Yoga Posture Recognition By Detecting Human Joint Points In Real Time Using Microsoft Kinect

Muhammad Usama Islam*, Hasan Mahmud†, Faisal Bin Ashraf‡, Iqbal Hossain§ and Md. Kamrul Hasan ¶
*Department of Computer Science and Engineering, Asian University of Bangladesh , Email:usamaislam@iut-dhaka.edu
†Department of Computer Science and Engineering, Islamic University of Technology , Email:hasan@iut-dhaka.edu
‡Department of Computer Science and Engineering, Uttara University , Email:faisalashraf@iut-dhaka.edu
§Department of Computer Science and Engineering, Islamic University of Technology , Email:iqbalhossain@iut-dhaka.edu
¶Department of Computer Science and Engineering, Islamic University of Technology , Email:hasank@iut-dhaka.edu

***Abstract*—Musculoskeletal disorder is increasing in humans due to accidents or aging which is a great concern for future world. Physical exercises can reduce this disorder. Yoga is a great medium of physical exercise. For doing yoga a trainer is important who can monitor the perfectness of different yoga poses. In this paper, we have proposed a system which can monitor human body parts movement and monitor the accuracy of different yoga poses which aids the user to practice yoga. We have used Microsoft Kinect to detect different joint points of human body in real time and from those joint points we calculate various angles to measure the accuracy of a certain yoga poses for a user. Our proposed system can successfully recognize different yoga poses in real time.**

Keywords—Human Posture Recognition, Real Time, Microsoft Kinect, Yoga, Pose detection

I. Introduction

Human by nature is vulnerable and subject to wide ranging health diseases of which musculoskeletal disorders is an important arena and needs urgent attention. Every year a wide range of people are affected from various types of musculoskeletal disorders due to accidents or aging problem. Yoga can promote positive physical change. Studies have shown that yoga is effective in managing symptoms associated with musculoskeletal disorders including osteoarthritis carpal tunnel syndrome, hyper kyphosis and low back pain. Additionally, Yoga shows significant improvements in motor skills and physiological methods that includes blood pressure, heart speed rate and human body weight have been noted. Research also suggests that cardiopulmonary benefits of yoga include improvement of cardiorespiratory wellness as well as improved forced expiratory volume and increased vital capacity. Traditionally, yoga is done in a yoga center in the presence of a yoga trainer who can guide the patient through therapeutic assistance.Our solution suggests a distant yoga training approach without the trainer where the patient can stand in front of a device and perform yoga poses correctly without the need of trainer or being present at the yoga training center. Real time human pose detection is a growing and important area which specializes on realizing and understanding the human pose from depth images. With the introduction of Microsoft Kinect it became easier to understand human pose. Till now a wide range of work is done on human pose detection , physical rehabilitation using Kinect. But, Yoga an important aspect of physical rehabilitation remained untouched. In our work we want to detect yoga poses by human from Microsoft Kinect and cross check the poses against true data . We have set a reference model of true data for certain poses with respect to which we have detected the poses.The main challenge was the detection of body co-ordinate joints.Using the co-ordinates of Kinect capture,the poses are detected. The purpose is to remotely monitor the yoga poses for patients. The system will work as a personal guidance system. The purpose of remotely monitoring yoga poses is to increase fitness with minimal monetary investment.As, at a certain time doctors or physicians can't look after all the patients,the will benefit both the patients and doctors in terms of efficiency.

II. Related Works

In the literature, there are a number of works that have been proposed for human posture recognition [1]–[7] . Among them, most of the works use color information captured from traditional cameras. For human gesture recognition the use of kinect is explored in details with various data mining classification methods such as backpropagation neural network (BPNN), support vector machine (SVM), decision tree (DT), and naive Bayes (NB) [1].Coaching of elderly personnel through kinect has been of prior importance because of the emergence of tele-medicine and distant therapeutic services.While the discussion of the potential of kinect as a health monitoring system has been discussed [2] along with its limitation around occluding parts of the body or objects that are not relevant to the scene.

