Recognizing infants and toddlers over an on-production fingerprint database

Vanina Camacho [1] Guillermo Garella [2]Francesco Franzoni [3]Luis Di Martino [4]Guillermo Carbajal [5]Javier Preciozzi [6]Alicia Fernández [7]

Abstract: It is widely known that biometric systems based on adults fingerprints have reached an outstanding performance when compared against other biometric traits. This explains their extensive use by governmental agencies in charge of citizen identification. Nevertheless, the performance is highly degraded when fingerprints of newborns or toddlers are used. In this work, we analyze the performance of existing solutions (both at sensor and matching level) using 45000 infants fingerprints taken from an on-production civilian database. We also propose a solution by zooming the input fingerprints with an interpolation factor based on ridges distances. The developed solution shows improvements in both fingerprint quality (NFIQ 2.0) as well as recognition performance.

Keywords: fingerprint, biometric, recognition, newborns, infants, toddlers, id, interpolation

1 Introduction

Fingerprints are commonly preferred over other biometric traits for their inherent features, such as *distinctiveness*, *permanence*, and *performance* [RFJ08]. This explains its extensive use by national IDs/passports issuance offices and borders control among others. Fingerprint matching solutions are very mature and achieve a very good performance. Nevertheless, most available systems and research focus on adult fingerprints. In recent years, several works were conducted to analyze the suitability of using fingerprints in children [Co13, Ja15, Ja16a, Ja16b]. To the best of our knowledge, the most extensive study was presented by the *Joint Research Center* of the *European Commission (UE)* [Co13]. In this work, a database of fingerprints obtained from $2,611$ children (in the 0-12 years old range) with 500 dpi scanners were used. This data was acquired by the Portuguese government passport issuance offices. The report concluded that it was difficult to identify

[1] Facultad de Ingeniería, Instituto Ingeniería Eléctrica, Ave. Julio Herrera y Reissig 565, Montevideo, vcamacho@fing.edu.uy

[2] Facultad de Ingeniería, Instituto Ingeniería Eléctrica, Ave. Julio Herrera y Reissig 565, Montevideo, guillermo.garella@fing.edu.uy

[3] Facultad de Ingeniería, Instituto Ingeniería Eléctrica, Ave. Julio Herrera y Reissig 565, Montevideo, francesco.franzoni@fing.edu.uy

[4] Facultad de Ingeniería, Instituto Ingeniería Eléctrica, Ave. Julio Herrera y Reissig 565, Montevideo, dimartino@fing.edu.uy

[5] Facultad de Ingeniería, Instituto Ingeniería Eléctrica, Ave. Julio Herrera y Reissig 565, Montevideo, carbajal@fing.edu.uy

[6] Facultad de Ingeniería, Instituto Ingeniería Eléctrica/ Dirección Nacional de Identificación Civil, Ave. Julio Herrera y Reissig 565, Montevideo, jpreciozzi@dnic.gub.uy

[7] Facultad de Ingeniería, Instituto Ingeniería Eléctrica, Ave. Julio Herrera y Reissig 565, Montevideo, alicia@fing.edu.uy

children with less than six years old. They also concluded that it is necessary to use higher resolution scanners. Recently, a longitudinal study done in a population of 309 individuals ranging from zero to five years old was presented [Ja16b]. It was reported for the first time, the feasibility of using fingerprint to identify children at an early age: very good results were obtained in children older than six months using special scanners of 1270 dpi. Good results were also obtained using standard adult fingerprint scanners of 500 dpi in children of at least one year old.
In this work, we conduct an extensive analysis using a considerable size dataset obtained from an on-production environment. The data was obtained from the Uruguayan National Identification Agency (DNIC[3]) which is responsible for issuing Uruguayan passports and ID cards. The main contributions of the present work are twofold. First, we continue the analysis done in previous studies [Co13, Ja15, Ja16a, Ja16b] but with a greater number of individuals. Secondly, we show the robustness of fingerprints to identify children on an on-production civilian database, where fingerprints were acquired during the standard ID card and Passport issuance processes.

The rest of the article is organized as follows. In Section 2 we describe the scenario in which the data was obtained. Section 3 describes the protocol used to perform the analysis. It includes a preprocessing proposal with a zoom stage, with a learned interpolation factor, based on the estimation of the ridge distance for an age range. Section 4 describes the experiment realized and results obtained. Finally, section 5 describes the conclusions and future work.

