VOL. 10, NO. 3, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

AN ANALYSIS ON GAIT RECOGNITION OF TWINS

Wan-Noorshahida Mohd-Isa, Junaidi Abdullah, Chikkanan Eswaran, Amalina Ibrahim
Faculty of Computing and Informatics, Multimedia University (MMU), Cyberjaya, Selangor, Malaysia
E-Mail: noorsha@mmu.edu.my

ABSTRACT

The aim of this paper is to investigate the viability to recognise an individual by gait where this individual is having a twin sibling where he/she is sharing similar facial features, hence may render a challenge to face recognition systems. To facilitate the investigation, supervised classification experiments are devised to compare the gait of both twins and non-twins siblings and additionally compare both of these with the inclusion of an early SOTON data set to synthesise a small population data set. An important component of this investigation is to firstly develop data sets consisting of twins and non-twins siblings, which are video-based data sets. After extraction of lower limbs kinematics signals from the videos, to learn the recognition viability performance, this paper performs classification with a leave-one-out cross validation on the data sets. The best correct classification rates using the proposed descriptor is 95%, 76%, 82%, and 74% respectively for data sets containing of only non-twins siblings, only twins, both non-twins and twins, and both non-twins and twins including SOTON data set.

Keywords: gait biometric, twins siblings, classification, small sample validation.

INTRODUCTION

Gait is a biometric based on the way human walks. It is believed to be unique to every person (Murray, 1967) (Stevenage et al., 1999) and every person is said to have his/her own idiosyncratic way of walking. Gait and face biometrics are biometrics suitable for use in surveillance systems. These are non-invasive biometrics via computer vision technologies that do not require an individual to even realise that he/she is being recorded.

However, for a security system that employs face biometric, the issue of facial similarity still poses a risk and thus interest in research on face biometric for recognising between twins has increased (Rychlik et al., 2008), (Park & Jain, 2010), (Sun et al., 2010), (Phillips et al., 2011), (Strikland et al., 2011), (Jain et al., 2012), (Srinivas et al., 2012), (Dinakardas et al., 2013). One central issue in these twins biometric literatures is to distinguish one twin from another via the facial features.

So far, no research on the gait biometric recognition of twins has been found except the ones by us. Similarly, there is no preliminary research being done to analyse gait among siblings let alone a gait biometric research involving twins in a sample population. This paper aims to research this issue since it is unknown if an individual with a twin can be recognised based on his/her gait via computer vision.

This work adopts a data-driven machine learning approach via classification to investigate its data set. Thus, the methodology section that follows is divided into subsections that describe the data sets collection, feature extraction and selection, and finally classification. After that this paper presents and discusses its results and finally presents its conclusion.

METHODOLOGY

Data sets collection

In this work, there are two supervised gait data sets that are investigated; one is a 12 pairs of twins data set (TW) and two is a 10 pairs of non-twins siblings data set (NT). These are volunteered young adult subjects, age 16-28, declared as healthy without known gait problem. To reduce the effect of body weights, the subjects in the data sets are having almost similar height-to-weight ratio. This carefully establishes a better measure than just relying on weight measurement alone.

Following the setup of early SOTON database (Nixon et al., 2006), which has successfully been used in previous gait biometric research on unique individual recognition, each subject in each data set is to have recorded videos of them walking from left-to-right and right-to-left. This is to ensure data are invariant to walking leg and body sides. Subjects are to walk their usual walk and at their normal pace to ensure invariability to walking speed. At least three steps are recorded in each video. There are at least two videos per subject recorded for the left-to-right and right-to-left walking. Each video is then converted to image-frames.

Additionally, two extra data sets are created by combining both the TW and NT data sets, known as siblings data set (SB) and by combining the SB data set with an early SOTON data set of 10 unique individuals (AP) that has been used here (Mohd-Isa, 2005). The early SOTON data set is from the *Southampton Human ID at a Distance* project, which was supported by DARPA. This paper has used the *Large DB* data set that can be found here:
http://www.gait.ecs.soton.ac.uk/database/large_db.php3.
The data sets dictionary is as described in Table-1.

VOL. 10, NO. 3, FEBRUARY 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

Table-1. Data sets dictionary of this paper.

Acronyms	Description
TW	Twins of 24 individuals (12 pairs)
NT	Non-twins siblings of 20 individuals (10 pairs)
SB	Combination of TW and NT data sets
AP	Combination of SB data set and 10 unique individuals from early SOTON database.