SVM technique is used to classify body posture from a 3D visual hull constructed from a set of silhouette input data in [8] which returns the recognized human postures in thumbnail images form.A multiclass SVM is used to classify human posture using human skeleton, angles of six sticks in the skeleton and object motion vectors to recognize human behavior analysis in [9]. In [10] a centroid radii model as shape descriptor is used to represent human posture in each frame. In [11], a system is proposed for simultaneous people tracking and posture recognition in cluttered environments in the context of human-robot interaction. In [12], a combination of MPEG-7 contour-based shape descriptor and projection histogram was used to recognize the main postures and the view of a human based on the binary object mask obtained by the segmentation process. Great efforts have been devoted to vision-based human posture recognition [8]–[12]. However, building a robust human posture recognition system with

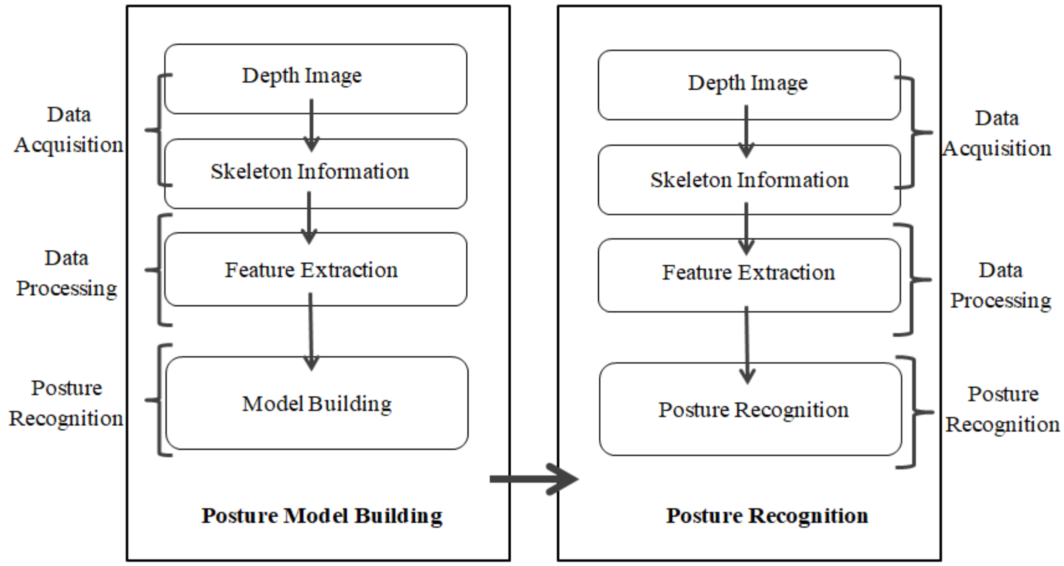


Fig. 1: Steps of the proposed system

superior performance still remains to be a challenging task. Conventional approaches in this area were generally based on regular sensor such as RGB camera, which are not only computationally costly but also easily affected by brightness variations and background clutters [3]. Recently there has been an increasing interest in using Microsoft 3D sensor Kinect for vision-based human posture recognition [9-11]. In addition to a RGB camera, Kinect contains a set of built-in IR sensor system, which can capture distance information between object and the IR sensor and offers depth image at the frame rate of 30fps with a resolution of 480*320. Moreover, the accuracy and robustness of Microsoft Kinect to detect the joint points of human body is perfect which aids to detect complex human postures [13]–[16].

III. Kinect Based Human Posture Detection

Our architecture system consists of mainly three steps - Data acquisition, Data processing and Yoga pose detection as depicted in Fig. 1. We can acquire different types of data in data acquisition step such as color, depth and skeleton information. The next step involves further processing of data for our respective goals. In this paper, we focus to use the skeleton information for yoga pose recognition.

A. Data Acquisition

For capturing video data, Kinect is well equipped with its in built infrared laser projector that is combined with a monochrome CMOS sensor. This device also has a multi-array microphone and a RGB camera that facilitates to capture image. Apart from ability of Kinect to facilitate color image and depth image, it also has a skeletal tracking tool that can recognize 20 joints of a human body as shown in Fig. 2.

Each joint point saves three information- index value, co-ordinates along x, y and z direction and status of joint. The index value is unique for each joints. The co-ordinates provides the value in x,y and z direction of the joint as shown in Fig.3.

