Proceedings of 8th IOE Graduate Conference
Peer Reviewed
ISSN: 2350-8914 (Online), 2350-8906 (Print)
Year: 2020 Month: June Volume: 8

Model Based Gait Recognition Using Weighted KNN

Shyamdeep Sah [a], Sanjeeb Prasad Panday [b]

[a, b] *Department of Electronics and Computer Engineering, Pulchowk Campus, IOE, TU, Nepal*
Corresponding Email: [a] sspsah3@gmail.com, [b] sanjeeb@ioe.edu.np

Abstract
Gait recognition is to identify humans based on their way of walking. Unlike standard biometric recognition techniques, gait recognition is a non-intrusive technique. Both data collection and classification processes can be done without a subject's cooperation. In this work, a new modelbased gait recognition technique that used Weighted KNN is proposed. The technique consists of three elements: transformation of Kinect skeletal coordinates into CoB (Centre of Body) relative coordinates, posture-based features extraction and posture-based classification that uses weighted-KNN. Before in this work, gait recognition have been implemented by using algorithms such as KNN, SVM, MLP, 1R, C4.5 decision tree, Naive Bayes and Extra-trees. Among these algorithms, KNN produced the best accuracy. In this work, gait recognition has been implemented using Weighted KNN which is extended version of KNN and it was found that W-KNN, compared to KNN, requires less number of frames to produce result at given accuracy.

Keywords
gait recognition, Kinect, Centre of Body coordinate, weighted-KNN Classification

1. Introduction

Gait recognition is a biometric identification technique that uses the way that people walk to identify a person. It focuses on limb movements during a walk. Unlike other biometric identification techniques, gait data can be unobtrusively collected from afar without a subject's awareness. It is also much harder to alter a person's gait features continuously. Moreover, gait recognition can be performed at a distance [1].

Gait recognition technology can be classified into two types: model-free (appearancebased) methods and model based methods [2]. Model-free gait recognition technologies use gait features extracted directly from images captured from human walking sequences. Most of these model-free techniques use silhouettes or partial silhouettes of bodies to analyse gait information. Gait Energy Image (GEI) is a model-free gait recognition technique which uses the average image of silhouettes as gait features [3]. Model-based gait recognition techniques classify based on human body structures and skeletal data. Before advances in sensor devices, not much had been done in model-based techniques. However, that has changed with recent developments, especially the development of the Microsoft Kinect and Microsoft Kinect SDK [4].

2. Related Work

Model-based gait recognition system took its development pace with the introduction of Kinect by Microsoft. The first remarkable work was done by J. Preis et al. in 2012 [5]. 13 biometric features for the identification of a person: The height, the length of legs, torso, both lower legs, both thighs, both upper arms, both forearms, the step-length, and the speed were used. The design was implemented with three different classifiers: 1R, a C4.5 decision tree and a Naive Bayes. The best accuracy (91%) came with Naive Bayes classifier with 9 human subjects for fixed direction walk.

In 2013, Cheewakidakarn et al. defined a distance between two walking sequences using static features, such as average length of right arm, and kinematic features such as body tilt angle, with weight factors [6]. Euclidean distance was used to calculate "static distance" and Dynamic Time Warping was used to measure "kinematic distance" between two walking sequences. KNN was used in their implementation and the technique accuracy was 51.77with 17 human subjects for free style walk.

N. Jianwattanapaisarn et al. also represented a human walking sequence in a sequential data form. The two techniques, model-based and model-free, were combined in their design [7]. A distance between two walking sequences was measured with three different types of features: static distance, kinematic distance, and mass vector distance, which was based on binary silhouettes of a body. KNN was used and their technique accuracy was 92.56fixed direction walk. All of them have different height, body shape and weight. Andersson and Araujo created a dataset where 160 participants were asked to walk in a half-circle path where a Kinect was placed in the centre of the circle and rotated to follow a participant [8]. 60 gait attributes and 20 Anthropometric Attributes were used. The best result came from mainly using static body (anthropometric) features. Kinematic parameters alone perform poorly. The design was implemented with three different classifiers: KNN (K=5), SVM and MLP. The best accuracy was (87.7%) using KNN.

