A Hybrid Approach to Human Posture Classification during TV Watching

Jonathan H. Chan[1*], Thammarsat Visutarrom[1], Sung-Bae Cho[2], Worrawat Engchuan[1],

Pornchai Mongolnam[1], and Simon Fong[3],

[1] Data and Knowledge Engineering Laboratory,

School of Information Technology, King Mongkut's University of Technology Thonburi, Bangkok, Thailand

jonathan@sit.kmutt.ac.th, thammarsat@gmail.com, worrawat.eng@sit.kmutt.ac.th, pornchai@sit.kmutt.ac.th

[2] Department of Computer Science, Faculty of Engineering

Yonsei University, Seoul, Korea

sbcho@cs.yonsei.ac.kr

[3] Department of Computer and Information Science, Faculty of Science and Technology

University of Macau, Taipa, Macau SAR

ccfong@umac.mo

[*]Corresponding author's email: jonathan@sit.kmutt.ac.th

Abstract. Human posture classification in near real time is a significant challenge in various fields of research. Recently, the use of the Microsoft Kinect system for 3D skeleton detection has shown to be of promise. This work compares four common classifiers and the use of a hybrid approach for classification. The results show that the use of a hybrid genetic algorithm and random forest classifier is able to provide fast and robust human posture classification. Finally, to aid in further development of posture detection, a comprehensive human posture data set while watching television has been generated in this work for benchmarking purpose and made available publicly at http://dlab.sit.kmutt.ac.th/index.php/human-posture-datasets.

Keywords: hybrid approach; human posture classification; Kinect; random forest; genetic algorithm; television watching; benchmarking

1. Introduction

The identification of human posture is of significant interest in numerous disciplines of research. There is an increasing need to be able to classify human posture in near real time circumstances, especially in the fields of medical science and sport science. With the advent of rapid development in computer technology, it allows for much improved human posture classification. Since the release of the Microsoft Kinect system with its game console back in 2010, many researchers have used it for human postural classification and human body detection in general. With multiple sensors, including a depth sensor and an infrared sensor, Kinect is able to identify and separate the human body even from a complex background. Currently, Kinect also offers a computer-based software library with built-in support to identify skeletal joints, thus allowing the whole skeleton to be analysed. In particular, Kinect can detect up to 60 raw attributes of a human user in each static image or dynamic video frame. That is, there are a total of 20 skeleton joints with each having three positional coordinates X, Y and Z. (See Fig. 1.)

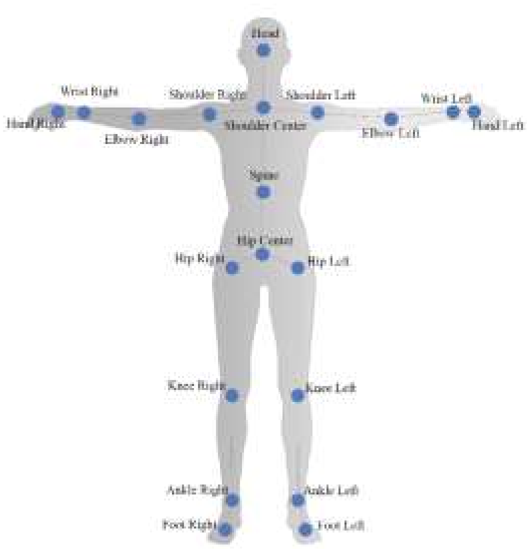


Figure 1. Position of the twenty skeleton joints from Kinect.

The detection of human posture may be done either in 2D or 3D[1-3]. Traditionally, 2D image processing is done using pixel colour information. That is, human posture classification would depend heavily on environmental factors such as light conditions and background complexity. The use of 3D image processing can reduce the background complexity issue that constraints 2D images. The main reason is that 3D image processing is able to provide depth information to better identify each object. For the Kinect system, it combines 3D image processing with infrared sensors that could be used in even dark areas, thus making it ideal for complex human postural analyses.

There has been many previous works to detect and classify human posture. Prior to the introduction of the Kinect system, Cohen and Li[4] developed a SVM model for human postural classification using a 3D visual-hull constructed from a set of silhouette input data. The system provided classified human body postures in form of thumbnail images. In 2007, Wu and Aghajan[5] proposed a method of human posture estimation using multiple cameras based on the concept of an opportunistic fusion framework that comprised three dimensions of space, time, and feature levels to obtain a 3D human skeleton.

The Kinect system[6-9] was launched by Microsoft in 2010. It was designed to be a 3D camera game controller to determine the human gestures of a player. Kinect has three input components: depth image sensors, RGB camera, and a multi-array microphone. With the ability to separate the human body from a complex background and in analysing the human skeletal joints, Kinect has been used by many researchers from the onset. Initially, the OpenNI software library was developed to interface a computer system to the Kinect. Subsequently, a Microsoft SDK was also made available as a software library for analysing skeletal joints. The OpenNI library can provide up to 15 main skeleton joints of a body. The latest Microsoft SDK library provides two detection modes. The default is stand tracking that can analyse up to 20 joints of the human body. The second mode is for seated tracking that can analyse 10 upper shoulder joints of the human body.

