Studies of One-to-One Fingerprint Matching with Vendor SDK Matchers

NISTIR 7221

April 22, 2005

Craig Watson, Charles Wilson,
Karen Marshall, Mike Indovina, &
Rob Snelick

Studies of One-to-One Fingerprint Matching with Vendor SDK Matchers

Craig Watson, Charles Wilson, Karen Marshall, Mike Indovina, Rob Snelick
National Institute of Standards and Technology

Abstract

NIST has conducted testing of one-to-one SDK (Software Development Kit) based COTS (Commercial Off-The-Shelf) fingerprint matching systems to evaluate the accuracy of one-to-one matching used in the US-VISIT program. Fingerprint matching systems from eleven vendors not used in US-VISIT were also evaluated to insure that the accuracy of the matcher tested was comparable to the most accurate available COTS products. The SDK based matching application was tested on 20 different single finger data sets of varying difficulty. The average true accept rate (TAR) at a false accept rate (FAR) of 0.01% was better than 98% for the two most accurate systems while the worst TAR at a FAR of 0.01% was greater than 94%. The data sets used and the ranking of the systems are discussed in detail in the report. A copy of this report and appendices are available at **http://fingerprint.nist.gov/SDK**.

Introduction

On February 4, 2003 a report titled "Use of Technology Standards and Interoperable Databases With Machine-Readable, Tamper-Resistant Travel Documents" [1] was submitted to the Congress jointly by the Attorney General, Secretary of State, Department of Homeland Security, and NIST . (This report is informally referred to as the 303A Report and was mandated by [2] and [3]). It discusses measurements of the accuracy of both face and fingerprints as they relate to U.S. border entry and exit. This study is part of the work undertaken for the US-VISIT program to measure the one-to-one accuracy of fingerprint matching systems used for one-to-one verification. Eight different one-to-one matching systems, including the one used by US-VISIT, are compared in this report.

The results of the one-to-one fingerprint-matching tests performed at NIST using vendor supplied SDK libraries are discussed in this report. These tests were designed to evaluate the current one-to-one matching technology for the US-VISIT program and to allow the comparison of vendors not in the US-VISIT program to the US-VISIT vendor. Initially NIST performed one-to-one matching tests using the verification test bed (VTB) bozorth98 matcher [4]. The results were promising but this matcher is not state of the art technology. To get a good assessment of the current capabilities of COTS matchers for US-VISIT, NIST was asked to test performance of current commercial products. The main result obtained from this testing was an estimate of how well commercial products performed one-to-one matching for verification over a wide range of fingerprint image qualities. The relative accuracy of thumbs and index fingers was also investigated.

SDKs were requested from several vendors including all highly ranked vendors in the FpVTE fingerprint test [5]. Currently NIST has results from twelve different vendors including the vendor currently use in the US-VISIT system (some vendors submitted multiple SDKs). The SDK testing is similar to the medium scale test (MST) performed during the fingerprint vendor technology evaluation (FpVTE) [5]. The FpVTE tests used data that was randomly selected from the same larger dataset for DOS-C, DHS2-C, BEN, and OHIO. These different datasets are discussed later and in reference [4].

There are two key differences between the SDK tests and FpVTE. First, the SDK tests were performed on larger samples of each of these datasets. Second, the FpVTE testing gave each vendor a set of N images and asked them to return a matrix of NxN matched scores. The program to perform the testing in FpVTE was written by each vendor and run on the vendors own computer hardware. In SDK testing, each call to the matcher function compares only two fingers at a time. The application to perform these matches was written by NIST using the SDK libraries supplied by each vendor and run on NIST computer hardware.

SDK Requirements

Each fingerprint biometric vendor was requested to deliver an SDK library, as specified below, in object-code format and appropriate for large-scale verification testing (i.e. one-to-one matching) in a Windows 2000 PC-based environment. The specifications called for this library, or set of libraries, to implement the core functionality necessary for verification testing; namely feature extraction and matching operations, with specified time limits. The matching operation was specified to return a (non-normalized) similarity score only.

After the initial contacts, a series of questions were often posed by the vendors regarding such things as the type fingerprint images to be tested, how the images were acquired (sensors), timing constraints, etc. These questions had to be addressed before the libraries could be obtained. Often the initial interaction would only yield the "stock" vendor SDK, which wasn't appropriate for large-scale batch verification testing due to limitations such as the inclusion of GUI components. Once the appropriate technical contacts within each vendor's organization were established, the proper libraries were delivered.

