Gait Analysis for Human Authentication using Depth Sensor

[1]Arya Ramesh, [2]T. Karthi , [3.] C. Thamilarasi
Applied Electronics SVHEC, Erode, India

Abstract—The prominence of system for automatic person identification has rise increasingly during the past years. Bio- metric features from an information security perspective are a process by which the identity of a person can be confirmed. Use of biometric features for authentication is one of the three most widely used approaches. Gait analysis is an important biometric technique for recognizing humans. Unlike other biometric fea- tures, human gait can be captured at a distance which makes it a self-effacing method for recognition. This paper focus on an authentication security system which is more oriented towards behavioural characteristics like gait analysis, rather just physical characteristics of humans. This blend of both characteristics helps evolving more powerful authentication system as compared to existing biometric security systems. An unconstrained gait recognition algorithm is proposed which uses 3D skeleton infor- mation and angles between skeleton joint points. Kinect sensor captures depth image is used to generate 3D skeleton structure. The temporal tracking of skeleton joints angles formed due body motion kinetics is used for gait analysis. The angles from the 2D representation of 3D data are computed by calculating the vectors between the three joints pairs using atan2 (the arc tangent function with two arguments).These skeleton joints angles forms the trajectories and hence gait model. The gait is acknowledged by calculating the minimum difference measure between the gait models of the training data and the testing data.

Index Terms—Gait analysis, Microsoft Kinect, Skeletal structure, Physical Authentication, Behavioral Authentication, Depth sensor, Human Skeletal Joints Model.

I. Introduction

In recent years, identity authentication has become increasingly important. People are required to verified or identified as a legitimate claimed individual to be ready to access secured assets or information. Biometrics allows an individual to be identified and authenticated through support of a group of recognizable and verifiable data, which is unique and specific to them. It also can be defined because the measure of physical body. Biometric identification is the process of comparing data for the person's characteristics thereto person's biometric ”template” to work out resemblance. The reference model is first stored during a database or a secure portable element sort of a open-end credit. The info stored is then compared to the person's biometric data to be authenticated. Here it's the person's identity which is being verified against his/her stored traits. Biometric authentication consists of determining the identity of an individual. The objective is to represent item(s) of biometric data from this person. It are often a photograph of their face, a record of their voice, or a picture of their fingerprint. This data is then compared to the biometric data of several other persons kept during a database. There are two categories of biometric technologies: Physiological measurements and Behavioural measurements. Physiological measurements are often either morphological or biological. These mainly contains fingerprints, the form of the hand, of the finger, vein pattern, the attention (iris and retina) and therefore the shape of the face, for morphological analyses. Voice, signature dynamics, keystroke dynamics, the way objects are used, gait, gestures, etc. are behavioural measurements. Currently, biometric system generally makes use of fingerprints, voice prints, face characteristics, facial features, iris features, retina image, hand geometry, palm print, etc.

Now-a-days it is not that difficult to fool even the iris based biometrics. An article online on Forbes stated that researchers tricked Iris Scanner using forged eye images, where Galbally‘s research showed that nearly 80 percentage of a commercial scanner validated his forged iris prints as real. There are many ways which make it easy for attackers to forge fingerprints. With the help of 3D printing other body parts can be forged if used for authentication like ear lobes. Fake contact lens can also be used to bypass the iris scan. The face can also be masked easily as it is not limited to what is seen in movies or some vfx effects but can be created in real. Hence it is seen that these systems are mostly based on one or more physical body parts which is the limitation of such existing systems. But what if the whole body anatomy can be taken as a security feature along with some behaviour features unique to every individual at some level like for example gait- the way a person walks. Such combination of physical as well as behavioural characteristics of humans will result in a very powerful authentication system which is very difficult to bypass or fool. These systems trace high level of uncertainties among humans. When such system is used with other security features (on which existing security systems work upon) then that will be a security add-on, resulting in the most promising biometric security system.

This research specially focuses on potential application of Microsoft Kinect- a depth sensor for human tracking and identification capabilities. Recent researches show that it is not difficult to fool even the traditional biometric security systems. It is easy to forge fingerprints, iris etc. Through this paper it is demonstrated how Microsoft‘s Kinect state-of-art ability to track human skeleton and its dimensions can provide a better security and enhanced authentication, with low false acceptance or rejection rate. It‘s a two level security system. Level 1 authentication is based on skeletal body features of any individual while Level 2 enhances this system by combining behavioural characteristics such as gait. Gait analysis and re-cognition is based on angles formed in between various three joints pairs computed from vector of lines, angle formed between those vectors with respect to their directions that represents body kinetics during motion. It provides better security and is autonomous in nature as well as inexpensive compared to existing security solutions. It turns out to be an effective intrusion detection or authentication system which could potentially be commercialized.