The inclusion of this early SOTON data set as the AP data set can be equated to synthesizing a real-world population that consists of unique individuals and twins. This strategy is chosen as an approach to validate the small sample size data, which is due to the scarcity of twins data.

Feature extraction

The lower limb orientation has been proposed as the descriptors to represent the kinematics of gait for the TW and NT video data sets. This is because according to research on motion perception, it had been shown that human observers perceive gait by relying on features from lower limbs motion (Todd, 1983). Other similar studies in similar domain also support the use of the proposed descriptor (Giese et al., 2008), (Roether et al., 2009) (Thurman et al., 2010). Early research on recognition of individual by gait has been successful as well (Huang, 2001), (Yam et al., 2004), (Bouchrika & Nixon, 2008).

To extract this kinematics attribute from a video, raw data in the form of coordinates are gathered and tracked at each coloured image-frame. The coordinates are the manually located positions of the hip (S_H), knee (S_K), and ankle (S_A) on the lower limbs of a human object in an image-frame. By triangle trigonometry, the kinematics measurements by the thigh (α) and lower leg(β) orientation angles can be calculated between the coordinates at the hip (S_H), knee (S_K), and ankle (S_A) locations at each image-frame as illustrated in Figure-1.

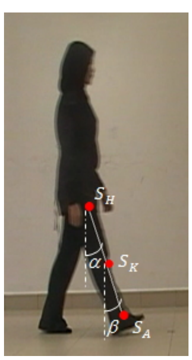

Figure-1. Coordinates (S_H, S_K, S_A) gathering and kinematics measurements (α, β) at an image-frame of a video.

Once all α and β values are gathered from all relevant image-frames in a video (as illustrated in Figure-2), they form two representative signals (α and β) for each video containing walking individual,

$$\boldsymbol{\alpha} = \{\alpha_{[n]}\} \qquad n \in \mathbf{Z}^{+} \qquad \alpha \in \left[-\pi/2, \pi/2\right] \tag{1}$$

$$\boldsymbol{\beta} = \{\beta_{[n]}\} \qquad n \in \mathbf{Z}^{+} \qquad \beta \in \left[-\pi/2, \pi/2\right] \tag{2}$$

n = image-frame number

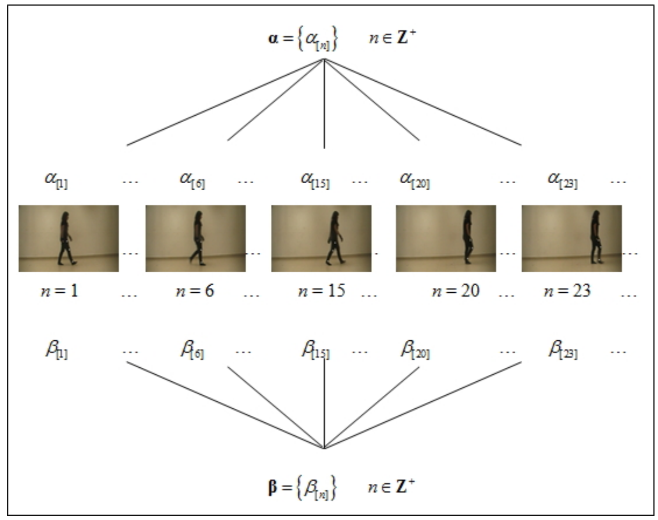

Figure-2. Collection of $\alpha_{[n]}$ and $\beta_{[n]}$ measurements at each image-frame form signals of α and β for a video.

Feature selection

Different person walks at different speed and thus the number of image-frames produced by each video varies. There is a need to select features to make further analysis invariant to the effect of speed. According to Murray (1967) gait is periodic and can be defined by a gait cycle as shown in Figure-3.

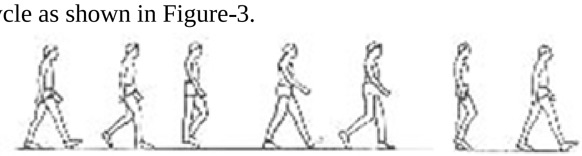

Figure-3. A complete gait cycle that starts from a heel strike and ends at a heel strike of the same leg (Murray, 1967).