Yang et al. used relative distance features and anthropometric features adding the 5 standard deviation to distance-based gait features [9]. The same dataset, which was created by Andersson and Araujo, was used [8]. The design was implemented using KNN and the technique accuracy was 95.4most remarkable work in model based gait recognition using Kinect was done by Nirattaya Khamsemanan et al. in 2018 [10]. The concept of Centre of Body (CoB) which addressed the multi-viewpoint issue of gait recognition was proposed. Posture-based feature extraction and posture-based classification was used in the design. The feature vector consists of 20 coordinates of a frame plus distances between coordinates of two consecutive frames of a walking sequence. The design was implemented with three different classifiers: KNN (K = 1), Extra-Trees and MLP. The best accuracy was (97.0 ± 0.5%) using KNN with 140 human subjects for fixed-direction walk.

Gait recognition is a current research field where the struggle has been for accuracy improvement. The previous research works show that the best accuracy has come with KNN and up to 97% accuracy has been achieved. The main objective of this research is to see how Weighted-KNN performs over simple KNN. Weighted-KNN has produced better accuracy in many research works [11, 12, 13].

3. Research Methodology

3.1 Block Diagram

Training:

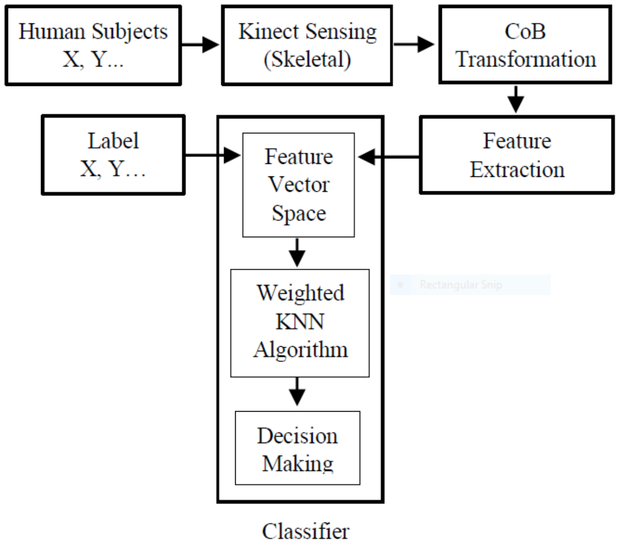


Figure 1: Block diagram of training steps

Testing:

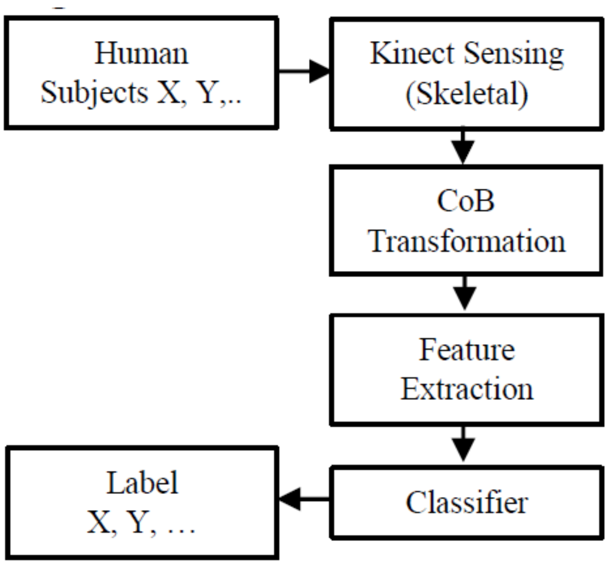


Figure 2: Block diagram of testing steps

3.2 Centre-of-Body Relative Coordinates

Gait data extracted in each frame are viewdependent. This means that directions of one frame to another frame in the same walking sequences are different. To address this issue, a concept of coordinates, called Centre-of-Body (CoB) relative coordinates, proposed by Nirattaya Khamsemanan et al. is used [10]. In one frame, Kinect coordinates of joints are transformed into CoB relative coordinates, based on their positions relative to 4 joints: hip-centre, spine, hip-right, and hip-left joints. This transformation can be done in two steps: finding a new basis for a frame and calculating CoB relative coordinates of joints based on these new axes. Intuitively speaking, CoB relative coordinates are obtained from moving an observation point of view so that all frames are monitored from the same angle. The Kinect system mirrors a user facing the sensor by default. One of the limitations of the Kinect skeletal system is that it cannot detect if a person is facing a camera or walking away. Consequently, in some frames, Kinect coordinates of 8 left and right joints are switched. To eliminate this limitation, a consecutive frame was used to determine the direction of a frame. If a subject is walking away from the camera, we reverse Kinect coordinates of all left and right joints before starting the process of CoB relative coordinates.

3.2.1 Finding a new basis:

In each frame, suppose Kinect coordinates of hip-centre, spine, hip-right, and hip-left joints are (x_0, y_0, z_0), (x_1, y_1, z_1), (x_2, y_2, z_2) and (x_3, y_3, z_3), respectively. Define vector v_s and v_h as shown in the equations (1) and (2).

$$V_s = \begin{bmatrix} x_1 - x_0 \\ y_1 - y_0 \\ z_1 - z_0 \end{bmatrix} \tag{1}$$

$$V_h = \begin{bmatrix} x_3 - x_3 \\ y_3 - y_3 \\ z_3 - z_3 \end{bmatrix} \tag{2}$$

The vector vs is a vector obtained from the difference between the hip-centre joint and the spine joint. The vector v_h is a vector obtained from the difference between the hip-right joint and hip-left joint. These two vectors form a plane S = span(v_s, v_h) with the hip-centre joint as the origin, as illustrated in Figure 3.

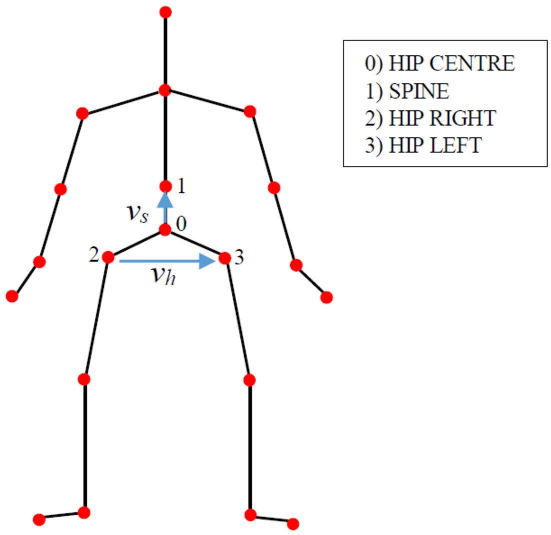


Figure 3: Illustration of vectors v_s and v_h

However, vectors vs and vh may not be perpendicular to each other. Gram-Schmidt process is applied to obtain an orthogonal basis $\{\bar{v}_s, \bar{v}_h\}$ that also spans the plane S. This new basis can be obtained by equations (3) and (4).

$$\bar{v}_s = v_s \tag{3}$$

$$\bar{v}_h = v_h - \frac{\overline{v}_s \cdot v_s}{\overline{v}_s \cdot \overline{v}_s} \cdot \overline{v}_s \tag{4}$$

Since Kinect absolute coordinates are 3- dimensional data and to also ensure that all frames are facing the same direction, a third vector is needed and is defined by equation (5).

$$\overline{v}_d = \overline{v}_s \times \overline{v}_h \tag{5}$$

However, vectors vh , vs , vd may not be unit vectors. Normalization of all three vectors are done to obtain an orthonormal basis, B = $\{u_h, u_s, u_d\}$. The normalization is needed so that all frames have the same units. This can be done by equations (6), (7) and (8).

