Biometric Identification of Identical Twins: A Survey

Kevin W. Bowyer and Patrick J. Flynn
Department of Computer Science and Engineering, University of Notre Dame
kwb@nd.edu, flynn@nd.edu

Abstract

The ability of biometric techniques to distinguish between identical twins is of interest for multiple reasons. The research literature touching on this topic is spread across a variety of areas. This survey pulls together the literature to date in this area, identifies available datasets for research, points out topics of uncertainty and suggests possible future research.

1. Introduction

Identical (monozygotic or MZ) twins are the result of the division of a single zygote (a fertilized egg) early in gestation. At the moment after splitting, the two fertilized eggs have identical DNA. MZ twins typically have strongly similar anatomy and facial appearance. Significant differences in appearance between MZ twins can develop due to both behavioral influences and epigenetic influences.

It was long assumed the MZ twins could not be distinguished by DNA matching; e.g., "By definition, identical twins cannot be distinguished based on DNA" [22]. However, in recent years it has become clear that DNA analysis does have the potential to distinguish between MZ twins [5,62]. A technique called "ultra-deep next-generation sequencing" was used to resolve a paternity test between MZ twins as possible parents [62]. The essential point is that random mutations accumulate over time and can be mapped by the new technique and used to distinguish MZ twins.

At least four reasons contribute to interest in using biometrics to distinguish between MZ twins [4,50]. One is that MZ twins are rare, even exotic, so that the topic naturally attracts curiosity. A second reason is that the legal system sometimes finds it is necessary to reliably distinguish between MZ twins. A third reason is that distinguishing between MZ twins is seen as a "hardest possible case" for biometrics, and so it becomes a strong argument for the more general validity of a biometric. Lastly, twin births are increasing in frequency [10], so that these issues are becoming more important over time. This paper summarizes the research literature to date on distinguishing between MZ twins using face, fingerprint, iris, speaker ID and handwriting. In the concluding section, we suggest open questions and possible future research topics.

2. Face Recognition

Facial appearance is a composite of skeletal, muscular, and dermal components. The first is genetically driven and the others are determined by genetics but modified by environment and behavior. Depending on the latter factors' differential influence, MZ twins may have relatively different facial appearances. Research reveals that face recognition systems are challenged by identical twins, and performance can degrade when probe and gallery samples are similarly posed and lit. As a mechanism for assessing face recognition algorithms, studying a data set composed of MZ twins, with appropriate structuring of the machine experiment (match, twin match, nonmatch), can help to characterize the marginal and distinctive effect of twin subjects on accuracy, and thus the margin for misclassification of "minimally different" individuals.

Sun *et al.* [58] published the first study on distinguishing between MZ twins using automatic face recognition. They experimented with data from 51 pairs of identical twins and 15 pairs of non-identical twins, acquired at a Beijing Twins Festival. They acquired face, fingerprint and iris data. The twins status was not recorded from the subjects, but was assigned retrospectively "based on observing whether the facial images of a set of twins were very similar or not" [58]. They conclude that the impostor distribution for matches between identical twins is more like a typical genuine distribution than like a typical impostor distribution.

Klare *et al.* [25] explored the ability of several types of local features to distinguish MZ twins. Their experiments used the ND-Twins dataset [47]. This work uses a taxonomy of facial features, in which Level 1 features are global descriptors such as eigenfaces, Level 2 features are coarsely localized or regional features such as SIFT descriptors, and Level 3 features are fine details such as facial marks and scars. Their results highlight the value of facial components and also the intrinsic challenges of identical twin discrimination.

Phillips *et al.* [47] published a study of MZ twins discrimination incorporating data captured at the Twins

Days 2009 and 2010 festivals, with data from over 100 sets of twins in each year. This dataset, the ND-Twins dataset, has been used by other researchers [25,30,31,35]. Phillips *et al.* present error rates and performance curves for multiple commercial face matching algorithms, distinguishing between twins under a variety of conditions, including images taken with a year of time lapse, and breakdown by age and gender. Later publications go into additional detail analyzing the accuracy of distinguishing between twins [44,48].

Mozaffari and Behravan [42] describe a dataset of twins face images scraped from the web, associated with different twins festivals around the world, including.images from 1,902 pairs of twins. Faces were detected with a Viola-Jones algorithm and resized to 256x256. Results of recognition using PCA, LDA and LBP recognition algorithms are given, with comparison to the same algorithms' performance on FERET and AR. Their database appears to be intended for distribution, but there is no information on how to obtain it.

Le *et al.* [30,31] develop an approach to distinguish between identical twins using features associated with face aging. The idea is that furrows that develop on the sides of the eyes ("crow's feet"), nose and mouth can be distinctive between twins. Their experiments use a Local Fisher Discriminant Analysis algorithm and the ND-Twins dataset [44,47]. They are able to obtain as high as 96% discrimination between twins, in the case of subjects making a smile expression and not wearing glasses.

Mahalingam and Ricanek [35] experiment with a component-based algorithm for distinguishing between twins, using both the ND-Twins dataset [44,47] and the CASIA dataset [58]. They report no significant difference in accuracy for distinguishing between male and female twins, and report that discrimination between twins becomes more accurate with increasing age.

