Discovering Correspondences between Fingerprints based on the Temporal Dynamics of Eye Movements from Experts

Chen Yu, Thomas Busey, John Vanderkolk

Indiana University
1101 East 10th Street, , Bloomington, IN, 47405
{chenyu,busey}@indiana.edu

Abstract. Latent print examinations involve a process by which a latent print, often recovered from a crime scene, is compared against a known standard or sets of standard prints. Despite advances in automatic fingerprint recognition, latent prints are still examined by human expert primarily due to the poor image quality of latent prints. The aim of the present study is to better understand the perceptual and cognitive processes of fingerprint practices as implicit expertise. Our approach is to collect fine-grained gaze data from fingerprint experts when they conduct a matching task between two prints. We then rely on machine learning techniques to discover meaningful patterns from their eye movement data. As the first steps in this project, we compare gaze patterns from experts with those obtained from novices. Our results show that experts and novices generate similar overall gaze patterns. However, a deeper data analysis using machine translation reveals that experts are able to identify more corresponding areas between two prints within a short period of time.

Keywords: fingerprint, Cognitive and Behavioral Studies, Computational data analysis, Data mining

1 Introduction

The goal of our study is to use computational techniques derived from machine translation to explore the temporal dynamics of complex visual processing tasks in fingerprint examinations. These examinations involve a process by which a latent print, often recovered from a crime scene, is compared against a known standard or sets of standard prints. Despite advances in automatic pattern matching technology [2,3], latent prints are still compared by human experts. In the United States and in many other countries there is no fixed number of matching points or details that is mandated by the courts or forensic science community [1]. This implicitly gives the examiners some latitude in terms of the details they choose to use in order to determine whether the two prints come from the same source. For example, instead of relying on matching minutiae, the examiner is free to use what details and patterns they feel are relevant, including what is known as first level detail of general direction of ridges, second level specific ridge paths, and third level detail or the texture and shape of individual ridge elements.

While most computational fingerprint projects focus on building automatic pattern recognition systems, the goal of the present study is different. Here we aim at a better understanding of human fingerprint expertise which will not only provide useful evidence to justify the use of fingerprints in count but also provide principles and insights to develop better automatic recognition systems to reach expert-level performance (especially for latent prints). To achieve this goal, we developed a mobile sensing system to collect moment-by-moment eye gaze data from fingerprint experts, and then develop and use computational approaches to analyzing such data to gain a better understanding of their fingerprint examination practices. From the scientific perspective, compared with other forensic evidence (e.g. DNA), there is less known about the information content in fingerprints or statistics such as the likelihood of a random correspondence. On the one hand, we know fingerprint images are information-rich: even twins with identical DNA have different fingerprints [4]. On the other hand, experts may choose to rely on different sources of information depending on the circumstances and their training, which may raise issues with respect to the nature of the evidence presented in court. There exist no explicit standards for what information shall be used in latent print examinations. A recent National Academy of Sciences report [1] was somewhat critical of the language used by examiners when testifying about their results, and called for more training and research on the nature of the latent print examinations. The report revealed weaknesses in our current knowledge about what information experts rely on when performing identifications and exclusions. Part of the difficulty resides in the fact that much of the processes of perception are unconscious and can be difficult to translate into language and examinations may be subject to extra-examination biases [5]. The aim of the present study is to systematically study fingerprint experts to understand better what cognitive and perceptual processes support their matching decisions. We are particularly interested in the temporal dynamics of the search process as one marker of expertise. From engineering and applied perspectives, understanding where experts look would also provide useful insights on how to build computational systems that can either perform similar tasks or assist human experts to better perform such tasks.

2. Related Work

There are several different approaches to automatic fingerprint recognition [2,3]. Su et al. [6] presented a novel individuality evaluation approach to estimating the probability of random correspondence (PRC) based on the distribution of three fingerprint features: ridge flow, minutiae and minutiae together with ridge points. Ratha et al. [8] generated multiple cancelable identifiers from fingerprint images and those identifiers can be cancelled and replaced when compromised, showing that feature-level cancelable biometric construction is practicable in large biometric deployments. Tuyls et al. [7] applied template protection schemes to fingerprint data by splitting the helper data in two parts, one part determines the reliable components and the other part allows for noise correction on the quantized representations.