$$u_h = \frac{\overline{v}_h}{|\overline{v}_h|} \tag{6}$$

$$u_s = \frac{\overline{v}_s}{|\overline{v}_s|} \tag{7}$$

$$u_d = \frac{\overline{v}_d}{|\overline{v}_d|} \tag{8}$$

Where $|\cdot|$ is the norm of a vector. This new set of basis $B = u_h, u_s, u_d$, as illustrated in Figure 4, spans the original frame. It allows us to compare each frame of a walking sequence to others with the same units and direction.

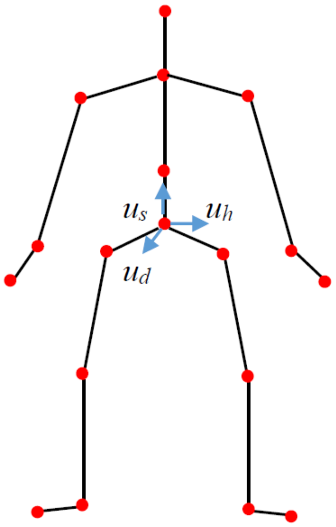


Figure 4: Illustration of $B = u_h, u_s, u_d$

3.2.2 Calculating CoB relative coordinates:

Suppose the Kinect coordinate of a joint is (x, y, z). A relative coordinate $(c_1, c_2, c_3)B$ of this joint with respect to an orthonormal basis B, defined above, can be obtained by solving the equation (9).

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{bmatrix} = c_1 u_h + c_2 u_s + c_3 u_d \tag{9}$$

The CoB relative coordinate of a joint, $(c_1, c_2, c_3)B$, represents a scenario when the origin of a 3-dimensional frame is always at the hip-centre joint, and the observation angle and the unit length coordinates of all frames are the same. Figure 5 shows an example of skeletons from Kinect coordinates (mirror) in the left column and CoB relative coordinates in the right column. Notice that, in the top two frames in the left column, a subject walks toward the front then changes direction to the left in the third frame. In the right column, skeletons are observed from the same angle in all frames.

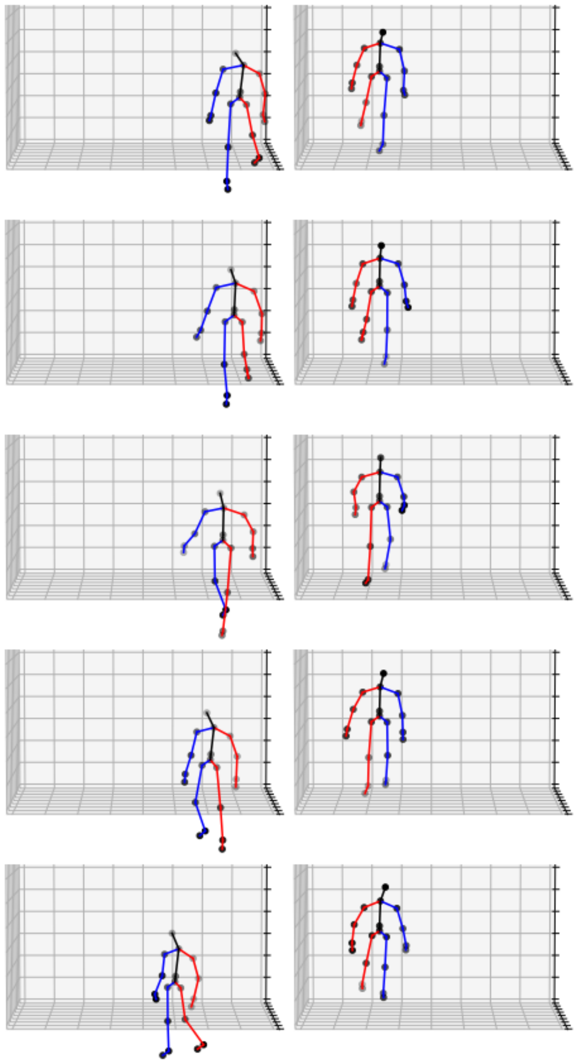


Figure 5: An example of side-by-side skeletons from Kinect coordinates (left column) and CoB relative coordinates (right column)