Juefei-Xu and Savvides [23] propose a method to identify MZ twins based on an Augmented Linear Discriminant Analysis. Their experiments also use the ND-TWINS dataset [47].

Hu *et al.* [68] investigated the matching of identical twin faces using a data set containing 455 pairs of MZ twins extracted from Internet photos. Interestingly, the twin pairs were extracted from a single photo, which enforces some similarity in imaging conditions. This dataset is used with an experimental protocol where the classification decision is "same person" versus "twin sibling" and training was performed using either supervised or unsupervised approaches. The dataset is available.

Lamba *et al.* [29] consider a generalization of the identical twin problem, namely the task of distinguishing 'look-alikes' such as impersonators (in this context, twins are considered 'biological look-alikes' by the authors). To support the study, the authors created a database containing a balanced set of images of celebrities and either "genuine or intentional" look-alikes. This work included a human discrimination study employing 50 volunteer raters who were asked to distinguish self-matches from look-alike matches under a time constraint. Interestingly, the results from the human study suggested unbalanced performance on data from different genders and origin. This study included baseline automatic face recognition evaluations using seven different matchers. Algorithms trained on a general (non-lookalike) database and evaluated on a look-alike database exhibited performance very close to random. This poor performance motivated a new method for verification using an SVM with phase features extracted from face regions with a neural network. The proposed method performed at a level well above the random performance of the baseline methods.

Zhang *et al.* [64,65] report on a "talking profile" approach to distinguishing between twins. The idea is to use signatures derived from facial motion analysis in video. They report on experiments with video of 39 pairs of twins speaking. They work with six different types of face motion in a talking profile, and find that it is important to sample multiple types of face motion. They also report on experiments with non-twins datasets.

Srinivas *et al.* [55,56] report on experiments with a manually annotated set of facial marks to distinguish between twins. They find that "the position of certain facial marks appears to be similar for twins" [55] but that facial marks can still be useful in discriminating between twins.

There is also a substantial literature on how human observers distinguish between MZ twins based on facial appearance [2,15,36,37,51,57,63]. Due to space limits, we do not attempt to discuss this literature here.

3. Fingerprint and palmprint matching

In the context of fingerprint, "level one features" are the overall flow pattern of the ridges. Recognition keys on matching "level two" features, which are the "minutia points", the points where ridges terminate or bifurcate. Palmprint is a recent topic of substantial interest in the forensic and biometric research communities, and there has also been work on palmprint of MZ twins. The pattern of ridges and minutiae on fingers and palms appears to have a coarse genetic basis influencing print type but a significant random element as well; this tends to make the prints of MZ twins distinguishable.

The earliest serious study of fingerprint recognition and MZ twins that we know of is that by Lin *et al.* [33]. However, they note that Wilder [63] found unusual similarity in ridge patterns of identical twins in 1919, and that this similarity does not extend to the minutia pattern. These two observations by Wilder have been borne out through a number of modern studies.

Lin *et al.* [33] analyzed data from 108 pairs of MZ twins, 88 pairs of fraternal twins and 59 siblings of these twins. They categorized the ridge patterns of the prints into eight classes: arch, tented arch, ulnar loop, radial loop, plain whorl, central pocket loop, double loop and accidental. Because this work was done in the early days of fingerprint analysis becoming automated, there was still a significant manual element, and so "only 38 pairs of fingerprints were arbitrarily selected" for minutia-point analysis [33]. They found that (a) corresponding fingers of MZ twins had the same ridge pattern category over 87% of the time, (b) same-sex fraternal twins had the same ridge pattern category about 60% of the time, (c) same-sex non-twin siblings had the same ridge pattern category about 50% of the time, and (d) unrelated persons had the same ridge pattern category about 40%-50% of the time. Based on their minutia-point analysis of the 37 pairs of identical twins, they concluded that, "minutiae similarities between same-pattern and same-ridge-count fingerprints from MZ twins are significantly higher than minutiae similarities between random pairs of fingerprints" [33]. They hypothesized two factors that could potentially cause this effect – "Whether these similarities are associated with pattern and ridge count or result from a genetic relationship is not clear at this stage" [33].

Jain *et al.* [22] obtained a set of rolled fingerprint impressions for 100 pairs of identical twins. The fingerprints were originally recorded on paper and then scanned at 500 dpi. After discarding some data due to poor quality original images, they analyzed a dataset of index-finger fingerprint images from 94 pairs of identical twins. Their high-level results follow those of Lin *et al.* [33]. Considering five categories of ridge pattern – right loop, left loop, whorl, arch, and tented arch – they found that "there is only 0.2718 chance that two randomly chosen index fingers will have the same type which is much lower than the 0.775 chance that the fingerprints of two identical twins will have the same class label" [22]. However, they also found that identical twin, same-class fingerprints generally have different minutia patterns. Automated fingerprint matching can distinguish between identical twins with less than a factor of 5 increase of in the false match rate at the same false non-match rate.