Since the launch of the Kinect system, much works in postural detection have been done using such a system. A representative set of works is presented here. In 2011, Htike and Khalifa[10] developed a real-time gesture classification system to classify the dancing gestures from moving skeletal joint data obtained from Kinect. The accuracy of their system was 96.9% using a 4-second video capture record of the human skeletal motion. More recently, Dai et al. [11] proposed a machine learning and vision-based method for elderly fall detection using statistical human posture sequence modelling. A series of laboratory simulated falls and activities of daily livings (ADLs) were performed and recorded by the Kinect system. Hidden Markov Model was used for modelling the fall posture sequences and distinguishing different fall activities and ADLs. The average fall recognition rate was greater than 80%.

Unlike most other researchers, we use Kinect to classify human postures while watching television. In an earlier work, a system for simple postural detection and classification of elderly people while watching television was developed[12]. The experiments included four standard postures of stand, sit, sit on floor and lie down. A total of six models were used to compare the results of postural classification and the best classifier was selected by using 5-fold cross-validation. The six classification methods used were neural network, support vector machine, decision tree, logistic regression, random forest and naïve Bayes. The best accuracy was 97.88%, obtained by decision tree with Max-Min normalization technique. As a follow up work, a real-time identification system was developed to identify 18 human postures while watching television[13]. In particular, two classification architectures was evaluated: a single-stage classifier and a multiple-stage classifier. Three types of training sets were used, including raw skeletal training set, skeleton with attributes selection training set, and skeletal position transformation training set. The machine learning techniques compared were back propagation neural network (NN), naïve Bayes (NB), logistic regression (LR), and decision tree (J48). The best performance was obtained with an accuracy of 87.68 % by using the skeletal position transformation training set with neural network.

In this work, the ready availability of real-time raw skeleton data was exploited to develop a robust human posture classification model for television watching. To this effect, a large comprehensive data set with 30 human postures in various positions was collected systematically. This data set is provided publicly at the data repository of our lab at http://dlab.sit.kmutt.ac.th/index.php/human-posture-datasets.

This paper is organized as follows. First we give a brief summary of the workflow of human posture classification used in this work. Then we provide some backgrounds on the datasets, classification methods and performance measures used in this study. The results are presented next with appropriate discussions. Finally, we draw significant conclusions on the findings.

2. Methodology

The outline of the workflow is as follows. First, the experimental layout is given. Then the human posture dataset collected is described. Followed by details of the data preparation step. Finally, the methods of data classification and description of performance measures are provided.

Kinect layout

In our experiment, the recording data came from five positions as shown in Fig. 2a. Position 1 (P1) corresponds to 0 degree from the subject, i.e. directly facing the Kinect camera. Position 2 (P2) to Position 5 (P5) corresponds to 45, 90, 135 and 180 degrees from the subject. Each position generates one data set. In each data set, there are 63 attributes that include joint positions that analyse from Kinect's 60 attributes (20 joint * 3 axes of X, Y and Z), with the additional attributes being height and angle of the camera and distance from human to the Kinect system.

(a)

(b)

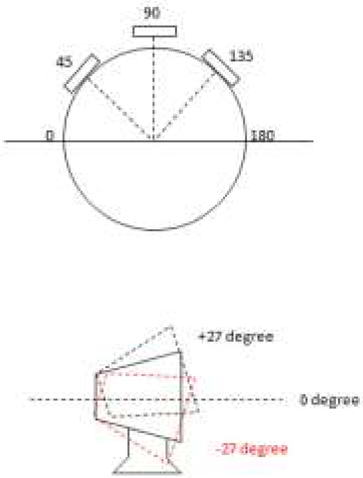


Figure 2. Recording positions and angles of the Kinect experiments.

In particular, height is distance of height from room floor to centre of Kinect camera. Three height levels of 1, 1.5, and 2 m were used. There are three camera angle that correspond with height, i.e. 1 m with 0 degree, 1.5 m. with 16 degree and 2 m with -27 degree. Kinect can rotate between + 27 degree to - 27 degree as shown in Fig. 2b.

The distance from the human subject to Kinect is measured differently depending on posture. For stand class, sit class, and sit floor class, it is measured from the foot of the subject to the base on the floor that the Kinect camera is located, as shown in Fig. 3a. For lean down (or lie down) class, it is measured from the centre of body point of room floor to that below Kinect camera, as shown in Fig. 3b.

