IDTESTINGLAB (IDTL)

CARLOS III UNIVERSITY OF MADRID

MULTI-SENSOR PERFORMANCE EVALUATION (MADRID-2)

PUBLIC REPORT (VERSION 2)

Leganes (Spain), November 19, 2018

VERSION CONTROL

Number	Date	Author	Comments
V0.1	2018/10/17	Raul Sanchez-Reillo	Draft extracted from Final Report
V1.0	2018/10/22	Raul Sanchez-Reillo	Final review and publication
V2.0	2018/11/19	Raul Sanchez-Reillo	Targets of Evaluation naming convention debugged and updated

Table of Contents

Figure Index

Table Index

Glossary

DET	*Detection Error Tradeoff*
EER	*Equal Error Rate*
FAR	*False Acceptance Rate*
FMR	*False Match Rate*
FMR1000	*FNMR at an FMR value of 0.1% or 1 in 1,000*
FMR10000	*FNMR at an FMR value of 0.01% or 1 in 10,000*
FMR100000	*FNMR at an FMR value of 0.001% or 1 in 100,000*
FNMR	*False Non-Match Rate*
FPC	*FPC 1020 Fingerprint Sensor*
FRR	*False Rejection Rate*
FTA	*Failure to Acquire*
FTE	*Failure to Enrol*
FVQ	*A Fingerprint Vendor Quality provider*
GTV	*Ground Truth Verification*
GUTI	*University Group for Identification Technologies – Carlos III University of Madrid*
IDX	*Idex sensor*
NEU	*Neurotechnology FingerCell Algorithm*
NFIQ	*NIST Fingerprint Image Quality Algorithm*
NFIQ2	*NIST Fingerprint Image Quality Algorithm (2nd version)*
NT	*Neurotechnology Quality Method*
NXL	*NB-3023-U2 Fingerprint sensor*
NXS	*NB-2024-U Demonstrator*
ppi	*Pixels per inch*
QAA	*Quality Assurance in Acquisition*
QAE	*Quality Assurance in Enrolment*
UC3M	*Carlos III University of Madrid*
UPG	*EikonTouch 710 fingerprint sensor*

1 Introduction

This public report is an excerpt of part of an Evaluation Report where the performance of the combination of 5 fingerprint sensors and 7 fingerprint recognition algorithms (both feature extraction and comparison) has been analysed. The evaluation has been requested and sponsored by NEXT Biometrics, but it has been performed independently following the best practices given in ISO/IEC 19795-1 and ISO/IEC 19795-2. This includes, for example, that the scanner sequence for capturing is randomized and neither the sponsor nor the components manufacturers have been disclosed to operator or test subjects.

This evaluation is conceptually a continuation of the one performed between 2013 and 2015, which published a set of partial reports. One of the publications was made available in May 11, 2015 at http://idtestinglab.uc3m.es/data/_uploaded/Publications/PUBLIC%20REPORT_589%20users_Fingerprint_v1_1_release.pdf. There is a second publication coming out of such evaluation, as a scientific paper in IET Biometrics, under the following reference:

> *Fernandez-Saavedra, Belen; Sanchez-Reillo, Raul; Ros-Gomez, Rodrigo; Liu-Jimenez, Judith: 'Small fingerprint scanners used in mobile devices: the impact on biometric performance', IET Biometrics, 2016, 5, (1), p. 28-36, DOI: 10.1049/iet-bmt.2015.0018. Available at IET Digital Library, http://digital-library.theiet.org/content/journals/10.1049/iet-bmt.2015.0018*

This report has benefitted from the lessons learned in such previous evaluation, in terms of acquisition of data, error detection and execution performance, as well as in personal data protection. As the sensors are different from the ones used in 2015, a new database has been acquired. Details will be provided in Sections 3 and 4.

In order to differentiate between both evaluations, this one will be called as "Madrid-2", as the previous one has been known as "Madrid".

This introduction is completed with the description of the Testing Laboratory and Key Personnel that has carried out this evaluation. The last part of this introduction outlines the structure of this report.

1.1 IDTestingLab (IDTL)

Carlos III University of Madrid (UC3M) (http://www.uc3m.es) is one of Spain's most prestigious technical Universities. Due to its public, non-profit nature, the exploitation and dissemination strategies of UC3M largely coincide on its main objective, which is to use research results to advance and progress scientific knowledge. Exploitation of research achievements is carried out along two activities: educational in which existing and well-established knowledge and methods are diffused among the attendants of the University lectures and activities, and research into advancements and extensions of the understanding of scientific disciplines. To this end, UC3M relies on a pool of expert human resources and its reputation, which is based on past achievements, helping to attract the top choice of prospective students and research associates.

Research at Carlos III University of Madrid has always been one of the basic pillars of the University's activities, both to improve teaching and to generate new knowledge and new lines of research.

Within UC3M, the Electronics Technology Dpt. has 5 Research Groups. Among them, the University Group for Identification Technologies (GUTI – http://guti.uc3m.es) has a great experience in Biometrics, Smart Cards and Security in Identification Systems. In detail, GUTI's expertise in its R&D lines is:

- Smart Cards, from R&D to final applications (active since 1989).
- Biometrics, having large experience in different biometric modalities such as hand geometry, iris recognition, fingerprint, vascular system and handwritten signature (active since 1994).
- Match-on-Card Technology, achieving the first ever Match-on-Card solution in 1999.
- Security Infrastructures, developing their own PKI using smart cards in 1997.
- Their work in all these lines has leaded to hold the Secretariat in the Spanish Mirror Subcommittee in Biometrics (AEN/CTN71/SC37) and the Chair in the Spanish Mirror Subcommittee in Identification Cards (AEN/CTN71/SC17). They are also experts in SC27.

As a result of this work, UC3M opened an Evaluation Laboratory for Identification Technologies, called IDTestingLab (http://idtestinglab.uc3m.es). IDTestingLab is equipped with all relevant instruments to perform technology and scenario evaluations, and its personnel are trained to carry out operational evaluation as soon as a customer requests that kind of work.

This laboratory has carried out several tests, both by Industry request and by R&D project requirements. For those tests, a variety of tools have been developed, as well as building scenarios for end-to-end evaluations (scenario evaluations). Several innovative methodologies have already been designed and developed, amongst which are a methodology to measure the environmental condition influence on biometric systems (which has led to the development of ISO/IEC 29197), and a methodology for measuring the influence of usability in the performance of biometrics.

1.1.1 Experience

Since UC3M – IDTestingLab begun its activities, this laboratory has conducted several technology and scenario biometric performance evaluations for different biometric systems by either Industry request or R&D projects requirement. These works can be seen through its dissemination in different international conferences as well as in relevant scientific journals.

Moreover, within its R&D lines, UC3M-IDTestingLab has worked on biometric evaluations, both in terms of performance (including environmental conditions and usability) and security, considering the alignment of biometric evaluations and security evaluations, specifically with Common Criteria. Some of this work has been published to both the industrial and scientific community:

- Raul Sanchez-Reillo, Ines Ortega-Fernandez, Wendy Ponce-Hernandez, Helga C. Quiros-Sandoval, *"How to implement EU data protection regulation for R&D in biometrics"*, Computer Standards & Interfaces, Volume 61, 2019, Pages 89-96, ISSN 0920-5489, https://doi.org/10.1016/j.csi.2018.01.007. (http://www.sciencedirect.com/science/article/pii/S0920548917303161)
- R. Blanco-Gonzalo, R. Sanchez-Reillo, I. Goicoechea-Telleria and B. Strobl, *"The Mobile Pass Project: A User Interaction Evaluation"*, in IEEE Transactions on Human-Machine Systems, vol. 48, no. 3, pp. 311-315, June 2018. doi: 10.1109/THMS.2018.2791571. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8315493&isnumber=8358712
- R. Sanchez-Reillo, J. Liu-Jimenez and R. Blanco-Gonzalo, *"Forensic Validation of Biometrics Using Dynamic Handwritten Signatures"*, in IEEE Access, vol. 6, pp. 34149-34157, 2018. doi: 10.1109/ACCESS.2018.2849503. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8391714&isnumber=8274985
- Ines Goicoechea-Telleria, Raul Sanchez-Reillo, Judith Liu-Jimenez, and Ramon Blanco-Gonzalo, *"Attack Potential Evaluation in Desktop and Smartphone Fingerprint Sensors: Can They Be Attacked by Anyone?"* Wireless Communications and Mobile Computing, vol. 2018, Article ID 5609195, 16 pages, 2018. https://doi.org/10.1155/2018/5609195.
- I. Goicoechea-Telleria, J. Liu-Jimenez, H. Quiros-Sandoval and R. Sanchez-Reillo, *"Analysis of the attack potential in low cost spoofing of fingerprints"*, 2017 International Carnahan Conference on Security Technology (ICCST), Madrid, 2017, pp. 1-6. doi: 10.1109/CCST.2017.8167798. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8167798&isnumber=8167784
- R. Sanchez-Reillo, H. C. Quiros-Sandoval, I. Goicoechea-Telleria and W. Ponce-Hernandez, *"Improving Presentation Attack Detection in Dynamic Handwritten Signature Biometrics"*, in IEEE Access, vol. 5, pp. 20463-20469, 2017. doi: 10.1109/ACCESS.2017.2755771. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8048499&isnumber=7859429
- Raul Sanchez-Reillo, *"Signature analysis in the context of mobile devices"*, Image and Vision Computing, Volume 55, Part 1, 2016, Pages 34-37, ISSN 0262-8856, https://doi.org/10.1016/j.imavis.2016.03.011. (http://www.sciencedirect.com/science/article/pii/S0262885616300476)
- Fernandez-Saavedra, Belen; Sanchez-Reillo, Raul; Ros-Gomez, Rodrigo; Liu-Jimenez, Judith: *'Small fingerprint scanners used in mobile devices: the impact on biometric performance'*, IET Biometrics, 2016, 5, (1), p. 28-36, DOI: 10.1049/iet-bmt.2015.0018. IET Digital Library, http://digital-library.theiet.org/content/journals/10.1049/iet-bmt.2015.0018
- Blanco Gonzalo, Ramón; Sánchez Reíllo, Raúl; Liu Jiménez, Judith. Usability in biometric recognition systems. PhD Thesis Universidad Carlos III de Madrid. Departamento de Tecnología Electrónica. 2016. URI: http://hdl.handle.net/10016/23210
- Ramon Blanco-Gonzalo, Raul Sanchez-Reillo, Oscar Miguel-Hurtado, Eric Bella-Pulgarin, *"Automatic usability and stress analysis in mobile biometrics"*, Image and Vision Computing, Volume 32, Issue 12, 2014, Pages 1173-1180, ISSN 0262-8856,

https://doi.org/10.1016/j.imavis.2014.09.003.
(http://www.sciencedirect.com/science/article/pii/S0262885614001383)

- B. Fernandez-Saavedra, R. Sanchez-Reillo, J. Liu-Jimenez and J. G. Ruiz, *"Best practices for the security evaluation of biometric systems"*, 2014 International Carnahan Conference on Security Technology (ICCST), Rome, 2014, pp. 1-6. doi: 10.1109/CCST.2014.6987034. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6987034&isnumber=6986962
- Blanco-Gonzalo R., Diaz-Fernandez L., Miguel-Hurtado O., Sanchez-Reillo R. (2014) *Usability Evaluation of Biometrics in Mobile Environments*. In: Hippe Z., Kulikowski J., Mroczek T., Wtorek J. (eds) Human-Computer Systems Interaction: Backgrounds and Applications 3. Advances in Intelligent Systems and Computing, vol 300. Springer, Cham
- Belen Fernandez-Saavedra, Raul Sanchez-Reillo, Judith Liu-Jimenez, Oscar Miguel-Hurtado, *"Evaluation of biometric system performance in the context of Common Criteria"*, Information Sciences, Volume 245, 2013, Pages 240-254, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2013.05.022. (http://www.sciencedirect.com/science/article/pii/S0020025513003976)
- Blanco-Gonzalo, Ramon; Sanchez-Reillo, Raul; Miguel-Hurtado, Oscar; Liu-Jimenez, Judith: *'Performance evaluation of handwritten signature recognition in mobile environments'*, IET Biometrics, 2014, 3, (3), p. 139-146, DOI: 10.1049/iet-bmt.2013.0044. IET Digital Library, http://digital-library.theiet.org/content/journals/10.1049/iet-bmt.2013.0044
- Fernández Saavedra, María Belén; Sánchez Reíllo, Raúl; Evaluation methodologies for security testing biometric systems beyond technological evaluation. PhD Thesis, Universidad Carlos III de Madrid. Departamento de Tecnología Electrónica. 2013. URI: http://hdl.handle.net/10016/17145
- R. Blanco-Gonzalo, O. Miguel-Hurtado, A. Mendaza-Ormaza and R. Sanchez-Reillo, *"Handwritten signature recognition in mobile scenarios: Performance evaluation"* , 2012 IEEE International Carnahan Conference on Security Technology (ICCST), Boston, MA, 2012, pp. 174-179. doi: 10.1109/CCST.2012.6393554. URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6393554&isnumber=6393529
- Raul Sanchez-Reillo, Raul Alonso-Moreno, Belen Fernandez-Saavedra, Young-Bin Kwon, *"Standardised system for automatic remote evaluation of biometric algorithms"*, Computer Standards & Interfaces, Volume 34, Issue 5, 2012, Pages 413-425, ISSN 0920-5489, https://doi.org/10.1016/j.csi.2012.01.006. (http://www.sciencedirect.com/science/article/pii/S0920548912000189)
- B. Fernandez-Saavedra, R. Sanchez-Reillo, J. Liu-Jimenez and I. Tomeo-Reyes, *"Common Criteria and Biometric Performance Testing"*, Proceedings of the International Biometric Performance Conference, NIST, Gaithersburg, Maryland, 2012.
- B. Fernandez-Saavedra, R. Sanchez-Reillo, R. Alonso-Moreno and I. Tomeo-Reyes, *"Security Evaluation of Biometric Systems in Common Criteria"*, Proceedings of the International Common Criteria Conference, Antalya, Turkey, 2010.

- B. Fernandez-Saavedra, R. Sanchez-Reillo, R. Alonso-Moreno and I. Tomeo-Reyes, *"Development of a Protection Profile for Biometric Systems Following ISO/IEC TR 15446"*, Proceedings of the International Common Criteria Conference, Tromso, Norway, 2009.
- R. Sanchez-Reillo, R. Blanco-Gonzalo, O. Miguel-Hurtado and A. Mendaza-Ormaza, Migrating Biometrics to Mobile Scenarios: Performance and Usability Evaluation, Proceedings of the International Biometric Performance Conference, NIST, Gaithersburg, Maryland, 2012.
- B. Fernandez-Saavedra, R. Alonso-Moreno, J. Uriarte-Antonio and R. Sanchez-Reillo, Evaluation methodology for analysing usability factors in biometrics, Aerospace and Electronic Systems Magazine, IEEE, 2010, 25(8), p. 20-31.
- R. Sanchez-Reillo, J. Uriarte-Antonio, I. Tomeo-Reyes and B. Fernandez-Saavedra,Testing Laboratory Interactions for Evaluating Biometrics,Proceedings of the International Biometric Performance Conference, NIST, Gaithersburg, Maryland, 2010.
- B. Fernandez-Saavedra, F. J. Diez-Jimeno, R. Sanchez-Reillo and R. Lazarick, Establishment of baseline performance for "end to end" biometric system evaluations, Security Technology (ICCST), 2010 IEEE International Carnahan Conference on, 2010.
- B. Fernandez-Saavedra, R. Alonso-Moreno, J. Uriarte-Antonio and R. Sanchez-Reillo, Usability Evaluation of Fingerprint Based Access Control Systems, Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on, 2010.
- R. Sanchez-Reillo, B. Fernandez-Saavedra, J. Liu-Jimenez and Y. B. Kwon, Changes to Vascular Biometric System Security & Performance, Aerospace and Electronic Systems Magazine, IEEE, 2009, 24(6), p. 4-14.