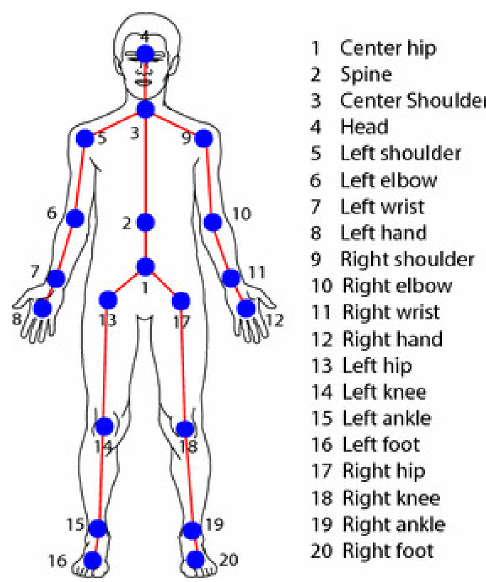


Fig. 2: Human Skeleton Information from Kinect

The status information provides whether the joint is tracked or not.

B. Data Processing

From the tracking of joint points of human skeleton by Kinect we can extract the coordinates of the joint points. As yoga positions are predefined and we intend to recognize specific three poses we can leave the joint points marked as 8,11,2,16,20 for further processing. For the pose that we are trying to detect, these joint points will not have any beneficial value for precise detection. For example, the joint point value of Left wrist and right wrist are enough for detection which is why the joint point value of left hand and right hand are unimportant and were not taken into consideration.The same

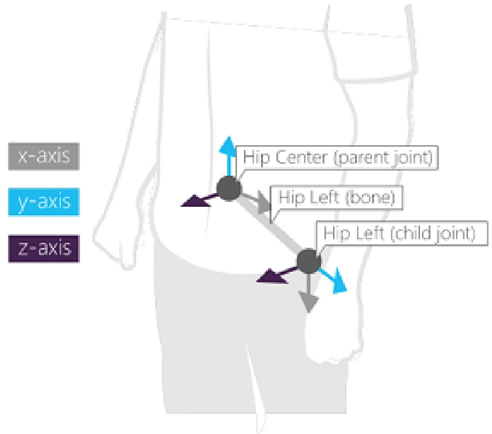


Fig. 3: Three coordinates of joint point in Kinect

instance can be given for Spine, left foot and right foot joint point. In this paper, we will find out the joint angles of the selected points for acquiring accuracy. In this context, we will use 15 points for further processing which are shown as marked in Fig. 4.The points that are used for further processing are marked blue within a black circle and the points that are not used are marked in blue circles.

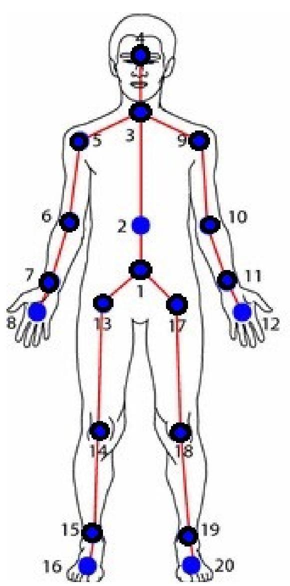


Fig. 4: Selected joint points for yoga pose detection

C. Yoga Posture Recognition

For yoga posture recognition we propose to use the following method where we will use the selected 10 joint points for calculating. All the x, y and z coordinates of the joint points determines the structure of every single yoga position.
We have built the reference structure for every yoga position by gathering information of the joint points from human posture of 5 gymnastics under supervision of certified physicist. Then we have manipulated the data of these 5 models and came up with a standard reference model by taking the average angular distances of the joint points which makes our reference frame perfect to compare with candidate frames.
To calculate the angular distance between three points in 3d axis we have followed the concept of vectors. For example, in Fig 4, three points 3 (Shoulder Center), 5 (Shoulder Left) and 6 (Elbow Left) will make an angle C. For calculating the value of angle 356 we will create two vectors $\overrightarrow{53}$ and $\overrightarrow{56}$. Let angle 356 is denoted as angle Θ and $\overrightarrow{53}$ is denoted as $\overrightarrow{XY}$ and $\overrightarrow{56}$ is denoted as $\overrightarrow{XZ}$ and then the angle between these two vectors will be found from the cosine formula for finding angle between vectors.The formula for finding the cosine angle for this instance is,

$$\Theta = cos^{-1}\frac{(\overrightarrow{XY} * \overrightarrow{XZ})}{(|\overrightarrow{XY}| * |\overrightarrow{XZ}|)} \tag{1}$$

Applying this formula, we have calculated all the angles from angle A to angle J. Let the test angle be T, the reference angle be denoted as R and the total number of angles are denoted as TA then the average deviation of a pose δ can be calculated as follows,

$$\delta = \frac{\sum(T - R)}{TA} \tag{2}$$

The recognition is said to be unsuccessful, if the average deviation or specific angle deviates abruptly from the usual.

