European Commission

JRC TECHNICAL REPORTS

Fingerprint Recognition for Children

Final Report

Institute for the Protection and Security of the Citizen

Digitial Citizen Security Unit

2013

Report EUR 26193 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for the Protection and Security of the Citizen

Contact information
Günter Schumacher
Address: Joint Research Centre, Via Enrico Fermi 2749, TP 36, 21027 Ispra (VA), Italy
E-mail: guenter-egon.schumacher@jrc.ec.europa.eu
Tel.: +39 0332 78 6085
Fax: +39 0332 78 5145

http://www.jrc.ec.europa.eu/

This publication is a Technical Report by the Joint Research Centre of the European Commission.

JRC 85145

EUR 26193 EN

ISBN 978-92-79-33389-7 (print)
ISBN 978-92-79-33388-0 (CD-ROM)
ISBN 978-92-79-33390-3 (pdf)

ISSN 1018-5593 (print)
ISSN 1831-9424 (online)

doi:10.2788/3086

Luxembourg: Publications Office of the European Union, 2013

Printed in Italy

Study on

Fingerprint Recognition for Children

Final Report

September 2013

Administrative Arrangement
ABAC No. 30-CE-0259474/00-10
JRC Contract No. 31216-2008-12 NFP ISP

European Commission Joint Research Centre
Institute for the Protection and Security of the Citizen
Digital Citizen Security Unit

Acknowledgement

The Joint Research Centre wishes to thank the Portuguese government, and in particular its Passport and Immigration Service (SEF), for the generous offer to access fingerprint data from its citizens for the purpose of this study. We also gratefully acknowledge the cooperative spirit of all people involved from SEF in helping us making the necessary arrangements.

We also would like to express our thanks to the research groups of Prof. Munk from the University of Göttingen and Prof. Maltoni from the University of Bologna for fruitful discussions and their technical assistance.

Table of Contents

List of Abbreviations

AFIS	Automated Fingerprint Identification System
BIODEV II	Field trials on operational aspects of VIS with regard to biometrics
BKA	Bundeskriminalamt (German federal police)
BMS	Biometric Matching System (of the VIS)
BSI	Bundesamt für die Sicherheit in der Informationstechnik (German federal agency for information security)
BVA	Bundesverwaltungsamt (German federal administration agency)
dpi	Dots (pixels) per inch
EER	Equal Error Rate
EP	European Parliament
FAR	False Acceptance Rate
FBI	Federal Bureau of Investigation
FRR	False Rejection Rate
FMR	False Match Rate
FNMR	False Non-Match Rate
FTA	Failure to Acquire
FTE	Failure to Enrol
FVC	Fingerprint Verification Competition
IQF	Image Quality of Fingerprint (see Appendix 1)
NFIQ	NIST Fingerprint Image Quality (see Appendix 1)
NIJ	(U.S.) National Institute of Justice
ROC	Receiver Operating Characteristic

RRR	Record Rejection Rate
SDK	Software Development Kit
SEF	Portuguese Passport Service
TAR	True Acceptance Rate
TNO	Toegepast-Natuurwetenschappelijk Onderzoek (Dutch research organisation)
U.S.	United States of America
VIS	Visa Information System

List of Figures

Executive Summary

The JRC study on the fingerprints of children

This report summarises the findings of a JRC study dedicated to the question of whether or not automated fingerprint recognition for children is feasible, that is, if the recognition rates obtained with this technology for children are similar to those reached for adults. If necessary, the minimum age threshold for the reliable use of automatic fingerprint recognition should be revised.

The question became relevant in the context of new legislation on security measures at European borders in which biometric controls have become an important element such as the EU passport or the Visa Information System. According to experiences gained with children in the context of initial experiments involving fingerprint verification, doubts were raised about the feasibility of such methods for this age group.

Despite such doubts, little knowledge was available about the reasons for the assumed higher error rates, at least for younger children, apart from speculations about the impact of the smaller structure size of children fingerprints or about the fact that the fingers are still growing at these early stages.

In 2009, the European Parliament called the Commission for clarification on this feasibility question. Soon it was realised that the study will face two major challenges in order to draw meaningful conclusions about the changes over time of children fingerprints:

- As the data which will be processed belongs to children, it was logically required to implement to most stringent safeguards in order to guarantee the highest level of care for preserving their rights.
- On the other hand, large amounts of data are required in order for those studies to be carried out in a statistically significant manner. Any quantity in the range of hundreds or even thousands of well selected test persons could already be helpful to present some initial findings, but tens of thousands of individuals would be required for a real performance analysis of state-of-the-art fingerprint recognition systems.

Eventually, a solution respecting these requirements was found with the determinant support of the Portuguese government which made available a large source of fingerprint data from children from their national repository of passport data[1].

---

[1] The JRC applied the highest security standards in dealing with this data source and received prior approval for the processing of these data from the relevant authorities.

These were not just single fingerprints (as this would be of minor value) but pairs of fingerprints of the same finger acquired during the first issuance and the renewal of a passport some years later. This unique feature of the database has provided the JRC with the necessary information to carry out this study and to address the questions related to fingerprints and children. These pairs of fingerprints, captured with some years apart from each other, allow for the first time to monitor the progression of the fingerprint patterns over time.

The major conclusions of the study

- **Growth has limited influence on fingerprint recognition.**

  Although the time difference was predicted to be the most important factor of child fingerprint recognition, all tested algorithms showed the same recognition rate regardless of the time between the fingerprints (of up to 4.5 years).

- **Size (in terms of the dimensions of the relevant fingerprint characteristics) does not constitute any theoretical barrier for automated fingerprint recognition.**

  Within the available investigation window of up to 4.5 years between the acquired fingerprints, there was no theoretical barrier observed for proper automated recognition by current matching algorithms – provided the images are of sufficient quality.

- **Image quality (in terms of low contrast and distortion effects) is the ultimate problem for child fingerprints, and image quality is strongly influenced by size.**

  Though the observed image issues are well-known also from adults, with the smaller structure sizes of child fingerprints, the issues get worse and the probability for it increase. Therefore, proper enrolment is the key factor for successful recognition.

- **Relevant quality metrics for fingerprints need revision with regard to the children case.**

  As far as quality metrics for fingerprint images assume feature dimensions for adults, adaption to child fingerprints will be necessary. Otherwise, the reported quality scores might mislead the acquisition process. The data and the result processed for this study can contribute to the revision of quality metrics accordingly.

- **Isotropic growth model may serve as a good approximation to cover changes over time.**

  The underlying data from Portugal with only two fingerprints per test item do not allow for a clean distinction of distortion from other effects. However, an isotropic model (i.e. linear growth of the fingerprint in all directions) seems to be sufficient to estimate the real level of impact that the growth effect has, if any.

- **Alternative acquisition devices for fingerprints should be seriously considered in the future.**

  Experiments with multi-spectral and touchless fingerprint capture devices, as well as with traditional devices with enhanced user guidance, gave promising indications on how the quality issues could be better managed – on top of already existing best practice guidelines for the improvement of quality in fingerprint acquisition.

These conclusions confirm that under appropriate conditions, ***fingerprint recognition of children aged between 6 and 12 years is achievable with a satisfactory level of accuracy.*** Further results from the study may help as well to quantify those conditions.

Recommendations

- **Image quality is the key.**

  A certain minimum level of training of operators and data subjects is necessary to acquire high quality images. Training needs to be designed for the particular setting in which the fingerprints acquisition will be carried out. Analysis of the context should be a strong prerequisite and guidelines for this purpose can be further elaborated and promoted.

- **Matching algorithms can be further improved.**

  Experiments with various versions of matching algorithms from a commercial vendor suggest that there is still some room for improvement, at least with regard to time differences well beyond 5 years which could not be investigated by this study. An earlier study had already demonstrated the benefits of such measures. Improvements can be made with respect to adaptations towards child feature dimensions and/or by applying the isotropic growth model. These improvements will need then to be tested and evaluated on a rigorous and fully independent manner.

- **Availability of relevant test data.**

  The important insights gained from the Portuguese data with respect to realistic automated recognition for child fingerprints emphasised clearly the need for long term availability of such data for relevant research and development. The key aspects of such a data repository would be permanence (as unique a EU wide reference), full compliance with security and data protection requirements and efficient usage (despite the security measures) with appropriate quality metrics .

- **Selection of acquisition devices.**

  Experiments with multi-spectral, touchless and novel four-finger capture devices, gave promising indications on how the quality issues could be better managed. These emerging technologies should be further explored

Remaining open questions

Despite the efforts of this study, some questions remain open due to the limitations of the available data:

- **Calibration with Adult Data.** The impact of the children specific aspects still need to be more clearly distinguished from the general quality degrading aspects. Therefore, the recognition performance of fingerprints of adults needs to be compared to that of child fingerprints where the adult data were acquired under similar conditions to those of the child data. This would allow to predict the performance loss for children in the absence of any particular compensation measure.
- **Enrolment Tests.** In order to quantify a practical age limit, given the best available technology, larger field trials on enrolment of children need to be conducted. These trials should further investigate and quantify the impact of certain enrolment devices and procedures.
- **Refined Growth Model.** The current results do not contradict the assumption of an almost isotropic growth model as suggested by an earlier study. At least, it seems suitable as a first order approximation for improving algorithms in cases where the time difference is not greater than the one considered in the present study (i.e. 4.5 years). However, it is desirable to draw conclusions for longer time windows (beyond 5 years) in order to give a clear message to developers of fingerprint recognition systems.

Operational Setting of the Study

On May 1, 2009, Administrative Arrangement (AA) No. 31216-2008-12, entitled "Fingerprint Recognition Study below the Age of 12 Years", entered into force between DG HOME and the Joint Research Centre (JRC). An Amendment signed on April 30, 2012, extended the activities of the JRC until August 31, 2012.

This document is the final report of the AA, including all findings and recommendations. It builds on and extends the work described in previous Interim Reports of the AA.

The report is structured and detailed according to the work plan as described in the Annex 1 of the AA, named "Specification for a Technical Feasibility Study".

This document was prepared by the Digital Citizen Security Unit of the Joint Research Centre:

Digital Citizen Security Unit
European Commission
Joint Research Centre, Ispra Establishment (TP.361)
Via E. Fermi, 2749
21027 Ispra (VA) – Italy
Tel. +39 0332 789515
Fax +39 0332 785145

Questions about this document should be addressed to:

Mr. Günter Schumacher
Tel. +39 0332 786085
Email: Guenter-Egon.Schumacher@jrc.ec.europa.eu

Chapter 1: **The Challenge and the Study’s Approach**

1.1 Introduction

Biometrics in general – and fingerprints in particular – have become increasingly important elements in the context of public security. Used historically for forensic purposes (but sometimes also for identification in the absence of any other mean), fingerprints are accepted as a mean to establish a unique link to a specific person. Modern electronic data processing allows for the possibility of comparing fingerprints very rapidly without the involvement of human fingerprint experts. Thus, the vast majority of existing security scenarios in which the identification of persons plays a key role can be potentially enhanced through the usage of fingerprints.

The proper usage of a technology implies precise knowledge of its limits and constraints. With regard to automated fingerprint recognition technology, there exist well-known limits related to *processing performance* (i.e. how fast it can be done), to *accuracy* (i.e. how reliable the result of a comparison is) and to *handling* (i.e. the level of expertise necessary for its use). However, there is also a limit with respect to *ageing*. Biometric identifiers (including fingerprints) have in common that they are based on physiological properties which may change over time. For the particular case of fingerprints, it is assumed that the characteristic pattern obtained from each finger is absolutely unique and unchanged for lifetime but at least the size of the pattern grows from childhood to adulthood [1].

Though fingerprint experts would be able, in principle, to compare the fingerprints obtained at different ages of a person, automated tools still need to be thoroughly tested under this scenario, especially when children´s fingerprints are taken into account. Before this study, there was no evidence that automatic fingerprint recognition systems would be able to correctly match samples of the same (juvenile) user acquired with several years difference. On the contrary, developers of fingerprint recognition systems already highlighted the existence of potential problems with child fingerprints.

1.2 The JRC study

1.2.1 Request from the European Parliament

When it came to relevant legislation in which automated fingerprint verification plays a role, there had been always a reservation with regard to children. Is it technically feasible to include age groups with potentially changing biometric features? In the absence of a clear scientifically based answer to this question, a supposedly “safe” age limit was usually introduced, as it was the case for the Visa Information System (VIS) where this limit was set to 12 years. Consequently, the European Parliament (EP) requested an investigation from the European Commission, and in May 2009 the Joint Research Centre (JRC) received the task to conduct it within an envisaged time frame of three years. The precise objective of the JRC study was:

> ***The Study's Objective[2]:***
>
> *The objective of this study is to carry out a thorough and integrated in-depth assessment of the technical feasibility of different age limits for fingerprint recognition - in particular of children aged between 6 and 12 years - in the context of large-scale databases. The study should give therefore an answer as to whether the change of size of fingerprints of this age group - related to the growth process of fingerprints – has a crucial impact on accuracy for verification and identification purposes.*

According to the Study's Objective, the focus was on age groups below 12 years even though there has never been a scientifically based justification of any age limit above 12 years at which the accurate use of fingerprint recognition for children can be safely assumed.

In simplified terms, the approach the study should follow was:

- review existing research results or related studies on the feasibility question (i.e., if under a certain age fingerprint recognition rates drop beyond acceptable values),
- close knowledge gaps by own investigations, and
- derive a clear statement about the feasibility question with respect to age limits.

1.2.2 The challenge about relevant data

It soon turned out that the knowledge and the available data to address the feasibility question was quite poor. Regarding pre-existing data, only few and in most cases inappropriate resources exist. In particular there was no data where the data subjects were tracked over a sufficiently long period of time (see the next chapter for a more detailed overview).

Generally, the recruitment of a set of (juvenile) test persons encompasses a series of difficulties. The required test individuals would be minors, some of them at a very young

[2] Excerpt from the Technical Annex of the Administrative Agreement between DG HOME and the JRC.

age. Appropriate data protection procedures need to be in place, approved by the data controller and the relevant authorities; the persons' biometric data need to be regularly acquired over a long time period; and there is a high risk that some of them would withdraw during the course of the study – the longer the study, the higher the risk (see section 2.2.3 about a study in the United States which confirms the relevance of that risk). Even if a significant number of persons would still be available, the minimum observation time needs to be large enough in order to have significant measurements (above noise) to draw robust conclusions.

1.2.3 The data from Portugal

Fortunately, the JRC received the offer from the Portuguese government to access fingerprint data of children obtained in the context of issuing passports. As Portuguese passports have a validity of 5 years for children above 4 years old and 2 years for children below, a second fingerprint of already registered children can be obtained once the child needs renewal of the passport. All these data is kept in a national register since the roll-out of the new e-Passports for which the fingerprint data is captured. Portugal offered in May 2010 to allow access to these data, provided such access would be compliant with the applicable data protection legislation.

However, there were some drawbacks from this approach:

- The approval process to access the data, necessary at national and European level, lasted some 18 months, i.e. half of the originally envisaged time frame for the study.[3]
- Although the processed data respects the specifications set out in the applicable regulation and in the European Commission decisions on minimum security features for passports [35], [36], [37], these rules only specify the data format and the resolution but very little on image quality. This has consequences on the scientific conclusions as will be explained in the next chapters.

Despite the mentioned obstacles and issues, the data from Portugal turned out to be of decisive importance for the success of the study. To the best of the author's knowledge, there is no similar dataset available for research on fingerprints.

1.3 Research Methodology

Apart from the review and analysis of the existing knowledge about child fingerprints and their automated recognition (chapter 2), the availability of the Portuguese data suggests the following 3-step investigation approach:

---

[3] This fact led eventually to an extension of the timeframe of the study by another year.

Step	Subject of investigation	Target result
1	To perform recognition experiments with current and prototype matching algorithms, targeted to analyse dependencies with respect to age groups and time difference between the acquired fingerprints. **See chapter 4.**	Differences in performance compared to adults. Vendors should be able to improve their algorithms based on test performed on the available data.
2	To analyse the growth effect by estimating the feature displacement in two corresponding fingerprints and to create a model to describe the feature displacement. **See chapter 5.**	The model for the change of the feature configuration within a fingerprint. This model shall be adoptable by fingerprint developers in order to improve the recognition rates when dealing with fingerprint samples captured with a large time difference.
3	To quantify issues related to the image acquisition through experiments with traditional and alternative acquisition devices. **See chapter 6.**	As the data from Portugal has been acquired under unknown conditions, further experimentation is necessary. These experiments shall enable a more precise quantification of issues during fingerprint acquisition and shall derive recommendations on best practices.

1.4 Structure of the report

Structure of the Report

Chapter 1 provides a more detailed description of the motivation behind this study and why it has been initiated. In particular, it describes the study's main objective.

Chapter 2 gives an overview of similar or related research activities and the results relevant for the current study.

Chapter 3 describes the Portuguese database which has been made available for investigation and testing.

Chapter 4 summarises the results and conclusions of the performance experiments carried out with different fingerprint recognition systems. It also highlights the crucial aspect of fingerprint image acquisition.

Chapter 5 describes the findings with respect to the growth effect, i.e. the modelling and quantification of the changes of the fingerprints over time.

Chapter 6 gives an overview of alternative fingerprint acquisition devices which are considered more appropriate for children.

Finally, **Chapter 7** summarises the conclusions achieved in this study.

The report also contains an **Appendix 1** on essential concepts related to fingerprint recognition, which may be found useful in order to better understand the different aspects discussed throughout the report.

Chapter 2: **Existing Knowledge and Related Studies**

2.1 Fingerprint Recognition and its Problem with Children

2.1.1 Fingerprints

A **fingerprint** is defined as a 2-dimensional impression (an imprint) of the friction ridges at human fingertips (Figure 1). There is usually the distinction between flat fingerprints (by simply touching a surface which should reflect the imprint) and rolled fingerprints (by additionally rolling the fingertip to the left and right side to increase the imprint). For the purpose of this study, only flat fingerprints are considered as this is the most common type to be used for authentication purposes like border control.

The recognition aspects related to these ridges are – for obvious practical reasons – studied on the basis of these impressions rather than on the fingers directly. However, this mapping of a 3-dimensional reality (the friction ridges) to a 2-dimensional image (the fingerprint) imposes already some problems to be taken into account.

Figure 1: Fingertip and fingerprint[4]

The usage of fingerprints is based on the assumption that they are unique, i.e. there are no two fingers of any human being on this planet which give the same fingerprint (see Annex 1 for more details on this uniqueness assumption).

[4] Source: Wikimedia Commons

2.1.2 Recognition process

With regard to fingerprint recognition, there exists the widely used distinction between **identification** and **verification**. Comparing one fingerprint against a reference sample in order to verify an identity claim (i.e., is this person who he claims to be?) is referred to as "verification". In contrast to this, "identification" refers to the process of potentially recognising a certain fingerprint within a large database of fingerprints (i.e., within this database, who does this fingerprint belong to?). The reader should be aware of the fact that this technical definition of identification may lead to confusion compared to its usual interpretation in establishing the identity of a person.

For the purpose of massive fingerprint identification, automated systems based on special hardware and software have been developed, called Automated Fingerprint Identification Systems (AFIS). Figure 2 depicts the automated process in more detail. Its main elements are:

- Acquisition of fingerprint image (by an acquisition device)
- Image processing (to enhance the relevant information)
- Feature extraction (to encode the found information)
- Storing and comparison of the fingerprints

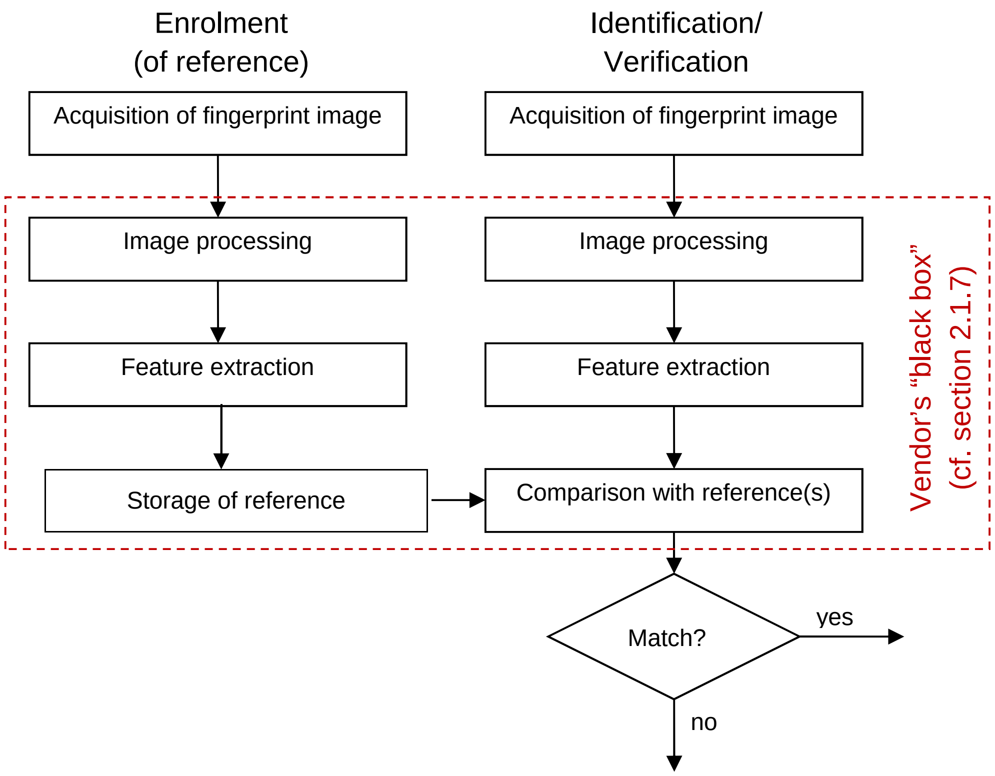

Figure 2: Fingerprint recognition process

2.1.3 Acquisition of fingerprint image

The well-known “ink and paper” approach, used previously by law enforcement, has been replaced by special fingerprint scanning devices (see Figure 3). The fingertip is placed on the specifically designed scanner and an image of the impression of the finger is produced. This approach usually allows only for flat fingerprints as the fingertip needs to be kept relatively still. Rolled fingerprints require a more sophisticated dynamic acquisition technology which is almost exclusively used for enrolling criminals.

Figure 3: Acquisition of a fingerprint from an index finger by a commercial scanning device

The result is typically an image with a resolution in the range of 500 dpi (dots/pixels per inch), with 256 greyscale levels and some 2-4 cm in horizontal and vertical direction. Other resolutions exist but are less widely used. **In any case, image acquisition is the most important step in the chain. Information loss at this phase can hardly be recovered in the following steps.**

It has to be stressed that the term “image of a fingerprint” is not identical to an image of the relevant part of the finger (apart from sides being flipped). As mentioned at the beginning of this chapter, the fingerprint is defined as the 2-dimensional impression (an imprint) of the finger on a flat surface. This impression is subject to a number of effects which we will describe briefly:

- Fingers can be subject to **conditions** such that certain type of acquisition devices may encounter problems to achieve proper quality. Critical conditions are usually extreme dryness or humidity, or the presence of certain substances (dirt, food remains, etc.). These conditions can lead to low contrast images where the clear identification of ridge lines (and thus fingerprint features) becomes very difficult.
- Depending on the way the finger is placed on the sensing device, certain types of **distortions** can occur. “Distortion” means that the ridge patterns of the imprints can vary in form and size from those in the real finger (see Figure 29 on page 61 for an example). This effect depends on the amount of pressure of the finger onto the device and how the pressure is distributed.

Figure 4: Wrong positioning of the finger on a scanning device
(too far turned to left side; too far to right side; too steep;
too much advanced; too less advanced)

- Another problem with image acquisition (apart from different pressure levels) is the **positioning of the finger** on the device. The finger is actually a 3-dimensional object from which different 2-dimensional imprints could be generated. For example, the right side of the fingertip gives a completely different fingerprint than its left side although both can still be considered as fingerprints of the same finger (cf. Figure 4). But such fingerprints, with different sides of the fingertip, would hardly ever be positively matched.

2.1.4 Image processing

While a human expert can directly analyse the features in the image, a computer program only "sees" pixels with different greyscale levels. It needs to find common structures starting from the single pixel level to certain clusters of pixels. In most cases, the first step is to eliminate low contrast effects in order to better distinguish between ridge lines and the space in-between (i.e., valleys). For this purpose, image processing techniques are used.

2.1.5 Feature extraction

Once the image is in its best form after being processed, common structures or patterns (called features) need to be found between the two compared fingerprints. Feature extraction usually starts with identifying ridge lines in order to follow them (pixel by pixel) until ends or bifurcations are found. To some extent, this step has similarity with a blind person “feeling” his way forward. Similarly to the blind person, the algorithm might be misled by artefacts.

The result of the feature extraction is a so-called “template” which mainly consists of a list of features, each of which is defined by its coordinates (relative to the image), its relative angle, its type (usually “bifurcation” or “ending” of a ridge line) and some confidence score.