Besides, UC3M-IDTestingLab was requested by Spanish Certification Body, i.e. Centro Criptológico Nacional (CCN), for developing a guide for the performance evaluation of biometric devices. This is the "CCN-STIC-492 Evaluación de parámetros de rendimiento en dispositivos biométricos" (https://www.ccn.cni.es/index.php?option=com_content&view=article&id=6&Itemid=9&lang=es). Unfortunately, this guide is not available to the general public.

In addition, UC3M-IDTestingLab is involved in national and international standardization activities being its personnel are editors or co-editors of several ISO/IEC JCT1 SC37 standards:

- ISO/IEC 30106-3 - Information technology -- Object oriented BioAPI -- Part 3: C# implementation (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=59555) 01/05/2016
- ISO/IEC 30106-2 - Information technology -- Object oriented BioAPI -- Part 2: Java implementation (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=59554) 01/04/2016
- ISO/IEC 30106-1 - Information technology -- Object oriented BioAPI -- Part 1: Architecture (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=53226) 01/04/2016

- ISO/IEC TR 30125 - Biometrics -- Biometrics used with mobile devices (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=53245) 01/04/2016
- ISO/IEC 19794-6:2011/Amd 1:2015 - Conformance testing methodology and clarification of defects (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=61097) 16/09/2015
- ISO/IEC 29794-6:2015 - Information technology -- Biometric sample quality -- Part 6: Iris image data (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=54066). 01/07/2015
- ISO/IEC TR 29194:2015 - Information Technology -- Biometrics -- Guide on designing accessible and inclusive biometric systems (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=45273). 21/05/2015
- ISO/IEC 19794-11:2013/Amd 1:2014 - Conformance test assertions (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=59434). 09/07/2014
- ISO/IEC TR 30117:2014 - Information technology -- Guide to on-card biometric comparison standards and applications (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=53237). 13/03/2014
- CEN/TS 15480-1:2012 Identification card systems. European Citizen Card. Physical, electrical and transport protocol characteristics (http://shop.bsigroup.com/ProductDetail/?pid=000000000030260014) 01/12/2012
- ISO/IEC 29109-6:2011 - Information technology -- Conformance testing methodology for biometric data interchange formats defined in ISO/IEC 19794 -- Part 6: Iris image data (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50908). 01/10/2011
- ISO/IEC 19794-6:2011 - Information technology -- Biometric data interchange formats -- Part 6: Iris image data (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=50868). 01/10/2011
- ISO/IEC 29164:2011 - Information technology -- Biometrics -- Embedded BioAPI (http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=45243). 29/09/2011
- ISO/IEC 19794-6 AMD1 Biometric data interchange formats - Part 6: Iris image data - Amendment 1: Conformance testing methodologies
- ISO/IEC 29109-6 Conformance testing methodology for biometric data interchange formats defined in ISO/IEC 19794 - Part 6: Iris image data
- ISO/IEC 29164 Embedded BioAPI

- ISO/IEC 29197 Evaluation methodology for environmental influence in biometric system performance
- ISO/IEC 30106-3 Object Oriented BioAPI - Part 3: C# Implementation

Finally, it is important to highlight that UC3M-GUTI has been an active participant in different European Projects (eEpoch, BioSec, BEST Network, MobilePass, EKSISTENZ, ORIGINS, PYCSEL and AMBER), as well as National and Regional funded projects (SIDECAR, PIBES, ESDIB, EMOCION, iSec, etc.) where identification scenarios were requested to be identified and/or developed. In all of those projects, UC3M-GUTI has been the partner in charge of such tasks.

1.1.2 Short profile of key personnel

Dr. Raul Sanchez-Reillo is currently Associate Professor at UC3M. He is the Head of the University Group for Identification Technologies (GUTI), involved in project development and management concerning a broad range of applications, from Social Security Services till Financial Payment Methods. He has taken part in European Projects like eEpoch, BioSec, BEST Network, EKSISTENZ, MobilePass and ORIGINS, carrying out leadership tasks. He is expert in Security and Biometrics and member of SC17, SC27 and SC37 Standardization Committees, holding the Spanish Chair in SC17 and the Secretariat in SC37.

1.2 Document Structure

The report is structured as follows. After this introduction, a brief explanation of the fingerprint sensors used is provided in Section 2. Then the procedure for collecting the database is explained in Section 3, providing the description of the database acquired in Section 4.

Once the acquisition process has been explained, the following section describes the approach for the off-line processing of the database with each algorithm. Also, Section 5 describes the enrolment input, the acquisition process and the ground truth validation, as well as the execution of the biometric comparisons for obtaining the performance.

Then, the results obtained for one of the algorithms used is provided. The results shown are of the only algorithm which has allowed the public dissemination of the results. The results have been obtained after an off-line execution with the whole database previously captured. The report will be finished with the conclusions obtained, as well as some new lessons learned.

2 Fingerprint Sensors

This evaluation report uses 5 different sensors for its content. Those 5 sensors are described briefly in this clause.

2.1 NB-3023-U2 Fingerprint sensor (NXL)

The NB-3023-U2 fingerprint sensor uses the patented NEXT Active Thermal™ sensing principle. The operation principle is different from capacitive sensors. Heat is applied to sensor pixels and the heat transfer into the finger in contact with the sensor surface is measured. The NB-3023-U2 is a commercial fingerprint sensor with USB interface and can connect to a Windows 7 or 8.1 or 10 host computer. It is based on the NEXT NB-2023-U2 sensor module in an ergonomic housing.

The sensor is flat and does not have distortion like some optical sensors. The sensor resolution is 385 ppi in both vertical and horizontal direction. A library delivering a plain image is available. The sensor produces 256 grey levels.

The NB-3023-U2 sensor has an active scanning area of 11.9 x 16.9 mm and 180 x 256 pixels.

For readability of this report, this sensor will be mentioned by the acronym NXL.

Figure 1 – NEXT NB-3023-U2 fingerprint sensor (named as NXL)

2.2 NB-2024-U Demonstrator (NXS)

This sensor is not a commercial product but a prototype using NB-2024-U2 module in an ergonomic housing with opening similar to typical notebook integration. In other words, it is a cropped version of the NXL sensor. The active sensing area is 11.9 x 11.9 mm while the emulation of the notebook integration may result in reduced touchable area. As the resolution is 385ppi, the dimension of the images acquired is 180 x 180 pixels, being acquired with 256 grey levels.

For readability this sensor will be named as NXS.

Figure 2 – NB-2024-U Demonstrator (on the right, named as NXS), compared with NXL (on the left)

2.3 EikonTouch 710 fingerprint sensor (UPG)

Initially manufactured by UPEK, currently is within the product line of Crossmatch, under product name TCS1. This sensor uses capacitive technology, with a resolution of 508ppi using an active area of 12.8 x 18.0 mm. The image acquired has 256 x 360 pixels, and the image is captured using 256 grey levels. The model used has gold colour surface coating removed and has FIPS 201 certification.

For readability of this report, this sensor will be mentioned by the acronym UPG.

Figure 3 – EikonTouch 710 fingerprint sensor (named as UPG)

2.4 FPC 1020 fingerprint sensor (FPC)

The FingerprintCards FPC 1020 fingerprint sensor is actually not a commercial product but part of an evaluation kit. This sensor uses active capacitive technology to obtain the images of the fingerprint. When a finger is in contact with the sensor area, a weak electrical charges is sent via the finger. Using these charges the sensor measures the capacitance pattern across the surface.

The sensor has an active area of 9.6 x 9.6 mm and with a resolution of 508ppi acquires images of 192 x 192 pixels with 256 grey levels.

For readability of this report, this sensor will be mentioned by the acronym FPC.

Figure 4 – FPC 1020 fingerprint sensor (named as FPC)

2.5 Idex sensor (IDX)

The Idex Eagle sensor is also a capacitive small area sensor, with a compact design to be integrated in any kind of authentication device, such as a Point of Service, or a USB-connected peripheral. Since it was not available directly from Idex, a third-party vendor USB device relying on the sensor was procured - using only the image interface of the solution. The sensor has an active area of 7.7 x 7.8 mm and a resolution of 373ppi. The images acquired are of 113 x 115 pixels in greyscale and the number of grey levels used is not specified.

For readability of this report, this sensor will be named IDX

Figure 5 – Idex fingerprint sensor (named as IDX), picture of USB reader not shown here.
Reader was not disassembled. Sensor picture taken from public information.

3 Database Collection Procedures

This section is written with the intention to set and clarify the concepts, rates, decisions and process, in order to acquire the Madrid-2 DB, before performing the final Madrid-2 Evaluation on 5 sensors and several algorithms.

This section only covers the acquisition process, both enrolment and visit acquisition. The whole process has considered the lessons learned from the acquisition of the "Madrid DB". Therefore, there are several differences between this acquisition and the one in the "Madrid DB". For those readers which are familiar with the "Madrid DB" evaluation, the main differences are indicated along this text, written in *green and italics*.

This kind of evaluation requires further definitions to those stated in the published version of ISO/IEC 19785-1 and -2. Therefore, the first two sub-sections cover the terminology to be used and the error rates needed.

3.1 Terms to be used for evaluations

This document copies the terms from ISO/IEC 19795-1, ISO/IEC 19795-2 and ISO/IEC 2382-37. Be aware that the most recent document is the one on vocabulary (ISO/IEC 2382-37) and that ISO/IEC 19795 parts are currently in revision and will have to adopt terms from the latest edition of ISO/IEC 2382-37.

3.1.1 ISO/IEC 19795-1

- **Sample:** user’s biometric measures as output by the data capture subsystem
    - EXAMPLE Fingerprint image, face image and iris image are samples.
    - NOTE In more complex systems, the sample may consist of multiple presented characteristics (e.g., 10-print fingerprint record, face image captured from different angles, left and right iris image pair.)
- **Presentation:** submission of a single biometric sample on the part of a user
- **Attempt:** submission of one (or a sequence of) biometric samples to the system

  - NOTE An attempt results in an enrolment template, a matching score (or scores), or possibly a failure–to-acquire.
- **Transaction:** sequence of attempts on the part of a user for the purposes of an enrolment, verification or identification
  - NOTE There are three types of transaction: enrolment sequence, resulting in an enrolment or a failure-to-enrol; a verification sequence resulting in a verification decision; or identification sequence, resulting in an identification decision.
- **genuine attempt:** single good-faith attempt by a user to match their own stored template

3.1.2 ISO/IEC 19795-2

- **enrolment attempt:** submission of one or more biometric samples for a Test Subject for the purposes of enrolment in a biometric system
  - NOTE 1 One or more enrolment attempts may be permitted or required to constitute an enrolment transaction. An enrolment attempt may be comprised of one or more enrolment presentations.
  - NOTE 2 See Annex B for illustration of the relationship between presentation, attempt, and transaction
- **enrolment presentation:** submission of an instance of a biometric characteristic for a Test Subject for the purpose of enrolment
  - NOTE One or more enrolment presentations may be permitted or required to constitute an enrolment attempt. An enrolment presentation may or may not result in an enrolment attempt.

3.1.3 ISO/IEC 2382-37

- **biometric sample:** analog or digital representation of biometric characteristics prior to biometric feature extraction
  - EXAMPLE A record containing the image of a finger is a biometric sample.
- **biometric probe / biometric query:** biometric sample or biometric feature set input to an algorithm for biometric comparison to a biometric reference(s)
  - Note 1 to entry: In some comparisons, a biometric reference might be used as the subject of the comparison with other biometric references or incoming samples used as the objects of the comparisons. For example, in a duplicate enrolment check, a biometric reference will be used as the subject for comparisons against all other biometric references in the database.
  - Note 2 to entry: Typically, in a biometric comparison process, incoming biometric samples serve as the subject of comparisons against objects stored as biometric references in a database.
- **biometric reference:** one or more stored biometric samples, biometric templates or biometric models attributed to a biometric data subject and used as the object of biometric comparison

    - EXAMPLE Face image stored digitally in a passport, fingerprint minutiae template in a National ID card or Gaussian Mixture Model for speaker recognition, in a database.
    - Note 1 to entry: A biometric reference may be created with implicit or explicit use of auxiliary data, such as Universal Background Models.
    - Note 2 to entry: The subject/object labelling in a comparison might be arbitrary. In some comparisons, a biometric reference might be used as the subject of the comparison with other biometric references or incoming samples and input to an algorithm for biometric comparison. For example, in a duplicate enrolment check a biometric reference will be used as the subject for comparison against all other biometric references in the database.
- **capture attempt:** activity with the intent of producing a captured biometric sample
    - Note 1 to entry: The capture attempt is the interface between the presentation by the biometric capture subject and the action of the biometric capture subsystem.
    - Note 2 to entry: The "activity" taken may be on the part of the biometric capture subsystem or the biometric capture subject.
- **capture transaction:** one or more capture attempts with the intent of acquiring all of the biometric data from a biometric capture subject necessary to produce either a biometric reference or a biometric probe
- **biometric presentation:** interaction of the biometric capture subject and the biometric capture subsystem to obtain a signal from a biometric characteristic
    - Note 1 to entry: The biometric capture subject may not be aware that a signal from a biometric characteristic is being captured.