The goal of the present study is not to develop an automatic fingerprint recognition system. Instead, we intend to address fundamental research questions on human fingerprint examination using computational analysis techniques to discover how human fingerprint experts conduct a pattern matching task given both inked and latent prints and what visual features on the fingerprints they used which will shed light on building expert-like automatic fingerprint systems. To the best of our knowledge, this is the first study to use eye tracking to understanding the expertise fingerprint examiners (to our knowledge), eye tracking techniques have been successfully applied in several other scientific fields to assess implicit knowledge from human experts in certain domains. The field of mammography research has adopted similar eye tracking methodologies. Krupinski et al. [9] have used eye tracking to investigate not only what features radiologists rely on when inspecting mammograms, but also to suggest cognitive mechanisms such as holistic processing when experts are viewing mammograms. A similar work with chest x-rays demonstrated that dwell times were longer on missed tumors than at other locations, suggesting that errors in radiology are due to identification problems rather than detection problems. The field of questioned documents has also benefited from an eye tracking approach [13] as well, which has helped to delimit the visual features that experts rely on when comparing signatures. All of those applications of eye tracking indicate the promise of using eye movement data in fingerprint examination studies.

3. Experiment and Data

The first challenge to achieve our research goal is to collect behavioral data from fingerprint experts. As noted earlier, assessing implicit knowledge from human experts is a non-trivial task as such knowledge cannot be inferred from survey or questionnaire by asking experts where they look and why they look at there. Humans generate about 3 eye fixations per second to gather visual information from their environment. If we assume that experts move their eyes equally frequently, they must produce a large number of eye fixations in a print examination and clearly they cannot recall precisely which areas in the print they just visit moment by moment. Meanwhile, we know their eyes are not random; instead they actively collect visual information that is then fed into their brain for their decision making. In light of this, our solution is to collect momentary eye movement data from latent print examiners and use advanced computational techniques to analyze such data to lead to a better understanding of their expertise.

Nowadays eye tracking becomes a more and more popular technique in behavioral studies as well as in human-computer interaction and marketing studies. There are several commercial eye tracking systems available in the market (e.g., Tobii eye tracking system, www.tobii.com). However, in practice, gathering data from experts poses a particular challenge since they work in different police or FBI branches throughout the U.S. and most commercial eye tracking systems can only be used in a laboratory environment.

. To solve this problem, we developed a portable eye tracking system that is based on an open-source hardware design [10], allowing us to recruit experts at forensic identification conferences and collect gaze data from them. Participants were seated approximately 60 cm (~24 inches) away from a 21” LCD monitor. The fingerprint images were presented side-by-side on a 21” LCD monitor at a resolution of 1580 x 759 pixels. The monitor itself was set to its native resolution of 1680 x 1050 pixels. As shown in Figure 1, participants wore our head-mounted eye tracker which used two small cameras to monitor the eye and the view of the scene respectively according to the hardware proposed by [10]. Both cameras are mounted and specially located on a pair of lightweight safety glasses. One infrared light is located next to the eye-camera in order to illuminate the eye properly. This light provides us a constant spot of white light known as the first corneal reflection, which will be used for further offline analysis using the *ExpertEyes* software, an open source approach for analyzing eye-tracker data http://code.google.com/p/experteyes/wiki/ExpertEyes) developed by our research group. This image processing software takes two video streams from two cameras and generate (x,y) coordinates indicating where a person is looking in the scene camera. Further, we developed another program to convert (x,y) coordinates in the scene camera into (x,y) coordinates on the print image by detecting the location of the monitor in the scene camera's view. In this way, our system captures, moment by moment, which area in a print a person is gaze at.