It shall be recalled here what was already mentioned about distortion and positioning of fingers during acquisition. This translates into the fact that characteristic features of fingerprints of the same finger have hardly ever the same coordinates or angle even when transformed into the same coordinate system. Distortion leads to **different positioning coordinates** of the two **sets of features** to be compared. In the worst case, the intersection of both sets of features is too small or the deviation of coordinates is too big to allow proper recognition. Thus, computer based comparison has to take into account and deal with this type of “non-similarities”.

2.1.6 Storing of and comparison with the reference fingerprint(s)

The comparison of two templates after feature extraction is the most vendor-dependent part of the process and is based on a so-called **matching algorithm**. The common property of each algorithm is the fact that it determines a similarity score between the two fingerprints to be compared. Given the unavoidable “non-similarities” between samples of even the same fingerprint, as just explained, the score is a statistical measure to what extent two samples come from the same person despite those disturbing elements. For a given **score threshold**, two fingerprints “match” if the score has at least the value of that threshold, they do not match if the score is lower.

2.1.7 Error rates

The selection of the matching threshold is a very sensitive task. It determines implicitly the two statistical measures:

- **False Reject Rate (FRR)**, the probability that the computed score for two fingerprints coming from the same finger is below the fixed matching threshold (and thus falsely considered as a “non-match”),
- **False Accept Rate (FAR)**, the probability that the computed score for two fingerprints produce by different fingers is above the considered matching threshold (and thus falsely considered as a “match”).

The FRR and the FAR are interrelated. If the selection of the threshold is done in order to keep the FRR low, it will necessarily increases the FAR, and vice versa. The relation between FRR and FAR usually characterises the behaviour of an algorithm and is depicted in the so-called **Receiver Operating Characteristic (ROC)** diagrams (see section 4.1 and Appendix 1 for examples). Depending on the application (e.g., high security *vs* high convenience for the user), the matching threshold may be set to obtain different pairs of values FRR/FAR. For instance, in high security areas a FAR of 0,1% (i.e. on average 1 out of 1000 impostor attempts would be accepted) could be considered appropriate. The ROC diagram then tells what FRR is associated to that choice. Typical values for "good" systems operating on "good" (or "high quality") data at a FAR of 0,1% is a FRR of 1-7% (i.e. 1-7 out of 100 genuine access attempts would be denied). "Good" system refers to a current state-of-the-art commercial system; "good" data means that the fingerprints have been obtained with sufficient care and methodology.

This kind of vague definition is one of the major problems in the deployment of large scale systems. Vendors tend to present figures obtained with "good" data while in practice – due to various constraints and conditions – real operational data tends to be "less good".

Another characteristic error rate is the **Equal Error Rate (EER)** that will be used in later chapters. EER corresponds to the score threshold for which FAR and FRR are equal. EER permits to characterize a system performance with just one value and not a pair of them (FAR and FRR) as both are implicitly included in it.

It is important to note that in Figure 2 the steps "image processing", "feature extraction" and "matching" are in all commercial solutions encapsulated in a **vendor specific "black box"**. The results of the intermediate steps are not visible to the user. The "black box" receives images from an acquisition device, stores in a database the derived templates during enrolment in a vendor specific format and delivers as results either a simple "match"/"non-match" or just the matching score. Everything else is hidden for competitive reasons as vendor specific know-how. **This situation makes it difficult to understand the impact of the particular child aspects on the behaviour of the individual steps inside the "black box".**

2.1.8 The Case of Children

After the description of automated fingerprint recognition in general, the particularities that arise when having children as subjects for the fingerprint recognition process will be addressed. Taking the general scheme of Figure 2, children present some major differences that can be observed with respect to the case of adults:

- The **sizes of the fingerprint most relevant structures** (in particular the distance between the ridges) are smaller than for adults, down to one third. As already mentioned ridge patterns are developed during the human foetus phase and remain a life-long constant [1]. Fingerprints of children have been studied by a large number of scientific publications for various reasons[5]. The average ridge distance of newborns is about 0,15 mm, i.e. about one third of an adult, for a ten-year old it is already 0,3-0,35 mm. At least one scientific paper established a link between the ridge distance and the body growth, more precisely the seated height [7]. From the empirical information found, we estimate the average ridge distance roughly according to Figure 5.

The smaller dimensions could create various types of problems:

- There could be resolution problems with the acquisition devices, given the fact that they operate at one fixed resolution (see chapter 6).
- There could be problems with certain assumptions used by recognition algorithms with respect to average ridge distances (confirmed by several vendors). Such assumptions are used to decrease the computation time for processing fingerprints.
- The smaller distances between fingerprint features could decrease the ability of algorithms to deal with the "non-similarities" introduced by distortion effects and ill-positioning of the finger. However, this more complicated aspect will be further discussed in section 4.2.

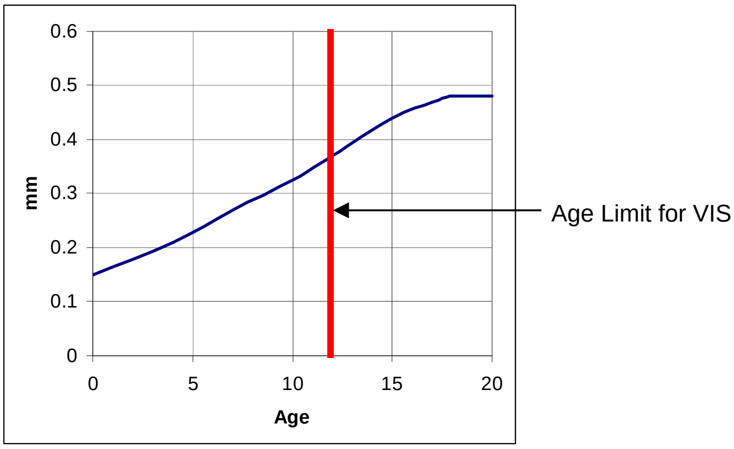

Figure 5: Estimated ridge distance vs. age of children

- Children can have a significantly **different behaviour** during the enrolment compared to adults due to lack of sufficient understanding of the process or simply due to their children-specific attitude. This behaviour can be considered as less

---

[5] Surprisingly, even an age determination method based on ridge breadth has been proposed [6] which estimates the age of a person (including children) with a mean error of 1,71 years.

"cooperative" with respect to the objective to obtain fingerprint images of adequate quality. Consequently, the problems illustrated in Figure 4 can occur more frequently.

- Fingers of children (and thus the corresponding fingerprint) grow at the same speed as the rest of the body. The **time elapsed** between the enrolment of the fingerprint and the acquisition of the test sample to be compared could be long enough to prohibit smooth matching. It was unknown so far (at least at the beginning of the study) to what extent this growth effect becomes relevant for recognition and whether algorithms take this effect into account.

Some studies have already confirmed these problems as will be explained in the next chapter. They suggest that fingerprint images taken under similar conditions to those of adults are statistically of poorer quality than those of grownups.

2.2 Related Studies

With regard to pre-existing knowledge or related research, the following activities have been identified as being the most relevant:

- A study done by **TNO** in 2004 on behalf of the Dutch Ministry of the Interior and Kingdom Relations in order to test the feasibility of including biometric identifiers in Dutch travel documents [17]
- The **BIODEV II** study performed between 2007 and 2009 by a number of European Member States in preparation of the biometric enrolment in the context of the then to be established Visa Information System (VIS) [33]
- A study conducted by **Ultra-Scan** and funded by the US National Institute of Justice (NIJ) between 2006 and 2009 which should elaborate a growth model for fingerprints of children [34]
- A study conducted between 2009 and 2010 by the **University of Göttingen** on request of the German Bundeskriminalamt (BKA), also focussing on the analysis and modelling of growth in fingerprints of children [28].

In the following, these projects, their scope and limits will be discussed along with their potential benefits for the current study.

2.2.1 TNO Study

The study was conducted in order to test the feasibility of including biometric identifiers in Dutch travel documents and focussed on the modalities face and finger [17]. With respect to children, a separate study was started which involved a number of 145 children between the age of 0 and 13 years, out of which 65 children were in the age group between 6 and 12 years. <u>It was concluded that fingerprinting of children below the age of</u>

4 is “virtually impossible”[6] due to the size of the ridge structures. Although it was not mentioned explicitly, it is assumed that all fingerprints (both for enrolment and verification) were done with a 500 dpi scanner.

Figure 6 displays the quality of the acquired fingerprints according to NIST’s Fingerprint Image Quality (NFIQ) metrics (see Appendix 1) with scores from 1=”excellent” to 5=”very bad”. The statistic shows a continuous decrease of quality along the age axis with the lowest values for elderly people above 65. In some contrast to the statement about infeasibility of fingerprinting below the age of 4, the statistics shows even for small children much higher NFIQ values than for adults above 50. However, this is due to the fact that good NFIQ values do not guarantee a high recognition rate.

Although the study admitted to the general feasibility of children fingerprinting from 4 years on, it has to be noted that the test persons and the overall scenario tends to be rather “cooperative”. Both the testers and the involved children have tried their best to achieve a good recognition rate which, on the one hand, still demonstrates the general feasibility but, on the other hand, neglects to a large extent the operational issues. This explains why the overall recognition rate for the mainly considered children age group (above 4 years) was almost 100%.

It also needs to be noted that the study did not address any significant time differences between the samples to be compared.

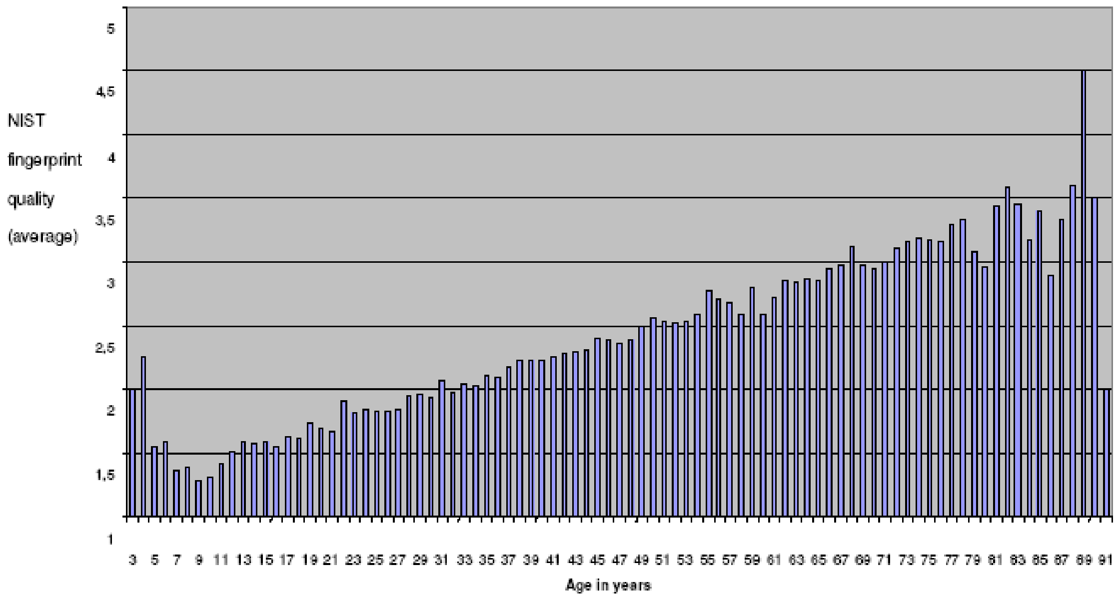

Figure 6: Average Fingerprint Quality by Age in TNO Study[7]

[6] See [17], page 25.

[7] Source: TNO

2.2.2 BIODEV II

This study (funded by the European Commission) was according to its title on "Experiment concerning the capture, storage and verification of biometric data for visa applicants conducted by Austria, Belgium, France, Germany, Luxemburg, Portugal, Spain and United Kingdom". The study encompassed field trials in selected embassies of the participating countries for the operational aspects of biometric enrolment and verification in the context of the Visa Information System (VIS).

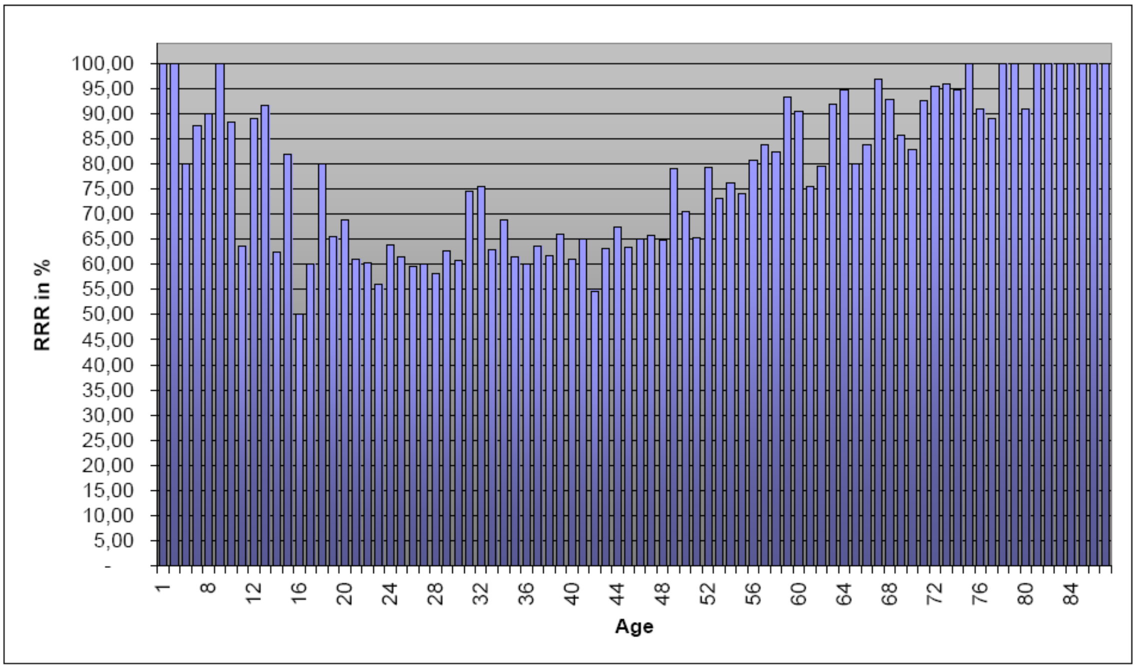

Figure 7: Record Rejection Rate in relation to age (German Damascus embassy)

Some results have been published by the German partner Bundesverwaltungsamt (BVA) who had test installations in Damascus (Syria) and Ulan Bator (Mongolia). After first experiments had been carried out on the Damascus site, relatively high Record Rejection Rates[8] (RRR) during enrolment in relation to age (Figure 7) had been observed. A record was rejected if it was not possible to enrol fingerprints of all ten fingers[9] at an NFIQ value between 1 and 3. Although the NFIQ values for all captured fingerprints show a typical distribution (Figure 8), the strong requirement on a complete data record led to a global RRR of more than 50% with an overall average of 70%. Furthermore, Figure 7

[8] RRR is special statistical measure introduced by that study in order estimate the performance of the system at person level rather than on individual fingerprint level.

[9] 4-finger acquisition devices were used which allow to capture four fingers of one hand at once and finally simultaneously both thumbs.

demonstrates the expected increased difficulties for children fingerprints (around the age of 12 downwards) and for elderly people.

On the other hand, the evolvement of this particular pilot demonstrated also in a typical way the problems to be encountered during operational testing: From a starting RRR of 70% the percentage was finally brought down to about 30% due to additional measures introduced for enrolment (training of users, adoption to environmental conditions, scanner position, etc.) and by using the appropriate quality metrics provided by the vendor of the Biometric Matching System. At a later stage of the testing, the vendor provided new quality metrics which further decreased the RRR to about 3% (for persons of at least the age of 12). Thus, enrolment of high quality fingerprint images require thorough understanding of the issues involved and the development of relevant best practice guidelines.

Finding:
Best practice is key

Enrolment of high quality fingerprint images require thorough understanding of the issues involved and the development of relevant best practice guidelines.

Fortunately, although not imposed by the VIS regulation, the German BIODEV II partner also enrolled children for test purposes. The result is a collection of fingerprints of about 300 children of non-European origin below the age of 12 years. Because all ten fingers were recorded, this amounts to a total of almost 3000 children fingerprints, each of them even in several versions of different quality. However, for each finger registered, there was only one fingerprint sample available.

2.2.3 Ultra-Scan/NIJ Study

In spring 2010 the JRC became aware of a study funded by the National Institute of Justice (NIJ) in the US about "Quantifying the Dermatoglyphic Growth Patterns in Children through Adolescence" [34]. The study was conducted between 2006 and 2009 by Ultra-Scan[10], a major U.S. fingerprint systems company in the health care and security market. The full report became available in December 2010.

The objective of the study was to determine whether a commonality of growth exists and to develop a mathematical model for predicting this change. The minutiae displacement seemingly did not follow the predicted linear (isotropic) transformation. There are serious reservations about this result. Although a 5 year observation period was originally envisaged, the major part of the 300 test subjects were followed over 2 years only. Moreover, and although the conductors of the study had the technical capabilities to do

[10] http://www.ultra-scan.com

better, the fingerprint images were acquired using 500 dpi scanners only. The ridge distance growths within 2 years roughly 0,04 mm (cf. Figure 5), which corresponds to less than 1 pixel. Therefore, potential differences in the fingerprints can hardly be measured precisely and distinguished from other effects like distortion. Even by using additional statistical tools to get rid of noise, it comes with no surprise that the predicted correlation could not be demonstrated with the data material in hand.

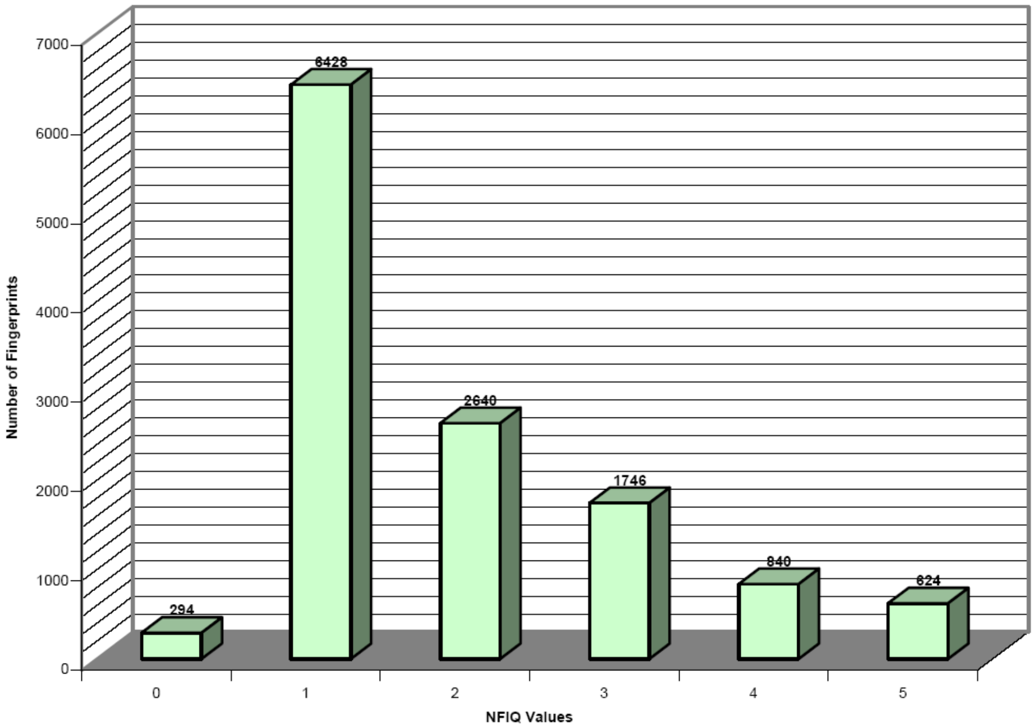

Figure 8: NFIQ values (German Damascus embassy)

2.2.4 U Göttingen/ BKA Study

The German Bundeskriminalamt (BKA) funded an investigation about the feasibility of recognising fingerprints of children after a certain number of years. The study was conducted by the Institute for Mathematical Stochastics at the University of Göttingen (Prof. Axel Munk) between 2009 and 2010. For this study, the BKA provided fingerprints of 48 reoffending juveniles that have been fingerprinted in criminal records between 2 and 48 (on average 4.5) times. All images were taken with 500 dpi.

The chosen approach to align the images taken at different ages was even simpler than in the Ultra-Scan/NIJ study but nonetheless quite efficient. Based on the assumption that the ridge distances follow the same growth chart as the body length, the researchers tried to modify the geometry of the level 1 and level 2 features of the older image in order to match with the newer one [28]. With the help of so-called shape analysis techniques[11] [29], they analysed the probability of an isotropic (i.e. linear in all directions) growth of the pattern.

The results for the mentioned age group are relatively positive, even if they also suggest that modifications of the algorithms might already be necessary for children above the age of 12 years.

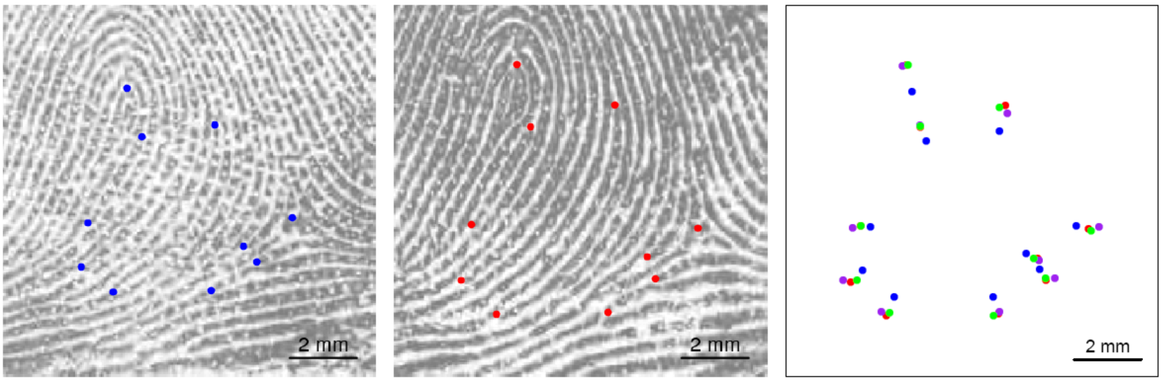

Figure 9: Example results of alignment for fingerprints of different age (Source: [28])

Figure 9 gives a typical example of the proposed alignment of fingerprints of different ages. The blue dots in the third picture are those of the older image, the red dots those of the newer image which also serves as the reference. The purple dots have been calculated by applying the linear transformation according to the predicted growth curve. In addition green dots have been added as refined positions for the red ones by eliminating the effect of distortion from the image through another statistical analysis.

This approach has been performed to all possible pairs of fingerprints at different times of the same finger. The performance of this approach is depicted in Figure 10. The leftmost dotted line depicts the mean distances of the pure distortion due to pressing the fingertip on the acquisition device. This corresponds to the corrected green dots in Figure 9. The graph shows the empirical cumulative distribution over all available images. The middle dotted line gives the distribution of mean distances of the corresponding red and blue dots according to Figure 9. Finally, the rightmost dotted line gives the distribution of mean distances of the corresponding purples and the reds in Figure 9.

[11] Shape analysis was applied to the so-called Delaunay triangulation of the fingerprint features, i.e. a mesh of triangles which have the feature coordinates as its corners. Such a mesh is usually used by fingerprint recognition algorithms.

The interpretation of this graph is as follows: The best practical performance is characterised by the leftmost dotted line because it is caused only by the distortion generated by the acquisition process itself. The "distance" of two images compared with the new approach comes very close to this leftmost dotted line while the "distance" of images compared by a normal algorithm (i.e. not adapted to ageing) has a completely different (i.e. worse) profile.

However, this positive result has to be seen under the aspect that mainly data of persons above 12 years was considered. The study had not significant data for the age group below and was thus unable to create a statement directly applicable for the main objective of the JRC study.

