

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/321121947>

Kinect gait skeletal joint feature-based person identification

Conference Paper · July 2017

DOI: 10.1109/ICCI-CC.2017.8109783

CITATIONS

8

READS

706

2 authors, including:



Md Wasiur Rahman

The University of Calgary

12 PUBLICATIONS 64 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Conferences and Journals Special Issues [View project](#)



Cognitive Intelligence and Cognitive Machine Learning [View project](#)

DTW-based kernel and rank level fusion for 3D gait recognition using Kinect

Faisal Ahmed · Padma Polash Paul · Marina L. Gavrilova

Abstract This paper presents a new 3D gait recognition method that utilizes the Kinect skeleton data for representing the gait signature. We propose to use two new features, namely joint relative distance (JRD) and joint relative angle (JRA), which are robust against view and pose variations. The relevance of each JRD and JRA sequence in representing human gait is evaluated using a genetic algorithm. We also introduce a dynamic time warping (DTW)-based kernel that takes a collection of JRD or JRA sequences as parameters and computes a dissimilarity measure between the training and the unknown sample. The proposed kernel can effectively handle variable walking speed without any need of extra pre-processing. In addition, we propose a rank level fusion of JRD and JRA features that can boost the overall recognition performance greatly. The effectiveness of the proposed method is evaluated using a 3D skeletal gait database captured with a Kinect v2 sensor. In our experiments, rank level fusion of joint relative distance (JRD) and joint relative angle (JRA) achieves promising results, as compared against only JRD and only JRA-based gait recognition.

Keywords Gait recognition · Kinect v2 sensor · Joint relative distance · joint relative angle · DTW-kernel · 3D skeleton

1 Introduction

In recent years, an increased interest in biometric research has resulted in a new class of fast and reliable identity management solutions that are being actively

deployed in both university and corporate settings [1–3]. Human biometric traits can roughly be divided into two categories: i) physiological biometrics, such as face, iris, ear, fingerprint, palm, etc. and ii) behavioral biometrics, such as voice, handwriting, signature, etc. In most cases, biometric recognition requires physical contact or cooperation of the individual being identified [4]. In addition, a number of biometric recognition methods require high quality image or video and critically depends on the accurate detection and extraction of features (e.g. face, iris, etc.). Gait is one of the few biometrics that can be recognized from a distance, without any cooperation of the individual. Gait analysis has applications in many fields, such as surveillance, virtual and augmented reality, 3D human body modeling and animation [5,6], motion and video retrieval [7], etc. Many gait recognition methods can work on low-resolution videos [8], since only extracting the person silhouette can serve as a good enough feature representation [9]. In addition, imitating other person's gait is very difficult, which makes biometric gait recognition a robust approach to unobtrusive person identification.

While biometric gait recognition has been studied for the past twenty years, the recent popularization and low cost of Kinect has contributed to the spike in the interest in gait recognition using Kinect data. Here, the main difference from traditional gait recognition methods is that the depth and skeleton data is available directly from the device, which renders unnecessary the time consuming video processing steps. This paper presents a new Kinect-based 3D gait recognition method that utilizes the skeleton data to construct a robust representation of human gait signature. The major contributions of this paper are summarized as follow:

Firstly, we propose two new features, namely the joint relative distance (JRD) and joint relative angle

F. Ahmed · P.P. Paul · M.L. Gavrilova
Department of Computer Science, University of Calgary, Calgary, AB, Canada
E-mail: {faahmed, pppaul, mgavrilov}@ucalgary.ca

(JRA) computed among different skeletal joints, which are robust against view and pose variations. The proposed method utilizes the spatio-temporal changes in different JRD and JRA features computed over a complete gait cycle to represent the gait signature. Intuitively, relative distances and angles among different joints over a complete gait cycle represent the relative motion patterns of the corresponding joint-pair. The proposed gait representation captures this motion information and any potential synchronization between two joints to discriminate among different individuals. Secondly, we present a genetic algorithm-based joint-pair selection approach which evaluates every joint-pair based on its relevance and distinctiveness for JRD and JRA-based gait signature representation. Thirdly, we introduce a new dynamic time warping (DTW)-based kernel for gait classification. The proposed kernel uses DTW to align two gait samples based on the selected gait signature fragments originated from different joint-pairs. Individual match scores are then combined to obtain the final dissimilarity measure. The proposed kernel can effectively handle the differences in walking speed, thus eliminating the need of extra pre-processing steps, such as resampling. Finally, we propose a rank level fusion of JRD and JRA features to further boost the recognition performance. The performance of the proposed method is evaluated using a Kinect skeletal gait database. In our experiments, fusion of JRD and JRA features yields promising recognition performance, as compared against only JRD, only JRA, and some recent Kinect-based gait recognition methods.

2 Related work

Different gait recognition methods found in literature can be divided into two categories: i) model-based approaches and ii) model-free approaches [10]. In model-based approaches, explicit models are used to represent human body parts (legs, arms, etc.) [11]. Parameters of these models are estimated in each frame and the change of the parametric values over time is used to represent gait signature. However, the computational cost involved with model construction, model fitting, and estimating parameter values makes most of the model-based approaches time-consuming and computationally expensive [11]. As a result, they are infeasible for a wide range of real-world applications. BenAbdelKader et al. [12] estimated two spatiotemporal parameters of gait, namely stride length and cadence for biometric authentication of a person from video. Later, Urtasun and Fua [13] proposed a gait analysis method that relies on fitting 3-D temporal motion models to synchronized video

sequences. Recovered motion parameters from the models are then used to characterize individual gait signature. A similar approach proposed by Yam et al. [14] models human leg structure and motion in order to discriminate between gait signatures obtained from walking and running. Although this method presents an effective way to view and scale independent gait recognition, it is computationally expensive and sensitive to the quality of the gait sequences [15].

