

GAIT RECOGNITION USING KINECT AND LOCALLY LINEAR EMBEDDING

D. Mustapha¹, A. Zabidi¹, R. Sahak¹, N. M. Tahir¹, I. M. Yassin^{1,*}, F. H. K. Zaman¹, Z. I. Rizman³, M. Karbasi² and M. M. M. Zan¹

¹Faculty of Electrical Engineering, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

²Kulliyah of Information and Communication Technology, International Islamic University Malaysia, 50728 Kuala Lumpur, Malaysia

³Faculty of Electrical Engineering, Universiti Teknologi MARA, 23000 Dungun, Terengganu, Malaysia

Published online: 10 September 2017

ABSTRACT

This paper presents the use of locally linear embedding (LLE) as feature extraction technique for classifying a person's identity based on their walking gait patterns. Skeleton data acquired from Microsoft Kinect camera were used as an input for (1). Multilayer Perceptron (MLP) and (2). LLE with MLP. The MLP classification accuracy result was used for comparison between both. Several MLP and LLE properties were tested to find the optimal number of setting that can improve the MLP performance. Based on the two methods used, the neural network implemented with LLE showed the better accuracy compared to the neural network alone.

Keywords: locally linear embedding; neural network; multilayer perceptron.

Author Correspondence, e-mail: ihsan.yassin@gmail.com

doi: <http://dx.doi.org/10.4314/jfas.v9i3s.58>



1. INTRODUCTION

Determining the identity of an individual is an important task in many fields such as law enforcement, counter-terrorism, fraud prevention and access control. Many research has been done in the field of identity verification in order to discover a fast, accurate, unobtrusive and computationally inexpensive method to verify a person's identity [9].

Biometrics is a study of unique, measurable characteristics of the human body that can be automatically be used to verify a person's identity. There are many types of biometric identification methods, among them are fingerprint, retina, and gait [9].

Gait analysis is defined as a method to determine a person's identity based on their walking patterns. It is an unobtrusive method for verifying a person's identity as it does not require the person to become in contact with the biometric device. Furthermore, advancements in computation and image processing has led to significant progress in this field.

In this paper, we present a Microsoft Kinect based gait recognition system. The system uses an infrared depth sensor in the Kinect device to generate a skeleton map consisting of 20 joint locations. We then apply the Locally Linear Embedding (LLE) algorithm to (1) transform the skeleton map for improved recognition performance, and (2) reduce the features used to represent the skeleton data. The contribution of this paper is the application of LLE for feature enhancement of the skeletal data prior to being classified by a Multi-Layer Perceptron (MLP) [12] classifier. We demonstrate the effectiveness of the LLE algorithm to improve the classification accuracy of the MLP, while reducing the number of features required to represent the data.

The remainder of this paper is organized as follows: Section II presents a review of some relevant research. This is followed by a description of the methodology in section III. Results and discussions are presented in section IV. Finally, concluding remarks are presented in section V.

2. LITERATURE REVIEW

2.1. Relevant Works

In [6], an identity verification system using Microsoft Kinect was presented. Features were extracted using distances measured from the upper body centroid relative to upper and lower

limb coordinates, Classification was performed using MLP neural network.

In [5] explored the capability of the MLP classifier to recognize differences in walking patterns between normal subjects and subjects suffering from Parkinson's Disease (PD). The research showed that the MLP was able to successfully identify the gait patterns between normal subjects and PD patients with 87.5% accuracy.

In [3], a study on limb movement during normal and fast walking was presented under different walking conditions. All of the walking conditions tested showed different gait cycles but with similar space extension in the walking cycle. Moderate and fast walks were found to have comparable element highlights in a step cycle, with reduced arm space expansion compared to those of the legs. Additionally, it was found that slope walking and even level have distinctive joint patterns.

In [4] presented a MLP based system for classification of human walking/running. Data acquisition was performed using Microsoft Kinect to capture joint data. The joint data was then used to train a MLP to differentiate between walking and running patterns. The results indicate that the MLP was able to distinguish between walking [14] and running patterns with above 90% accuracy.