3.1.4 Conclusions on Terminology

Starting with the vocabulary standard (which is the most recent one, and it will have to be adopted in the revision of ISO/IEC 19795), these are the conclusions:

- After going through the definitions, the chronological order of the terms is the following:
    - Presentation (user interaction)
    - Transaction (one or more attempts)
    - Attempt
    - Sample
    - Probe
    - Reference
- So the final result is a **sample**. When a sample is acquired during enrolment, it will be used to create the subject **reference**. When a sample is acquired during the acquisition phases, it will be considered a **probe**, which will be compared against the references.
- In order to obtain a sample, several attempts can be used
- A transaction is the process defined with the target of acquiring a sample. A transaction defines the maximum number of attempts, and other parameters such as time-out, etc.
    - Definition in ISO/IEC 19795-1 has to be updated

- Presentation is a user-interaction concept, which defines the act of a user to start and accomplish a transaction.
    - Definition in ISO/IEC 19795-1 has to be updated

3.1.5 Madrid-2 Evaluation definitions

- Enrolment:
    - Target:
        - 4 samples for enrolment
            - *In "Madrid DB" the target was to get 2 valid transactions, i.e. 2 samples*
    - 4 Transactions with operator assistance
    - Transaction definition:
        - Maximum of 3 attempts
        - Enrolment sample taken from the first attempt overcoming the enrolment quality and ground truth thresholds
            - For the 1st transaction, only the enrolment quality threshold
        - Independently of the sensor SDK, a time-out of 20 seconds is applied before declaring a FTA
            - *In "Madrid DB" no time-out was considered for enrolment*
        - Operator is able to control the whole process including denying enrolment of a finger
- Acquisition visits:
    - Target:
        - 1st visit:
            - 6 samples to be acquired by finger and sensor
        - 2nd visit:
            - 10 samples to be acquired by finger and sensor
                - *In "Madrid DB" the target for the 2nd visit was to get 6 samples*
    - A total of 16 transactions without operator assistance but with operator supervision
    - Transaction definition:
        - Maximum of 3 attempts
        - Sample taken from the first attempt overcoming the visit quality and ground truth thresholds
        - Independently of the sensor SDK, a time-out of 10 seconds is applied before declaring a FTA
        - Possibility to manually cancel the transaction due to sensor not detecting finger

3.2 Error Rates

The following will apply (some of them not included in ISO/IEC 19785-1):

- FTE – Failure To Enrol. Increments when not possible enrolment after finishing trying to capture a maximum of 4 enrolment samples
- FTA – Failure To Acquire. This rate is not going to be used in here, as it is better to improve granularity in the determination of error rates, using the following ones:
    - FTAea – Failure To Acquire a sample during Enrolment at an attempt. Increments if an attempt is discarded during enrolment (even with the 3rd attempt)
    - FTAe – Failure To Acquire a sample during Enrolment. Increments after failing in 3 attempts for that sample.
    - FTAaa – Failure To Acquire a sample at an attempt during Acquisition. Increments if an attempt is discarded during enrolment (even with the 3rd attempt)
    - FTAa – Failure To Acquire a sample during Acquisition. Increments after failing in 3 attempts for that sample.
    - FTPea – Failure To Process an attempt during enrolment
    - FTPap – Failure To Process an attempt during acquisition, due to the pure processing process (not the verification)
    - FTPav – Failure To Process an attempt during acquisition, due to not verification during GTV

3.3 File Labelling

The storage of the samples will be done in the following way

3.3.1 Tree Structure

The database will be stored using individual files at the following tree structure, which separates files according to the sensor:

- MADRID2DB
    - NXL
    - NXS
    - UPG
    - FPC
    - IDX

Within each of these folders, the following tree will be created:

- \enrolment
    - \errors
    - \samples
    - \references

        - \acquisition
            - \errors
            - \samples

3.3.2 File Name

Each sample is named using the following convention:

CCC_UUUUUU_TT_SSS_MMM_AA_OOO_PPP_QQQ_Z.extension

Where:

- CCC – is the capture device (i.e. the hardware). Available names are:
    - NXL: Next Large sensor
    - NXS: Next Small sensor
    - FPC: Fingerprint Card
    - IDX: Idex sensor
    - UPG: UPEK Golden plated
- UUUUUU – 6 character ID number of the test crew subject (filled with ‘0’ on the left)
- TT – trait number according to ISO. Possible values:
    - 01 – Right Thumb
    - 02 – Right Index
    - 03 – Right Middle
    - 06 – Left Thumb
    - 07 – Left Index
    - 08 – Left Middle
- SSS – Session identifier. Possible values are:
    - ENR – enrolment
    - V01 – visit 1
    - V02 – visit 2
- MMM – Sample number as a 3 character number (filled with ‘0’ on the left)
- AA – Attempt number for that sample. Possible values:
    - 01
    - 02
    - 03
- OOO – NFIQ2 quality score (filled with ‘0’ on the left)
- PPP – NT quality score (filled with ‘0’ on the left)
- QQQ – FVQ quality score (filled with ‘0’ on the left)
- Z – Status code. Possible values:
    - 0 – Sample OK
    - 1 – FTA due to quality
    - 2 – FTA due to processing
    - 3 – FTA due to verification
    - 4 – FTA due to timeout

  - …
  - 7 – FTA due to user rejection
  - 8 – FTA due to operator rejection
  - 9 – FTA other
  - N – Not for use, as not having a reference to provide GTV
- The extension will be the one applicable. For example:
  - PGM
  - PNG
  - BIR
  - *In "Madrid DB" the samples were captured only using BMP file format, although in the post-processing the ISO/IEC 19794-2 BIR format was also used.*

3.4 Internal Functions

For assuring a good database acquisition, the following internal functions will be used

3.4.1 Quality Assurance in Enrolment (QAE)

During enrolment, 3 quality algorithms will be executed, although only 2 of them will be used for determining the success (or not) of overcoming que quality criteria. The three algorithms will be:

- NFIQ2, which provides a score between 0 and 100, and will be coded as OOO. There are no suggestions for a threshold in enrolment
  - This quality algorithm has not been used for QAE for several reasons, including that the NFIQ2 version used is not initially intended to be used with this kind of fingerprint images, as they are obtained with non-optical sensors, and also with some of them not reaching the 500ppi resolution.
    - *Anyhow, this quality criteria has been calculated as to test its behaviour, and also compare it with the poor performance observed with NFIQ1 (known simply as NFIQ) in the "Madrid Evaluation"*
- NT (Neurotechnology), which provides a score between 0 and 100, and will be coded as PPP. The manufacturer states a recommended score for enrolment of 60 or above in final applications.
- FVQ (a Fingerprint Vendor Quality provider), which provides a score between 0 and 100, and will be coded as QQQ. The manufacturer states a recommended score for enrolment of 40 or above in final applications.

3.4.1.1 Thresholds and Quality criteria

It has been checked that for dry fingers with some of the sensors used, the manufacturer thresholds are too restrictive. In addition, establishing a too restrictive threshold may bias the results in favour of the algorithms from the manufacturers also providing the QAE. Therefore, it has been decided to use the following thresholds:

- the_nt = 48

- the_fvq = 30

The quality criteria to be used will be a logical AND on both quality assessment algorithms:

QAE = (Qnt >= the_nt) AND (Qfvq >= the_fvq)

3.4.2 Quality Assurance in Acquisition (QAA)

During acquisition (outside of enrolment), also the 3 same quality algorithms than in QAE will be executed. But there are some differences in the thresholds recommended and the ones finally used. The three algorithms will be:

- NFIQ2, as in QAE. The manufacturer does not provide any suggestion for a threshold in acquisition.
- NT (Neurotechnology), as in QAE, but for acquisition the manufacturer states a recommended score of 48 or above.
- FVQ (a Fingerprint Vendor Quality provider), also as in QAE, but also the manufacturer states a different score recommended for acquisition. In this case such score is 30 or above.

Note: while NFIQ and NFIQ2 are the required image quality algorithms in many government applications, is has proven superior in Madrid after-analysis to evaluate the quality metrics from the top-performing algorithms – in this case Neurotechnology and a non-disclosed vendor.

3.4.2.1 Thresholds and Quality criteria

For the same reasons used to state the threshold for enrolment, the one for acquisition has also been reduced:

- tha_nt = 40
- tha_fvq = 24

As not to bias the results obtained to encourage both algorithms, the quality criteria will be a logical OR on both quality assessment algorithms:

QAA = (Qnt >= tha_nt) OR (Qfvq >= tha_fvq)

3.4.3 Ground Truth Verification (GTV)

One of the lessons learned during the "Madrid Evaluation" was the need to build a ground truth mechanism that will guarantee that all samples are tagged correctly during the acquisition process, as to avoid data loss and incoherent results that will require double checking the database content offline.

In other words, it is necessary to check that the finger that is captured is the one relevant to the finger labelled and previously enrolled. Therefore, a GTV mechanism is implemented, using 2 commercial algorithms (NT and FVQ). The GTV mechanisms should not bias the evaluation, so a low threshold will be used, as well as a relaxed criteria.

3.4.3.1 Thresholds and acceptance criteria

The similarity score thresholds chosen for the GTV are the following:

- ths_nt = 40
- ths_fvq = 20

As not to bias the results obtained to encourage both algorithms, the acceptance criteria will be a logical OR on both scores:

GTV = (CSnt >= ths_nt) OR (CSfvq >= ths_fvq)

Being CSnt and CSfvq the comparison scores obtained with each of both GT comparison algorithms.

If the reference for one of the algorithms is not found, the result will be the one for the comparison of the other algorithm.

If none of both references are found, the result will be false, with Z = N, and storing the sample acquired in the errors folder.

3.4.4 Acquisition References Generation

During the DB acquisition process, the enrolment will be done not using the enrolment functionality from any algorithm manufacturer, but just the best sample acquired for each of the two commercial algorithms that are being used. Therefore, the process to generate the reference will be the following:

1. For NT:
   1.1. From all samples collected, pick the one with the highest PPP
   1.2. Process the sample to obtain the feature vector
   1.3. Store the feature vector in the reference folder, with the following format: CCC_UUUUUU_TT_SSS_NT_Z.extension
      1.3.1. SSS = REF
      1.3.2. Z = 0, if the processing is OK, and Z=2 if some error has occurred. If an error has occurred, then the file shall be empty, but with this filename.
2. For FVQ:
   2.1. From all samples collected, pick the one with the highest QQQ
   2.2. Process the sample to obtain the feature vector
   2.3. Store the feature vector in the reference folder, with the following format: CCC_UUUUUU_TT_SSS_IN.extension
      2.3.1. SSS = REF
      2.3.2. Z = 0, if the processing is OK, and Z=2 if some error has occurred. If an error has occurred, then the file shall be empty, but with this filename.

3.5 Program Flow

This section will explain how the DB acquisition program is designed.

3.5.1 Global Variables

For each capture device, the following variables will be calculated and stored:

- To be in the DB:
  - attempts_e: number of attempts during enrolment
  - fta_ea: number of acquisition errors per attempt during enrolment
  - ftp_ea: number of processing errors per attempt during enrolment
  - fta_e: number of acquisition errors at sample
  - fte: number of fingers not able to be enrolled
  - attempts_a: number of attempts during acquisition
  - fta_aa: number of acquisition errors per attempt during acquisition
  - ftp_ap: number of processing errors per attempt during acquisition
  - ftp_av: number of attempts not correctly verified during GTV
  - fta_a: number of acquisition errors at sample
- Within the execution:
  - C: capture device
  - U: user
  - T: finger
  - S: session
  - M: sample number
  - A: attempt number
  - O: NFIQ2 quality score
  - P: NT quality score
  - Q: FVQ quality score
  - Z: status

3.5.2 Flow description

The whole process will follow these steps:

1. Set U
2. Enrolment (4 samples, 3 attempts/sample)
   - 2.1. S = ENR
   - 2.2. Set next T
   - 2.3. If no more fingers available, jump to 3
   - 2.4. Set next C
   - 2.5. If no more capture devices available, jump to 2.2

- 2.6. M = 0
- 2.7. M++, A = 0
- 2.8. If M > 4
  - 2.8.1. If no sample has been obtained: fte++
  - 2.8.2. Else
    - 2.8.2.1. Generate_References
    - 2.8.2.2. If none of both references have been created, fte++ (the differences for each algorithm will be illustrated in the post-evaluation)
  - 2.8.3. Jump to 2.4
- 2.9. A++, attempt_e++
- 2.10. If A >3
  - 2.10.1. fta_e++
  - 2.10.2. jump to 2.6
- 2.11. Start Acquisition
- 2.12. If timeout
  - 2.12.1. Z=4
  - 2.12.2. Generate empty file (in errors folder) with qualities to NNN
  - 2.12.3. Jump to 2.9
- 2.13. If QAE is not successful
  - 2.13.1. Z=1
  - 2.13.2. fta_ea++
  - 2.13.3. Store file (in errors folder) with each quality score
  - 2.13.4. Jump to 2.9
- 2.14. Else
  - 2.14.1. Z=0
  - 2.14.2. Store file (in enrolment samples folder) with each quality score
  - 2.14.3. Jump to 2.9

3. Acquisition (6/10 samples, 3 attempts/sample)
   - 3.1. S = V0x (being x either 1 or 2)
   - 3.2. Set capture device (C) order and finger (T) order
   - 3.3. Set next T // next finger
   - 3.4. If no more fingers available, EXIT to new user
   - 3.5. M = 0;
   - 3.6. M++ // next sample
   - 3.7. If M> 6/10, jump to 3.3 // jumpt to next finger
   - 3.8. Set next C // next sensor
   - 3.9. If no more capture devices available, set C before the beginning and jump to 3.5
   - 3.10. A=0;
   - 3.11. A++, attempt_a++ // next attempt
   - 3.12. If A >3
     - 3.12.1. fta_a++
     - 3.12.2. jump to 3.8
   - 3.13. Start Acquisition
   - 3.14. If timeout

        3.14.1. Z=4
        3.14.2. Generate empty file (in errors folder) with qualities to NNN
        3.14.3. Jump to 3.11 // go to next attempt
    3.15. If QAA is not successful
        3.15.1. Z=1
        3.15.2. fta_aa++
        3.15.3. Store file (in errors folder) with each quality score
        3.15.4. Jump to 3.11 // go to next attempt
    3.16. If GTV is not successful due to processing error
        3.16.1. Z=2
        3.16.2. ftp_ap++
        3.16.3. Store file (in errors folder) with each quality score
        3.16.4. Jump to 3.11 // go to next attempt
    3.17. If GTV is not successful due to verification error
        3.17.1. Z=3
        3.17.2. ftp_av++
        3.17.3. Store file (in errors folder) with each quality score
        3.17.4. Jump to 3.11 // go to next attempt
    3.18. If GTV is successful
        3.18.1. Z=0
        3.18.2. Store file (in acquisition sample folder) with each quality score
        3.18.3. Jump to 3.8 // go to next sensor

4 Composition of the Database

This section of the document will describe the contents of the DB collected.

4.1 Test Crew Demographics

The test crew presents the following statistical data.

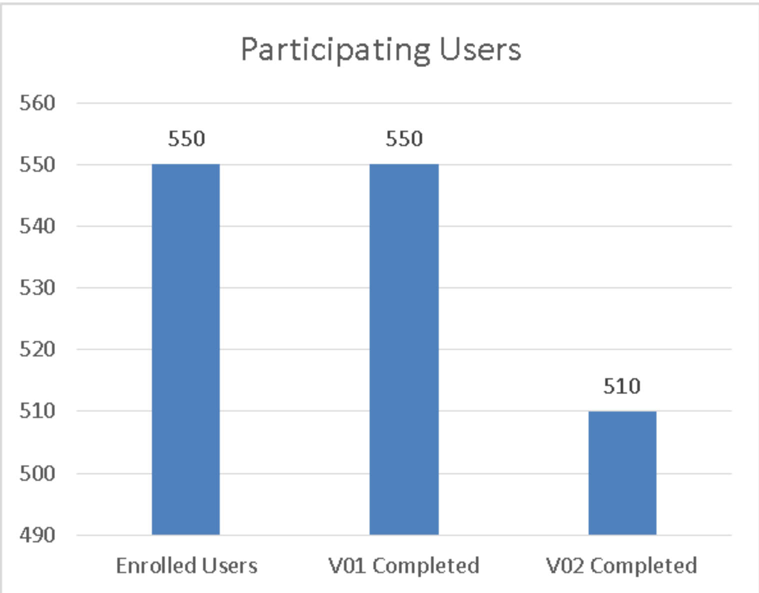


Figure 6 – Number of participating users

So only the 510 users having completed V02 will be used for the evaluation. The following figures provide the description of the composition of the database in terms of gender, age distribution, laterality, fingers damaged and previous knowledge.