Without exception, these libraries were delivered as a set of one or more Dynamic Link Libraries (DLLs), along with a set of one or more C language header files. Often, the vendors provided example code written in C or C++ to demonstrate calling their library. These example codes proved useful in filling in gaps in the vendor provided documentation, especially in determining things such as library initialization procedures, match comparison parameter order (in the cases of asymmetric matchers), etc.

Upon receiving the SDK library distribution from the vendor, the software and documentation were installed on a single PC dedicated to the task of integrating the library and the NIST verification test program. In order to test for proper integration, a set of 20 test fingerprint images (from the NIST public fingerprint database) were employed to produce a 100 score similarity matrix. These images were then sent to the vendor via email, with instructions on generating a (hopefully identical) score matrix, to be emailed back to NIST for comparison. If the matrices were not in agreement, NIST contacted the vendors and resolved the problem so that proper integration was assured.

Most technical problems with individual vendor SDKs were encountered and resolved during the integration phase on the dedicated integration PC. It should be noted that some problems were encountered during the initial test runs on the SDK testing platform (for example, abnormal termination, failures to enroll or match, etc.) which required further vendor interactions to resolve the problems.

Testing

After the vendor supplied NIST with the software and initial support to make sure it was being used correctly, all testing was performed by NIST staff on NIST owned computers. The systems used for testing were five dual-processor (3 GHz) PCs running windows 2000[1]. This allowed the SDK to run on several different datasets in parallel. All results were scored with the software package used in the VTB report [4].

A key difference between the SDK test and the FPVTE MST test is that the SDK test required the vendors to accept only two input fingerprint templates and return a matched score for those two templates. In FPVTE, vendors were given a set of N images and asked to return a matrix of NxN matched scores.

Evaluation Data

The fingerprint datasets used to evaluate the one-to-one matchers included rolled and plain fingerprints from inked paper and live-scan devices. The data is from several sources: US VISIT (VISIT_POE, VISIT_POE_BVA), Department of State (DOS-C), Department of Homeland Security (DHS2-C, DHS10, and BEN), Texas Department of Public Safety (TXDPS), Ohio web-check (OHIO), and NIST Special Database 29 (SD29, the only database available to the public). Except for VISIT, OHIO, and BEN, which were not available for the original VTB testing, a more detailed description of these datasets can be found in the VTB report [4]. Similar descriptions are included in this document for VISIT, OHIO, and BEN datasets.

A random sample of 6,000 subjects were selected from the larger datasets (VISIT_POE, VISIT_POE_BVA, DOS-C, DHS2-C, DHS10, BEN, TXDPS) and all the subjects that did not have segmentation problems were used in the smaller datasets (OHIO-885, SD29-180). The VISIT dataset used to different test on used two different captures at the Point of Entry (POE) as the probe and gallery (VISIT_POE) and the other used an instance from POE as the probe and an instance from Bio-VISA Application (BVA) as the gallery (VISIT_POE_BVA). The datasets were consolidated to find previously unknown mates and presumed mates that were not really mates. Consolidation used the output of various matches to help identify potential errors and then human experts examined those fingerprints to make a final determination. The human experts are fingerprint examiners with over twenty years of experience. Since DHS2-C was consolidated after testing had started, the final dataset size after consolidation was only 5,888 not 6,000. The four finger slaps in BEN, DHS10, TXDPS, OHIO, and SD29 were all segmented with the NIST segmentor as discussed in the VTB report [4].

[1] Specific hardware and software products identified in this report do not imply recommendation or endorsement by the National Institute of Standards and Technology. It was necessary to study and report on these products as they were being used at the time in US-VISIT and other relevant programs

VISIT, DOS-C and DHS2-C contained live-scan plain impressions of the left and right index fingers. They were captured in an operational environment and should give results of what to expect from real time data of plain to plain (p2p) impression matching.

DHS10 and TXDPS were mainly inked paper segmented plain and rolled impressions and BEN was mainly live-scan segmented plain and rolled impressions. These three datasets are most useful in evaluating plain to rolled (p2r) impression matching. DHS10 and TXDPS are examples of the currently available fingerprint databases and BEN representing data captured with newer live-scan methods with some control on quality.