Our signal values vary between positive angles (when a person's leg moves forward) and negative angles (when a person's leg moves to the rear of the body) continuously in a continuous video. There will be zero values that each signal crosses for a continuous video. These zero values represent the position when the leg is straight. Thus, we define our gait cycle to be between two consecutives zero values of the same leg as in Figure-4. Figure-4 is an extended version of Figure-3 with an addition of arrows to point to the position of zero values.

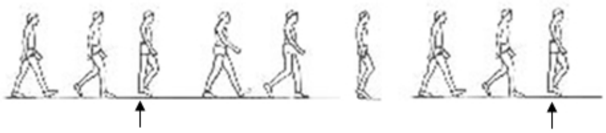

Figure-4. Our gait cycle that is defined between two zeros.

To deal with the unequal lengths of signal data, we resample the signals lengths to 30 data points by firstly interpolating each signal data and then resampling by a piece-wise cubic spline. The number 30 is chosen since it is the average lengths of all signals that can make up one gait cycle. Both processes of interpolation and resampling are done via the Support Vector Regression (SVR) framework. The mathematical formulation for β is as follows (the mathematical formulation for α is similar):

$$\hat{\beta}(t)=\sum_{j=1}^{n}\gamma_j K(t_{[j]},t)+\delta \qquad (3)$$

$\hat{\beta}(t)$ = SVR function estimate for $\{\beta_{[n]}\}$
$K(t_{[j]},t)$ = kernel function as in Equation 5
γ_j = support vectors
δ = bias term

γ and δ are found by minimising the regularised risk R,

$$R=\sum_{j=1}^{n}Loss\left(\beta(t),\hat{\beta}(t)\right)+\zeta\|\gamma\|^2 \qquad (4)$$

ζ = regularisation parameter

While the piece-wise cubic polynomial spline kernel is,

$$K(t_{[n]},t)=1+t_{[n]}\cdot t+t_{[n]}t\min(t_{[n]},t) -\frac{(t_{[n]}+t)}{2}\left(\min(t_{[n]},t)\right)^2+\frac{1}{3}\left(\min(t_{[n]},t)\right)^3 \qquad (5)$$

The resultant signals of new α and β are ilustrated in Figure-5 and Figure-6.

To sum up, for every video, there are two representative signals. Each person has four videos, and thus has in total eight representative signals as his/her kinematics descriptor. Additionally, a combined feature is formed that concatenates both α and β and is denoted as signal **w**. The **w** signal adds for more signals to each video. Thus there are in total 12 representatives signals.

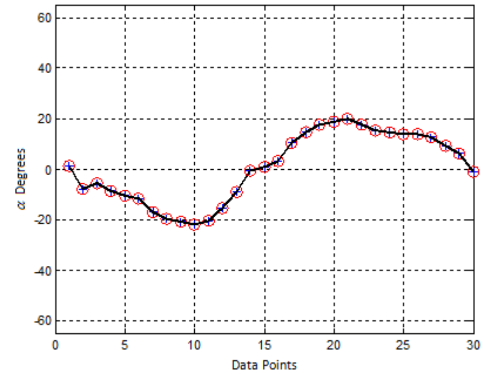

Figure-5. An example α signal after feature selection.

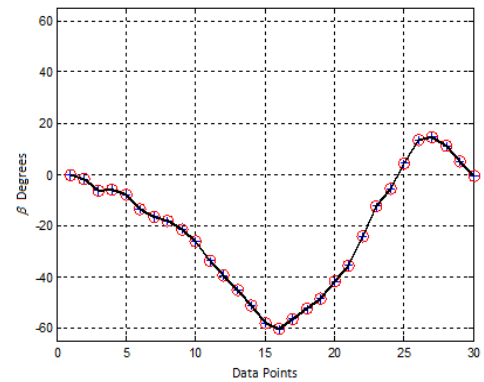

Figure-6. An example β signal after feature selection.

Classification analysis

To investigate the usefulness of the gait descriptor, machine learning method of supervised classification has been employed for calculating a correct classification rate (CCR). A high CCR value may indicate that a person has been successfully recognised by his/her kinematics descriptor and vice versa. In any supervised classification method, the data set is divided into a training and a test set. A classifier is then used to map an unlabelled exemplar from the test set to a labeled class using internal data structures information in the training set.