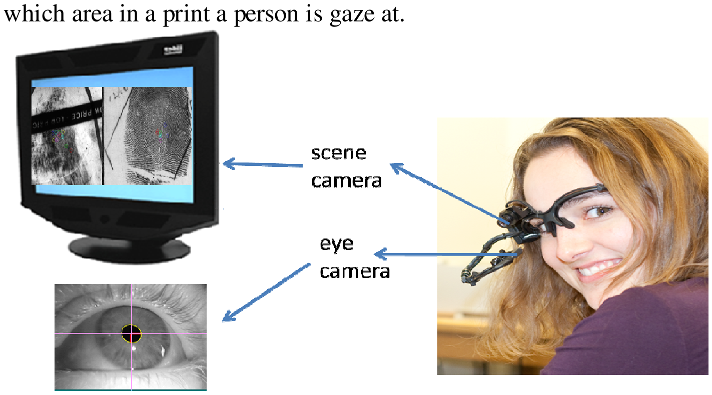

Figure 1: Our eye tracking system consists of two cameras -- the eye camera captures eye images while the scene camera captures the visual information from the first-person view (in this case, fingerprints on a computer screen). Our home-made software relies on image processing techniques to infer where a participant looks on the computer screen.

In our experiments, participants included both experts and novices. The reason to include novices is that as a first step in this project, we wanted to establish a baseline to compare with. Even though novices do not have fingerprint expertise, our human visual system is still quite powerful to detect the similarities between visual stimuli. Therefore, the data from experts would allow us to distinguish between general capabilities in human visual system and true fingerprint expertise in experts.
Human observers (e.g. experts or novices) were asked to visually examine a set of fingerprints one by one and decided whether the two fingerprints displayed simultaneously match with each other or not. There was no particular instruction about where they should look during the matching task so that they could freely move their eyes on fingerprint images. The typical latent print examination can take hours or even days to complete for difficult prints. Examiners will sometimes start with an inspection of the latent print, which may be augmented by photographs, notes or drawings. They then move on to the inked print, which helps prevent situations in which they may begin to see details in the latent print that are shared with the ink print. Experts often describe spending hours or even days with a problematic or difficult print. Our statistical analyses, however, require a relatively large number of

images in order to ensure reliable results, and we wanted to gather a complete dataset from each participant. As a result, we decided to limit the amount of time that each participant could spend on each fingerprint to avoid corrupting our database with uninformative eye movements as they waited for the next print.

There were two datasets used in the present study. Dataset 1 was collected from a task in which participants were asked to examine 35 pairs of inked and latent prints. Each trial of examination took 20 seconds and then they moved to the next image pair. The stimuli for this study were taken from National Institutes of Standards and Technology Special Database 27. The print stimuli in dataset 2 consisted of two clear images that were scanned at a 1000-dpi from clear inked prints collected from members of the Bloomington Indiana community. Participants were allowed to review two prints individually for 5 seconds on each, and they had 10 seconds to review two prints displayed simultaneously. There were 12 fingerprint experts and 12 novices in each study. Experts were recruited at forensic identification conferences in Nevada, Illinois and Indiana, while the novices were members of the Bloomington, Indiana community. Figure 2 illustrates an example stimulus in dataset 1, along with the eye fixations and eye trace for participants. Our eye tracker generates gaze data at the sampling rate of 30Hz. In total, there were approximately 504,000 gaze data points in each dataset.

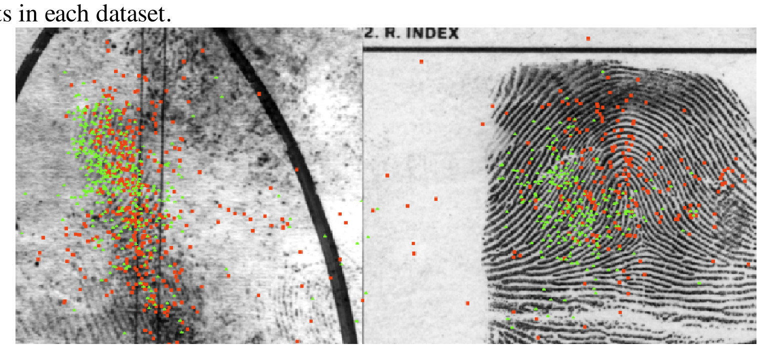

Figure 2. Latent/inked pair from dataset 1, with fixations from all experts overplotted as green dots, and fixations from all novices overplotted as red dots. The green dots tend to be clustered in the upper-right portion of the inked print (right image), which corresponds to the area of high detail in the latent print. However, novices have a much wider distribution of fixations, including in regions that have very poor image quality.