OHIO contained three sets of live-scan segmented plain impressions and one set each of inked paper and live-scan rolled impressions and was captured in a controlled environment. The inked paper rolled fingerprint data was not used in the SDK testing. This live-scan dataset is considered high quality data as it was captured in a controlled environment with every effort made to produce good quality fingerprints. OHIO should show peak performance of the matchers for plain to plain and plain to rolled impressions (p2p and p2r).

SD29 contained ten-print card mates of inked paper segmented plain and rolled impressions (p2p, r2r, p2r). The dataset is from legacy inked paper data.

The datasets are grouped by finger position, so only right index are compared to right index and so on for other fingers. Not all fingers were tested for DHS10, BEN, and TXDPS as the results for US VISIT were most meaningful for the thumb and index fingers. However, since OHIO and SD29 are smaller datasets it was possible to make some evaluation on other fingers with those datasets. Considering all the datasets and fingers, the total number of matches performed by each SDK was 758,638,238.

DHS Benefits Data (BEN)

Description

Data from the Department of Homeland Securities Benefits program.

Number of Subjects	**Instances per Subject**
~80,000 (of 100,000) 20,000 Background (rolled impressions where four-finger plain segmentation failed)	One 10-print card per person
Impression Type	**Finger Positions Captured**
Live-scan Rolled and Plain	10 fingers segmented from rolled impressions and 10 fingers segmented from plain impressions on the same 10-print card
Capture Device(s)	**Availability**
DBI 1133S	Government use only

Data Preparation

Segmentation of rolled impressions from the 10-print card was done prior to receipt by NIST.

For plain impressions, only successful automatic segmentation results were used. No manual correction of segmentation results was performed, so a very small number of bad-segmented results may be included. Automatic segmentation resulted in approximately 80% yield across all available 10-print cards.

All cards used in the SDK report were consolidated using matchers and human examiners.

Ohio Web Check Data (OHIO)

Description

Data from the Ohio WebCheck Program

Number of Subjects	**Instances per Subject**
925	Three sets of live-scan plain impressions per person. One set each of live-scan and inked rolled impressions.
Impression Type	**Finger Position Captured**
Live-Scan and Inked Rolled and Plain	10 fingers segmented from rolled impressions and 10 fingers from plain impressions
Capture Device(s)	**Availability**
Identix TP600 & TP2000 Cross Match 442 Smiths-Heimann LS2 Check	Government use only

Data Preparation

Segmentation of the rolled impressions was done prior to receipt by NIST.

Two sets of live-scan plain impressions were segmented before receipt by NIST. The other set was segmented by NIST.

U.S. VISIT Point of Entry Data (VISIT_POE) and Bio-Visa Application (VISIT_BVA)	
Description Data from U.S. VISIT captured from persons entering the U.S. at airport points of entry (POE) and at Consulates when applying for a U.S. VISA (BVA).	
Number of Subjects ~3.5 Million with one POE instance ~715K with 2 or more POE instances ~290K with POE and BVA instances	**Instances per Subject** One to many cases per person, with each case containing one right index finger impression and one left index finger impression.
Impression Type Live-scan Plain	**Finger Positions Captured** Right and Left Index
Capture Device(s) Cross Match 300A (POE) Smiths-Heimann ACC01394 (BVA)	**Availability** Government use only
Data Preparation All cases used in the report were consolidated using matchers and human examiners.	

Results

The number of ROC curves and tables for analyzing the ROC curves was too large to put in the body of this report so they are included in appendices. Appendix A has the rankings of each dataset by SDK and Appendix B has the ranking of each SDK by dataset. Appendix C contains the ROC curves for each SDK. Appendix D contains the ROC curves for each dataset. Appendix E has tables for each SDK with values of TAR at a FAR of 0.01% and 1%, and FAR at a TAR of 98%. Appendix F has tables with the same values as appendix E but sorted by dataset.

The tables in appendices E and F are useful for comparing the SDKs at the three different performance points on the ROC curve. Linear interpolation (log on the x-axis) was used to determine the readings at the set value points of FAR 0.01%, FAR 1%, and TAR 98%.

The ROCs show that the most accurate SDKs were consistent across all datasets while others had a wide range of results depending on which dataset was being used. This effect was also seen in the FpVTE tests [5]. The SDKs were ranked based on their average TAR for the different datasets at a FAR of 0.01%. The DOS-C, DHS2-C, VISIT_POE, and VISIT_POE_BVA datasets were grouped for ranking performance on plain to plain fingerprints and BEN, DHS10, and TXDPS were grouped for ranking on plain to rolled fingerprints. All OHIO results were grouped together and all SD29 results were grouped together.