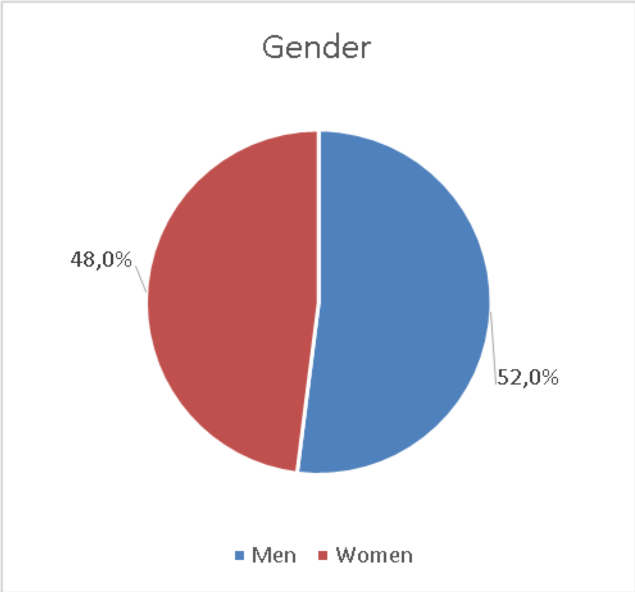


Figure 7 – Gender distribution

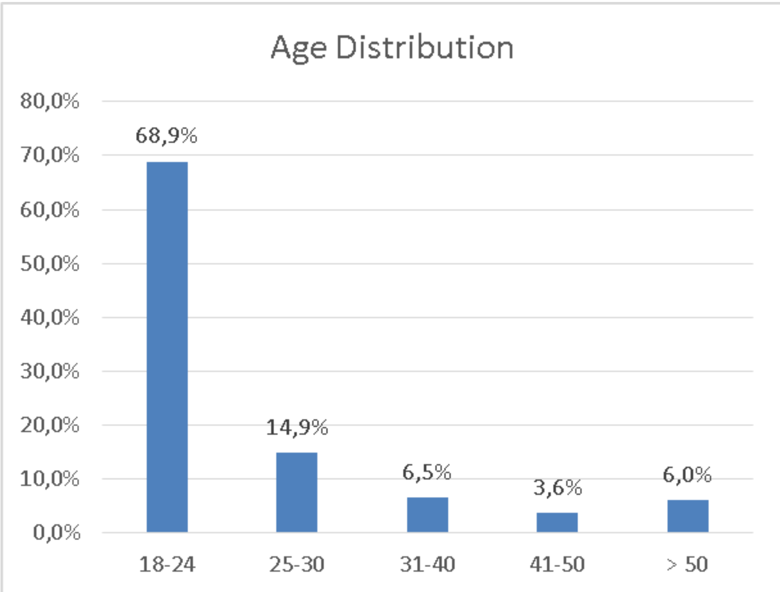


Figure 8 – Age Distribution

Only adult persons have been admitted participation in the database collection due to legal aspects. The age distribution is biased towards younger test subjects, because most of the volunteers were university students and their relatives.

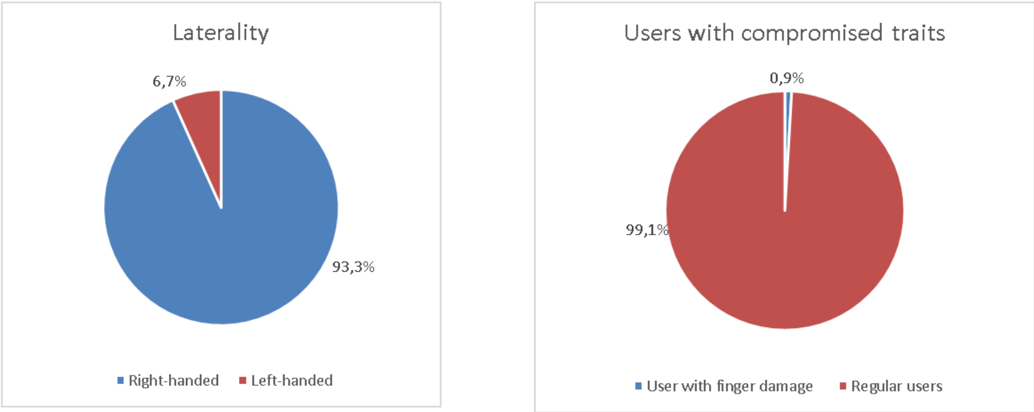


Figure 9 – Laterality and Fingers with Initial Damage

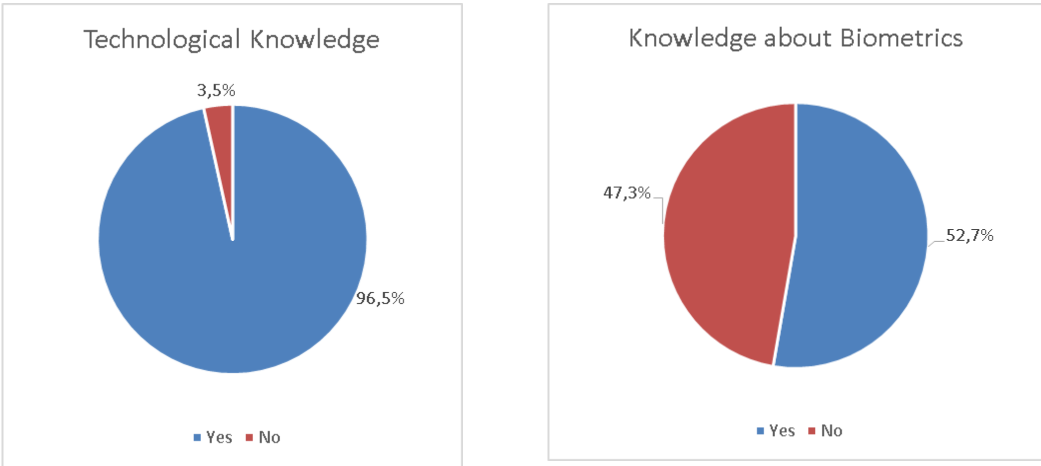


Figure 10 – Previous Knowledge

4.2 Visits and Acquisition Results

The following figure shows the distribution of the days between visits. It can be seen that nearly all users spent more than 14 days between visits, and few of them delayed the 2nd visit to more than 200 days.

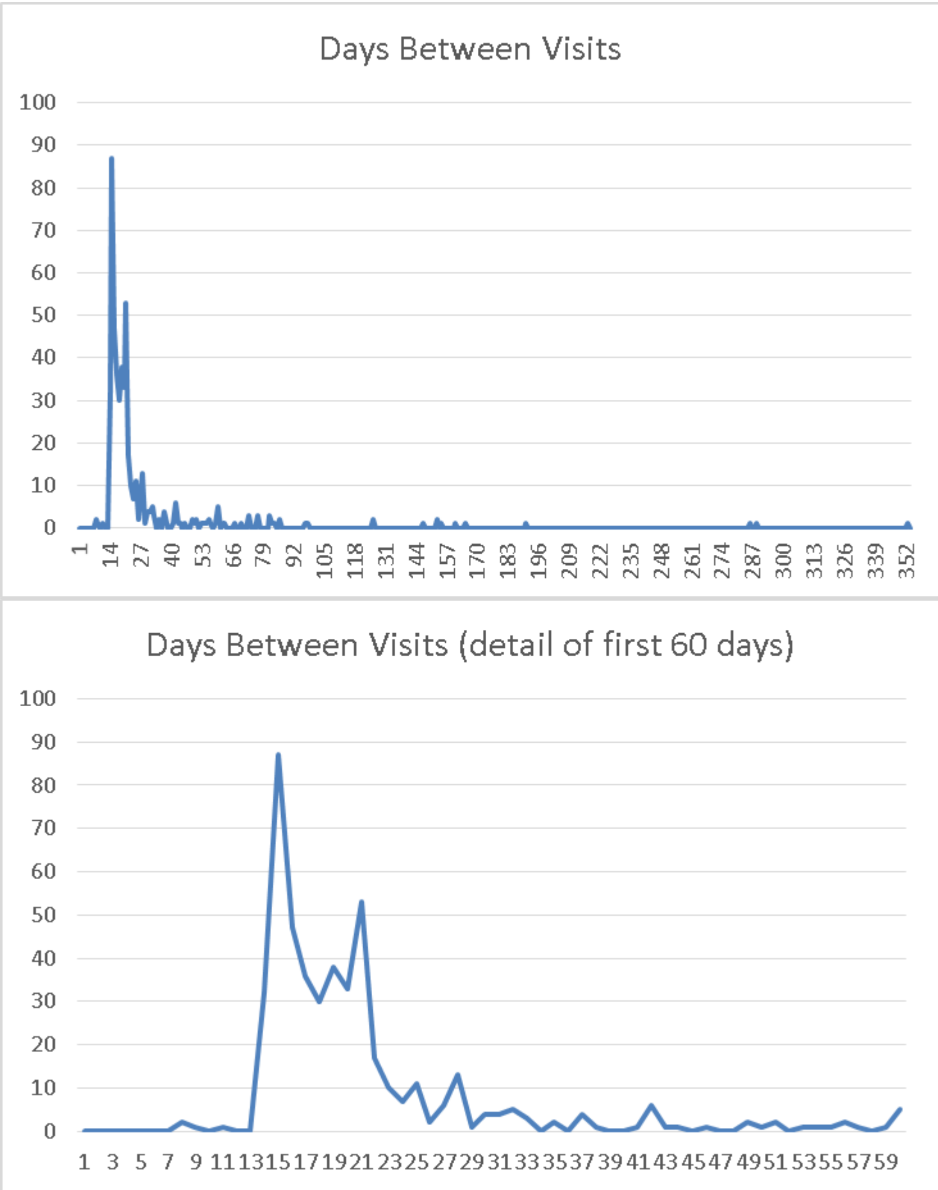


Figure 11 – Time Distribution between Visits

Considering the enrolment policy, the errors in enrolment are given in the following figure:

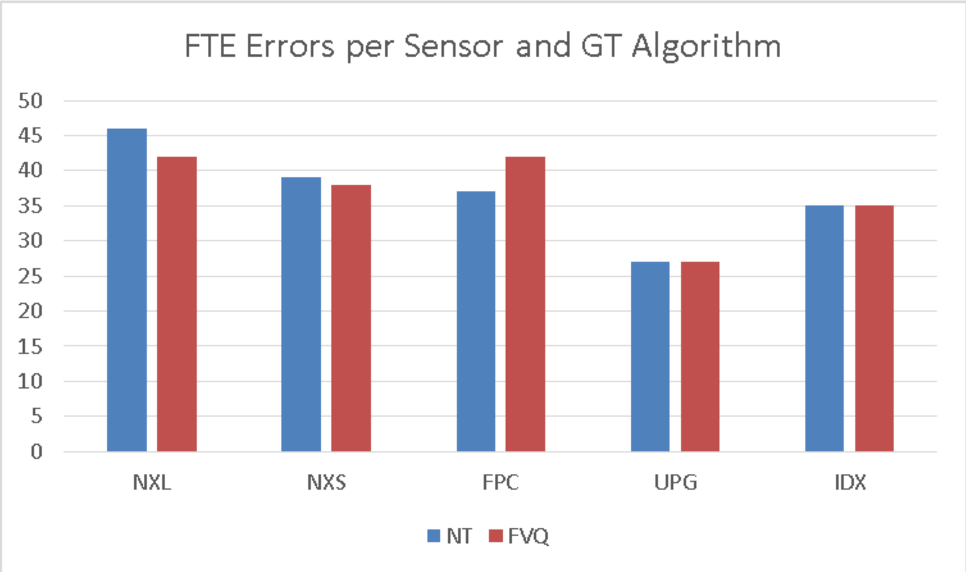


Figure 12 – Number of FTE per Sensor during DB Collection

4.2.1 Accepted Samples

The number of accepted samples per sensor and visit is given in the following figure:

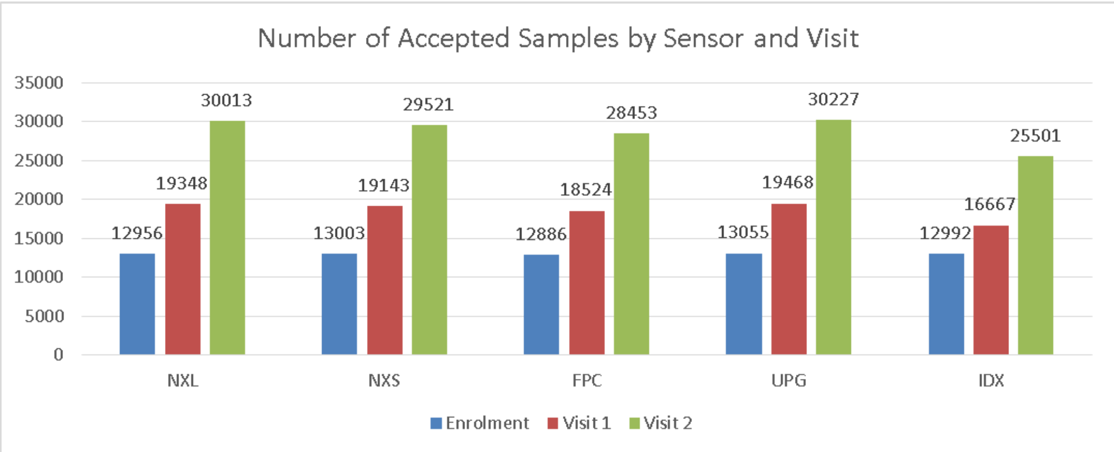
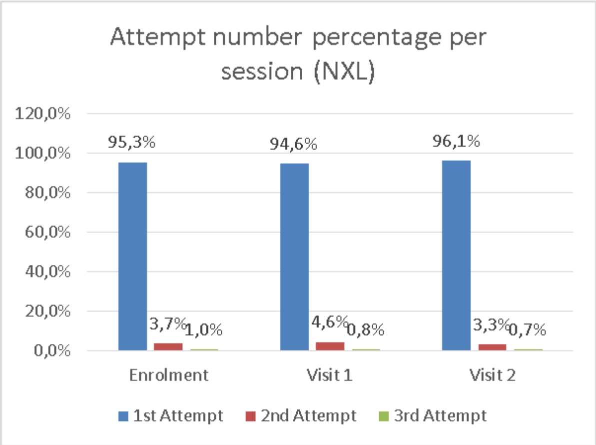


Figure 13 – Number of Accepted Samples per Sensor and Visit

The distribution of attempts needed to accept a sample can be obtained from the following figures:



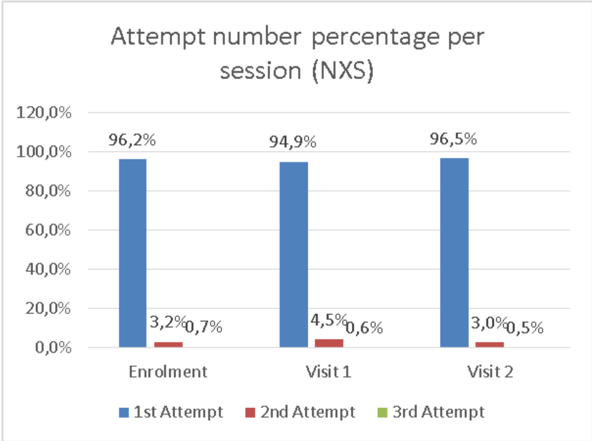
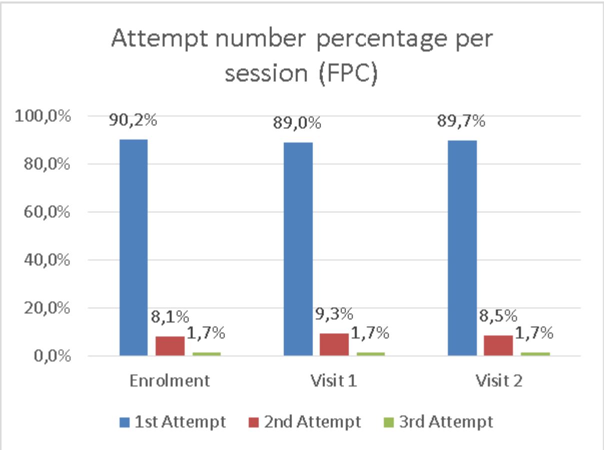


