Two Hand Gesture Based 3D Navigation in Virtual Environments

I. Rehman*, S. Ullah, M. Raees

University of Malakand, Department of CS & IT, Chakdara (Pakistan)

uniR

LA UNIVERSIDAD
EN INTERNET

Abstract

Natural interaction is gaining popularity due to its simple, attractive, and realistic nature, which realizes direct Human Computer Interaction (HCI). In this paper, we presented a novel two hand gesture based interaction technique for 3 dimensional (3D) navigation in Virtual Environments (VEs). The system used computer vision techniques for the detection of hand gestures (colored thumbs) from real scene and performed different navigation (forward, backward, up, down, left, and right) tasks in the VE. The proposed technique also allow users to efficiently control speed during navigation. The proposed technique is implemented via a VE for experimental purposes. Forty (40) participants performed the experimental study. Experiments revealed that the proposed technique is feasible, easy to learn and use, having less cognitive load on users. Finally gesture recognition engines were used to assess the accuracy and performance of the proposed gestures. kNN achieved high accuracy rates (95.7%) as compared to SVM (95.3%). kNN also has high performance rates in terms of training time (3.16 secs) and prediction speed (6600 obs/sec) as compared to SVM with 6.40 secs and 2900 obs/sec.

I. Introduction

THE effective and realistic interaction is a prerequisite to attain/explore physical/behavioral properties of a Virtual Environment (VE) and its constituent objects. The interaction in VEs ranges from keyboard and mouse where button clicks (key/mouse) carry no information [1], to touch based, which uses 2D gestures with finger/hand actions via touchscreens/interactive panels, and recently to touchless (natural) interfaces which involves 3D gestures in midair [1]. Hand gesture recognition plays a vital role in various applications such as sign language recognition [2-5], virtual and augmented reality [6-10], robotics [11-13], physical (health) sciences [14-16], natural sciences [17], computer engineering [18, 19], and industrial areas [20, 21] etc.

Gesture based interaction is the most attractive form of natural interaction due to the exclusion of physical contact with hardware where input/interaction devices need to be permanently connected (via some physical means) with computer. These physical media (cables) are used for delivery of input to the computer. These media lead to extra burden, need space, extra cost and complexity. Gesture based interaction offers more natural and intuitive HCI with various multimodal forms [22-26].

Different sensors have been used in computer vision and image processing for recognition purposes, such as Koller et al. [27], who used monocular camera for tracking in augmented reality applications. Jalal et al., presented different systems for Human Activity Recognition (HAR) based on depth video [28, 29] and depth imaging [30, 31]. Some authors proposed, Depth Images-based Human Detection and Activity Recognition [32], and Human Pose Estimation and Recognition from RGB-D Video [33].

Feature selection has a vital role in any recognition system. For face recognition different methods have been used. In Holistic Matching method [34, 35], the complete face area is taken as input to the recognition system. In Feature based method, the position and statics of nose, eyes and mouth are considered as input to the system. The Hybrid method uses the combination of both Holistic and feature based methods [36].

For action recognition, Lowe [37] introduced Scale Invariant Feature Transform (SIFT). Dense Sampling Scale Invariant Feature Transform (DSIFT) was used by different authors for action recognition [38-48]. Histogram of oriented gradient (HOG) was used by [41, 47-55]. Shape context (SC) was proposed by Belongie and Malik [56] for feature extraction and also was used by Wang et al. [57], Gupta et al. [51], and Yao and Fei-Fei [58]. For recognition of action from still images, GIST was proposed by Oliva and Torralba [59] and used by Gupta et al. [51], Prest et al. [60], and Li and Ma [48]. The Speeded Up Robust Features (SURF), is proposed by Bay et al. [61] and used by Ikizler et al. [62] to represent the human silhouettes for action recognition.

Different techniques are proposed by various authors for gait recognition. Ahmed et al. [63], used horizontal and vertical distances of selected joint pairs. Andersson et al. [64], calculated mean and standard deviation in the signals of lower joint angles. Ball et al. [65] used mean, standard deviation and maximum of the signals of lower joint angles. Dikovski et al. [66] proposed a set of seven different features, such as joint angles and inter-joint distances aggregated within a gait cycle, body parameters, along with various statistics. Kwolek et al. [67], used static body parameter, bone rotations, and the person's height. Preis et al. [68], presented 13 pose attributes. Sinha et al. [8], used multiple gait features, upper and lower body area, inter-joint distances and other

* Corresponding author.

E-mail address: inam.btk@gmail.com

features [65, 68]. Skin Joints Features were also proposed by [69, 70] for gait recognition.

Different approaches are proposed for hand recognition such as mount based sensors [71], multi-touch screen sensors[72, 73] and vision based sensors [74-76]. The depth based hand gesture recognition have three types i.e. static hand gesture recognition [77], hand trajectory gesture recognition [78, 79] and continuous hand gesture recognition [80, 81]. Most of the authors used computer vision and image processing methods [82-85], along with some newly introduced input devices such as Leap Motion [86-88] and Kinect [89]. For a natural interaction with AR environments, fiducial markers are used with fingers [90, 91]. Different computer vision techniques are used for detection of hand and fingertips for AR interaction [92-94]. These systems are commercially limited due to problems such as skin color and precise depth sense [94]. Different glove based techniques have been used for accurate interaction [95-97] but limited due to its cumbersome nature.

For gesture based navigation, different systems have been proposed so far, but have limited commercial application due to cumbersome or inaccurate nature, cost, or dependency/need for special devices and their limited range. Recent research mostly stresses to deal with these problems, but simple and intuitive interaction is still the major area than needs to be improved. The previous gesture based navigation techniques, mostly rely on the coarse/unrealistic alteration/shape of hands and fingers layout for transition among different gestures which results in an increased physical and mental load on the user.

In this paper, we propose a novel two hand gesture based 3D navigation technique for VEs with the objective of providing intuitive and easy navigation. Navigation includes 3D movement, i.e. forward, backward, up, down, left, and right along with an effective speed control mechanism. Computer vision techniques are used for detection of gestures (colored thumb, fingers) from the real scene while OpenGl is used as a front end for navigation in the VE. Machine learning tools such as SVM and kNN are used to assess the accuracy and performance of the proposed gestures.

The rest of the paper is organized as follows: section 2 presents related work, section 3 describes the proposed system, section 4 consists of experiments and evaluation and finally section 5 is related to conclusion and future work.

II. RELATED WORK

In daily life communication, hand gestures cover the gap of merely verbal information, and so is a necessary part of effective and meaningful communication with the receiver. In HCI, hand gesture based interaction is more valuable due its natural and attractive nature. Hand gestures refer to meaningful movement of the hand and fingers[86], which entails most valuable information[98].

In the past different navigation techniques in VEs have been proposed [99, 100]. Different types of sensing techniques have been used for recognition of patterns in gestures [101]. Different sensing systems have been proposed so far, such as glove based, vision based, along with some newly introduced devices such Leap Motion [102] and Kinect [89].