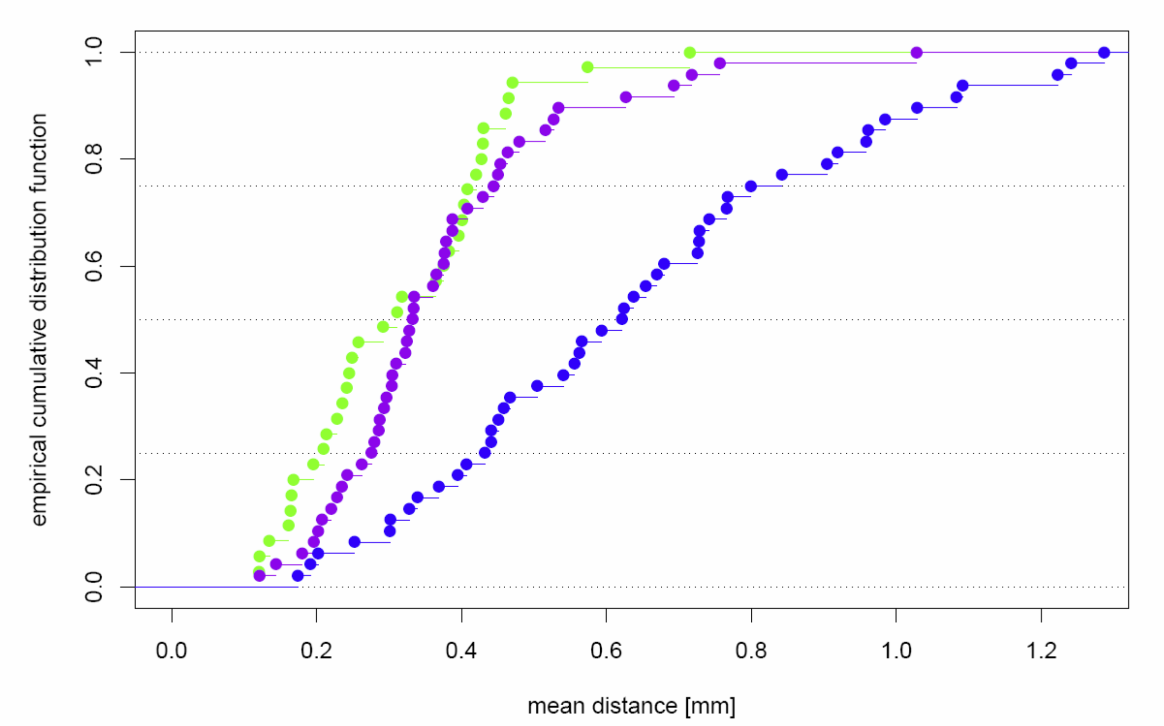

Figure 10: Performance of the linear approach
(Source: [28])

The statistical method they applied has several advantages relevant also for the current study:

- The performance graph according to Figure 10 can be immediately generated for any set of data with an ageing profile. This allows a relative assessment against traditional algorithms rather than an absolute assessment. It simply shows how a methodology to align images of different age compares to "ideal" matching (i.e. only dealing with acquisition distortions) and to matching without consideration of ageing.

- The methodology is open to replace the alignment approach by any other one. This helps to get a quick performance profile of any new approach to predict the displacement of the fingerprint features if this would become necessary.

2.3 Summary conclusion from pre-existing information

We summarise the conclusions drawn from all mentioned sources in the following table:

Related Study	Result	Consequence for this study
TNO Study	- Demonstrates feasibility of registration of children fingerprints at 500 dpi at a predicted minimum age of 4 years.	- Did not consider any ageing effects. - Performance of the used matcher needs to be revisited as the study dates back to 2004.
BIODEV II Study	- Demonstrates feasibility of registration of children fingerprints at 500 dpi and their quality according the NFIQ.	- Provided good indications about ridge distances from an onsite evaluation of the data.
Ultra-Scan/NIJ Study	- No real conclusion due to less significant data used - Design of the study questionable	- Demonstrates the need for data with a time distance of 3 years or more for the age group 6-12 or, alternatively, the need for high resolution images if time distance is shorter.
University of Göttingen/ BKA Study	- Suggests the need for addressing children related aspects even for persons above 12 years - Demonstrated successfully the performance of a relatively simple growth prediction model for the age group above 12 years.	- Adoption of the universal framework for an alignment and comparison of landmark configurations of related fingerprints.

Chapter 3: **The Data Source**

3.1 Child Fingerprints from Portugal

By means of an agreement between the JRC and the Portuguese Immigration Service (SEF)[12], 3264 pairs of fingerprints of children were handed over to the JRC. The data stems from a national registry of passport data in which copies of the fingerprints of the passport holder are also stored. Portugal captures fingerprints for passport purposes from all persons applying for a passport regardless of the age. Therefore, the repository contains fingerprints of citizens from 0 years on.

“Pair of fingerprints” means: two prints of the same finger taken at two different points in time. As always two index fingers were enrolled in the context of issuing a new passport, the 3264 pairs of fingerprints come from 1632 different persons. This is a relatively large set of fingerprints (compared to usual performance testing exercises with some 300-1000 persons) and the only database yet known with children's fingerprints taken at different points in time.

3.2 The Structure of the Data

Despite the origin of the data as acquired during the administrative act of issuing a passport, there is no further indication about its actual ground truth, i.e. the question whether the supposedly corresponding fingerprints are actually related. An exhaustive manual inspection of the fingerprints in order to verify its ground truth was therefore performed. The result of the analysis revealed some clear erroneous cases (1,9%) but also a large number of cases for which proper correspondence could not be immediately verified (16%). The erroneous cases were:

- For 30 individuals, left and right index fingers have been swapped, affecting 60 pairs of fingerprints. After re-swapping, the data could be used again.
- For 1 individual, the second enrolment used a different finger, affecting 1 pair of fingerprints. That pair has been excluded from further consideration.

For the remaining doubtful cases, the affected fingerprints have been taken out of consideration for the purpose of the study. This reduced the set of test fingerprint pairs from 3264 to actually only 2611, but still remains similarly large.

---

[12] http://www.sef.pt/portal/V10/EN/aspx/page.aspx

The age of the children at the time of fingerprinting ranges from 0 to 11 (i.e. under 12). The age differences between two corresponding fingerprints range from 24 to 54 months.

Apart from few exceptions, the statistical distribution of the data is presented in Table 2. It shows that the majority of data is from very young children with a time distance between the prints of 2-4 years. This fact is due to the particular passport issuing process as described before. This leads to a number of particularities (Table 1) that need to be considered:

Particularity	**Consequence**
Only some 150 pairs of fingerprints of children above 7 years.	Not much impact for the conclusions about the growth effect of children between 6 and 12 in general. There is still more data than for the Göttingen study but more data would be required for real performance testing of algorithms at a later stage.
No data for children above 12 years	Does not allow verifying the findings of the Göttingen study with a larger data set, but anyhow complements that study. However, data of children above 12 is desirable for performance testing with vendors.
Majority of data between 0 and 7 years age.	May focus the analysis on children below 6-7 years.
No data pair with oldest print of a child above 9 years.	Does not directly allow conclusions about recognisability of children above 9 years after 2 and more years have elapsed. Due to the upper limit of 12 years, there are no pairs for 9 year-old children after 3 years.
No additional metadata such as gender and body height	No direct conclusion possible on the link between these parameters and the growth effect. However, not essential for the modification of algorithms as these parameters would be hard to be taken into account.

Table 1: Particularities of the data and consequences

Age Difference (months)	Age Group										
	2	3	4	5	6	7	8	9	10	11	
24	0	6	34	38	0	2	0	0	8	4	92
25	2	17	36	40	6	0	8	4	8	8	129
26	2	16	22	38	5	6	2	0	2	2	95
27	4	15	26	32	9	1	1	2	3	0	93
28	6	19	29	32	12	4	2	4	4	6	118
29	11	18	22	30	13	0	0	0	2	4	100
30	2	19	28	23	12	2	2	6	4	5	103
31	2	16	20	25	20	6	0	6	4	1	100
32	3	24	24	23	34	7	7	1	0	4	127
33	0	24	19	23	28	0	12	0	2	4	112
34	0	19	21	24	32	0	2	0	6	0	104
35	0	18	9	25	35	4	0	2	5	8	106
36	0	9	19	19	46	2	0	3	4	6	108
37	0	6	22	20	35	12	2	8	4	2	111
38	0	9	18	27	21	9	1	8	0	2	95
39	0	6	25	24	26	12	0	2	0	2	97
40	0	5	31	20	18	11	2	2	0	0	89
41	0	6	16	26	17	11	1	4	2	0	83
42	0	3	18	13	20	13	2	2	2	0	73
43	0	1	22	28	11	22	6	0	4	2	96
44	0	0	23	16	27	15	2	2	2	4	91
45	0	0	19	30	9	21	2	5	4	8	98
46	0	0	12	28	8	21	2	2	4	0	77
47	0	0	20	24	19	12	5	2	6	2	90
48	0	0	15	26	20	18	6	8	4	6	103
49	0	0	5	15	12	8	7	2	6	7	62
50	0	0	4	5	9	1	4	4	2	6	35
51	0	0	1	9	5	3	2	0	0	0	20
52	0	0	0	0	0	1	0	0	0	2	3
53	0	0	0	0	0	0	0	0	0	0	0
54	0	0	0	2	0	0	0	0	0	0	2
	32	256	560	685	509	224	80	79	92	95	2612

Finding: Limitation of available data

The available data does not allow for seamless conclusions from birth to adulthood. Conclusions can only be made within the age group of 0-11 years.

Table 2: Distribution of fingerprint pairs over age (2-11) and time distance (24-54 months). Records at age x with time distance y means one pair of samples at age x and at age x-y.

3.3 Quality of the Data

With regard to quality, the set is relatively mixed. According to NFIQ (as the widely used quality metrics, though outdated and less appropriate for children fingerprints[13]), there is a clear increase of cases where both fingerprints of the same finger give the highest score of 1 (see Figure 11). Vice versa, the likelihood of having both fingerprints either NFIQ=4 or 5 (the two lowest quality scores given by NFIQ) decreases with age as depicted in Figure 12.

Figure 13 displays the NFIQ mean values for all age groups over the complete set of fingerprints. The trend is quite obvious; the older the children, the better the quality of the fingerprint images. The quality finding was also set in relation to those of the studies from

[13] see the notes on NFIQ in Appendix 1.

TNO[14] and BIODEFII[15]. For the latter, it has to be mentioned that the figures were obtained at the start of the study and before appropriate enrolment guidelines had been developed to increase the overall quality[16]. As with these measures the fingerprint quality increased dramatically [33], it can be expected that also for potentially enrolled children the quality would have been much better and comparable to that of the TNO study. However, such (improved) data was not available for comparison.

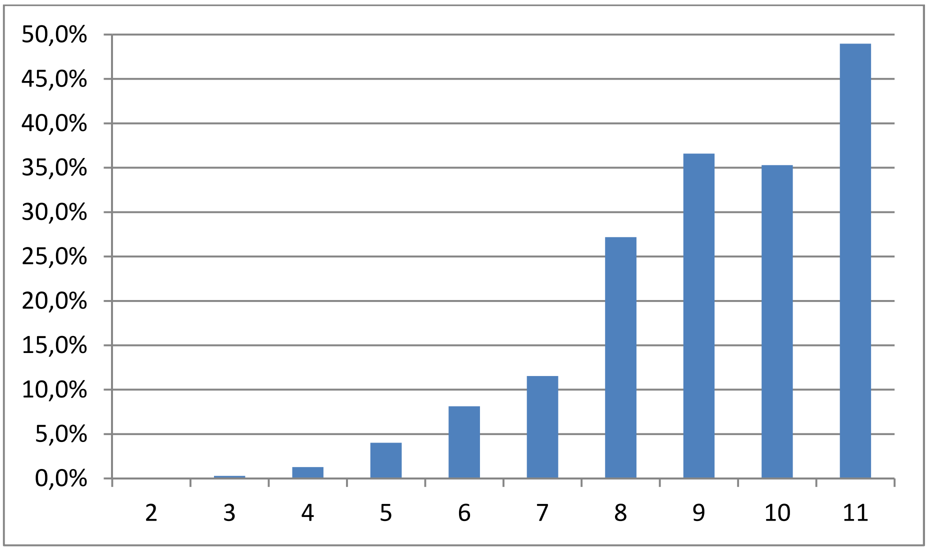

Figure 11: Percentage per age group for which both fingerprints are at NFIQ=1 (age grouping at the time of the second fingerprint)

Finally, Figure 14 shows the distribution of corresponding fingerprints with respect to the NFIQ value of the first (i.e. older) and second (i.e. newer) fingerprint.

As a further reference, the IQF quality metrics from MITRA [16] has been used (Figure 15) which is based on aspects like contrast, sharpness and detail rendition of a digital image. These metrics are mainly applied for latent fingerprints found at crime scenes and, in contrast to NFIQ, are less dependent on potential assumptions with regard to feature dimensions applicable only for adults. It is interesting to note, that for the age groups

[14] For TNO study, averages were taken from the publication [17].

[15] For BIODEV II, right index fingers only of 526 individuals. The fingerprints were obtained with 4-finger scanners rather than 1-finger scanner as for data from Portugal and for the TNO study. The number of individuals for each age group were: age 6: 29; age 7: 32; age 8: 38; age 9: 58; age 10: 58; age11: 66; age 12: 70; age 13: 85; age 14: 90.

[16] There were actually fingerprints of 3 individuals at age 4 and 5, all of with NFIQ values of 1. However, we dropped this data because of its weak statistical significance.

between 3 and 11 the average values increase only slowly from 70 to 80 (out of 100)[17], i.e. the quality differences according to IQF are less grave. Also, the statistical standard deviation decreases slightly with age.

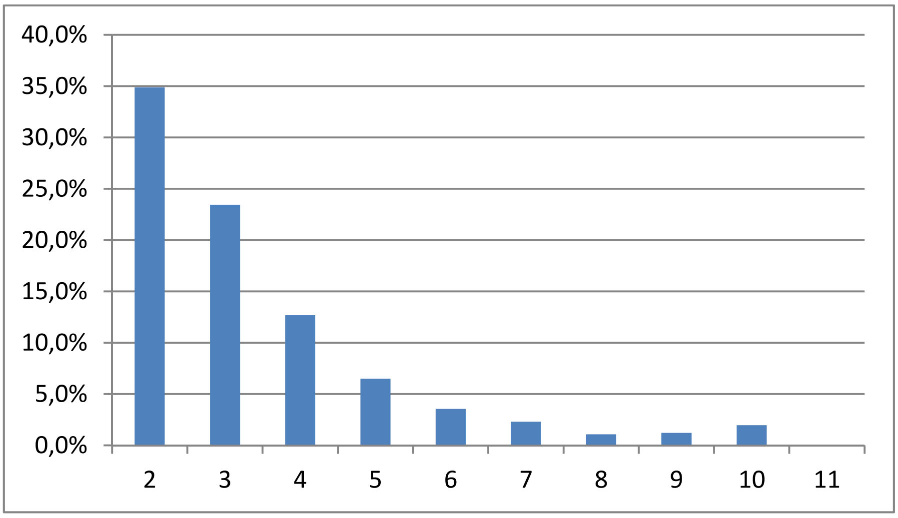

Figure 12: Percentage per age group for which both fingerprints are NFIQ=4 or 5 (age grouping at the time of the second fingerprint)

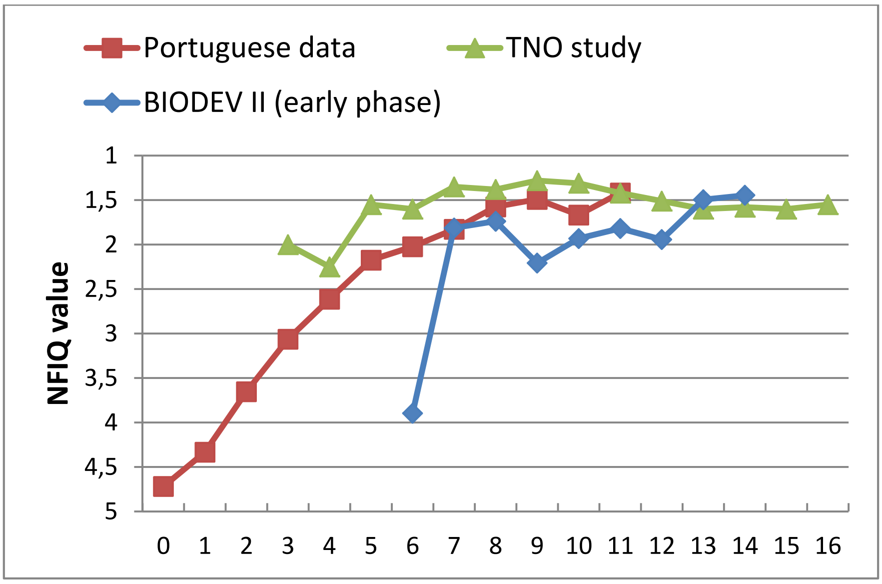

Figure 13: NFIQ mean value per age groups 0-16
(from 1=excellent to 5=very bad)

[17] The scores of IQF are usually grouped as follows: 0-25 unacceptable, 26-50 marginal, 51-75 adequate, 76-100 excellent.

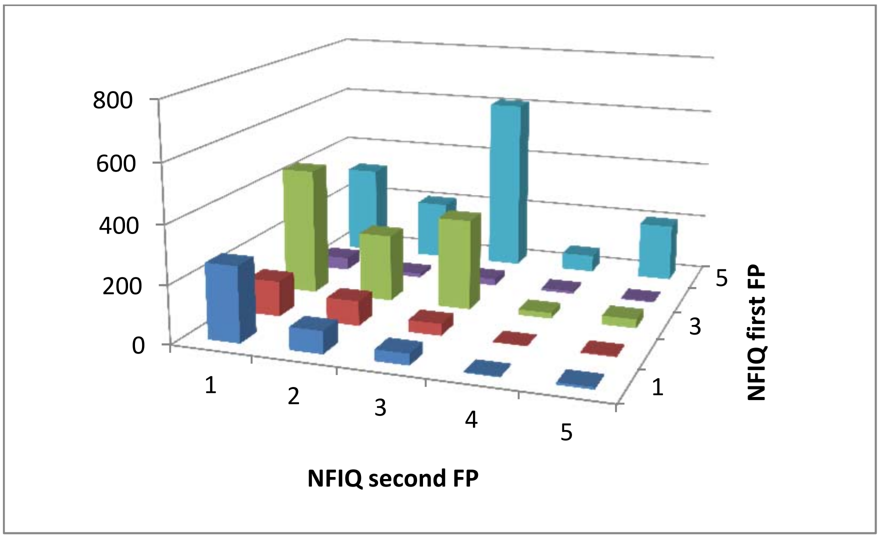

Figure 14: Distribution of fingerprints pairs according to NFIQ value of first and second print

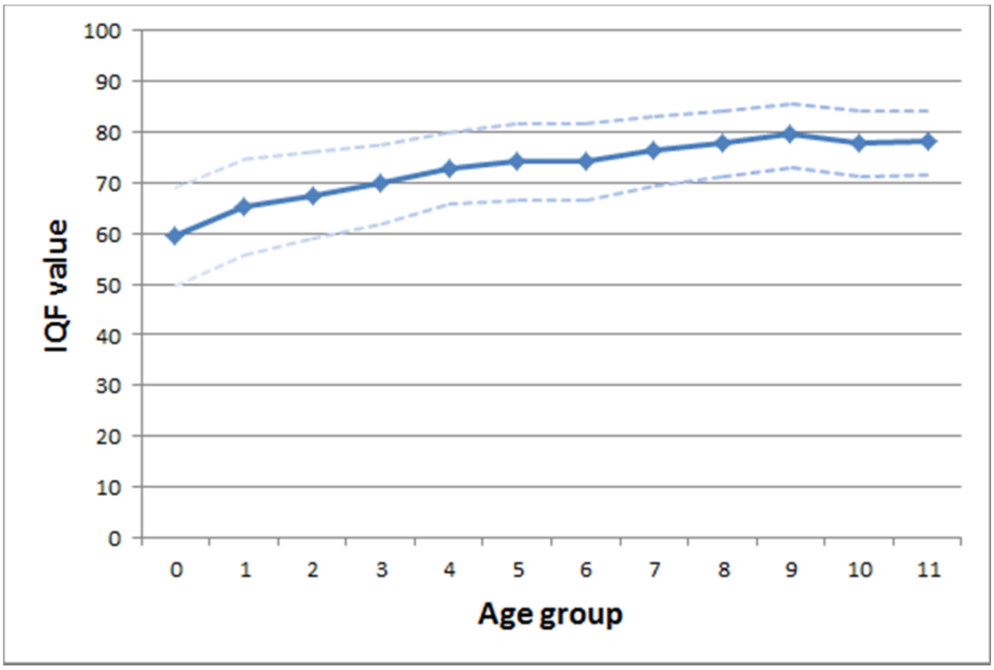

Figure 15: IQF values per age group, including standard deviation (dotted) (from 0=worst to 100=best)

Apart from the individual NFIQ and IQF quality scores, there is no further information available about the fingerprints. In particular, it is not possible to know what data from adults from the same source would produce as typical NFIQ/IQF values (in order to distinguish general quality aspects from children specific

aspects), or what recognition performance could be achieved with adult data (in comparison with the results of section 4.1).

In fact, it is very likely that the enrolment process has improved during the observed time window of 4.5 years due to habituation of the involved actors. Unfortunately, the available data was with no indication about the absolute calendar date at which it has been acquired.

The quality of most fingerprints was good enough for manual identification of a significant number of landmarks for further analysis, even if these fingerprints would less qualify for automated recognition. This point will be discussed in more detail in Chapter 6.

As can be seen in Figure 14, the majority of data (60%) has an NFIQ value of 1-3 for the second and 3-5 for the first fingerprint. For the purpose of testing fingerprint recognition algorithms, such data are usually considered “challenging”. However, as will be explained in section 4.1, the restriction to a subset with only NFIQ=1 values (some 200 or 6%) does not change results by complete orders of magnitude.

Chapter 4: **Analysis of the Individual Recognition Steps**

In this chapter, the findings in analysing the individual steps of the recognition process are presented.

4.1 Fingerprint Recognition Algorithms

4.1.1 Used matchers

With respect to matching algorithms, the following matchers were used:

- A free matching algorithm provided by the U.S. National Institute of Standards and Technology (NIST) called **“bozorth3”**. The algorithm is part of the free fingerprint processing tool kit NBIS[18]. The toolset includes also “nfiq” which is widely used to determine the NFIQ quality value of a fingerprint image. Although this algorithm is not considered as “leading edge technology”, it is often used in literature as a benchmark.
- **2 commercial matchers** (referred to as “Vendor 1” and “Vendor 2”) for which anonymity is applied on request of the vendors. Both vendors are major players in the fingerprint recognition market[19] and can be considered representative for state-of-the-art commercial solutions. One of the two vendors provided – apart from an off-the-shelf solution – also an experimental version with “children feature”, i.e. with adaptations for better recognising children fingerprints.

Generally, the JRC offers interested vendors the opportunity to perform trials on the particular data set as long as the JRC itself has access to this data. Due to the strict data protection conditions, the vendors must follow a rigorous protocol which does not allow direct contact with the data but only through authorised JRC personnel. The JRC accepts the business sensitivity of these tests and ensures anonymity of the vendors when requested.

All experimentation followed the same general methodology. The pairs of fingerprints were separated in two global sets “old” and “new”. “Old” consists of all first (i.e. older) fingerprints of the pairs and “new” of the second one. Then all members of “old” were compared with all members of “new”. There was no comparison of members of “old” (or

---

[18] http://www.nist.gov/itl/iad/ig/nbis.cfm

[19] The list of main players include 3M Cogent, Cross Match, Dermalog, Lumidigm, Morpho, NEC, Neurotechnology, Suprema.

“new”) against members of the same group. Within “old” and “new” subgroups with particular profile were established, e.g. all members of “old” at age 2 and their counterparts in “new”. For all matching tests, the respective ROC diagrams as explained in section 2.1 and Appendix 1 was developed. For each matching test there was only one “genuine user” comparison (i.e. old sample against new). All other samples were considered “imposters”.

For the separation in age groups, it is important to bear in mind that towards the end of the age scale for the first print (i.e. 9 years) there is no second print with a time difference to the first one of more than 3 years. For 8 years old, the time difference is less than 4 years, but this is less important due to the overall smaller number of pairs with such a large distance.

In all ROC diagrams the Equal Error Rate[20] (EER) can be found along the dotted blue lines.

4.1.2 Results with NIST’s bozorth3

The tests with bozorth3 are displayed in the following three figures. Figure 16 shows the individual ROC curves for the various age groups within “old”. Obviously, there is a clear relation between the age group and the matching performance, i.e. the older the children, the better the recognition result. For the oldest available age group of children at 9 (at the time of the first fingerprint), the EER is about 10% or even less.

Figure 17 shows results for certain quality profiles and the combination with the age group 8-9 (the two oldest available within the Portuguese data). It clearly suggests a relationship between NFIQ value and recognition performance. A value of NFIQ=1 for both fingerprints gives better results compared to the case with only the first print having NFIQ=1. On the other hand, if restricted to the group 8-9, a NFIQ value of 1 clearly gives better results than for the average of all others with both prints at NFIQ=1.

As mentioned before, the age group 9 needs to be considered with some care. Performance is clearly the best but there is no pair with time distance of more than 3 years inside.

With regard to absolute performance, a FAR of 0,1% would roughly imply FRR in the order of 60% or higher (cf. section 2.1). However, applications including children could have other target FAR/FRR combinations, at least with respect to children. The EERs are between 10% and 60% with an average of 35%. It is worthwhile noting the relatively good results for the joint age group of 8 and 9 with the best NFIQ value for all fingerprints. If the trend would continue for age groups above 9 (with same quality profile), then the corresponding error rates could reach acceptable ranges.