(a)

(b)

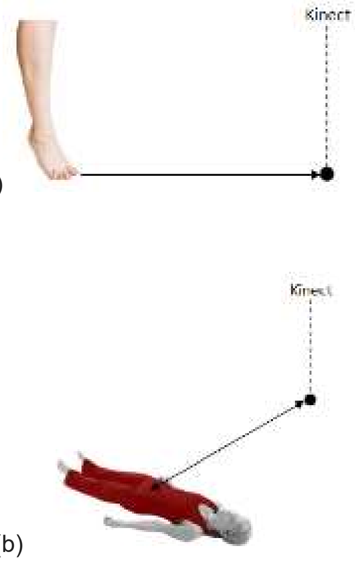


Figure 3. Configuration for measurement of distance from Kinect.

Human posture dataset

This part describes the preparation of training and testing data sets used in developing the human posture classification model. We used Kinect XBOX 360 with the Microsoft SDK library version 1.8 with the default stand tracking mode to detect a human body having 20 skeletal joints. The distance for test subjects is between 1.8-3.0 meters away from Kinect during the data collection process. This distance has been added as an additional attribute as well as height and angle of the Kinect camera.

Training and testing sets were recorded systematically as described earlier. The number of instances in the five training and five testing data sets are 13,500 each (except for two cases with 13,499 and 13,501). The details of number of instances in each posture for both training set and testing set are summarized in Table I. In summary, there are a total of 30 postural class memberships with four main class memberships of stand, sit, sit on floor and lean down.

Data preparation

This step is to prepare the training sets before being used in learning of each model. In our experiment, we considered two types of classifier architectures. First is a single-stage classifier and second is a multiple-stage classifier.

The single-stage classifier model uses all training data packaged in one set. Therefore, the single-stage classifier has only one model in the prediction data. Consequently, if the training data are large, it is very time consuming in creating the model.

On the other hand, the multiple-stage classifier would separate the data into multiple training sets, and that would help to reduce the time to create the model, depending on the number of separated training sets. Therefore, the multiple-stage classifier requires many models working together in the prediction data. For this architecture, we considered both single classifier for both stages and the hybrid approach combining genetic search algorithm with classifier.

For single-stage classification, training and testing were done with all 63 attributes initially. This is labelled as “all”. In order to test the robustness of the Kinect system, the three additional attributes of height and angle of Kinect and distance of subject to Kinect are removed. This is labelled as “60 points”. Finally, the multi-stage with hybrid of GA and classifier is labelled as “GA”.

Data classification

In this work, we select four learning models, including neural network (NN), naïve Bayes (NB), support vector machine (SVM), and random forest (RF); all of which are available in the Weka data mining tool[14]. Detail of feature selection by Genetic Algorithm (GA) and the classifiers[15-16] used in this study are described as follows.

Neural network used in the experiment is the simple multilayer perceptron (MLP) which uses the back-propagation algorithm in learning. Structure of the neural network in the experiment comprises three layers, including input layer, hidden layer, and output layer. The input and output layer values depend on the number of attributes and the number of class memberships of each training set. Naïve Bayes is a simple probability classifier based on the Bayes' theorem with the assumption of independence between every pair of features. The nodes in the Bayesian model are created from the given training data. Each node counts the number of rows per attribute value per class for nominal attributes and calculates the Gaussian distribution for numerical attributes. SVM is another popular supervised-machine learning algorithm, which builds an n-hyperplanes for n-features to separate each distinct class apart with maximal margin. Random Forest is an ensemble classification algorithm. A RF classifier was implemented with the subsampling and sub-spacing scheme. So, the RF model is robust to the over-fitting issue[17-18]. The RF have been widely and successfully applied in the classification in many fields[19-20].

The feature selection was done by WEKA 3.6.11. The feature evaluation for GA was performed by Correlation-based Feature Subset Selection (CfsSubsetEval)[21]. The population size and number of generations were set at the default value of 20.

For implementation of the classifiers, MATLAB version 2010 was used to build a neural network model with number of nodes in hidden layer equal to 10. The R language was used to build the other three classifiers. In particular, NB and SVM were applied using "e1071" package[22] while RandomForest package[23] was used to perform RF classification.

3. Results and discussion

The results of human posture classification using various approaches are shown here. This is followed by insightful interpretation and discussion of the results.

The first results are shown on Table II which provides a comparison of the classification performance using the different positions of the Kinect system to the human subject. The ensemble classifier RF was used because it has been found by various researchers to provide accurate and robust performance. As expected, we found the performance is very good if using the same position for training the RF model and subsequent testing. However, the performance dropped off substantially when a different position is used for testing. Consequently, a master data set that combines all five Kinect positions was generated for both training and testing purposes.

Due to the large size of the combined data sets, computational time requirements increased significantly especially for more vigorous classifiers such as SVM and NN. Thus, a hybrid approach using GA for attribute selection was used and compared with the single stage classification approach. The results of GA attribute selection is shown in Table III. It is interesting to note that only the Y component was selected along with the 10 upper body skeleton joints. These feature points are shown in Fig. 4.