Glove based devices use movement based approach with high performance in some applications such as recognition of sign language [103]. CyberGlove, a type of data gloves is used for tracking hand gestures [104]. Cooper et al. [105] used color coded glove for tracking of hand movement, but the system needs wearable gloves which decrease user's experience in the environment. Kim et al. [106], used a Cyber Glove, a wearable device for recognition of hand gestures and performed different navigation tasks as shown in Fig. 1.
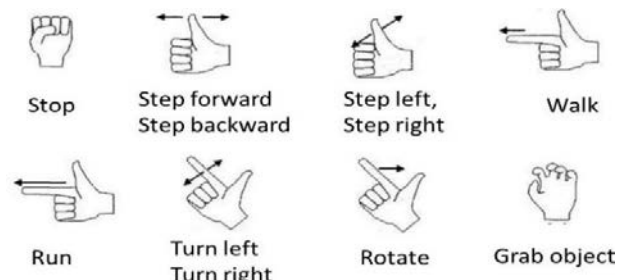

Fig. 1. Gestures with Cyber Glove.

Although wearable devices were used mostly for gesture based interaction in the past, the cumbersome and costly nature of gloves, limit its widespread use in HCI [101].

Chen et al. [107], used computer vision for the detection of hand gestures from the video taken by a webcam. Two types of hand gestures were proposed i.e. appearance based which used bare hand and marker based (with colored markers on a black glove) 3D hand model. For moving a virtual car, different gestures were proposed as shown in Fig. 2. The system is unable to provide complete 3D movement and speed control. The system has less degree of correspondence with real world navigation.

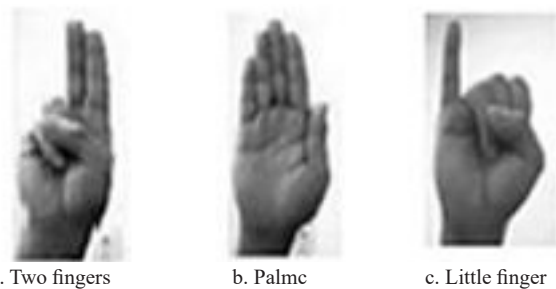

a. Two fingers   b. Palmc   c. Little finger

Fig. 2. Gestures for virtual car navigation [107].

Krum et al. [108], presented a navigation interface (earth 3D visualization) using verbal and hand gestures. The system used image processing techniques for detection of hand gestures taken from a Gesture Pendant video camera. Multiple infrared emitting LEDs were used for illuminating the hand gestures in front of camera. Different types of hand gestures were used for navigation.

Shao Lin [109], used Leap Motion controller for detection of hand gestures. The proposed technique used both hands for three different types of gestures as shown in Fig. 3. Similarly clockwise and anti-clockwise rotation of hands produced the same rotation in the VE but the technique has no mechanism for up, down movements and speed control.

Kerefeyn et al. [110], used different gestures for controlling and manipulation of virtual objects in a VE using Leap Motion controller. The system proposed five different gestures for interaction with virtual objects using right hand (see Fig. 4). The system does not provide complete 3D navigation i.e. forward, backward, up, down, left, and right as well as there is no mechanism for speed control.

Khundam et al. [111], proposed single hand gestures via Leap Motion controller for navigation in VE as shown in Fig. 5. Forward movement is done by raising the hand with straight palm. While backward movement is achieved via turning palm direction or reversing hand facing. Pushing the palm to left causes step movement to left side while its movement to right results in step right. Grasping or moving out of display area causes hold position. Forward speed depends on advance movement of palm while pulling of palm to body controls backward speed. More or less movement towards left or right side causes more or less speed. The proposed gestures are hard to learn.
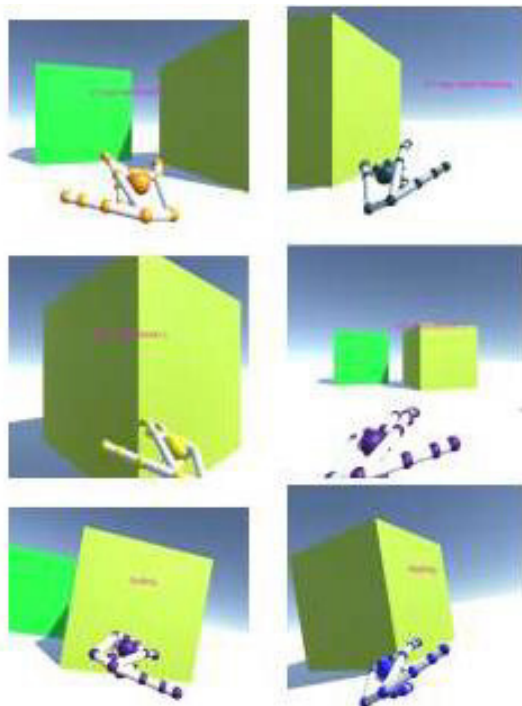
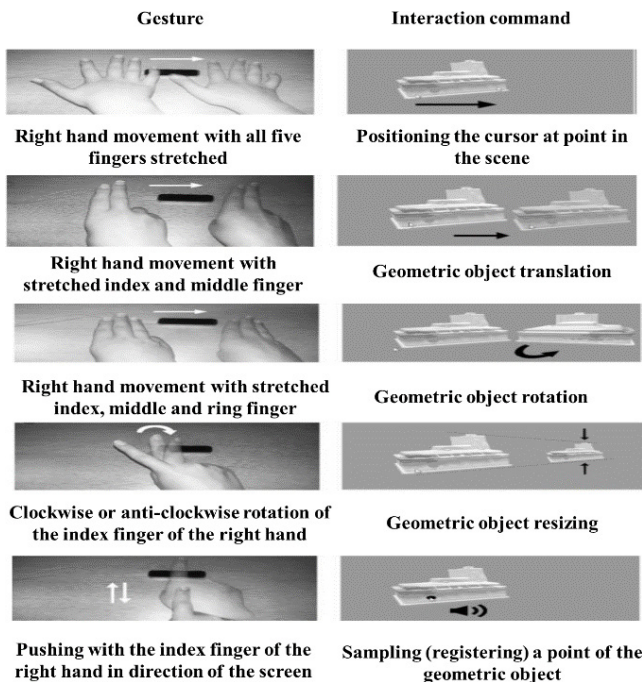Fig. 3. Gesture based navigation by Shao Lin [108].

Fig. 4. Gesture based interaction by Kerefeyn et al. [110].

Fig. 5. Hand gestures proposed by Khundam et al. [111].

Batista et al. [112], used Leap Motion controller and proposed different hand gestures as shown in Fig. 6, for controlling an avatar in a virtual tour but there is no mechanism for up/down movement and turning left/right.
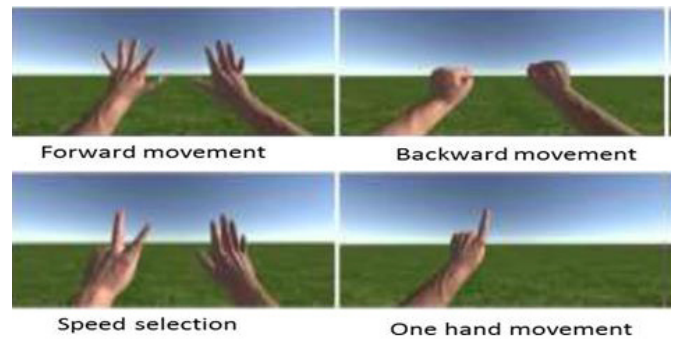

Fig. 6. Gestures proposed by Batista et al. [112].