Table1 shows the SDK ordering for each of the data groupings and tables 2-6 show the average, minimum, and maximum TAR values for the SDKs in each data group at a FAR of 0.01%. The highest ranked SDKs were R (99.01%), H (98.97%), I (98.73%), O (98.61%), and Q (98.20%). These SDKs were mostly ranked 1-5 and performed consistently well across all the datasets with a difference between their minimum and maximum TAR scores of 4.45% (R), 4.4% (H), 5.71% (I), 5.16% (O), and 8.14% (Q). SDKs P (97.72%), J (97.54%), F (97.53%), G (96.92%) and D (96.37%) were a close second tending to share the 6-10 rankings. The performance variation across the datasets started to increase in this group P (8.25%), J (8.05%), F (9.03%), G (9.82%), and D (10.91%). SDKs C (95.16% TAR, 16.43% min./max. TAR difference), N (94.78%, 10.92%), and K (94.18%, 13.75%) shared the 11-13 rankings and L (91.43%, 20.26%) ranked 14th. Next came SDK A (85.66%, 47.68%) and the VTB (84.9%, 29.44%), then SDKs B (80.81%, 33.11%), E (76.08%, 53.49%) and M (56.1%, 80.61%). A point of interest in these rankings is that there are only seven unique vendors present in SDKs R, H, I, O, Q, P, J, F, G, D, C, N, K, and L and five unique vendors in SDKs A, VTB, B, E, and M.

Not surprisingly, the datasets perceived to have better quality (i.e. BEN and OHIO) performed better than the other datasets. Of the three p2r (plain to rolled) datasets (BEN, DHS10, and TXDPS), BEN right thumb (RT) and left thumb (LT) gave the highest TAR scores for all the SDKs. The average TAR (appendix B) over all the SDKs for each dataset also shows that BEN data (97.68% RT, 96.43% LT, 91.3% RI, 88.29% LI) performs better than DHS10 (89.79%, 88.85%, 84.75%, 82.77%) and TXDPS (91.18%, 90.30%, 90.89%, 89.19%). The difference in performance was not as great for the index fingers; which may be attributed to the fact that the segmented plains had larger variations in rotation that could have affected the performance of some matchers.

OHIO data was considered the highest quality data available in the test and all the SDKs had their most accurate performance on this data. Obviously the most accurate SDKs had little room to improve but the improvement in the less accurate SDKs was significant as shown by comparing the average TAR of all the SDKs for different datasets and fingers. For example the right index (RI) fingers for DOS-C 92.87%, DHS2-C 83.80%, VISIT_POE 93.5%, VISIT_POE_BVA 92.94%, BEN 91.3%, DHS10 83.75%, and TXDPS 90.89% compared to OHIO 94.82%.

Looking at the performance of thumbs vs. index in Appendix B (SDKs ranked by datasets), the thumbs in datasets BEN, DHS10 and TXDPS generally have higher accuracy than index fingers and right fingers on average had higher accuracy than left fingers. The average TAR performance for RT was 97.68% (BEN), 89.79% (DHS10), and 91.18% (TXDPS) and for LT was 96.43%, 88.85%, and 90.3%. For the corresponding RI fingers 91.3%, 83.75%, and 90.789% and left index (LI) was 88.29%, 82.77%, and 89.19%. Right index were more accurate than left in the DOS-C (92.87% RI, 83.8% LI), slightly more accurate in VISIT_POE (93.5%, 91.33%) and VISIT_POE_BVA (92.94%, 89.76%), but about the same in DHS2-C (83.8%, 83.23%). In the OHIO data, where image quality was very good, left and right thumbs were better than index fingers but this difference appears to be caused by SDK M as it performed ok on thumbs but horribly on index fingers. The average TARs for OHIO were 98.27% RT, 97.54% LT, 94.82% RI, 94.91% LI.

The ROC curves for the OHIO datasets (Appendix C) and OHIO TARs ranked by SDK (Appendix A) show that plain to plain (p2p) was slightly better than plain to rolled (p2r) on similar high quality data. The results are much closer than previously reported as SDK M performed significantly better on p2r versus p2p. This is also shown when looking at the average TAR performance of OHIO p2p versus p2r: RT (98.27% p2p, 96.98% p2r), RI (94.82%, 95.72%), RM (94.0%, 94.8%), RR (92.25%, 92.07%), RL (88.29%, 87.95%), LT (97.54%, 96.12%), LI (94.82%, 94.91%), LM (92.85%, 93.58%), LL (91.51%, 90.36%), LL (85.47%, 86.74%). Again the majority of improvement is seen in the lower scoring SDKs.