Jain *et al.* make an interesting observation about degradation in fingerprint matching for twins and the fact that twins are more likely to have the same fingerprint class. They show that the degradation when matching non-twin fingerprints of the same class is similar to that when matching twins' fingerprints. Their conclusion is that "… the minutiae-based similarity of identical twin fingerprints, is of the same order as the similarity between unrelated people who have the same fingerprint class label. Hence, we conclude that the larger similarity observed in identical twins is due to the high class correlation in their fingerprint types" [22].

Han *et al.* [16] analyze live-scan, ten-print images from 51 pairs of identical twins, 15 pairs of fraternal twins and the non-twin relationships in 52 families. They find that the similarity of fingerprints of identical twins is higher than that of non-identical-twin siblings, and that the similarity of fingerprints of non-identical-twin siblings is higher than the similarity of fingerprints of parents and children.

Srihari *et al.* [53] analyze ten-print, live-scan rolled fingerprint images of 224 sets of MZ twins and 74 sets of fraternal twins (there is some uncertainty about these numbers, as different counts are given at different points in the paper). Identical twins were found to have the same class label for the corresponding finger 57% of the time, fraternal twins 39% of the time, and unrelated persons 32% of the time. Minutia-based fingerprint matching resulted in an EER for non-twins of 2.9%, and 6.2% for twins. One interesting element of their study that does not have a counterpart in previous work concerns fraternal versus identical twins in minutia-based matching – "The similarity of fingerprints of identical twins is the same as the similarity between fingerprints of fraternal twins. This indicates that genetic influence on the formation of minutiae in identical twins is the same as the influence among fraternal twins." [54]. This claim is more surprising in light of the fact that fraternal twins had much lower agreement in class labels (39%) than did identical twins (57%).

Liu and Srihari analyzed live-scan ten-print scans from 188 sets of MZ twins and 39 sets of fraternal twins [34]. They found that MZ twins had the same category of ridge pattern on corresponding fingers 66.7% of the time, and fraternal twins 46.4% of the time, and unrelated persons 26.5% of the time. In the minutia-based matching study, they found that the EERs for verification of identical twins' fingerprints, fraternal twins' fingerprints, and non-twin fingerprints were 5.09%, 4.88%, and 3.33% respectively (see Table 1 of [34]).

Sun *et al.* [58] analyzed a twins fingerprint dataset where the status of twins as fraternal or identical was not recorded at acquisition, and so the authors "… derived this information based on observing whether the facial images of a set of twins were very similar or not" [58]. Analyzing images from 51 pairs of identical twins and 15 pairs of fraternal twins, they reached the same general conclusions as found in previous works.

Tao *et al.* [59] analyze a dataset representing 83 pairs of identical twins. Their dataset is a superset of that analyzed by Sun *et al.* [58], with more persons and multiple (same-session) images per finger. One distinctive aspect of their results is that they find a quite small accuracy degradation for minutia-based matching of identical twins: "We also find that the automatic fingerprint verification matcher VeriFinger 6.1 SDK can distinguish between identical twins with a slightly lower accuracy than in non-twins (5.8333% vs. 5.3843%)". It would be interesting if this result was reproduced on a dataset with stronger ground-truth identification of identical versus fraternal twins.

Mishra *et al*. reported good accuracy on a proposed fingerprint matching algorithm [38] – "We tested the proposed method on thumbprint images of an identical twin pair generated by using Incept H3 T&A Terminal and fifty pairs of identical twins of FVC04, and FVC06 datasets. …" [38]. However, the FVC 04 [11] and FVC 06 [12] datasets are not known to contain any twins' fingerprints. Based on this discrepancy, we do not consider results from this paper.

In the only work known to us that considers palmprints of twins, Kong *et al*. [26, 27] analyze palmprints from 53 pairs of identical twins. At a matching threshold that results in 97% genuine acceptance and 0.000044% false accept rate (FAR) for unrelated persons, the FAR for identical twin imposters is 0.02%. They also perform experiments to investigate elements of the palmprint that are most genetically related, and they consider the relation between left and right palmprints.

Two conclusions seem clear from the substantial body of knowledge around the topic of fingerprint recognition of identical twins. One, identical twins have a much higher frequency of the same class of ridge pattern on corresponding fingers than do unrelated persons or fraternal twins. Two, state-of-the-art minutiae-based fingerprint recognition algorithms can distinguish between identical twins with good accuracy, though performance is reduced compared to that obtained when recognizing unrelated persons.

4. Iris recognition

The formation and layout of tissue and tissue irregularities in the iris is driven by a gestational process with significant randomness in the finer details. The limited research to date suggests that iris recognition is not noticeably challenged by distinguishing MZ twins. Daugman first observed this on images from three sets of twins [7] – "A set of six pairwise comparisons among the eyes of actual monozygotic twins also yielded a result (mean HD = 0.507) expected for unrelated eyes. It appears that the phenotypic random patterns visible in the human iris are almost entirely epigenetic".

