IET Biometrics

Research Article

ISSN 2047-4938
Received on 19th December 2017
Accepted on 23rd February 2018
E-First on 18th April 2018
doi: 10.1049/iet-bmt.2017.0282
www.ietdl.org

Steps to solving the infant biometric problem with ridge-based biometrics

Johannes Kotzerke[1] ✉, Stephen A. Davis[1], Jodie McVernon[2], Kathy J. Horadam[1]

[1]*School of Science, RMIT University, Melbourne, Australia*

[2]*The Peter Doherty Institute for Infection and Immunity, The University of Melbourne and Royal Melbourne Hospital, Melbourne, Australia*

✉ *E-mail: johannes.kotzerke@rmit.edu.au*

Abstract: The pressing infant biometric problem is to find a biometric means to identify infants cheaply, reliably, and automatically. Physical traits of infants are tiny, delicate, and grow rapidly. The authors focus on a novel area of friction-ridge skin as a potential answer: the ball under the big toe. The ballprint is readily accessible, with more features and larger ridges than a fingerprint. The authors followed 54 newborns for 2 years, capturing their ballprints with an adult fingerprint scanner within 3 days of birth, at 2 months, at 6 months, and at 2 years. The authors show the growth of the ballprint is isotropic rather than affine during infancy. The isotropic growth rate from birth can be measured by the change in inter-ridge spacing, which the authors show precisely mirrors change in physical length from birth, as recorded by World Health Organisation for large, diverse infant populations. From 2 months of age, by using isotropic scaling to compensate for growth, the authors successfully matched good quality images with 0% equal error rate using existing adult fingerprint technology, even for captures 22 months apart. These findings flag the value of ballprints as a practical means of infant identification, by themselves, or together or sequentially with other biometrics.

1 Introduction

In 2013, UNICEF estimated that nearly 230 million children under the age of 5 worldwide are not formally registered at birth, and without identification, may be denied healthcare, education, and political and economic rights [1]. The ILO has estimated that each year, 1.2 million children are trafficked for slave labour or illegal adoption [2]. The pressing infant biometric problem [3, 4] is to find a biometric means to identify infants cheaply, reliably, and automatically. An important motivation is to be able to link an infant's physical trait with its personal information rather than rely on paper-based records which can be lost, damaged, or stolen. Proper identification may also prevent accidental swapping of babies in hospitals. In the case of an infant's vaccination history, it could serve a critical role in improving vaccination uptake and delivery in developing countries.

Biometrics are physical traits used to identify us or to verify our identity. Commonly used traits such as face [5], iris [6], fingerprint, or other areas of ridge skin such as palmprint have been shown to work reliably for adults [7]. Attempts to use them for infants (children aged between 0 and 24 months) have so far failed to give similarly acceptable levels of accuracy. The task is more challenging because of an infant's rapid growth, its inability or unwillingness to hold still to allow the trait to be captured reliably, and the tininess of the physical structures to be captured by a sensor. For newborns, the problems of uncooperative subjects and tiny, delicate physical structures are compounded by parental protectiveness and, for friction-ridge biometrics, by drying, cracking, and flaking skin in the days after birth.

Table 1 Alignment error

Visits	Rotation and translation only	Similarity transformation	Affine transformation
V2 and V3	5.45	2.03 (62.8%)	1.02 (81.3%)
V3 and V4	15.36	3.44 (77.6%)	1.29 (91.6%)
V2 and V4	20.67	3.25 (84.3%)	1.48 (92.8%)

Average median error in pixels between corresponding minutiae, and the percentage reduction when using similarity and affine transformations compared to rotation and translation alone.

The most frequently tested biometric options for infants have been: face; the friction-ridge based fingerprint, footprint, and palmprint; and more recently, ear shape. A survey of the literature to 2014 appears in [8, Table 1], and an update in [9]. For none of the favoured adult traits has verification of identity over the first years of life been categorically demonstrated in the literature. Longitudinal data sets have been available only for studies of footprint and fingerprint, with study numbers typically between 20 and 300 individuals. The longest longitudinal study of any infant biometric until 2015 has been 2 months for footprints [10]. The most recent account of child fingerprint recognition studies appears in [11], where 309 children in the age range 0–5 years were tracked over periods ranging from 2 months to 1 year with up to four sample periods. A purpose-built high-resolution scanner [12] was employed to acquire thumbprints (only) and encouraging results obtained for children over 6 months old. For children 0–6 months old, verification and search performance dropped dramatically, with the authors noting that capturing good quality prints in this age range remains a notable challenge. The number of newborns was ~20, of whom at most 5 were tracked for the full year [11, Table III].

Here, we report on the longest longitudinal infant biometric study of which we are aware, where 54 infants are tracked from birth to 2 years old with four sample periods: newborn and at 2, 6, and 24 months of age.