4. Method

Given the high spatial and temporal resolutions of our eye tracking data, the technical challenge is to discover meaningful patterns from those rich datasets. Our computational data analysis consists of four steps as illustrated in Figure 3: 1) temporal fixation finding: reducing the continuous time series of raw gaze data into a sequence of eye fixations defined mostly by the speed of eye movements over time; 2) spatial clustering to calculate Regions-of-Interest (ROIs): clustering (x,y) gaze data points into several clusters/ROIs based on the spatial distribution of gaze data on the prints; 3) alignment: segmenting the ROI sequences into ink-latent fixation pairs based on temporal proximity; 4) using a machine translation method to compute the correspondences between ROIs in the inked and latent prints. As a result, we extract the patterns of which corresponding areas that experts examine back and forth

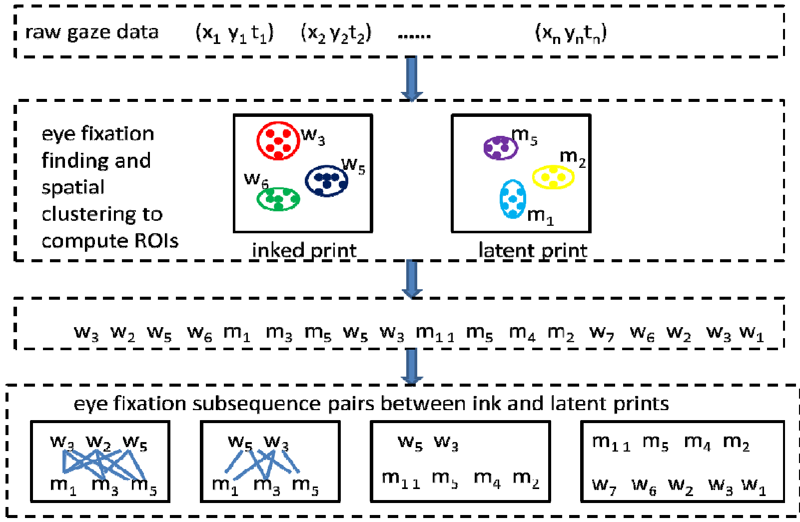

Figure 3: The overview of our data analysis approach consisting of four components.

between two prints and which areas that novices pay attention to when conducting the same matching task. In the following, we will provide technical details of each component briefly described above.

4.1. Temporal Fixation finding

We have developed our own algorithm of eye fixation finding which is composed of four steps: 1) First, we computed the magnitude of velocity from raw eye movement data (x,y); 2) we next used a pre-defined threshold to segment the whole continuous stream into several big segments that correspond to dramatic eye location changes; 3) we analyzed each big segment and re-segmented each into individual segments that correspond to small eye position changes. Those small segments may or may not correspond to fixations; 4) Finally, we took spatial information into account by merging small segments (detected from Step 3) if they were spatially close to each other (e.g. eyes moving around an area with a certain speed). After the above four steps, we successfully segmented a continuous eye movement stream into several eye fixations by integrating both temporal (the speed of eye movement, etc.) and spatial (the overall spatial changes of eye gaze location) information.

4.2 Spatial Clustering.

Given raw (x,y) coordinates from Step 1, the next step is to group those data points into several clusters – the regions that participants frequently visit. We used Hierarchical agglomerative clustering [11]. The basic idea is to treat each eye fixation location as a singleton cluster at the beginning and then successively merge (or *agglomerate*) pairs of clusters until all clusters have been merged into several prototype clusters. We took the centroid of each cluster as a ROI. The criterion used to terminate the clustering algorithm is the minimal distance required to group two data points. In the present study, the distance was set to be 20 pixels which match to a visual angle of 0.5 degrees.