Sun *et al*. [59] experimented with iris images from 51 sets of MZ twins, using a different matching algorithm than Daugman. They conclude that "[t]he identical twin impostor distribution is very similar to the general impostor distribution. However, the peaks that are present in the identical twin impostor distribution tail *may* indicate that the irises of identical twins have some correlation." (italics added)

Hollingsworth *et al*. [17,18] observed a similar result for 76 pairs of self-reported MZ twins. They found that the peaks of the impostor histogram for unrelated persons and for MZ twins occur at essentially the same fractional HD. They report that "a Komogorov-Smirnov test comparing the two histograms finds small but statistically significant evidence to reject the null hypothesis that the two sets of scores are from the same distribution" [18]. However, they also found that when images that resulted in segmentation problems were removed from the data, there was no longer a statistically significant difference [18].

Hollingsworth *et al*. [18] also studied human matching of iris images. In this context, it is important to realize that humans organize and interpret iris texture differently from algorithms. The Daugman-style iris code represents iris texture as the result of a set of coarsened Gabor filter phase responses sampling the iris at uniform increments of angular and radial distance. Humans interpret iris texture more holistically in terms of the overall texture pattern. Similarities in overall pattern do not necessarily translate to similarities in iris code.

The dataset used by Hollingsworth *et al*. [17,18] is available, and the dataset used by Sun *et al*. [59] may also be available. However, there is very little work involving twins and iris recognition.

5. Speaker Recognition

The gross anatomy of the larynx and vocal cords is determined genetically, and so MZ twins tend to have similar speaking fundamental frequencies. Speech formation similarities may have both genetic and developmental components due to proximity during childhood. Thus, in identical voice capture environments we would anticipate that MZ twins would be more challenging to distinguish than unrelated individuals [70]. The limited research conducted to date suggests that the error rate for automatic speaker recognition in distinguishing MZ twins is higher than the error rate for the general population [1].

Speaker recognition has a rich research history [9]. The earliest work known to us using automated speaker ID on more than one set of MZ twins is [6]. They used data from nine sets of male, Hebrew-speaking, MZ twins. For each person, the data included three sets of two approximately 20-second segments of reading different prescribed text. Recordings for a given person were acquired in a single session. Six features were computed from the speech samples, a minimum distance classifier used, and 100% correct identification of twins was possible. They also report that, "[p]reliminary tests with human expert listeners, in a closed experiment of five twins, yielded 100% identification" [6].

Homayounpour and Chollet [19,20] report on experiments in a text-independent speaker scenario, with the person speaking over a telephone line. Each person was recorded on three different occasions, with at least one week between sessions, reading three different one-page texts. The dataset includes nine sets of MZ twins, two sets of fraternal twins and 27 other speakers. Several approaches to automatic speaker identification were experimented with, and an EER as low as 22% was achieved for distinguishing between MZ twins [19]. They also report on several different types of tests [19]. One involves persons listening to two six-second samples and deciding if they are from the same person or not. Testing on sample pairs that did not come from twins or siblings, listeners with some familiarity with twins had a mean error rate of 15.6% and listeners without familiarity with distinguishing between twins had a mean error rate of 15.7%. This difference could not be considered meaningful. Testing on sample pairs that come from twins, listeners who do not have familiarity with distinguishing twins had a mean error rate of 26.1% whereas listeners who do have familiarity with distinguishing twins had a mean error rate of 18.3% for distinguishing between the twins with which they are familiar. Lastly, in another human-listener experiment, family members of MZ twins listened to one six-second audio and were asked to identify which of the twins was the speaker, with a resulting EER of 8.2%.

Patil and Basu [45] report on speaker ID experiments involving 17 pairs of twins from five dialectical zones in the Maharashtra region of India. The audio for each person involved 10 repetitions of answers to scripted prompts and some free-response speech. The dataset was organized into training segments of duration 30, 60 and 90 seconds, and testing segments of 1, 3, 5, 10, 12 and 15 seconds. Matching based on linear predictive cepstral coding (LPCC) and on Teager energy operator Mel frequency cepstral coefficients (MFCC) were compared. The MFCC approach generally results in higher accuracy, being able to distinguish between MZ twins with up to 91% accuracy for the longest-duration test samples (15 seconds). In another experiment, they [46] explore the more difficult problem of distinguishing between twins when the training sample is recorded speaking one language and the testing sample is recorded speaking a different language. In this case, the two languages are Marathi and Hindi. The accuracy for distinguishing twins in this scenario reaches 60 to 65%, far below when the train and test samples are in the same language. Their results also suggest that the problem is asymmetric, in that training with Hindi samples and testing with Marathi samples results in higher accuracy than the reverse.

Revathi *et al.* [49] experiment with a variety of approaches to speaker recognition, including LPCC, line spectral frequency (LSF) and Discrete Cosine Transformed Cepstrum. They used a dataset from just four pairs of MZ twins. They find that LSF achieves the highest accuracy (81%) in correctly distinguishing between twins.

Ariyaeeinia *et al.* [1] investigate the ability of speaker recognition technology to differentiate between MZ twins using data from 49 identical twin pairs. Each person contributed two audio recordings, one of about 60 seconds duration and one of about 5 seconds. For each person, the reference or enrollment model is created using the first 30 seconds of the long sample. The remainder of the long sample is then used as a test sample and the reading of the date of birth is used as a "short" test sample. With unconstrained cohort normalization, the EER for the short test samples was 1% and for the long samples was 0%.