In [7], a Kinect-based intelligent intrusion detection system was presented. Kinect was used to capture useful gait information of the intruder and compare it to the normal gait patterns. A MLP was trained with gait patterns of normal subjects compared with intruders. The system was able to discriminate between intruders and normal subjects with 94.4% training accuracy and 81.1% recognition accuracy under simulated conditions.

LLE has been used for human activity identification in a manifold-based framework [10], where LLE was used to extract important information from motion sensors placed on subject bodies. LLE was used to capture the inherent structure of the data and construct a nonlinear manifold for each movement type. Experiments performed showed that the LLE method outperformed ISO-map and Eigen-map techniques as it made fewer assumptions on the movement flags and reduced execution time. In [8], LLE was shown to be able to reduce dimensionality in gait analysis application.

2.2. Locally Linear Embedding (LLE)

Figure 1 presents the locally linear embedding (LLE) method which fundament is based on

simple geometric institute. In essence, the calculations attempt to resolve a low-dimensional fixing which is approximately similar to the high-dimensional features (the fundament is improved to retain local settings of the nearest neighbours so that affinity co-located with regard to other low dimensional space). Accordingly, in very good situation it is possible to get an embedding solely from the geometric situation without recourse to scale, distance or connectivity between interval data [1].

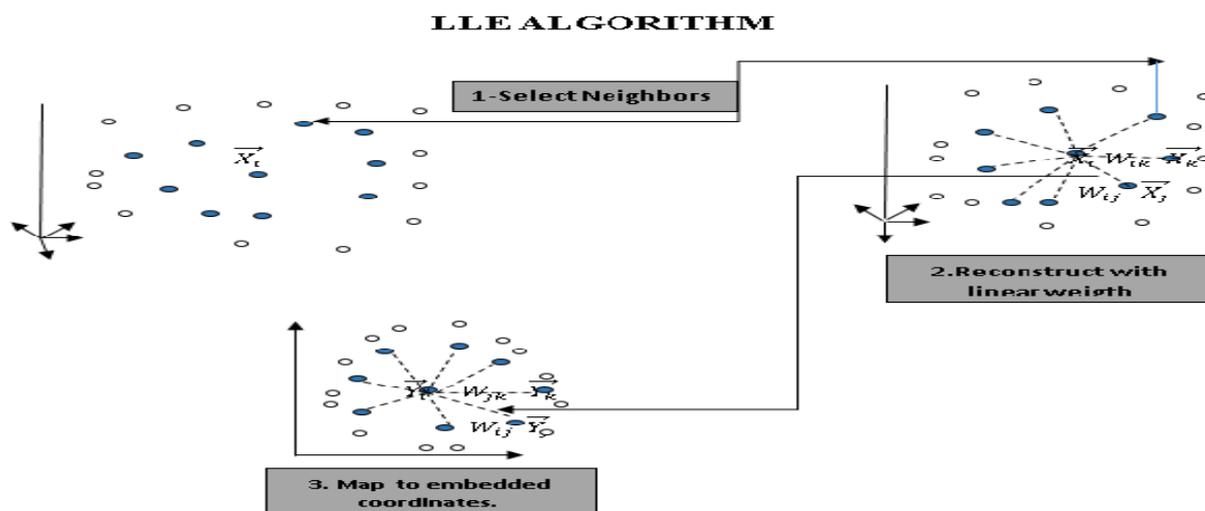


Fig.1. Three mapping high dimensional input \vec{x}_i to low dimensional outputs \vec{y}_i via local linear reconstruction weights W_{ij}

Summary of the LLE algorithm

1. Compute the neighbours of each data point, \vec{x}_i .
2. Compute the weight W_{ij} that best reconstructs each data point \vec{x}_i from its neighbors, minimizing the cost in equation 3.21 by constrained linear fits.
3. Compute the vectors \vec{y}_i best reconstructed by the weights W_{ij} , minimizing quadratic form in Equation 1 by its bottom nonzero eigenvectors.