Two standard and established classifiers, the *k*-nearest neighbor (*k*-NN) and linear discriminant analysis (LDA) are chosen for the task. Both are similarity-based linear classifiers that consider successful classification based on nearest distance of a test sample (unknown class) to a training sample with a known class. The difference between the two is that *k*-NN measures distance nearest to some *k* numbers of samples while LDA measures distance nearest to a discriminant function.

For the TW data set, there are in total 96 samples of 24 subjects (12 pairs) with four video samples each; two samples of left-to-right walking and two samples of right-to-left walking. For the NT data set, there are 80 samples of 20 subjects (10 pairs). Altogether the SB data set contains 176 samples of 44 subjects (22 pairs). The early SOTON data set has 40 samples of 10 unique subjects. Altogether the AP data set contains 216 samples of 54 subjects. This is summarised in Table-2.

Table-2. Size of data sets.

Name	Size		
	Subjects	**Pairs**	**Total samples**
TW	24	12	96
NT	20	10	80
SB	44	22	176
AP	54	0	216

Due to the difficulty in getting volunteered subjects, this gives rise to the small sample size as seen in Table-2, validation technique is required for confirming the best estimator for classification performance. In a validation technique, data set is split into training and test sets. For each run of classification (or known as classification run), a test set is set aside containing some number of samples, other samples become the training set. At each run, the classifier is re-trained from scratch with combinations of training and test sets. The performance estimate (via CCR) is obtained as the average of the classification performance considering all runs.

The leave-one-out cross validation (LOO) takes out one sample as its test set and others become its training set at each classification run. This allows for a very accurate estimate of performance due to the large number of splits that is based on the total number of samples in a data set. The trade-off is the computational expense. However, since our data sets are small, the LOO is our preferential choice so that the classifier runs on as many total samples in our data set as possible.

RESULTS AND DISCUSSIONS

Results

Table-3–Table-5 summarise the results of classification for all data sets using **w**, α, and β signals, respectively. The standard Euclidean and City-Block distance measures are used with the *k*-NN classifier. The average CCR values in all three tables are the results when the test sample is tested for its individuality and hence gait viability as a descriptor.

In Table-3, with the *k*-NN classifier of $k = 1$ and via City-Block distance measure, an individual gait may be recognised as unique with an average CCR of 95% for NT and 76% for TW, when using the **w** data, as highlighted in Table-3. The next highest result in NT is through $k = 1$ classification via *k*-NN classifier with the Euclidean distance at 87.5%, while the third highest is the LDA classifier that gives out 75% average CCR. For TW, the second highest results of 65.6% are from $k = 3$ of *k*-NN classifier of both Euclidean and City-Block distance. LDA gives out the worst performance at 43.8% for TW data set.

Table-4 and Table-5 give out the best results when using the 1-NN City-Block classifier with values of 80% and 63.5% for NT and TW data in Table-4 while Table-5 lists 70.8% and 61.3% for NT and TW data sets, respectively for each α and β kinematics.

To further look at *how well* the descriptors can be discriminative, extra data sets known as SB data set and AP data set have been created. This AP data set can be equated to synthesising a real-world population that consists of unique individual and twins. The results are presented in the last two columns of Table-3 – Table-5. Comparing these two columns in all three tables, the best results are from 1-NN via City-Block. In Table-3, the results are 82.4% for SB data set and 74% for AP data set when using the **w** signal. Table-4 is for the α descriptors where, the results are 63.1% for SB and 55.6% for AP. Table-5 shows the results of β descriptors where, the results are 61.4% for SB and 56.5% for AP data set.

Table-3. Average CCR values for all data sets using **w** signal (Concatenation of thigh and lower leg orientation).

w signal	**Average CCR** (%)			
	TW	**NT**	**SB**	**AP**
LDA	43.8	75.0	77.8	70.8
1-NN via Euclidean	76.0	87.5	75.0	66.7
3-NN via Euclidean	65.6	65.0	60.8	56.5
5-NN via Euclidean	45.8	55.0	52.8	48.1
1-NN via City-Block	76.0	95.0	82.4	74.0
3-NN via City-Block	65.6	76.3	73.9	68.1
5-NN via City-Block	58.3	70.0	65.9	62.5

Table-4. Average CCR values for all data sets using α Signal (thigh orientation signal).