4.3 Temporal Alignment

Now that we have a sequence of ROIs extracted from participants' gaze data – some over one image and the rest on the other. Our goal is to calculate correspondences between gazed regions in one image with gazed regions in the other image as participants conducted the matching task. To do so, we view this task as similar to machine translation in natural language processing [12]. The general idea of machine translation is this: assume that we have parallel texts from two languages, for example, "Harry Potter and the Order of the Phoenix" in both English and French, the goal of machine translation is to infer which two words in the two languages correspond. This inference can be done based on statistical information, such as how frequent "egg" in English and "oeuf" in French co-occur together and how frequent "egg" appears without "oeuf". Intuitively, if a word in English always co-occurs with another word in French and that word in English appears only when the other word in French appears, then those two words are likely to correspond to each other. Most often an assumption in machine translation is a sentence-level assignment – which sentence in English maps to which one in French is known. Say it in other way, we have sentence pairs from two languages and use this data to infer word correspondences.

In the fingerprint matching task, we conceptualize ROIs from one image as words in English, and ROIs on another print as words in French. Based on this conceptualization, the aim here is to find which gazed region in one print maps to which gazed region in the other print. To achieve this, we also need to segment continue gaze data generated by participants into "sentence" pairs. This is done based on the observation that participants may generate a few fixations on one image, switch to examine another image with more fixations to search for corresponding areas on the other image. In light of this, and as showed at the bottom of Figure 3, we first divided a whole sequence into several subsequences using the visual attention switches between two prints as breaking points, and then grouped those subsequences into several pairs based on temporal proximity. The outcome of this alignment is a set of fixation sequence pairs from which we further calculated which fixated area in one image maps to what fixated area in the other image in the next step. We call each pair of two fixation subsequences on two prints a searching instance as we assume that participants were comparing and matching regions between two prints through those eye fixations on both prints. Figure 3 shows four searching instances extracted from a continue ROI sequence. To the degree to which experts will find matching features in both prints we will be able to discover these through machine translation.

4.4 Machine translation

The general setting of the final step is as follows: suppose we have one ROI set from one image $\boldsymbol{X} = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots \ldots, \boldsymbol{w_N}\}$ and the other ROI set from the other image $\boldsymbol{Y} = \{\boldsymbol{m_1}, \boldsymbol{m_2}, \ldots \ldots, \boldsymbol{m_M}\}$, where N is the number of ROIs in one print and M is the number of ROIs in the other print. Let S be the number of searching instances. All gaze data are in a dataset $\boldsymbol{\chi} = \left\{(\boldsymbol{S_w^{(s)}}, \boldsymbol{S_m^{(s)}}), \boldsymbol{1} \leq \boldsymbol{s} \leq \boldsymbol{S}\right\}$, where for each searching instance, $\boldsymbol{S_w^{(s)}}$ consists of r ROIs $\boldsymbol{w_{u(1)}}, \boldsymbol{w_{u(2)}}, \ldots \ldots, \boldsymbol{w_{u(r)}}$, and $\boldsymbol{u(i)}$ can be selected from 1 to N. Similarly, the corresponding gaze sequence on the other print $\boldsymbol{S_m^{(s)}}$ includes $\boldsymbol{l}$ possible ROIs $\boldsymbol{m_{v(1)}}, \boldsymbol{m_{v(2)}}, \ldots \ldots, \boldsymbol{m_{v(l)}}$ and the value of $\boldsymbol{v(j)}$ is from 1 to M. In the example in Figure 3, there are four searching instances in which every ROI in one image can potentially be mapped with any co-occurring ROIs in the other image. The computational challenge here is to build several one-to-one mappings from many-to-many possible mappings within multiple searching instances as not all of the ROIs (generated by participants) within an instance can reliably map to the other ROIs on the other image. We suggest that to figure out which ROI in one image goes to which ROI in the other image, a good solution shouldn't consider the mapping of just a single ROI-ROI pair, but instead we should estimate all these possible mappings simultaneously. Thus, we attempt to estimate the mapping probabilities of all of these pairs so that the best overall mapping is achieved. In doing so, the constraints across multiple searching instances and the constraints across different ROI-ROI pairs are jointly considered in a general system which attempts to discover the best ROI-to-ROI mappings based on the overall statistical regularities in the whole eye fixation sequence.