IV. Experimentation

A. Reference Model Building

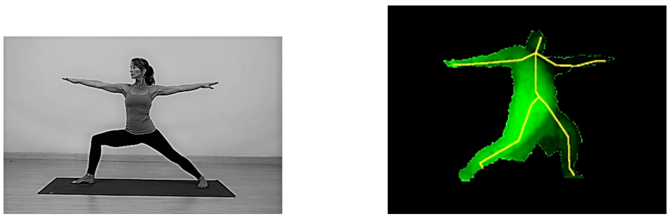
We have built a system in UNITY 3D using Microsoft Kinect which can record the object and will calculate the coordinates all these 15 necessary joint points and from these coordinates the angle between different body parts which are enough to distinguish any yoga posture from others.
The Kinect is set on 110cm measure height. The angle of Kinect for 4 postures is 0 degree. The testing room uses neon light (but actually does not effect on the result of recording skeleton data because Kinect uses IR camera which works very well without light). The interface is shown in Fig. 5, Fig. 6 and Fig. 7. We captured three postures of 5 people(3 men and 2 women) who are categorized by thin, normal,overweight and obsessed for Goddess Squat, Warrior and Reverse Warriors pose. For each person, we record 3 times, each time the duration is about 5 to 8 seconds. Subject is asked to do the posture with the following scenarios and the necessary angles between different body parts are calculated which generate our reference models.

1) Pose 1: Subject mimics sitting by bending its knees so that the arms and elbows become parallel to the head and each legs stays in parallel position as shown in Fig. 5(a) and we get the reference angle values for calculating the pose which is obtained from the Kinect 5(b).

2) Pose 2: In this special position, the subject spreads its left leg as much as possible while the right leg is vertical to the ground and spreads two hand so that they fall in a line as shown in Fig. 6(a) and we get the reference angle values for calculating the pose which is obtained from the Kinect 6(b).

(a) The pose (b) Captured in Kinect

Fig. 5: Goddess Squat



(a) The pose (b) Captured in Kinect

Fig. 6: Reverse Warrior

3) Pose 3: In this pose, Subject spreads it right leg as much as possible while the left leg is vertical to the ground and keeps the right hand on right knee and the left hand is placed parallel to the head and vertical to the ground as shown in fig 7(a) and we get the reference angle values for calculating the pose which is obtained from the Kinect 7(b).

(a) The pose (b) Captured in Kinect

Fig. 7: Warrior II

V. Result and Discussion

For calculating the results from the experiments we have selected the major angles which differs the poses and calculated every angle for every experiment. As we have selected 15 important joint points, we have considered all the 10 angles between these joint points which are shown in Fig. 8.

We experimented each pose on 5 different person and recorded all the angles shown in Fig. 8. We calculated the deviation of the angles from the reference model which tells us how much close to the pose the person got. We have considered up to 2.5 degree of deviation in any angle. If a person misses an angle by deviating greater than 2.5 degree we have considered the pose as not detected. Otherwise we have considered the pose as detected as achieving a pose with 100% accuracy is

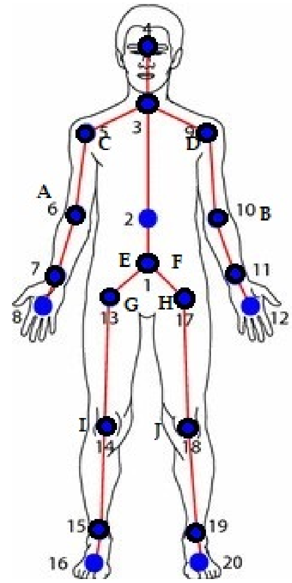


Fig. 8: Selected angles for calculating pose accuracy

quite difficult and rare. A little deviation is always accepted by the physicists.

In Table I, we can see the deviation of each person in every angle from the reference model for Posel-1. If we further explore, we can see that the pose of Person 2 was not detected as he deviated more than 2.5 degree at angle A and B. Though his average deviation is quite low, he failed to achieve a part of the pose which is why the status of the pose was not detected. Other persons could achieve the pose with a good accuracy and little average deviation of the angles.