In 2011, the Bill & Melinda Gates Foundation called for biometric solutions for improved uptake and coverage of childhood vaccinations in developing countries, using low-cost cell phones [4]. Using exploratory funding from the Foundation and with a strict 1-year timeframe, we established the ‘Happy Feet’ project to study the infant footprint as a possible solution. The use of infants' footprints for verification is widely accepted by parents, hospital personnel, and law enforcement because they have been used as such for decades, for forensic purposes, or taken as a keepsake when a baby is born [13–15]. The authors and collaborators collected samples from 54 infants at newborn, 2 and 6 months old within the 1-year period 30 May 2012–22 May 2013. The samples are of each whole foot, including flexion creases and some ridge patterns, and of each ballprint [16] – the friction ridge skin of the hallucal region in the pad area under each big toe.

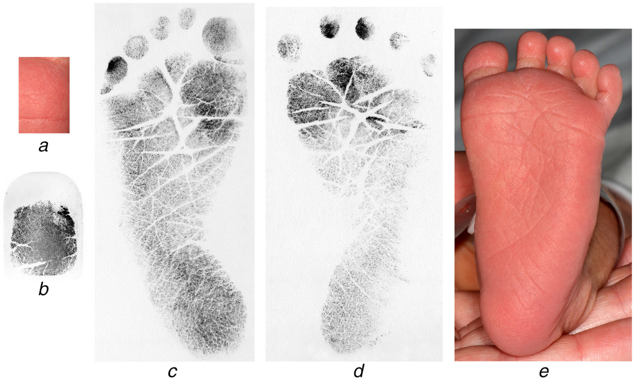

Fig. 1 *Infant 021's feet and left ball captured with different methods at 2 days of age. As prints capture impressions in contrast to photographs, the photographs have been mirrored along the vertical axis to match their corresponding prints visually*
(a) Ball (camera), ***(b)*** Ballprint (fingerprint scanner), ***(c)*** Left footprint (inkless paper), ***(d)*** Right footprint (inkless paper), ***(e)*** Right foot (camera)

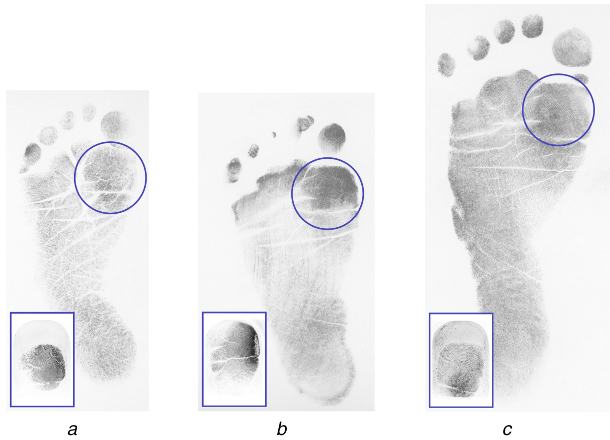

Fig. 2 *Left footprints of infant 021 at the age of 2 days, and 2 and 6 months (captured using inkless paper). The area of the ball is circled and the corresponding ballprint appears in the lower left-hand corner. The change in size due to growth is evident*
(a) V1: 2 days, ***(b)*** V2: 2 months, ***(c)*** V3: 6 months

To complete the Happy Feet database, ballprint samples were captured from the same 54 infants when they reached 2 years old.

1.1 Prior work

Results based on this database for the 0–6-month period appear in [8, 9, 17, 18]. Fig. 1 shows some of the samples from a newborn.

1.1.1 Foot crease pattern: For the prints of the whole foot (captured with an FBI-certified inkless paper fingerprint system), we asked first whether the flexion crease pattern might potentially be a good choice for an infant biometric, as it has been claimed to be unaffected by drying [13] and persistent over time [19]. In [17], we developed an algorithm for extracting crease patterns and performed a small manual experiment to test this claim. In [9], we obtained police forensic expert opinions on the comparison footprints in [17] and performed automatic crease extraction and matching experiments. We found no evidence to support the claim of persistence for flexion creases in these infants over the first 6 months of life.

Hence, we have concentrated on the ballprint. Compared to a fingerprint, the ballprint exhibits greater feature richness [16, 20, 21] leading to improved differentiability between individuals, and larger physical structures [16, 22], allowing the infant ballprint to be captured using off-the-shelf adult fingerprint sensors. The growth and expansion of the ballprint can be seen clearly in Fig. 2.

1.1.2 Ballprints of newborns: It is widely recognised that it is very difficult to capture friction-ridge skin prints from newborns which show minutiae reliably. Many captures are of such poor quality that they cannot be used for this purpose.