Formally, given a dataset χ, we use the machine translation method proposed in [12] to maximize the likelihood of generating/predicting one set of ROIs from one image given a set of ROIs from the other image:

$$P\left(S_m^{(1)}, S_m^{(2)}, \ldots, S_m^{(S)} \middle| S_w^{(1)}, S_w^{(2)}, \ldots, S_w^{(S)}\right)$$
$$= \prod_{s=1}^{S} \sum_a p(S_m^{(s)}, a | S_w^{(s)})$$
$$= \prod_{s=1}^{S} \frac{\epsilon}{(r+1)^l} \prod_{j=1}^{l} \sum_{i=0}^{r} p(m_{v(j)} | w_{u(i)}) \qquad (2)$$

where the alignment a indicates which ROI in one image is aligned with which ROI in the other image. $p\left(m_{v(j)} \middle| w_{u(i)}\right)$ is the mapping probability for a ROI-ROI pair and ϵ is a small constant.

To maximize the above likelihood function, a new variable $c\left(m_m \middle| w_n, S_w^{(s)}, S_m^{(s)}\right)$ is introduced which represents the expected number of times that any particular ROI w_n in one subsequence $S_w^{(s)}$ generates any specific ROI m_m in the other subsequence $S_m^{(s)}$:

$$c\left(m_m \middle| w_n, S_w^{(s)}, S_m^{(s)}\right) = \frac{p\left(m_{v(j)} \middle| w_{u(i)}\right)}{p\left(m_m \middle| w_{u(1)}\right) + \cdots + p\left(m_m \middle| w_{u(r)}\right)}$$

$$\times \sum_{j=1}^{l} \delta(m_m, v(j)) \sum_{i=1}^{r} \delta(w_n, u(i)) \qquad (3)$$

where δ is equal to 1 when both of its arguments are the same and equal to zero otherwise. The second part in Equation (3) counts the number of co-occurring times of w_n and m_m. The first part assigns a weight to this count by considering it across all the other ROIs in the same searching instance. By introducing this new variable, the computation of the derivative of the likelihood function with respect to the mapping probability $p(m_m|w_n)$ results in:

$$p(m_m|w_n) = \frac{\sum_{s=1}^{S} c\left(m_m \middle| w_n, S_w^{(s)}, S_m^{(s)}\right)}{\sum_{m=1}^{M} \sum_{s=1}^{S} c\left(m_m \middle| w_n, S_w^{(s)}, S_m^{(s)}\right)} \qquad (4)$$

As shown in Algorithm 1, the method sets an initial $p(m_m|w_n)$ to be flat distribution, and then successively compute the occurrences of all ROI-ROI pairs $c\left(m_m \middle| w_n, S_w^{(s)}, S_m^{(s)}\right)$ using Equation (3) and the mapping probabilities using Equation (4). In this way, our method runs multiple times and allows for re-estimating ROI-ROI mapping probabilities. The detailed technical descriptions can be found in (Brown et al., 1994).

Algorithm 1 Estimating ROI-ROI mapping probabilities

Assign initial values for $p(m_m|w_n)$ based on co-occurrence statistics.
repeat
    E-step: Compute the counts for all ROI-ROI pairs using equation 3.
    M-step: Re-estimate the mapping probabilities using equation 4.
until the mapping probabilities converge.

---

5. Results

As the first steps of this project, we focus on comparing gaze patterns between experts and novices. The first set of results is to compare overall gaze fixation statistics between the expert and novice groups. In dataset 1, we found no differences in terms of the average duration of each fixation for the two groups (M_{expert}= 185.51ms M_{novice}=183.50; p>0.5). In addition, we measured the proportion of the overall fixation duration (within a 20-second trial) and again the results showed no differences (M_{expert}= 11.04 sec M_{novice}=10.96 sec; p>0.5). The results derived from dataset 2 are almost identical with those from data set 1. Thus, we couldn't distinguish between experts and novices based on their overall eye movement patterns. One plausible reason is that participants in both groups were actively engaged in the pattern matching task and therefore their overall eye movement pattern were driven by low-level visual saliency of fingerprints and were controlled by the same neural architecture in the brain [14]. However, if this is the case, we expect that a deeper computational data analysis based on machine translation described earlier may reveal the differences between two groups. In particular, our research questions are 1) whether experts' gaze patterns are more consistent than those from novices; 2) whether experts can identify more ROI-ROI pairs within a short period of time; and 3) whether those pairs identified by experts or novices are actually correct.