From Table II, we get the deviation of angles for every person posing the Warrrior II yoga pose. Person 2 fails to achieve the pose at angle A which is the angle between left arm and left elbow.So his pose will be marked as not detected as it is not the perfect or close to perfect pose. Other four person could achieve the pose with very little deviation which are detected as perfect pose.

From Table III, deviation of angles for every person posing the reverse warrior yoga pose can be found. If we analyze the results, we can see that person 2 failed to achieve the yoga positions as some of the angles deviated more that the acceptable deviation. Other persons could get the positions with acceptable deviation.

Fig 12,13 and 14 shows the deviation of angles with reference value from accuracy for every person for three poses. Fig 9,10 and 11 shows the accuracy at every angle for every person in every pose.If we further explore the infographs, we can see that if any angle fails to acquire the accuracy of greater than 97% we come to conclusion that the person could not get the pose perfectly.

From the experimental results we can see that our proposal of taking 10 angles between body parts is sufficient enough to detect the yoga poses. We have experimented on 5 different persons and our approach could detect all the poses with a great accuracy. We have set the threshold deviation to 2.5 degree

TABLE I: Deviation of angles from reference model for Pose 1

Experiment No	A	B	C	D	E	F	G	H	I	J	Avg. Deviation	Status
Person 1	1.53	0.05	0.80	1.02	2.01	2.13	0.06	0.50	0.02	0.03	0.815	**Correct**
Person 2	2.56	3.02	1.20	0.45	0.36	.096	0.57	0.69	0.36	1.36	1.067	**Not Correct**
Person 3	0.10	0.06	0.87	0.96	0.54	0.19	1.02	0.06	1.35	1.25	0.640	**Correct**
Person 4	1.23	1.26	0.08	0.64	1.02	0.64	1.25	1.87	2.01	2.08	1.208	**Correct**
Person 5	0.06	0.08	1.06	0.98	1.26	1.47	0.60	0.57	1.68	1.97	0.973	**Correct**

TABLE II: Deviation of angles from reference model for Pose 2

Experiment No	A	B	C	D	E	F	G	H	I	J	Avg. Deviation	Status
Person 1	0.80	0.95	1.54	1.65	0.05	2.03	1.58	0.66	2.15	0.06	1.147	**Correct**
Person 2	1.05	0.52	1.15	1.78	1.36	1.60	0.58	0.99	1.64	0.67	1.231	**Correct**
Person 3	0.00	0.05	2.97	1.87	2.64	0.89	1.37	0.45	1.05	1.69	1.298	**Not Correct**
Person 4	0.27	1.01	1.28	0.78	2.02	1.65	0.85	0.87	1.05	1.02	1.080	**Correct**
Person 5	0.01	0.00	0.18	0.45	0.26	1.07	0.25	0.55	2.08	1.78	0.663	**Correct**

TABLE III: Deviation of angles from reference model for Pose 3

Experiment No	A	B	C	D	E	F	G	H	I	J	Avg. Deviation	Status
Person 1	0.53	0.14	0.47	0.92	1.31	1.45	0.01	1.78	0.03	0.53	0.717	**Correct**
Person 2	3.06	2.42	0.20	0.05	0.36	1.96	0.83	0.72	0.26	0.41	1.027	**Not Correct**
Person 3	1.20	0.86	1.07	0.96	0.34	0.45	1.22	0.16	2.35	1.85	1.046	**Correct**
Person 4	0.23	0.27	0.08	0.24	1.22	0.74	0.67	1.82	1.01	1.56	0.784	**Correct**
Person 5	1.00	0.08	0.58	1.98	1.46	0.47	0.65	1.57	1.08	1.34	1.021	**Correct**