$$E(W) = \sum_i |\vec{x}_i - \sum_j W_{ij} \vec{x}_j|^2 \tag{1}$$

To sum up the square distance between entire data points and their reconstruction of the weight W_{ij} , the contribution of the j th data point to the i th reconstruction is highlighted. To calculate the weight, cost function will be minimized in the equation above items to two controllable sections: a small controllable section and an invariance controllable section. The

small controllable section is where each data point \vec{X}_i is reconstructed only from its neighbours, enforcing $W_{ij}=0$ if \vec{X}_j does not belong to this set. The invariance controllable section is the chain of weight matrix added to one:

$$\sum_j W_{ij} = 1 \tag{2}$$

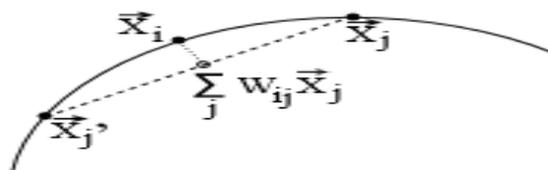


Fig.2. A data point \vec{X}_i , its neighbours \vec{X}_j and its locally linear reconstruction $\sum_j W_{ij} \vec{X}_j$. The

reconstruction weights are constrained to satisfy $\sum_j W_{ij} = 1$

The goal is to find low dimensional outputs \vec{Y}_1 that are reconstructed by the same weights W_{ij} as the high dimensional inputs \vec{X}_1 . The embedding cost function in the Equation 3 defines a quadratic form in the outputs \vec{Y}_1 . Subject to constraints that make the problem well-posed, the cost function has a unique global minimum. This unique solution for the outputs \vec{Y}_1 is the result returned by LLE as the low dimensional embedding of the high dimensional inputs \vec{X}_1 . The embedding cost function can be minimized by solving a sparse $N \times N$ eigenvalue problem.

$$\Phi(Y) = \sum_i |\vec{Y}_i - \sum_j W_{ij} \vec{Y}_j|^2 \tag{3}$$

3.METHODOLOGY

Data acquisition was implemented considering indoor situations (white penda flour lighting). The Kinect device was configured at zero-degree elevation. Subjects are required to wear long pants and cover shoes during data acquisition. A total of 10 different subjects were instructed to walk at normal pace ten times towards the Kinect device. Therefore, the total number of samples was 100. The Kinect device was able to acquire skeletal data at approximately 30 frames per second. For each of the 100 samples, the number of frames acquired was 78.

After the data was collected, we trained two different MLPs. MLP1 was trained using the raw skeletal data, while the MLP2 was trained using skeletal data preprocessed using the LLE

method. Table 1 shows the parameters used to train the MLPs. The number of hidden units, LLE number of neighbours (k) and maximum embedding dimensionality (dmax) were varied to discover the optimal parameters to obtain the best accuracy. As both MLPs were used for pattern classification, the hidden and output activation functions were set to Tangent-Sigmoid, and the training algorithm used was the Scaled Conjugate Gradient algorithm, which has been proven to be optimal for pattern classification tasks [2].

In order to avoid overfitting, the Early Stopping algorithm was used. The algorithm divides the dataset into training, validation and testing datasets. During training, the training set was used to guide the optimization [15] of weights in both MLPs, while the validation set was used to monitor the training progress. The errors for the two datasets are expected to go down when training progresses. When overfitting occurs, the validation error would start to increase, while the training error would still decrease. It is during this time the MLP is deemed to over fit, and training is automatically stopped [2].

Table 1.MLP1 and MLP2 training parameters

Parameter	MLP1	MLP2
Number of Hidden Units	Varied between 5 to 30.	
Hidden Activation Function	Tangent-Sigmoid	
Output Activation Function	Tangent-Sigmoid	
Overfitting Avoidance Technique	Early Stopping Algorithm	
Training Algorithm	Scaled Conjugate Gradient Algorithm	
Input Features Used	Raw skeletal data	Skeletal data preprocessed using LLE. Two parameters for LLE, dmax (maximum

embedding dimensionality) and k (number of neighbours). Many combinations of k and d_{max} was tested ($k=5$ to $k=95$ and $d_{max}=5$ to $d_{max}=95$). These parameters have significant influence on classification accuracy as it affects how the data is transformed.