Liang et al. [113] presented a system using Leap Motion controller for the detection of hand gestures. There were different types of modules for navigation such as single hand gestures, designed for children which causes avatar in VE to fly left, right, up, and down as shown in Fig. 7.

Hand gesture	Movement	Target action
	Move right	Fly to the right
	Move left	Fly to the left
	Move down	Fly down
	Move up	Fly up
	Stretch	Hover
	Stretch to grip	Grasp stick
	Grip to stretch	Drop stick

Fig. 7. Hand gestures proposed by Liang et al. [113].

Leap Motion controller is unable to detect all fingers of hand, specially middle and pinky finger, moreover, it has a limited working space [109]. Other problems include misdetection during overlapping of hands [114], crossing of field boundaries, and varying lighting conditions [112]. Leap Motion gives inaccurate results for hands beyond 250mm upside the controller as well as it gives unstable sampling frequency [115]. A comparative study conducted by [114] states that Kinect provides inaccurate but comprehensive data while Leap Motion gives comparatively accurate results. Nabiyouni et al. [114] stated that Leap Motion fails to recognize cross over fingers or if they are next to each other. Rotation of palm or fingers more than 80 degrees causes failure in tracking. Moreover it produces significant fatigue.

Kinect is used for detection of full human body gestures to interact with virtual objects [117]. Kumerburg et al. [116] used Microsoft Kinect as input device for navigation in a VE. Different gestures for navigation were proposed such as raising both arms upside for fly and forward movement as shown in Fig. 8. The proposed navigation gestures were hard to learn and use.

Gesture	Command	Characteristic
	Tilt up	Bend backward with hands down
	Turn left	Raise the left hand above shoulder sideways with the right hand down
	Pan left	Move left hand across the body and right hand down
	Fly, move foward	Raise both arms up
	Fly pan right	Raise both arms and move right hand across head

Fig. 8. Gestures proposed by Kumerburg et al. [116] using Microsoft Kinect.

Vulture et al. [118], also used Microsoft Kinect and proposed gesture for navigation using both arms as shown in Fig. 9.

Gesture	Characteristic
	Foward navigation
	Backward navigation
	Navigation to the right
	Navigation to the left
	Up to a higher level
	Down to a lower level

Fig. 9. Gestures proposed by Vulture et al. [118].

The previous techniques for gesture based navigation mostly depend on variant shapes/layouts of hands, fingers, and arms which lead to extra mental and physical load on user in learning and usage. They are limited in use due to less realistic nature.

III. Proposed System

We propose a new, two hand gesture based navigation technique for VEs with a close resemblance to car steering driving. The relative position of both thumbs determines various gestures. These gestures are used for 3D navigation in the VE. Navigation includes Forward, Backward, Up, Down, and Left and Right side movement. The green and yellow color caps made of paper/rubber are used with thumbs. The VE consists of different 3D objects as shown in Fig. 10. Identification of different gestures leads the virtual object/camera to navigate accordingly.
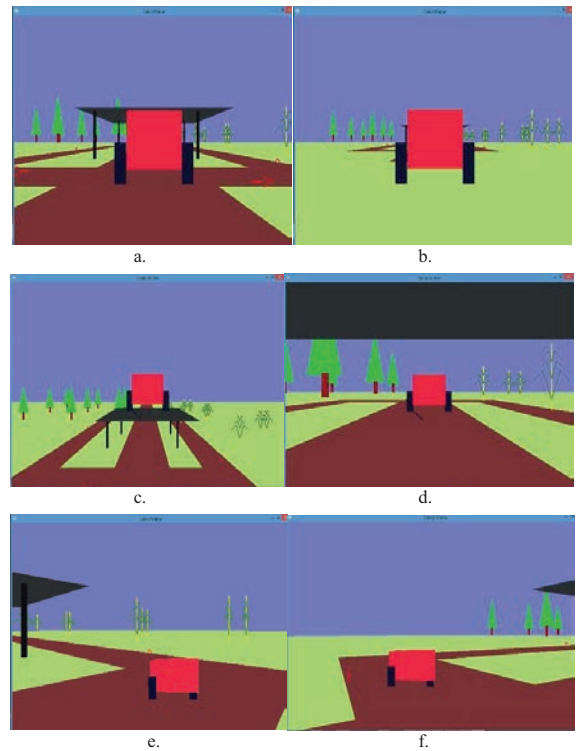

Fig. 10. Scenario of VE (from different perspectives). (a). forward movement, (b). backward movement, (c). Upward movement, (d). Downward movement, (e). Right turn, (f). Left turn.

A. System Architecture

The proposed system uses OpenCV as backend and OpenGl as frontend tool. OpenCV is used for image processing, consists of different phases such as image acquisition, conversion of image to HSV (Hue, Saturation, Value), thresholding for green and yellow colors, and finally the calculation of 2D position (x, y) of the specified colors as shown in Fig. 11. OpenGL is used for designing and interaction of the VE. OpenGL, based on position of colors (Green and Yellow), identifies various gestures that leads to different navigation tasks in the VE.
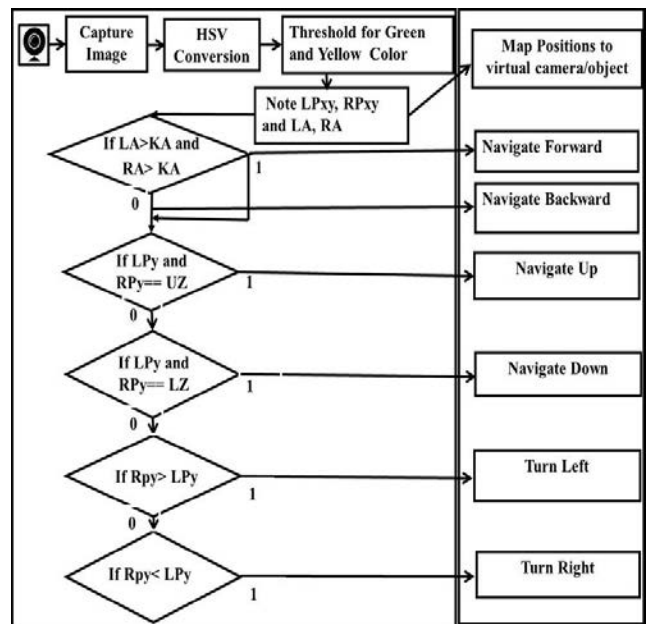

Fig. 11. (a) Backend (Image processing), (b) Frontend (VE).

First of all OpenCV performs image acquisition via a camera. The image is then converted to HSV for realistic performance which is then thresholded for Green and Yellow colors. In the last stage, position of both colors is calculated dynamically from the image. OpenGL receives the positions of both thumbs colors. On the basis of these positions, different gestures are identified which lead to 3D navigation in the VE.