Instead of modeling individual body parts, the model-free approaches utilize the silhouette as a whole in order to construct a compact representation of walking motion [11]. Gait energy image (GEI) [16] and motion energy image (MEI) [17] are two of the most well-known model-free gait recognition methods. The basis of the MEI representation is a temporal vector image. Here, each vector point holds a value, which is a function of the motion properties at the corresponding sequence image [17]. On the other hand, GEI accumulates all the silhouette motion sequences in a single image, which preserves the temporal information as well [16]. Many of the recent model-free gait recognition methods extend GEI to a more robust representation. For example, Chen et al. [18] proposed frame difference energy image (FDEI), which utilizes denoising and clustering in order to suppress the influence of silhouette incompleteness. Li and Chen [19] fused foot energy image (FEI) and head energy image (HEI) in order to construct a more informative energy image representation. Another approach proposed by Lam and Lee [20] employs motion silhouette image (MSI) - a gray level image representation that embeds the critical spatio-temporal information of a gait signature. Although model-free approaches are computationally inexpensive, they are sensitive to view and scale changes and therefore, not suitable in uncontrolled environments.

The release of the Kinect sensor and the software development kit (SDK) by Microsoft has changed the research landscape in model-based gait recognition. Kinect is a low-cost consumer-level device made up of an array of sensors, which includes i) a color camera, ii) a depth sensor, and iii) a multi-array microphone setup. In addition, Kinect sensor can track and construct a 3D virtual skeleton from human body in real-time [21]. All these functionalities of Kinect have led to its application in different real-world problems, such as home monitoring [22], healthcare [23], surveillance [24], etc. The low-computation real-time skeleton tracking feature has encouraged some recent gait recognition methods that extract features from the tracked skeleton model. One of the pioneer studies conducted by Ball et al. [25] used Kinect for unsupervised clustering of gait samples. Features were extracted only from the lower body

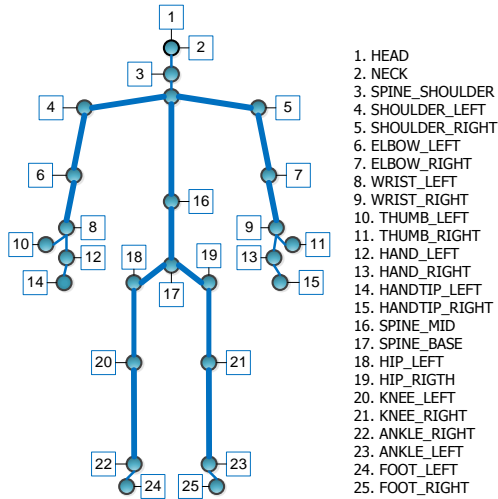


Fig. 1 3D skeleton joints tracked by the Kinect v2 sensor.

part. Preis et al. [26] presented a Kinect skeleton-based gait recognition method based on 13 biometric features: height, the length of legs, torso, both lower legs, both thighs, both upper arms, both forearms, step-length, and speed. However, these features are mostly static and represent individual body structure, while gait is considered to be a behavioral biometric, which is more related to the movement patterns of body parts during locomotion. Gabel et al. [27] used the difference in position of these skeleton points between consecutive frames as their feature. However, their proposed method was only evaluated for gait parameter extraction rather than person identification.

3 Proposed method

The proposed model-based gait recognition approach utilizes the 3D skeleton data obtained from the Kinect v2 sensor. Robustness to view and pose changes are the main advantages offered by the proposed method. Released in mid-July 2014, Kinect v2 offers a greater overall precision, responsiveness, and intuitive capabilities than the previous version [28]. The v2 sensor has a higher depth fidelity that enables it to see smaller objects more clearly, which results in a more accurate 3D object construction [28]. It can track a total of six people and 25 skeletal joints per person simultaneously [28]. In addition, while the skeleton tracking range is broader, the tracked joints are more accurate and stable than the previous version of the Kinect [28].

There are several steps involved in the proposed gait recognition method. First, complete gait cycles are

detected from video sequences recorded by the Kinect sensor. Since gait is a cyclic motion, features extracted from a single gait cycle can effectively represent the gait signature. Next, the joint relative distance (JRD) and joint relative angle (JRA) features are computed over the complete gait cycle. We consider the spatio-temporal JRD and JRA sequences originated from different joint pairs as individual fragments of the gait signature. However, not all joint-pairs are relevant in gait representation. Selection of relevant JRD and JRA features and classification based on the selected features are difficult tasks since the length of the JRD or JRA sequences might vary for the same person depending on walking speed. As a result, applying traditional statistical feature selection or classification methods requires resampling of features to obtain equal sized feature vectors extracted from different videos of the same person. However, resampling of time-sequence data involves deleting information or adding pseudo information, which might affect the selection or recognition performance. We propose a new dynamic time warping (DTW)-based kernel that can process variable length sequences of JRDs and JRAs and compute a dissimilarity measure between the training and the unknown sample without any resampling. We also propose a genetic algorithm-based joint pair selection approach that evaluates the original JRD and JRA sequences based on their distinctiveness. The fitness evaluation is performed using the proposed DTW-kernel based classifier. The gait signature fragments (JRD and JRA sequences) originated from the selected joint-pairs are then used as the training samples. In addition, we propose a rank level fusion of JRD and JRA in order to further boost the recognition performance. To our knowledge, this is one of the first studies that fuses distance and angle-based features in rank level for Kinect-based gait recognition. Figure 2 shows the components of the proposed gait recognition method.