Künzel reports on an experiment with 35 pairs of MZ twins [28]. The audio recording for each individual consists of two samples. One is a reading of the German version of the poem "The North Wind & the Sun", which is 35 – 45 seconds in length. The other is the person talking spontaneously about their life, limited to about 80 seconds in length. Künzel analyzes the data from the male and female twin pairs separately. For each of the males and the females, there are two experiments. For the male twins in Experiment A, the system was able to distinguish between twins with no errors. For the male twins in Experiment B, the equal error rate for distinguishing between twins was 11%. For female twins in Experiment A, the equal error rate for distinguishing between twins was 19%. For female twins in Experiment B, the equal error rate for distinguishing between twins was 48%.

The most striking elements of Künzel's study are the difference in accuracy for distinguishing between male and female twins, and the large difference in results based on the length and type of the audio samples. Künzel also provides a useful summary of relevant studies on speech and language acquisition that have been done with twins, and of studies on physiological and anatomical structures.

6. Handwriting

The limited research to date suggests that there may be a performance degradation for MZ twins in handwriting-based authentication. Various studies looked at the similarity of twins' handwriting before automated analysis was introduced [3, 8, 13, 52]. The first study using automated handwriting analysis for verification of twins' identity appears to be Srihari *et al.* [54]. The operating parameters of their CEDAR-FOX system were tuned on handwriting samples from 1,000 persons not involved in the twins study. A sample is an intentionally-designed English text called the "CEDAR letter". The letter is designed so that the "top half" and "bottom half" can be used as separate samples. Samples were acquired for 31 pairs of self-identified MZ twins, 169 pairs of fraternal twins, and six pairs who were uncertain. (At different places in the paper, the numbers of MZ and fraternal twins are reversed).

The error rate in distinguishing MZ twins using automatic handwriting analysis was 17% when matching samples with the same content (e.g., top half of CEDAR letter to top half) and 24% when matching samples with different content (e.g., top half to bottom half). For fraternal twins, the error rate was 11% in both cases. This compares to an error rate for non-twins of 3.15% and 4.24%, respectively. Thus the error rate for fraternal twins is more than twice that of non-twins, and the error rate for MZ twins is about twice that of fraternal twins.

Mohammed and Shamsuddin [39] report on experiments with 390 handwriting samples from 13 pairs of MZ twins, giving 15 handwriting samples for each person. The length and content of the writing samples is not described. Dividing the dataset into 60% to train the parameters and 40% for testing, they were able to achieve 100% recognition. However, it appears that the train and test division is not subject-disjoint. That is, subjects whose handwriting samples appear in the test set likely also have samples in the training data, and so the reported accuracy is likely "optimisitic". Thus the accuracy level reported cannot fairly be compared to that in other works.

7. Conclusions

Table 1 summarizes the MZ twins datasets that appear to be available to the research community, with a selected reference that uses the dataset, the type of biometric, the number of MZ twins, and an indication of the ground truth of MZ status. One major need to advance research in this area is a large, readily available multi-modal dataset that includes both MZ and fraternal twins. Ideally, the "ground truth" for MZ / fraternal would be established in some formal manner.

Table 1. Summary of available twins datasets.

ref	biometric	number of twin pairs	MZ ground truth
[44]	face	126 + 120	self-reported
[58]	face, iris, fingerprint	51	assigned by looking at face images
[18]	iris	76	self-reported
[61]	3D face	107	self-reported
[68]	face	455	?(scraped from web)
[42]	face	1902	?(scraped from web)

Distinguishing between MZ twins using face images has received more attention than other biometric modalities. There are two face image datasets that have been used by multiple research groups, ND-Twins-2009-2010 [66] and CASIA-Twins [67]. Avenues for face matching with increased accuracy for MZ twins include the use of facial marks [56] and the use of regions of the face that develop age-related characteristics [30]. In the special case of "mirror identical" twins [14], asymmetry analysis may be especially advantageous.

There appears to be disagreement in the literature on the degree of similarity between MZ versus fraternal twins' fingerprints. Srihari *et al*. state that "[t]he similarity of fingerprints of MZ twins is the same as the similarity between fingerprints of fraternal twins" [34], even though they and others find that fraternal twins have a much lower similarity in class of ridge pattern than do MZ twins.

Other issues in fingerprint revolve around the lack of any large and standard dataset. All MZ twins studies reviewed involve relatively small numbers of pairs of MZ twins. Thus it is possible that there are some rare cases that simply have not been seen in any of the studies to date. Also, it is not clear that any of the twins fingerprint databases in the works reviewed are currently available to researchers. There is only one instance of two different research groups studying the same dataset [58, 59], and that is only a partially overlapping dataset.

In some respects, iris recognition has the most unusual results for distinguishing between MZ twins. There is agreement that, with respect to the industry-standard, Daugman-style approach to iris matching, MZ twins' iris codes are as different as those of unrelated persons. This is a stronger statement than can be made about other modalities. At the same time, there is also evidence that human observers can classify pairs of iris images as (a) belonging to twins or (b) belonging to unrelated persons at accuracy far above chance [17, 18].