Indeed, the results of the machine translation analysis applied to experts and novices are very clear. Our method produced all of the possible ROI-ROI mappings between fixations on the two images. We chose two criteria to select reliable ROI-ROI pairs. First, the co-occurring frequency is at least 2, meaning that a participant at least looked at one region in one image and subsequently look at another region in the other region twice. Second, the overall mapping probability $p(m_m|w_n)$ needs to be greater than 0.4. Based on this selection, the experts have an average of 17.1 reliable mappings/links found, while the novices have an average of 7.1 links found ($t(34)=-6.73$; $p<0.001$, sd = 8.84) from dataset 1. For dataset 2 with clean prints, we found a similar result. The machine translation found an average of 11.1 links for experts and 8.3 links for novices ($t(29)=-3.18$; $p<0.01$, sd = 4.59). Figures 4 (from dataset 1) and 5 (from dataset 2) show examples from both experts and novices. This demonstrates that the temporal dynamics for experts are much better as input to the machine translation algorithm in terms of assigning corresponding links between the two images. Within a short period of time, experts can manage to find more corresponding pairs than novice do.

Are those pairs visually spotted by either experts or novices are actually correct correspondences? Did expert do a better job in finding correct correspondences than novices did? To address those questions, we asked an expert to independently place corresponding marks on the pairs of clean prints in dataset 2 (the latent/inked pairs from dataset 1 did not have sufficient details to allow this procedure with high accuracy). We then used a second order polynomial function to map every point on the left print to a corresponding point in the right print for each print pair. As shown in Figure 5, each link (highlighted by red lines) has a ground-truth location (indicated by green dots) on the right print that is the matching location for the left side of the link. We computed the distance between this true matching location and the location obtained for each correspondence pair discovered by the machine translation algorithm. We found that both groups produce similar deviations. The mean for experts was 53.8 pixels, and the mean for novices was 51.2 pixels. These values were not significantly different ($t(29)<1$). This deviation is about 1 degree of visual angle which for our images corresponds to about 2 ridge widths in distance. This is perhaps a surprisingly small number given that the machine translation algorithm does not know about space directly.

Taken together with the results of the number of reliable pairs found, the converging evidence is that both experts and novices can identify matching locations (corresponding ROIs, etc.) between two (clear) prints, However, experts can find more of those pairs than novices do. In addition, our results also demonstrate the power of our data analysis method based on machine translation. Given large datasets of eye movement data, our method can successfully extract meaningful patterns that are not apparent from simple methods (e.g. computing the average fixation during or the total fixation time). Specially, this approach allows us to compute matching locations between two fingerprints based on gaze data generated by participants. Thus, the method fits quite well with the fingerprint matching task which may explain the reason of its success.

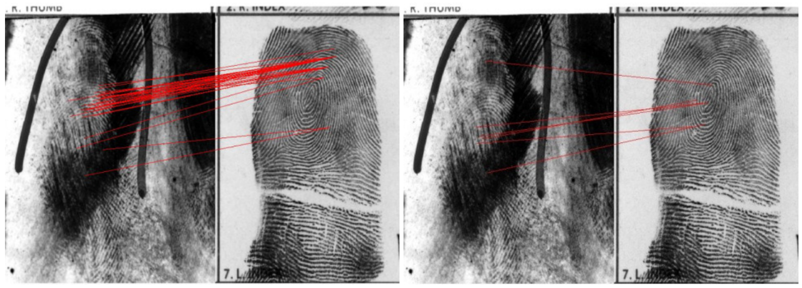

Figure 4: the corresponding regions from Data Set 1 with inked and latent prints. Left: an example result from experts. Right: an example result from novices.

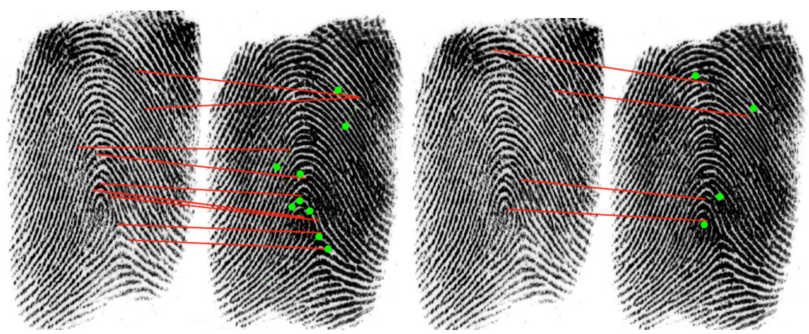

Figure 5: the corresponding regions from Dataset 2. Left: an example result from experts. Right: an example result from novices.