OpenCV performs image acquisition via ordinary webcam. Finger caps of green and yellow colors are used for left and right thumbs of both hands. Skin color is omitted to achieve best results as it varies from person to person. A rang of Hue, Saturation and Values are selected for green and yellow colors to detect thumbs in stable lighting conditions. First the region of interest of the image (RI_img) is extracted from the Frame Image (F_img) to avoid false detection of the background green and yellow colors. The IR_img is then segmented from the F_img based on the skin area which is the most probable area to get the thumb fingers. YCbCr model has the capability to distinguish between skin and non-skin colors [119].

$$F-Img-YCbCr \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.1 & 128 & 24.8 \\ -37.3 & -74 & 110 \\ 110 & -93.2 & -18.2 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$
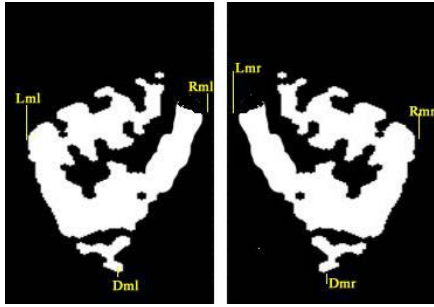

Fig. 12. Binary images of both hands after conversion to YCbCr.

After getting the binary image of F-Img (see Fig. 12), *RI-Img* with rows *'m'* and column *'n'* is extracted from *F-Img* using algorithm [120] as,

$$RI\_Img(m,n) = \left(\bigcup_{r=Dml}^{F\_Img.Row(0)}(F\_Img), \bigcup_{c=Lml}^{F\_Img.Colum(Rml)}(F\_Img) \right) \quad (2)$$

Where *Lml*, *Rml* and *Dml* represent Left-most, Righ-most and Down-most skin pixels of the left hand.

$$RI\_Img(m,n) = \left(\bigcup_{r=Dmr}^{F\_Img.Row(0)}(F\_Img), \bigcup_{c=Lmr}^{F\_Img.Colum(Rmr)}(F\_Img) \right) \quad (3)$$

Where *Lmr*, *Rmr* and *Dmr* represent *Left-most*, *Righ-most* and *Down-most* skin pixels of the right hand.

The segmented image RI-Img is then thresholded for green and yellow color simultaneously, using HSV color space (see Fig. 13).

$$RI\_Img(x,y)= \begin{cases} 1 \text{ if } RI\_Img.H(x,y) \geq 63 \wedge RI\_Img.H(x,y) \leq 98 \\ 1 \text{ if } RI\_Img.S(x,y) \geq 98 \wedge RI\_Img.S(x,y) \leq 255 \\ 1 \text{ if } RI\_Img.V(x,y) \geq 37 \wedge RI\_Img.V(x,y) \leq 255 \\ 0, \text{ Otherwise} \end{cases} \quad (4)$$

$$RI\_Img(x,y)= \begin{cases} 1 \text{ if } RI\_Img.H(x,y) \geq 65 \wedge RI\_Img.H(x,y) \leq 80 \\ 1 \text{ if } RI\_Img.S(x,y) \geq 110 \wedge RI\_Img.S(x,y) \leq 255 \\ 1 \text{ if } RI\_Img.V(x,y) \geq 145 \wedge RI\_Img.V(x,y) \leq 255 \\ 0, \text{ Otherwise} \end{cases} \quad (5)$$
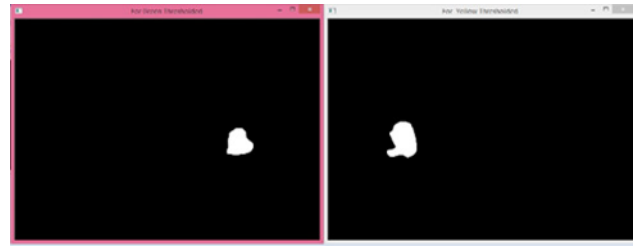

Fig. 13. Detected thumbs of both hands.

B. Navigation

Navigation is the movement towards the desired position in a VE. OpenCV computes positions of both thumbs in 2D. The z-axis movement is deduced from the area variation of the detected thumbs. As the area of the thumbs increases with inward and decreases with outward movement in z-axis (towards the camera eye) as shown in Fig. 14 (a, b). So increase from a predefined (threshold) area KA results in forward navigation while decrease results in backward navigation. The value of KA is half of the fully detected thumb area (near to camera eye) which divides the navigation space (z-axis) into two zones i.e. Forward and Backward zone, as shown in Fig. 15. For accurate navigation in the VE, it is necessary that both thumbs should be visible to camera. Movement of both thumbs forward (where detected thumb area >KA) towards camera eye results in forward navigation while moving thumbs backward (where detected thumb area <KA) produces backward navigation.

LPx and LPy represent position of left hand thumb (with green cap) in x and y-axis, RPx and RPy are position of right hand thumb (with yellow color) in x and y-axis, UZ is the upper zone, LZ is the lower zone, and LA and RA represent the detected thumb areas.

The algorithm for forward and backward movement is given below:

If(LA>KA AND RP>KA)

Forward navigation

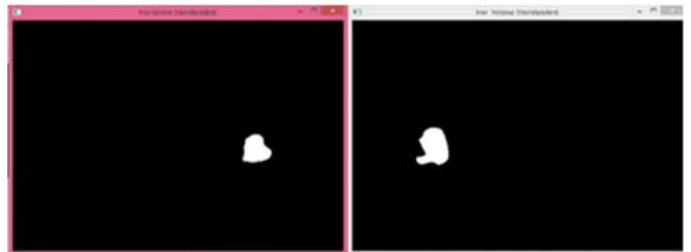If(LA<KA AND RA<KA)

Backward navigation

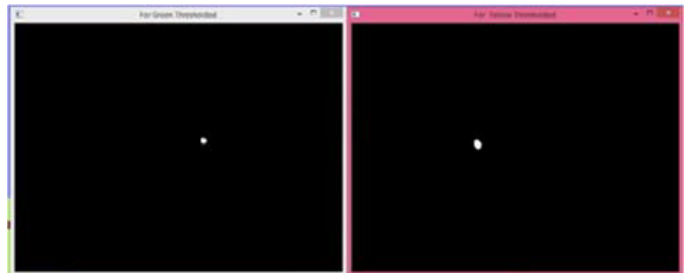Fig. 14. (a) Detected (Left and Right hand) areas in forward navigation.

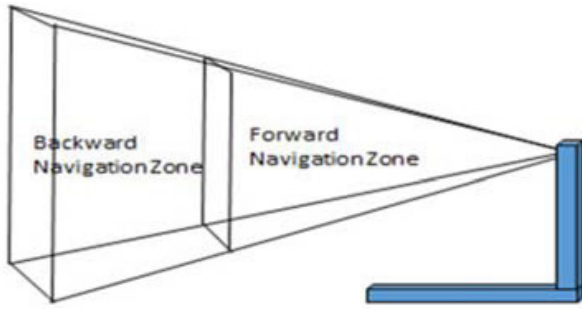Fig. 14. (b) Detected (Left and Right hand) areas in backward movement.

Fig. 15. Forward and backward navigation zones.

For navigation in y-axis, the thumbs of both hands use UZ or LZ. If both thumbs simultaneously move to upper zone (UZ), upward navigation (along y-axis) is produced, while at lower zone (LZ), it leads to downward navigation as shown in Fig. 16.
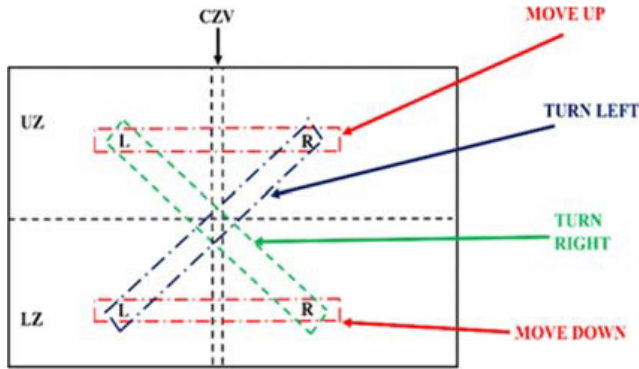

Fig. 16. Navigation space for gestures.