3.1 Gait cycle detection

The first component of any gait recognition method is to isolate a complete gait cycle so that salient features can be extracted from it. Regular human walking is considered to be a cyclic motion, which repeats in a relatively stable frequency [11]. Therefore, features extracted from a single gait cycle can represent the complete gait signature. A gait cycle is composed of a complete cycle from rest (standing) position-to-right foot forward-to-rest-to-left foot forward-to rest or vice versa (left foot forward followed by a right foot forward) [29]. In order to identify gait cycles, the horizontal distance between the ANKLELEFT and AN-

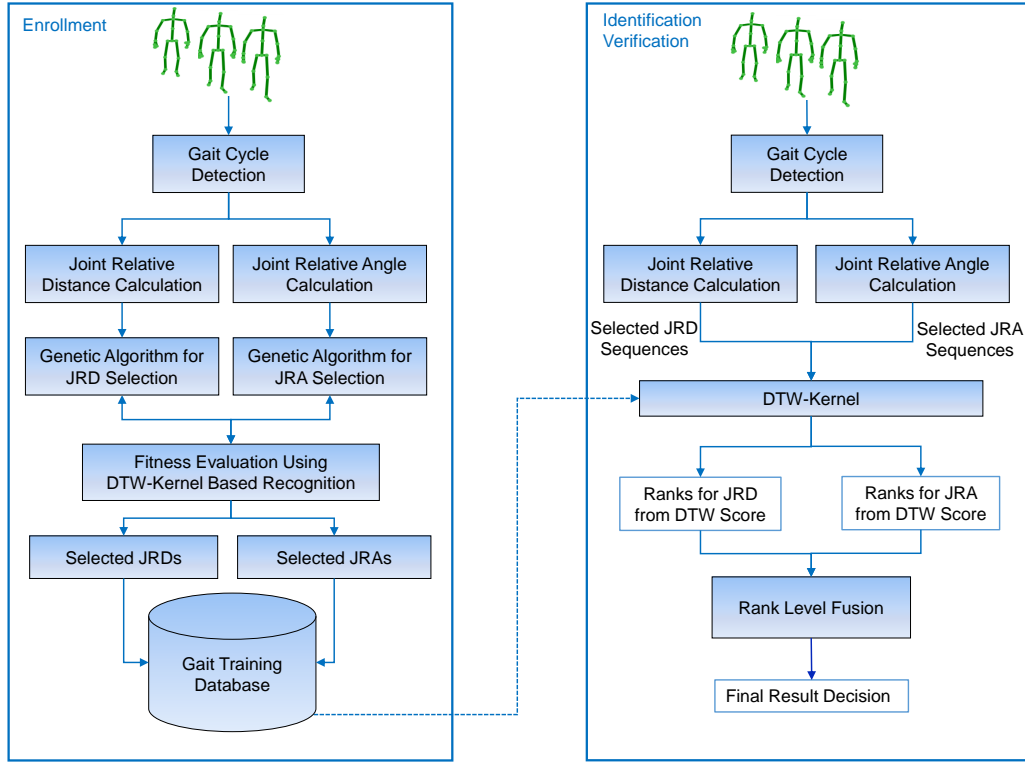


Fig. 2 Overview of the proposed gait recognition method.

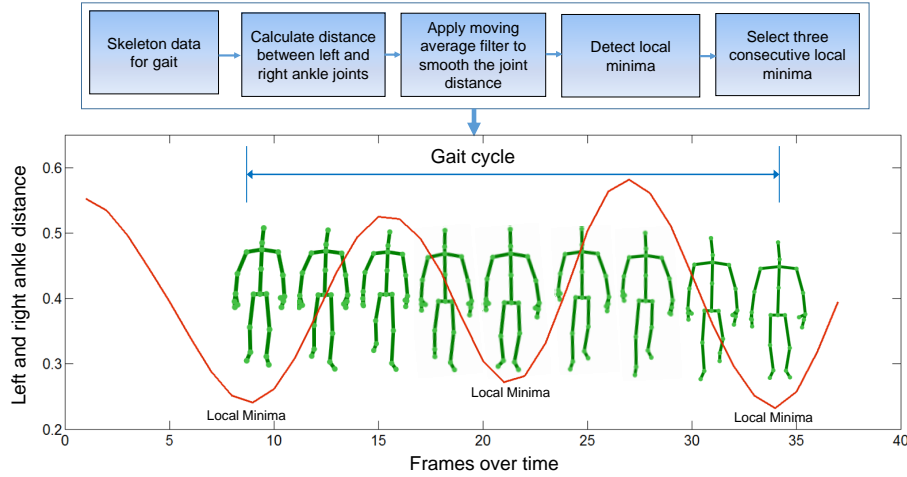


Fig. 3 Detection of a complete gait cycle by tracking the horizontal distance between the left and right ankle joints.

KLE.RIGHT joints was tracked over time, as shown in Fig. 3. A moving average filter was used to smooth the distance vector. During the walking motion, the distance between the two ankle joints will be the maximum when the right and the left legs are farthest apart and will be the minimum when the legs are in the rest (standing) position. Therefore, by detecting three subsequent minima, it is possible to find the three sub-

sequent occurrences of the two legs in the rest position, which corresponds to the beginning, middle, and ending points of a complete gait cycle, respectively [30].

3.2 Gait feature representation

In this paper, we propose two different gait feature representations, one is based on the joint relative distance

(JRD) and the other one is based on the joint relative angle (JRA) measures. This section presents a detail discussion on JRD and JRA.

Joint relative distance (JRD) between any two skeletal joints $p_1(x_1, y_1, z_1)$ and $p_2(x_2, y_2, z_2)$ can be defined as the Euclidean distance between these two joints in a 3D space [31]:

$$\delta(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

JRDs computed over time provide an intuitive representation of the relative movements of the joints involved. This representation is particularly useful in gait recognition, since it can capture the notion of synchronization between different joints. For example, movement of the left ankle during walking is synchronized with the movement of the right ankle. JRDs between these two joints can effectively capture this trait. Previously, Tang and Leung [32] used variance of JRD (VJRD) measures to represent certain human motions in their proposed motion retrieval system. In this paper, we show that, JRDs computed over a complete gait cycle for certain joint pairs can effectively be used to represent the gait pattern. Here, instead of using the variance of JRDs, we treat the JRDs as individual gait signature fragments originated from the corresponding joint pairs.