Figure 14 – Attempt number to accept a sample for sensors NXL and NXS



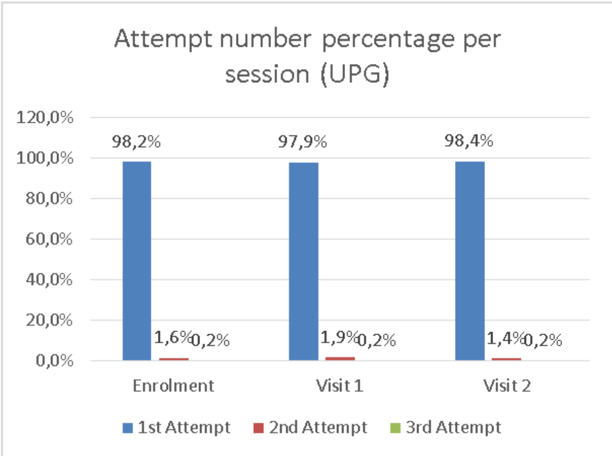


Figure 15 – Attempt number to accept a sample for sensors FPC and UPG

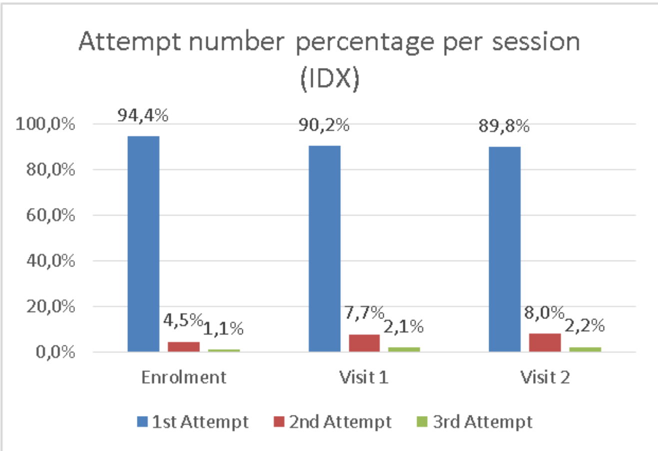


Figure 16 – Attempt number to accept a sample for sensor IDX

4.2.2 Acquisition Errors

In order to generate these accepted samples, the following errors have occurred:

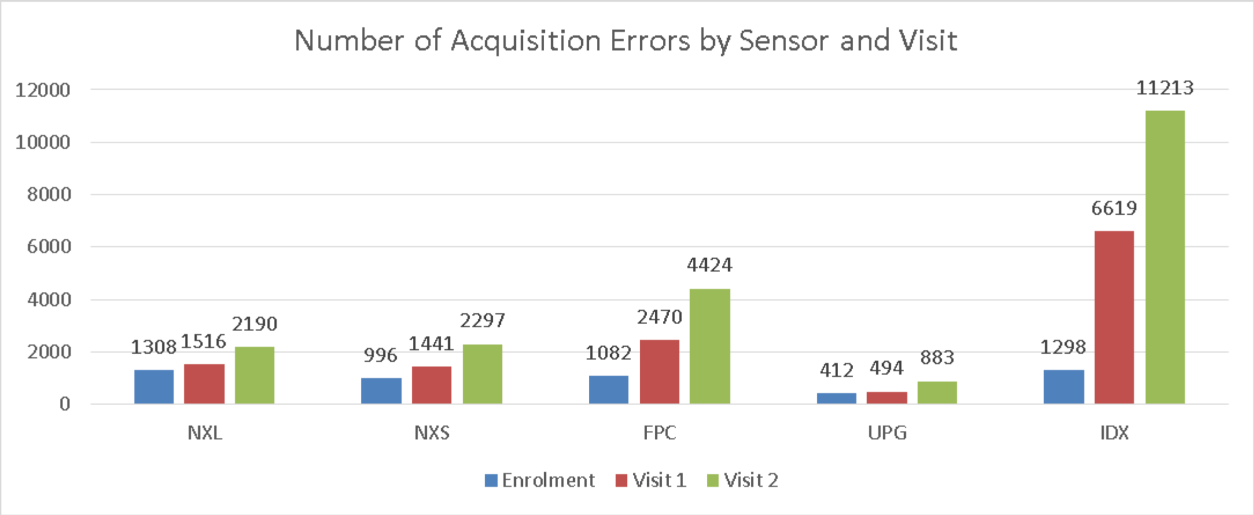


Figure 17 – Number of Acquisition Errors by Sensor and Visit

The reason for those errors and the comparison among each sensor can be obtained through the following figure:

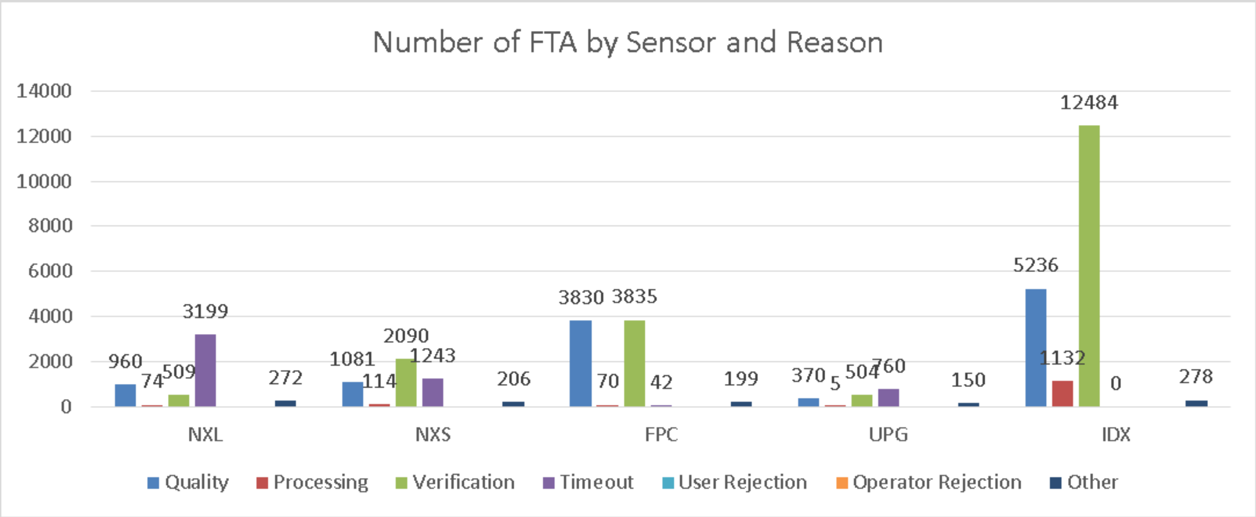


Figure 18 – Number of FTA by Sensor and Reason

The proportion of errors for each of the sensors can be analysed by using the following pie charts:

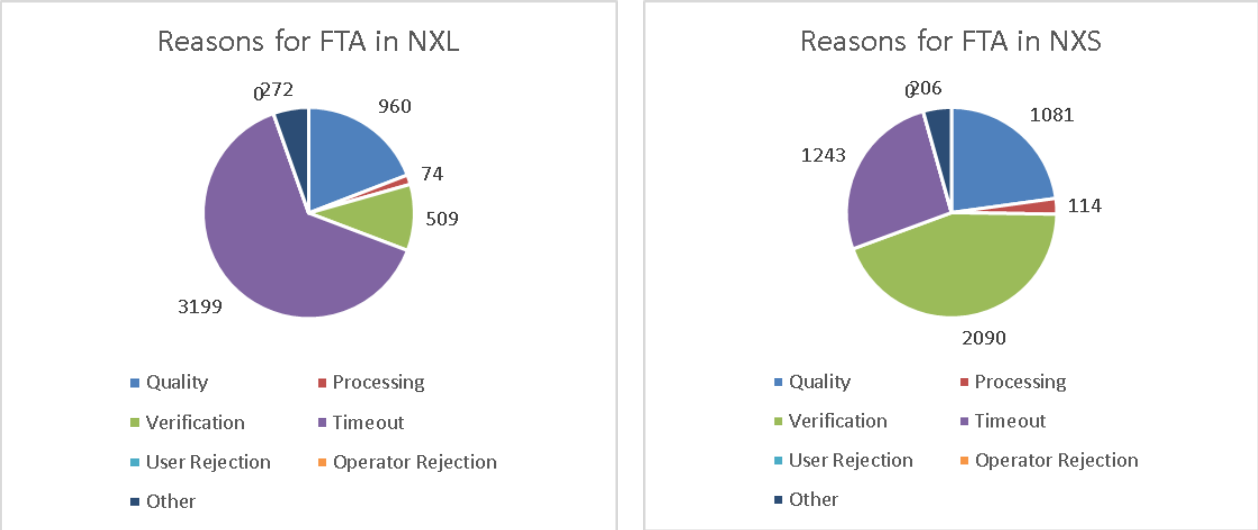


Figure 19 – FTA reasons for sensors NXL and NXS

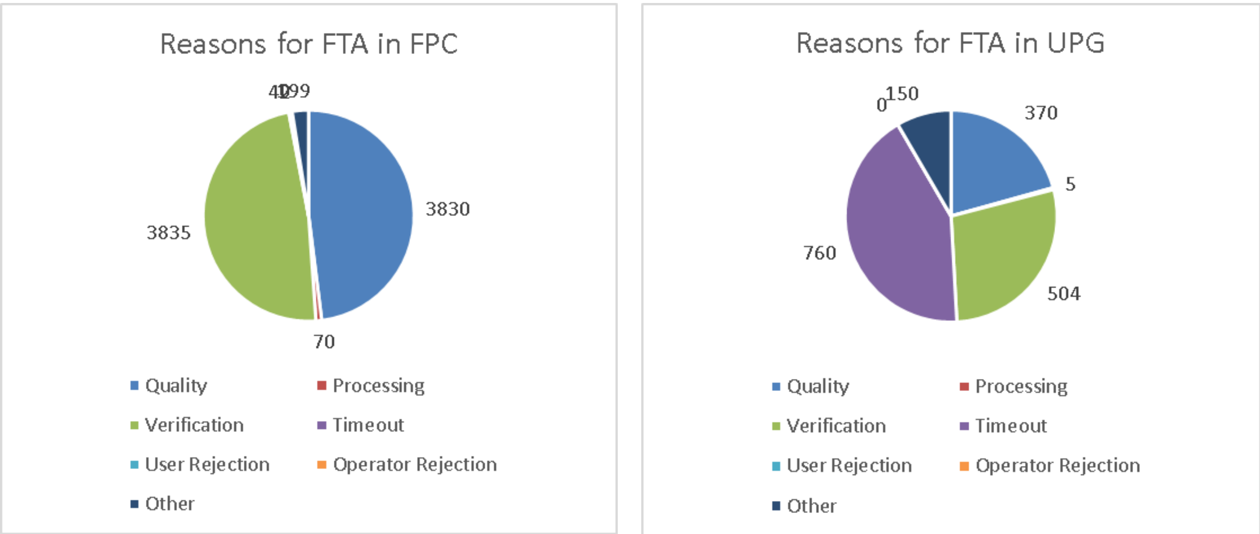


Figure 20 – FTA reasons for sensors FPC and UPG

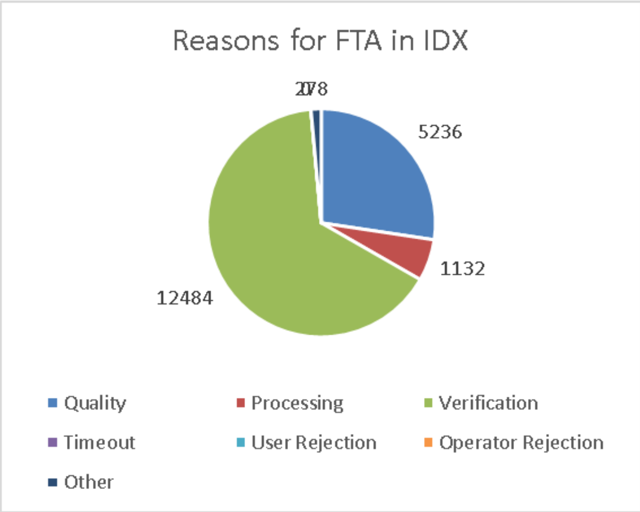


Figure 21 – FTA reasons for sensor IDX

Note: not all sensor vendor SDKs provided a reliable call-back mechanism to determine time-out cases during acquisition. Therefore, sensors NXL, NXS and UPG may suffer a relative disadvantage in the FTA scores compared to FPC and IDX.

4.3 Quality Analysis

As already mentioned, during the collection process, two quality algorithms have been used: NT and FVQ. This section provides the distribution of both quality scores, both for the case of accepted samples, and for the one related to errors. In addition, during post-process, NFIQ2 has been used. It has to be mentioned that NFIQ2 is not intended / not trained for samples coming from semiconductor sensors, as the implementation currently available is intended to be used with optical sensors with a resolution of 500ppi.