Dataset Division

70 (training) : 15 (validation) : 15 (testing)

4. RESULTS AND DISCUSSION

Fig. 3 and Fig. 4 shows the classification accuracy versus the number of MLP hidden units. The accuracy changes when the number of hidden unit is adjusted. The accuracy range of training data was from 63.3% to 100%, while the testing accuracy ranged from 33.3% to 93.3%. The testing accuracy was generally lower compared to the training set. This observation is expected since the testing set consist of data previously unseen during training. Because of this, the MLP is more likely to misclassify data inside the testing set.

The best MLP1 accuracy was obtained when the number of hidden unit was 25. Because of this the 25 hidden unit was chosen as the baseline for comparison with features train using LLE. The confusion matrix in Figure 5 and Figure 6 show the classification performance for 25 hidden units for MLP1. Using 25 hidden units, MLP1 managed to obtain 100% training accuracy and 93.3% testing accuracy.

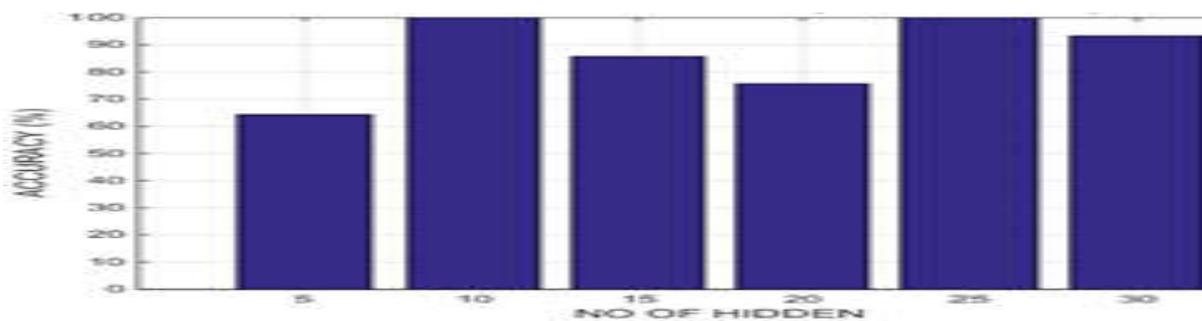


Fig.3. No of hidden unit vs training accuracy for MLP1 (without LLE)

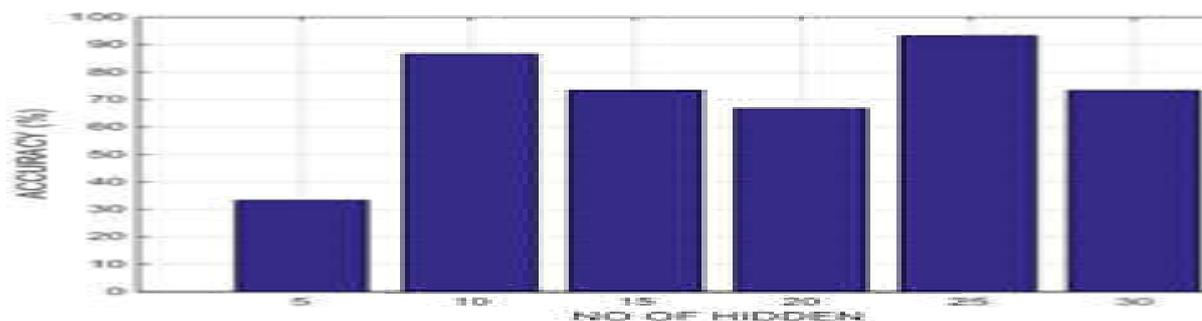


Fig.4. No of hidden unit vs testing accuracy for MLP1 (without LLE)



Fig.5. Training set confusion matrix for MLP1 (without LLE)



Fig.6. Testing set confusion matrix for MLP1 (without LLE)

Once the best MLP structure [11] was obtained, the setup were used in the second experiment

(MLP2). LLE was used to improve the skeletal features used in the previous MLP1 experiment.