On the other hand, joint relative angle (JRA) between two joints p_1 and p_2 can be defined as the angle formed by p_1 and p_2 with respect to a reference point r . Given the coordinates of 3 points p_1 , p_2 , and r in a 3-D space, the angle Θ_{p_1, p_2} formed by $p_1 \rightarrow r \rightarrow p_2$ using the right hand rule from r can be calculated as:

$$\Theta_{p_1, p_2} = \cos^{-1} \frac{\overrightarrow{p_1 r} \cdot \overrightarrow{r p_2}}{\|\overrightarrow{p_1 r}\| \|\overrightarrow{r p_2}\|} \quad (2)$$

Here, $\overrightarrow{p_1 r} = r - p_1$, $\overrightarrow{r p_2} = p_2 - r$, the dot(.) represents dot product between two vectors, and $\|\overrightarrow{p_1 r}\|$ and $\|\overrightarrow{r p_2}\|$ represent the length of $\overrightarrow{p_1 r}$ and $\overrightarrow{r p_2}$, respectively. The SPINE.BASE joint was selected as the reference point, since this joint remains almost stationary during walking. Just like JRDs, changes of JRA values over the full gait cycle also provide an intuitive representation of the relative movements among different joints. In this study, we consider JRDs or JRAs originated from a particular joint-pair as a small fragment of the gait signature. Thus, the full gait signature can be defined as a collection of JRD or JRA values over a complete gait cycle originated from different joint-pair combinations of the human skeleton. For the 25 skeletal joints, there is a total of 300 possible joint-pair combinations, which is a high-dimensional feature space. In addition, not all joint-pair is relevant in gait feature representation. For example, JRDs or JRAs between the

SPINE_SHOULDER and the SPINE_MID joints does not represent any information related to human gait, since both these joints remain almost stationary when a person walks. Therefore, identifying the skeletal joint-pairs that are relevant to human gait motion is imperative for the proposed gait recognition method.

We propose a genetic algorithm-based relevant joint-pair selection method for gait feature representation. Genetic algorithm (GA) [33] is an evolutionary search heuristic that works with a pool of individuals. These individuals can be represented using different data structures, such as strings, trees, graphs, etc. In the proposed method, the individuals are represented as a fixed-length string of decimal values, ranging from 1 to 300, without any duplication. Here, each value represents a particular joint-pair combinations. The length of the string was set to the number of the most relevant joint-pairs we are searching for.

Genetic algorithm works in an iterative fashion. At each iteration, a population of individuals is generated. The algorithm starts with a randomly generated set of individuals. For each individual in a population, a fitness value is then calculated. For this study, we evaluate the fitness of each individual based on the discriminative capability of the corresponding joint-pairs encoded by the individual. We use the collection of gait signature fragments originated from the corresponding joint-pair combinations represented by the individual as the whole gait signature representation (feature descriptor). This representation of gait feature is then used for the classification task and the obtained recognition rate is used as the fitness value for that particular individual. For the classification task, we propose a DTW-based kernel, that generates individual match scores from the fragments of the gait signature and combine the result to obtain the overall match score. The individuals are then ranked according to their fitness value and a number of best individuals are selected to generate a set of offspring through the application of crossover and mutation operations on them. The off-spring are added to the population of individuals and only the individuals having high fitness value are selected for the next iteration. This process is continued for a fixed number of iterations and the best-so-far individual is selected as the most relevant set of joint-pairs.

3.3 DTW-kernel for gait recognition

Joint relative distances (JRD) or joint relative angles (JRA) for different joint-pairs computed over a full gait cycle essentially represent sequences of time-series data. Alignment of such temporal gait data is a challenging

task due to variation of walking speed, which might result in variable length JRD/JRA sequences for the same person. Therefore, applying traditional classifiers in this scenario requires extra pre-processing steps, such as resampling. However, resampling of time-sequence data involves deletion or adding new data, which might affect the recognition performance. On the other hand, non-linear time sequence alignment techniques can effectively reduce the effect of variable walking speed by warping the time axis. Dynamic time warping (DTW) is a well-known non-linear sequence alignment technique. Originally proposed for speech signal alignment [34], recent DTW applications are mostly verification-oriented, such as offline signature verification [35]. In this paper, we propose to utilize DTW to design a kernel for gait recognition that takes a collection of JRD/JRA time series data originated from different joint pairs as the parameter and outputs the dissimilarity measure between two given gait samples. Use of DTW in this case allows the alignment of different length JRD/JRA sequences, which enables the classifier to match gait samples without any intermediate resampling stage.

Given the set of all joint relative distances $JRD = \{\delta_1, \delta_2, \dots, \delta_p\}$, where each δ_i represents JRDs for two particular joints computed over a full gait cycle, we obtain a subset of JRD by selecting relevant joint-pairs using genetic algorithm:

$$\delta = \{\delta_i | i = 1, 2, \dots, N \text{ where } \delta_i \in JRD\} \quad (3)$$

Similarly, given the set of all joint relative angles $JRA = \{\theta_1, \theta_2, \dots, \theta_q\}$, where each θ_i represents JRAs for two particular joints with respect to the reference point computed over a full gait cycle, the selected subset after applying genetic algorithm can be represented as:

$$\theta = \{\theta_i | i = 1, 2, \dots, M \text{ where } \theta_i \in JRA\} \quad (4)$$

Let, δ_{train} and δ_{test} are two JRD sequences from the same joint-pair computed over a complete gait cycle, where the length of δ_{train} and δ_{test} are represented as $|\delta_{train}|$ and $|\delta_{test}|$, respectively.