2 Description of the Uruguayan scenario

Uruguay has a population of 3.42 millions people, with a born rate of 14.15% [4]. The agency in charge of national identity management is the National Civil Identification Agency (DNIC) and is responsible of issuing ID cards and Passports. The identification process is based on fingerprint comparison, which, as usual, is used in two different scenarios: enrollment and renewal. Uruguay is a particular country regarding identification since the enrollment is done at birth: by law, parents have 45 days to obtain the identity card for the newborn. This is done since its creation in 1978. Due to the difficulty of matching the fingerprints of newborns, this information is only stored but not used for identity verification or de-identification. When the child is 5 years old, a complete ten fingerprint is obtained and these fingerprints are stored for this individual throughout his life. This procedure is in execution since 1978 which means that all the Uruguayans born before 1973, are actually enrolled with their 5 year old fingers (nearly 2.4 millions). It is worth noting that during the enrollment and the renewal process, an adult fingerprint may be compared with a fingerprint corresponding to a 5 year old child and this is done systematically and as part of the actual process. Until 2011, all the fingerprints were obtained using ink on paper. As part of the enrollment process, these templates were scanned at 500 dpi, segmented on each finger, and stored digitally to further visualization. It was not until 2010

[3] in Spanish: Dirección Nacional de Identificación Civil
[4] https://en.wikipedia.org/wiki/Uruguay

that an AFIS system was installed on DNIC, which was filled with all these previously scanned fingerprint image. From 2011 until now, the ten fingerprints were acquired using fingerprint scanners. Given the fact that acquisition devices operate at one fixed resolution, there are certain disadvantages while working with children, mostly because of the small size fingerprint. Also, there are some problems with certain assumptions used by recognition algorithms with respect to average ridge distances. Another obstacle with smaller distances between fingerprint features is the decreased ability of algorithms to deal with the non-similarities introduced by distortion effects and bad positioning of the finger at capture.

3 Protocol Specification

Because of the national-wide characteristics of the biometric database, the number of individuals at our disposal is far more bigger than any other work done before.

3.1 Dataset

One of the criteria used to select the individuals and data for this study was obviously to have at least two fingerprints for the same finger. The total size of the dataset is 45000 pairs of fingerprints that were grouped by age as detailed on Table 1, where also the total number of individuals on each group is described.[5]. The variability on the number of individuals per group deserves an explanation. As explained before, it is mandatory for all children born in Uruguay to obtain an identity after 45 days of born. For this reason we have more than 13000 individuals on the group ranging from 0 to 1 month. In a typical scenario (if the child was registered in time and the document was not lost or stolen), the child has to renew the document at the age of 5. This is the reason we have more than 14000 individuals on the group between 5 and 6 years old. The other cases are exceptions on the typical identity management, and correspond to late enrollments or lost/stolen documents (it can also correspond to passport emission, which is also included in the same database). For these reasons, we have different number of fingerprints captures for each group. In this work, data is distributed according to the age at the time of the first capture. In table 1 we show the distribution of data and the number of fingerprints pairs for each set. All fingerprints were acquired by a well known commercial scanner model at 500 dpi, used by DNIC in all its offices (35 in total, distributed all along the country). Finally, and to compare some of the results with public databases, we also present some results using the adults NIST database MFCP2 [WF16].

3.2 Preprocessing method

One of the reasons why children identification is challenging is that most of the commercial systems are implemented and configured to work with adult fingerprints. This was

[5] Due to privacy regulations, DNIC data cannot be published.

already reported in [Ja16b] using NFIQ 2.0 as a quantitative measure of the quality of the fingerprints. A similar result was obtained in this work, as can be seen in Table 2. In order to use existing commercial AFIS systems, we need to preprocess the children fingerprints in such a way that the resulting image is well suited for these AFIS systems. This preprocessing process consists of two steps: an interpolation (to resize children fingerprints to an adult size) and segmentation (to reduce errors on minutiae extraction). Both steps are explained in the next sections.