For ballprints of newborns (captured with scanner and camera), we asked if there was a 'window of opportunity' shortly after birth when the friction ridge skin pattern could reliably be captured. We found no evidence of such a period [9]. For these babies, the earliest age at which the skin dryness and flakiness settles enough for capture of prints of adequate quality (i.e. in which minutiae can be identified reliably) appears to depend on the individual.

In [18, 23, 24], an algorithm, EVA, for estimating the evidential value (EV) of an adult fingermark is developed. (See Section 2.4 for more details.) We wondered if the EV score $q \in [0, 1]$ for fingermarks could be used to assign an image quality value to newborn ballprints. Potentially, this would allow poor quality prints to be discarded at the time of capture and new ones to be acquired until enough of acceptable quality are collected. In [18], the EVA algorithm was tested on the newborn ballprints from 101 feet (prints were not available for all 108 feet because some of the newborns were too uncooperative). The fingermark classifier used was chosen experimentally to be *k*-NN, optimised at FMR5. It was shown that an EV score could be assigned to all these newborn ballprint images (mean 0.48, standard deviation 0.12). In closed-set identification, rank 1 retrieval improved from 13.52 to 17.27% when the images scoring below the median were removed.

1.1.3 Ballprints of 2 and 6-month olds: In [8], we investigated the growth of the ballprint as the 54 infants aged from 2 to 6 months. Pairs of prints, at most one pair per foot, were chosen manually to have sufficient image quality and reasonable overlap to ensure minutiae matching was possible. Of the 108 feet, 65 had such pairs. A researcher marked up corresponding minutiae in these 65 pairs; the number of corresponding minutiae ranged from 5 to 52 with median 17. These minutiae were used for growth analysis and to optimise the capture resolution parameter entered for automatic extraction using an off-the-shelf commercial adult fingerprint matcher (Neurotechnology Verifinger 6.7). Though growth is a complex process, we found that over this 4-month period, it can be reasonably well estimated by assuming it is isotropic. Verification test sets contain the prints of those 65 pairs and additional ones similar in overlap and quality to them, resulting in 94 and 98 prints at 2 and 6 months, respectively. In total, there were 44 intra 2-month genuine comparisons, 52 intra 6-month genuine comparisons, and 131 genuine comparisons inter 2 and 6 months. The same number of imposter comparisons was randomly sampled and was consistent for each experiment. After we reduce the capture resolution set at template creation in Verifinger, it performs an internal scaling of images with an unspecified scaling factor. Matching with Verifinger gave equal error rates (EERs) for intra-age verification of 0% at 2 months and at 6 months, and of 7.8% for comparisons between infants aged 2 and 6 months. For more details, see [8].

1.2 Contribution of this paper

Here, we work with the entire Happy Feet database, building on the work in Section 1.1. We revisit the modelling of growth as the infants age from 2 to 24 months, and conclude that the first-order growth of ballprints is isotropic rather than affine over this period.

Under the assumption of isotropic growth, we investigate the inter-ridge spacing (IRS) as a simple linear quantity that can be measured on a ballprint in order to determine a scaling factor for isotropic growth as a function of age. Even for our newborns, the IRS can be measured automatically from their ballprints, whereas minutiae may not be reliably identifiable. We demonstrate that, over the whole timespan from birth to 2 years old, the IRS for our 54 infants is closely correlated by age with the World Health Organisation (WHO) length charts derived from large international population studies. Thus, we base isotropic scaling on the corresponding ratios of IRSs.

Then we test ballprints at 2, 6, and 24 months for verification within and across age groups, both without and with isotropic

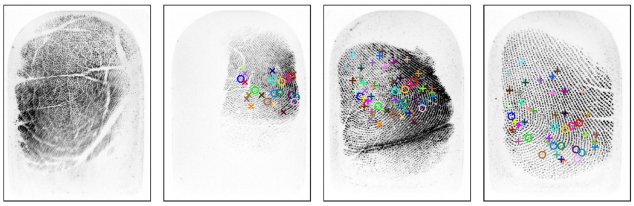

Fig. 3 *Comparison of four left foot ballprints from infant* 019 *captured shortly after birth, at 2, 6 months, and 2 years and their manually marked minutiae; corresponding ones share the same colour. The different markers indicate for which visits the correspondence could be established: circle:* $\{V2, V3, V4\}$, *multiplication:* $\{V2, V3\}$, *plus:* $\{V3, V4\}$. *The correspondence between V2 and V4 has been induced via the correspondences of* $\{V2, V3\}$ *and* $\{V3, V4\}$. *The V1 print is not marked up because, owing to its poor quality, we were unable to detect minutiae reliably*

scaling. A subset of ballprints with sufficient overlap from age to age is used. We achieve complete separation within each of the three age groups, but if no scaling is used, the verification performance across age groups drops dramatically. If scaling is used, competitive EERs across each of the three age differences are obtained.