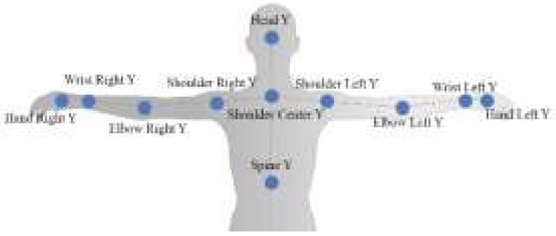
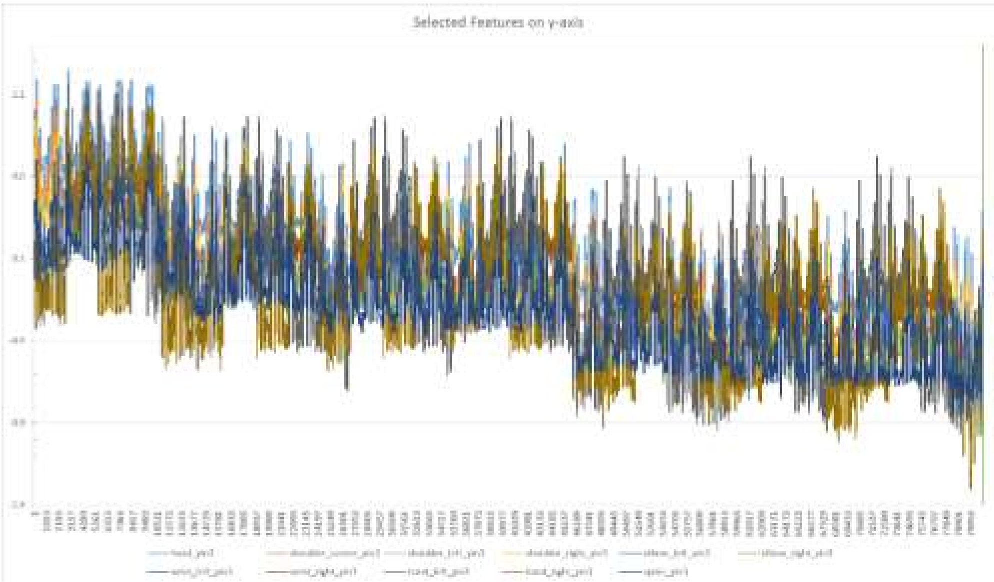
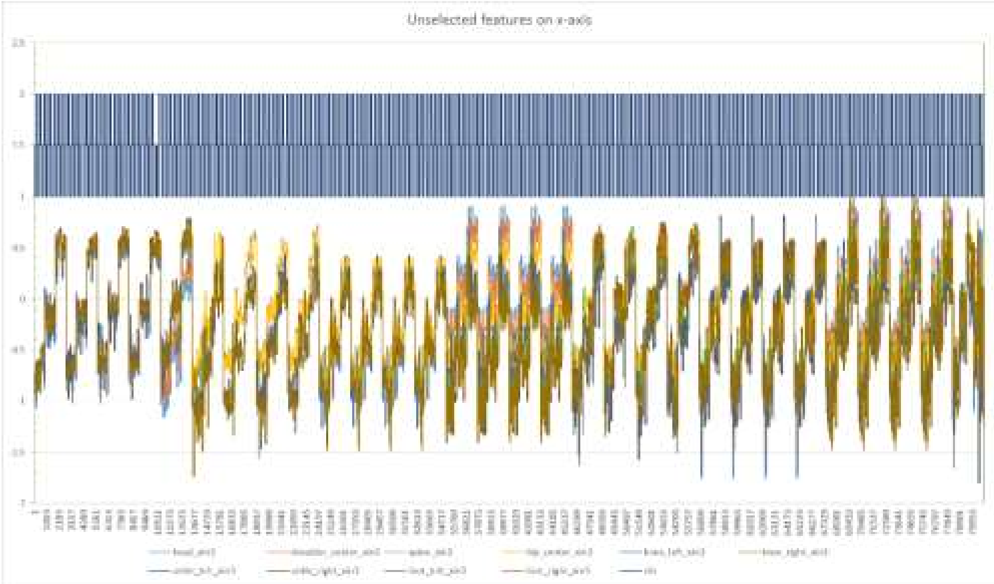