II. RELATED WORKS

In recent years new classification techniques have evolved for gait based systems which can be analysed using WEKA analyser which is type of analyser used in machine learning. Johannes Preis, Moritz Kessel, Martin Werner, Claudia Linnhoff-Popien have presented a paper 'Gait recognition with Kinect' describing about gait analysis using Microsoft Kinect. The system use skeletal data for recognition and did not use the depth- and color-images directly. The proposed system consist of three components: the first component record the skeleton information and is fed into second component for feature extraction. The third component is machine learning framework WEKA to identify a person on the basis of previously recorded training data. Classification based on step length and speed still yielded 55.2 percentage success rate using either Naive Bayes or the decision tree. They have evaluated a number of body features together with step length and speed, their relevance for person identification, and present the results of an empirical evaluation of the system, where the system was able to accomplish a success rate of more than 90 percentage with nine test persons [1]. Machado molina M, Ingrid Bonnunger, Malay Kishore Dutta, Tobias Kutzner, Carlos M use another framework method in WEKA to improve the information extraction. The system use OpenNi framework for feature extraction and for classification, uses different techniques of machine learning of the WEKA software to find the best method for verifying people as authorized or not in safety areas. WEKA analysis is done in two steps: first is to identification in a closed set and identification in an open set [2].

Generally the techniques used for gait recognition can be broadly divided into two main classes: model free and model based. In model free technique the gait recognition algorithms usually depend on silhouette of the human body and using that it generates normalized grayscale images called gait energy images (GEIs). Typically, for a person identification principal component analysis is performed on these GEIs, which is followed by image classification [3]. Mori et al. [4] in their research used silhouettes in order to extract gait information for recognition. The model free methods work well when used for classification but such methods implicitly assume that data is noise free. When these methods are applied to real problems, noise invariably reduces the classification accuracy.

Model based methods are not only suitable for extracting features for signal processing techniques but they also exploit the oscillating movement of the lower body during a walk. However model based methods are restricted to gaits where a person walks along a linear path parallel to the image plane as other paths would provide variant data to the classifier. The method proposed here is also a model based technique. Depth information is an important cue when human recognizes objects because the objects may not have consistent colour and texture but must occupy an integrated region in space. Range images have several advantages over 2D intensity images: range images are robust to the change in colour and illumination. Also, range images are simple representations of 3D information [5]. Kinect is a newly evolved technology. The human skeleton is approximated by Shotton et al. [6] algorithm that works in the Kinect background. There is no restriction on a person‘s walk that it has to be relative to the sensor viewpoint. 3-Dimentional data is being used to extract gait data. Lu and Chia in their paper had discussed about human detection had used features based on gradients, such as histograms of oriented gradients (HOG), or extract interest points in the image, such as scale-invariant feature transform (SIFT), etc. In their paper, they had also presented a novel human detection method using depth information from the Kinect. They had proposed a model based approach, which detects humans using a 2-D head contour model and a 3- D head surface model. A segmentation scheme to segment the human from his/her surroundings and extract the whole contours of the figure based on detection points. They had also explored the tracking algorithm based on the detection result. The methods were tested on a database taken by the Kinect in their lab and presented superior results [7]. Ball et al. [8] had directly used the skeleton point trajectories from the three dimensional skeleton data. Although the method proposed was based on unsupervised learning. 71 walk half- cycles were captured from 4 people but there the walk cycles were manually extracted out by segmentation. The lower body portion was only used for clustering with the help of K- means. The overall accuracy of the research was approximately 43.5percentage.