The matcher speeds are included (table 7) to give a relative measure of speed between the different SDKs. The matcher speeds were measured by computing the average number of matches that were performed in 5 minutes. The timing was done on a 3.06GHZ Pentium dual processor PC but with only one matcher job running. The general observation is that as matcher performance increases speed decreases. This is not a surprise since most vendors said if given more time they could always produce more accurate matches.

DOS-C / DHS2-C / VISIT	BEN / DHS10 / TXDPS	OHIO	SD29	ALL
R	R	H	R	R
I	H	R	H	H
H	I	I	O	I
O	O	Q	I	O
Q	J	O	Q	Q
F	Q	F	J	P
P	P	P	P	J
J	F	D	G	F
G	G	J	F	G
D	D	G	D	D
C	N	C	C	C
N	K	N	N	N
K	C	K	K	K
L	L	L	L	L
B	VTB	A	A	A
VTB	A	VTB	VTB	VTB
A	B	B	B	B
E	E	E	E	E
M	M	M	M	M

Table 1: SDK rankings at FAR 0.01% for the different datasets.

SDK	Avg. TAR	Min TAR	Data Set	Max TAR	Dataset
R	99.01	95.55	dhs10_li	100.00	sd29_r2r_rt
H	98.97	95.60	dhs10_ri	100.00	sd29_r2r_rt
I	98.73	94.29	dhs10_li	100.00	sd29_r2r_rt
O	98.61	94.84	dhs10_ri	100.00	sd29_r2r_rt
Q	98.20	91.86	dhs10_li	100.00	sd29_r2r_rt
P	97.72	91.75	dhs10_li	100.00	sd29_r2r_lm
J	97.54	91.95	dhs10_li	100.00	sd29_r2r_li
F	97.53	90.86	dhs10_li	99.89	ohio_p2p_lt
G	96.92	89.91	dhs10_li	99.73	sd29_r2r_lt
D	96.37	88.90	dhs10_li	99.81	ohio_p2p_lt
C	95.16	83.00	txdps_rt	99.43	ohio_p2p_rt
N	94.78	87.97	dhs10_li	98.89	sd29_r2r_rt
K	94.17	86.25	dhs10_li	100.00	sd29_r2r_rt
L	91.27	79.18	dhs10_li	99.44	sd29_r2r_rm
A	85.66	50.05	dhs2-c_li	97.73	ohio_p2p_rt
VTB	84.90	66.76	ohio_p2r_ll	96.20	ben_rt
B	80.81	61.57	txdps_rt	94.68	ohio_p2p_rt
E	76.08	40.33	dhs2-c_ri	93.82	ohio_p2p_rt
M	56.10	12.48	ohio_p2p_ll	93.09	ben_rt

Table 2: SDK rankings for all the datasets combined at FAR 0.01%.

SDK	Avg TAR	Min TAR	Data Set	Max TAR	Dataset
R	98.57	97.64	dos-c_li	99.63	bva_ri
I	98.51	97.17	dos-c_li	99.50	bva_ri
H	98.34	97.31	dhs2-c_li	99.40	bva_ri
O	98.05	96.80	dhs2-c_li	99.21	bva_ri
Q	97.58	96.61	dhs2-c_ri	98.68	bva_ri
F	97.48	95.32	dos-c_li	98.89	bva_ri
P	97.14	95.37	dhs2-c_li	98.69	bva_ri
J	96.72	94.40	dos-c_li	98.22	bva_ri
G	96.40	93.91	dos-c_li	98.18	bva_ri
D	96.05	93.53	dos-c_li	97.99	bva_ri
C	94.81	90.86	dhs2-c_ri	97.63	bva_ri
N	93.35	88.96	dos-c_li	96.18	bva_ri
K	92.12	88.17	dhs2-c_li	95.64	bva_ri
L	85.78	80.13	dhs2-c_li	91.22	bva_ri
B	82.88	74.07	dhs2-c_li	91.03	dos-c_ri
VTB	80.23	75.68	dos-c_li	83.66	poe_ri
A	76.23	50.05	dhs2-c_li	89.44	bva_ri
E	64.85	40.33	dhs2-c_ri	79.18	dos-c_ri
M	52.41	19.84	dhs2-c_li	78.68	poe_ri

Table 3: SDK rankings for DOS-C/DHS2-C/VISIT datasets at FAR 0.01%.