6. Conclusions

The focus on the present study is not to build an automatic fingerprint system. Instead, we intend to address fundamental questions on fingerprint examination. Most fingerprint examination practices still heavily rely on human expertise to confirm and double check the results produced by automatic systems, especially in harder cases with latent prints. A more complete understanding of human fingerprint expertise will serve at least two important purposes. First, the results can be used as scientific evidence to justify human fingerprint practices. Second, the insights and principles gained by computational analyses of expert's gaze data can be incorporated into automatic recognition systems to improve the performance of those systems. Toward this goal, we developed and used an eye tracking device to record momentary gaze data from both fingerprint experts and novices. We then applied machine learning techniques to extract gaze patterns from both groups. We found that experts are able to identify more corresponding pairs than novices do within a short period of time and use that information to make correct judgments, showing the promise of this research direction. In our future work, we plan to further analyze image regions identified by our current system to infer what visual features are encoded in those regions. In addition, the technical contribution of the present paper is to introduce a machine translation method to compute correspondences between two prints based on gaze data generated by human experts. In this way, our approach can be viewed as human-guided machine learning as experts' gaze are used as supervisory signals to the automatic correspondence detection system.

Acknowledgement

This research was supported by NIJ grants NIJ 2005-MU-BX-K076 and NIJ 2009-DN-BX-K226. The authors would like to thank Dean Wyatte, Francisco Parada, and Ruj Akavipat for their help with data collection and analysis.

References

1. National Academy of Sciences: Strengthening Forensic Science in the United States: A Path Forward. Washington D.C.: The National Academies Press, (2009)
2. Maltoni, D., Maio, D., Jain, A., Prabhakar. S., Handbook of Fingerprint Recognition, Springer-Verlag New-York, (2003).
3. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K: FVC2000: Fingerprint Verification Competition, IEEE Trans. PAMI, 24, 3, 402-412 (2002).
4. Srihari, S.N., Cha, S., Arora, H., Lee, S.J.: Discriminability of Fingerprints of Twins. Journal of Forensic Identification 58(1), 109–127 (2008)
5. Dror, I. E., Charlton, D., Peron, A. E.: Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International, 156*(1), 74-78 (2006).
6. Su, C., Srihari, S.R.: Probability of Random Correspondence for Fingerpriints, Proc. Third International Workshop on Computational Forensics, The Hague, Netherlands, Springer (2009).
7. Tuyls, P., Akkermans, A. H. M., Kevenaar, T. A. M., Schrijen, G.-J., Bazen, A. M., Veldhuis, R. N. J.: Practical biometric authentication with template protection. In Proceedings of AVBPA, vol. 3546 of LNCS, pp. 436–446, Hilton Rye Town, USA (2005).
8. Ratha, N. K., Chikkerur, S., Connell, J. H., Bolle, R. M.: Generating cancelable fingerprint templates. IEEE Trans. pattern analysis and machine intelligence, 29(4):561–572 (2007).
9. Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Academic Radiology, 3*(2), 137-144.
10. Babcock, J.S., Pelz, J.: Building a lightweight eye tracking headgear, ETRA 2004: Eye Tracking Research and Applications Symposium, 109-113 (2004).
11. Jain, A.K., Dubes, R.C.: Algorithms for Clustering Data, Prentice Hall: New Jersey (1998).
12. Brown, P.F., Stephen A., Pietra, D., Pietra, V.J.D., Mercer, R.L: The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19:263–311 (1994).
13. Dyer, A. G., Found, B., & Rogers, D. (2006). Visual attention and expertise for forensic signature analysis. *Journal of Forensic Sciences, 51*(6), 1397-1404.
14. Duchowski, A. T. (2007). *Eye tracking methodology : theory and practice* (2nd ed.). London: Springer.