Interpolation: In [Ja15, Ja16b], a fixed interpolation value of 1.8 was used in all cases, for fingerprints acquired at 500 dpi. In this work, we try to obtain a scale factor that depends on the age, and apply this obtained scale factor to resize the image to adult size. Even when we compare two children fingerprints, we rescale both of them to an adult size, enabling the use of existing AFIS systems. To determine the scale factor for each age, we follow [Kh11]: knowing the local ridge orientation, distance between ridges is measured by projecting the gray value levels along an orthogonal direction to the local ridge orientation and finding the minimum values. For each one of the ages, the median of the distances between ridges was selected, which was later compared with the distance between adult ridges on a 500 dpi image (9 pixels). The relation is given by Equation 1.

$$f_{oi} = \text{distance between ridges on adults}/\text{distance between ridges on age group}(dc_i) \quad (1)$$

Table 1 shows the result of the ridge distance analysis, divided by age group. As expected, median value augment as age increase. Table 1 also includes the final interpolation values for each age, which are the ones used to interpolate fingerprints in all the experiments done in this work.

Age group at first capture	Total of fingerprints	Average age at first capture	Average age at second capture	Interpolation Factor	Distance between ridges
Newborns	13050	18 d	8 m	1.52	5.92
1-2 m	5395	1m 12 d	8 m	1.63	5.52
2-3 m	902	2m 12 d	8 m	1.65	5.43
3-4 m	349	3m 13d	8 m	1.58	5.68
5-6 m	195	5m 13d	10 m	1.60	5.62
6-12 m	627	8m 27d	1 y 2 m	1.54	5.86
1-2 y	988	1y 6m	2 y 2 m	1.49	6.06
2-3 y	1164	2y 6m	3 y 7 m	1.47	6.14
5-6 y	14836	5y 6m	7 y 2 m	1.32	6.80
6-7 y	2234	6y 6m	7 y 6 m	1.29	6.96
7-8 y	1397	7y 6m	8 y 8 m	1.26	7.13
8-9 y	1471	8y 8m	9 y 10 m	1.24	7.27
9-10 y	2787	9y 9m	10 y 7m	1.22	7.40

Tab. 1: Description of database, interpolation factor and distance between ridges for each set.(d=days, m = months, y = years, newborns $<$ 1 m)

Segmentation: In order to eliminate the background acquisition noise on paper captures, fingerprints are segmented by looking for the first n aligned points whose values are black enough, assuming that they belong to the fingerprint, and cropping the image to this bounds. Values *n* and the threshold to assume that pixels are enough black were learned from FVC 2004 [Fi04].

4 Experiments and results

4.1 Performance Evaluation Metrics

In order to analyze the system performance, we use the the usual metrics: *True acceptance rate* (TAR) which is the percentage of times that the system correctly verifies a true claim of identity, and *False Acceptance Rate* (FAR), which is the probability that the system incorrectly matches the input pattern to a non-matching template in the database. *Receiver Operating Characteristics* (ROC) curve is plotted as TAR vs FAR at different thresholds (from 0 to 1) to indicate the verification performance.

4.2 Quality vs accuracy

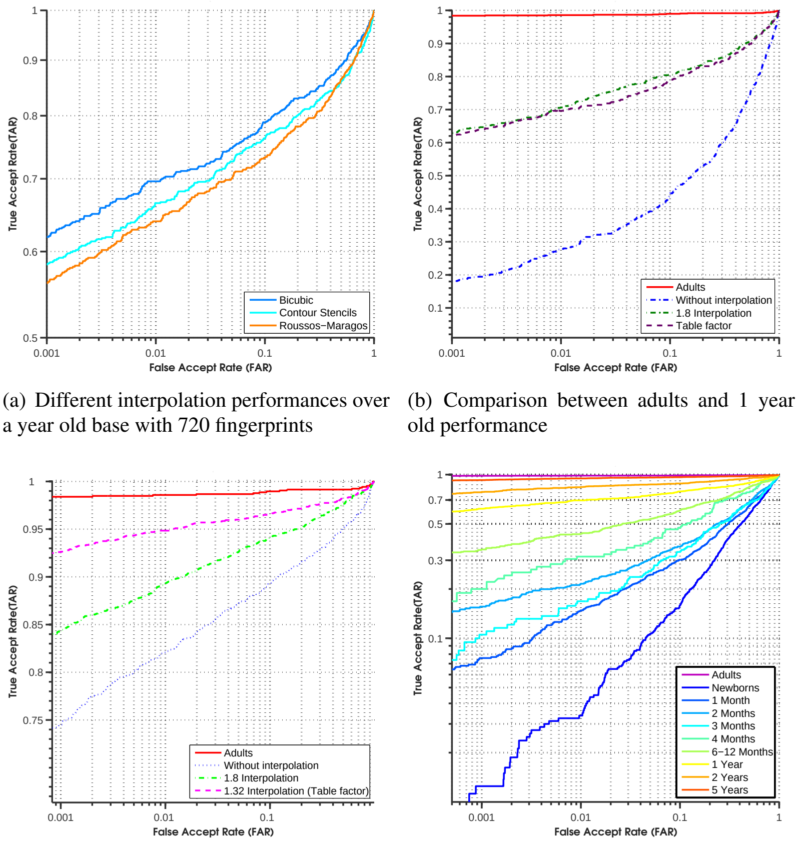
Performance of biometric systems depends to a great extent on the quality of data. Therefore, quality indicators can be used as a way to compare the effectiveness of different preprocessing methods. In this work we used NFIQ 2.0[De], which delivers a number from 0 to 100 directly related with the performance prediction of the matcher evaluating a single fingerprint. In table 2 NFIQ 2.0 data quality is shown.