R. Venkatesh Babu and Naresh Kumar M S had proposed an efficient gait recognition approach based on extraction of skeleton joints trajectories obtained from a Kinect sensor camera. Kinect as a depth sensor provides a new dimension of features for gait analysis. It provides a more robust 3- D skeleton information since the depth information provides an accurate silhouette sequence. The recognition rate for this proposed approach was above 90 percentage. The robustness of the proposed approach is quantified for fixed and moving camera scenarios [9]. Another research by Tata Consultancy Services Innovation Labs team in which they had presented few static and hybrid model based features for gait recognition using Microsoft Kinect. In addition to gait cycle, the new gait features that were taken into account are the areas formed between upper and lower body part. That involves a few static and hybrid models for gait recognition. The upper body area showed the static features while the area formed by the distances between the body centroid and the centroids derived from the upper and lower limbs joints formed the dynamic features. An Artificial Neural Network based system by N.

R. Pal and K. K. Chintalapudi, was used to perform features selection and classification. All this was done in batch mode. Results indicate that the static features get higher importance compared to the dynamic features. Moreover, the hybrid features proposed in this paper are having a great importance compared to dynamic feature sets as proposed in the ANN based earlier feature selection processes. [8]Another related work where gait analysis of complete body was done with the help of Kinect. This analysis can be used in health monitoring, diagnosis and rehabilitation. The approach was non-intrusive and economical where comprehensive gait information was collected using all body parts such as arm kinematics, stride information, as well as other parameters to improve the stride intervals measurements. Also no expensive markers or battery based devices were used except inexpensive Kinect sensor, which have reinstalled base to capture human skeletal data and generate virtual skeleton as an input to supervised learning model. [10]

Most of these researches contrast on the limitations due to noise in extracted dynamic features and normalization of the skeletal forms otherwise the results could be much better and more robust systems could be formed.

III. BUILDING THE DATABASE

For this research 30 subjects were selected to create a training database such that the heights of selected subjects range from 5‘ to 6‘ feet (Table 1). The Kinect sensor records the data approximately at 30 frames per second. Each of these thirty subjects walked ten times, resulting into 300 total walks observed using Kinect and stored in the database. For the first 8 walks the subjects were asked to move in a straight line towards the Kinect sensor. For the remaining two walks the subjects were asked to shift the movement by some degree to left and right from the straight path.

In each walk an average of 150 frames were recorded and any frame in which any of the 21 joints(out of 25 joints not considering hand tip joints and thumb joints) was missing was ignored. In a skeleton frame sixteen angles are formed from various joints, which are computed from the vectors.

IV. PROPOSED METHOD

This system is a human authentication system that performs skeletal measurements and gait analysis to automatically distinguish between the owner (known persons) and the intruder (unknown person). There are two levels of security in this system described below:

TABLE I: VARIABLE HEIGHT TEST SUBJECTS ENTERED INTO THE KNOWN PERSON DATABASE.

Number of Persons	Heights
2	5’ or below
13	5’1” to 5’6”
12	5’7” to 6’
3	6’ or above

A. Level 1-Physical Authentication Phase

In Level 1 the Euclidean distance between the pairs of selected joints (skeletal joints pairs that are mentioned below) are computed. Now for the effective authentication of an unknown person, the sum of the differences between those Euclidean distances of the unknown person to that of a known person‘s skeletal measurements stored in the database must be within a pre-specified threshold (i.e. 14 cm).The joints pairs that are considered by the system for physical static features are as follows: - Left and right fore-arm (distance between wrist to elbow joint) - Left and right upper-arm (distance between elbow to shoulder joint) - Left and right thigh (distance between knee to hip joint) - Left and right shin (distance between ankle to knee joint) - Shoulder width (distance between left shoulder to right shoulder joint) If the sum of these nine measurements does not match with any of the known persons‘ skeleton dimensions in the database then authentication fails. Kinect keeps a track of all the necessary joints in form of tracked, inferred, or not- tracked state. Joints with the tracked state are only considered. On average the length of the above mentioned nine body segments (nine measurements) for any known candidate (test data) is about m while 0.14 m is the threshold used by the system. The difference between two persons must be less than 0.14 m out of the total length of nine body segments before the system can match an unknown person skeleton among the known persons’ skeletons.

$$(14/348) * 100 = 4 \quad (1)$$

Therefore, 4 percentage (approx.) tolerance is allowed. The measurements as drafted by Kinect vary from one frame to other frame but if any of these frames lies within 4 percentage of difference then authentication at initial level i.e. level-1 is successful. With the help of Microsoft Kinect the length of the segments between two joints can be calculated. Segments between joints likes head and shoulder-centre, right-hand and right-wrist etc. which are in red colour as shown in figure 2 have been skipped as their impact is low in finding the differences between two individuals. While the 9 segments of a skeleton which are shown (Figure 2) in blue color are only considered. Those lengths are computed using Euclidean Distance = d' between any two joints using following formula:

EUCLIDEAN DISTANCE

$$(x2 - x1)^2 + (y2 - y1)^2 + (z2 - z1)^2 \quad (2)$$

delta: total difference epsilon: maximum delta to be allowed
so, if delta is less than epsilon then authentication is successful

Where

$$delta = \sum (d1_i + d2_i) \quad (3)$$

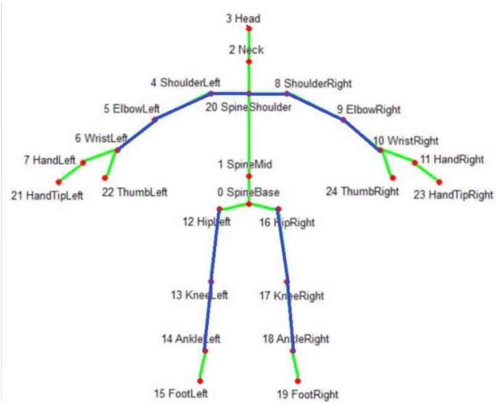

Fig. 1. Skeletal structure with various joints and segments

- i vary from 0 to 9
- d1 is the distance between 2 joints of the detected person
- d2 is the distance between 2 joints of the known person

It is an area of future improvement if the variations between two frames can be further smooth-en. The Kinect functions at approx. 30 frames per seconds. The lookup for each image is done against the stored biometric database. It takes 0.4 milliseconds of time for the authentication algorithm to execute and perform a lookup. Thus level-1 matches with the speed requirements and make this algorithm more efficient.

B. Level 2 - Behavioural Authentication phase

After the current user's successful authentication at Level 1, the user enters the gait analysis stage. Level 2 consists of two phase-recording phase and comparison phase. The approach used for recording gait is to store the unknown user's respective walk frames for deriving the test walk data and for comparison-body kinetics during human motion represented using vectors.

- Recording Phase

In recording phase the Kinect tracks a person's skeleton, it stores the 16 body angles formed at various instance of time. Each frame is stored in form of these 16 feature vectors into a txt file. This takes place for the entire duration of walk. After gait vectors data is recorded which forms the test data, this data is input into an array. Complete details of a frame are stored at a particular location in an array. But question arises how the angles between the following bodies segments could be found in real time.The figure 2 shows the joints identified by Microsoft Kinect.

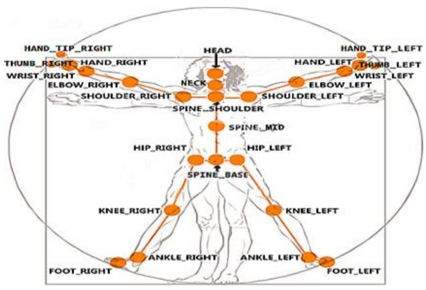

Fig. 2. Joint types

The following 16 joint angles has been considered for gait analysis

1) Right Wrist - Right Elbow- Right Shoulder
2) Left Wrist - Left Elbow -Left Shoulder
3) Right Elbow- Right Shoulder-Shoulder Center
4) Left Elbow - Left Shoulder- Shoulder Center
5) Right Shoulder-Shoulder Center-Spine mid
6) Left Shoulder-Shoulder Center-Spine mid
7) Right Wrist-Right Shoulder- Spine mid
8) Left Wrist-Left Shoulder- Spine mid
9) Right Hand – Right Wrist- Right Elbow
10) Left Hand- Left Wrist-Left Elbow
11) Right Ankle – Right Knee - Spine Base
12) Left Ankle – Left Knee – Spine Base
13) Right Knee- Right Hip- Spine Base
14) Left Knee- Left Hip-Spine Base
15) Right Hip-Spine Base-Spine Mid
16) Left Hip –Spine Base- Spine Mid

From these 16 joints pairs, vectors can be calculated by joining these joints pairs and thus respective angles between those pair of vectors can be calculated w.r.t their directions. The method used to calculate an angle between any 3 joints pair is explained here under.

After calculating the angles of each frame, they are saved into a .txt file which can later be import into an array so that they can be compared with test walk frames angles.