$$\delta_{train} = a_1, a_2, a_3, \dots, a_{|\delta_{train}|} \quad (5)$$

$$\delta_{test} = b_1, b_2, b_3, \dots, b_{|\delta_{test}|} \quad (6)$$

Here, a_t and b_t are the JRD values of δ_{train} and δ_{test} at time t , respectively. Given these two time series, DTW constructs a warp path $W = w_1, w_2, w_3, \dots, w_L$, where $\max(|\delta_{train}|, |\delta_{test}|) \leq L \leq |\delta_{train}| + |\delta_{test}|$. Here, L is the length of the warp path between the two JRD sequences. Each element of the path can be represented as $w_l = (x, y)$, where x and y are two indices from the δ_{train} and δ_{test} , respectively. There are a number of constraints that DTW must satisfy. Firstly, the warp path

must start at $w_1 = (1, 1)$ and end at $w_L = (|\delta_{train}|, |\delta_{test}|)$. This in turn ensures that, every index from the both time series is used in path construction. Secondly, if an index i from δ_{train} is matched with an index j from δ_{test} , it is prohibited to match any index $> i$ with any index $< j$ and vice-versa. This restricts the path from going back in time. Given these restrictions, the optimal warp path can be defined as the minimum distance warp path $dist_{optimal}(W)$:

$$dist_{optimal}(W) = \min \sum_{l=1}^L \{dist(w_{li}, w_{lj})\} \quad (7)$$

Here, w_{li} and w_{lj} are two indices from δ_{train} and δ_{test} , respectively and $dist(w_{li}, w_{lj})$ is the Euclidean distance between w_{li} and w_{lj} .

We extend this basic DTW formulation to a kernel in order to compute the dissimilarity between a training and a testing gait sample, each of which is a collection of JRD/JRA sequences of different joint-pairs. The proposed DTW-kernel aligns the training and testing JRD/JRA sequences of the same joint-pair with each other and computes a match score between them. Summation of all the match scores obtained from the different joint-pair JRD/JRA sequences from the training and testing samples is treated as the final dissimilarity measure. Formally, the proposed DTW kernel Δ for JRD-based gait representation can be defined as:

$$\Delta(\delta, \delta') = \sum_{n=1}^N \{ \min \sum_{l=1}^L \{dist(w_{n,li}, w_{n,lj})\} \} \quad (8)$$

Here, $\delta = \{\delta_1, \delta_2, \dots, \delta_N\}$ and $\delta' = \{\delta'_1, \delta'_2, \dots, \delta'_N\}$ are collections of JRD sequences from N different joint-pairs and $\min \sum_{l=1}^L \{dist(w_{n,li}, w_{n,lj})\}$ represents the minimum warp path distance between the n -th joint pair JRDs of δ and δ' . The same formulation can also be used for the DTW kernel Δ for JRA-based gait representation:

$$\Delta(\theta, \theta') = \sum_{m=1}^M \{ \min \sum_{l=1}^L \{dist(w_{m,li}, w_{m,lj})\} \} \quad (9)$$

Here, $\theta = \{\theta_1, \theta_2, \dots, \theta_M\}$ and $\theta' = \{\theta'_1, \theta'_2, \dots, \theta'_M\}$ are collections of JRA sequences from M different joint-pairs and $\min \sum_{l=1}^L \{dist(w_{m,li}, w_{m,lj})\}$ represents the minimum warp path distance between the m -th joint pair JRAs of θ and θ' .

JRD and JRA represents features from different domains: one is distance-based features and the other one is angle-related features. Feature level fusion of JRD

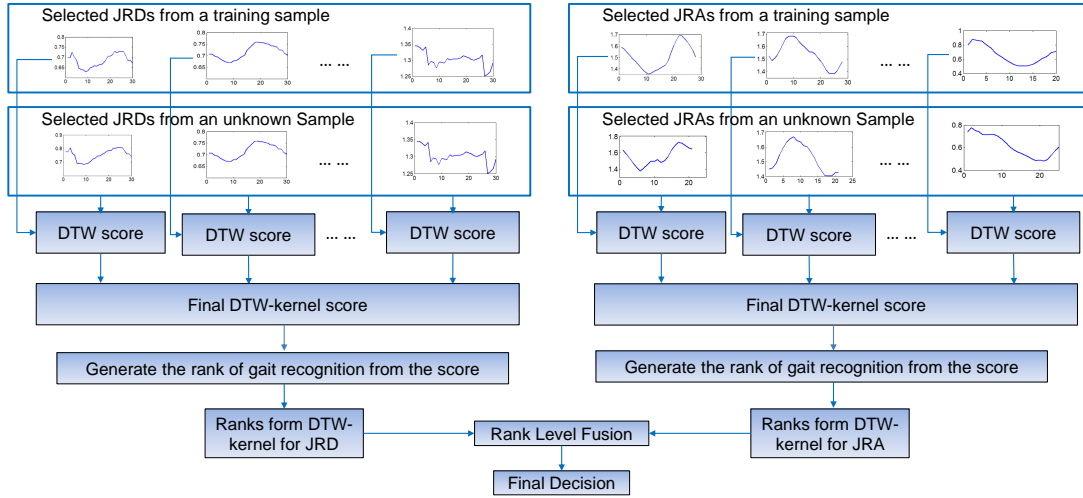


Fig. 4 Proposed rank-level fusion of JRD and JRA features based on DTW-Kernel.

and JRA is a difficult task due to the differences in feature domain. Therefore, we propose to use rank-level fusion of JRD and JRA to improve the recognition performance at the decision level. In the proposed fusion, both JRD and JRA sequences are used to generate match scores individually and the top 3 candidates from both are selected for voting. The candidate with the highest number of votes is selected as the match. Figure 4 illustrates the proposed method.

4 Experiments and Results

4.1 Experimental setup and dataset description

The performance of the proposed method is evaluated using a Kinect skeletal gait database, provided by the SMART Technologies ULC, Calgary, Canada. The gait database comprises 20 participants (14 male, 6 female), from around 20 to 35 years old. For each person, a series of 3 videos was recorded in a meeting room environment. The position of the Kinect was fixed throughout the recording session. Each of the video scenes contains a participant entering the meeting room, walking toward a chair, and then sitting on the chair. Figure 5 shows a frame of a sample video from the gait database. We conducted a 3-fold cross-validation in order to evaluate the effectiveness of the proposed method. In a 3-fold cross-validation, the whole dataset is randomly divided into 3 subsets, where each subset contains an equal number of samples from each category. The classifier is trained on 2 subsets, while the remaining one is used for testing. The average classification rate is calculated after repeating the above process for 3 times.