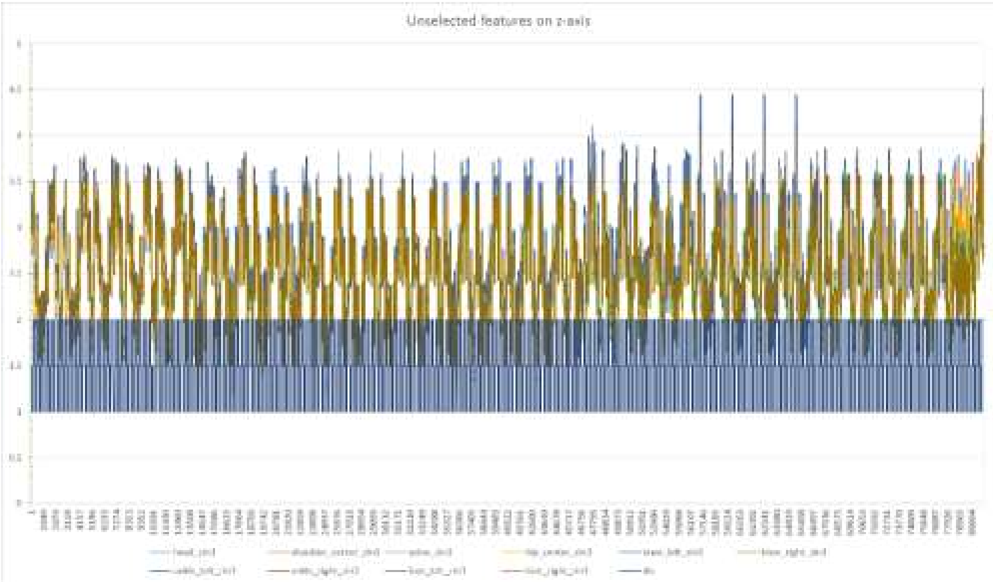
Figure 4. Selected features by Genetic Algorithm-based feature selection.

For the sake of investigating why only the Y components of the features are selected by the feature selection algorithm, three groups of features are visualized: selected features on the Y-axis, non-selected features on the X-axis and non-selected features on the Z-axis. These are displayed in Figures 5a, 5b and 5c respectively.

(a)




(b)



(c)

Figure 5. Visualization of (a) selected features on Y-axis by Genetic Algorithm-based feature selection; (b) non-selected features on X-axis; and (c) non-selected features on Z-axis.

Intuitively one can observe the characteristics of the patterns of the feature values that lead to the selection by the genetic-based feature selection algorithm. For instance, in Figure 5a, the selected feature patterns are non-stationary, and they possess certain unique shapes. These unique shapes of the patterns are relatively easier to be differentiated by pattern recognition algorithms; hence they are relevant to posture classification and were selected by the feature selection methods. On the other hand, the patterns of the non-selected features as shown in Figure 5b and Figure 5c are stationary and periodic. The feature values are more or less about the same, which do not form a prominent pattern in the context of pattern recognition.

To further reinforce this speculation, the coefficient of variations (CV) were computed for these three groups of feature values. The results are tabulated in Tables IV, V and VI respectively. The absolute values of CV are known to relate to the relevance of features with respective to producing an accurate classification model. In other words, CVs can be used as a measure in evaluating whether the features should be chosen or not in feature selection. From the Tables, it can be seen and confirmed that the selected group of features have higher CV indices compared to the other two groups. The average CV values for the selected group was 15.92, while the non-selected groups of features on the X-axis and Z-axis were 3.06 and 0.22 respectively. It essentially shows that movements along the Z-axis have little variations, thus little contribution to classifying the types of postures which are of concern in this study. To sum up, postures which are characterised by features of low degree of CV[24] are mediocre in action (or variation), and they are not easy to be distinguished by machine learning algorithms.

In addition, in order to assess the robustness of the Kinect system, the three additional attributes of height and angle of Kinect and distance from human subject were dropped to develop another reduced model. A summary of the results for these three approaches is shown in Table VII. It can be seen that RF provided consistently good performance for all three approaches. Whereas NB provided the worst performance.

A more detailed analysis of the single stage classifier results for each human posture is shown in Table VIII. The performance measure shown is Recall since it is of particular interest in the biomedical field. Note that the results are similar for other measures (not shown) such as true positive, false positive, false negative, false discovery rate, and so on. As can be seen in Table VIII, the overall performance of RF is good with most postures at 1.0 but the performance drops to as low as 0.94 for the group of human postures of sit on floor and lean back (sitfloorLB). To further improve this performance, a two-stage approach was tested using the sitfloorLB postures as the second stage. The results shown that the Recall improved to 1.0 for all postures.

Finally, in order to provide a better recommendation for human posture classification while watching television, the computational times for each classifier for single stage classification are shown in Table IX. Referring to these results, it is clear that RF provided the best performance and the hybrid GA-RF approach further improved on the speed of classifier development. Thus the hybrid GA-RF approach is a good candidate for further scale-up analysis of the big data challenge.

In summary, the ensemble classifier RF was able to provide a fast and reliable prediction of the human postures, and the hybrid GA-RF approach is promising for big data analysis[25]. Nonetheless, for practical implementation, postural transitions should also be considered in order to detect and predict accurate postures for real-time demands while watching television.

4. Conclusions

In this work, we are interested in the performance of using Kinect for human posture classification for user subjects while watching television. A comprehensive data set consisting of 30 human postures was generated using Kinect and deposited in a data repository for benchmarking purpose. The results showed that the use of RF provided the best combination of performance and computational efficiency for single stage architecture. However, there was a minor problem with the sit on floor and lean back posture. A two-stage RF was able to solve this problem. Finally, the hybrid GA-RF approach further improved the speed and robustness of classification by using only the 10 upper body skeleton points selected by GA.