3.2.3 Feature Extraction

In this work, algorithm used by Nirattaya Khamsemanan et al. in their work has been used for feature extraction [10]. The Kinect version-1 produces skeletal model of the human body that consists of 20 joints as shown in the figure 6. Each joint is represented by 3 dimensional coordinates thus there are 20 coordinates corresponding to 20 joints.

By transforming the Kinect coordinates into the CoB relative coordinates, each frame represents the posture of a human subject rotated to face the same direction. Thus, each frame provides two types of information: dimensions of the human subject's body parts, and positions of the body parts. These two types of information provide meaningful and unique features

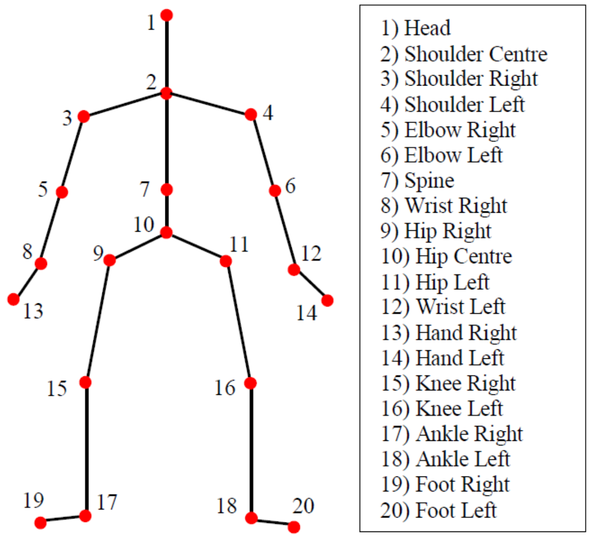


Figure 6: Kinect Skeletal Model

to identify the human subject. The dimensions of the body parts vary among human subjects. The positions of the body parts form a posture that is often unique. Hence, the CoB coordinates of frame are directly used as features.

However, each frame does not provide the information related to the body movement of the human subject. Therefore, 20 features have been added to show the movement. These features are the Euclidean distances of the same joint from one frame to its next frame. These features show how the joints move when the subjects walk.

Let $w = \{f_1, f_2, \ldots, f_n\}$ represent the a walking sequence, composed of n frames. Each frame is a sequence of 20 joints, i.e. $f_i = \{j_{i,1}, j_{i,2}, \ldots, j_{i,20}\}$. Each joint is a 3-dimensional CoB relative coordinate showing the position of the joint. $n-1$ posture-based feature vectors are generated. Each vector $x_i = [x_{i,1}, x_{i,2}, \ldots, x_{i,80}]$ is composed of 80 features. The first 54 features are from the 3-dimensional CoB relative coordinates of the 20 joints of the frame f_i..

The other 20 features are from the Euclidean distances showing the joint movements, i.e.

$$x_{i,60+k} = |j_{i,k} - j_{i+1,k}| \tag{10}$$

3.2.4 Weighted KNN

From past research works, it can be said that KNN is the best choice classifier for gait recognition [8], [9], [10]. The idea of KNN can be extended such that feature vector, among k nearest neighbours which is particularly closer to the testing feature vector, should get a higher weight in the decision than such neighbours that are far away from the testing feature vector [11]. This modified version of KNN is known as Weighted-KNN. The Weighted-KNN has produced more accurate result than simply KNN in many research works [11], [12] and [13].

