Infant Footprint Recognition

Eryun Liu
ISEE, Zhejiang University
38 Zheda Rd., Hangzhou 310017, China
eryunliu@zju.edu.cn

Abstract

Infant recognition has received increasing attention in recent years in many applications, such as tracking child vaccination and identifying missing children. Due to the lack of efficient identification methods for infants and newborns, the current methods of infant recognition rely on identification of parents or certificates of identity. While biometric recognition technologies (e.g., face and fingerprint recognition) have been widely deployed in many applications for recognizing adults and teenagers, no such recognition systems yet exist for infants or newborns. One of the major problems is that the biometric traits of infants and newborns are either not permanent (e.g., face) or difficult to capture (e.g., fingerprint) due to lack of appropriate sensors. In this paper, we investigate the feasibility of infant recognition by their footprint using a 500 ppi commodity friction ridge sensor. We collected an infant footprint dataset in three sessions, consisting of 60 subjects, with age range from 1 to 9 months. We proposed a new minutiae descriptor based on deep convolutional neural network for measuring minutiae similarity. The descriptor is compact and highly discriminative. We conducted verification experiments for both single enrolled template and fusion of multiple enrolled templates, and show the impact of age and time gap on matching performance. Comparison experiments with state of the art algorithm show the advantage of the proposed minutiae descriptor.

1. Introduction

The Global Vaccine Action Plan (GVAP\(^1\)) is a roadmap to prevent millions of deaths through more equitable access to vaccines. Countries are aiming to achieve vaccination coverage of over 90% nationally and over 80% world wide by 2020. According to statistics from the World Health Organization (WHO), global vaccination coverage has remained steady for the past few years. For example, during 2014, about 115 million (86%) infants worldwide received three doses of diphtheria-tetanus-pertussis (DTP3) vaccine, protecting them against infectious diseases that can cause serious illness and disability or be fatal. By 2014, 129 countries out of a total of 193 had reached at least 90% coverage of the DTP3 vaccine. Despite improvements in global vaccine coverage during the past decade, WHO also points out that regional and local disparities continue to exist due to 1) limited resources, 2) competing health priorities, 3) poor management of health systems, and 4) inadequate monitoring and supervision\(^2\). Better management of health systems requires a more accurate and efficient child recognition system. It is estimated that vaccine wastage rates are higher than 50% in some of the most challenging geographies\(^3\). Vaccinations are not administered to the children in need due to the lack of an effective method to keep track of which children have been vaccinated and which vaccines have been administered to each child. There is an urgent demand in immunization management system to track vaccination schedule of each child reliably and efficiently. Besides tracking child vaccination, infant recognition is also urgently desirable in many other applications such as identifying missing children, preventing baby swaps\(^4\) and other child welfare applications.

Biometric recognition has been actively researched for many decades with the aim to solve the problem of person identification. Many of the techniques have been successfully deployed in a wide range of civil and forensic applications. However, almost all of the techniques are designed for adults or teenagers, not for infants. These techniques may not be useful for infants directly since some of the biometric modalities, e.g. face, are not stable shortly after birth for the purpose of automatic recognition. DNA is supposed to be the most accurate means of child recognition (except

\(^1\)http://www.who.int/immunization/global_vaccine_action_plan/GVAP_doc_2011_2020/en/

\(^2\)http://www.who.int/mediacentre/factsheets/fs378/en/

\(^3\)http://vaxtrac.com

in the case of identical twins). However, the current DNA based identification method is not suitable for real-time application and may also be viewed as an invasive procedure due to the inconvenience in sample capture. Further, the use of DNA may raise some privacy concerns.

Skin friction ridge patterns appearing on the surface of our hands and feet are a special kind of biometric modality which is presumably stable even before birth [19]. Prior work on infant recognition has investigated some of the popular biometric traits, e.g., fingerprint and face [8], [9], [7]. In this paper, we study the footprint as a means of infant recognition. Footprints have several advantages over other biometric traits for identifying newborns. For example, in some countries, taking newborns’ footprint images is a routine procedure for the certificate of birth before infant is released from the hospital. Thus footprints have a lower degree of parental concerns than fingerprints. Additionally, newborns often keep their fists closed, making it difficult to capture fingerprint images.