[20] Cf. section 2.1.7 on page 27.

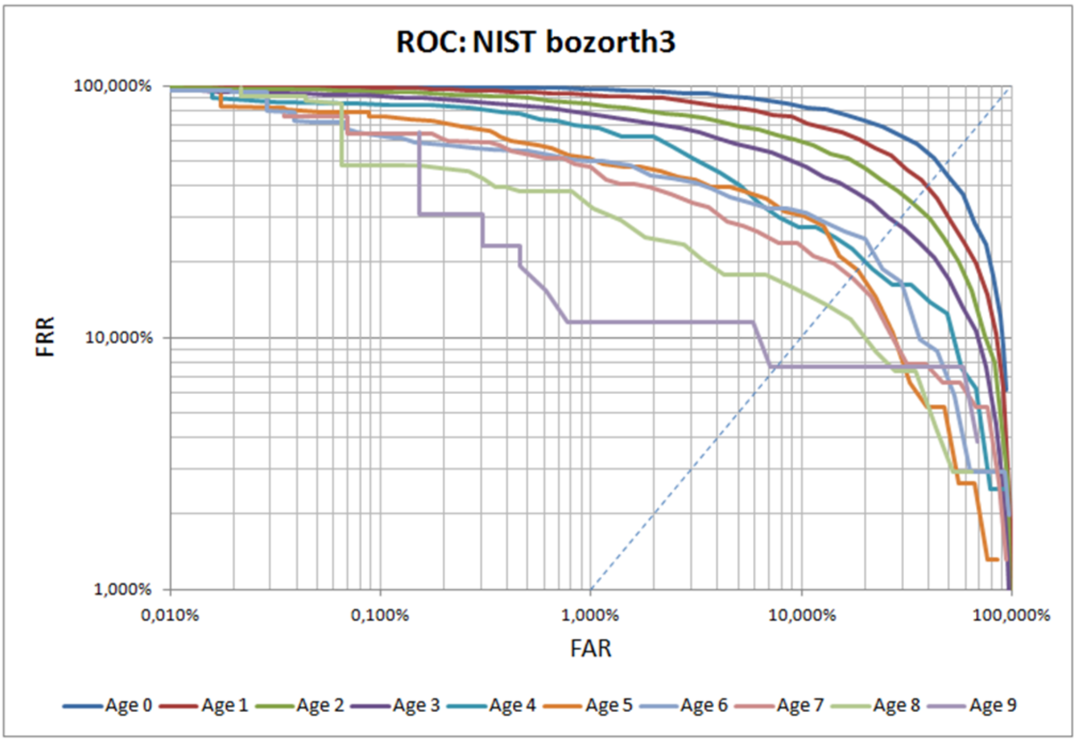

Figure 16: ROC diagram for NIST bozorth3 for different age groups (logarithmic scale) (classification according to oldest fingerprint of a pair).

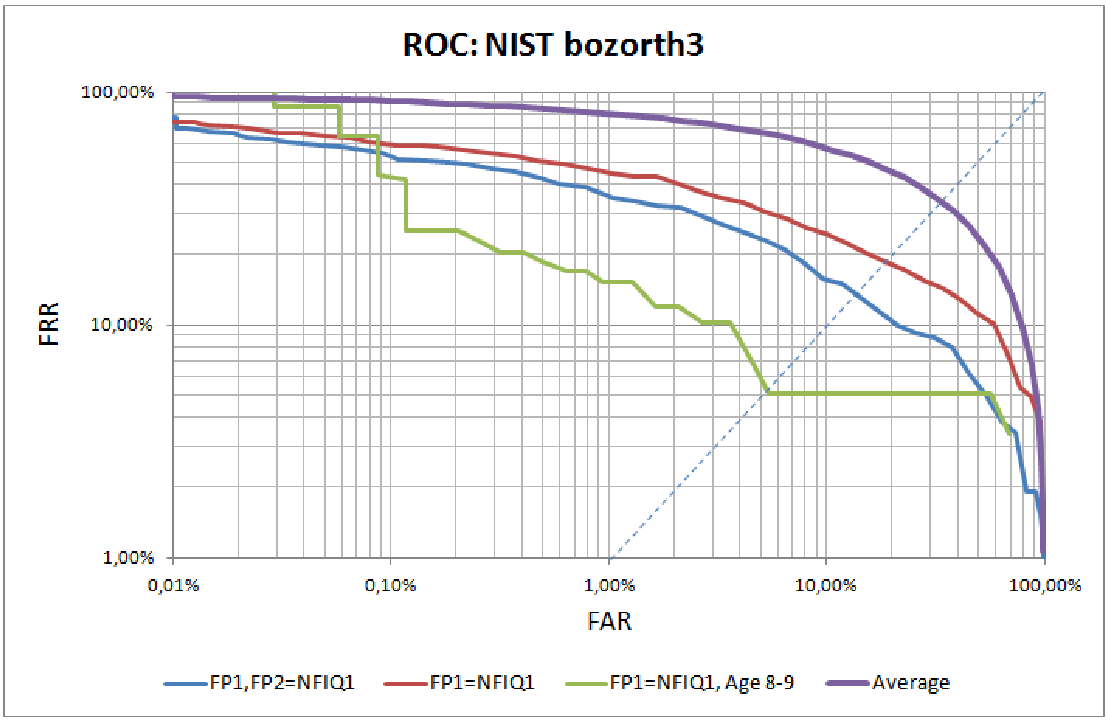

Figure 17: ROC diagram for NIST bozorth3 for different quality profiles (logarithmic scale)

Figure 18 provides a relative surprise. When comparing various groups of time differences between corresponding fingerprints there is no clear trend visible. The performance of pairs with a time difference between 24 and 30 months is almost the same as those for the difference between 49 and 54 months. However, it has to be taken into account that for the latter there is much less statistical material available within the data set.

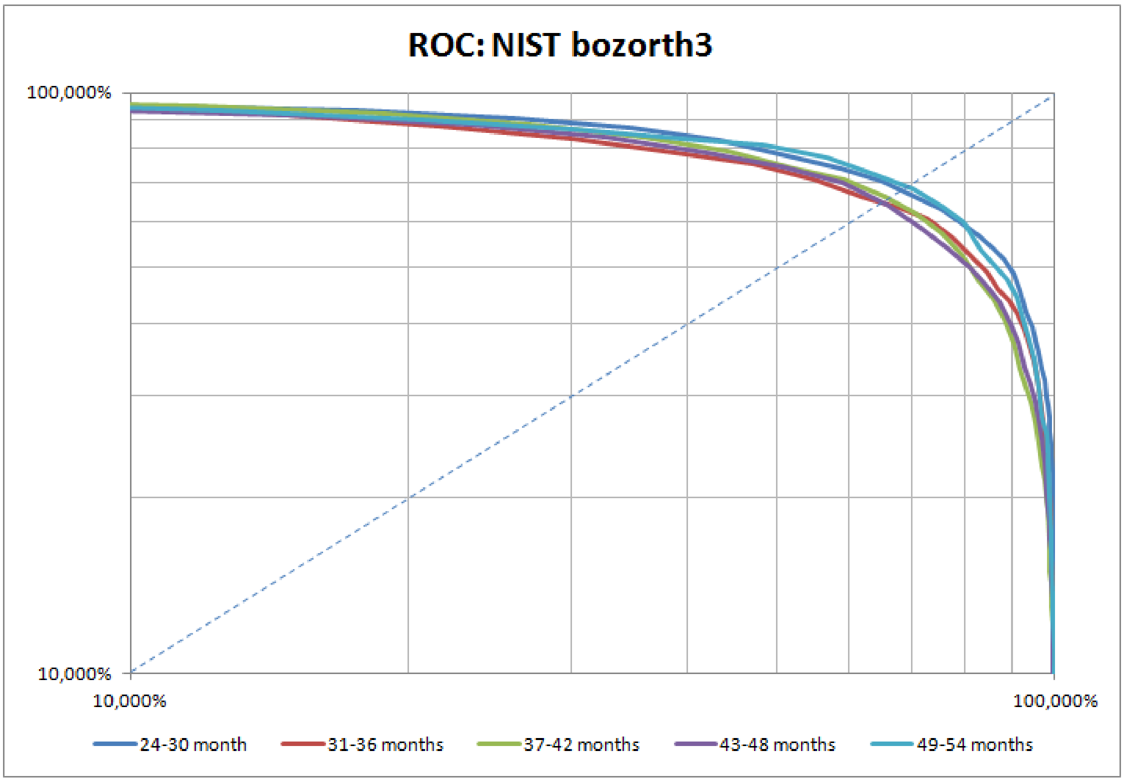

Figure 18: ROC diagram bozorth3 for various time differences between fingerprints (logarithmic scale)

4.1.3 Results with Vendor 1

The results with bozorth3 could be qualitatively confirmed, even though the performance of Vendor 1 is better than that of bozorth3, at least on average. Again, the FRR at a FAR of 0,1% is on average above 60%. The following three figures show the results for the same settings as used for bozorth3 in the previous section.

Figure 19 shows the individual ROC curves for the various age groups within “old”. Again, there is a clear relation between the age group and the matching performance. For the oldest available age group of children at 9 (at the time of the first fingerprint), the EER is about 10% as for bozorth3.

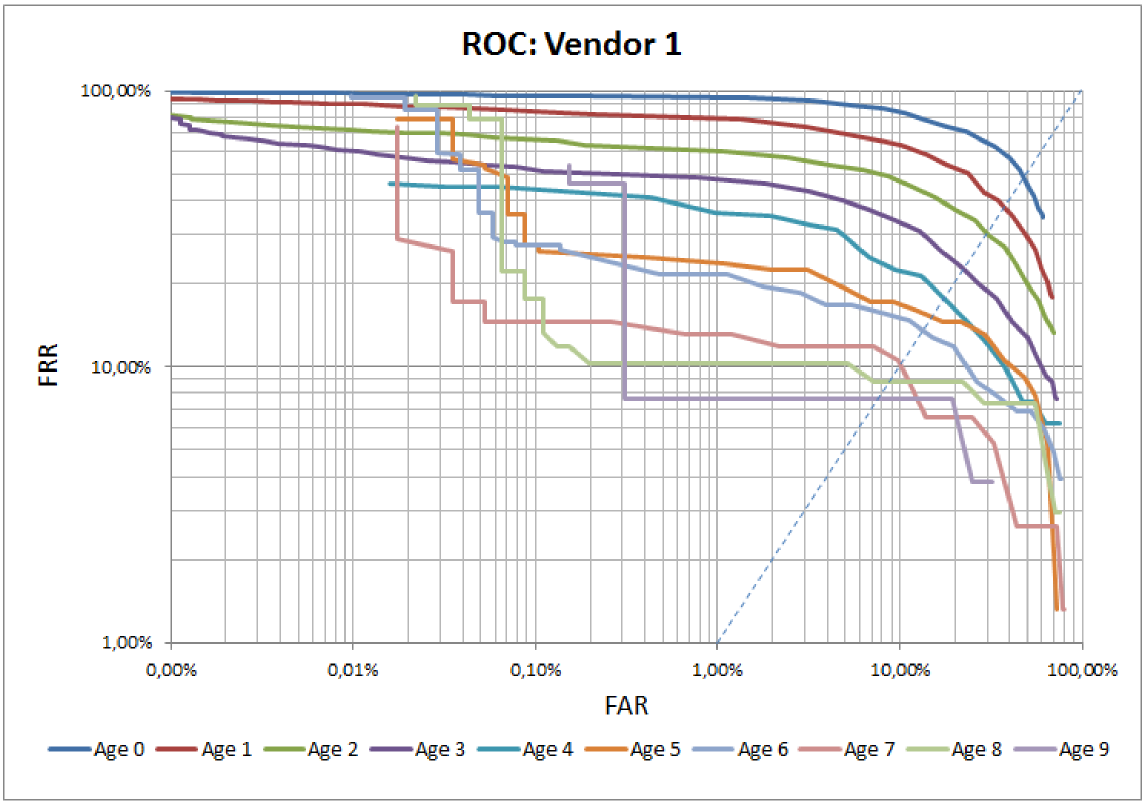

Figure 19: ROC diagram for Vendor 1 for different age groups (logarithmic scale) (classification according to oldest fingerprint of a pair).

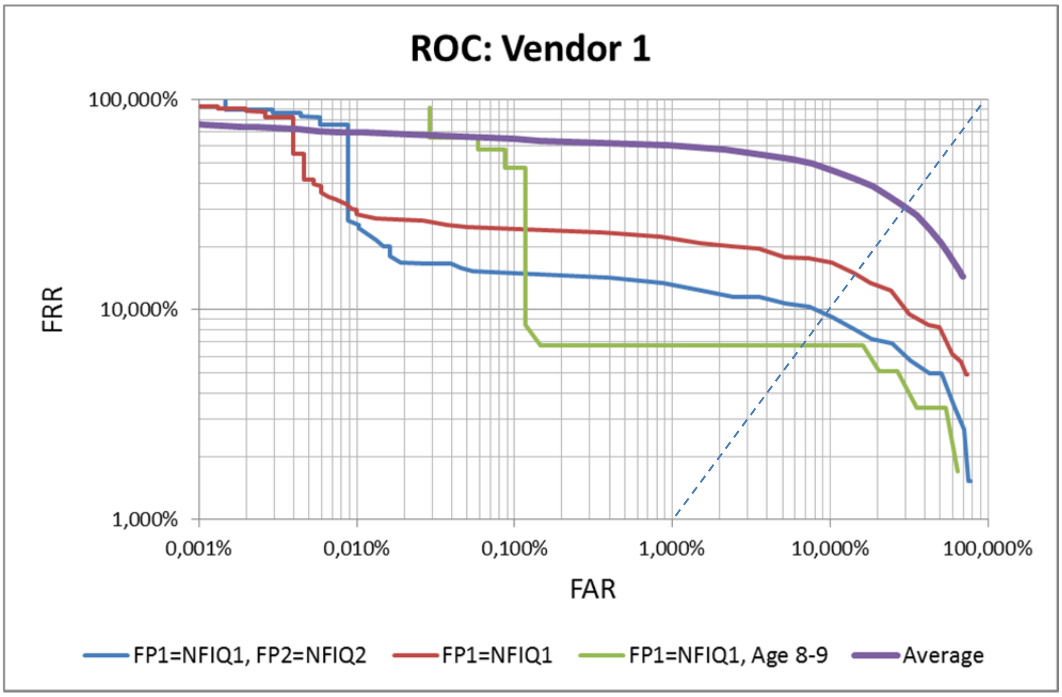

Figure 20: ROC diagram for Vendor 1 for different quality profiles (logarithmic scale)

Figure 20 shows results for certain quality profiles and the combination with the age group 8-9 (the two oldest available within the Portuguese data). It clearly suggests a relationship between the NFIQ value and recognition performance and similar conclusions as for bozorth3.

Finally, Figure 21 shows also for Vendor 1 a relative independence with respect to time differences between corresponding fingerprints. Here, only the time difference between 24 and 30 months deviates significantly from the averages of the other groups.

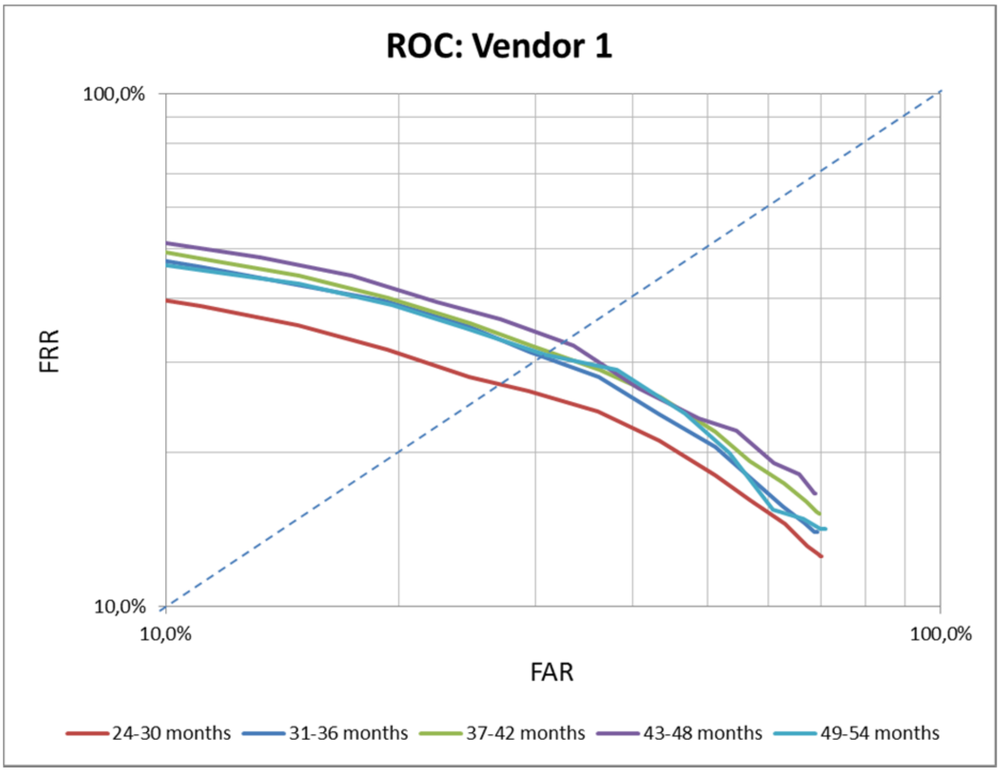

Figure 21: ROC diagram Vendor 1 for various time differences between fingerprints (logarithmic scale)

4.1.4 Results with Vendor 2

Vendor 2 provided a couple of versions to be tested from which we present three: This prediction is displayed in for the following versions:

- **Version 1:** the current commercial matching algorithm
- **Version 2:** an experimental version which is supposed to be more suitable for child fingerprints.

Though generally the results are very similar to those of vendor 1, Figure 22 shows some clear performance gains from version 1 to version 2 of vendor 2. The EER of version 1 is (along the dotted blue line) around 40 %, for version 2 around 27 %, a gain of about 30% in performance The conclusion of these experiments is that the vendor has made steps into the right direction but there might be still room for improvement.

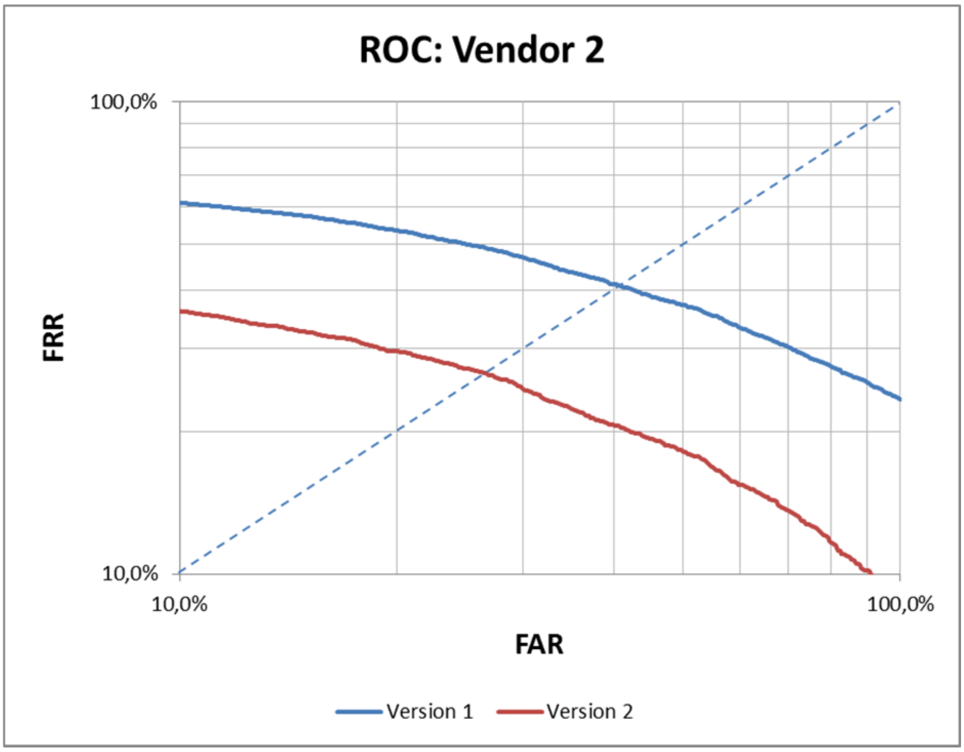

Figure 22: ROC diagram for versions 1 and 2 of Vendor 2
(all data sets)

Unfortunately, this analysis cannot be refined by age groups because the ROC diagrams have been produced under the vendor's own control. It is only possible to display the matching results for a given score threshold. The threshold chosen is the one the vendor indicated to use for practical installations in order to achieve a FAR of 0,1 %. Figure 23 shows the percentages of comparisons of corresponding fingerprints (i.e., "genuine" scores) with a score value above the threshold (i.e. positive "matches") for the two versions mentioned. As a certain anomaly, version 1 performs better for the age groups up to 3 years. For all others, version 2 has almost twice as many "matching cases". In any case, Figure 23 also shows that successful recognition (i.e. matching well above 0%) is possible even for very young children.

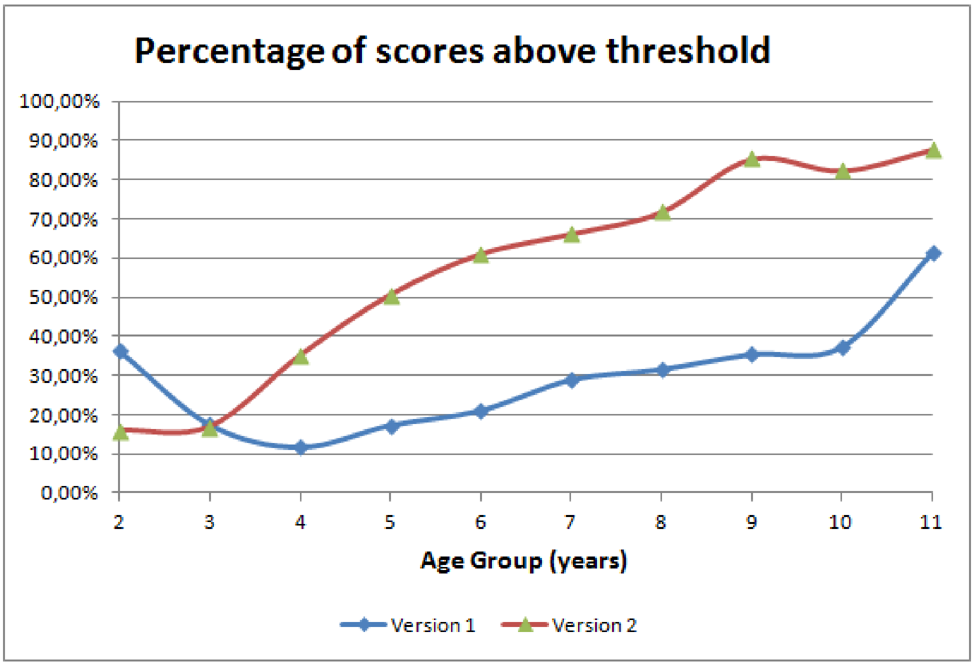

Figure 23: Comparison of "genuine" scores above threshold of Vendor 2 (i.e. corresponding fingerprints only, FAR=0,1%)

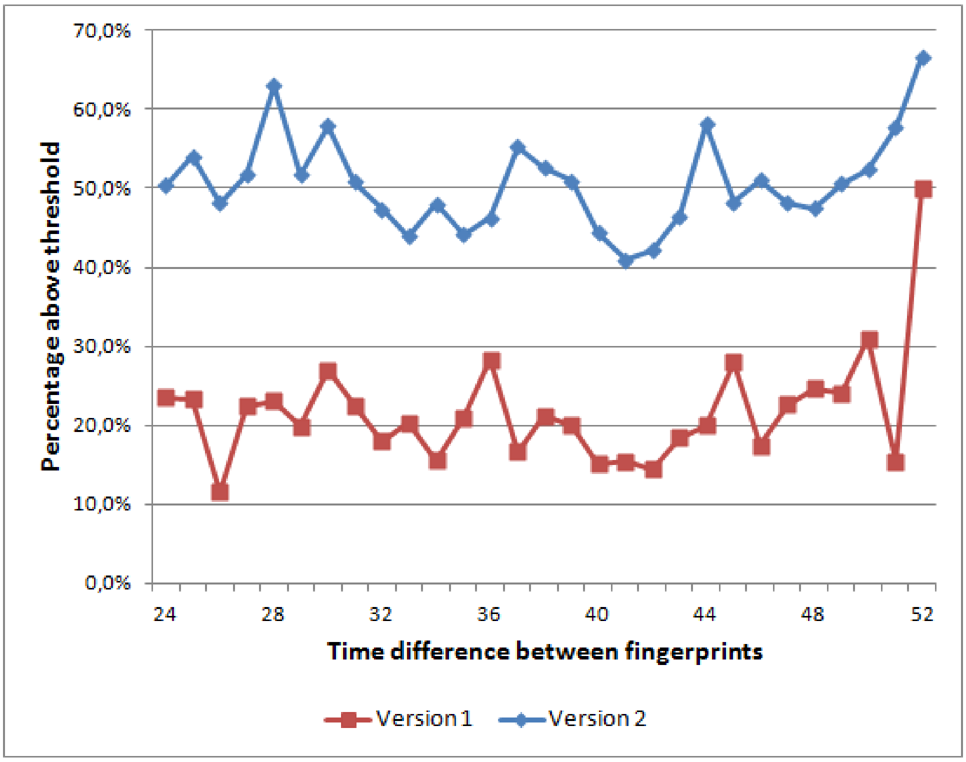

Figure 24: Dependency on time difference between fingerprints of Vendor 2 (i.e. corresponding fingerprints only, FAR=0,1%))

Finding:
Young children recognisable

Fingerprint recognition is possible even for very young children (cf. Figure 23) though at lower recognition rate than for older children.