SDK	Avg TAR	Min TAR	Data Set	Max TAR	Dataset
R	98.19	95.55	dhs10_li	99.54	ben_rt
H	98.14	95.60	dhs10_ri	99.65	ben_rt
I	97.69	94.29	dhs10_li	99.52	ben_rt
O	97.56	94.84	dhs10_ri	99.40	ben_rt
J	96.84	91.95	dhs10_li	99.36	ben_rt
Q	96.54	91.86	dhs10_li	99.21	ben_rt
P	96.11	91.75	dhs10_li	99.16	ben_rt
F	96.00	90.86	dhs10_li	99.23	ben_rt
G	95.88	89.91	dhs10_li	98.93	ben_rt
D	94.53	88.90	dhs10_li	98.75	ben_rt
N	93.88	87.97	dhs10_li	98.19	ben_rt
K	92.63	86.25	dhs10_li	98.18	ben_rt
C	91.57	83.00	txdps_rt	98.81	ben_rt
L	90.10	79.18	dhs10_li	97.81	ben_rt
VTB	85.58	73.51	dhs10_li	96.20	ben_rt
A	82.95	74.07	dhs10_li	96.08	ben_rt
B	75.13	61.57	txdps_rt	93.30	ben_rt
E	74.73	57.15	dhs10_li	91.48	ben_rt
M	56.59	33.65	dhs10_li	93.09	ben_rt

Table 4: SDK rankings for BEN/DHS10/TXDPS datasets at FAR 0.01%.

SDK	Avg TAR	Min TAR	Data Set	Max TAR	Dataset
H	99.40	98.68	p2p_ll	99.89	p2p_rt
R	99.39	98.33	p2p_ll	99.89	p2r_rt
I	99.27	97.71	p2p_ll	99.89	p2p_rt
Q	99.20	97.88	p2p_ll	99.77	p2p_rm
O	99.12	97.97	p2r_lr	99.89	p2p_lt
F	98.73	96.03	p2r_ll	99.89	p2p_lt
P	98.66	96.87	p2r_lr	99.60	p2p_ri
D	97.94	94.39	p2r_ll	99.81	p2p_lt
J	97.85	94.25	p2p_ll	99.60	p2p_lt
G	97.51	92.61	p2p_ll	99.66	p2p_lt
C	97.28	90.95	p2p_ll	99.43	p2p_rt
N	95.65	89.38	p2p_ll	98.56	p2r_ri
K	95.37	89.25	p2r_ll	98.59	p2p_ri
L	94.07	79.81	p2p_ll	99.19	p2p_rt
A	89.52	76.04	p2p_ll	97.73	p2p_rt
VTB	86.86	66.76	p2r_ll	95.99	p2p_rt
B	85.68	73.72	p2r_ll	94.68	p2p_rt
E	82.41	64.32	p2p_ll	93.82	p2p_rt
M	50.89	12.48	p2p_ll	92.78	p2p_rt

Table 5: SDK rankings for OHIO dataset at FAR 0.01%.

SDK	Avg TAR	Min TAR	Data Set	Max TAR	Dataset
R	99.32	97.22	p2p_li	100.00	r2r_rt
H	99.32	98.33	p2p_li	100.00	r2r_rt
O	98.98	97.22	p2p_li	100.00	r2r_rt
I	98.90	96.25	p2p_ri	100.00	r2r_rt
Q	98.41	93.97	p2p_ri	100.00	r2r_rt
J	98.04	94.03	p2p_li	100.00	r2r_li
P	97.98	96.67	p2p_lm	100.00	r2r_lm
G	97.17	91.67	p2p_li	99.73	r2r_lt
F	97.17	93.89	p2p_rm	99.44	r2r_rm
D	95.90	90.84	p2p_rm	98.89	r2r_rt
C	95.25	92.22	p2p_lm	98.33	r2r_rm
N	95.03	89.83	p2p_lm	98.89	r2r_rt
K	94.75	89.37	p2p_lt	100.00	r2r_rt
L	91.24	82.78	p2p_rm	99.44	r2r_rm
A	87.24	79.76	p2p_ri	93.95	r2r_ri
VTB	84.20	75.04	p2p_ri	95.58	r2r_rm
B	78.53	72.22	r2r_lt	86.11	p2p_rt
E	74.53	61.11	p2p_lm	86.67	r2r_rt
M	63.83	23.89	p2p_lm	90.56	r2r_rm

Table 6: SDK rankings for SD29 dataset at FAR 0.01%.