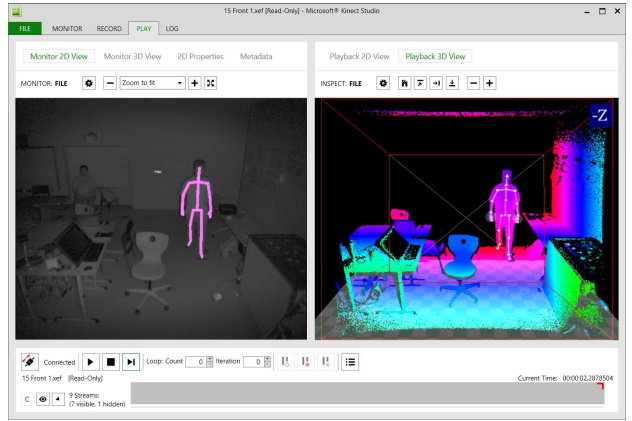


Fig. 5 Sample video frame from the gait database captured using Kinect v2 sensor.

Since the database comprises 3 videos per person, in each fold, two videos were used for the training and the rest was used for testing.

4.2 Results and discussions

For the 25 skeletal joints tracked by the Kinect sensor, there is a total of 300 possible JRD or JRA sequences that can be extracted from a gait cycle. In order to identify relevant joint-pair features for gait representation, we first apply the proposed genetic algorithm (GA)-based JRA and JRD selection method. We discard a number of joints beforehand in order to reduce redundancy. For example, the hand, thumb, and handtip joints are redundant with respect to wrist. These four joints are almost at the same position and the

movements of these 4 joints are almost identical during walking. Therefore, we discard the hand, thumb, and handtip joints from the both hands and consider only the wrist. The same can be said for ankle and foot joints. In our observation of the database, we found that, the foot joint movements are noisy in most cases. Therefore, only ankle joints are considered. The GA was run separately for JRD and JRA features. For both cases, the population size was set to 20, where each individual was evaluated based on the recognition performance of the corresponding joint-pairs encoded by the individual. With this setup, the simulation was run for 20 times and the set of joint-pairs encoded by the best individual was selected as the final outcome.

For joint relative distance, the GA was applied to find the best 10, 15, 20, 25, and 30 JRD sequences. Figure 6 shows the cumulative match characteristic (CMC) curve for the selected JRD sequences. It can be observed that, the best recognition performance is obtained for the subset of 25 joint relative distances. Further increasing the size of the subset does not contribute in the performance, rather results in a decrease. In addition, while the selected 25 JRD sequences achieve a recognition rate of 100% at rank 5, the other subsets achieve at most 95%. Based on the analysis of the selected JRD sequences, Fig. 7 shows a graphical representation of the most-relevant joint-pairs for JRD-based gait signature representation. Here, each skeletal joint is considered as a node in a graph and an edge between two joints represents a relevant joint-pair combination. This representation of joint-pair relationships provide a better understanding of the underlying motion patterns of human gait. From the figure, we can identify some clusters of joint-pair relations based on human body structure. They are: i) edges between upper-left and lower-right joints, ii) upper-right and lower-left joints, iii) upper-left and lower-left joints, iv) upper-right and lower-right joints, v) lower-left and lower-right joints, and v) center and lower-left/lower-right joints. In a broader view, JRDs between i) upper body and lower body joints and ii) lower-left and lower-right joints are the most relevant in JRD-based human gait representation.

For joint relative angle, the GA was applied to find the best 10, 15, 20, and 25 JRA sequences. Figure 8 shows the cumulative match characteristic (CMC) curve for the selected JRA sequences. In this case, the best recognition performance is obtained for the subset of 15 joint relative angles. Increasing the size of the subset any further results in a decrease in recognition performance. For the selected 15 JRA sequences, a recognition rate of 100% is achieved at rank 5, while for the other subsets of JRA sequences, the highest recognition rate is around 98%. Figure 9 shows the graphical

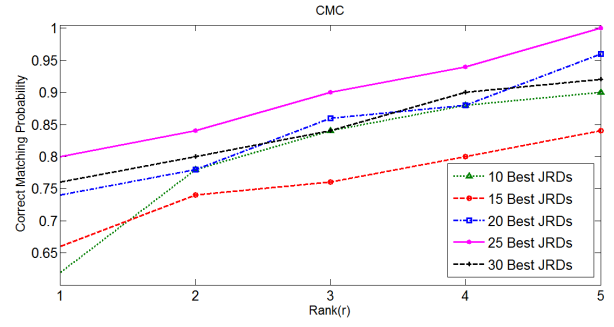


Fig. 6 Cumulative match characteristic (CMC) curve for different subsets of selected JRD sequences.

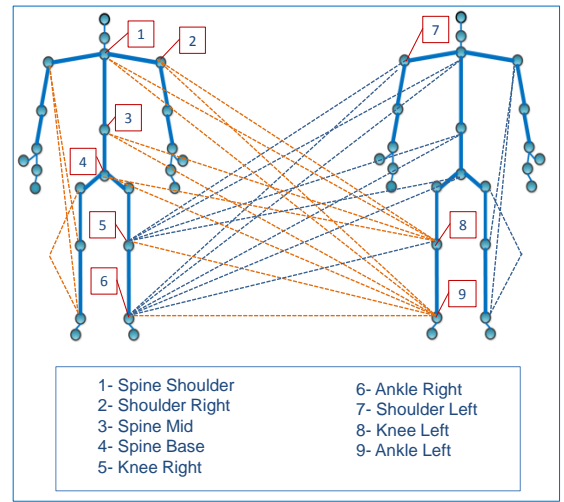


Fig. 7 A graphical representation of the selected best JRD joint-pairs.

representation of the most relevant joint-pairs for JRA-based gait signature representation. It can be observed that, the joint-pair relations for JRAs are almost similar to JRDs. However, for JRA-based representation, while no joints from the body center (spine section) was selected, relations between the i) left shoulder and left wrist joints and the ii) right shoulder and right wrist joints are found informative.