Footprint recognition has long been studied in forensic applications for solving crime cases [21]. However, a fully automatic footprint recognition system only appeared in recent years [13]. However, it has not yet been extensively studied for newborn and infant recognition.

Table 1 briefly reviews the literatures on infant footprint recognition according to the type of features used for matching. There are two types of footprint images which have been studied in the literature. The first one is low resolution (around 100 ppi) footprint images where the friction ridge information could not be captured. The main features used are creases which can be captured at this low 100 ppi resolution. The second type is high resolution (≥ 500 ppi) footprint images where the friction ridge information, i.e., minutiae, is captured.

Several algorithms to match low resolution footprint images of newborns were proposed in [13], [12] and [11]. They collected footprint images from 101 feet within 2 days in one session, after birth using Cannon Powershot SX110 IS camera. The best Equal Error Rate (EER) was reported to be 1.34%. In [1], the authors collected 240 footprint images from 40 newborns within 2 days after birth using a Canon EOS 7D camera. Among these 40 newborns, 22 subjects had a special background setup and 18 subjects had normal background condition during footprint capture. Authors achieved a 65% rank-1 identification accuracy with only 40 images in the gallery. In general, the low resolution footprint image does not contain ridge friction information which can be observed at a finer scale. The main features that can be extracted in these images are creases which may not be stable yet.

In [16], Kotzerke et al. proposed an algorithm for creases feature extraction from newborn footprints, but matching performance was not reported. Footprints have much larger area than fingerprints and palmprints. As such footprints have a larger number of minutiae than palmprints and fingerprints. For recognition purposes, such a large amount of friction ridge area or minutiae is actually not necessary. In [15], Kotzerke et al. proposed to only use ridge structure features of ballprint (the hallucal area under the big toe) for recognition. They collected ballprint images from 54 newborns. The infants’ right and left ballprints were captured at the ages of 2 days (the first session), 2 months (the second session), and 6 months (the third session). In their experiment, the ballprints captured at the first session were discarded due to poor quality, and 192 ballpoints were selected manually from the other two sessions. The EER of matching ballprint images between session 2 and session 3 was 7.28%.

In this paper, we investigate the use of high resolution footprint image as a means of infant recognition. An infant footprint dataset was collected. The dataset includes 60 infants with both left and right footprints captured; each infant participated in all three sessions of data collection, most of which were at the ages of 1 month old, 3 months old, and 6 months old. We proposed a new minutiae descriptor based on deep convolutional neural network (CNN). Considering the lack of large footprint dataset, the transfer learning strategy is adopted. The CNN model is first trained on millions of hard samples of minutiae pairs selected from NIST SD14 [18] by a state of the art fingerprint matcher [4], then fine-tuned on a small set of infant footprint minutiae pairs. With the proposed minutiae descriptor, a modified match propagation algorithm [17] is proposed to match infant footprint images.

We conducted verification experiments for both single enrolled template and fusion of multiple enrolled templates, and showed the impact of age and time gap on performance. The Genuine Accept Rate (GAR) values at a False Accept Rate (FAR) of 0.01 of our footprint matching system are 61%, 55%, and 83% when matching session 1 vs session 2, session 1 vs session 3, and session 2 vs session 3, and fusing left and right footprint images. Comparison experiments with traditional handcrafted minutiae descriptor show the advantage of proposed descriptor.

2. Infant Footprint Matching

2.1. Ridge Width Normalization

Since the ridge spacing of infants’ footprints is much smaller than that of adults’ footprints, we appropriately adjusted the resolution of the infant footprint image before performing feature extraction. For simplicity, we model the
When the desired ridge width is used to establish a match and the desired ridge width is greater than the ridge width of the comparison, the ridge is called the "ridge crease alone." After thresholding and thinning the enhanced image, the enhanced footprint image is obtained by Gabor filtering [6].

\[ w = \alpha m + \beta, \]  

where \( \alpha \) and \( \beta \) are parameters of linear model which can be estimated from the training data. The scale factor \( s \) is then a function of infant’s age in months:

\[ s(m) = \frac{\hat{w}}{\alpha m + \beta}, \]  

where \( \hat{w} \) is the desired ridge width. In this paper, \( \hat{w} \) is set to 10 for footprint image with 500ppi resolution.