Several points are worth making about the body of work in distinguishing between twins using speaker identification. All of the studies to date involve relatively small numbers of pairs of MZ twins, less than 50 pairs. Almost no attention has been paid specifically to fraternal twins. The effect of the twins' age on the ability to distinguish between them has not been investigated. The possible gender-based difference in accuracy in distinguishing between twin pairs has been investigated by just one study, with rather surprising results.

There has been relatively little work in using handwriting to distinguish between MZ twins. The more optimistic accuracy estimates reported are likely not from subject-disjoint train and test experiments.

Studies involving MZ twins have been done with other biometric modalities than those discussed above, including ear [43], gait [40,41] and 3-D face scans [32,60,61].

References

[1] A. Ariyaeeinia, C. Morrison, A. Malegaonkar and S. Black, A test of the effectiveness of speaker verification for differentiating between identical twins, *Science & Justice* 48 (4), 182-186, 2008.

[2] S. Biswas, K.W. Bowyer and P.J. Flynn, A study of face recognition of identical twins by humans, *IEEE Int.*

Workshop on Info. Forensics and Security (WIFS), 1-6, 2011.

[3] D. Boot, An investigation into the degree of similarity in the handwriting of identical and fraternal twins in New Zealand, *J. Am. Soc. Quest. Doc. Exam.* 1(2):70–81, 1998.

[4] K. W. Bowyer, What surprises do identical twins have for identity science?, *Computer* 44(7):100-102, 2011.

[5] B. Budowle, Editors' pick: Molecular genetic investigative leads to differentiate monozygotic twins, *Investigative Genetics* 5:11, 2014.

[6] A. Cohen and T. Vaich, On the Identification of Twins By Their Voices, *ESCA Workshop on Automatic Speaker Recognition, Identification and Verification*, 213-216, 1994. http://www.isca-speech.org/archive

[7] J. Daugman, How iris recognition works, *IEEE Trans. on Circuits and Systems for Video Technology* 14 (1), 21-30, 2004.

[8] T. Dziedzic, E. Fabianska, Z. Toeplitz, Handwriting of Monozygotic and Dizygotic Twins, *Problems of Forensic Sciences* LXIX, 30–36, 2007.

[9] A. Fazel and S. Chakrabartty, An Overview of Statistical Pattern Recognition Techniques for Speaker Verification, *IEEE Circuits and Systems Magazine* 11 (2), 62 – 81, 2011.

[10] D.B. Fell and K. Joseph, Temporal trends in the frequency of twins and higher-order multiple births in Canada and the US, *BMC Pregnancy and Childbirth* 12:103, 2012.

[11] FVC2004: the Third International Fingerprint Verification Competition, http://bias.csr.unibo.it/fvc2004/.

[12] FVC2006: the Fourth International Fingerprint Verification Competition, http://bias.csr.unibo.it/fvc2006/

[13] D. J. Gamble, The handwriting of identical twins. *Can. Soc. Forensic Sci. J.* 1980;13:11–30.

[14] L. Gedda, G. Brenci A. Franceschetti, C. Talone and R. Ziparo, Study of mirror imaging in twins', *Progress in Clinical and Biological Research* 69A, 167-168, 1981.

[15] B. Guyuron, D.J. Rowe, A.B. Weinfeld, Y. Eshraghi, A. Fathi and S. Iamphongsai, Factors contributing to the facial aging of identical twins. *Plast. Reconstr. Surg.* 123(4):1321-31, 2009.

[16] Y. Han, C. Ryu, J. Moon, H. Kim, H. Choi, A Study on Evaluating the Uniqueness of Fingerprints Using Statistical Analysis, in *Information Security and Cryptology*, LNCS Volume #3506, 467-477, 2005.

[17] K. Hollingsworth, K. Bowyer, and P. Flynn, Similarity of iris texture between identical twins, *IEEE Conf. on Computer Vision and Pattern Recognition Workshops*, pp. 22–29, 2010.

[18] K. Hollingsworth, K.W. Bowyer, S. Lagree, S.P. Fenker and P.J. Flynn. Genetically Identical Irises Have Texture Similarity That Is Not Detected By Iris Biometrics, *Computer Vision and Image Understanding* 115, Pages 1493–1502, 2011.

[19] M. M. Homayounpour and G. Chollet, Discrimination of voices of twins and siblings for speaker verification, in *4th European Conf. Speech Communication and Technology* (EUROSPEECH '95), Madrid, Spain, pp. 345–348, 1995.

[20] M. M. Homayounpour and G. Chollet, A Study of Intra- and Inter-speaker Variability In Voices of Twins for Speaker Verification, *XlII Int'l Congress of Phonetic Sciences* (ICPhS 95), Stockholm, Sweden, pp. 298–301, 1995.

[21] A. K. Jain, S. Prabhakar, S. Pankanti, Twin Test: On Discriminability of Fingerprints, Audio- and Video-Based Biometric Person Authentication, Springer LNCS #2091, 211-217, 2001.

[22] A. K. Jain, S. Prabhakar, S. Pankanti, On the similarity of identical twin fingerprints, *Pattern Recognition* 35:2653 – 2663, 2002.