RPx,y is the right hand position (x,y), LPx,y left hand position (x,y), UZ upper zone, UL lower zone, and CZV is central zone vertical.

The algorithm for moving in vertical (y-axis) direction is given below:

If(RPy== UZ AND LPy== UZ)

Upward movement (along +ve y-axis) (see Fig. 17.a)

If(RPy== LZ AND LPy== LZ)
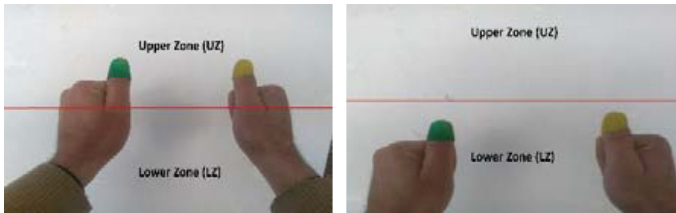
Downward movement (along -ve y-axis) (see Fig. 17.b)

a.   b.

Fig. 17. Gestures for movement (a) Up (b) Down.

C. Speed Control

Speed control (Sp) is an important requirement in any navigation technique. In the proposed technique, speed can be controlled via changing the horizontal (x-axis) distance between both thumbs in navigation and turning. The speed remains normal for maximum distance while it increases with decrease in relative distance between both thumbs i.e. the speed is inversely proportional to the relative distance between both thumbs.

Mathematically it can be written as:

$$Sp \propto 1/abs(RPx - LPx) \tag{6}$$

D. Turning

Left or right turn is taken by simply comparing the y-position of both thumbs. For left turn, the right hand thumb will be in upper zone (UZ) while left hand thumb will be in lower zone (LZ). For right turn, the right hand thumb will be in lower zone (LZ) while left hand thumb will be in upper zone (UZ) as shown in Fig. 16.

The algorithm for turning right and left is given below:

If(RPy> LPy)

Turn left (see Fig. 18.a)

If(RPy< LPy)

Turn right (see Fig. 18.b)

a.   b.

Fig. 18. Gestures for Turning (a) Left (b) Right.

E. System Implementation

The proposed system was implemented using Visual Studio 2013 with corei3 laptop having 1.7 GHz processor, 4GB RAM, and 640x480 resolution low cost built in camera. The OpenCV, as a backend tool, performs different image processing tasks such as acquisition of image from camera, identification of thumbs colors (green and yellow), and dynamic area and pose calculation of these colors. OpenGL, as a frontend tool is responsible for creation and interaction with VE as shown in Fig. 19. The system allows users to interact (navigate) with VE using his/her both hands thumbs having colored (green, yellow) caps. The left hand thumb uses green color while right hand thumb uses yellow color cap. The combination and relative position of both colored thumb makes different gestures which are used to perform 3D navigation in the VE.
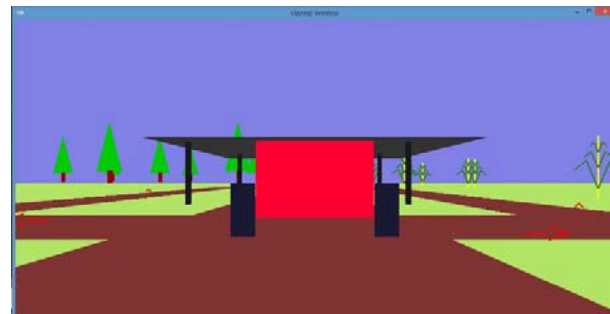

Fig. 19. Screen shot of the experimental scenario.

IV. Experiments and Evaluation

We performed objective and subjective evaluation to assess the accuracy and effectiveness of the proposed navigation technique in VEs. We also used machine learning models i.e. SVM and kNN to assess the accuracy and performance of the proposed gestures.

A. Protocol and Task

Forty (40) volunteer male students participated in the experimental study. Their ages were in the range of 22 to 35 years. All of them had gaming experience with keyboard, mouse, and touch screen but had no experience with gesture based VEs. In the training phase all the students were demonstrated about the use of proposed system and gestures. After that they performed different navigation tasks (Forward, Backward, Up, Down, and turning Left, and Right) using the specified gestures in the VE. Each student performed five pre-trails of each navigation task. After the training session, each participant performed four trails of three different navigation tasks in three different lighting conditions making a total of 1440 trails.

1) Interaction Routes

The experimental environment consisted of three different routes from start to stop position as shown in Fig. 20.
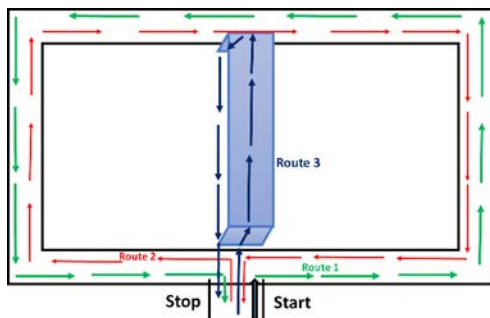

Fig. 20. Complete 2D route model in VE.

2) Interaction task

The students need to perform four (4) trials of each of the following three tasks:

- Task 1. The first task was to follow the Route 1 which covers forward navigation (five times), one right turn and four left turns as shown in Fig. 20.
- Task 2. The second task was to follow the Route 2, which covers five forward movements, one left and four right turns.
- Task 3. The third task was to follow the Route 3 which covers one upward, two forward, one downward, and one backward movement.

3) Lighting conditions

The proposed system uses colored thumbs for interaction with the VE. The detection thumb is highly dependent on the surrounding light. So we performed all the tasks in three different lighting intensities.

- Low light intensity (3-7 Lux) Dark limit of civil twilight under a clear sky [128].
- Medium light intensity (50-70 Lux) Family living room lights (Australia, 1998) [129] to office building hallway/toilet lighting [130-131].
- High light intensity (200-500 Lux) Office lighting [129] [132-133], sunrise or sunset on a clear day.

After performing the tasks, task completion time (TCT) and errors for each task were recorded for objective analysis. Misdetection and deviation from the specified route were considered as errors. Finally, each participant filled a questionnaire for subjective analysis.

B. Result Analysis

In this section, we performed objective analysis (task completion time and errors) and subjective analysis (questionnaire) to assess the accuracy and effectiveness of the proposed navigation technique in VEs.

Objective Analysis

1) Task Completion Time and Errors

The Mean and SD of time and errors for task 1 is shown in Fig. 21 and 22. The Mean time and SD is (89.01, 20.02), (82.30, 19.19), and (61.98, 15.16) for low, medium, and high lighting intensity respectively. The Mean and SD of errors is (3.31, 1.49), (2.19, 1.37), (1.71, 0.97) for low, medium, and high light intensities. The Mean for both time and errors are minimum for high lighting intensity as compared to other. It means that the technique has good performance in higher lighting intensity.