The selected JRD and JRA sequences are used in the proposed rank level fusion in order to boost the recognition performance. We compare the performance of the proposed fusion against only JRA and only JRD-based gait recognition. Figure 10 shows the CMC curve for the recognition performance using JRA-JRD rank level fusion, only JRA, and only JRD. It can be observed that, fusion of JRD and JRA features at the rank level increases the recognition performance. For rank 1, the JRD-JRA fusion achieves a recognition rate of 92%, while the recognition rates for only JRD and

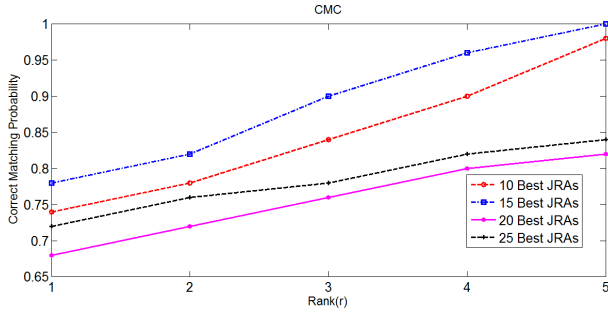


Fig. 8 Cumulative match characteristic (CMC) curve for different subsets of selected JRA sequences.

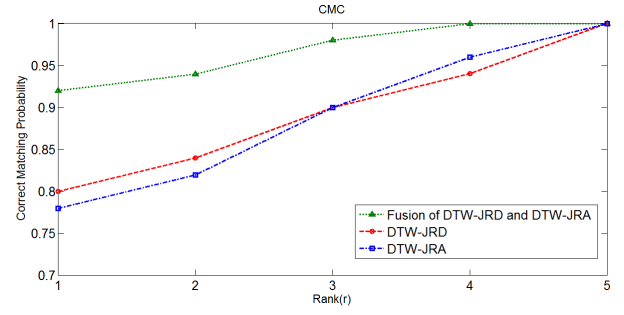


Fig. 10 Cumulative match characteristic (CMC) curve for rank level fusion of JRD and JRA, only JRD, and only JRA.

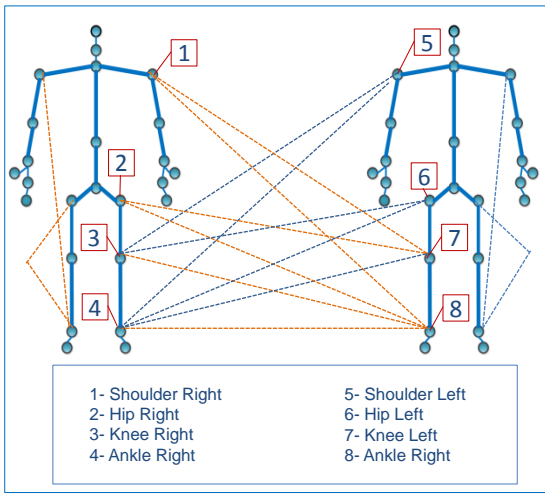


Fig. 9 A graphical representation of the selected best JRA joint-pairs.

only JRA are 80% and 78%, respectively. In addition, while the JRD-JRA fusion achieves 100% recognition rate at rank 4, individually both JRD and JRA reaches 100% recognition rate at rank 5.

Table 1 Recognition rates of different methods (for 3-fold cross validation).

Method	Recognition Rate (%)
Fusion of JRD and JRA + DTW-Kernel	92.1
Ball et al. [22]	66.7
Preis et al. [23]	84.2

Finally, we compare the performance of the proposed method against some recent Kinect skeleton-based gait recognition methods. We have selected two studies and tested their performance on our gait database. Details of the selected two methods can be found in [25]

and [26]. Table 1 shows the recognition performance of these methods for rank one. From the experimental results, it can be said that, gait recognition based on the DTW-kernel and rank level fusion of JRD and JRA sequences is more robust and achieves higher recognition performance than some of the existing gait recognition methods. The superiority of the proposed method is due to the utilization of view and pose invariant relative distance and angle features in a rank level fusion and non-linear alignment of variable length feature sequences using the DTW-kernel.

5 Conclusion

We proposed a new gait recognition method that uses an inexpensive consumer-level sensor, namely the Microsoft Kinect v2. The proposed method utilizes the spatio-temporal changes in relative distances and angles among different skeletal joints to represent the gait signature. To determine the relevance of a joint pair in representing human gait, we used a genetic algorithm that evaluates the gait signature fragments originated from different joint-pairs based on their distinctiveness. We also proposed a new DTW-based kernel that combines individual DTW match scores computed for a collection of JRD or JRA sequences. Here, the use of DTW makes the proposed kernel robust against variable walking speed and thus eliminates any need of extra pre-processing. Finally, the recognition performance was further boosted by a rank level fusion of JRD and JRA. Extensive experiments using a Kinect skeletal gait database showed promising results for the fusion of JRD and JRA features, compared against only JRD and only JRA-based recognition. In the future, we plan to extend the proposed method for action recognition and motion retrieval.

6 Acknowledgment

The authors would like to thank NSERC DISCOVERY program, URG, NSERC ENGAGE, AITF, and SMART Technologies ULC, Canada for partial support of this project.