For the next step, we would like to further validate our results by creating an even larger dataset from many representative subjects and then apply our methodology to the postural classification of elderly people whose main pastime is watching television. The eventual goal is be able to determine the behaviour of the elderly to assist in the interaction with their family members, physicians, and other health practitioners, thus improving the elderly's quality of life in general.

Acknowledgement

This work was supported in part by the Industrial Strategic Technology Development Program, 10044828, Development of Augmenting Multisensory Technology for Enhancing Significant Effects on the Service Industry, funded by the Ministry of Trade, Industry and Energy (MI, Korea). This work had been partially supported by the Higher Education Research Promotion and National Research University (NRU) Project, Thailand's Office of the Higher Education Commission. We also wish to acknowledge Amontep Wijitcharoen for his assistance in running some of the scripts. Finally, the authors would like to thank Katlin Kreamer-Tonin for proofreading the paper.

REFERENCES

[1] Y. Sun, M. H. Li, J. S. Hu and E. L. Wang, 2D Recovery OF Human Posture, Proceedings of the International Conference on Machine Learning and Cybernetics, **(2002)** Nov. 4-5; Beijing, China, pp. 1638 - 1640.

[2] K. Takahashi and S. Sugakawa, Remarks on human posture classification using self-organizing map, Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, **(2004)** Oct. 10-13; Netherlands, pp. 2623 - 2628.

[3] C. C. Li and Y. Y. Chen, Human Posture Recognition by Simple Rules, Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, **(2006)** April 10-12; Southampton, United Kingdom, pp. 351 - 356.

[4] I. Cohen and H. Li, Inference of Human Postures by Classification of 3D Human Body Shape, Proceedings of the IEEE International Workshop on Analysis and Modeling of Faces and Gestures, **(2003)** October 17; Nice, France, pp. 74 – 81.

[5] C. Wu and H. Aghajan, Model-based Human Posture Estimation for Gesture Analysis in an Opportunistic Fusion Smart Camera Network, Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, **(2007)** September 5-7; London, England, pp. 453 – 458.

[6] Microsoft, Kinect for Windows Architecture, [Online] http://msdn.microsoft.com/en-us/library/jj131023.aspx, [Sep. 18, 2014].

[7] M. Ruta, F. Scioscia, M. Di Summa, S. Ieva, E. Di Sciascio and M. Sacco, Semantic matchmaking for Kinect-based posture and gesture recognition, Proceedings of the IEEE International Conference on Semantic Computing, **(2014)** August 1-3; Hue Royal City, Vietnam, pp. 309 - 312.

[8] Y. Liu, Z. Zhang, A. Li and M. Wang, View Independent Human Posture Identification using Kinect, The 5th International Conference on Biomedical Engineering and Informatics, **(2012)** October 16-18; Chongqing, China, pp. 1590-1593.

[9] H. P. H. Shum, E. S. L. Ho, Y. Jiang and S. Takagi, Real-Time Posture Reconstruction for Microsoft Kinect, IEEE Transactions on Cybernetics, 43(5), 1357 – 1369 **(2013)**.

[10] K. K. Htike and O. O. Khalifa, Comparison of supervised and unsupervised learning classifiers for human posture recognition, Proceedings of the International Conference on Computer and Communication Engineering, **(2010)** May 1-12; Kuala Lumpur, Malaysia, pp.1-6.

[11] X. Dai, M. Wu, B. Davidson, M. Mahoor and Z. Jun, Image-Based Fall Detection with Human Posture SequenceModeling, Proceedings of the IEEE International Conference on Healthcare Informatics, **(2013)** September 9-11; Philadelphia, Pennsylvania, USA, pp. 376 - 381.

[12] T. Visutarrom, P. Mongkolname, and J. H. Chan, Postural Classification using Kinect, Proceedings of the 18th International Computer Science and Engineering Conference, **(2014)** Jul. 30 - Aug. 1; Khon Kaen, Thailand, pp. 346 - 351.

[13] T. Visutarrom, P. Mongkolnam, and J. H. Chan, , Multiple-stage classification of human poses while watching television, Proceedings of the International Symposium on Computational and Business Intelligence, **(2014)** December 7-8; New Dehli, India, pp. 10-16.

[14] The University of Waikato, Weka 3-Data Mining with Open Source Machine Learning Software in Java, [Online] http://www.cs.waikato.ac.nz/ml/weka/, [May 1, 2014].

[15] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning Data Mining, Inference, and Prediction, 2nd Edition, Springer Series in Statistics, Springer **(2008)**, pp. 119-603.

[16] J. Han and M. Kamber, "Data mining concepts and techniques," 2nd Edition, Morgan Kaufmann Publishers, **(2006)**, pp. 285-378.

[17] M. A. Hall, Correlation-based Feature Subset Selection for Machine Learning, Ph.D. thesis at The University of Waikato, Hamilton, New Zealand, April **(1999)**.