Fig. 21. Mean and SD of time for Task 1.
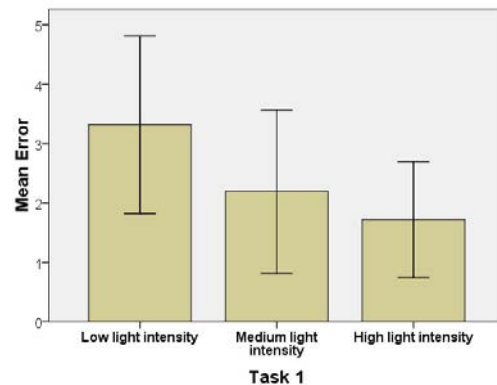
Fig. 22. Mean and SD of errors for Task 1.

The Mean and SD of time and errors for task 2 is shown in Fig. 23 and 24. The Mean time and SD is (68.24, 17.80), (62.53, 15.00), and (51.48, 14.06) for low, medium, and high light intensity respectively. The Mean and SD for errors is (2.31, 1.33), (1.96, 1.14), (1.67, 1.02) for low, medium, and high light intensities. The Mean and SD for both time and errors decreases with increase in lighting.

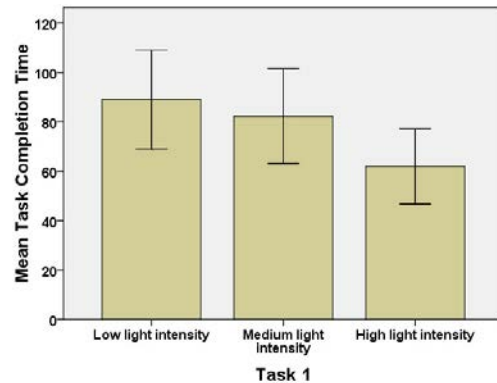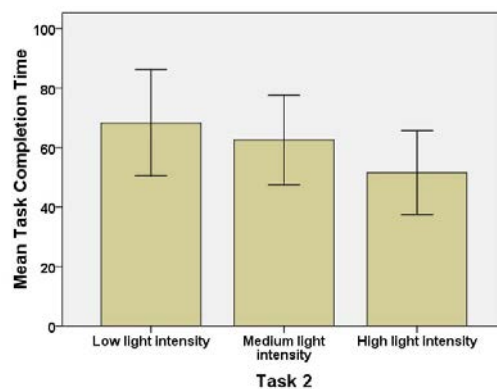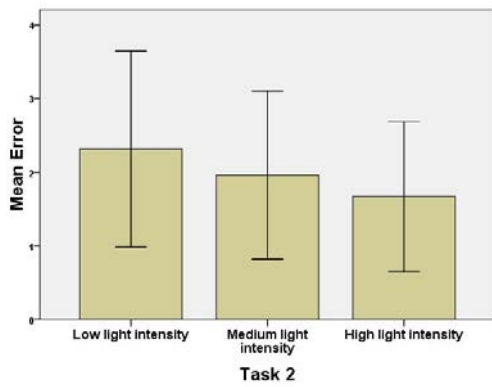Fig. 23. Mean time and SD for Task 2.

Fig. 24. Mean error and SD for Task 2.

The Mean and SD of time and errors for task 3 is shown in Fig. 25 and 26. The Mean and SD of time is (62.01, 14.11), (56.28, 13.58), and (50.24, 13.89) respectively for low (3-7 Lux), medium (50-70 Lux), and high (200-500 Lux) light intensity.
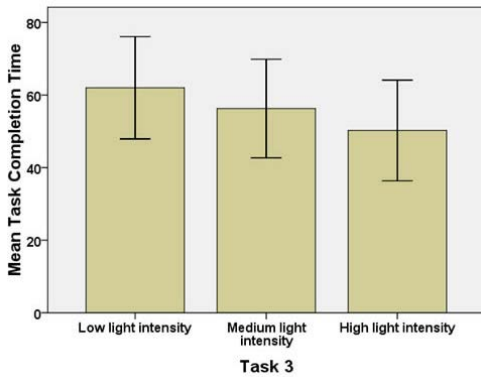
Fig. 25. Mean time and SD for Task 3.
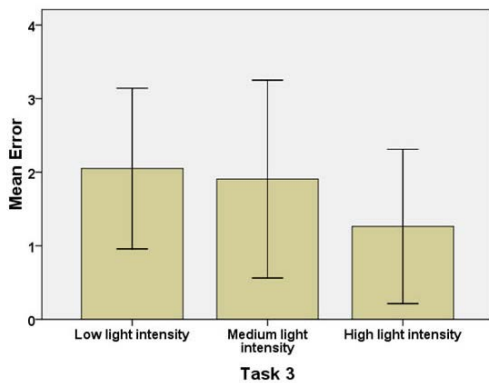
Fig. 26. Mean error and SD for Task 3.

2) Task Learning

The mean task completion time of each task is shown in Fig. 27, 28, and 29. The results show that the mean time decreases with each successive trial which leads to improved task learning. So it means that task learning improves with experience. The system intrinsically works better in high lighting conditions, while the repetition of task improves learning of task performance. The enhanced learning effect is obvious due to simple and realistic nature of gestures.
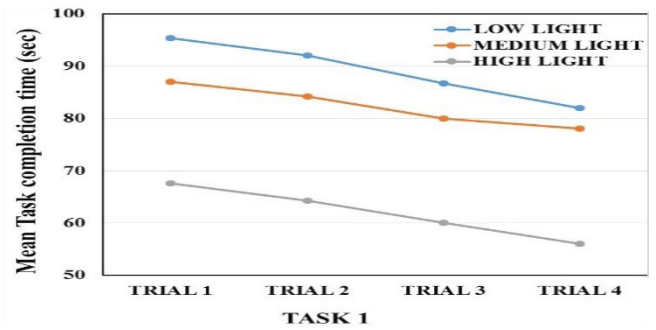
Fig. 27. Mean time of each trial for Task 1.
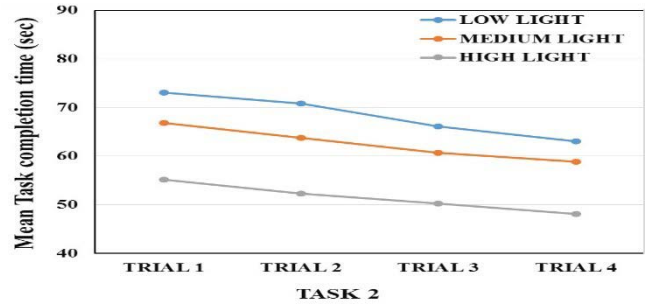
Fig. 28. Mean time of each trial for Task 2.
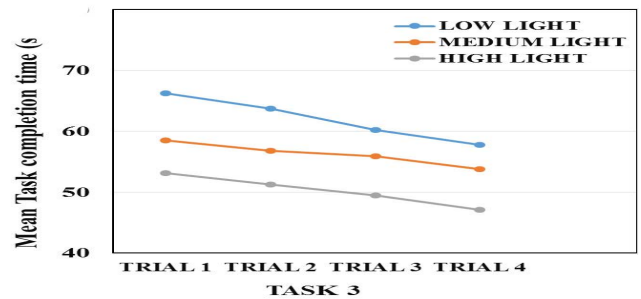
Fig. 29. Mean time of each trial for Task 3.