[23] F. Juefei-Xu and M. Savvides, An Augmented Linear Discriminant Analysis Approach for Identifying Identical Twins with the Aid of Facial Asymmetry Features, *IEEE Conference on Computer Vision and Pattern Recognition Workshops* (CVPRW), 56-63, 2013.

[24] C.H. Kau, A Zhurov, R Bibb, L Hunter and S Richmond, The investigation of the changing facial appearance of identical twins employing a three-dimensional laser imaging system, *Orthod. Craniofac. Res.* 8(2):85-90, 2005.

[25] B. Klare, A. A. Paulino, and A. K. Jain, Analysis of Facial Features in Identical Twins, *Int. Joint Conference on Biometrics* (IJCB), 1-8, 2011.

[26] A. W. Kong, D. Zhang, and G. Lu, A study of identical twins palmprints for personal authentication, *Advances in Biometrics,* Springer, 2005.

[27] A. W. Kong, D. Zhang, and G. Lu, A study of identical twins palmprints for personal verification, *Pattern Recognition*, 39(11), 2149–2156, 2006.

[28] J. Künzel, Automatic speaker recognition of identical twins, *Int. J. of Speech, Language and the Law*, 17.2, 251–277, 2010.

[29] H. Lamba, A. Sarkar, M. Vatsa, R. Singh and A. Noore, Face Recognition for Look-Alikes: A Preliminary Study, *Int. Joint Conference on Biometrics*, 1-6, 2011.

[30] T. H. N. Le, K. Luu, K. Seshadri and M. Savvides, A Facial Aging Approach to Identification of Identical Twins, *IEEE Fifth Int. Conference on Biometrics: Theory, Applications and Systems* (BTAS), 23-27, 2012.

[31] T. H. N. Le, K. Seshadri, K. Luu and M. Savvides, Facial Aging and Asymmetry Decomposition Based Approaches to Identification of Twins, *Pattern Recognition* 48(12):3843-3856, 2015.

[32] H. Li, D. Huang, L. Chen and Y. Wang, A group of facial normal descriptors for recognizing 3D identical twins, *IEEE Fifth Int. Conference on Biometrics: Theory, Applications and Systems* (BTAS), 23-27, 271 – 277, 2012.

[33] C. H. Lin, J. H. Liu, J. W. Osterburg and J. D. Nicol, Fingerprint Comparison. I. Similarity of Fingerprints, *J. of Forensic Sciences* 27 (2), 290-304, 1982.

[34] Y. Liu, S. N. Srihari, A Computational Discriminability Analysis on Twin Fingerprints, in Computational Forensics: *Third Int. Workshop on Computational Forensics* (IWCF 2009), Springer, 43-54, 2009.

[35] G. Mahalingam and K. Ricanek, Investigating the Effects of Gender and Age group based Differences in Identical Twins, *Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics* (NCVPRIPG), 18-21, 2013.

[36] M. Martini, I. Bufalari, M. Antonietta Stazi, S. Maria Aglioti, Is That Me or My Twin? Lack of Self-Face Recognition Advantage in Identical Twins, *PLoS One* 10(4):e0120900, 2015.