Finally, if the lowest quality images are removed from the test set (bottom 33.3% of images), complete separation across each of the three age differences is achieved.

The remainder of the paper is organised as follows. More details of the Happy Feet database and our methodology are given in Section 2. Description of our experiments and results appears in Section 3 and the conclusion and discussion are in Section 4.

2 Database and methodology

Here, we describe in more detail than Section 1.1 the database and methods we used.

2.1 Database

The Happy Feet database is the only longitudinal footprint and ballprint database from newborn through the 2 years of infancy, to the best of our knowledge, in the world.

All data were collected by the same nurse. Data collection was carried out over 2.75 years, with data being collected at four ages: at ~2 days, 2, 6 months, and 2 years. These times were selected because they are regular vaccination ages. Visit 1 (V1), from 30 May 2012 to 6 December 2012, was to the newborns (ranging from the age of 11–683 h with an average age of 67 h). Visit 2 (V2), from 12 July 2012 to 13 February 2013, was to the young infants (ranging from the age of 31–84 days with an average age of 55 days). Visit 3 (V3), from 7 December 2012 to 22 May 2013, was to the infants (ranging from the age of 161–211 days with an average age of 183 days). Visit 4 (V4), from 11 August 2014 to 25 February 2015, was to the older infants (ranging from the age of 702–894 days with an average age of 780 days). All images were captured in one session per child per visit. The first three visits were held in the private hospital rooms of the mothers' obstetricians and visit 4 involved home visits.

The database contains captures of both balls (right and left) for each of the 54 participants. The total number of ballprint images per visit are 589 (V1), 625 (V2), 628 (V3), and 1147 (V4), totalling 2989 images.

We used a commercial off-the-shelf single adult fingerprint scanner, the NEC PU900-10, to capture typically 6 (V1, V2, V3) or 12 (V4) ballprint impressions per visit and individual foot. The increased number of captures at visit 4 is to compensate for the infant's growth, to maximise potential overlap of the prints between visits.

There are three main reasons to choose a commercial off-the-shelf single fingerprint scanner. Firstly, its small, round design and soft capture area minimises the infant's risk of injuries. It also increases both the parents' and their infant's acceptance of the scanner. Secondly, its high native sensor resolution (1000 ppi), even though the output is downsampled in hardware to 500 ppi, maximises the captured details. The print quality ranges from very poor to very good due to the well-known problems of uncooperative subjects, dry and flaking skin in newborns, and tiny, malleable ridges. Thirdly, the use of commercial 'standard' technology is cost-effective and it has already proven its suitability for the purpose for which it is marketed. This is not necessarily the case for specialised hardware which has to be developed, especially for a study.

In order to capture an infant's ballprint, we found that it was best to sit the infant on its parent's or the researcher's lap and to press the fingerprint scanner against the ball area. This way, it is possible to control the infant more easily regardless of its mood and encourage it to cooperate. The infant's feet were cleaned with baby wipes before each session.

All data were collected in accordance with the ethics protocols of RMIT University, The University of Melbourne, and St Vincent's Private Hospital Research Committee.

2.2 Growth model

The simplest first-order model for expansion of the ballprint, and hence adjustment for physical growth, is isotropic scaling, in which compared images scale by the same factor in any direction. This corresponds to a similarity transformation where rotation and translation are used to align two prints. An affine transformation introduces the possibility of different scaling factors for different directions. We asked whether isotropic or affine growth is the better first-order model. To test this, we identified minutiae in good quality prints and looked for correspondence (see Fig. 3) across V2, V3, and V4. We excluded prints from V1 due to the lack of visible minutiae and we established correspondence between V2 and V4 via V3. The marked-up ballprints from V2 and V3 were those described in Section 1.1.3; there were 65 feet, but for a few of these, suitable V4 correspondences could not be found. In all, minutiae matches were manually marked up by the same researcher for 63 different feet from 37 infants across the three visits, giving 63 sets of three ballprints.

2.3 IRS algorithm

The IRS is a global level characteristic of ridge skin that varies across any area of ridge skin and across the body (e.g. the adult ballprint IRS is ~25% larger than adult fingerprint IRS). It depends on gender, ethnicity, and age [25, 26], but it is distinctive enough to allow broad classification of individuals from the same population [27]. We employed the radially averaged power spectrum (RLAPS) algorithm of [23] to measure the IRS of the Happy Feet ballprints. The algorithm uses the RLAPS for a given angle and it takes into account that the IRS varies depending on where and how it is determined. The IRS decreases around singular points [20], so it is often measured in areas where none are present [25]. Distortions lead to an uneven and non-circular energy distribution at singular points. RLAPS uses a radial limitation around the strongest frequency peak to increase robustness against this. Additionally, it chooses the spectral peak corresponding to the highest frequency but with an amplitude close to the maximum to take this into account. Details appear in [24].