α signal	**Average CCR** (%)			
	TW	**NT**	**SB**	**AP**
LDA	52.1	66.3	53.4	46.8
1-NN via Euclidean	43.8	72.5	52.8	46.3
3-NN via Euclidean	42.7	60.0	48.3	43.1
5-NN via Euclidean	41.7	27.5	32.4	31.9
1-NN via City-Block	63.5	80.0	63.1	55.6
3-NN via City-Block	53.1	70.0	60.2	55.1
5-NN via City-Block	54.2	46.3	50.6	46.8

Table-5. Average CCR values for all data sets using β signal (Lower leg orientation signal).

β signal	Average CCR (%)			
	TW	**NT**	**SB**	**AP**
LDA	63.5	60.0	59.7	56.0
1-NN via Euclidean	61.5	60.0	56.8	51.4
3-NN via Euclidean	51.0	50.0	45.5	41.7
5-NN via Euclidean	38.5	38.8	35.8	32.9
1-NN via City-Block	70.8	61.3	61.4	56.5
3-NN via City-Block	54.1	60.0	53.4	50.0
5-NN via City-Block	46.9	48.8	44.3	41.2

Figure-7 shows the best classification performance of all three developed descriptors across all data sets. The figure is tabulating values from the 1-NN City-Block classifier from Table-3 to Table-5. The average CCR are plotted against four different data sets, which are the NT, TW, SB, and AP. This figure can be used for looking at recognition trends which has shown a decreasing trend whenever the data set sample (population) increases. As described earlier, the NT data set has 20 classes (20 persons), TW has 24 classes (24 persons), SB has 44 classes, and AP has 54 classes (20 + 24 + 10 persons). These results are useful to predict the generalisation ability of the gait signal descriptors in a large population.

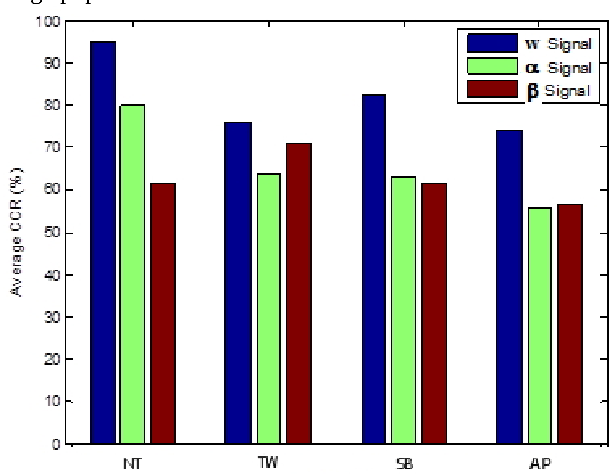

Figure-7. Recognition performance trend across data sets.

DISCUSSIONS

It is apparent that the highest results are mostly from the value of k = 1 for the k-NN classifier. This situation arises in all three NT, TW, and SB data sets. One possible causes of this is perhaps due to the sample size limitation. As mentioned in Raudys & Jain (1991) the number of neighbours among others can affect the evaluation of a finite number of sample data set. Due to the small sample size, of only four, the number of nearest neighbours of k = 3 and k = 5 does not affect the results much. This can be seen from the results where the difference between the average CCRs of 1-NN is sometimes so much higher than its 3-NN and 5-NN counterparts.

On the results between the LDA classifier and k-NN, the LDA has not been performing very well for most data sets when compared to k-NN. The LDA calculates the CCR by considering the statistics of the training set population and thus creates the decision function for classification; as opposed to the k-NN classifier where it looks at the location of the data in a classification sample space. Again the number of samples is believed to affect the results of the LDA classifier. Since statistical measure on the small size data set may be inaccurately estimated.

It may be expected that the performance of the classifiers to drop as the number of samples increases, which means an increase of the sample population. As has been seen in Figure-7, this effect has been shown by the decreasing trends occurring in the figure. Also, the performance does increase with combined signals descriptor (**w**) than any independent signal descriptor (α) and (β).

Is gait a viable biometric for use in person recognition involving twins? From the promising results in Table-3 – Table-5, perhaps TW data set can be said to contain such singularity that may allow a twin to be recognised by his/her gait. However, the results can be said to depend much on its proposed descriptor.