SDK	Average # of Matches per Second
M	2000
P	1380
E	900
N	890
Q	770
C	720
A	640
O	610
G	560
K	520
D	500
J	480
F	300
B	240
H	220
VTB	180
R	80
I	80
L	15

Table 7: Timing of SDKs.

Conclusions

One of the major results of this test is that it demonstrates that NIST in-house testing using vendor supplied biometric software (SDKs) is a practical, accurate and cost effective alternative to public competitions such as FpVTE [5] or FRVT 2002 [8]. Once the SDK specification was written, the interaction between NIST and the vendors worked very effectively. Comparison with the medium scale test done in [5] shows that similar accuracy results and better speed results can be obtained using SDK testing. The process is also much more cost effective than public competitions. The cost to run these tests was about 50% of the cost conducted in [5].

Each vendor's performance and ranking on both the SDK and FpVTE tests were compared. Only one vendor performed with significantly lower accuracy on the SDK test. This vendor was not one of the three highest ranked vendors in FpVTE. All other vendors had similar performance and ranking on both the SDK and FpVTE tests. This provides independent confirmation that the systems used in FpVTE contained substantially the same algorithms as discussed in this report.

It should be noted that the datasets (except SD29) were government use only so the vendors were not given a “training” set to tune software algorithms before testing.

Results show there are SDKs that perform consistently well across all the datasets. As expected this level of performance results in lower matcher speeds than those discussed in references [6] and [7]. The matcher used in US VISIT is competitive with other COTS matchers, based on accuracy and speed for the given datasets tested. This is discussed in more detail in reference [7]. In general, thumbs matched better than index fingers and right fingers were better than left.

If the data quality is good enough, as shown with the OHIO data, a faster matcher could do almost as well as the slower matchers. Also, thumbs and index fingers performed equally well on the high quality Ohio dataset.

References

[1] "Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability," PDF document at http://www.itl.nist.gov/iaui/894.03/pact/pact.html ("NIST Appendix A" link), November 2002.

[2] Public Law 107-56 (USA PATRIOT ACT); 107th United States Congress, Washington, D.C.; 26 October 2001

[3] Public Law 107-173 (Enhanced Border Security and Visa Entry Reform Act of 2002); 107th United States Congress, Washington, D.C.; 14 May 2002

[4] C L Wilson, C I Watson, M D Garris, A Hicklin; "Studies of Fingerprint Matching Using the NIST Verification Test Bed (VTB)" NIST IR 7020; National Institute of Standards and Technology; Gaithersburg Maryland , 07 July 2003 at ftp://sequoyah.nist.gov/pub/nist_internal_reports/ir_7020.pdf

[5] Charles Wilson, R. Austin Hicklin, Harold Korves, Bradford Ulery, Melissa Zoepfl, Mike Bone, Patrick Grother, Ross Micheals, Steve Otto, and Craig Watson, "Fingerprint Vendor Technology Evaluation 2003," DRAFT NIST IR, June 2004.

[6] Stephen S. Wood and Charles L. Wilson, "Studies of Plain-to-Rolled Fingerprint Matching Using the NIST Algorithmic Test Bed (ATB)," NIST IR 7112; National Institute of Standards & Technology, Gaithersburg Maryland at ftp://sequoyah.nist.gov/pub/nist_internal_reports/ir_7112.pdf

[7] C. L. Wilson, M. D. Garris, and C. A. Watson ,"Matching Performance for the US-VISIT IDENT System Using Flat Fingerprints," NIST IR 7110; National Institute of Standards & Technology, Gaithersburg Maryland, May 2004, at ftp://sequoyah.nist.gov/pub/nist_internal_reports/ir_7110.pdf

[8] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone, "Face recognition vendor test 2002," NIST IR 6965, National Institute of Standards & Technology, Gaithersburg Maryland, March 2003, results at http://www.itl.nist.gov/iaui/894.03/face/face.html

[9] Other related NIST fingerprint work may be viewed at http://fingerprint.nist.gov.