2.4 Image quality algorithm

The EVA EV algorithm [23] is based on image features extracted from captures of adult fingermarks and a ground truth. A pseudo fingermark database of 1428 marks from 4 adult individuals was captured three ways: by camera, mobile phone, and flatbed scanner. Ground truth is the majority vote of three police fingerprint experts who decided for each mark, if it is of sufficient EV. They made a binary decision as to whether the mark was of value for identification or not; this is a decision made at the analysis stage of the Analysis–Comparison–Evaluation–Verification (ACE-V) framework for fingerprint assessment. Best performance by EVA was achieved for feature sets which fused the NFIQ2 feature set, and the number of minutiae and quality value determined by the commercial extractor Verifinger. For

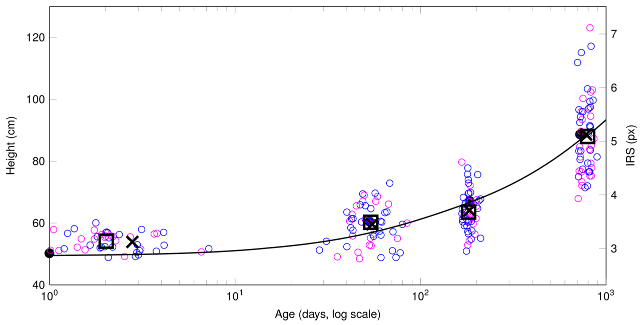

Fig. 4 *WHO growth chart presents physical length (and from day 731 onward height) versus the infant's age on a log scale (black curve). We overlay it with the median IRS per infant, the median (square) and mean (multiplication) per visit are indicated; the infant's gender is colour-coded (girls: pink, boys: blue). The measured IRS for the sample ballprints shown in Fig. 3 is highlighted by a black dot*

classification using EVA, a classifier is trained on the scanner images and its parameters are chosen via the lowest error at a fixed false match rate (FMR) for the camera and phone images. The classifiers employed were support vector machine, discriminant analysis, and k-nearest neighbours (k-NN). EVA can be modified to output an EV score $q \in [0, 1]$ instead of a binary decision. For more details, see [18, 23].

We propose that the EV output from EVA can be considered a basic image quality estimate for friction-ridge images. This is based on the assumption that sufficient EV and image quality are correlated. This hypothesis is reasonable because in [23], it is shown that for fingermarks, the EV can be derived from a set of image features. Therefore, the contrary argument should hold true as well and the EV provides a basic image quality estimate. This proposal is supported by results of testing on several challenging databases in [18].

We now work on the hypothesis that this evidential quality algorithm, trained on ground truth for adult fingermarks, can be successfully applied to infant ballprints for identification purposes, since both are friction-ridge skin images having highly variable quality.

3 Experiments and results

As is the experience for fingerprint [11, 12], acquisition of ridge skin pattern for infants aged less than a few weeks is particularly difficult, with the skin of a newborn drying, cracking, and flaking on exposure to air. Extraction of local-level features (minutiae) from V1 images of ballprints was only possible in a small number of cases. However, the global level IRS could be measured for each newborn.

3.1 First-order growth modelling

We aligned the 63 pairs of marked-up ballprints (see Section 2.2) by: (i) finding a rotation and translation that minimised the squared error between corresponding minutiae, (ii) estimating a similarity transformation, and (iii) estimating an affine transformation. For the latter two, a random sample consensus (RANSAC) approach was used to achieve alignment. For each of the three possible pairings (V2 and V3, V3 and V4, and V2 and V4) and each of the three approaches, the median value of the errors between corresponding minutiae was averaged over the 63 sets of ballprints and is shown in Table 1. A sharp drop in the magnitude of the median error is observed when using a similarity transformation, followed by a slighter further drop when using an affine transformation.

As was previously shown in the growth error analysis for growth from V2 to V3 [8], there is similar correlation in the spatial error for growth from V2 to V4 and from V3 to V4. Again, statistical measures (Moran's I and Geary's C) indicate that the correlation is local rather than global, which could be due to the skin being stretched at the edges of the scanner.

While we observe improvement in average median error when using an affine transformation, the relative performance of the similarity transformation does not get worse for larger age gaps between captures; in fact, the reverse is true (84% reduction in error when there is a 22-month age gap versus 78% reduction in error when the age gap is 18 months, see Table 1). If the observed error differences were due to affine growth then the effect should strengthen over longer time periods as the balls of the infants' feet continue to grow at different rates in different directions. We conclude from these results that the first-order growth of ballprints is isotropic rather than affine over the first 2 years of life.