[37] E. McKone, N. Kanwisher and B. C. Duchaine, Can generic expertise explain special processing for faces?, *Trends in Cognitive Sciences* 11(1):8-15, 2006.
[38] K. Nath Mishra, P. C. Srivastava, A. Agrawal, V. Tripathi and R. Garg, Minutiae Distances and Orientation Fields Based Thumbprint Identification of Identical Twins, *Int. J. Image, Graphics and Signal Processing* 5(2), 51-59, 2013.
[39] B.O. Mohammed and S.M. Shamsuddin, Improvement in twins handwriting identification with invariants discretization, *EURASIP J. on Advances in Signal Processing* 2012:48, 2012.
[40] W.-N. Mohd-Isa, J. Abdullah and C. Eswaran, Distinguishing Twins by Gait via Jackknife-Like Validation in Classification Analysis, *First Int. Conf. on Advanced Data and Information Engineering (DaEng-2013) Lecture Notes in Electrical Engineering* Volume 285, 301-308, 2014.
[41] W.-N. Mohd-Isa, J. Abdullah and C. Eswaran, Classification of Gait Biometric on Identical Twins, *J. of Advanced Computer Science and Technology Research* 2 (4), 166-175, 2012.
[42] S. Mozaffari and H. Behravan, Twins Facial Similarity Impact on Conventional Face Recognition Systems, *19th Iranian Conf. on Electrical Engineering* (ICEE), 1-6, 2011.
[43] H. Nejati, L. Zhang, T. Sim, E. Martinez-Marroquin and G. Dong, Wonder Ears: Identification of Identical Twins from Ear Images, *21st Int. Conf. on Pattern Recognition,* 1201-1204, 2012.
[44] J. R. Paone, P. J. Flynn, P. J. Philips, K. W. Bowyer, R. W. Vorder Bruegge, P. J. Grother, G. W. Quinn, M. T. Pruitt, and J. M. Grant, Double Trouble: Differentiating Identical Twins by Face Recognition, *IEEE Trans. on Information Forensics and Security* 9(2), 285-295, 2014.
[45] H. A. Patil and T. K. Basu, Detection of Bilingual Twins By Teager Energy Based Features, *Int. Conf. on Signal Processing and Communications* (SPCOM), 32-36, 2004.
[46] H. A. Patil and T. K. Basu, Teager Energy Mel Cepstrum for Identification of Twins in Marathi, *IEEE India Annual Conference* 2004 (IndiCon 2004), 58-61, 2004.
[47] P. J. Phillips, P. J. Flynn, K. W. Bowyer, R. W. Vorder Bruegge, P. J. Grother, G. W. Quinn, M. Pruitt, Distinguishing Identical Twins by Face Recognition, *IEEE Int. Conf. on Automatic Face & Gesture Recognition and Workshops* (FG), 185 – 192, 2011.
[48] M.T. Pruitt, J.M. Grant, J.R. Paone, P.J. Flynn, Facial recognition of identical twins, *Int. Joint Conf. on Biometrics*, 1-8, 2011.
[49] A. Revathi, R. Chinnadurai and Y. Venkataramani, Effectiveness of LP derived features and DCTC in twins identification-Iterative speaker clustering, *Int. Conf. on Computational Intelligence and Multimedia Applications*, 535-539, 2007.
[50] K. Ricanek and G. Mahalingam, Biometrically, How Identical Are Identical Twins?, *IEEE Computer* 46(3):94-96, 2013.
[51] L. Saether and B. Laeng, On facial expertise: processing strategies of twins' parents. *Perception* 37: 1227–1240, 2008.
[52] R. Saudek, The Handwriting of Identical Twins Reared Apart, *Character and Personality* 1(4):268-285, 1933.
[53] S. N., Srihari, H. Srinivasan and G. Fang, Discriminability of fingerprints of twins, *J. of Forensic Identification* 58(1):109-127, 2008.
[54] S. Srihari, C. Huang and H. Srinivasan, On the Discriminability of the Handwriting of Twins, *J. Forensic Science* 53(2):430-446, 2008.
[55] N. Srinivas, G. Aggarwal, P. J Flynn and R. W. Vorder Bruegge, Facial Marks as Biometric Signatures to Distinguish between Identical Twins, *Computer Vision and Pattern Recognition Workshops* (CVPRW), 106-113, 2011.
[56] N. Srinivas, G. Aggarwal, P. J Flynn and R. W. Vorder Bruegge, Analysis of Facial Marks to Distinguish Between Identical Twins, *IEEE Trans. on Information Forensics and Security* 7(5):1536 – 1550, 2012.
[57] S. V. Stevenage, Which twin are you? A demonstration of induced categorical perception of identical twin faces, *British J. of Psychology* 89(1):39-58, 1998.
[58] Z. Sun, A. Paulino, J. Feng, Z. Chai, T. Tan, and A. K. Jain. A study of multibiometric traits of identical twins. *Proc. SPIE 7667: Biometric Technology for Human Identification VII, 76670T, 2010*.
[59] X. Tao, X. Chen., X. Yang, J. Tian, Fingerprint Recognition with Identical Twin Fingerprints, *PLoS ONE* 7(4): e35704, 2012.
[60] V. Vijayan, K. Bowyer and P. Flynn, 3D Twins and Expression Challenge, *ICCV Workshops: Workshop on Benchmarking Facial Image Analysis Technologies*, 2100-2105, 2011.
[61] V. Vijayan, K. Bowyer, P. Flynn, D. Huang, L. Chen, M. Hansen, O. Ocegueda, S. Shah and I. Kakadiaris, Twins 3D Face Recognition Challenge, *Int'l Joint Conf. on Biometrics*, 1-7, 2011.
[62] J. Weber-Lehmann, E. Schilling, G. Gradl, D. C. Richter, J. Wiehler, R. Burkhard, Finding the needle in the haystack: Differentiating ''identical'' twins in paternity testing and forensics by ultra-deep next generation sequencing, *Forensic Science International: Genetics* 9, 42–46, 2014.
[63] H. H Wilder, Physical Correspondence in Two Sets of Duplicate Twins, *J. of Heredity* 10(9):410-422, 1919.
[64] L. Zhang, N. Ye, W. Martinez Marroquin, D. Guo, T. Sim, New hope for recognizing twins by using facial motion, *IEEE Workshop on Applications of Computer Vision* (WACV), 209-214, 2012.
[65] L. Zhang, K. Ma, H. Nejati, L. Foo, T. Sim and D. Guo, A talking profile to distinguish identical twins, *Image and Vision Computing* 32(10):771–778, 2014.
[66] http://www.nd.edu/~cvrl
[67] http://biometrics.idealtest.org
[68] J. Hu, L. Lu, Y.-P. Yan, Fine-Grained Face Verification: Dataset and Baseline Results, *Int. Conf. on Biometrics*, 79-84, 2015.
[69] M. Kucken and A.C. Newell, Fingerprint formation, *J. Theor. Biol.* 235(2005):71-83
[70] W.D. Van Gysel, J. Vercammen and F. Debruyne, Voice similarity in identical twins, *Acta Otorhinolaryngol Belg.* 2001;55(1):49-55.