### 2.2. Feature Extraction

The local ridge structure of footprint is similar to that of palmprint. For a given footprint image after ridge width normalization, the palmprint feature extraction algorithm proposed in [10] was then used to compute feature components of footprint images in a block-wise manner in the frequency domain. For each block of size \( 16 \times 16 \), six peak points were selected from the magnitude map in the frequency domain. Then, a region growing algorithm is applied onto these feature maps to group orientation and frequency features; high quality groups with local features being consistent are combined together. With the estimated orientation fields, an enhanced footprint image is obtained by Gabor filtering [6]. After thresholding and thinning the enhanced image, minutiae were extracted from the skeleton image. These minutiae form the footprint template.

### 2.3. Minutia Embedding

Minutia point is a feature point of footprint image, which contains location and direction information. However, a minutia with only location and direction is not distinguishable from other minutiae. For the minutiae pairing purpose, it is desirable to describe a minutia with more discriminative features, which we call minutiae descriptor. The traditional minutiae descriptors are all handcrafted (e.g., MCC [5], orientation descriptor [24], local minutiae structure descriptor [14]). Recently, deep convolutional neural network has shown its power in supervised feature representation learning [2]. In this paper, we design a new minutia descriptor for infant footprint recognition based on deep learning.

Given a minutia \( m = (x, y, \theta) \), where \( (x, y) \) and \( \theta \) are the minutia coordinate and direction, respectively, we represent \( p_m \) as an image patch that centered at \((x, y)\) and around

<table>
<thead>
<tr>
<th>Author</th>
<th>Dataset</th>
<th>Sensor</th>
<th>Performance</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blake [3], 1959</td>
<td>1,388 newborn footprints collected immediately after birth</td>
<td>Inked</td>
<td>&quot;99% of the original footprints could be identified by flexure crease alone&quot;</td>
<td>Need manual processing</td>
</tr>
<tr>
<td>Shepard et al. [20], 1966</td>
<td>51 newborns printed immediately after birth and 5–6 weeks later</td>
<td>Inked</td>
<td>19.6% of the newborns matched correctly</td>
<td>Need manual processing</td>
</tr>
<tr>
<td>Thompson et al. [23], 1981</td>
<td>100 full-term infant footprints and 20 footprints from premature infants; for premature babies, a second set was captured at discharge 4-8 weeks later</td>
<td>Inked</td>
<td>Accuracy for full-term infant and premature infants was reported to be 11% and 0%, respectively</td>
<td>Need manual processing</td>
</tr>
<tr>
<td>Jia et al. [13], [12] and [11], 2012</td>
<td>1968 footprint images from 101 feet captured within 2 days after birth</td>
<td>Cannon Power-shot SX110 IS camera</td>
<td>EER = 1.34%</td>
<td>Low resolution images captured at one session by researchers</td>
</tr>
<tr>
<td>Balamenakshi and Sumanthi [1], 2013</td>
<td>240 images from 40 newborns, collected within 2 days after birth</td>
<td>Canon EOS 7D camera</td>
<td>65% rank-1 identification accuracy with background size of 40 images</td>
<td>22 of the subjects had special background setup during image capture</td>
</tr>
<tr>
<td>Kotzerke [16], 2013</td>
<td>54 subjects within 3 days after birth, 41 subjects at 8 weeks and 4 subjects at 6 months; 4 impressions of each foot were captured</td>
<td>Neossoa Printed Products Identifier and HP Scanjet G4010</td>
<td>No matching experiments conducted</td>
<td>Offline image capture, including cleaning feet, wiping with the inkless towelette, pressing on paper, and scanning</td>
</tr>
<tr>
<td>Kotzerke [15], 2014</td>
<td>54 newborn footprint been collected in three sessions (age: 2 days, 2 months and 6 months)</td>
<td>NEC PU900-10 (1000 ppi)</td>
<td>EER = 7.28% in matching footprints captured at 6 months and 2 months age</td>
<td>The first session (2 days old) images were not used due to their low quality. 192 images from the second and third sessions were manually selected for matching</td>
</tr>
<tr>
<td>This work</td>
<td>756 images from 42 infants, captured in three sessions (age: 1 month (session 1), 3 months (session 2) and 6 months (session 3))</td>
<td>Watson Mini (500 ppi) from Integrated Biometrics, Inc.*</td>
<td>GAR values (at FAR=0.01) are 61%, 55%, and 83% when matching session 1 vs. session 2, session 1 vs. session 3, and session 2 vs. session 3, and fusing left and right footprint images</td>
<td>Images were captured by doctors did not have any special training in footprint capture; no special background was setup during image capture</td>
</tr>
</tbody>
</table>