Subjective Evaluation

For the subjective evaluation, each student filled a questionnaire. The questionnaire was consisted of different questions related to the following topics:

1. Cognitive load during interaction
2. Technique learning
3. Fatigue during interaction

Likert scale is used for scaling purposes where 1 is for lowest level and 5 for highest level.

1) Cognitive Load During Interaction

The main goal of the proposed technique is to use simple and realistic alignment of fingers or hands in order to get an easy to learn and use interaction. The responses for the question concerned to cognitive load on users during interaction is shown in Fig. 30. The results show that most of the students opted for lowest (52.5%) and low level (35%) of mental load created during interaction.
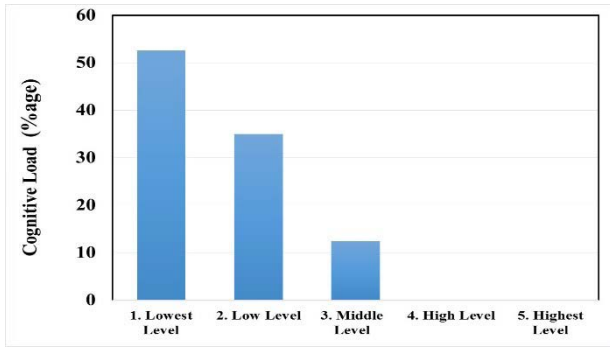
Fig. 30. Cognitive load on user during interaction in VE.

2) Technique Learning

The navigation technique is developed with the aim of ease in learning. Most of the students selected the proposed technique as easy to learn as shown in Fig. 31. This figure shows that 35% and 40% students opted the highest and high level learning. We used real word phenomena/experience for the proposed navigation technique.
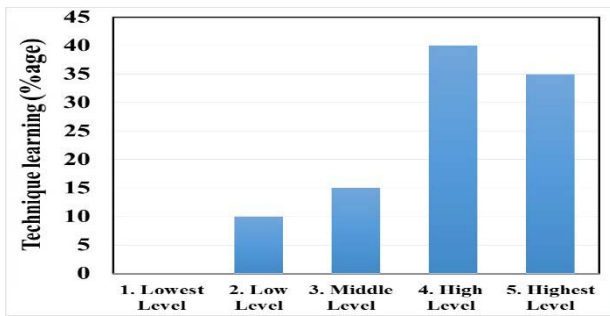
Fig. 31. Technique learning.

3) Fatigue

The students responses concerning the created fatigue during interaction is shown in Fig. 32. The results show that there is a considerable fatigue (40% students opted for high level and 10% for highest) developed during interaction. So we can conclude that fatigue (physical) increases gradually with time in restless operations while the main reason for creation of fatigue is due to baseless/midair operation of hands gestures for long times.
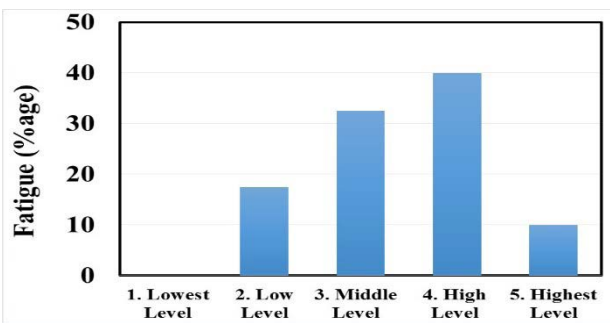
Fig. 32. Fatigue in interaction.

Performance Evaluation of Gesture Recognition

Various authors used different gesture recognition engines for recognition such as HMM [121-124], Advanced HMM [125], Combined HMM and SVM [126], and kNN [127]. We used SVM and kNN models to assess the accuracy of the proposed gestures. To evaluate the performance of the proposed system, 20 participants were

selected to perform 10 trials of each of the six (6) different gestures i.e. Forward, backward, upward, downward, left and right. It makes a total of 1200 gestures.

The confusion matrix for SVM (see Table I) shows 95.3% mean accuracy of the proposed system with 97% (maximum) accuracy for right gesture and 93% (minimum) accuracy for left gesture.

TABLE I. Confusion Matrix For Svm

The confusion matrix for the proposed system with kNN model shows (see Table II) mean accuracy of 95.7% with 99.0% (maximum) accuracy for upward gestures and 94% (minimum) accuracy for left and right gesture.

TABLE II. Confusion Matrix for KNN

1) Comparison of SVM and kNN

Comparison of both SVM and kNN in terms of accuracy and performance is shown in Table III. The results show that kNN has high recognition accuracy (95.7 %) as compared to SVM (95.3%). The kNN also has high performance rates in terms of training time (3.16 secs) and prediction speed (6600 obs/sec) as compared to SVM with 6.40 secs and (2900 obs/sec).

TABLE III. Comparison of SVM and KNN Based on Recognition Accuracy, Training Time, and Prediction Speed

Method	Recognition accuracy	Training Time	Prediction Speed
SVM	95.3%	6.40 secs	2900 obs/sec
kNN	95.7%	3.16 secs	6600 obs/sec

V. Conclusion and Future Work

In this paper, we proposed a novel two hand gesture based interaction technique for 3 dimensional (3D) navigation in virtual environments (VEs). The system uses computer vision techniques for the detection of hand gestures (colored thumbs) from real scene and performs different navigation (forward, backward, up, down, left, and right) tasks in the VE. The proposed technique also allows users to efficiently control speed during navigation. This is implemented via a VE for experimental purposes. Forty participants tested the proposed technique in different lighting scenarios. Experiments revealed that the technique is feasible in normal lighting conditions, easy to learn and use, having less cognitive load on users. Its performance is evaluated using gesture recognition engines i.e. SVM and kNN. kNN achieves high accuracy rates 95.7% as compared to SVM (95.3). kNN also has high performance rates in terms of training time (3.16 secs) and prediction speed (6600 obs/sec) as compared to SVM with 6.40 secs and (2900 obs/sec).

The proposed system is sensitive to lighting conditions.

In future we will compare our proposed camera based system with Motion leap in terms of accuracy, interaction space, update rate, and positional distortion.

References

[1] Saffer, D., *Designing gestural interfaces: Touchscreens and interactive devices*. 2008: " O'Reilly Media, Inc.".

[2] Yang, R., S. Sarkar, and B. Loeding, *Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming*. IEEE transactions on pattern analysis and machine intelligence, 2010. **32**(3): p. 462-477.

[3] Zafrulla, Z., et al. *American sign language recognition with the kinect*. in *Proceedings of the 13th international conference on multimodal interfaces*. 2011. ACM.

[4] Uebersax, D., et al. *Real-time sign language letter and word recognition from depth data*. in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. 2011. IEEE.

[5] Pugeault, N. and R. Bowden. *Spelling it out: Real-time ASL fingerspelling recognition*. in *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on*. 2011. IEEE.

[6] Wickeroth, D., P. Benolken, and U. Lang. *Markerless gesture based interaction for design review scenarios*. in *Applications of Digital Information and Web Technologies, 2009. ICADIWT'09. Second International Conference on the*. 2009. IEEE.

[7] Frati, V. and D. Prattichizzo. *Using Kinect for hand tracking and rendering in wearable haptics*. in *World Haptics Conference (WHC), 2011 IEEE*. 2011. IEEE.