3.2 Scaling factor determination

Assuming isotropic growth, we next used the IRS as a simple linear measurement taken from the ballprint images to determine an isotropic scaling factor as a function of age (cf. Fig. 3). The IRS was measured automatically for all the ballprints (including newborns') in the database using the RLAPS algorithm of [23] (see Section 2.3). That is, the IRS was measured for all images in the database, regardless of quality, for each individual at each visit. The median IRS per infant and visit is shown in Fig. 4. In Fig. 4, we overlay our IRS samples at the different ages on the WHO growth chart of median height (length). The WHO data show the relationship between a child's median length from shortly after birth up to the age of 5 years in steps of single days. (We restrict to the first 1000 days here.) The median is derived from a large sample study involving several countries and ethnic groups [28] and ~8500 children. We observe that the median and mean IRS fit tightly to the WHO curve and that the individual IRS samples follow the curve well. The correlation coefficient of the median IRS at each of the visits and the length obtained from the WHO curve at the median age per visit is 0.9874. The correlation coefficient of the mean IRS and the lengths at the average age is 0.9855.

We conclude from these results that the first-order growth of the infant ballprint can be determined directly from the IRS over the whole 24 months of infancy. Furthermore, we infer that the median and mean IRS at non-capture ages track the WHO length curve of Fig. 4 and may therefore be estimated from 4.

3.3 Intra- and inter-age group verification performance

In [8], we showed that unknown internal scaling based on the capture resolution we provided to a commercial adult fingerprint matcher gave EERs for intra-age verification of 0% at 2 months, 0% at 6 months, and of 7.8% for comparisons between infants aged 2 and 6 months, using prints with sufficient overlap.

To test whether isotropic scaling could improve on these results and succeed in verification over longer timespans, we again limited our database to prints with clearly visible ridge pattern and sufficient overlap between the different captures, selecting 94 from V2, 98 from V3, and 178 from V4. The V2 and V3 prints selected are the same as in [8]. Instead of random selection of equal numbers of imposter comparisons as genuine comparisons, as in the earlier experiment, here we used all imposter comparisons available (cf. Table 2). Our scaling factor for each pair of visits is the ratio of the median IRS per visit (the values at squares in Fig. 4).

The use of an off-the-shelf commercial matcher (Neurotechnology Verifinger 7.0) developed for adult fingerprints

Table 2 Verification comparison numbers

	V2	V3	V4
V2	44/(94 × 93)	131/(94 × 98)	246/(94 × 178)
V3	—	52/(98 × 97)	245/(98 × 178)
V4	—	—	474/(178 × 177)

Fraction of the number of genuine comparisons on the total number of comparisons across different visits. Self-comparisons for intra-visit experiments are excluded, but symmetric comparisons are considered to be different.

Table 3 Ballprint verification EER scores

Method	V2V2, %	V3V3, %	V4V4, %	V2V3, %	V3V4, %	V2V4, %
no scaling	4.55	3.85	0.00	24.66	40.82	49.98
scaling	0.00	0.00	0.00	2.29	1.29	2.30

Ballprint EER scores achieved for all age gaps with and without adjustment for growth.

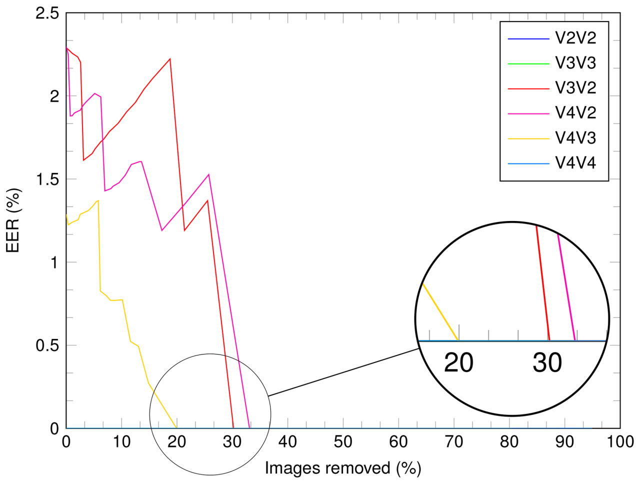

Fig. 5 *This diagram illustrates how the accuracy, measured by EER, improves when low-quality images are removed from the test set. The verification is performed error-free when the lowest-quality third of the images is removed. The image quality of the infant ballprints is derived from a quality algorithm for adult fingermarks*

delivers competitive results if the ballprints are isotropically scaled (see Table 3). In particular, the EER for inter-visit verification from 2 to 6 months drops from 7.8 to 2.29% when scaling using IRS is used. If the ballprints are not scaled, it is evident that as soon as the ballprints change in size due to the infant's growth, the matching accuracy, indicated by EER, decreases dramatically. Matching accuracy is little better than chance over the 22- and 18-month intervals.