LLE depends on two parameters, which are no of Neighbors (k) and Maximum Embedding Dimensionality (dmax). A brute-force method was used to discover the optimal settings for k and dmax, ranging from k=5 to k=95 and dmax=5 to dmax=95. The k and dmax parameters listed in Table 2 obtained the best MLP2 classification results.

Table 2. Best four LLE parameter combinations for MLP2

	k=55, dmax=50	k=62, dmax=40	k=70, dmax=50	k=90, dmax=50
Training accuracy (%)	100	100	100	100
Testing accuracy (%)	100	100	100	100

Since the four parameter combinations generated 100% accuracy for training and testing sets, k = 55 and dmax = 50 was arbitrarily chosen for further analysis. For dmax = 50, the classification accuracy versus the number of neighbors (k) is shown in Fig. 7. As can be seen, the classification accuracy for both training (blue line) and testing (red line) sets were generally high for k=5 to k=50, and the best results were obtained when k=55 and dmax =50.

In testing set, it can be seen that when number of neighbours (k) was increased up to 50, the classification accuracy was improved. But when the number of neighbour (k) were increased further (k more than 50), the classification performance showed signs of degradation. The selection of k is a tradeoff between the richness of features and classifier complexity. A small number of neighbours may be insufficient to represent critical information in dataset, while high numbers may relate to a larger feature set that may increase classifier complexity.

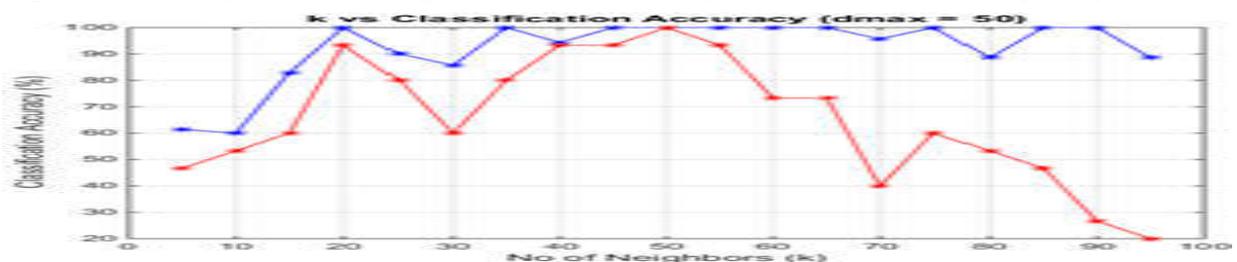


Fig.7. No of neighbors (k) vs classification accuracy (maximum embedding dimensionality, dmax = 50)

The confusion plots (Fig. 8 and Fig. 9) show the classification result for the best combination of k and dmax. Both training and testing sets registered 100% accuracy. This result indicates that when LLE was used to transform the gait features prior to classification, the results were improved by 6.7% in testing set compared to MLP1. Therefore, we can conclude that the use of LLE has a positive effect on classification accuracy (Table 3).



Fig.8. Training set confusion matrix for MLP2 (with LLE)



Fig.9. Testing set confusion matrix for MLP2 (with LLE)

Table 3. Summary of classification accuracy (with and without LLE)

Item	MLP1 (without LLE)	MLP2 (with LLE, k=55,dmax=50)
Training accuracy	100%	100%
Testing accuracy	93.3%	100%
Hidden units	25	25

5. CONCLUSION

LLE was used in combination with MLP to preprocess Kinect skeletal features for gait recognition [13]. This method has been used with many combination of k and dx parameters to obtain optimal accuracy. The use of LLE had improved the overall testing accuracy by 6.7 % over using raw features.

6. ACKNOWLEDGEMENTS

The authors would like to graciously acknowledge the Ministry of Higher Education and UniversitiTeknologi MARA for supporting this research work through Grant No: 600-RMI/DANA 5/3/ARAS(4/2015).