Finding:
Independence of Time Difference

Recognition rate appears to be largely independent from the time difference between the fingerprints to be compared

The dependency about the time difference between the fingerprints to be compared was also analysed. The algorithms of this vendor confirmed the almost independence from this parameter as shown in Figure 24. Though the percentage of comparisons with scores above the threshold is higher for Version 2, both versions do not show any dependency on the time difference. Figure 24 also demonstrates, that efforts to improve algorithms with respect to children aspects pay off.

Finding:
Algorithms can be improved

State-of-the-Art matching algorithm can be improved by adaptation towards children specific aspects.

4.1.5 Relation between Image Quality Scores and Recognition Rate

As Figure 17 and Figure 20 suggest, recognition performance increases with quality according to NFIQ. Vice versa, NFIQ can serve to some extend as a relative prediction to recognition performance. Figure 25 shows how recognition rate (at a commonly chosen FAR of 0,1%) of corresponding fingerprints increases when the NFIQ value of both fingerprints is equal or better than 3, 2 and 1, respectively. The strongest dependency is visible for Vendor 1 but also for the algorithm version 2 of Vendor 2 (cf. 4.1.4).

With regard to the IQF quality metrics, the results are displayed in Figure 26. The quality categories are chosen for IQF values larger than 65, 70, 75, and 80, respectively. Again, the strongest dependency relates to the algorithm of Vendor 1. For all other algorithms, the increase in recognition rate is almost marginal.

Finding:
Quality metric could be inadequate

The quality metrics according to NFIQ and IQF as an indicator for recognition performance could be inadequate children fingerprints

The results suggest that both NFIQ and IQF in its current form could be partially inadequate to predict recognition rate of a matching algorithm in the case of child fingerprints.

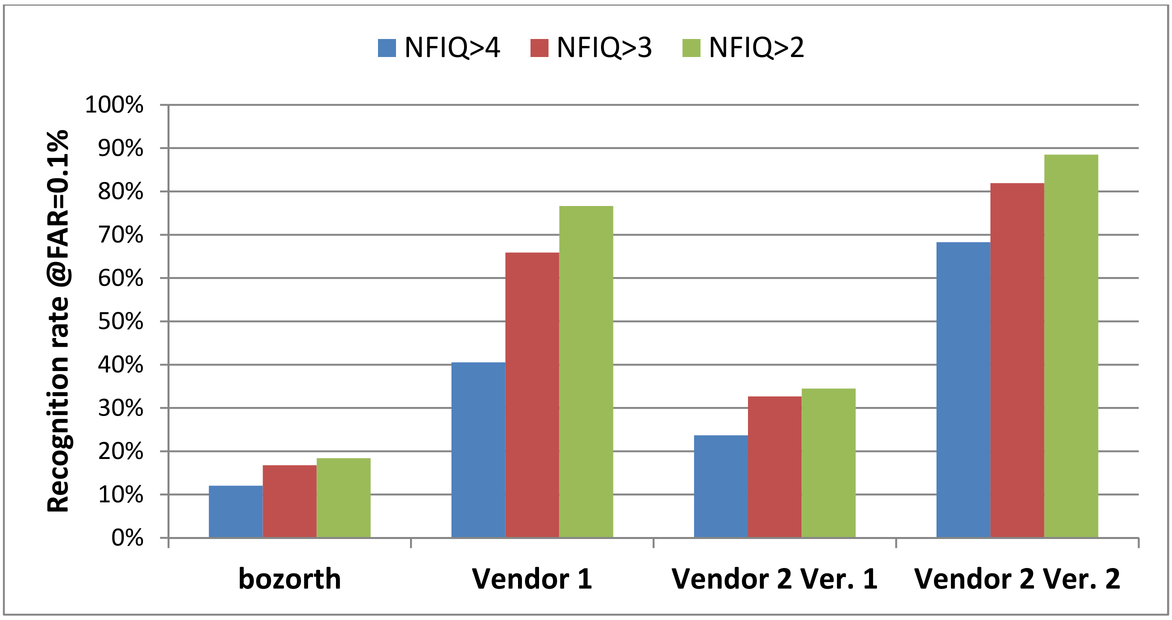

Figure 25: Recognition rates vs NFIQ of the tested algorithms

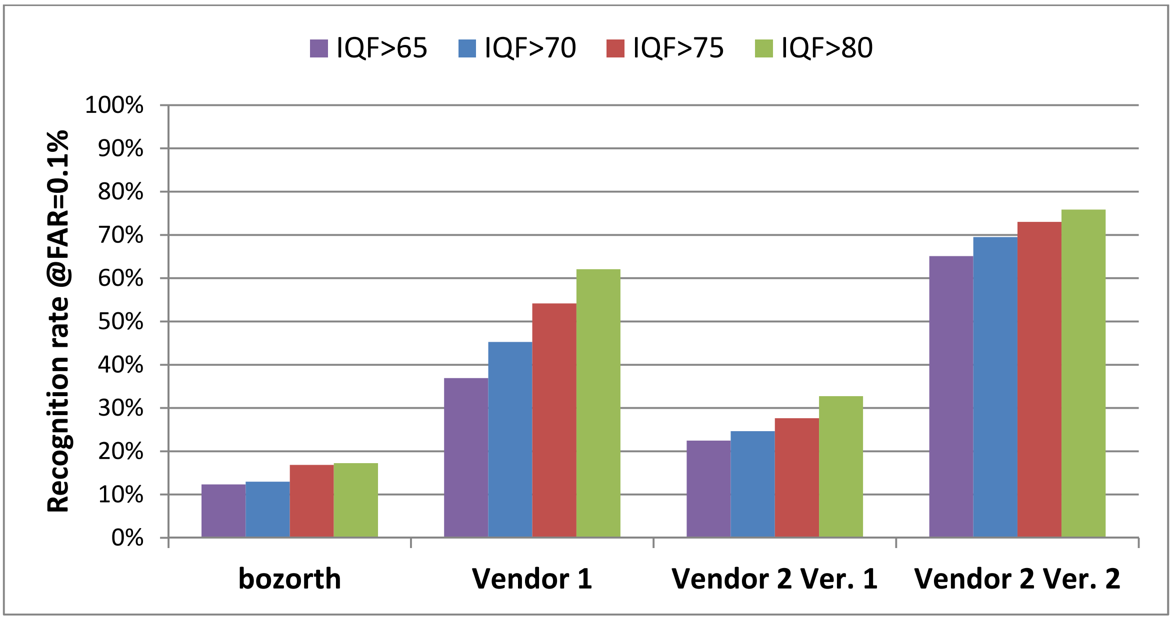

Figure 26: Recognition rates vs IQF of the tested algorithms

Figure 25 and Figure 26 also demonstrate that a certain metric might be more appropriate in combination with a certain matching algorithm than with another. Very often, commercial systems come with their own (vendor specific) quality metrics. This makes it difficult to draw any general conclusion about the relationship between quality and matching performance. This result calls strongly for the establishment of an EU-wide dataset and independent quality metrics.

4.2 Image Acquisition

4.2.1 General observations

The images from the Portuguese dataset have been acquired under conditions for which only a few quantifiable criteria where known. The images have been captured with off-the-shelf fingerprint scanners of 500 dpi. However, the quality constraints have not been very stringent in the absence of commonly agreed guidelines or best practices for the case of passport data. As the NFIQ value distribution suggests, also “bad” images with NFIQ 4 or 5 have been accepted for registration. Even an image with no fingerprint at all was found.

In general, the manual inspection of the fingerprints from Portugal revealed a number of issues which relate to the problem of image acquisition:

- A high percentage of images with larger low contrast areas. The subjectively estimated percentage is about 50% of all fingerprints.

- There seems to be also a high percentage of images with quite obvious distortions due to different ways of pressing the finger onto the scanning device. If the relevant area of fingertip is not pressed correctly onto the scanner, distortions with respect to the planar coordinates of the features may occur.

These two main observations, low contrast and distortion, deserve further analysis. As the information on the particular conditions under which the images had been acquired was not available, tentative reproductions of these results by JRC experiments were conducted. However, the problem of low contrast depends highly on the used scanning device. The impact of these devices will be discussed in more detail in chapter 6.

4.2.2 Quantification of distortion

Apart from the well-known problems with dry or wet fingers [31], in the main point of interest was the level of possible distortion. In order to get a first idea of the order of magnitude of this effect, the following experiments were conducted. Some fingerprints from 3 JRC staff volunteers were landmarked by hand and the landmark coordinates were extracted before deleting the fingerprints.[21] The test person benefited from an arbitrary freedom in the way he/she liked to press the finger on the scanning device. These “free” images were compared with images acquired under the most ideal conditions[22] possible. Some sample results are displayed in Figure 27 and were obtained using methods of shape analysis [29]. They show how strong the selected set of landmarks (connected by lines for better visibility) can be distorted through ill-positioning of the finger onto the device. The approach to align the two sets of landmarks follows the methodology of University of Göttingen (see Figure 9 of section 2.2.4).

The remarkable observation is that the order of magnitude of the distortion is at the same level as the observed deviations of mapped landmark configurations when applying the isotropic growth model in the next chapter. In other words, the deviation from an isotropic growth model can be explained by distortions during the capturing of the fingerprint.

Such strong distortions (even with no time difference between the fingerprints) may prevent matching algorithms to establish sufficient similarities. In the cases shown here, the matching scores for the two fingerprints to be compared were well below the recommended threshold for a “match”, both for bozorth3 and for Vendor 1, with the exception of test person 1 whose distorted fingerprint of the index finger still matched well with Vendor 1.

---

[21] Immediate deletion of the fingerprints was necessary to avoid problems of data protection with regard to the personal data of the volunteers. From the landmark positions only (without indication about the nature of the landmark) the real fingerprint can no longer be reconstructed.

[22] Fingers neither too dry or too wet; pressure on device optimised to receive best contrast with minimal distortion.

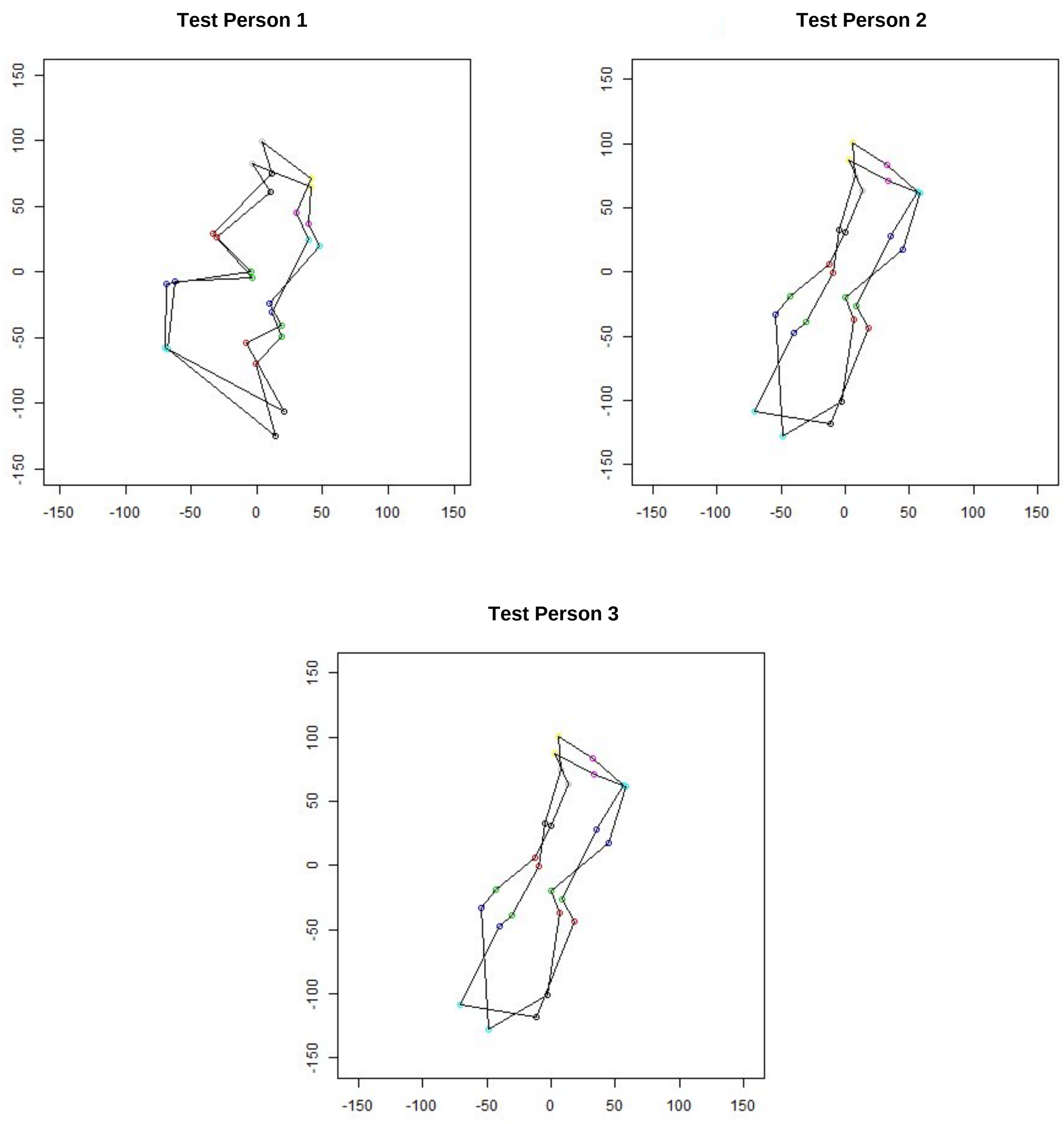

Figure 27: Sample distortions of landmark configurations through ill-positioned finger onto the scanning device (scale in pixel)

4.2.3 Potential impact of distortion

In the landmark configuration of test persons 2 and 3, the distances between corresponding landmarks are sometimes of the same size as the distances to neighbouring landmarks as illustrated in Figure 28. If a matching algorithm is able to cope with the a deviation of the selected landmark within the dotted circle (as a kind of “tolerance area”) then other landmarks within that circle could get in conflict with the chosen matching strategy. Although the precise matching strategy may vary from matcher to matcher, it can be assumed that such an ambiguity has an influence on the matching scores.

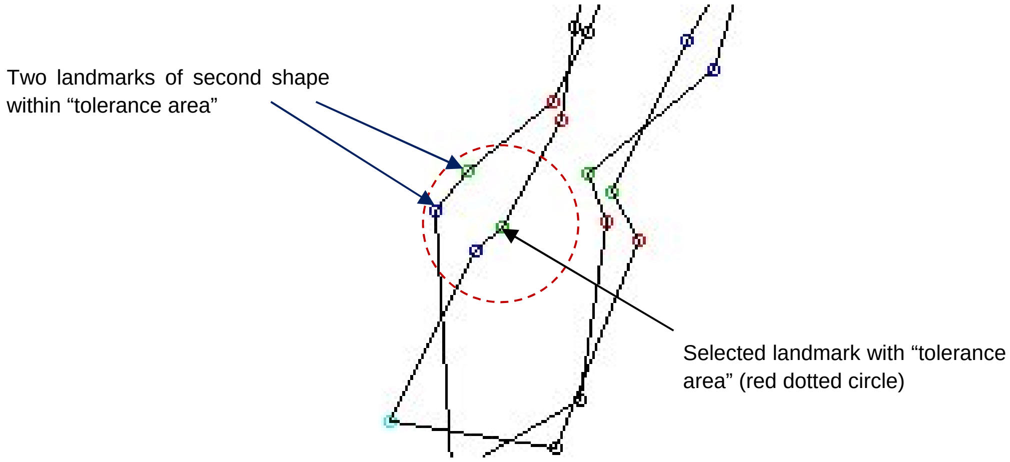

Figure 28: Example of ambiguous landmark correlation
Excerpt from Figure 27 (Test person 3)

It is therefore possible to conclude that within average smaller distances of fingerprint features of children, the problem becomes worse because this type of “tolerances” are measured in absolute values rather than in relative ones. The relevant values are algorithm dependent and are usually derived from considerations about the available resolution, the assumed ridge distance (for adults) and possible other factors. In any case, these factors are fixed and not dependent on the size of the finger in question.

The situation becomes additionally worse, if the images have also larger low contrast areas. In those areas, the precise localisation of landmarks (even if distorted) is very difficult if not impossible.

It can safely concluded, that the probability of such cases as depicted in Figure 28 rises the younger the children are, making distortions likely an even more significant issue for children fingerprints.

An extreme distortion example is the following: The two fingerprint images in Figure 29 are obtained under normal contact with the scanning device (left image) and with strong force in forward direction (right image). The result is a scaling of the common (and thus comparable) fingerprint region by a factor of 0,82. This factor is in the range of the growth effect as can be seen in the next chapter.

Figure 29: "Normal" and strongly distorted fingerprint image of the same finger
The white arrows denote the distance between the same landmarks in both prints.[23]

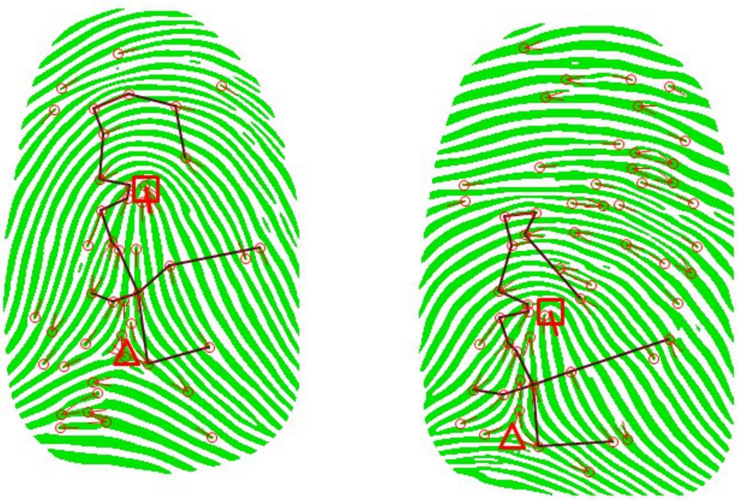

Figure 30: Extracted features and alignment for the prints of Figure 29 by Vendor 1

---

[23] Fingerprint not from Portuguese child data.

Another observation is that strong distortion can be connected with the reduction of contrast as is also illustrated by Figure 29.

Despite this obvious distortion, a matching algorithm like “Vendor 1” is still able to do an alignment with relatively high score (i.e. considered as “match”) as shown in Figure 30.

Finding:
Factors determining image quality

Image quality is strongly influenced by the combination of distortion with other quality decreasing factors like low contrast.

The following conclusions can be made:

- Strong distortions of fingerprint images are possible both for adults and children (the latter concluded from the observation of the Portuguese data) but might have a stronger impact for children.
- Matching is anyhow possible even in the presence of such distortions; however, there are obvious limits.
- Low contrast (cf. chapter 5) further increases the impact of distortion. Improving contrast (through image processing techniques or alternative types of devices) would therefore improve the situation. In chapter 6, the possibilities of existing alternative devices will be discussed.

Chapter 5: **Analysis of the Growth Effect**

The analysis of the growth effect was based on manual selection of landmarks in a limited set of 54 fingerprint pairs. It turned out that this was not as easy as expected, even leaving aside the time consumption for such an exercise. The problem was again the quality of the data. Unlike automated tools which can afford to work with less reliable landmarks (i.e. minutiae positions), this manual annotation needs absolute precision. Otherwise, the results would be distorted by additional factors and would need further processing.

The particular issues for this exercise have been:

- The "ground truth" of the data is based on the information provided by SEF only, in particular about the age of the children at the time when images have been acquired.
- The different quality of the two corresponding images always lead to a residual inaccuracy of the order of 2-4 pixels or 0,1 - 0,2 mm. This is because of different thicknesses of ridge lines in the fingerprint as illustrated in Figure 31.

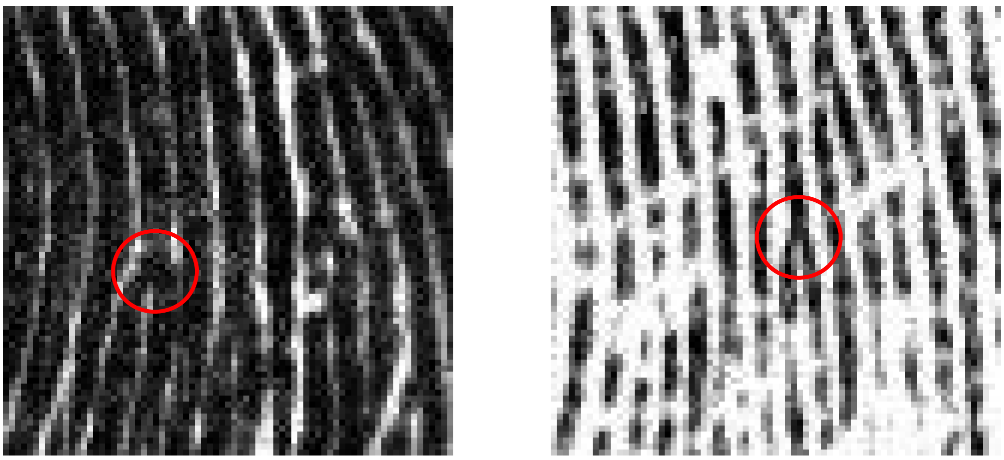

Figure 31: (Partial) fingerprints with corresponding landmarks [24]

Again, different hand encodings of the same set of landmarks with the same technique as used for Figure 27 were compared in order to ensure the mentioned residual inaccuracy is not further increased. A typical result is shown in Figure 32.

[24] The samples are 75 x 75 pixels each, or 3,75 x 3,75 mm.

It illustrates that the error through this approach is much less than the deviations observed otherwise (i.e. through growth effect or distortions).

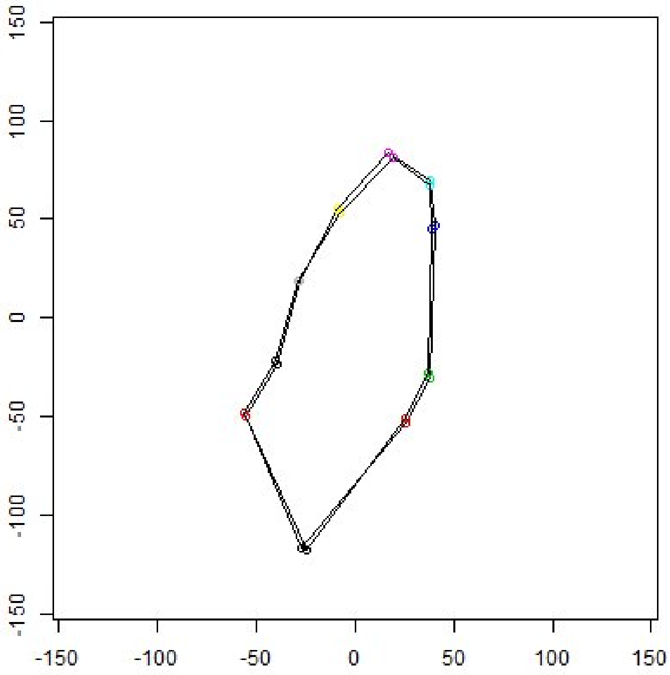

Figure 32: Experimental observation of deviation in individual landmark annotations

- A number of fingerprints have significant large low contrast parts, i.e. parts of the image where no distinction between ridge lines can be made (Figure 33). Such areas might still be used for matching (though with lower confidence) but are not precise enough for the manual landmark annotations. In some cases, such images had to be completely excluded from this type of analysis.
- For some of the fingerprint pairs, there was only a small common fraction due to different positioning of the finger during the two enrolments. In worst cases, the number of possible landmarks was too small to be used for further analysis.