Age Group	Numbers of fingerprints	Initial Quality	Preprocessing quality	Variance of preprocessing quality
Newborns	2264	1,68	2,62	4,37
1-2 m	2176	2,25	6,98	9,19
2-3 m	733	2,66	9,83	11,55
3-4 m	288	3,46	8,17	11,95
5-6 m	161	3,59	10,37	9,97
6-12 m	482	6,12	14,16	15,07
1-2 y	712	14,50	30,23	20,53
2-3 y	784	23,55	42,83	22,96
5-6 y	2963	36,03	48,56	25,48
Adults (NIST MFCP 2)	1086	45,98	-	-

Tab. 2: NFIQ 2.0 data quality

4.3 EXPERIMENTS

We start our set of experiments by analyzing interpolation. We compare a classic bi-cubic interpolation with two other interpolation methods: Interpolation with Geometric Contour Stencils [Ge11a] and Tensor-Driven Diffusion for Image Interpolation [Ge11b]. Figure 1(a) presents the results obtained on a one year old database with 720 fingerprints. In all cases, we use the interpolation factor described in Table 1. We can see that there is no significative difference between the different methods and in fact, bi-cubic obtains the better results. For the rest of the experiment, we use bi-cubic as the interpolation method. In the next experiment, we analyze the results on a database of one and five years old and compare them with the performance on an adult database (NIST MFCP2 database [WF16]). Table 3, Figures 1(b) and 1(c) presents the results. In this case, each pair of fingerprint is considered an identity (we present the results considering two fingerprints per identity later). It is clear from the results that interpolation is mandatory to obtain good results. What is more, applying the correct interpolation factor improves the results in the case of five years, as we can see when we compare the results obtained using the interpolation factor from Table 1 and the one obtained with an interpolation factor of 1.8. In the case of one year, we obtain almost the same performance for both interpolation factors. Since selected minutia extractor works with a default image size, using the proposed inteprolation factor we ensure looking for minutias over the whole fingerprint. In our final



(a) Different interpolation performances over a year old base with 720 fingerprints

(b) Comparison between adults and 1 year old performance

(c) Comparison between adults and 5 year old performance

(d) Performance of every set interpolating by table I factors

Fig. 1

	IF=1	IF proposed	IF=1,8	Fusion IF proposed	Fusion IF=1.8
TAR (%) five years	74,41	92,64	84,35	98.33	90.65
TAR (%) one year	18.12	61,88	62.34	79.28	81.42
TAR(%) Adults	98.39	-	-	-	-

Tab. 3: Performance given by TAR for a fixed FAR in 0.1% for interpolation factor of 1.8 and the proposed, with and without fusion for two fingers, (IF = interpolation factor)

experiment, we compare the results obtained using the corresponding interpolation factor obtained from Table 1 for different databases grouped by age. Figure 1(d) and Table 4 present the results. We also include the results obtained for an adult database. From the results, we can see that from 5-6 years old ($92,64\%$) are comparable to the ones obtained for adults ($98,39\%$) and even 1-2 years and 2-3 years present good results ($61,88\%$ and $78,37\%$). We recall that in all these experiments, we consider that each identity has only one fingerprint. In order to compare our results with the one obtained in [Ja16b], we perform a last experiment where we consider two fingerprints for each individual (right thumb and right index). In the five years old database, where we have 599 individuals, we obtain a TAR of 98.33% at a fixed FAR in 0.1%. From a total of 111 subjects in one year database, we obtain a TAR of 79.28% at a fixed FAR in 0.1%. In [Ja16b], authors reported a TAR of 100% for a fixed FAR in 0.1% for children from one to five years old. When we replicate the experiment with our dataset (applying 1.8 factor), we obtain a TAR 90.65% for the five years old and a TAR of 81.42% for one year old database, in both cases with a FAR of 0.1%. We believe that the main differences with the result reported in [Ja16b] is obviously the source of the dataset. In our case, the data was obtained directly from the on-production environment, without any participation on the way fingerprints were acquired. We consider that the results obtained from the fusion experiment (which is in fact the usual scenario on identification, where in general we have more than one fingerprint per individual) are very illustrative and confirms that fingerprints can be used to identify children starting from one year old. This claim is supported with the data used in this work, obtained directly from an on-production system.