3.4 Verification performance with EV scoring

Finally, we asked to what extent, image quality affects verification matching performance. For the ballprint images in the test set, we applied the EVA algorithm (Section 2.4) with a k-NN classifier with parameters optimised at FMR10. Note that the algorithm was not trained or optimised on the ballprint images themselves but on fingerprint expert decisions on the EV of adult fingermarks. After scaling the ballprint images to 500 ppi, the same image features as for the fingermarks are extracted from the ballprint images (NFIQ2 features, Verifinger minutiae number, Verifinger quality score). The classifier is applied to these and outputs for each ballprint image an EV q in the range [0, 1]. We interpret q as the image's quality, or, the amount of information that increases distinctiveness. We rank the ballprints by q.

During the quality-thresholded verification experiments, all comparisons involving a print with q lower than a threshold are removed and the EER is re-computed.

Eventually, the test set could be restricted to such sufficiently high-quality captures that complete separation was achieved across all different age comparisons. Error-free verification was achieved when the lowest-quality third (33.3%) of the images contained in the test set were removed. Fig. 5 illustrates how the EER decreases as the percentage of images removed increases. In larger populations, we suggest that pre-filtering of images by EV score will deliver competitive verification performance.

4 Conclusion and discussion

In conclusion, we have made four discoveries about the infant ballprint with this database. Firstly, we found that a linear measure on the ballprint, the IRS, measurable from birth, is highly correlated with the physical body length of an infant, which is a well-established and accepted measure for growth. Secondly, the physical changes a ballprint undergoes while the infant grows can be compensated for by isotropic scaling. The median IRS of a population is a sufficient measure to compensate for these changes directly and efficiently. Thirdly, when the growth is compensated for, low EERs are achieved, and when low-quality images are rejected or excluded, significantly lower EERs are possible. Finally, the friction ridge skin on the infant ballprint is sufficiently similar to that on the adult fingerprint that mature adult fingerprint technology can be used directly for image capture and, after scaling, for feature extraction and matching. This provides a practical and cost-effective technical solution to the infant biometric problem.

These discoveries are based on a small database. Nonetheless, it is comparable in size with many infant biometric studies reported in the literature, and it is unique as a longitudinal study, covering the whole of infancy for the same infant cohort for 2 years. It will be made publically available. Fig. 4 supports our view that the Happy Feet database is representative of much larger infant populations. Our results demonstrate that ballprints may be acquired as early as at 2 months of age for verification of identity later in life. We suggest this will also be true for even younger infants, perhaps as young as 2 weeks old. We suggest our results will scale up to larger populations and also may translate to other ridge-based biometrics proposed for infant identification.

The newborn period of the first 2 weeks or so remains the most challenging for the capture of good quality friction ridge skin prints. If capture is combined with a quality assessment algorithm such as we use, to filter low-quality prints out until sufficiently good ones are obtained at the first enrolment period, this problem may be avoided.

We conclude that the large ridge structure, predictable growth, and feature richness of the ballprint may make ballprints a practical means of infant identification, by themselves, or together or sequentially with other biometrics.

5 Acknowledgments

We thank Paula Nathan for her hard work and persistence with data collection, and we thank the families of the infants who participated in the study. The data collection for infants aged 0–6 months was funded by the Bill & Melinda Gates Foundation (OPP1046031); the research was funded by the Australian Research Council (DP120101188).

6 References

[1] United Nations Children's Emergency Fund.: ‘Every child's birth right: inequities and trends in birth registration’, 2013. Available from: http://www.unicef.org/publications/index_71514.html

[2] International Labour Office: ‘*Every child counts: new global estimates on child labour*’ (International Labour Organization, Geneva, Switzerland, 2002)

[3] United Nations Children's Emergency Fund.: ‘Expanding immunization coverage’, 2011. Available from: http://www.unicef.org/immunization/index_coverage.html

[4] Bill & Melinda Gates Foundation.: ‘Create low-cost cell phone-based solutions for improved uptake and coverage of childhood vaccinations’, 2011. Available from: http://www.grandchallenges.org/MeasureHealthStatus/Topics/CellPhoneApps/Pages/Round7.aspx

[5] Best-Rowden, L., Jain, A.K.: ‘Longitudinal study of automatic face recognition’, *IEEE Trans. Pattern Anal. Mach. Intell.*, 2018, **40**, (1), pp. 148–162

[6] Grother, P., Matey, J.R., Tabassi, E., *et al.*: 'IREX VI – ITL biometrics overview'. NIST Interagency Rep. No. 7948, National Institute of Standards and Technology, Gaithersburg, Maryland, 2013

[7] Yoon, S., Jain, A.K.: 'Longitudinal study of fingerprint recognition', *Proc. Natl. Acad. Sci. USA*, 2015, **112**, (28), pp. 8555–8560, Available from: http://www.pnas.org/content/112/28/8555.abstract