Landmarks were chosen where there is high confidence on the precise position. An average of some 12 landmarks was defined per pair of fingerprints but not less than 7. Landmarks were also positioned in order to cover the widest area of the common fraction of the two fingerprints.

Figure 33: Example of low contrast fraction of an image

The purpose of the statistical analysis following the annotations is:

- To test to what extent isotropic growth models can be applied to the pairs encoded.
- To understand the critical parameters which explain the reason why some pairs could be matched with the matchers tested so far and why others not.
- To test the potential improvement of matching results when isotropic growth is assumed.
- To estimate any kind of age limit for the parameters to be examined in more detail.

With regard to the first point (isotropic growth model), some examples are given in Figure 34. It shows for 4 sample pairs of fingerprints how close landmarks could be mapped into a common coordinate system using only translation, rotation and scaling (isotropic mapping). Same colours are used for corresponding landmarks (however, only 8 different colours available!). The approach to align the two sets of landmarks follows the methodology of University of Göttingen (see Figure 9 of section 2.2.4) which is based on the shape analysis tool of the statistical tool set R [29].

The samples have been chosen to illustrate four different extreme cases:

1. Large distance between the remapped landmarks and large time distance between the corresponding fingerprints (Figure 34). Scaling factor: 0,808
2. Large distance between the remapped landmarks and small time distance between the corresponding fingerprints (Figure 35). Scaling factor: 0,937
3. Small distance between the remapped landmarks and large time distance between the corresponding fingerprints (Figure 36). Scaling factor: 0,914
4. Small distance between the remapped landmarks and small time distance between the corresponding fingerprints (Figure 37). Scaling factor: 0,968

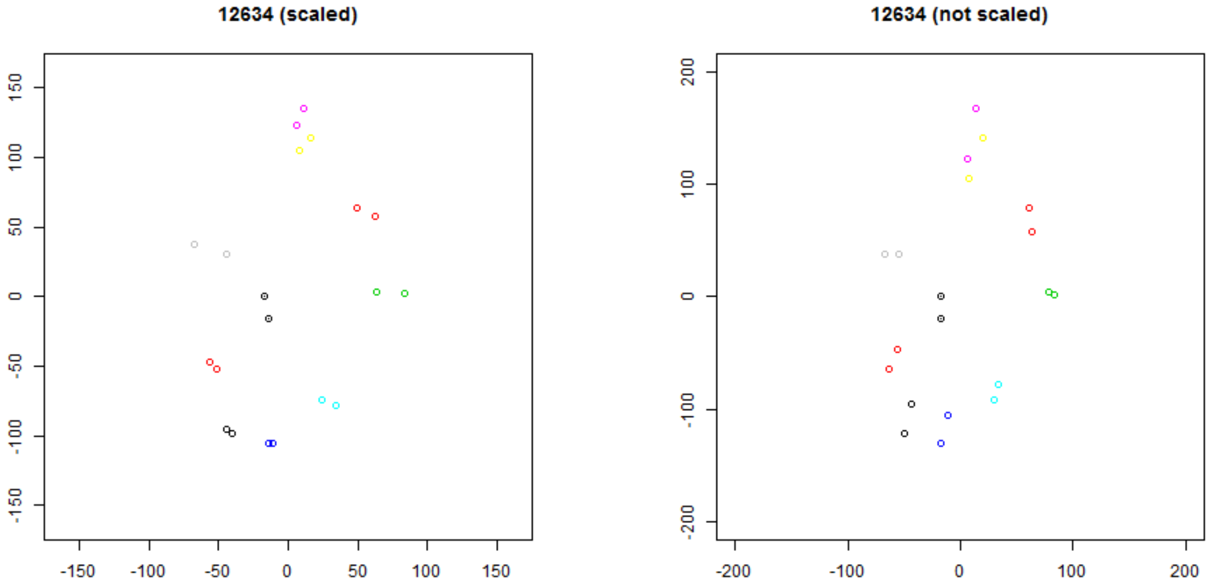

Figure 34: Closest configuration of landmarks after isotropic mapping
Example with large distance between landmarks and large time distance (42 month)
(scaled and not scaled, units in pixel)

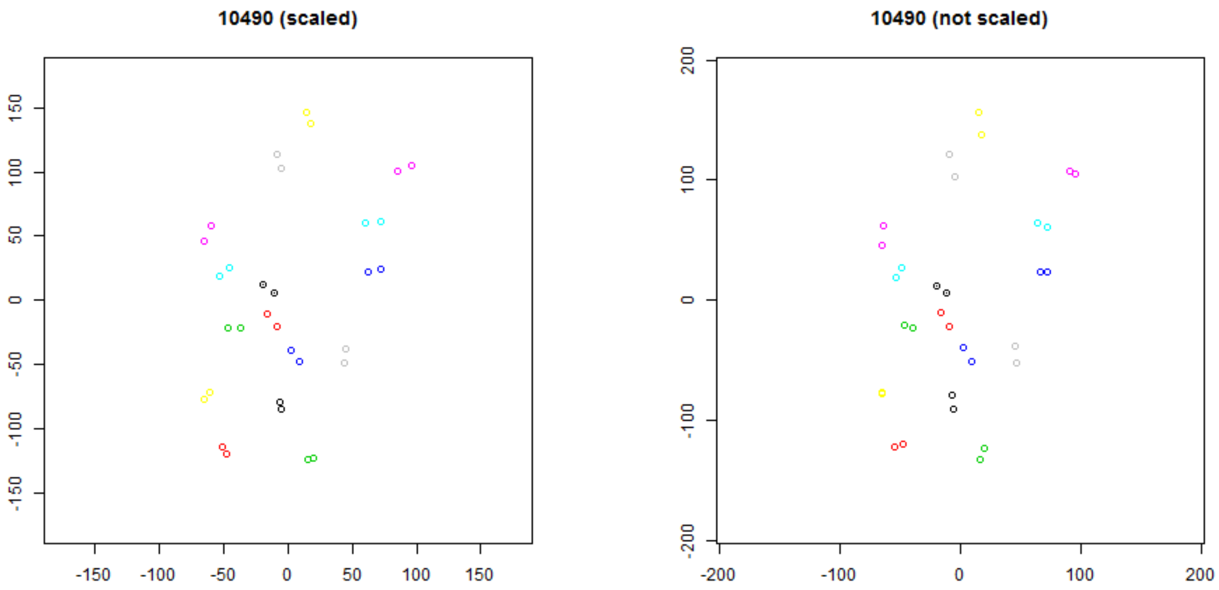

Figure 35: Closest configuration of landmarks after isotropic mapping
Example with large distance between landmarks and small time distance (25 month)
(scaled and not scaled, units in pixel)

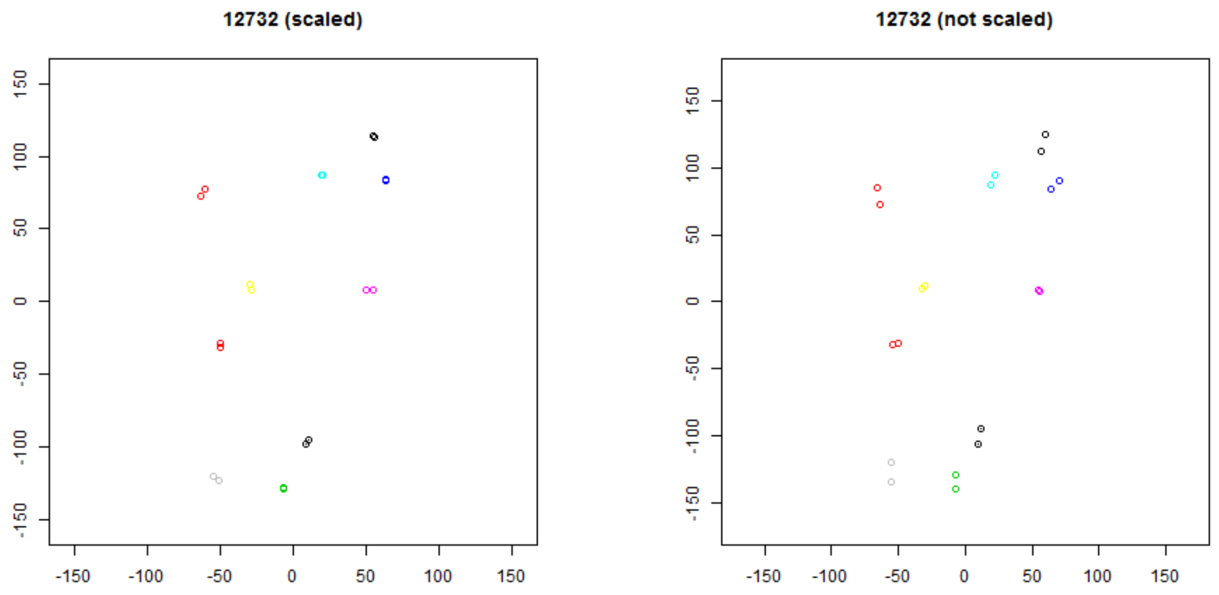

Figure 36: Closest configuration of landmarks after isotropic mapping
Example with small distance between landmarks and large time distance (50 month)
(scaled and not scaled, units in pixel)

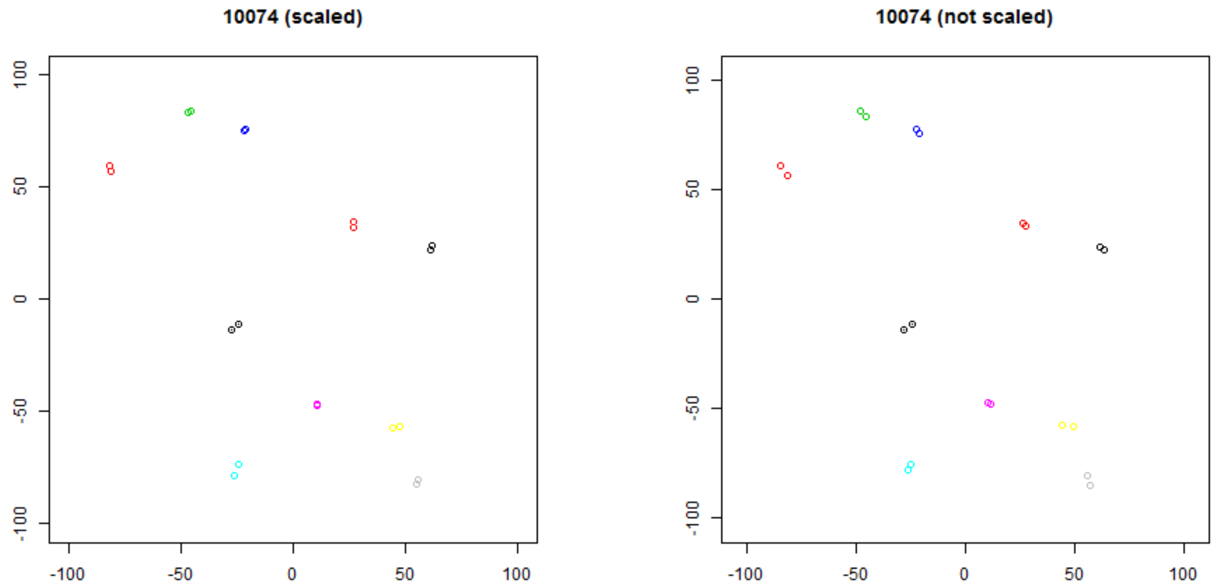

Figure 37: Closest configuration of landmarks after isotropic mapping
Example with small distance between landmarks and small time distance (25 month)
(scaled and not scaled, units in pixel)

In all cases, the reported scaling factor is estimated by the R tool as the optimal factor to scale the landmark configuration of the newer fingerprint towards the older one (apart from appropriate translation and rotation). The estimated scaling factor may therefore depend on the particular selection of landmarks. However, as the individual selections should cover a wide area of the fingerprint, the impact of distortion to the scaling factor should be kept at a minimum.

The unit of the coordinates are pixels, where 1 pixel equals 0,05 mm. The samples have been chosen with the additional property that the pairs of fingerprints behind could be matched with the matchers mentioned before, i.e. NIST's "bozorth3" and "Vendor 1". In particular the first sample illustrates that even under strong scaling (i.e. growth effect) from one fingerprint to the other recognition was still possible.[25]

For all cases, the scaled (with given scaling factor) and non-scaled configurations are given.

The example of Figure 36 is of particular importance. It shows that the fingerprint pattern of different size (with an age difference of 50 months) can be mapped closer together than any of the distortion samples of Figure 27. This observation is confirmed by many other examples with no counterexample so far. However, due to the limited number of "usable" fingerprints as explained above, this conclusion still needs to be treated with care. More data would be definitely desirable. **In addition, the limitation to just two fingerprints per finger did not allow further statistical analysis similar to what was done by University of Göttingen [28] where up to 48 fingerprints per finger were available.** In particular, the definition of a "mean point configuration" (as done in [28]) would not have made sense here.

Another observation from the presented and similar examples is: When comparing the scaled and non-scaled configuration, they either look very similar or they look very much like two differently distorted images. As all these pairs "matched" with the tested matching algorithms, it can be concluded that **the necessary scaling (or "growth adaption") challenges the matchers in the same way and order of magnitude as "normal" distortion can do** (cf. Figure 29 and Figure 30 of section 4.2). This can explain why the samples can be matched even with algorithms that do not take into account any growth aspects. It is important to stress that any conclusion in this way is only valid within the observation window of 24 to 54 months of time difference.

A confirmation of the prediction of the scaling factor of an isotropic mapping as suggested by the study of University of Göttingen was also tried. The deviation of the predicted scaling factor on the basis of growth charts of children[26] is shown in Figure 38. The average of deviations is 4.4% (standard deviation 0.27%), with only a few reaching up to

[25] The reader should note that for matching more features than the chosen landmarks are considered.

[26] See http://www.cdc.gov/growthcharts/

10% or even 20%. The larger deviations can be explained by distortion effects (within the fingerprint) and by deviations in the individual growth development of children. The statistical analysis therefore supports the Göttingen prediction.

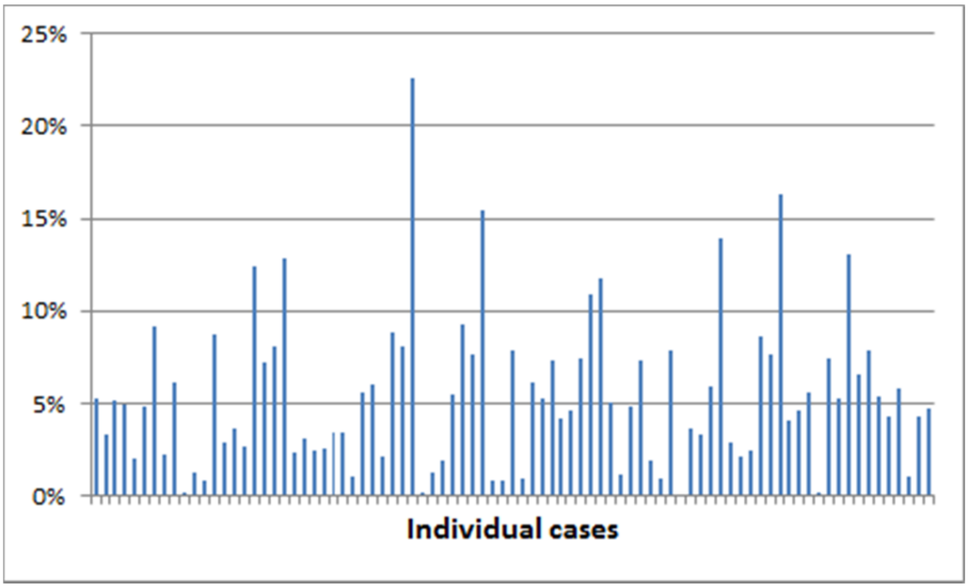

Figure 38: Relative deviation (in %) of scaling factor from prediction (average 4.4%, standard deviation 0.27%)

The improvement in the recognition rate when applying the growth factor as suggested by the researchers from Göttingen was as well verified. However, such tests were only possible where the coordinates of the templates used were transformed by the matching algorithms, in this case only NIST's bozorth3. Other matchers use vendor dependent templates which we cannot access directly.

Finding: Isotropic Growth Model (1)

Growth of fingerprints can be described by an isotropic model at good accuracy.

Alternatively, the images themselves could have been scaled but this possibility was rejected due to unknown and hardly traceable effects in the images' digitised configurations. Scaling in portions of pixel units would lead to transformations within the greyscale approximation of the original, and this effect would make the interpretation of the tests questionable. However, this is a promising experiment to be conducted in cooperation with interested vendors.

The results with bozorth3 were rather disappointing. Figure 39 shows the comparison of matching scores obtained without such a scaling factor (blue dots) and with the scaling factor suggested (red dots). Bozorth3 was applied to all corresponding fingerprint pairs. The cases were ordered on the horizontal axis according to the results without scaling (blue dots) in increasing order with respect to matching scores. For each case, the corresponding match result with scaling (in red) is displayed at the same horizontal position as the match results without scaling (in blue). In this way, the red dots illustrate the deviation from the normal match results. Obviously, the deviation can be positive or negative, with no clear trend visible.

A more detailed analysis of the individual cases revealed the image quality as reason. Matching for the non-scaled fingerprints failed mainly due to wrong minutiae detection. If such wrong minutiae positions are transformed they remain wrong and cannot improve matching.

Finding: Isotropic Growth Model (2)

Application of an isotropic growth model to fingerprint recognition does not necessarily improve recognition rate.

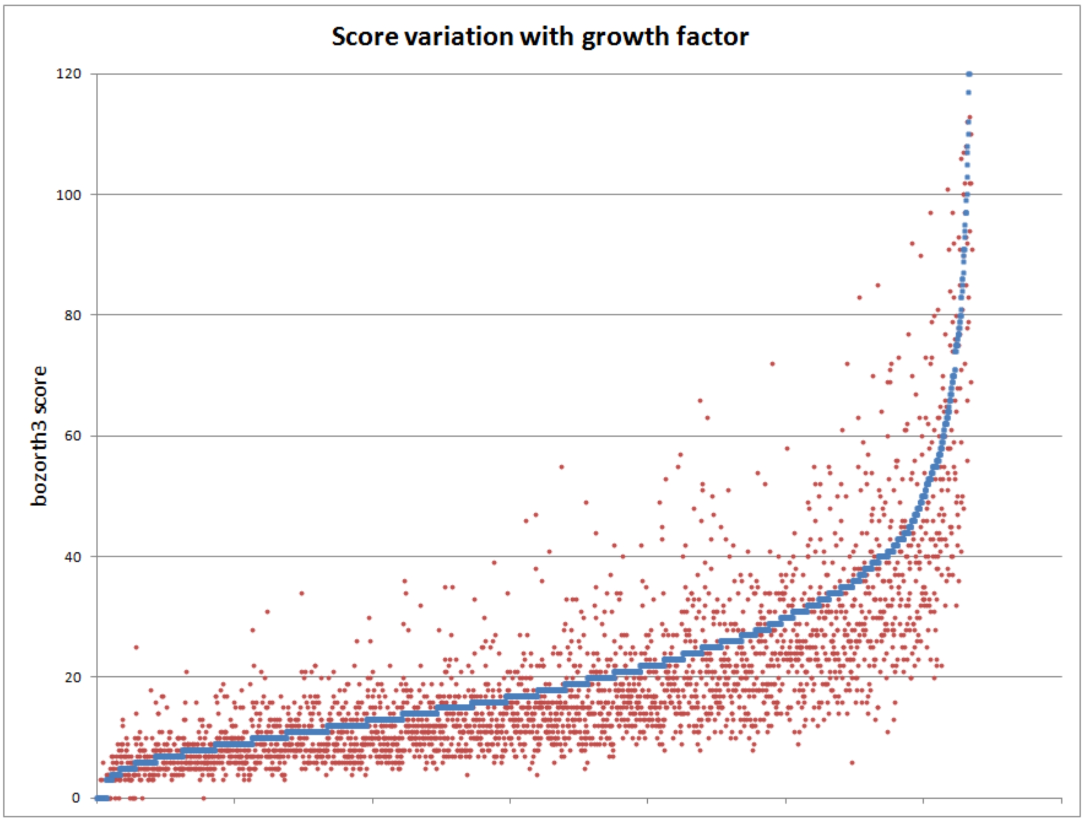

Figure 39: Comparison of matching scores with and without growth scaling factor.
(blue dots: without scaling, red dots: with scaling)

Chapter 6: **Impact of Acquisition Devices**

6.1 Problems with Optical Fingerprint Scanners

The images in the Portuguese database showed to a large extent typical deficiencies of fingerprints acquired with state-of-the-art optical fingerprint scanners (see Annex 1 for an overview of scanner types). These devices are almost exclusively used in such critical applications like border control that motivated this study.

The problems with this type of image acquisition were already discussed in section 2.1.3. Almost all of these problems have their origin at the point of contact of the finger with the device, usually a glass plate with optional coating.

- The pressure of the finger onto the device and the distribution of the pressure within the area of contact affect **contrast** and the **level of distortion**. On the other hand, sharpness is usually not an issue due to the fixed distance of the contact area to the imaging sensors. There is usually no built-in guidance for the person in question how he/she may best position the relevant finger on the device to achieve the best possible image. In many cases, not even the operator controlling the device can provide more assistance because the firmware of the device decides automatically when to take the image.
- The position of the finger (as a 3-dimensional object) on the device defines the portion of skin ridge pattern which is actually stored as fingerprint. As shown in Figure 4 on page 26, this portion may be different for each individual fingerprint taken from a specific finger, thus leading to **less overlapping areas** with less probability of matching.
- Finally, the **condition of the finger** itself has a strong impact on the image quality. If the finger is too dry, the ridge lines in the image tend to be discontinuous, introducing wrong predictions of ridge line endings and bifurcation points. If the finger is too wet, ridge lines in the image get thicken up to the level where individual lines can no longer be distinguished. Certain substances on the finger, even invisible for the eye, can create similar situations as with dry and wet fingers or additionally worsen such bad preconditioning.

Although these cases are well-known, the likelihood of occurrence increases further with children:

- The right usage of the scanner, both with respect to pressure and positioning, is already difficult to communicate to any test person and requires particular training of the test person and the operator (see sections 2.2.1 and 2.2.2). With children, communication and training is even more difficult.

- The typical behavior of children to touch anything with less care and aftercare clearly increases the probability that a child has critical substances at the finger to be scanned.

In conclusion, the case of the Portuguese data suggests that either more effort needs to be spent to develop best practice guidelines for the usage of optical scanners with children, and/or alternative devices should be considered.

6.2 Potential Alternative Devices

With regard to alternatives, three different types have been examined and compared to the traditional solution:

1. The single-finger multispectral devices from Lumidigm
2. The single-finger touchless devices from TBS Biometrics
3. The multi-finger Guardian device for "civil applications" from Cross Match, with improved user online user guidance

The three alternative devices have been tested against the following traditional devices:

4. The Fingerprint Live Scanner ZF1 from Dermalog (single-finger)
5. The 2D Enroll fingerprint scanner from TBS Biometrics (single-finger)

6.2.1 Multispectral sensors from Lumidigm

From the overall handling, multispectral fingerprint sensors offered by Lumidigm are the closest to the (single finger) optical devices explained before (cf. Figure 3 and Figure 40, resp.). Also here, the fingertip is placed on a small area with a glass plate. However, unlike traditional optical sensors, the optical unit "captures multiple images of the finger under different illumination conditions that include different wavelengths, different illumination orientations, and different polarization conditions" [32]. All individual images and its information are then merged together to derive what is called an "image equivalent to that produced by a conventional fingerprint reader". However, as this only "equivalent", the sensor misses important certification for use in law enforcement (see Appendix 1) which makes is also less attractive for border control.

On the other hand, thanks to its multispectral approach, the vendor claims stronger independence from the problems listed in section 6.1. The devices unfold their robustness in scenarios like large scale amusement park access, in replacement of tickets and with the inclusion of children. For a pure authentication application, this advantage could rule out the lacking law enforcement ability, provided the achievable error rates would be in an acceptable range.

For testing in section 6.3, the **"SBV-100 Fingerprint Reader"** has been used.

Figure 40: Multispectral sensor from Lumidigm[27]

6.2.2 Touchless sensors from TBS Biometrics

In contrast to the sensor types mentioned before, the relevant part of the fingertip does not touch the device even though the rest of the finger is placed in a special guide to facilitate the process (Figure 41).

Figure 41: Touchless sensor of TBS[28]

[27] Source: http://www.lumidigm.com/

[28] Source: http://www.tbs-biometrics.com

The touchless device produces an image which is a 2D projection of the real finger rather than an imprint. As the reference is always the traditional fingerprint (as described on page 23), the 2D projection is mapped to a simulation of the imprint. Again, this has an impact on certification (see Appendix 1) which has excluded also this type of devices from border control.

The device is similarly robust against preconditions of the finger as the multispectral device explained before. In contrast to that, the touchless device features also complete independence with regard to the positioning of the finger.