* http://www.integratedbiometrics.com
the direction $\theta$ of minutia $m$. In this paper, the size of the cropped image patch is $160 \times 160$ pixels.

A minutia embedding function $f$ takes an image patch $p$ as input, and outputs a feature vector $h$ (also called minutia descriptor). The minutia embedding process is represented as $h = f(p)$. To obtain the embedding function $f$, a deep neural network is constructed. We train the DNN with two sets of minutiae pairs, one large fingerprint minutiae dataset set for coarse training and one small infant footprint minutiae dataset set for fine tuning. Given a minutiae pair $(m, m', c)$, where $c = 1$ if $m$ and $m'$ from the same minutia of footprint, otherwise $c = 0$, we first apply network $f$ on $m$ and $m'$ to obtain the feature vector $h = f(p_m)$ and $h' = f(p_{m'})$, respectively.

Then, we construct two tasks to train the network. The first is to model the problem as a regression problem. The inner product, $q = h \otimes h'$, is computed to predict the similarity of $m$ and $m'$. The second task is binary classification task. The feature vector $h$ and $h'$ are concatenated and then followed by a fully connected layer to classify whether the two minutiae are mated or not. The flowchart of the proposed minutiae embedding network is shown in Fig. 1. The weights of the embedding network are shared by both feature extraction of image patches. The VGG-16 model [22] is used as the embedding function $f$, and the size of output feature vector is 2048. The stochastic gradient descent (SGD) optimization method is used in training.

### 2.4. Template Matching

Currently, there is no automatic matching algorithm specifically designed for high resolution footprint images. Given that the ridge characteristics of footprints are similar to those of palmprints, we modified the palmprint matcher proposed in [17] for our footprint matching problem. The modified algorithm matches two footprint image templates using the following steps:

1. Given a set of minutiae along with their local orientation descriptors [24], a set of orientation descriptor centroids is obtained by k-means clustering algorithm (with $K = 32$). This step is performed offline.

2. In matching two templates of footprint, we first classify each minutia into one of the centroids.

3. Form a set of initial minutiae pairs by matching the proposed minutiae descriptors between the minutiae of the same type of orientation centroid. The top $N$ minutiae pairs with highest similarity are selected as initial minutiae pairs.

4. From an initial minutiae pair with the highest similarity, local match propagation is performed to find additional minutiae pairs iteratively. A match score is computed based on the number of matched minutiae and their similarities. The minutiae similarity is computed as the inner product of the proposed minutiae descriptor. Since the number of minutiae in footprint is large, the match propagation strategy is helpful to reduce the overall computation time.

5. The final match score is the maximum score in match propagation. In our experiments, five initial minutiae pairs were selected (i.e., $N = 5$).

The difference between the proposed footprint template matching algorithm and the one proposed in [17] is that we replace the local minutiae structure descriptor with the proposed fixed length minutiae descriptor in selecting initial minutiae pairs and in computing the final match scores.