[18] D. Meyer, E. Dimitriadou, K. Hornik, A. Weingessel, F. Leisch, C. Chang, and C. Lin, Misc Functions of the Department of Statistics (e1071), TU Wien, R package, **(2015)**.

[19] D. Liaw, M. Wiener, Breiman and Cutler's random forests for classification and regression, R package, **(2015)**.

[20] L. Breiman, Random Forest, Machine Learning Journal, 45, 5-32 **(2001)**.

[21] L. Breiman, L., J. H. Friedman, R. A. Olshen, C. J. Stone, Classification and regression trees, Wadsworth, Belmont, California **(1984)**.

[22] W. Engchuan, K. Dhindsa, A. C. Lionel, S. W. Scherer, J. H. Chan, J. H. and D. Merico, Performance of case-control rare copy number variation annotation in classification of autism, BMC Medical Genomics, 8(Suppl 1), S7 **(2015)**.

[23] P. O. Gislason, J. A. Benediktsson, J. R. and Sveinsson, Random Forests for land cover classification, Pattern Recognition Letters, 27(4), 294-300 **(2006)**.

[24] S. Fong, J. Liang, R. Wong, M. Ghanavati, A novel feature selection by clustering coefficients of variations, Proceedings of 2014 Ninth International Conference on Digital Information Management, **(2014)** Sept. 29 - Oct. 1; Phitsanulok, Thailand, pp.205-213.

[25] J. Lu, J. Zhou, H. Ruan, and G. Luo, Establishing a University Library-Based Health Information Literacy Service Model in the Age of Big Data, J. Med. Imaging Health Inf. 6, 260-263 **(2016)**.

TABLE I. NUMBER OF INSTANCES IN EACH POSTURE.

Posture	Number of Instances	
	Train	Test
leandown	13500	13500
sitLb	13500	13500
sitLbBHP	13500	13500
sitLbLHP	13500	13500
sitLbRHP	13501	13500
sitLf	13500	13500
sitLfBHP	13500	13500
sitLfLHP	13500	13500
sitLfRHP	13500	13500
sitS	13500	13500
sitSBHP	13500	13500
sitSLHP	13500	13500
sitSRHP	13499	13500
sitfloorLB	13500	13500
sitfloorLBBHP	13500	13500
sitfloorLBLHP	13500	13500
sitfloorLBRHP	13500	13500
sitfloorLf	13500	13500
sitfloorLfBHP	13500	13500
sitfloorLfLHP	13500	13500
sitfloorLfRHP	13500	13500
sitfloorS	13500	13500
sitfloorSBHP	13500	13500
sitfloorSLHP	13500	13500
sitfloorSRHP	13500	13500
stand	13500	13500
standBHP	13500	13500
standLHP	13500	13500
standRHP	13500	13500
standleanforward	13500	13500

*(Lb/Lf: lean back/forward; RHP/LHP/BHP: right/left/both hand up; S: straight)

TABLE II. CLASSIFICATION RESULT (ACCURACY) USING EACH POSITION AS TRAINING SET

Train (Position)	Test (Position) 1	2	3	4	5
1	**99.99**	53.01	47.68	29.48	36.65
2	46.83	**99.99**	72.08	45.64	39.93
3	41.31	62.7	**96.66**	51.73	48.13
4	38.27	52.55	60.52	**99.85**	49.16
5	30.47	28.85	43.04	39.73	**100**

*The highest accuracy of each training set is highlighted in bold.

TABLE III. SELECTED FEATURES BY GA ALGORITHM AND CFSATTRIBUTEEVAL USING WEKA

Found in		Feature ID	Feature Name
9	Folds	2	head_yin3
9	Folds	5	shoulder_center_yin3
9	Folds	8	shoulder_left_yin3
8	Folds	11	shoulder_right_yin3
9	Folds	14	elbow_left_yin3
9	Folds	17	elbow_right_yin3
10	Folds	20	wrist_left_yin3
10	Folds	23	wrist_right_yin3
10	Folds	26	hand_left_yin3
10	Folds	29	hand_right_yin3
8	Folds	32	spine_yin3

TABLE IV. STATISTICS OF SELECTED FEATURES VALUES ON Y-AXIS

head_yin3	0.219937323	0.373679657	1.699027938
shoulder_center_yin3	0.073074038	0.364593213	4.989367262
shoulder_left_yin3	-0.025636896	0.375148953	14.63316596
shoulder_right_yin3	-0.015356577	0.347347602	22.61881722
elbow_left_yin3	-0.033477611	0.37032402	11.06184138
elbow_right_yin3	-0.118728611	0.323232847	2.722451184
wrist_left_yin3	-0.007566191	0.428861032	56.68123465
wrist_right_yin3	-0.096745039	0.384861797	3.978103703
hand_left_yin3	0.009157107	0.458254788	50.04361802
hand_right_yin3	-0.086529187	0.416564946	4.814155319
spine_yin3	-0.176231644	0.33259765	1.887275416
		Average CV:	15.92082346