- Comparison Phase

In comparison phase all the frames angles of the test walk and the training walks are retrieved and differences are found between the test walk frames angles and each of the training data sets- 300 walks (30 users with 10 walks each). Since every walk has approximately 150 frames. So there are around 150 person if existed in the database on basis of gait recognition. To improve algorithm results against mismatch a subject with highest number of matched walks among the 30 subjects with the test walk is taken as a successful match to the current user's test walk.

```
Vector3D a1 = new Vector3D(Joint0.Position.X, Joint0.Position.Y, Joint0.Position.Z);
Vector3D a2 = new Vector3D(Joint1.Position.X, Joint1.Position.Y, Joint1.Position.Z);
Vector3D a3 = new Vector3D(Joint2.Position.X, Joint2.Position.Y, Joint2.Position.Z);

//Calculate vectors joining the two joints
Vector3D b1 = a1 - a2;//Joint0toJoint1
Vector3D b2 = a3 - a2;//Joint1toJoint2

//Normalise the above vectors to unit vector
b1.Normalize();
b2.Normalize();

//Find a dot-product of these two vectors
dotProduct = Vector3D.DotProduct(b1, b2);

//Find the cross product of these two vectors
Vector3D crossProduct = new Vector3D();
crossProduct = Vector3D.CrossProduct(b1, b2);
double crossProdLength = crossProduct.Length;

//cal angle formed using atan2 in radians
double angleFormed = Math.Atan2(crossProdLength, dotProduct);

//calc angle formed in degree
double angleInDegree = angleFormed * (180 / Math.PI);
double rounddAngle = Math.Round(angleInDegree, 2);

return rounddAngle;
```

TEST RESULT

Testing plays an important role to check whether the system performance is satisfactory or not, in order to meet any research goal. Various tests to find the efficiency and robustness of this system are performed.

A.Testing Level 1

TABLE II: RESULT MATRIX TO COMPUTE FALSE ACCEPTANCE AND FALSE REJECTION RATE AT LEVEL1

	Sub1	Sub2	Sub3	Sub4	Sub5
Sub1	8	2			
Sub2		9			
Sub3	1		9		
Sub4			1	9	
Sub5			1		8

16 vectors pertaining to each walk specific to each user. The comparison is made in such a way that the 16 angles in a test walk frame are compared with the 16 angles of every frame of a walk of a specific person in training data i.e. a frame of test walk of an unknown person is matched with every frame of a train walk of a specific user. During comparison of two frames, small deviation among angles occurs but is negligible. For this reason a tolerance of -/+ 6.5 degree is allowed for each angles difference computed. If all the angles in a frame lie within this tolerance range then that frame is fully matched. The comparison continues with every frame of test walk which is matched against the 10 walks frames of every user in training database.

The numbers of test walk frames satisfying above constraint i.e. matched successfully are counted. If at least 90 percentage of the test walk frames match with a training walk frames of a particular subject from the training database then the unknown user is successfully authenticated against the matched subject in the training database. In order to improve the performance and reduce time complexity if after the 15 percentage com- pletion of the comparisons between frames of a train walk and the test walk, if even a single frame is of the test walk is not matched with a train walk then that training data walk is skipped as it can‘t guarantee a 90 percentage or above match. So here not only the user is classified based on nearest neighbour but also on a threshold, that at least 90 percentage of test walk frames should match with the training database walks frames of a particular subject, for successful authentication. Initially the skeletal details of 30 people were added in the database of known persons. The system is able to detect and match most of the times an unknown person to the known,

Table 2 represent the test result of level 1

$$43/50 * 100 = 86 \tag{4}$$

From the above data success rate(efficiency) of the system at Level1 only in restricted domain* is 86 percentage.

* here restricted domain means that the process is done for 5 subjects in a lab at SVHE College where each subject was asked to enter into a classroom ten times for authentication process. The observations were made in indoor ambience lightening conditions.

A. Testing Level 1 and Level 2

TABLE III: RESULT MATRIX TO COMPUTE SUCCESSFUL AUTHENTICATION RATE

	Sub1	Sub2	Sub3	Sub4	Sub5	Sub6	Sub7	Sub8	Sub9	Sub10	SubX
Sub1	5										
Sub2		4									1
Sub3			5								
Sub4				5							
Sub5					4						
Sub6						4					
Sub7							5				
Sub8								3			1
Sub9									5		
Sub10										4	

The table 3 represent the output of successful authentication. Here subject X refers to mismatch.

$$44/50 * 100 = 92 \tag{5}$$

From the above data overall success rate(efficiency) of the system is 92 percentage.