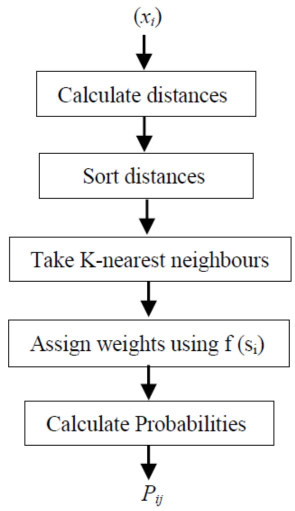


Figure 7: Weighted KNN steps

For transforming distances to weights, Weighted-KNN uses a function $f(\cdot)$. Since $f(\cdot)$ assigns higher weight to neighbour closer to testing feature vector, it must be monotonously decreasing function. Thus the function $f(\cdot)$ has following properties:

1) $f(s_i) \geq 0$, for all s_i

2) $f(s_i)$ gets its maximum for $s_i = 0$

3) $f(s_i)$ descents monotonously for $s_i \rightarrow \infty$

for i= 1, . . . , k

Where s_i is distance between testing feature vector and i$^{\text{th}}$ neighbour and R is the set of real numbers. Theoretically, there are infinite number of function $f(s_i)$ following the above three properties. The three functions mentioned below covers almost all the transformation possible and have been used in our work.

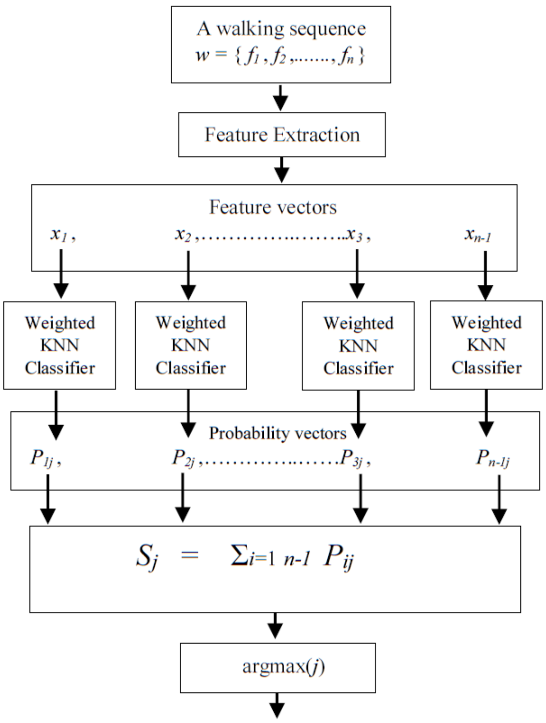
$$\begin{aligned} &1)\ f(s_i) = 1/(1+as_i) \\ &2)\ f(s_i) = 1-as_i{}^2 \\ &3)\ f(s_i) = 1-as_i \end{aligned} \tag{11}$$

Where the minimum value of parameter 'a' is Zero and its maximum value is selected so that weight is always positive. It can be noticed that if $a = 0$, $f(s_i)$ equals to 1 and the Weighted-KNN algorithm is reduced to KNN algorithm.

Let xi be the ith feature vector extracted from a walking sequence w with length of n and $C = \{C_1, C_2, \dots, C_m\}$ be the set of all classes. Then the Weighted-KNN classifier gives vector of class probabilities, $P_{ij} = P(x_i \in C_j)$ where $j = 1, 2, \dots, m$, when the input to the classifier is feature vector x_i. Probabilities P_{ij} are calculated using the weights generated by transformation function $w_i = f(s_i)$ as given by the equation (12).

$$P_{ij} = P(X_i \in C_j) = \frac{\sum_{w_i \in C_j} W_i}{\sum_{i=1}^{k} W_i} \tag{12}$$

For final prediction, after generating probabilities $P(X_i \in C_j)$ for all i and j, the probabilities belonging to class j are added to obtain sum S_j and the class with maximum sum value is the final output. The concept is shown in the figure 8. The figure also shows the use of posture based feature extraction and classification from the work of Nirattaya Khamsemanan et al. (2018).