7. REFERENCES

- [1] Saul L K, Roweis S T. Think globally, fit locally: Unsupervised learning of low dimensional manifolds. *Journal of Machine Learning Research*, 2003, 4:119-155
- [2] Demuth H, Beale M. *Neural network toolbox for use with Matlab: Computation, Visualization, Programming: User's guide version 4*. Massachusetts: The MathWorks Inc., 2000
- [3] Honggui L, Xingguo L. Gait analysis using LLE. In *7th IEEE International Conference on Signal Processing*, 2004, pp. 1423-1426
- [4] Maierdan M, Watanabe K, Maeyama S. Human action recognition based on the angle data of limbs. In *15th IEEE International Symposium on Soft Computing and Intelligent Systems Joint 7th International Conference on and Advanced Intelligent Systems*, 2014, pp. 140-144
- [5] Manap H H, Tahir N M, Yassin A I, Abdullah R. Anomaly gait classification of parkinson

disease based on ANN. In IEEE International Conference on System Engineering and Technology, 2011, pp. 5-9

[6] Sinha A, Chakravarty K, Bhowmick B. Person identification using skeleton information from kinect. In International Conference on Advances in Computer-Human Interactions, 2013, pp. 101-108

[7] Wei Z, Chakraborty G. Construction of intelligent intrusion detection system based on KINECT. In IEEE International Joint Conference on Awareness Science and Technology and Ubi-Media Computing, 2013, pp. 81-87

[8] Xu J, Ding C, Cong W, Li J, Wang Y. Gait recognition based on LLE. In IEEE International Conference on Measurement, Information and Control, 2012, pp. 825-828

[9] Yassin A I. Face detection using multilayer perceptrons trained on min-max features and optimized using particle swarm optimization. In 3rd International Conference on Signal Processing and its Application, 2007

[10] Zhang M, Sawchuk A A. Manifold learning and recognition of human activity using body-area sensors. In 10th IEEE International Conference on Machine Learning and Applications and Workshops, 2011, pp. 7-13

[11] Azlee Z, Nooritawati M T, Ihsan M Y, Zairi I R. The performance of binary artificial bee colony (BABC) in structure selection of polynomial NARX and NARMAX models. International Journal on Advanced Science, Engineering and Information Technology, 2017, 7(2):373-379

[12] Ihsan M Y, Azlee Z, Rozita J, Megat S A M A, Rahimi , Abu H A H, Zairi I R. Comparison between cascade forward and multi-layer perceptron neural networks for NARX functional electrical stimulation (FES)-based muscle model. International Journal on Advanced Science, Engineering and Information Technology, 2017, 7(1):215-221

[13] Che Z C H, Rozita J, Nooritawati M T, Ihsan M Y, Zairi I R. Automated classification of autism spectrum disorders gait patterns using discriminant analysis based on kinematic and kinetic gait features. Journal of Applied Environmental and Biological Sciences, 2017, 7(1):150-156

[14] Nor M N M, Rozita J, Nooritawati M T, Ihsan M Y, Zairi I R, Rahmat H. EMG signals analysis of BF and RF muscles in autism spectrum disorder (ASD) during walking.

International Journal on Advanced Science, Engineering and Information Technology, 2016, 6(5):793-798

[15] Ihsan M Y, Azlee Z, Megat S A M A, Nooritawati M T, Hasliza A H, Husna Z A, Zairi I R. Binary particle swarm optimization structure selection of nonlinear autoregressive moving average with exogenous inputs (NARMAX) model of a flexible robot arm. International Journal on Advanced Science, Engineering and Information Technology, 2016, 6(5):630-637

How to cite this article:

Mustapha D, Zabidi A, Sahak R, Tahir N M, Yassin I M, Zaman F H K, Rizman Z I, Karbasi M, Zan M M M. Gait recognition using kinect and locally linear embedding. *J. Fundam. Appl. Sci.*, 2017, 9(3S), 755-767