4.3.1 Accepted Samples

Quality score distributions for the accepted samples are presented in following figures:

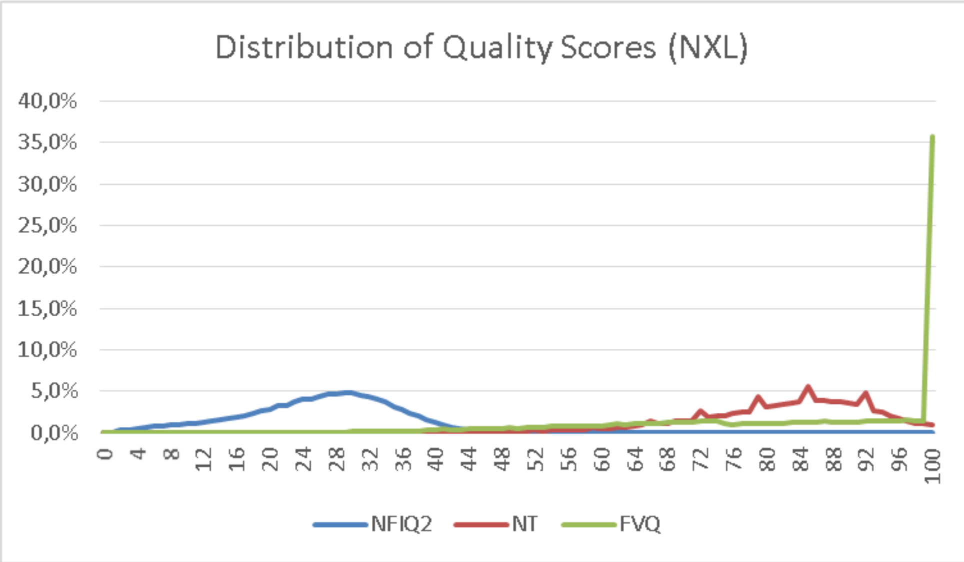


Figure 22 – Distribution of Quality Scores for NXL sensor

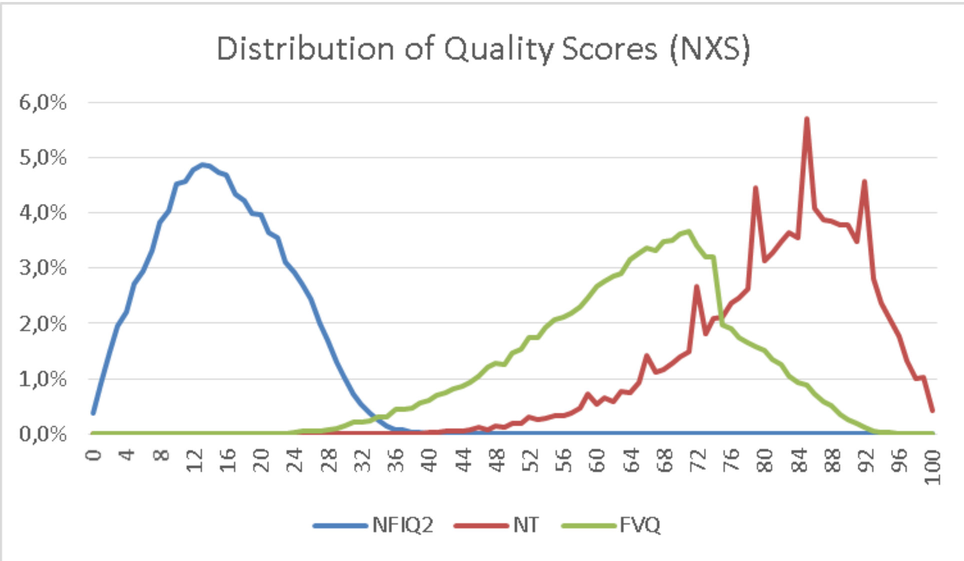


Figure 23 – Distribution of Quality Scores for NXS sensor

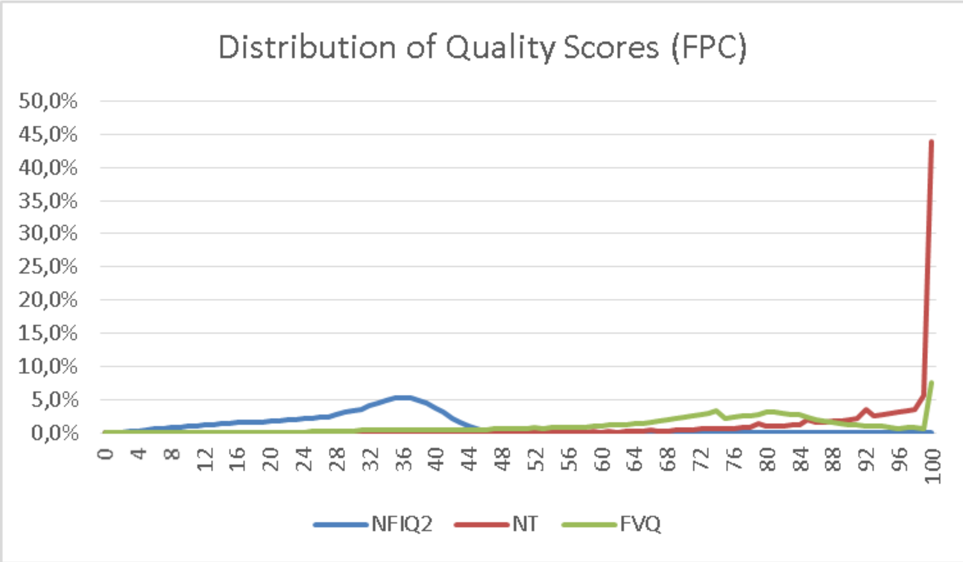


Figure 24 – Distribution of Quality Scores for FPC sensor

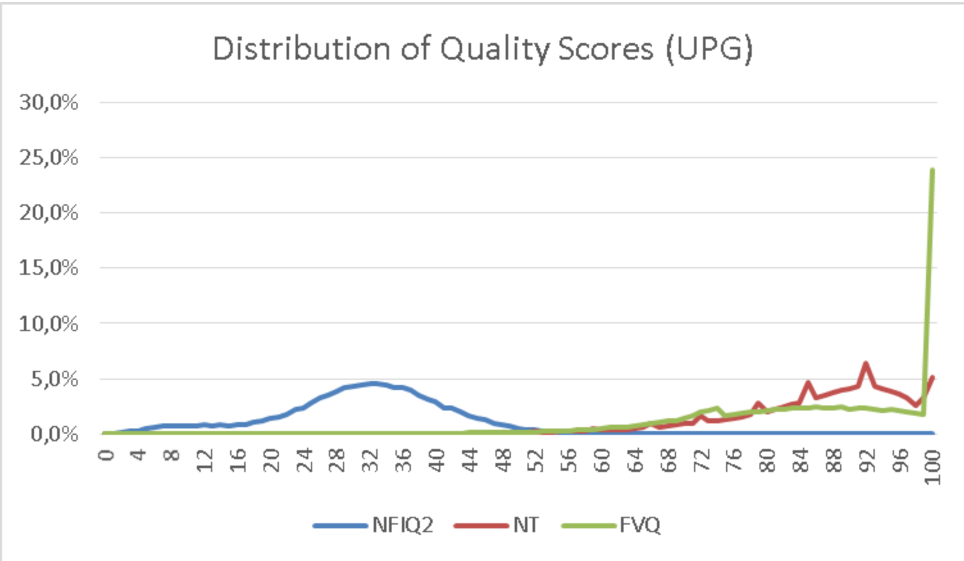


Figure 25 – Distribution of Quality Scores for UPG sensor

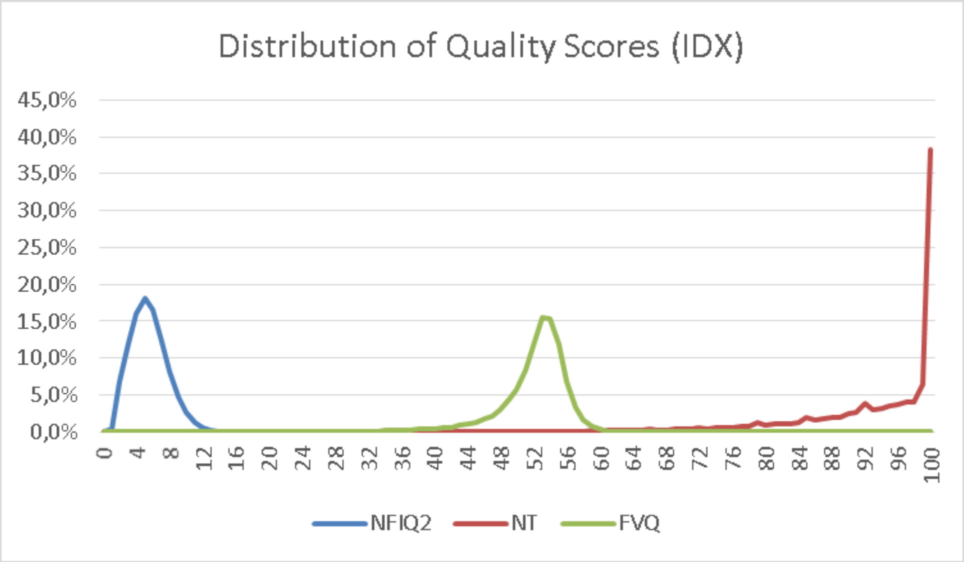


Figure 26 – Distribution of Quality Scores for IDX sensor

4.3.2 Errors in DB Collection

For the errors, the following figures show the quality score distribution, for those errors outside of the time-out ones. In these figures, only NT and FVQ algorithms are analysed, as NFIQ2 was used only during the post-processing (i.e., not at acquisition time).

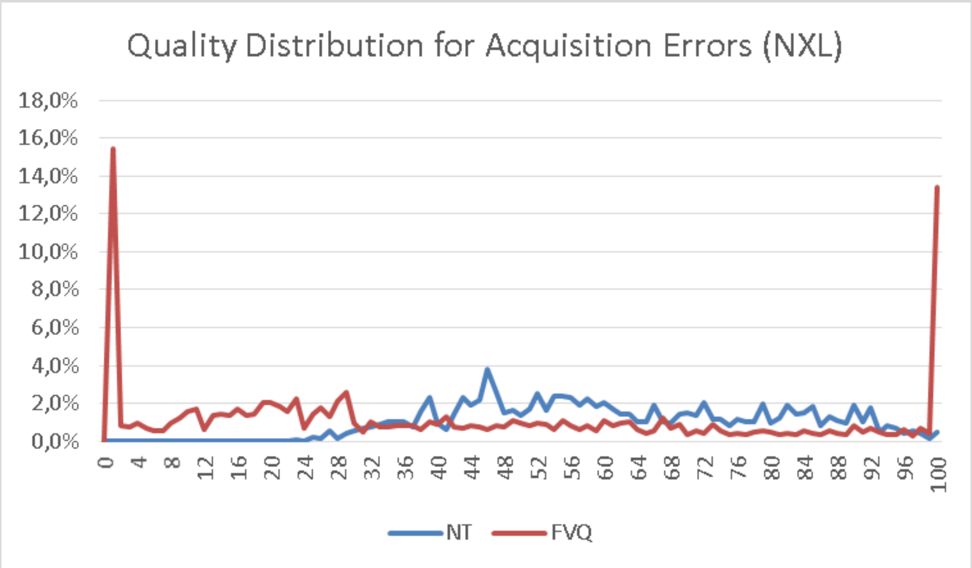


Figure 27 – Distribution of Quality Scores for Acquisition Errors in NXL sensor

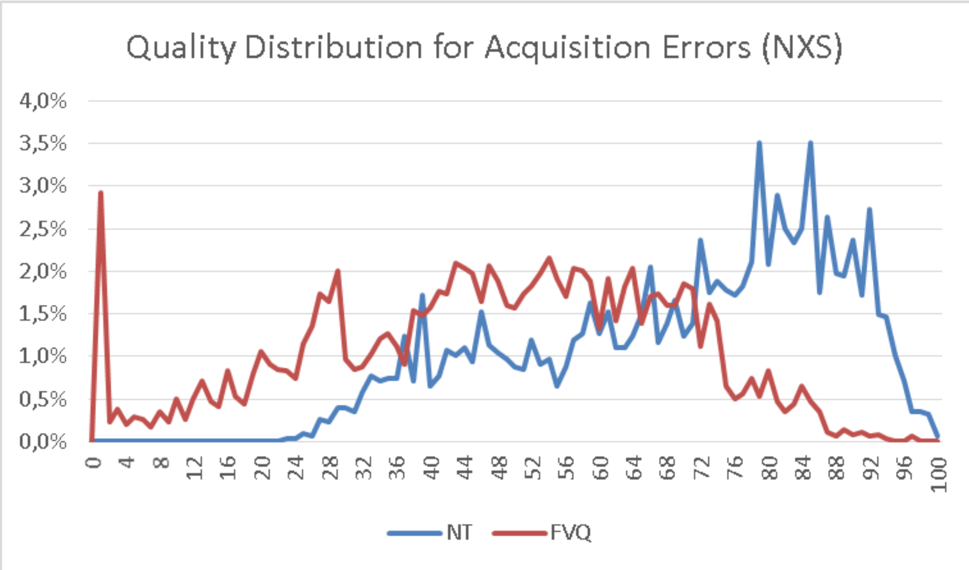


Figure 28 – Distribution of Quality Scores for Acquisition Errors in NXS sensor

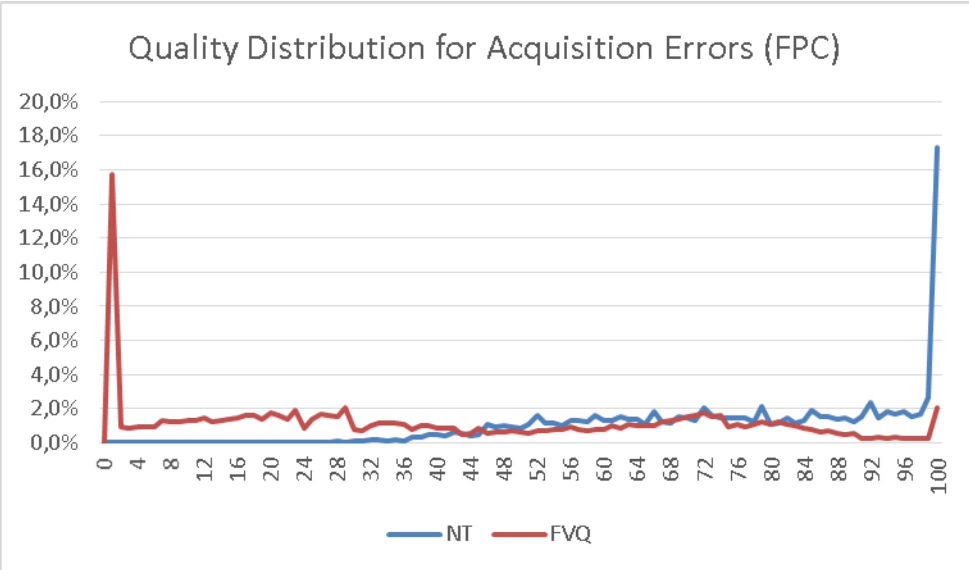


Figure 29 – Distribution of Quality Scores for Acquisition Errors in FPC sensor

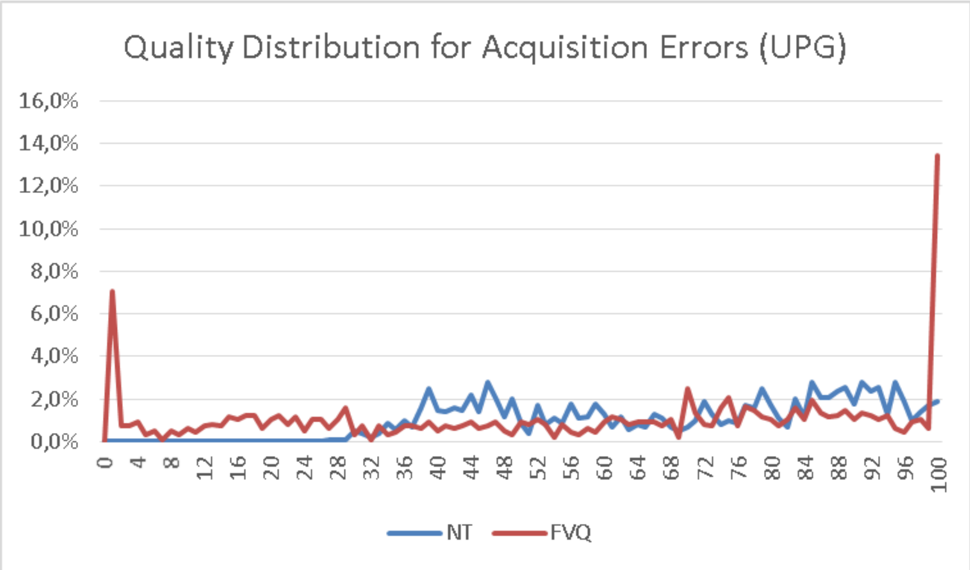


Figure 30 – Distribution of Quality Scores for Acquisition Errors in UPG sensor

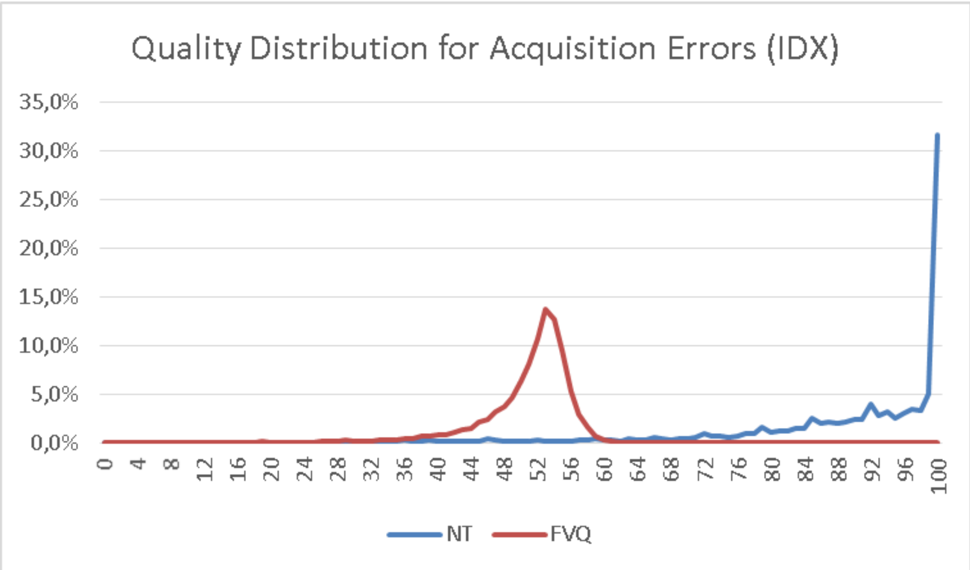


Figure 31 – Distribution of Quality Scores for Acquisition Errors in IDX sensor

5 Post-Processing and Results Generation

After capturing fingerprint image data, each of the evaluated algorithms has to be executed to extract features (i.e. algorithm specific template generation), and to perform all mated and non-mated comparisons. This Section describes how this process has been performed, as well as the impact of the whole process into the results generated.