(Class with maximum S_j)

Figure 8: Final Prediction Process

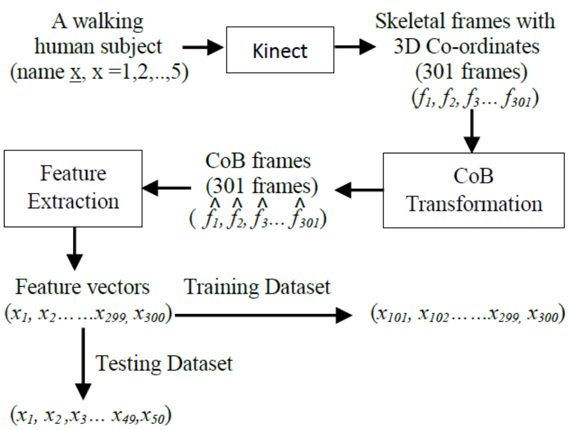


Figure 9: Dataset A creation procedure

4. Experiments and Results

4.1 Gait Dataset

To carry out the experiment, dataset of five human subjects was used. Two types of dataset, Dataset A (self-created) and Dataset B (created by other researcher) were used. Both datasets consist of sequence of feature vectors extracted from the sequence of the frames $(f_1, f_2, f_3, \dots, f_n)$, generated by Kinect. The dataset A creation process is illustrated in figure 9. The sequence in Dataset A $\{x_1, x_2, x_3, \dots, x_{49}, x_{50}\}$ was further divided into five groups.

Dataset B is dataset of five persons (Person020-Person024) created from Kinect dataset of V. O. Andersson and R. M. Araujo [8]. Dataset B was used varying number of frames from 50 to 200 in testing and 200 to 800 in training.

4.2 Output

The experiment was performed with the dataset A for all three transformation functions f (.) and varying the value of parameters 'K' and 'a'. The best result came from the transformation function $f(s_i) = 1 - as_i$ and is shown in the table 1. When $K = 1$ or $a = 0$, the weighted-KNN algorithm reduces to simple KNN algorithm, thus the first raw and first column values are output for simple KNN.

Table 1: Identification Accuracy (Dataset A)

a \ K	1	3	5	7	9	11	13	15
0	60%	64%	68%	64%	68%	60%	60%	64%
0.1	60%	64%	68%	64%	64%	60%	64%	64%
0.2	60%	64%	68%	68%	68%	60%	64%	64%
0.3	60%	64%	68%	68%	68%	60%	64%	64%
0.4	60%	64%	68%	64%	64%	56%	68%	56%
0.5	60%	60%	68%	68%	60%	60%	60%	60%
0.6	60%	64%	64%	64%	56%	60%	60%	60%
0.7	60%	64%	68%	68%	64%	60%	60%	60%
0.8	60%	60%	80%	72%	68%	64%	64%	64%
0.9	60%	68%	68%	72%	76%	72%	64%	56%
1	60%	72%	72%	60%	64%	68%	56%	52%

Tables 2 to 5 shows output for dataset B for varying number of frames (N).

Table 2: Identification Accuracy (N = 50, Dataset B)

a \ K	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	84%	84%	84%	84%	84%	84%	84%	84%	84%	84%	84%
3	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
5	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
7	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%

Table 3: Identification Accuracy (N = 75, Dataset B)

a \ K	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
3	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
5	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
7	88%	88%	88%	88%	92%	92%	88%	88%	88%	88%	88%

Table 4: Identification Accuracy (N = 100, Dataset B)

a \ K	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
3	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%	88%
5	88%	88%	88%	92%	88%	88%	88%	88%	88%	88%	88%
7	92%	88%	88%	88%	92%	92%	92%	92%	92%	92%	92%

Table 5: Identification Accuracy (N = 150, Dataset B)

a \ K	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%
3	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%
5	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%
7	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%	92%

5. Conclusion

From the output of the experiment, a curve can be plotted between accuracy and number of frames as shown in figure 10. It can be concluded from the curve that Accuracy (WKNN) $\geq$ Accuracy (KNN) for given number of frame (N) and W-KNN requires comparatively less number of frames to produce result at given accuracy.