Age Group	TAR (%)	TAR (%) with preprocessing
Newborns	NA	1,25
1-2 m	NA	7,57
2-3 m	NA	15,61
3-4 m	NA	10,53
5-6 m	NA	20,00
6-12 m	2,53	34,88
1-2 y	18,12	61,88
2-3 y	27,24	78,37
5-6 y (2000)	74,41	92,64
NIST MFCP 2	98,39	98,39

Tab. 4: Performance given by TAR for a fixed FAR in 0.1%

5 Conclusions and future work

In this work, we present an analysis of using fingerprints for children identification and verification. We perform all the study on a production database, where fingerprints were acquired on usual ID card and Passport. The results show that fingerprints can be used without any additional hardware starting from one year old. As we can see in Table 4, performance improves in accordance with children's growth. We also show that applying the corresponding interpolation factor, we obtain similar or better results than using a fixed interpolation size. We conclude that preprocessing fingerprint according to their age is a necessary step that deserves more research.

In future works, we plan to determinate the system performance using the interpolation factor corresponding to each fingerprint ridges distance more than to a range according to age. We also want to acquire fingerprints with a scanner with a higher resolution in order to analyze the feasibility of using fingerprints for children below one year old. Because we have access to the full fingerprint database at DNIC, we are planning to repeat the experiments with far more individuals including matching between children and adults.

6 Acknowledgment

The authors would like to thank the DNIC agency for its collaboration and for granting us the permission to access their valuable data. This work has been supported by an investigation grant provided by the ANII (Uruguay Agency of Investigation and Innovation).

References

[Co13] Commission, European: Fingerprint Recognition for Children. Technical report, Joint Research Centre, 2013. Available: https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/fingerprint-recognition-children.

[De] Development of NFIQ 2.0. Available: https://www.nist.gov/services-resources/software/development-nfiq-20.

[Fi04] Fingerprint Verification Competition (FVC), 2004. Available: http://bias.csr.unibo.it/fvc2004/.

[Ge11a] Getreuer, Pascal: Image Interpolation with Geometric Contour Stencils. Image Processing On Line, 1:98–116, 2011.

[Ge11b] Getreuer, Pascal: Roussos-Maragos Tensor-Driven Diffusion for Image Interpolation. Image Processing On Line, 1:178–186, 2011.

[Ja15] Jain, Anil K; Arora, Sunpreet S; Best-Rowden, Lacey; Cao, Kai; Sudhish, Prem Sewak; Bhatnagar, Anjoo: Biometrics for Child Vaccination and Welfare: Persistence of Fingerprint Recognition for Infants and Toddlers. arXiv preprint arXiv:1504.04651, 2015.

[Ja16a] Jain, Anil K.; Arora, Sunpreet S.; Best-Rowden, Lacey; Cao, Kai; Sudhish, Prem S.; Bhatnagar, Anjoo; Koda, Yoshinori: Giving Infants an Identity: Fingerprint Sensing and Recognition. In: Proceedings of the Eighth International Conference on Information and Communication Technologies and Development. ICTD '16, ACM, New York, NY, USA, pp. 29:1–29:4, 2016.

[Ja16b] Jain, Anil K; Arora, Sunpreet S; Cao, Kai; Best-Rowden, Lacey; Bhatnagar, Anjoo: Fingerprint Recognition of Young Children. IEEE Transactions on Information Forensics and Security, 2016.

[Kh11] Khan, Muhammad Ali: , Fingerprint image enhancement and minutiae extraction, 2011.

[RFJ08] Ross, Arun A.; Flynn, Patrick J.; Jain, Anil K.: Handbook of Biometrics. Springer, 2008.

[WF16] Watson, Craig; Flanagan, Patricia: NIST Special Database 14 Mated Fingerprint Card Pairs 2 WSQ Compressed Images. 2016.