5.1 Post-Processing Overview

After the lessons learned from the "Madrid Evaluation" the whole post-process phase has been redesigned. The main objectives searched are:

- As the execution of the whole set of non-mated comparisons takes extremely long, the post-process is designed in a way that it can be gradually executed. In other words, as soon as there are samples captured, it can be executed, and when new samples are captured, the comparisons are resumed, avoiding the executions of those comparisons already calculated.
- Also, to avoid double calculations, for each of the algorithms under evaluation, the biometric reference calculated with the enrolment samples will be calculated once and stored for future comparisons.
- Additionally, each sample will have to be compared with all different biometric references. Therefore, for each sample, the algorithm specific template will be obtained and stored, as to only capture this once.
- Intermediate data shall be stored and logged at all times to be able to react in case of any execution problem, such as a computer hang or a black-out.

With these objectives in mind, the post-processing follows, for each of the evaluated algorithms, these steps:

1. Biometric reference generation:
    a. For each subject and finger that has completed the enrolment session, it is checked if the biometric reference has already been computed
    b. If it exists, then skip to the next finger and/or subject

   c. If it does not exist, then calculate the biometric reference using all valid enrolment samples from that user/finger. Annotate the existence of the new biometric reference.
2. Samples' feature extraction:
   a. For each sample, check if the template has been generated for such sample with the algorithm under evaluation.
   b. If it exists, then skip to the next sample.
   c. If it does not exist, calculate the template of such sample and annotate its existence.
3. Comparisons:
   a. For each biometric reference:
      i. For each sample's template:
         1. If the comparison has already been computed, skip it
         2. If the comparison is new, calculate the comparison, and obtain the similarity score
            a. If the sample belongs to the same user/finger of the reference, annotate the comparison, the execution time and its score in a CSV file for mated comparisons
            b. If not, annotate the comparison, the execution time and its score in a CSV file for non-mated comparisons
4. Result calculations:
   a. Using the CSV files calculate the throughput data and the error rates, both in numbers and plots.

5.2 Biometric Reference Generation

During the post-processing of the Madrid-2 DB the biometric references are calculated in two different ways, depending on the algorithm capabilities.

For those algorithms with template-merging capabilities, all enrolment samples are merged and sent to the particular enrolment method of such algorithm, in order to generate a single biometric reference. *This is completely different from what it was done during the "Madrid Evaluation", in which the biometric reference was obtained by choosing the enrolment samples with the highest quality.*

If the algorithm doesn’t have such capabilities, one enrolment sample will be chosen as the biometric reference, using the following procedure:

1. Perform a cross-comparison operation between all the available enrolment samples.
2. Keep the pair of samples that achieves the highest comparison score.
3. From those two, select the sample with the highest quality, according to both NEU and FVQ algorithms.

This reference will be the one used to perform the mated and non-mated comparisons during the database processing. The following table summarizes, for each algorithm used during the DB processing, the procedure used to generate the reference.

Table 1 – Reference generation method per algorithm

	Reference generation method
NEU	Template merging

5.3 Impact on Results Reported

The use of a Ground Truth Verification (GTV) has a potential impact on the results obtained. By rejecting those samples that are identified of not being properly captured, or not corresponding to the tagged subject/finger, there is a lower limit that cannot be reached in the FNMR. This makes that the FNMR curve will be flat up to a certain similarity score, and from there FNMR will evolve properly. At the same time, FMR will present a regular curve.

This situation provides an impact on DET and ROC curves, as not all combinations of FNMR vs. FMR are present. This will be seen as a discontinuity in those two curves. For example, in DET curves, it may happen that the points $(0, 10^0)$ or $(10^0, 0)$ may not be reached within a continuous evolution of the curve.

Another potential impact is that results could be biased in favour of the algorithms used for the GTV. This will be more important as the number of rejected samples increases. But after analysing the results shown in Section 3, the percentage of rejected samples is minimal. For example, considering NXL sensor, the number of attempts excluding timeouts was 63 860, but the final number of samples are 62 317. 509 of those attempts were rejected due to obtaining a similarity score below the threshold. This means a 0.8%, which can be considered insignificant.

Regarding quality, 960 of those attempts were rejected for this reason, making it a 1.5% of the attempts. This is slightly higher than the rejections due to verification, but still a small percentage to consider the results to be biased.

Finally, the way the throughput statistics are calculated, provides a lack of precision due to being calculated using the operating system resources of the computers performing the post-processing. The operating system allows to measure time only in units of milliseconds, and the huge time needed for the whole processing of the database does not allow to execute each comparison a large number of times as to obtain a more precise calculation of the time.

6 Performance with Neurotechnology-FingerCell (NEU)

This section explains performance results when processing the database using Neurotechnology-FingerCell algorithm. In particular, error rates and throughput rates will be shown. Within this report this algorithm will be named as NEU.

Regarding error rates, these metrics are given separately for enrolment (FTE error) and acquisition process (FTA error). For the comparison process, the following curves will be shown:

- FNMR vs. FMR curves per each fingerprint sensor
- ROC curves for the five fingerprint sensors
- DET curves for the five fingerprint sensors
- Additional rates: EER, FMR1000, FMR10000 and FMR100000

Error rates obtained are reported in two different sub-sections. The first one will focus on the enrolment and acquisition rates (FTE and FTA) when applying the algorithm to the acquired and validated "Madrid-2 DB". The second one will focus on comparison rates, such as FNMR and FMR for the samples that overcome the acquisition process.

6.1 Enrolment and Acquisition Results

In addition to the FTE and FTA rates obtained during acquisition and reported in Section 4, the following two tables represent the FTE and FTA achieved with this algorithm.

Table 2 – Enrolment Error Rates using NEU

	NXL	NXS	FPC	UPG	IDX
Number of enrolled users in DB	3 300	3 300	3 300	3 300	3 300
FTE errors	48	71	240	20	20
Number of references generated	3 252	3 229	3 060	3.280	3 280
FTE rate	1.45%	2.15%	7.27%	0.61%	0.61%

Table 3 – Acquisition Error Rates using NEU

	NXL	NXS	FPC	UPG	IDX
Number of correct samples in DB	49 361	48 664	46 977	49 695	42 168
FTA errors	2 580	4 343	6 174	5 719	5 412
Number of samples processed	46 781	44 321	40 803	43 976	36 756
FTA rate	5.23%	8.92%	13.14%	11.51%	12.83%

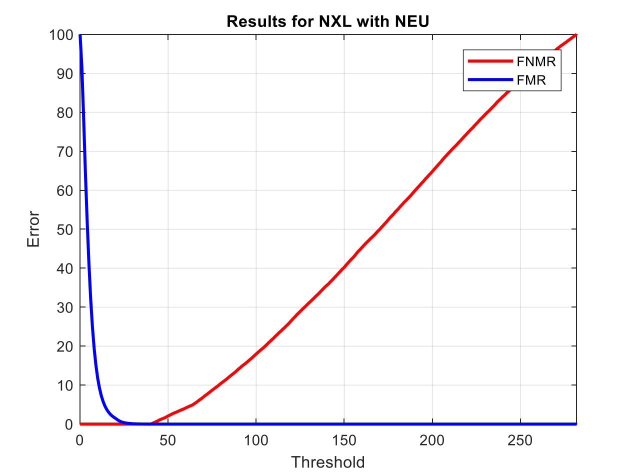
6.2 Comparison Results

With all the set of samples overcoming the FTA, the following table shows the number of executions of both mated and non-mated comparisons.

Table 4 – Number of comparisons conducted using NEU

	NXL	NXS	FPC	UPG	IDX
Mated comparisons	46 781	44 321	40 803	43 976	36 756
Non-mated comparisons	152 085 031	144 678 150	127 754 193	144 197 304	120 743 460

These comparisons provide with the following FMR vs FNMR figures. For each of the sensors the full figure is given, plus a zoomed version close to the EER (i.e. the crossing point).



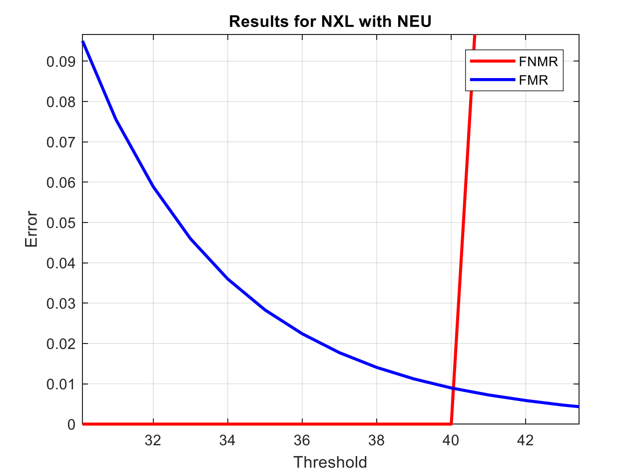


Figure 32 – FMR vs. FNMR curves for NXL sensor using NEU (Error in %)

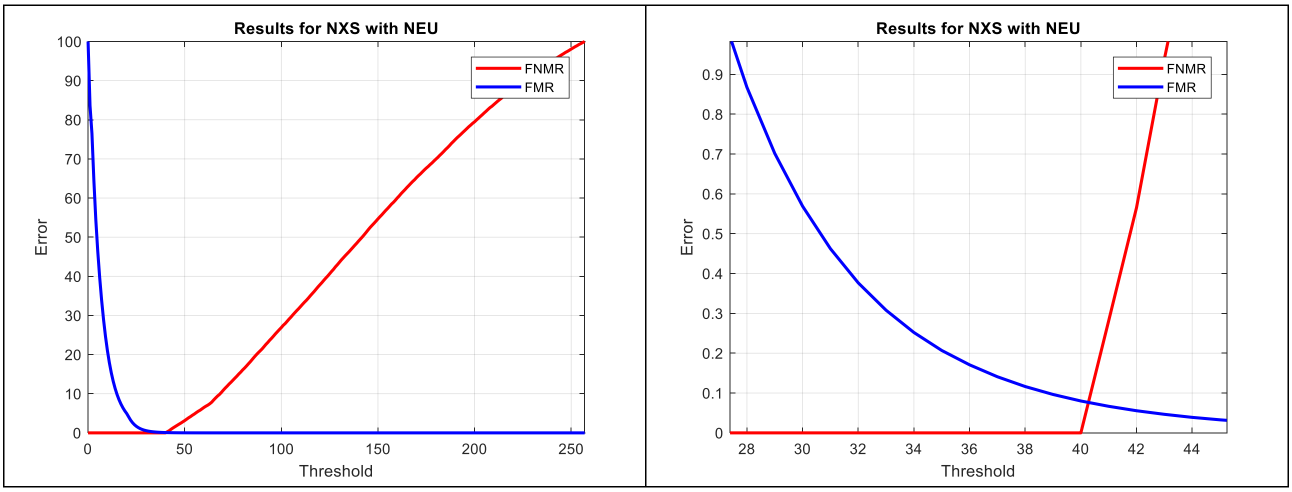


Figure 33 – FMR vs. FNMR curves for NXS sensor using NEU (Error in %)

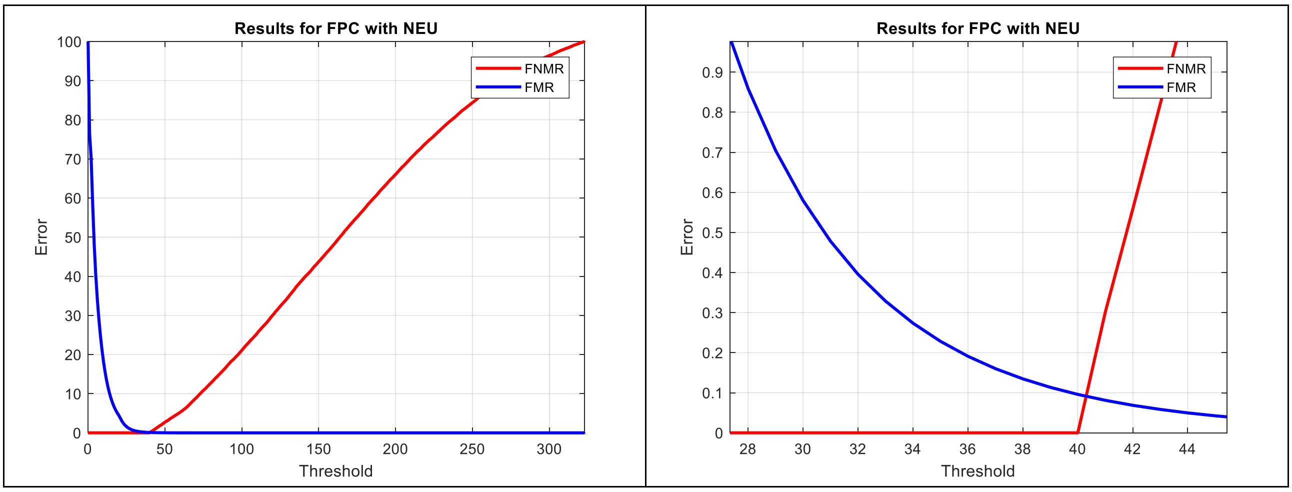


Figure 34 – FMR vs. FNMR curves for FPC sensor using NEU (Error in %)