For testing in section 6.3, the **“3D Enroll”** of TBS Biometrics has been used. The device comes with a useful capture software which helps the user to position the finger inside the optical unit.

Figure 42: New Guardian from Cross Match[29]

6.2.3 New Guardian from Cross Match

At first glance, the new **Guardian** from Cross Match is very similar to the multi-finger devices offered by the same vendor for years. The main difference to the previous devices is the stronger interaction with the test person. A small screen displays a live image of the fingers to be enrolled, allowing for immediate reaction and change of finger

[29] Source : Cross Match Technologies

positioning and pressure on the device. In this way, the test person can help to increase the image quality.

6.3 *Experiments*

6.3.1 Selection of test persons

Even though this study was on child fingerprints, the following experiments have been conducted with adults only. The reason is – as explained in section 6.1 – that the actual quality issues are observable with all age groups, children and adults. Therefore, there was no particular need or advantage to recruit juvenile test persons for investigating qualitatively general aspects of fingerprint acquisition devices with respect to their resilience against negative factors. However, for real quantitative assessments (as in the case of preparing for large deployments), test persons of all age groups would be required.

The test persons have been selected exclusively from JRC staff and with the following minimum criteria:

- The test persons must be able to provide good fingerprints under ideal conditions in order not to bias the results by general inabilities (like insufficient ridge profile or the like).
- The test persons must be able to deliver similarly challenging fingerprints when exposed to same set of critical preconditions.

There were in total 6 test persons, three males and three females, aged between 31 and 53. Single-finger capturing was limited to left and right index fingers. As the number of test person was very small, the experiments could only reveal some trends but represent no performance evaluation whatsoever.

6.3.2 Set of experiments

The relevant fingers of the test persons have been preconditioned according to the following set of basic categories:

1. **"Best"** The most favourable conditions under which the test person could give fingerprints. This was achieved by interactively capturing and inspecting the achieved fingerprints. The target was always NFIQ=1 and IQF > 80.

2. **"Humid"** Extreme humidity of the fingers (simulated with skin lotion). This condition makes ridge lines generally very broad up to the level that differentiating between ridge lines becomes impossible. Thus, certain features can no longer be identified.

3. **“Sugar”** Presence of sugar on the finger (moistening of the finger with sugar solution, followed by the careful drying). This results in similar issues as for humidity.

4. **“Dirt”** Presence of fine granular ash on the finger (touching of ash with the finger, followed by erasing loose particles with a kleenex). This condition can lead to similar effects known from dry fingers as the latter would be difficult to simulate.

Despite the pressure, all test persons were trained to place the individual fingers in a way to avoid any additional distortion.

The experiments follow the simple experimentation protocol for each of the test persons:

For each of the conditions “best”, “humid”, “sugar”, “dirt”:
  Prepare the finger accordingly
  For each of the devices 1-5 from page 72:
    Acquire 3 images of the finger;
    Select the best image from “best” as reference for matching;
    Calculate the NFIQ values and the match result with the selected reference using NIST’s bozorth3 (see section 4.1.1)

In addition to this simple rationale, the following safeguards have been introduced:

- The device has been carefully cleaned after each and every fingerprint acquisition
- The preconditioned finger(s) were tested with device 4 (Dermalog sensor) before any acquisition with any other device was done in order to be sure the test with that device used still the same preconditioned finger(s).

6.3.3 Results

The variety of different quality fingerprints obtained with traditional fingerprint scanners is enormous (see Figure 43 for some examples). Depending on the chosen preconditioning, all NFIQ quality levels between 1 (best) to 5 (worst) can be achieved. While matching can be observed in quite good correlation with NFIQ for images acquired with traditional scanners (including the test device from Cross Match), it turned out to be misleading for the multispectral device from Lumidigm and the touchless from TBS.

Qualitatively, as the exercise was, the results can be summarised in the following Table 3. Reference to “recognition rate” is made in the sense that the sample was compared to the reference sample as explained before. “Weak recognition” refers to bozorth3 scores of less than 40, “strong recognition” to scores well above 60.

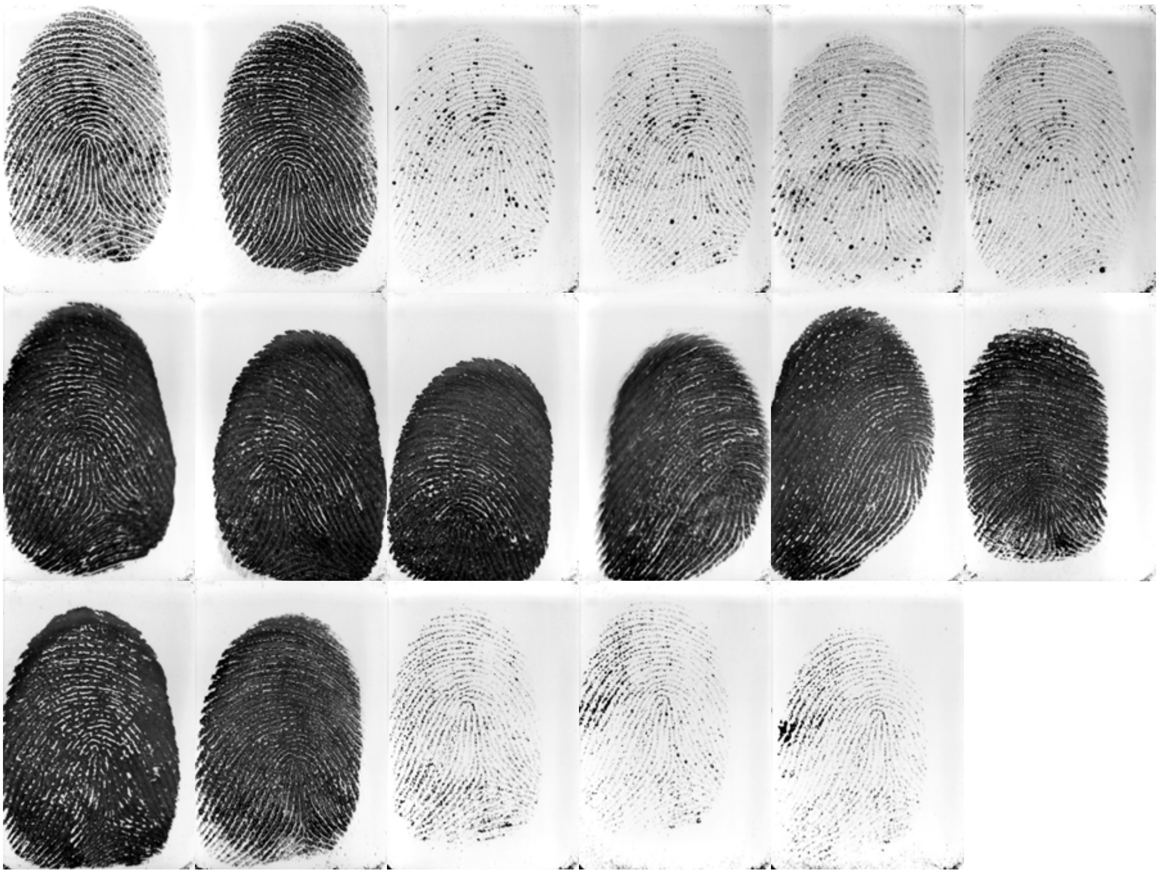

Figure 43: Fingerprint samples of different quality from the same finger

	Traditional (Dermalog/TBS 2D/ Cross Match)	Multispectral (Lumidigm)* * (see comment below)	Touchless (TBS)
Best	Strong recognition at NFIQ 1	Weak recognition though NFIQ was 1-2.	Strong recognition at NFIQ 1-3
Humid	Weak recognition with NFIQ at 4-5.	Weak recognition though NFIQ was 1-2.	Weak to strong recognition rate at NFIQ 3-4.
Sugar	Recognition mostly weak at NFIQ 3-5.	Weak recognition though NFIQ was 1-2.	Strong recognition at NFIQ 1-3
Dirt	Weak to strong recognition at NFIQ 4-5	Weak recognition rate low at NFIQ of mostly 1.	Strong recognition at NFIQ 1-2

Table 3: Qualitative result of fingerprint scanner comparison
(According to the conditions defined in section 6.3.2)

The immediate conclusion is that:

- The preconditioning very well simulated the observed problems with the Portuguese data
- The images obtained with the multispectral device are obviously not compatible with the images obtained with traditional scanners. This is in so far surprising as the corresponding NFIQ values were quite good.
- The images produced by the touchless device are largely compatible with images of traditional scanners. Moreover, the device is able to increase recognition in most of the cases, except for those characterized by “humid”. Unfortunately the quality indication of NFIQ was misleading in many cases.
- The Cross Match device was not much different from its image qualities than the other traditional devices. However, the improved user guidance has to be seen as beneficial for finding the right position and pressure onto the device in order to improve the quality.

Regarding the Lumidigm device it has to be noted that the images do match very well with images of the same device, regardless of the preconditioning (measured again with bozorth3). To that extent, the claims of the vendor about the resilience against such influences could be verified. However, the images have the mentioned incompatibilities with images of traditional scanners.

In summary, both the touchless device of TBS and the multispectral device of Lumidigm have promising abilities to deal with critically preconditioned fingers. The development of best practices with these devices (especially with children) could help to overcome some of the incompatibilities with current state-of-the-art fingerprinting techniques, in particular with regard to the nature of the images and the quality metrics.

Finding: Alternative acquisition devices

Touchless and multispectral devices have promising abilities to deal with critically preconditions fingers. Relevant best practices should be developed.

Chapter 7: **Conclusions and Recommendations**

7.1 Conclusions

The findings described in chapters 4-6 can be consolidated to the following overall conclusions:

- **Growth has limited influence on fingerprint recognition.**

  Although the time difference was predicted to be the most important factor children fingerprint recognition, all tested algorithms showed the same recognition rate regardless of the time between the fingerprints for the observation window of up to 4.5 years (see the finding about “Independence of Time Difference” on page 55). This suggests that the problems reported in the past and within that time window would come from a different factor.

- **Size (in terms of the dimensions of the relevant fingerprint characteristics) – and implicitly age – does not constitute any theoretical barrier for automated fingerprint recognition.**

  Again, within the available investigation window of up to 4.5 years between the acquired fingerprints, there was no theoretical barrier observed for proper automated recognition by current matching algorithms – provided the images are of sufficient quality (see the finding about “Young children recognisable” on page 55). There was no age group for which recognition rate was zero or close to zero. Nevertheless, the quality problem becomes more important the smaller the relevant structures to be compared are.

- **Image quality (in terms of low contrast and distortion effects) is the ultimate problem for children’s fingerprints, and image quality is strongly influenced by size.**

  Though the observed image issues are well-known also from adults, with smaller structure sizes of child fingerprints, the issues get worse and the probability for it increase. Therefore, proper enrolment is the key factor for successful recognition (see the finding about “Factors determining image quality” on page 62). This aspect remains as the major reason for the previous observation that fingerprint recognition of children gets worse the younger the children are. If it is neither the structure size as such (and thus the age) nor the age difference, then the image quality has to become the main focus.

- **Relevant quality metrics for fingerprints need revision with regard to the children case.**

  As far as quality metrics for fingerprint images assume feature dimensions for adults, adaption to children fingerprints is necessary. Otherwise, the reported quality scores might mislead any prediction about later recognisability and might in this way also mislead the acquisition process (see the finding about "Quality metric could be inadequate" on page 56). It is highly recommendable that similar data sources as used in this study are available for the adaption.

- **Isotropic growth model may serve as a good approximation to cover changes over time.**

  The underlying data from Portugal with only two fingerprints per test item do not allow for a clean distinction of distortion from other effects. However, the linear (isotropic) model seems to be sufficient to estimate the real level of impact that the growth effect has, if any (see the finding about "Isotropic Growth Model" on page 69). If algorithms are modified under this assumption, the recognition rate should rise.

- **Alternative acquisition devices for fingerprints should be seriously considered in the future.**

  Experiments with multi-spectral, touchless and novel four-finger capture devices, gave promising indications on how the quality issues could be better managed – on top of well-known best practice guidelines for the improvement of quality in fingerprint acquisition (see the finding about "Alternative acquisition devices" on page 78).

These conclusions confirm that ***fingerprint recognition of children aged between 6 and 12 years is achievable with a satisfactory level of accuracy***, provided appropriate best practice guidelines are developed and achievable within certain constraints of technical and organisational nature. Further results from the study may help to develop these guidelines

7.2 Recommendations

According to the findings of this study, a number of recommendations can be given for the future development of children fingerprinting techniques:

- **Image quality is the key.** As experienced during the BIODEV II study, good fingerprint image quality is not straightforward to achieve (see the finding about "Best practice is key" on page 33). A certain minimum level of training of operators and data subjects is necessary to acquire high quality images. However, training needs to be designed for the particular setting in which the fingerprints acquisition will be carried out. This encompasses the selected technology but also the particular operation environment and the particular data subjects to be enrolled. Good practices developed by projects like BIODEV II should be carefully studied and further developed .

- **Matching algorithms can be further improved.** The experience with Vendor 2 (see section 4.1.4) suggests that there is still some room for improvement in the field of fingerprint matching (see the finding about "Algorithms can be improved" on page 55). Potential improvements have also to be seen for time differences well beyond 5 years which could not be investigated by this study. However, the results from the University of Göttingen (see section 2.2.4) clearly demonstrate the benefits of certain adaptations to matching algorithms. Therefore, it is recommended all developers and vendors of fingerprint recognition algorithms to use JRC's offer to test their current and potential new algorithms in a rigorous and independent manner.

- **Availability of relevant test data.** The important insights gained from the Portuguese data with respect to realistic automated recognition for children fingerprints emphasised clearly the need for long term availability of such data for relevant research and development. Such EU wide data repository needs to be accessible permanently as a unique independent reference and benchmark, while securing the access to this type of sensitive data under solid conditions and stringent modalities.

- **Selection of acquisition devices.** Experiments with multi-spectral, touchless and novel four-finger capture devices, gave promising indications on how the quality issues could be better managed – on top of well-known best practice guidelines for the improvement of quality in fingerprint acquisition (see the finding about "Alternative acquisition devices" on page 78). These emerging technologies should be further explored and validated.

7.3 Open questions

Despite the efforts of this study, some questions remain open due to the limitations of the available data:

- **Calibration with Adult Data.** The impact of the child specific aspects still need to be more clearly distinguished from the general quality degrading aspects. Therefore, the recognition performance of fingerprints of adults needs to be compared to that of child fingerprints where the adult data were acquired under similar conditions to those of the children's data (see the finding about "Limitation of available data” on page 41). This would allow to predict the performance loss for children in the absence of any particular compensation measure.

- **Enrolment Tests.** In order to quantify a practical age limit, given the best available technology, larger field trials on enrolment of children need to be conducted (see the finding about “Best practice is key” on page 33 and the finding about “Alternative acquisition devices” on page 78). These trials should further investigate and quantify the impact of certain enrolment devices and procedures.
- **Refined Growth Model.** The current results do not contradict the assumption of an almost isotropic growth model as suggested by an earlier study (see the finding about “Isotropic Growth Model” on page 70). At least, it seems suitable as a first order approximation for improving algorithms in cases where the time difference is not greater than the one considered in the present study (i.e. 4.5 years). However, it is desirable to draw conclusions for longer time windows (beyond 5 years) in order to give a clear message to developers of fingerprint recognition systems.

Appendix 1: Important Concepts Related to Fingerprints

Fingerprints as Unique Identifiers

The idea of using fingerprints (i.e. the skin ridge structure) for uniquely identifying a person is mainly based on Sir Francis Galton's published work "Finger Prints" [2] in which he gave for the first time a statistical model of fingerprint analysis and identification and encouraged its use in forensic science. Although the uniqueness of fingerprints was assumed already long before, it was Galton who first provided a detailed analysis of this phenomenon.

Identification using fingerprints is based on the assumption that ridge structures do not change from birth to death, except with injury or disease and they possess an infinite variety of detail that is not repeated in other areas of the friction skin [2]. Ridge patterns are developed during the human foetus phase [1]. From this assumption it is concluded that two persons cannot have the same fingerprint patterns[30].

In order to compare such ridge structures, a classification of its elements has been developed which consists of 3 levels [31]:

1. **Basic fingerprint patterns**: arches, loops and whorls (Figure 44) including the number of ridges involved.

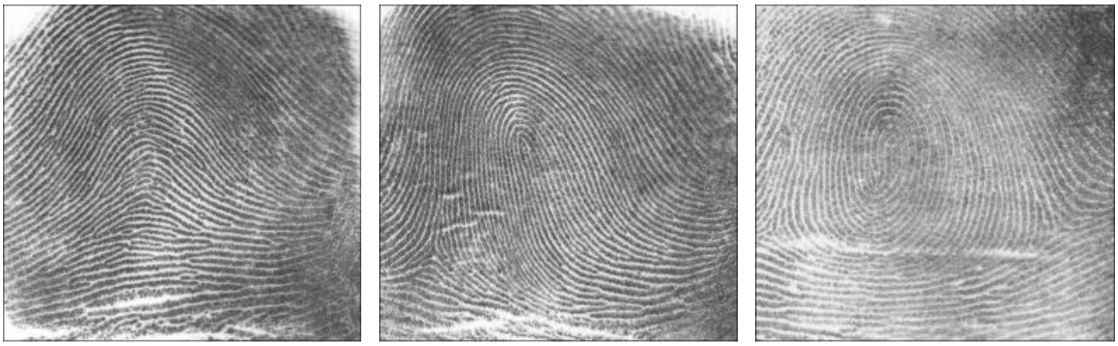

Figure 44: Arch, loop and whorl[31]

---

[30] This assumption is not undisputed. Until 2002, there were already 20 court cases in the US in which fingerprint evidence had been challenged (cf. [5]).

[31] Source: Wikimedia Commons

2. **Minutiae**: mainly endings and bifurcations of ridge lines (Figure 45). Fingertips usually have about 100 minutiae out of which only about 30-60 are considered when taking flat fingerprints.

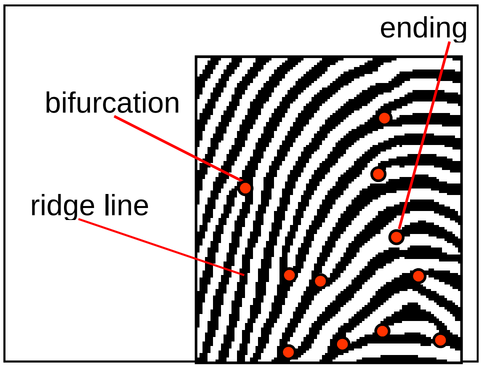

Figure 45: Minutiae[32]

3. **Pores and ridge contours** (Figure 46): Here, the finer structures along the ridges are considered and the distribution of pores along the ridges. There are about 6 pores/mm^2.

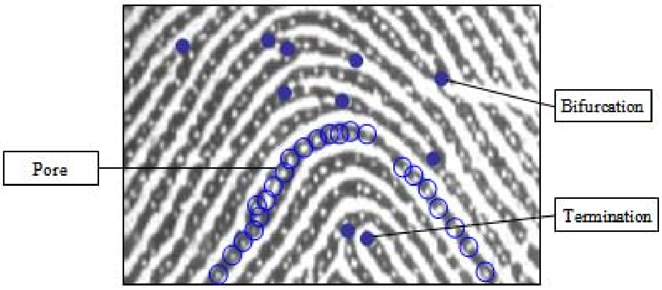

Figure 46: Pores on ridges[33]

Comparison is usually based on level 1 and 2 features only, also because level 3 features require much higher resolution images in order to be identified.

Technically, fingerprint recognition is the process of **comparing two fingerprints** with respect to a "sufficient" number of "identical" features. In the beginning of fingerprint recognition, such evidence has been established by a particularly trained person, a fingerprint (or dactyloscopy) expert. "Sufficient" usually translates into a fixed number of level 2 features (minutiae); "identical" means they should be at the same position. However, if completely different level 1 features are observed, this type of level 2 comparison is not necessary.

[32] Source: University of Bologna

[33] Source: University of Bologna

Automated Fingerprint Recognition

The automated comparison (by an algorithm) of two fingerprints is usually done on the basis of the extracted (local) features, i.e. the minutiae (see section 2.1). However, there exists also a broad range of other approaches to estimate the similarity of two fingerprints [31]. We concentrate here on minutia-based algorithms because they are most widely deployed and minutiae comparison is usually required for providing legal evidence.

Let A be the set of features of the first fingerprint and B the set of features of the second fingerprint (both expressed as a vector of minutiae parameters as explained above: type, position, angle, and optional quality score), then the matching can be expressed as a scoring function s(A,B). The scoring function determines “how close” A is to B (the “distance”) and delivers usually a value between 0 (=”absolute different”) and 1 (=”absolute identical”) or between any other fixed bounds. If the score is above a given threshold, then A “matches” B, i.e. the fingerprints are supposed to be from the same person. If the score is below that, the fingerprints do not match.

It is here where the statistical nature of fingerprint recognition becomes obvious. As the process of arriving at the sets A and B is distorted by a couple of aspects (as explained before), we can only expect with a certain likelihood that enough criteria can be found to clearly separate between a match and a non-match. On top of this, the scoring function itself is a well-preserved company secret although some elements might be known (e.g. the minimum number of minutiae that makes a match). Further aspects which may influence the scoring are:

- The numeric tolerances for the minutiae parameters.
- The consideration of the scoring parameter for a minutia
- The consideration of logic dependencies between minutiae.
- The consideration of found non-matching minutiae.
- The consideration of global parameters (level 1 features).

Thus, if both fingerprints have been enrolled with some degree of “noise”, then this may lead to non-recognition of certain minutiae or to the introduction of phantom minutiae. Furthermore, even in case the fingerprints should match, noise will lead to slightly different coordinate and angle values. Therefore, if the matching score is expressed in percentages of “similarity”, then even for the samples of the same persons the score will rarely reach 100% but is statistically disturbed.

A typical distribution of score values is depicted in Figure 47. It shows the characteristic difference between three types of comparisons:

- genuine users against the reference samples stored in a database
- inter-template comparisons, i.e. samples against each other (but not against themselves)
- impostors against any sample in the database

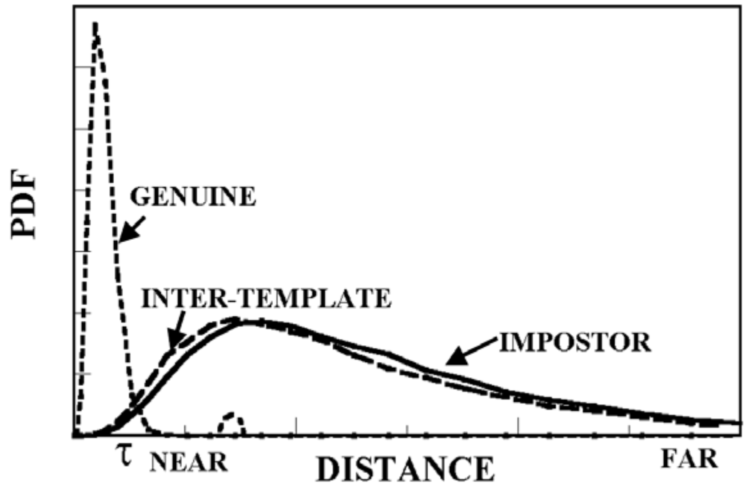

Figure 47: Example of distribution of score values of biometric matching[34]

The "Distance" axis from "near" to "far" denotes the score values in reverse order from 100 to 0. PDF denotes the probability distribution function.

Error Rates

To deal with the statistical nature of automated fingerprint recognition (and biometrics in general), a number of characteristic error rates have been introduced in order to measure the performance:

- False Acceptance Rate (FAR): The probability of being identified as another person[35].
- False Rejection Rate (FRR): The probability of not being recognised[36].
- Failure to Enrol (FTE): The probability of not being able to register a fingerprint[37].
- Failure to Acquire (FTA): The probability of the system's inability to capture or locate an image of sufficient quality.

If only single transactions are considered (without the possibility of repeated trial), the terminology "False Match Rate (FMR)" and "False Non-Match Rate (FNMR)" are used

---

[34] Source: [20]

[35] i.e. the score function delivers a value which is considered as "match" but the fingerprint is actually from a different finger.

[36] i.e. the score function delivers a value which is considered as "non-match" but the fingerprint is actually from the same finger.

[37] e.g. in cases where the finger skin is inappropriate for fingerprinting.

instead of FAR and FRR. There is also a dependency between the FAR and the FRR: If the FAR is brought down, the FRR usually increases, and vice versa.

Referring to Figure 47, the FMR is caused by the portion of the impostors' score distribution which is to the left of the threshold τ. The FNMR is caused by the portion of the genuine users' score distribution to the right of the threshold τ. Hence, FMR and FNMR depend both on the chosen threshold τ. For that reason, corresponding values of FMR(τ) and FNMR(τ) (or the more comprehensive error rates FAR and FRR) are depicted in a graph which is called Receiver Operating Characteristic (ROC) and which best reflects the recognition performance of a biometric system (see Figure 48).