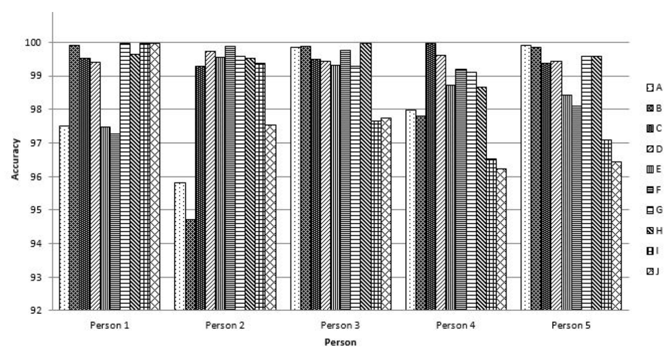


Fig. 9: Accuracy at every angle for pose 1

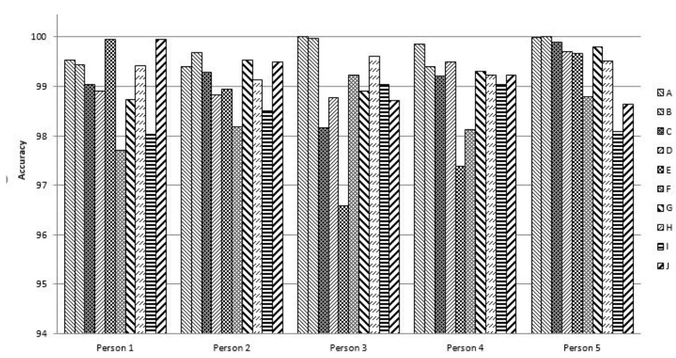


Fig. 10: Accuracy at every angle for pose 2

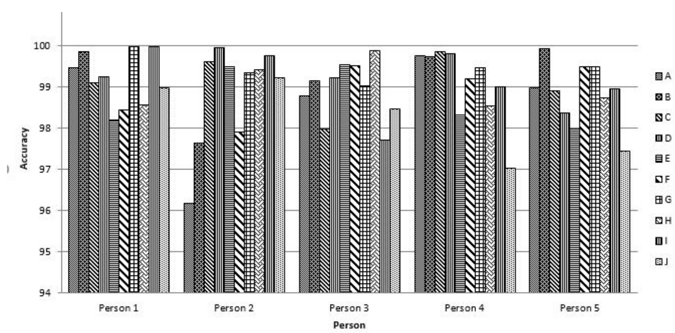


Fig. 11: Accuracy at every angle for pose 3

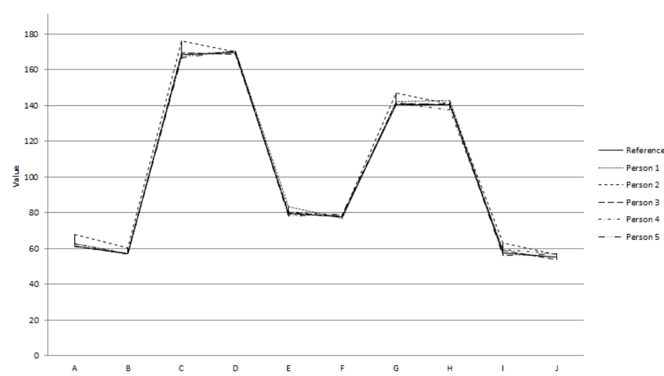


Fig. 12: Accuracy deviation at every angle for pose 1

because successive experimentation shows that if a person deviates more than 2.5 degree which ensures more than 97% accuracy in any angle of a certain pose, he/she cannot acquire the yoga pose perfectly.