### 2.5. Training Sample Selection

The minutiae pairs for training the neural network is critical to the success of training. We collected two datasets for training the model. The first dataset is a set of fingerprint minutiae pairs selected from NIST SD14 database [18]. This database includes 27000 fingers with each finger containing two impressions. Manual minutiae pairing is very time consuming and expensive. In this paper, the fingerprint matcher proposed in [4] was used to select the training minutiae pairs. Given two fingerprint images, the matcher outputs a match score indicating the similarity of these two images and a list of paired minutiae with corresponding minutiae similarities. For the database of NIST SD14, we have 27000 genuine matches and $\frac{27000 \times (27000 - 1)}{2}$ impostor matches in total. If the genuine match score greater than a threshold, then the paired minutiae are selected as positive training samples. The threshold is to make sure the output minutiae pairs are truly mated. The number of impostor matches is too big for our task. We randomly selected 4 millions of impostor matches, from which the output minutiae pairs are selected as negative training samples. To simulate the scaling problem of infant footprint images, we augment the positive training samples by resizing the image patches with a random scaling factor between 0.8 and 1.2. Finally, we generated 4 millions positive training samples and 4 millions negative training samples.

The dataset of fingerprint minutiae pairs is used as a coarse training of our model. To adapt the model specifically to the problem of infant footprint recognition, we collected another dataset for fine tuning the coarse model. We have collected footprint images from 18 infants. Each infant participated three sessions with each session provided three footprint impressions of both left and right feet. The palmprint matcher proposed in [17] was used to select minutiae pairs. For any genuine footprint match, we visually checked the minutiae correspondences side by side and determined whether the minutiae pairs are correct or not. If the matching is successful, all the minutiae pairs are kept as positive
3. Experiments

3.1. Infant Footprint Data Collection

We collected the footprint data in Changqing Chaoming community hospital in Hangzhou City, Zhejiang Province, China. The infants come to the hospital with parents for regular physical examination. Typically, the first infant visit to the hospital should be at the age of 1 month, with subsequent visits at 3 months and 6 months. The doctors/nurses working in the hospital collected the footprint images of each infant during the three visits of the infants. For reasons unknown to us, some of the subjects could not complete all three capture sessions. These subjects were not considered in our matching experiments.

We used Watson Mini fingerprint reader, manufactured by Integrated Biometrics company\(^7\), to collect footprint data. This device is designed for dual fingerprint capture. Given that there are no specially designed live footprint scanners available in the market due to the large sensing area of Watson Mini, we use it for footprint collection.

The total number of footprint images captured is 1080 from 60 infants. Three impressions were captured for each foot at each session. The dataset is divided into two sets, i.e., training and testing datasets. The training dataset contains 18 infants and testing dataset contains 42 infants.

It should be noted that the hospital staff who were collecting the images did not have any expertise or background related to footprints or biometric recognition. We trained them with some basic guidelines about how to collect footprint images before they formally started. There is no special setup of background environment at the hospital for capturing footprints. Thus, this dataset is more realistic than datasets used in prior works which were taken by experts and under controlled conditions [11, 15, 16].

3.2. Experiment Protocols

According to the number of footprint images enrolled in the system, we conducted three types of matching experiments:

1. **Scenario A**: In this scenario, only one footprint image is used as template, and one footprint image as query.

2. **Scenario B**: We assume there are three enrolled images of the same foot for each subject (left foot and right foot are regarded as two different subjects). The matching is conducted between one query footprint and three enrolled templates. The resulting three match scores were combined using the maximum fusion strategy to arrive at the final match score.

3. **Scenario C**: The third experiment assumes that there are six enrolled templates for one subject, with three templates for each foot. The matching is conducted between two queries, one for each foot, and their respective enrolled templates. Thus, we have six match scores and the final score is the fusion of six match scores. In this experiment, we used maximum fusion strategy.
3.3. Matching Accuracy