With respect to the general perception of a universally applicable concept of identification, it is interesting to note that in 2002 Pankanti et al. [5] stated that:

"Our results show that 1) contrary to the popular belief, fingerprint matching is not infallible and leads to some false associations, 2) while there is an overwhelming amount of discriminatory information present in the fingerprints, the strength of the evidence degrades drastically with noise in the sensed fingerprint images, 3) the performance of the state-of-the-art automatic fingerprint matchers is not even close to the theoretical limit, and 4) because automatic fingerprint verification systems based on minutia use only a part of the discriminatory information present in the fingerprints, it may be desirable to explore additional complementary representations of fingerprints for automatic matching."

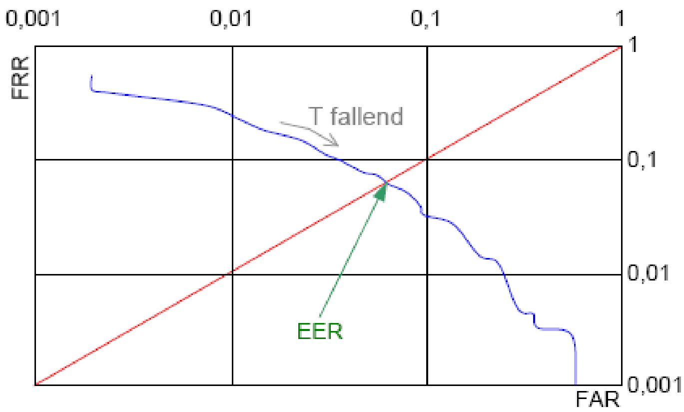

Figure 48: Example of a ROC curve[38]

Due to the mostly unknown approach of a matching system (apart from performance figures derived from tests on large fingerprint databases), it can only be estimated experimentally to what extent this statement from 2002 is still valid.

[38] Source: BSI

From a legal point of view, it has also to be noted that the considerable lack of transparency of matching algorithms implies also a lack of transparency why the decision for a match or a non-match has been taken. Even if it is considered as a “black decision box” with a certain acceptable probability of failure, it has to be questioned whether this probability can be estimated in practice.

Image Quality Metrics NFIQ

Even scanners in accordance with the quality certification mentioned in section 0 may capture images of different quality from the same finger. Differences exist mainly with respect to greyscaling, the introduction of (wrong) artefacts or the blurring of complete regions due to different treatment of surface conditions of the finger. Distortion of the finger skin while the image is taken may lead to wrong minutiae characterisation (i.e. wrong coordinates, wrong angles, even wrong type).

In the lack of a suitable international standard for fingerprint image quality, the ultimate reference is the NIST Fingerprint Image Quality (NFIQ) algorithm [15] which is basically a prediction of the performance of minutiae-based fingerprint matching systems. The statistical rationale behind the 5 quality levels (from 1=high to 5=low) is depicted in Figure 49 in which the confidence intervals[39] with respect to the True Acceptance Rate (TAR) for all 5 levels are displayed. For instance, the TAR of level 4 images is with a probability of 95% roughly between 91,5 % and 94,5 %.

Usually, this algorithm is applied during the registration process of fingerprints in order to accept or reject certain samples and to select the best possible image for further processing. However, recent experience in the BIODEV II study (see section 2.2.2) have demonstrated a significant inappropriateness of this algorithm for current deployments for which we had tried to find an explanation. In fact, the NIST algorithm has as one of its major elements a so-called neural network which is used to approximate a complicated non-linear function (in this case, the behaviour of an “average” matching algorithm). This neural network had been “trained” by a couple of vendor specific matching algorithms available at the time of development (2004). Since then, many improvements have been made to those algorithms and new algorithms have appeared on the market. Therefore, the neural network part of the NIST algorithm would need to be continuously updated or at least additionally trained towards the actual matching algorithm in use[40].

---

[39] The interval in which the true value is expected with a probability of 95%.

[40] In fact, NIST is currently conducting a revision of NFIQ towards a “NFIQ 2.0” (see http://www.nist.gov/itl/iad/ig/development_nfiq_2.cfm)

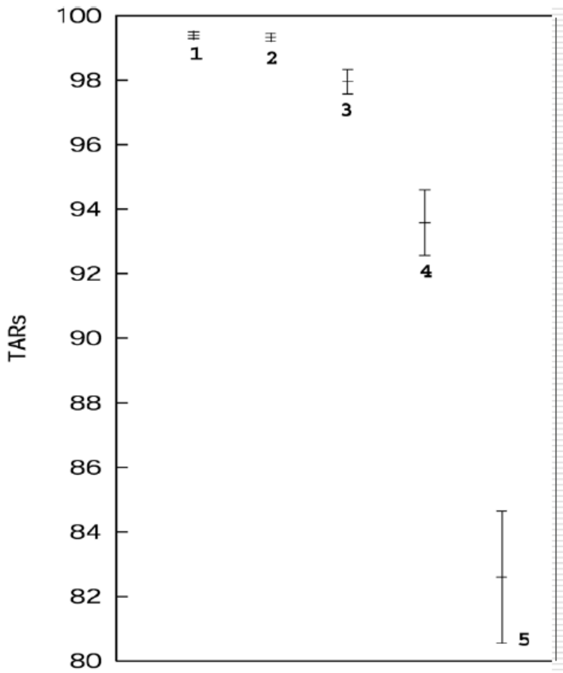

Figure 49: NFIQ Statistical Rationale

For each quality level, NIST calculated 95% confidence intervals of TARs[41]@ FAR[42]=0.1% for six matchers and sixteen datasets.[43]

Another important question is the appropriateness of NFIQ (even a successor version) for the case of children fingerprints. If existing matching algorithms are used to train the software, then their problems with children's fingerprints will be inherited to the NFIQ algorithm. As explained in section 2.1, matching algorithms may assume for performance reasons a certain distance between ridge lines and therefore, in some cases, not even recognise a child's fingerprint as being a genuine fingerprint at all.

---

[41] TAR = True Acceptance Rate (the percentage of correct identifications)

[42] FAR = False Acceptance Rate (the percentage of false identifications)

[43] Source: NIST

Image Quality Metrics IQF

The Image Quality of Fingerprint (IQF) software application has been designed to measure the visual quality of a digital fingerprint image, i.e., the apparent quality of the softcopy displayed image presented to a human observer who is knowledgeable in fingerprint assessment [16]. As such, it has been developed by the MITRE Corporation[44] mainly for latent fingerprints found at crime scenes in order to provide the examiner with a tool to measure the image quality. However, IQF is also applicable to livescan images from fingerprint scanners.

IQF is a fingerprint-tailored version of MITRE's general purpose Image Quality Measure (IQM). IQF and IQM compute image quality based on the two-dimensional, spatial frequency power spectrum of the digital image. The power spectrum, as the square of the magnitude of the Fourier transform of the image, contains information on the sharpness, contrast, and detail rendition of the. In IQF, the power spectrum is normalized by image contrast, average gray level (brightness), and image size; a visual response function filter is applied, and the pixels per inch resolution scale of the fingerprint image is taken into account. The fundamental output of IQF is a single-number image quality value which is the sum of the filtered, scaled, weighted power spectrum values.

In a comparison of the IQF values of these test images with their visual image quality, the IQF magnitudes appear to loosely correspond to the 4 usability classifications defined by the Biometric Application Programming Interface Consortium for biometric quality metrics having a 0 to 100 magnitude range [38]. For 500 dpi fingerprint images this correspondence is [16]:

BioAPI Quality Score (~IQF value):	BioAPI Biometric Utility:
0 - 25	unacceptable
26 - 50	marginal
51 - 75	adequate
76 – 100	excellent

Fingerprint Scanner Types

The acquisition of the fingerprint images to be compared is done either by scanning the impressions of inked fingers on paper (as done over years by law enforcement) or by the direct use of a fingerprint scanner. For the latter, mainly three different approaches are in use for larger scale applications:

---

[44] http://www.mitre.org

- *Capacitive sensors:* The finger is placed on a silicon chip and the different capacitive characteristics of fingerprint valleys and ridges generate the image electronically (Figure 50). The advantage is low cost and small size but the durability of the silicon plate is limited.

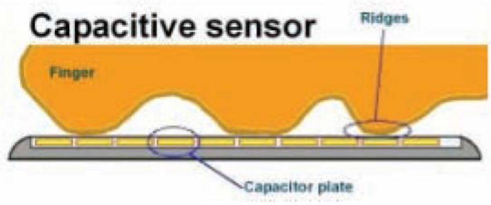

Figure 50: Principle of capacitive sensor[45]

However, this type of scanner is rarely used for the new type of governmental biometric authentication.

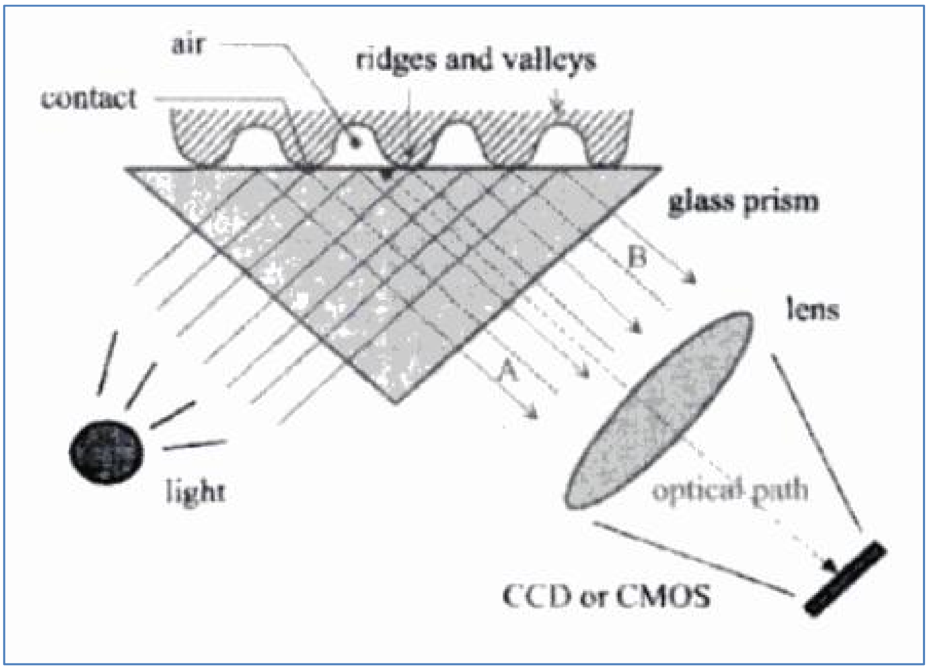

Figure 51: Principle of optical sensor[46]

---

[45] Source: Bundesamt für Sicherheit in der Informationstechnik (BSI)

[46] Source: [12]

- *Optical sensors:* A CCD[47] or CMOS[48] sensor registers the image of the fingerprint, usually when pressing the finger against a coated glass or plastic platen (see Figure 51). The sensor generates a digitised image on which the ridges and valleys appear as black, grey and white lines. Optical sensors tend to deliver better quality images, but are of larger size and higher cost than capacitive sensors. For applications like border control, this type of device is almost exclusively used (see also Figure 3 for an example).

- *Multispectral sensors:* This relatively new type of sensor overcomes the need of capacitive and optical sensors for a complete contact between the fingerprint and the sensor. By using multiple spectrums of light and advanced polarisation techniques the sensor extracts fingerprint characteristics from both the surface and subsurface of the skin (Figure 52). As a by-product, multispectral sensors are less vulnerable against spoofing attacks. The price is substantially higher than for a traditional optical sensor.

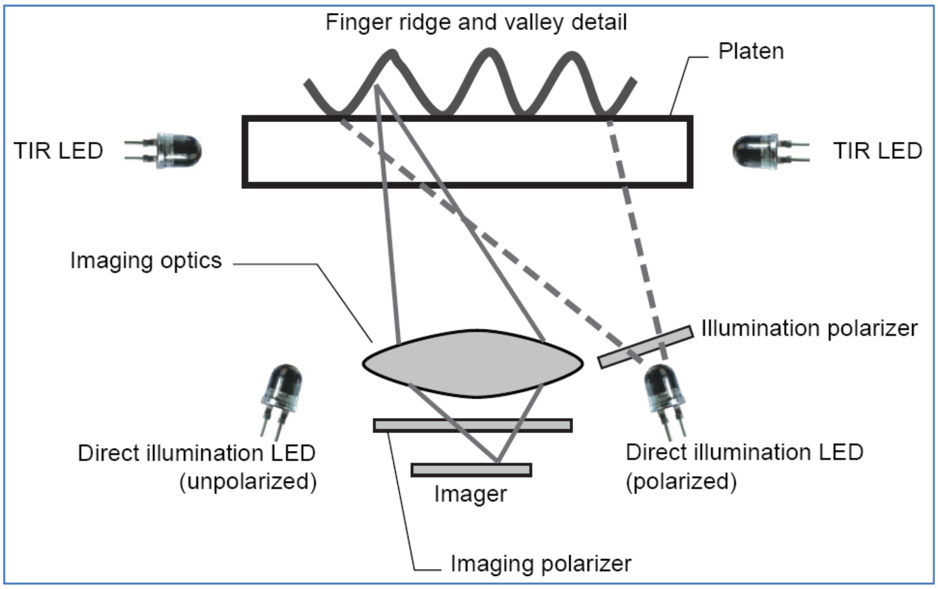

Figure 52: Schematic principle of multispectral sensors[49]

- *Touchless sensors:* In contrast to the sensor types mentioned before, the relevant part of the fingertip does not touch the device even though the rest of the finger is placed in a special guide to facilitate the process (Figure 53). Touchless sensors use special optical systems which take an image of the fingertip. As this image is

[47] Charge-coupled device

[48] Complementary metal–oxide–semiconductor

[49] Source: [31]

not identical with a “flat” fingerprint (as an imprint of the fingertip), the shape of the fingerprint is calculated from the image in order to be comparable with “real” fingerprints. The clear advantage of this type of devices is that they avoid all problems the other sensors have with a contacting plate (distortion, dirt, humidity, etc.). On the other hand, they require stronger computing power locally or remotely (an important cost factor) in order to calculate the fingerprint.

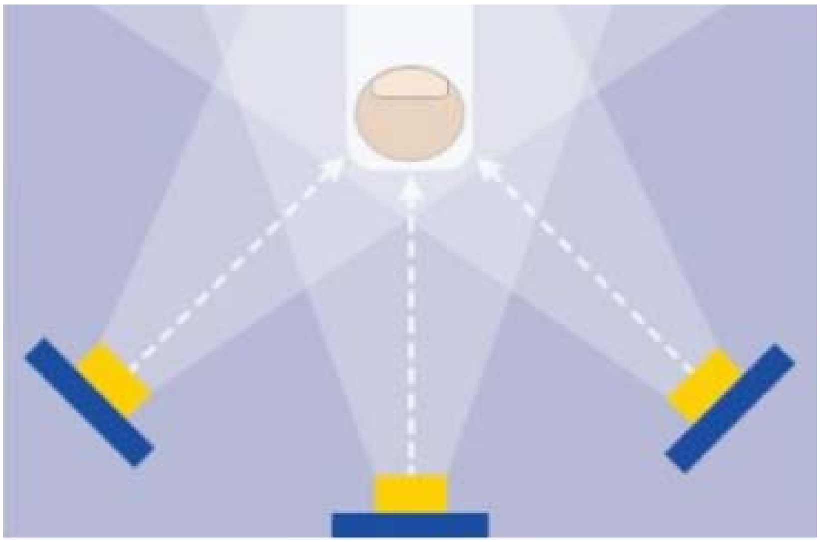

Figure 53: Schematic View Touchless Fingerprint Scanner[50]

Relevant Standards for Fingerprint Readers

In order to guarantee that the scanning devices do not introduce themselves additional distortions to the image, the following set of quality parameters have been defined by various institutions:

- FBI Image Quality Specifications (IQS) for fingerprint scanners, established for the US Personal Identification Verification program, whose aim is to improve the identification and authentication for access to US Federal facilities and information systems.
- PassDEÜV established by the Bundesamt für Sicherheit in der Informationstechnik (BSI) for the capture and quality assurance of fingerprints by the passport authorities and the transmission of passport application data to the passport

[50] Source: TBS Biometrics

manufacturers[51]; the PassDEÜV requirements are identical to the relevant FBI requirements [13] except for the acquisition area, which can be smaller.

- CNIPA-A/B/C: these three new set of specifications were developed by the biometric research group of University Bologna on behalf of the Italian CNIPA (the Italian National Center for ICT in the Public Administration) for inclusion within the guidelines for the Italian public administrations involved in biometric projects.

In a recent paper of Alessandroni et al. [14], it was demonstrated that the CNIPA quality criteria have the best balance between cost of the devices and their performance. Nevertheless, all the mentioned standards cover only the precision of the device but do not specify any particular countermeasure for the critical issues listed in section 2.1 and section 6.1.

[51] http://www.bsi.de/english/publications/techguidelines/tr03104

Appendix 2: List of References

[1] H. Cummins: The Topographic History of the Volar Pads in the Human Embryo. Contributions to Embryology; No. 394; Carnegie Institution of Washington: Washington, DC, pp 103-126 (1929).

[2] F. Galton: Finger Prints. MacMillan and Co, London (1892).

[3] E.R. Henry: Classification and uses of finger prints. Republished by Xlibris Corporation (2008)

[4] D.R. Ashbaugh: Ridgeology – Modern Evaluative Friction Ridge Identification. Royal Canadian Mounted Police, 1999. Available at http://www.onin.com/fp/ridgeology.pdf

[5] S. Pankanti et al.: On the Individuality of Fingerprints. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 8, pp. 1010-1025 (2002).

[6] M. Kralik, V. Novotny: Epidermal ridge breadth – An Indicator of Age and Sex in Paleodermatoglyphics. Variability and Evolution, Vol. 11, pp 5-30 (2003).

[7] A.F. Hecht: Über das Hand- und Fußflächenrelief von Kindern. Research in Experimental Medicine, Vol. 39, No. 1, pp. 56-66 (1924).

[8] D.C. Rife: Dermatoglyphics as Ethnic Criteria. Am J Hum Genet. Vol. 6, pp 319-327 (1954).

[9] D.C. Rife: Finger Prints as Criteria of Ethnic Relationship. Am J Hum Genet. Vol. 5(4), pp 389-399 (1953).

[10] H.F. Swofford: Fingerprint Patterns: A Study on the Finger and Ethnicity Prioritized Order of Occurrence. Journal of Forensic Identification, Vol. 55, No. 4 (2005)

[11] S. Nanavati et al.: Biometrics – Identity Verification in a Networked World. Wiley Computer Publishing, New York (2002)

[12] G. v. Grävenitz: Erfolgskriterien und Absatzchancen biometrischer Identifikationsverfahren. LIT Verlag, Berlin-Hamburg-Münster (2006)

[13] Department of Justice, FBI: “Electronic Fingerprint Transmission Specification”. CJIS-RS-0010 (V7), January 1999.

[14] A. Alessandroni et al.: Definition of Fingerprint Scanner Image Quality Specifications by Operational Quality. In: Biometrics and Identity Management, Lecture Notes in Computer Science, Volume 5372, pp 29-36 (2008)

[15] E. Tabassi et al.: Fingerprint Image Quality. NISTIR 7151, National Institute of Standards and Technology (2004)

[16] IQF (Image Quality of Fingerprints) Software Application. MITRE Technical Report MTR 070053. MITRE (2007)

[17] Evaluation Report Biometrics Trial: 2b or not 2b. TNO (2005)

[18] A.J. Mansfield, J. Wayman: Best Practices in Testing and Reporting Performance of Biometric Devices. National Physical Laboratory (2002)

[19] T.A. Louis: Confidence intervals for a binomial parameter after observing no successes. The American Statistician, Vol. 35(3), p 154 (1981)

[20] J.L. Wayman: Technical Testing and Evaluation of Biometric Identification Devices. In: Biometrics: Personal Identification in a Networked Society, Kluwer Academic Publishers, Dordrecht (1999)

[21] S. Dass et al.: Statistical Models for Assessing the Individuality of Fingerprints. Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05), 2005.

[22] G. Casella, R. L. Berger: Statistical Inference. Wadsworth & Brooks/Cole, Belmont (1990).

[23] T. Dunstone, N. Yager: Biometric System and Data Analysis: Design, Evaluation, and Data Mining. Springer, New York (2008)

[24] M. Okajima: Development of dermal ridges in the fetus. Journal of Medical Genetics, Vol. 12, pp 243-250 (1975).

[25] L. S. Penrose, P.T. Ohara: The Development of the Epidermal Ridges. Journal of Medical Genetics, Vol. 10, pp 201-208 (1973).

[26] Evaluierung biometrischer Systeme Fingerabdrucktechnologien – BioFinger. Bundesamt für Sicherheit in der Informationstechnik (2004)

[27] Untersuchung der Leistungsfähigkeit von biometrischen Verifikationssystemen – BioP II. Bundesamt für Sicherheit in der Informationstechnik (2005)

[28] C. Gottschlich, T. Hotz, R. Lorenz, S. Bernhardt, M. Hantschel, A. Munk: Modelling the growth of fingerprints improves matching for adolescents. arXiv, 1008.1150v1 (available at http://arxiv.org/abs/1008.1150v1)

[29] I. L. Dryden and K. V. Mardia: Statistical Shape Analysis. Wiley, Chichester (1998)

[30] T. Mansfield: Sample size in planning a biometric evaluation. National Physical Laboratory, U.K. (May 2010)

[31] D. Maltoni et al.: Handbook of Fingerprint Recognition. 2nd edition, Springer (2009)

[32] R.K. Rowe et al.: Multispectral Fingerprint Image Acquisition. Whitepaper, Lumidigm Inc. (2008) (available at http://www.lumidigm.com/download/Multispectral-Fingerprint-Image-Acquisition.pdf)

[33] F. Rahmun, O. Bausinger: Improving performance by improving fingerprint image quality – Experiences from pilot project BioDEVII. Presentation at the International Biometrics Performance Conference, Gaithersburg, USA, 2 March, 2010 (available at http://biometrics.nist.gov/cs_links/ibpc2010/pdfs/Rahmun_Fares_BausingerOliver_20100303_BestPracticeFingerprintEnrolmentVIS.pdf)

[34] Quantifying the Dermatoglyphic Growth Patterns in Children through Adolescence. Report NJI contract 2005-DD-BX-K056, UltraScan Corp. (2010) (available at http://www.ncjrs.gov/pdffiles1/nij/grants/232746.pdf

[35] Council Regulation (EC) No 2252/2004 of 13 December 2004 on standards for security features and biometrics in passports and travel documents issued by Member States. See: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2004:385:0001:0006:EN:PDF

[36] EU–Passport Specification. Working document (28.06.2006) http://ec.europa.eu/dgs/home-affairs/e-library/docs/comm_native_c_2006_2909_f_en_annex_draft_not_notified.pdf

[37] Commission Decision C(2011) 5499 amending Commission Decision C(2006) 2909 laying down the technical specifications on the standards for security features and biometrics in passports and travel documents issued by Member States http://ec.europa.eu/dgs/home-affairs/e-library/docs/comm_native_c_2011_5499_f_en.pdf

[38] BioAPI Specification Version 1.1, 16 March 2000. The BioAPI Consortium (see http://www.bioapi.org/DownloadsPage1.html)

European Commission
EUR 26193 – Joint Research Centre – Institute for the Protection and Security of the Citizen

Title: Fingerprint Recognition for Children

Author(s): Günter Schumacher

Luxembourg: Publications Office of the European Union

2013 – 98 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1018-5593 (print), ISSN 1831-9424 (online)

ISBN 978-92-79-33389-7 (print)
ISBN 978-92-79-33388-0 (CD-ROM)
ISBN 978-92-79-33390-3 (pdf)

doi:10.2788/3086

Abstract
This report summarises the findings of a JRC study dedicated to the question of whether or not automated fingerprint recognition for children is feasible, that is, if the recognition rates obtained with this technology for children are similar to those reached for adults. If necessary, the minimum age threshold for the reliable use of automatic fingerprint recognition should be revised.
The conclusions of the report are:

- Growth has limited influence on fingerprint recognition.
- Size (in terms of the dimensions of the relevant fingerprint characteristics) does not constitute any theoretical barrier for automated fingerprint recognition.
- Image quality (in terms of low contrast and distortion effects) is the ultimate problem for child fingerprints, and image quality is strongly influenced by size.
- Relevant quality metrics for fingerprints need revision with regard to the children case.
- Isotropic growth model may serve as a good approximation to cover changes over time.
- Alternative acquisition devices for fingerprints should be seriously considered in the future.

LB-NA-26193-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

ISBN 978-92-79-33390-3

9 789279 333903

Publications Office