CONCLUSIONS

This paper has presented a gait biometric recognition analysis on twins and non-twins siblings and extrapolate its analysis to include some unique individuals to simulate a real world population sample. Throughout this paper, the problem of person recognition by gait is considered. This paper investigates the viability of gait as a biometric in recognising person when there are individuals with genetic similarities included. Guided by the nature of gait, a lower limb orientation has been adopted as its descriptor. The recognition measure is learnt from classification performances on the descriptors of four data sets, which are small in size. The results cannot be verified in complete absence of a validation method. A leave-one-out cross-validation with multiple runs and multiple partitions may allow the confirmation of the accuracy of the developed descriptors. It is hoped that the results would be a contribution to the twins biometric research domain.

Future work may consider explorations on a more holistic description of gait such as the silhouette image data as was the approach by many in the domain of gait recognition on unique individual.

ACKNOWLEDGEMENTS

This research is supported by the Exploratory Research Grant Scheme (ERGS) from the Ministry of Education, Malaysia.

REFERENCES

Bouchrika I., and Nixon M. S. (2008, December). Gait recognition by dynamic cues. In Proceedings 19th International Conference on Pattern Recognition (ICPR 2008), 1–4. IEEE.

Dinakardas C. N., Sankar S. P. and George N. (2013). MultiModal Identification System in Monozygotic Twins. International Journal of Image Processing (IJIP), 7(1), 72–90.

Giese M. A., Thornton I. and Edelman S. (2008). Metrics of the perception of body movement. Journal of Vision, 8(9):13, 1–18.

Huang P. S. (2001). Automatic gait recognition via statistical approaches for extended template features. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, , 31(5), 818–824.

Jain A. K., Klare B. and Park U. (2012). Face matching and retrieval in forensics applications. IEEE MultiMedia, 19(1), 20–28.

Mohd-Isa W. N. (2005, October). Analysis on spatial and temporal features of gait kinematics. In Proceedings Fourth IEEE Workshop on Automatic Identification Advanced Technologies, (AutoID 2005), 130–133.

Murray M.P. (1967). Gait as a total pattern of movement. American Journal of Physical Medicine, 46 (1), 290–332.

Nixon M.S., Tan T. and Chellappa R. (2006) Gait databases. Human Identification Based on Gait, International Series on Biometrics, 4, 17–34.

Park U. and Jain A. K. (2010). Face matching and retrieval using soft biometrics. IEEE Transactions on Information Forensics and Security, , 5(3), 406–415.

Phillips P. J., Flynn P. J., Bowyer K. W., Bruegge R. V., Grother P. J., Quinn G. W. and Pruitt M. (2011, March). Distinguishing identical twins by face recognition. In Proceedings 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops (FG 2011), 185–192.

Raudys S. J. and Jain A. K. (1991, March) Small sample size effects in statistical pattern recognition: recommendations for practitioners. IEEE Transactions on Pattern Analysis and Machine Intelligence, 13(3), 252–264.

Roether C. L., Omlor L., Christensen A. and Giese M. A. (2009). Critical features for the perception of emotion from gait. Journal of Vision, 9(6):15, 1–32.

Rychlik M., Stankiewicz W., Morzynski M. (2008) Method of numerical analysis of similarity and differences of face shape of twins. In Proceedings 13th International Conference on Biomedical Engineering, 23, 1854–1857.

Srinivas N., Aggarwal G., Flynn P. and Vorder Bruegge, R. W. (2012). Analysis of Facial Marks to Distinguish Between Identical Twins. IEEE Transactions on Information Forensics and Security, 7(5), 1536–1550.

Stevenage S.S., M.S. Nixon. and K. Vince (1999). Visual analysis of gait as a cue to identity. Applied Cognitive Psychology, 13, 513–526.

Strickland E. (2011, March 23) Can biometrics ID and identical twin? IEEE Spectrum Magazine.

Sun Z., Paulino A. A., Feng J., Chai, Z., Tan T. and Jain A. K. (2010). A study of multibiometric traits of identical twins. In Proceedings of the SPIE, Biometric Technology for Human Identification VII, 7667, 76670T-12.

Thurman S. M., Giese M. A. and Grossman E. D. (2010). Perceptual and computational analysis of critical features for biological motion. Journal of Vision, 10(12), 1–14.

Todd J. T. (1983). Perception of gait. Journal of Experimental Psychology: Human Perception and Performance, 9(1), 31–42.

Yam C., Nixon M. S. and Carter J. N. (2004). Automated person recognition by walking and running via model-based approaches. Pattern Recognition, 37(5), 1057–1072.