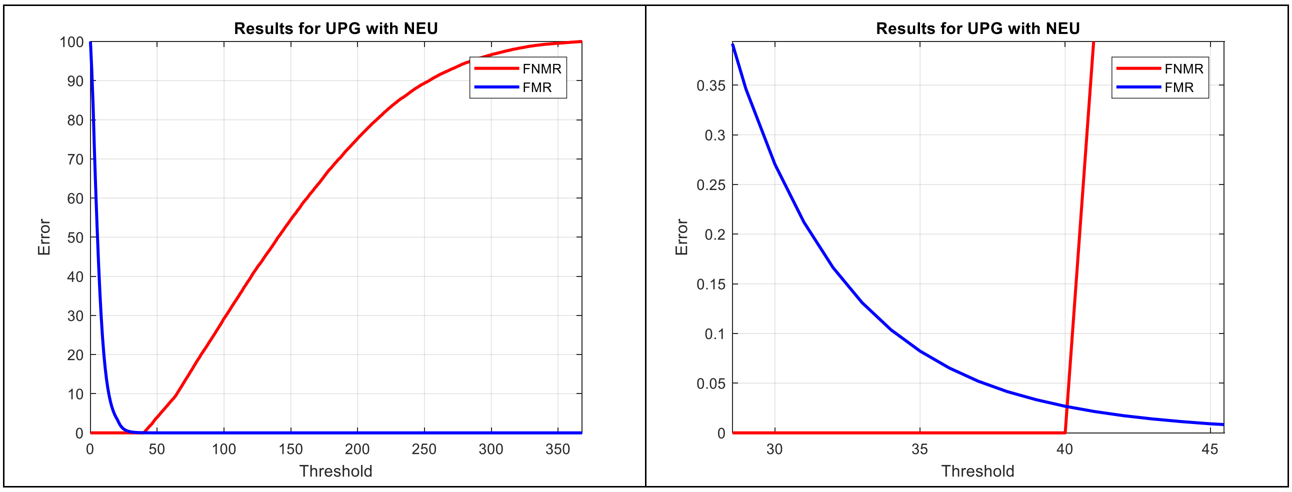


Figure 35 – FMR vs. FNMR curves for UPG sensor using NEU (Error in %)

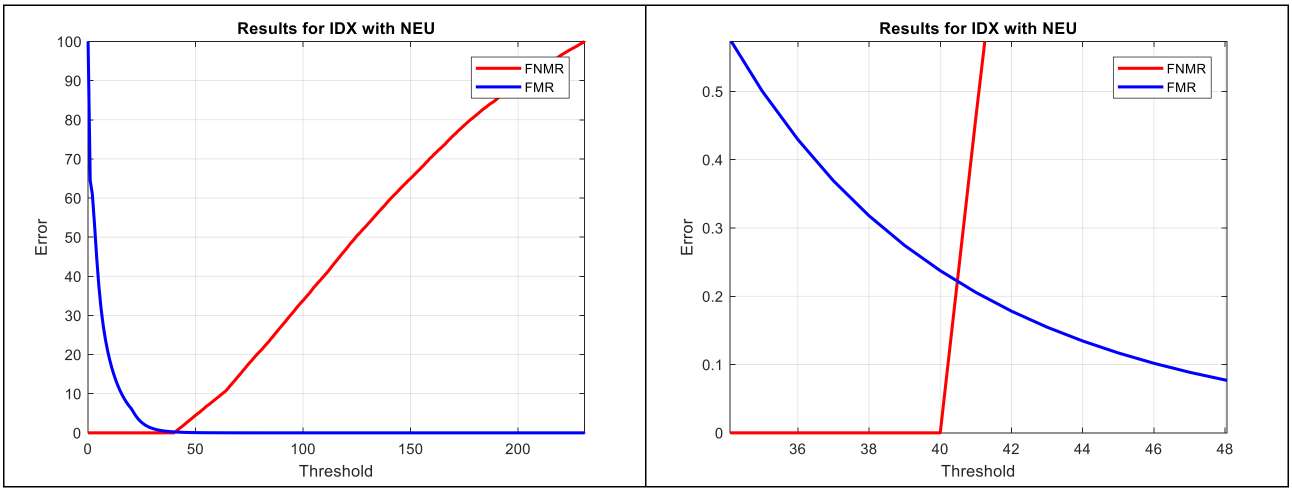


Figure 36 – FMR vs. FNMR curves for IDX sensor using NEU (Error in %)

These figures allow to calculate the DET and ROC curves. They are plotted in a combined way, so as to allow a better comparison among all sensors. Curves are presented in the whole range, as well as zoomed as to show them in the range of FMR between 10^{-4} (0.01%) and 10^{-6} (0.0001%). Both, DET and ROC are provided as to help the reader in understanding the results according to the plots the reader is more used to analyse.

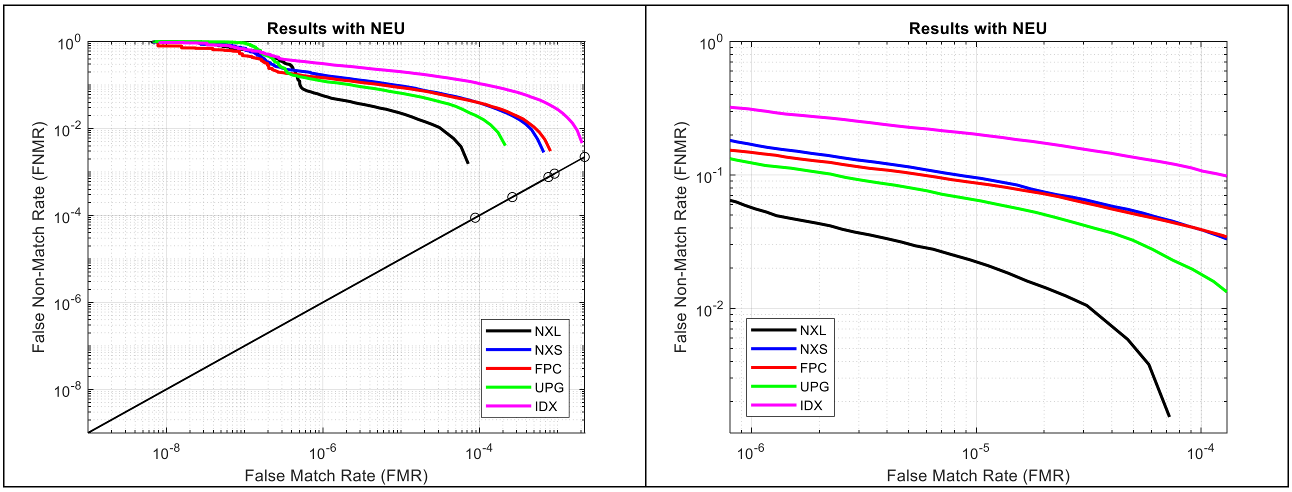


Figure 37 – DET curves for all sensors using NEU

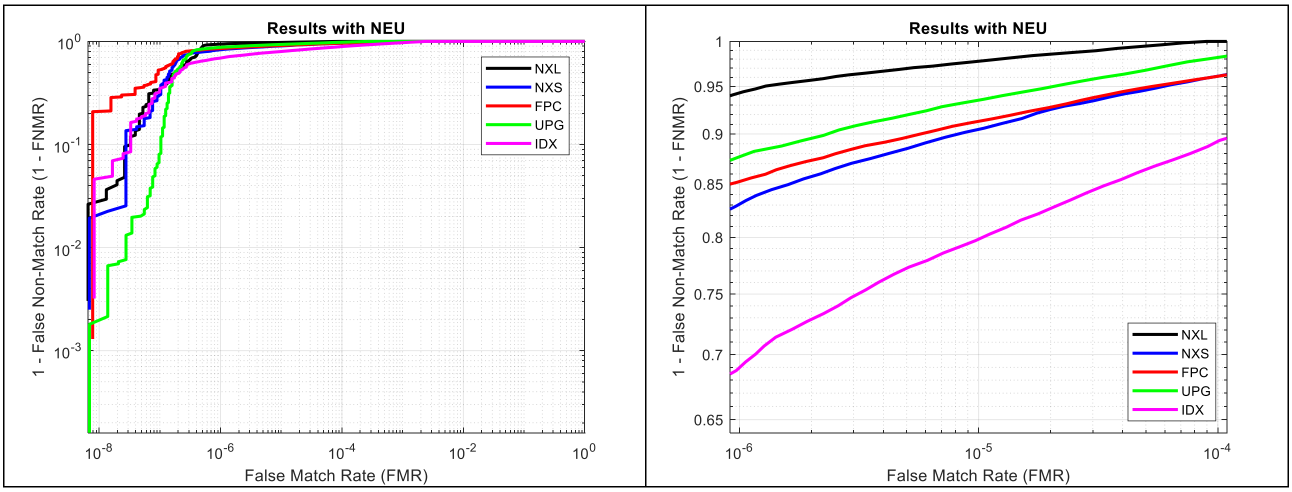


Figure 38 – ROC curves for all sensors using NEU

In terms of numbers, the following table provides the EER as well as the FNMR for the requirement of FMR being 0.1%, 0.01% and 0.001%, respectively.

Table 5 – Comparison Error Rates using NEU

	NXL	NXS	FPC	UPG	IDX
EER	0.009%	0.076%	0.091%	0.026%	0.222%
FNMR @ FMR=0,1%	0.000%	0.000%	0.000%	0.000%	2.682%
FNMR @FMR=0,01%	0.000%	3.846%	4.015%	1.963%	10.668%
FNMR @FMR=0,001%	2.314%	9.412%	8.827%	6.376%	20.304%

Results giving 0.000% FNMR, mean that for those conditions, no false non-match has been found using the evaluation database. With other conditions or a different database, it may be possible that some false non-accept comparisons appear.

7 Conclusions

This report has presented the results of an exhaustive evaluation of 1 (out of 7) fingerprint comparison algorithm and 5 fingerprint sensors. In order to achieve these results a new database, called Madrid-2 DB, has been acquired. The database has captured a total of 550 users. For each of these users, a maximum of 4 enrolment samples have been acquired, as well as 16 acquisition samples taken in 2 sessions with a time-gap between sessions of more than 15 days for most users. A total of 6 fingers were captured from each user (i.e. thumb, index and middle fingers from both hands). 510 users finished both acquisition sessions.

Considering each finger from each user as a single test subject, a total of 3 300 subjects have been enrolled. In terms of acquired samples, and after removing all acquisition errors, the total number of images is 236 865, being 49 361 for NXL, 48 664 for NXS, 46 977 for FPC, 49 695 for UPG and 42 168 for IDX.

Gender distribution is 52% for male subjects, while the age distribution is highly biased in the group of 18-24 year-old people, although more than 30% is for ages above 25 years old, including a 6% of people older than 50. Most of the people is right handed and even a higher majority do not present injuries in their fingers. 96% of the users are familiar with IT, being only 52% aware about biometrics.

Within the DB acquisition, a quality assessment and a ground truth verification has been used to guarantee that all samples are correctly tagged. Such mechanisms were able to detect some errors related to quality and to lack of possibility of verifying that it corresponds to the tagged finger. Those errors are completely insignificant for UPG and very low for NXL and NXS. Unfortunately, IDX presents significant errors in both quality and verification, reaching a 7% of erroneous attempts due to quality and 17% related to verification.

When analysing the quality of the acquired samples (i.e. those samples accepted), the results differ quite a lot depending on the quality algorithm used. With NT algorithm, the behaviour of NXL, NXS and UPG is similar and the accepted quality values are diverse and wide, while for FPC and IDX, the behaviour is extremist, getting all accepted samples a very high-quality score; anything with lower quality is rejected due to ground truth verification. Considering FVQ quality algorithm, the extremist behaviour is with NXL and UPG, being the behaviour with IDX non-deterministic.

The recently released NFIQ2 quality algorithm has been tested, even though the algorithm is only trained for optical sensors with 500ppi resolution. After the execution of the algorithm with all accepted samples, it can be confirmed that the algorithm is not prepared for semiconductor-based images, as the quality scores are always below 50 (out of 100) and even having many samples close to a quality score of 0, no matter whether the images are 500ppi (e.g. UPG) or lower (e.g. NXL).

Due to the low percentage of images rejected (except for IDX), the ground truth verification mechanisms can be considered validated, providing a minimal bias to the results achieved.

Focussing on the biometric algorithm included in this report, the following conclusions can be extracted. First of all, NEU algorithm presents different behaviour with each of the sensors used in terms of accepting the samples for processing. When the algorithm is used for enrolment, the FTE rate is reasonable low for UPG and IDX (below 1%), and somehow acceptable for NXL with 1.45%, raising above 2% with NXS, and increasing up to 7.27% for FPC. Therefore, the enrolment performance yield rate is not so much dependent on active area size (only one small sensor, FPC, is getting high error rates), but on the images acquired and the combination of the up to 4 samples used for enrolment.

Considering FTA, the error rates become higher for all sensors, behaving the best (i.e., lowest FTA) for NXL, and once again worse for FPC with IDX and UPG getting closer (i.e. in these three cases FTA was higher than 10%).

When analysing the results in terms of comparison performance, the DET curves show that NXL provides the best performance, followed by UPG at a significant distance. FPC and NXS present a similar performance, while IDX get a much worse performance. Considering the requirement of FMR=0.001%, NXL presents a FNMR as low as 2.3%, while all the others present FNMR>6%, with the very bad performance of IDX samples which get FNMR beyond 20%.

But if the initial requirement is FNMR=0.01%, no false non-match has been detected for NXL (i.e., FNMR=0%), and a rate below 2% is found for UPG, while the small area sensors get worse error rates, getting up to 10.7% for IDX as the worst case.

Last, but not least, it is important to detail why there are no throughput rates in this report. The reason is the lack of reliability on the numbers obtained due to the three factors. First, the algorithms have become fast enough so as to be able to perform comparisons in less than 1 millisecond. Second, the single comparison resolution provided by the PC platform is 1 millisecond, therefore not being able to get an accurate number for each individual comparison. Third, as the evaluation has been carried out under Windows 10, there have been a lot of internal thread interactions and pausing of the evaluation processes. These interferences became frequent enough so as to remove any significance to the timing results obtained.

This problem has been the major lesson learned during this evaluation. For future works, time will be computed independently of the comparison process. A low number of test subjects (e.g. 10 subjects) will be randomly selected, and with such low number both, the mated and non-mated comparisons will be executed a large number of times (e.g. 1000 times) counting the time spent for each of those multiple executions of each comparison. Dividing such time by the number or repetitions, an accurate timing is obtained, and with such number from all mated and non-mated comparisons for such small database subset, the throughput metrics will be calculated.

A final word on algorithms: This study evaluated a total of 7 different algorithms and 5 different sensors. Publication of algorithm performance results is subject to agreement with vendors. Therefore, only Neurotechnology FingerCell (NEU) results have been published in this report representing one of the top-performing algorithms. Other fingerprint algorithms can be categorized in high-performance and poor performance. Only two algorithms evaluated have shown poor results equivalent to at least an order of magnitude higher error rates compared to e.g. NEU. The other high-performance algorithms do behave different in detail while showing similar relative results compared to NEU.

References

[1] ISO/IEC IS 19795, "Information technology - Biometric performance testing and reporting", ISO/IEC, Geneva.

[2] Fernandez-Saavedra, Belen; Sanchez-Reillo, Raul. "Public Report on an Evaluation of 3 fingerprint sensors and 2 algorithms". IDTestingLab. May 11, 2015. http://idtestinglab.uc3m.es/data/_uploaded/Publications/PUBLIC%20REPORT_589%20users_Fingerprint_v1_1_release.pdf

[3] Fernandez-Saavedra, Belen; Sanchez-Reillo, Raul; Ros-Gomez, Rodrigo; Liu-Jimenez, Judith: 'Small fingerprint scanners used in mobile devices: the impact on biometric performance', IET Biometrics, 2016, 5, (1), p. 28-36, DOI: 10.1049/iet-bmt.2015.0018. Available at IET Digital Library, http://digital-library.theiet.org/content/journals/10.1049/iet-bmt.2015.0018