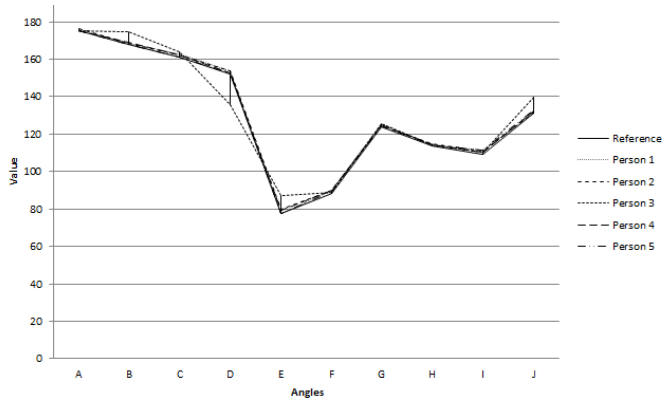


Fig. 13: Accuracy deviation at every angle for pose 2

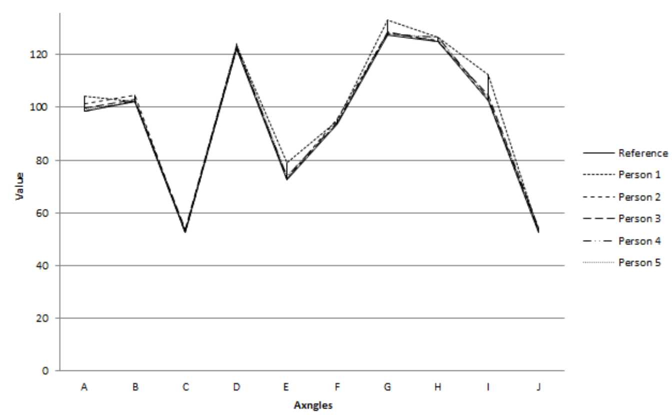


Fig. 14: Accuracy deviation at every angle for pose 3

VI. Conclusion

In this paper, we have proposed a system to recognize three major yoga pose by detecting of human joint points using Microsoft Kinect. We have detected the yoga poses considering above 97% accuracy in every angle between different body parts. Our system can also be used to recognize other yoga poses from the reference model of each pose. Therefore, we hope our proposed approach will help to practise yoga without trainer.

References

[1] O. Patsadu, C. Nukoolkit, and B. Watanapa, "Human gesture recognition using kinect camera," in *Computer Science and Software Engineering (JCSSE), 2012 International Joint Conference on*. IEEE, 2012, pp. 28–32.

[2] S. Obdržálek, G. Kurillo, J. Han, T. Abresch, R. Bajcsy *et al.*, "Real-time human pose detection and tracking for tele-rehabilitation in virtual reality," *Studies in health technology and informatics*, vol. 173, pp. 320–324, 2012.

[3] K. Rector, C. L. Bennett, and J. A. Kientz, "Eyes-free yoga: an exergame using depth cameras for blind & low vision exercise," in *Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, 2013, p. 12.

[4] T.-L. Le, M.-Q. Nguyen *et al.*, "Human posture recognition using human skeleton provided by kinect," in *Computing, Management and Telecommunications (ComManTel), 2013 International Conference on*. IEEE, 2013, pp. 340–345.

[5] Z. Zhang, Y. Liu, A. Li, and M. Wang, "A novel method for user-defined human posture recognition using kinect," in *Image and Signal Processing (CISP), 2014 7th International Congress on*. IEEE, 2014, pp. 736–740.

[6] J. Charles and M. Everingham, "Learning shape models for monocular human pose estimation from the microsoft xbox kinect," in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1202–1208.

[7] F. Pedersoli, S. Benini, N. Adami, and R. Leonardi, "Xkin: an open source framework for hand pose and gesture recognition using kinect," *The Visual Computer*, vol. 30, no. 10, pp. 1107–1122, 2014.

[8] I. Cohen and H. Li, "Inference of human postures by classification of 3d human body shape," in *Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003. IEEE International Workshop on*. IEEE, 2003, pp. 74–81.

[9] H.-C. Mo, J.-J. Leou, and C.-S. Lin, "Human behavior analysis using multiple 2d features and multicategory support vector machine." in *MVA*, 2009, pp. 46–49.

[10] H. Zhao, Z. Liu, and H. Zhang, "Recognizing human activities using non-linear svm decision tree," *Intelligent Computing and Information Science*, pp. 82–92, 2011.

[11] A. Chella, H. Dindo, and I. Infantino, "People tracking and posture recognition for human-robot interaction," in *proc. of International Workshop on Vision Based Human-Robot Interaction, EUROS-2006*, 2006.

[12] L. Goldmann, M. Karaman, and T. Sikora, "Human body posture recognition using mpeg-7 descriptors," in *Proceedings of SPIE*, vol. 5308, 2004, pp. 177–188.

[13] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, "Accuracy of the microsoft kinect sensor for measuring movement in people with parkinson's disease," *Gait & posture*, vol. 39, no. 4, pp. 1062–1068, 2014.

[14] Š. Obdržálek, G. Kurillo, F. Ofli, R. Bajcsy, E. Seto, H. Jimison, and M. Pavel, "Accuracy and robustness of kinect pose estimation in the context of coaching of elderly population," in *Engineering in medicine and biology society (EMBC), 2012 annual international conference of the IEEE*. IEEE, 2012, pp. 1188–1193.

[15] H. Mousavi Hondori and M. Khademi, "A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation," *Journal of Medical Engineering*, vol. 2014, 2014.

[16] B. Lange, S. Koenig, E. McConnell, C.-Y. Chang, R. Juang, E. Suma, M. Bolas, and A. Rizzo, "Interactive game-based rehabilitation using the microsoft kinect," in *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*. IEEE, 2012, pp. 171–172.