3.3.1 Scenario A

In the first scenario, we conducted verification across different sessions. The receiver operating characteristics (ROC) curves are reported in Fig. 2. The ROC curves are classified into three cases: Session 1 vs Session 2 (S12), Session 1 vs Session 3 (S13), and Session 2 vs Session 3 (S23). The numbers of genuine matches are 756 for all three cases. The number of impostor matches is 22,788. From the results, we observe that S23 (i.e., Session 2 vs Session 3) outperforms the other two cases. This is expected because ages of infants associated with these templates are older than that of S12 and S13. The templates in S12 and S13 were all collected in the first session. Still the performance of S12 is much better than S13 because the time gap in S12 is shorter than that in S13 by about 1 month. The average time gaps in S12, S13, and S23 are 1.95 months, 4.81 months, and 2.86 months, respectively. We do not have exactly the same time gaps because some subjects did not come for physical examination at their scheduled time. The GARs (at FAR=0.01) of S12, S13 and S23 are 48%, 33%, and 64%, respectively.

3.3.2 Scenarios B & C

For scenarios B and C, we fused the three (scenario B) or six (scenario C) match scores to get a final score by the maximum fusion strategy. The ROC curves of scenarios B and C are shown in Fig. 3 and Fig. 4, respectively. Similar trends as in Fig. 2 are observed. However, the curves are not as smooth as those in scenario 1 due to fewer number of match scores.

We again use S12, S13 and S23 to represent the three cases as that in scenario A. For scenario B, the numbers of genuine matches are 252 for all three cases. The GAR values (at FAR=0.01) of S12, S13 and S23 are 53%, 42%, and 74%, respectively. For scenario C, the number of genuine matches are 126 for all three cases. The GAR values (at FAR=0.01) of S12, S13 and S23 are 61%, 55%, and 83%, respectively.

In Fig. 5, we compare the matching performance of all three scenarios. The fusion of multiple templates improve the overall performance. The overall GAR values (at FAR = 0.01) of scenario A, B and C are 60%, 50% and 70%, respectively.
3.4. Comparison

The proposed minutiae descriptor is compared with the local minutiae structure (LocalMntStruct) descriptor proposed in [17]. The proposed descriptor is learned from a large dataset while the LocalMntStruct is hand designed. We compared the Scenario A, B and C, separately. The ROC curves are shown in Fig. 7. We see that our descriptor outperforms the LocalMntStrut descriptor for all three scenarios.

3.5. Computation Time

We implemented the footprint matching algorithm on a Macbook Pro laptop with 2.6 GHz Intel Core i7 CPU, 16 GB 1600 MHz DDR3 RAM, 1 TB SSD disk, NVIDIA GeForce GT 750M GPU and OS X Yosemite operating system. The GPU was used for minutiae descriptor extraction. The average computation time for enrolling one footprint image is about 20 seconds, including preprocessing, enhancement, and minutiae detection; and the average computation time for one genuine and one impostor matching are 130 ms and 330 ms, respectively.

4. Conclusions

Infant recognition is an urgently needed technology in many applications, such as tracking child vaccination, identifying missing children, and child welfare. Due to the rapid growth rates of babies, there are large variations in many biometric modalities, especially face. Like fingerprint and palmprint, footprint is a friction ridge pattern that is formed in gestation. Footprint capture has some advantages over fingerprint or palmprint. For example, babies tend to keep their fists closed, and there is relatively lower level of parental concerns for footprint acquisition because it is already captured for newborns in many countries.

In this paper, we proposed a new minutiae descriptor based on deep convolutional neural network for infant footprint recognition. To train this network, two sets of minutiae pair datasets were constructed. The first dataset, including 8M samples of fingerprint minutiae pairs, is used for coarse model training, and the second dataset, including 60K samples of infant footprint minutiae pairs, is used for fine tuning the model. The experimental results on a database collected by hospital staff (without any expertise in footprints) have shown both the feasibility as well as challenges of using footprint as a biometric for infants. Comparison experiment with the state of the art palmpint matcher also shows the advantage of the proposed minutiae descriptor for footprint.

References

Figure 6. An example of successful genuine matching. The match score is 0.08 (threshold at FAR=0.01 is 0.0013, score ranges in [0,1]) with 102 pairs of corresponding minutiae.

Figure 7. Comparison of ROC curves of the proposed descriptor and LocalMntStruct descriptor for three simulated matching scenarios.