References

1. Deutschmann I., Nordstrom P., Nilsson L.: Continuous Authentication Using Behavioral Biometrics. *IT Professional*, 15(4), 12-15 (2013)
2. Zhou X., Bhanu B.: Integrating face and gait for human recognition at a distance in video. *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, 37(5), 1119-1137 (2007)
3. Prabhakar S., Pankanti S., Jain A.K.: Biometric recognition: security and privacy concerns. *IEEE Security and Privacy*, 1(2), 33-42 (2003)
4. Munsell B.C., Temlyakov A., Qu C., Wang S.: Person Identification Using Full-Body Motion and Anthropometric Biometrics from Kinect Videos. In: *Proc. European Conf. on Computer Vision Ws/Demos, LNCS 7585*, 91-100 (2012)
5. Zhang Y., Zheng J., Magnenat-Thalmann N.: Example-guided anthropometric human body modeling. *The Visual Computer (CGI 2014)*, 1-17 (2014)
6. Bae M.S., Park I.K.: Content-based 3D model retrieval using a single depth image from a low-cost 3D camera. *The Visual Computer (CGI 2013)*, 29, 555-564 (2013)
7. Zhou L., Zhiwu L., Leung H., Shang L.: Spatial temporal pyramid matching using temporal sparse representation for human motion retrieval. *The Visual Computer (CGI 2014)*, 30, 845-854 (2014)
8. Das Choudhury S., Guan Y., Chang-Tsun L.: Gait recognition using low spatial and temporal resolution videos. In: *Proc. Intl. Work. on Biometrics and Forensics*, 1-6 (2014)
9. Wang L., Tan T., Ning H., Hu W.: Silhouette analysis-based gait recognition for human identification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(12), 1505-1518 (2003)
10. Han J., Bhanu B.: Statistical Feature Fusion for Gait-based Human Recognition. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2, 842-847 (2004)
11. Wang J., She M., Nahavandi S., Kouzani A.: A review of vision-based gait recognition methods for human identification. In: *Proc. IEEE Intl. Conf. on Digital Image Computing : Techniques and Application*, 320-327 (2010)
12. BenAbdelkader C., Cutler R., Davis L.: Stride and cadence as a biometric in automatic person identification and verification. In: *Proc. IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, 372-377 (2002)
13. Urtasun R., Fua P.: 3D Tracking for Gait Characterization and Recognition. In: *Proc. IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, 17-22 (2004)
14. Yam C., Nixon M.S., Carter J.N.: Automated person recognition by walking and running via model-based approaches. *Pattern Recognition*, 37, 1057-1072 (2004)
15. Sinha A., Chakravarty K., Bhowmick B.: Person Identification using Skeleton Information from Kinect. In: *Proc. Intl. Conf. on Advances in Computer-Human Interactions*, 101-108 (2013)
16. Han J., Bhanu B.: Individual recognition using gait energy image. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 28, 316-322 (2006)
17. Bobick A.F., Davis J.W.: The recognition of human movement using temporal templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23, 257-267 (2001)
18. Chen C., Liang J., Zhao H.: Frame difference energy image for gait recognition with incomplete silhouettes. *Pattern Recognition Letters*, 30, 977-984 (2009)
19. Li X., Chen Y.: Gait Recognition Based on Structural Gait Energy Image. *Jour. of Computational Information Systems*, 9(1), 121-126 (2013)
20. Lam, T.H.W., Lee, R.S.T.: A new representation for human gait recognition: Motion silhouettes image (MSI). In: *Proc. Intl. Conf. on Biometrics*, 612-618 (2006)
21. Shotton J., Fitzgibbon A., Cook M., Sharp T., Finocchio M., Moore R., Kipman A., Blake A.: Real-time human pose recognition in parts from single depth image. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1297-1304 (2011)
22. Stone E.E., Skubic M.: Evaluation of an inexpensive depth camera for passive in-home fall risk assessment. In: *Proc. Intl. Pervasive Computing Technologies for Healthcare Conf.*, 71-77 (2011)
23. Chang Y.J., Chen S.F., Huang J.D.: A Kinect-based system for physical rehabilitation: A pilot study for young adults with motor disabilities. *Research in Developmental Disabilities*, 32(6): 2566-2570 (2011)
24. Popa M., Koc A.K., Rothkrantz L.J.M., Shan C., Wiggers P.: Kinect Sensing of Shopping Related Actions. *Communications in Computer and Information Science*, 277: 91-100 (2012)
25. Ball A., Rye D., Ramos F., Velonaki M.: Unsupervised Clustering of People from 'Skeleton' Data. In: *Proc. ACM/IEEE Intl. Conf. on Human Robot Interaction*, 225-226 (2012)
26. Preis J., Kessel M., Linnhoff-Popien C., Werner M.: Gait Recognition with Kinect. In: *Proc. Work. on Kinect in Pervasive Computing* (2012)
27. Gabel M., Gilad-Bachrach R., Renshaw E., Schuster, A.: Full body gait analysis with Kinect. In: *Proc. Annual Intl. Conf. of the IEEE Engineering in Medicine and Biology Society*, 1964-1967 (2012)
28. Kinect for windows features [Online, accessed on October 10, 2014, Available: <http://www.microsoft.com/en-us/kinectforwindows/meetkinect/features.aspx>]
29. Kale A., Sundaresan A., Rajagopalan A.N., Cuntoor N.P., Roy-Chowdhury A.K., Kruger V., Chellapa R.: Identification of Humans Using Gait. *IEEE Trans. on Image Processing*, 13(9), 1163-1173 (2004)
30. Sarkar S., Phillips P.J., Liu Z., Vega I.R., Grother P., Bowyer K.W.: The humanID gait challenge problem: data sets, performance, and analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 27(2), 162-177 (2005)
31. Tang J.K., Leung H., Komura T., Shum H.P.: Emulating human perception of motion similarity. *Computer Animation and Virtual Worlds* 19, 211-221 (2008)
32. Tang J.K., Leung H.: Retrieval of logically relevant 3D human motions by Adaptive Feature Selection with Graded Relevance Feedback. *Pattern Recognition Letters*, 33, 420-430 (2012)
33. Holland J.H.: *Adaptation in Natural and Artificial Systems*, University of Michigan Press (1975)
34. Kruskal J.B., Liberman M.: The symmetric time-warping problem: from continuous to discrete. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparisons*. Addison-Wesley, Reading, Massachusetts (1983)
35. Shanker A.P., Rajagopalan, A.N.: Off-line signature verification using DTW. *Pattern Recognition Letters*, 28, 1407-1414 (2007)