[8] Kotzerke, J., Arakala, A., Davis, S., *et al.*: 'Ballprints as an infant biometric: a first approach'. 2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proc., 2014, pp. 36–43

[9] Kotzerke, J., Davis, S.A., Hayes, R., *et al.*: 'Newborn and infant discrimination: revisiting footprints', *Aust. J. Forensic Sci.*, 2017, pp. 1–14, Available from: https://doi.org/10.1080/00450618.2017.1324582

[10] Thompson, J.E., Clark, D.A., Salisbury, B., *et al.*: 'Footprinting the newborn infant: not cost effective', *J. Pediatr.*, 1981, **99**, (5), pp. 797–798

[11] Jain, A.K., Arora, S.S., Cao, K., *et al.*: 'Fingerprint recognition of young children', *IEEE Trans. Inf. Forensics Sec.*, 2017, **12**, (7), pp. 1501–1514

[12] Jain, A.K., Arora, S.S., Best Rowden, L., *et al.*: 'Giving infants an identity: fingerprint sensing and recognition'. Proc. of the Eighth Int. Conf. on Information and Communication Technologies and Development. ICTD '16. ACM, 2016, pp. 29:1–29:4. Available from: http://doi.acm.org/10.1145/2909609.2909612

[13] Blake, J.W.: 'Identification of the new-born by flexure creases', *Ident News*, 1959, **9**, (9), pp. 3–5

[14] Stapleton, M.E.: 'Best foot forward: infant footprints for personal identification', *FBI Law Enforcement Bull.*, 1994, **63**, (11), pp. 14–17

[15] Sinclair, R., Fox, C.: 'Infant-to-adult footprint identification', *J. Forensic Identif.*, 2007, **57**, (4), pp. 485–492

[16] Uhl, A., Wild, P.: 'Personal identification using eigenfeet, ballprint and foot geometry biometrics'. 2007. First IEEE Int. Conf. on Biometrics: Theory, Applications, and Systems, 2007. BTAS. IEEE, 2007, pp. 1–6

[17] Kotzerke, J., Davis, S., Horadam, K., *et al.*: 'Newborn and infant footprint crease pattern extraction'. Proc. of 2013 IEEE 20th Int. Conf. on Image Processing (ICIP 2013). IEEE, 2013, pp. 4181–4185

[18] Kotzerke, J., Hao, H., Davis, S.A., *et al.*: 'Identification performance of evidential value estimation for ridge-based biometrics', *EURASIP J. Inf. Sec.*, 2016, **2016**, (1), p. 24

[19] Massey, S.: 'Persistence of creases of the foot and their value for forensic identification purposes', *J. Forensic Identif.*, 2004, **54**, (3), pp. 296–315

[20] Maltoni, D., Maio, D., Jain, A.K., *et al.*: '*Handbook of fingerprint recognition*' (Springer Publishing Company, Incorporated, New York, USA, 2009, 2nd edn.)

[21] Okajima, M.: 'Epidermal-ridge minutia in the hallucal area', *Mitteilungen der Anthropogischen Gesellschaft in in Wien*, 1977, **107**, pp. 135–139

[22] Jain, A.K., Cao, K., Arora, S.S.: 'Recognizing infants and toddlers using fingerprints: increasing the vaccination coverage'. 2014 IEEE Int. Joint Conf. on. Biometrics (IJCB), IEEE, 2014, pp. 1–8

[23] Kotzerke, J., Davis, S.A., Hayes, R., *et al.*: 'Discriminating fingermarks with evidential value for forensic comparison'. 2015 Int. Workshop on Biometrics and Forensics (IWBF), IEEE, 2015, pp. 1–6

[24] Kotzerke, J., Davis, S., Hayes, R., *et al.*: 'Identification performance of evidential value estimation for fingermarks'. 2015 Int. Conf. of the Biometrics Special Interest Group (BIOSIG), IEEE, 2015, pp. 1–6

[25] Acree, M.A.: 'Is there a gender difference in fingerprint ridge density?', *Forensic Sci. Int.*, 1999, **102**, (1), pp. 35–44

[26] Gutièrrez Redomero, E., Alonso, C., Romero, E., *et al.*: 'Variability of fingerprint ridge density in a sample of spanish caucasians and its application to sex determination', *Forensic Sci. Int.*, 2008, **180**, (1), pp. 17–22

[27] Mundorff, A.Z., Bartelink, E.J., Murad, T.A.: 'Sexual dimorphism in finger ridge breadth measurements: a tool for sex estimation from fingerprints', *J. Forensic Sci.*, 2014, **59**, (4), pp. 891–897, Available from: http://dx.doi.org/10.1111/1556--4029.12449

[28] World Health Organization.: 'Child growth standards', 2006. Available from: http://www.who.int/childgrowth/standards/en/