhani, S., and Vetter, T. (2002). Face identification across different poses and illuminations with a 3d morphable model. In *Proc. of the 5th Int. Conf. on Automatic Face and Gesture Recognition*, pages 202–207.
- [Banz and Vetter, 1999] Blanz, V. and Vetter, T. (1999). A morphable model for the synthesis of 3d faces. In *Computer Graphics Proc. SIGGRAPH'99*, pages 187–194, Los Angeles.
- [Duda et al., 2001] Duda, R., Hart, P., and Stork, D. (2001). *Pattern Classification*. John Wiley & Sons, New York, 2nd edition.
- [Vetter and Poggio, 1997] Vetter, T. and Poggio, T. (1997). Linear object classes and image synthesis from a single example image. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(7):733–742.