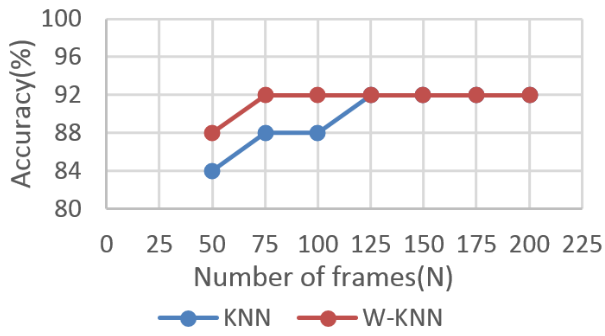


Figure 10: Accuracy curve

References

[1] J. E. Boyd and J. J. Little, “Biometric gait recognition,” in Advanced Studies in Biometrics. Springer, 2005, pp. 19–42.

[2] Darko S. Matovski, Mark S. Nixon, John N. Carter,“ Gait recognition”, Computer Vision, Springer, US, 2014, pp. 309–318.

[3] P.J. Phillips, S. Sarkar, I. Robledo, P. Grother, K. Bowyer, “The gait identification challenge problem: data sets and baseline algorithm”, 2002 Proceedings. 16th International Conference on Pattern Recognition, 2002, pp. 385—388

[4] Microsoft Corporation. (2017, March 5). Kinect for Windows. [Online]. Available: https://developer.microsoft.com/en-us/windows/kinect.

[5] J. Preis, M. Kessel, M. Werner, and C. Linnhoff-Popien, “Gait recognition with Kinect,” in 1st International Workshop on Kinect in Pervasive Computing. New Castle, UK, 2012, pp. P1–P4.

[6] A. Cheewakidakarn, N. Khamsemanan, and C. Nattee, “View independent human identification by gait analysis using skeletal data and dynamic time warping,” in Proceedings of the 14th International Symposium on Advanced Intelligent Systems (ISIS), 2013.

[7] N. Jianwattanapaisarn, A. Cheewakidakarn, N. Khamsemanan, and C. Nattee, “Human identification using skeletal gait and silhouette data extracted by Microsoft Kinect,” in Proceedings of the Joint 7thInternational Conference on Soft Computing and Intelligent Systems (SCIS 2014) and 15th International Symposium on Advanced Intelligent Systems (ISIS 2014), 2014, pp. 410–414.

[8] V. O. Andersson, and R. M. Araujo, "Person Identification Using Anthropometric and Gait Data from Kinect Sensor," in Proceedings of the Twenty- Ninth AAAI Conference on Artificial Intelligence, January 2015, pp. 425–431.

[9] K. Yang, Y. Doe, S. Lv, F. Zhang, and Q. Lv, "Relative distance features for gait recognition with Kinect," Journal of Visual Communication and Image Representation, vol. 39, pp. 209–217, 2016.

[10] Nirattaya Khamsemanan, Cholwich Nattee and Nitchan Jianwattanapaisarn "Human identification from freestyle walks using posture-based gait feature", IEEE Transactions on Information Forensics and Security, Vol. 13, No. 1, January 2018, pp. 119-128.

[11] K. Hechenbichler and K. Schliep "Weighted k-Nearest-Neighbor Techniques and Ordinal Classification", https://www.researchgate.net/, 2004.

[12] Xuesong Yan, Wei Li, Wei Chen, Wenjing Luo, Can Zhang, Qinghua Wu and Hammin Liu, "Weighted K-Nearest Neighbor Classification Algorithm Based on Genetic Algorithm", TELKOMNIKA Vol. 11, No. 10, October 2013 : 6173 – 6178.

[13] Ming Zhao and Jingchao Chen, "Improvement and comparison of Weighted K Nearest Neighbors classifiers for model selection", Journal of Software Engineering, 10(1): 109-118, 2016.