B. Testing Level 2 gait recognition algorithm alone

To test the accuracy of gait recognition algorithm of pro- posed system, one walk of a user is chosen at a time from 1 to 10 walks and matched against the remaining 9*30 walks of every user. For example suppose walk 2 is taken then it

is compared with walk 1,3,4,5,6,7,8,9,10 from user 1 to user 30
i.e. total 9*30 walks. This process repeated with each walk (1-10) of every user. Tble 4 give the result of this test.

TABLE IV: RESULT MATRIX TO COMPUTE SUCCESSFUL AUTHENTICATION RATE

Walk1	Walk2	Walk3	Walk4	Walk5	Walk6	Walk7	Walk8	Walk9	Walk10
93.33	96.66	100	100	100	93.33	93.33	96.66	96.66	96.66

VI. CONCLUSION

In this paper, a system is designed and tested to identify and authenticate the gait system of human. To extract the gait system a depth sensor Microsoft Kinect is used. Authentication is done is two phase: physical and behavioural authentication phase.The algorithm created calculate angles formed using specific joins while walking and tries to match with the data base. It is found that this approach give a success rate of 92 percentage. It won't take long to realize that, a gait recognition system can be used as a total security solution if merge with existing biometric systems that take face, voice, thumb prints as security features. The Kinect is such a device which is cheap enough to be included as standard hardware if its robustness and reliability can be proved. Security isn't a simple game. Whenever something new is developed or any advancement is made to it like in case of a new advanced biometric system, sooner ways are found to bypass or trick it. Even for this proposed system a human body clone could be generated using 3D printing technique and few gears and motorised servos, a servos controlling circuits could make the clone body move different parts is sufficient enough to bypass even the gait recognition of the system. But there is no doubt that still these human physical and behavioural characteristics based systems are better than the existing biometric systems.

REFERENCES

[1] Johannes Preis,Moritz Kessel, Martin Werner, Claudia Linnhoff-Popien (2017),'Gait Recognition with Kinect'.

[2] Machado molina M, Ingrid bönninger, Malay Kishore dutta, Tobias kutzner, Carlos M. travieso (2015), 'Gait-based recognition of humans using Kinect camera',Recent Advances in Computer Engineering, Com- munications and Information Technology

[3] R. Zhang and C. Vogler —Human gait recognition,|| in IEEE Comp. Soc. Conf. on Computer Vision, 2004 and Pattern Recognition Workshops, 2004, pages 18–18.

[4] A. Mori, Y. Makihara and Y. Yagi. —Gait recognition using period-based phase synchronization for low frame-rate videos,|| in 20th Int. Conf. on Pattern Recognition, 2010, pages 2194 – 2197

[5] BC. Vemuri, A. Mitichea and J. K. Aggarwal —Curvature- based Representation of Objects from Range Data,|| in Image and Vision Computing Vol. 4, No. 2, May 1986, pages 107-114

[6] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp,M. Finocchio, R. Moore,A. Kipman, and A. Blake. —Real-time human pose recognition in parts from single depth images|| In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 2, 2011 pages 1297–1304

[7] Lu Xia, Chia-Chih Chen and J.K. Aggarwal —Human detection using depth information by Kinect,|| in CVPRW, IEEE Computer Society Conference on, 20-25 June 2011.

[8] Adrian Ball, David Rye, Fabio Ramos and Mari Velonaki. —Unsuper- vised clustering of people from skeleton data,|| in HRI ,Proceedings of the seventh annual ACM/IEEE international conference on Human- Robot Interaction, 2012, pages 225–226.

[9] M. S. Naresh Kumar, Venkatesh Babu, —Human gait recognition using depth camera: a covariance based approach,|| in Eighth Indian Con- ference on Computer Vision, Graphics and Image Processing, Article No.20, 2012,doi:10.1145/24253- 33.2425353

[10] Aniruddha Sinha, Kingshuk Chakravarty and Brojeshwar Bhowmick —Person Identification using Skeleton Information from Kinect,|| in The Sixth International Conference on Advances in Computer-Human Interactions, February 24, 2013

[11] M. Gabel, R. Gilad-Bachrach, E. Renshaw and A. Schuster,—Full body gait analysis with Kinect,|| in Annual International Conference of the IEEE, 2012,.