TABLE V. STATISTICS OF NON-SELECTED FEATURES VALUES ON X-AXIS

Non-selected features on X-axis	Mean	Standard deviation	Coefficients of variation (CV)
head_xin3	-0.100577626	0.524566428	5.21553798
shoulder_center_xin3	-0.067500352	0.494701274	7.328869497
spine_xin3	-0.078971691	0.456942986	5.786161835
hip_center_xin3	-0.085288281	0.450383247	5.280716656
knee_left_xin3	-0.313402711	0.483115009	1.541515095
knee_right_xin3	-0.256697441	0.489513615	1.90696726
ankle_left_xin3	-0.356972506	0.497918443	1.394836953
ankle_right_xin3	-0.269206361	0.520405137	1.933108618
foot_left_xin3	-0.378240594	0.488667355	1.291948465
foot_right_xin3	-0.30761767	0.512305462	1.665396734
dis	1.499986522	0.408250811	0.272169653
		Average CV:	3.056111704

TABLE VI. STATISTICS OF NON-SELECTED FEATURES VALUES ON Z-AXIS

Non-selected features on Z-axis	Mean	Standard deviation	Coefficients of variation (CV)
head_zin3	2.451523062	0.518857192	0.211646874
shoulder_center_zin3	2.472070552	0.52141789	0.210923547
spine_zin3	2.506281178	0.527435813	0.210445587
hip_center_zin3	2.507005727	0.532409058	0.212368505
knee_left_zin3	2.667176386	0.547363428	0.205222058
knee_right_zin3	2.461543348	0.543557814	0.220819924
ankle_left_zin3	2.658481117	0.546819596	0.205688727
ankle_right_zin3	2.526529533	0.559229678	0.22134302
foot_left_zin3	2.634289906	0.547242364	0.207738094
foot_right_zin3	2.514885382	0.564266992	0.224370858
dis	1.499986522	0.408250811	0.272169653
		Average CV:	0.218430622

TABLE VII. CLASSIFICATION RESULT (ACCURACY %) USING ALL POSITION COMBINED AS TRAINING SET

	Accuracy			
Classifier	**RF**	**SVM**	**NN**	**NB**
Train all	99.29	98.65	92.43	37.82
Train (60 points)	99.22	98.94	94.18	37.98
Train with GA	99.2	81.34	75.54	35.25

TABLE VIII. CLASSIFICATION RESULT (RECALL) USING ALL POSITION COMBINDED AS TRAINING SET SEPARATED BY EACH POSTURE

Classifier	RF	SVM	NN	NB
leandown	1.00	1.00	0.89	0.49
sitLb	1.00	1.00	0.98	0.20
sitLbBHP	1.00	1.00	0.92	0.28
sitLbLHP	1.00	1.00	0.97	0.32
sitLbRHP	1.00	1.00	0.96	0.27
sitLf	1.00	1.00	0.95	0.27
sitLfBHP	1.00	0.98	0.92	0.38
sitLfLHP	1.00	0.99	0.96	0.35
sitLfRHP	1.00	0.98	0.96	0.36
sitS	1.00	1.00	0.97	0.31
sitSBHP	1.00	1.00	0.95	0.36
sitSLHP	1.00	1.00	0.97	0.42
sitSRHP	1.00	1.00	0.98	0.36
sitfloorLB	0.95	0.85	0.81	0.30
sitfloorLBBHP	0.95	0.98	0.84	0.32
sitfloorLBLHP	0.95	1.00	0.82	0.41
sitfloorLBRHP	0.94	0.98	0.88	0.27
sitfloorLf	1.00	1.00	0.90	0.34
sitfloorLfBHP	1.00	1.00	0.92	0.50
sitfloorLfLHP	1.00	0.98	0.95	0.43
sitfloorLfRHP	1.00	1.00	0.93	0.47
sitfloorS	1.00	0.99	0.94	0.25
sitfloorSBHP	1.00	1.00	0.87	0.29
sitfloorSLHP	1.00	0.98	0.87	0.34
sitfloorSRHP	1.00	0.99	0.94	0.31
stand	1.00	1.00	0.99	0.62
standBHP	1.00	0.97	0.85	0.48
standLHP	1.00	0.98	0.95	0.50
standRHP	1.00	0.99	0.95	0.48
standleanforward	1.00	1.00	0.97	0.20

*(Lb/Lf: lean back/forward; RHP/LHP/BHP: right/left/both hand up; S: straight)

TABLE IX. CLASSIFICATION RESULT (ACCURACY %) USING ALL POSITIONS COMBINDED AS TRAINING SET

	RF		SVM		NN		NB	
	Acc	Time	Acc	Time	Acc	Time	Acc	Time
Train all	99.29	0:45:00	98.65	2:22:00	92.43	1:03:00	37.82	0:00:06
Train (60 points)	99.22	0:43:00	98.94	2:30:00	94.18	1:03:00	37.98	0:00:06
Train with GA	99.2	0:11:00	81.34	2:00:00	75.54	0:33:00	35.25	0:00:02