[8] Choi, J., H. Park, and J.-I. Park. *Hand shape recognition using distance transform and shape decomposition*. in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. 2011. IEEE.

[9] Guo, Z.-m. *Research of hand positioning and gesture recognition based on binocular vision*. in *VR Innovation (ISVRI), 2011 IEEE International Symposium on*. 2011. IEEE.

[10] Raees, M., S. Ullah, and S.U. Rahman, *VEN-3DVE: vision based egocentric navigation for 3D virtual environments*. International Journal on Interactive Design and Manufacturing (IJIDeM), 2018: p. 1-11.

[11] Wu, X.-H., M.-C. Su, and P.-C. Wang. *A hand-gesture-based control interface for a car-robot*. in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. 2010. IEEE.

[12] Jalal, A., et al., *Human activity recognition via recognized body parts of human depth silhouettes for residents monitoring services at smart home*. Indoor and Built Environment, 2013. **22**(1): p. 271-279.

[13] Zhan, Y. and T. Kuroda, *Wearable sensor-based human activity recognition from environmental background sounds*. Journal of Ambient Intelligence and Humanized Computing, 2014. **5**(1): p. 77-89.

[14] Jalal, A. and Y.A. Rasheed. *Collaboration achievement along with performance maintenance in video streaming*. in *Interactive Computer Aided Learning, IEEE Conference on*. 2007.

[15] Jalal, A. and A. Shahzad. *Multiple facial feature detection using vertex-modeling structure*. in *Proceedings of the IEEE Conference on Interactive Computer Aided Learning, Villach, Austria*. 2007.

[16] Jalal, A., S. Kim, and B. Yun. *Assembled algorithm in the real-time H. 263 codec for advanced performance*. in *Enterprise networking and Computing in Healthcare Industry, 2005. HEALTHCOM 2005. Proceedings of 7th International Workshop on*. 2005. IEEE.

[17] Jalal, A. and S. Kim, *Advanced performance achievement using multi-algorithmic approach of video transcoder for low bit rate wireless communication*. ICGST International Journal on Graphics, Vision and Image Processing, 2005. **5**(9): p. 27-32.

[18] Jalal, A. *Security architecture for third generation (3g) using gmhs cellular network*. in *Emerging Technologies, 2007. ICET 2007. International Conference on*. 2007. IEEE.

[19] Jalal, A. and M.A. Zeb, *Security enhancement for e-learning portal*. International Journal of Computer Science and Network Security, 2008. **8**(3): p. 41-45.

[20] Jalal, A. and S. Kim. *The mechanism of edge detection using the block matching criteria for the motion estimation*. in *Proceedings of the Conference on Human Computer Interaction, Daegu, Korea*. 2005.

[21] Jalal, A. and S. Kim. *Algorithmic implementation and efficiency maintenance of real-time environment using low-bitrate wireless communication*. in *Software Technologies for Future Embedded and Ubiquitous Systems, 2006 and the 2006 Second International Workshop on Collaborative Computing, Integration, and Assurance. SEUS 2006/ WCCIA 2006. The Fourth IEEE Workshop on*. 2006. IEEE.

[22] Cassell, J., *A framework for gesture generation and interpretation*. Computer vision in human-machine interaction, 1998: p. 191-215.

[23] Freeman, W.T. and C.D. Weissman. *Television control by hand gestures*. in *Proc. of Intl. Workshop on Automatic Face and Gesture Recognition*. 1995.

[24] Hong, P., M. Turk, and T.S. Huang. *Gesture modeling and recognition using finite state machines*. in *Automatic face and gesture recognition, 2000. proceedings. fourth ieee international conference on*. 2000. IEEE.

[25] Moeslund, T.B., M. Störring, and E. Granum. *A natural interface to a virtual environment through computer vision-estimated pointing gestures*. in *International Gesture Workshop*. 2001. Springer.

[26] Streitz, N.A., et al., *Roomware: Towards the next generation of human-computer interaction based on an integrated design of real and virtual worlds*. Human-Computer Interaction in the New Millenium, Addison Wesley, 2001: p. 551-576.

[27] Koller, D., et al. *Real-time vision-based camera tracking for augmented reality applications*. in *Proceedings of the ACM symposium on Virtual reality software and technology*. 1997. ACM.

[28] Jalal, A., M.Z. Uddin, and T.-S. Kim, *Depth video-based human activity recognition system using translation and scaling invariant features for life logging at smart home*. IEEE Transactions on Consumer Electronics, 2012. **58**(3).

[29] Jalal, A., et al. *Human activity recognition via the features of labeled depth body parts*. in *International Conference on Smart Homes and Health Telematics*. 2012. Springer.

[30] Jalal, A., J.T. Kim, and T.-S. Kim. *Development of a life logging system via depth imaging-based human activity recognition for smart homes*. in *Proceedings of the International Symposium on Sustainable Healthy Buildings, Seoul, Korea*. 2012. Citeseer.

[31] Jalal, A. and S. Kamal. *Real-time life logging via a depth silhouette-based human activity recognition system for smart home services*. in *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on*. 2014. IEEE.

[32] Kamal, S., A. Jalal, and D. Kim, *Depth images-based human detection, tracking and activity recognition using spatiotemporal features and modified HMM*. J. Electr. Eng. Technol, 2016. **11**(3): p. 1921-1926.

[33] Jalal, A., Y. Kim, and D. Kim. *Ridge body parts features for human pose estimation and recognition from RGB-D video data*. in *Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on*. 2014. IEEE.

[34] Turk, M.A. and A.P. Pentland. *Face recognition using eigenfaces*. in *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*. 1991. IEEE.

[35] Jalal, A. and S. Kim, *Global security using human face understanding under vision ubiquitous architecture system.* World academy of science, engineering, and technology, 2006. **13**: p. 7-11.

[36] Parmar, D.N. and B.B. Mehta, *Face recognition methods & applications.* arXiv preprint arXiv:1403.0485, 2014.

[37] Lowe, D.G. *Object recognition from local scale-invariant features.* in *Computer vision, 1999. The proceedings of the seventh IEEE international conference on.* 1999. Ieee.

[38] Li, L.-J. and L. Fei-Fei. *What, where and who? classifying events by scene and object recognition.* in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on.* 2007. IEEE.

[39] Yao, B. and L. Fei-Fei. *Grouplet: A structured image representation for recognizing human and object interactions.* in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.* 2010. IEEE.

[40] Delaitre, V., I. Laptev, and J. Sivic. *Recognizing human actions in still images: a study of bag-of-features and part-based representations.* in *BMVC 2010-21st British Machine Vision Conference.* 2010.

[41] Shapovalova, N., et al. *On importance of interactions and context in human action recognition.* in *Iberian conference on pattern recognition and image analysis.* 2011. Springer.

[42] Yao, B., A. Khosla, and L. Fei-Fei. *Combining randomization and discrimination for fine-grained image categorization.* in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on.* 2011. IEEE.

[43] Koniusz, P. and K. Mikolajczyk. *Soft assignment of visual words as linear coordinate coding and optimisation of its reconstruction error.* in *Image Processing (ICIP), 2011 18th IEEE International Conference on.* 2011. IEEE.

[44] Koniusz, P. and K. Mikolajczyk. *Spatial coordinate coding to reduce histogram representations, dominant angle and colour pyramid match.* in *Image Processing (ICIP), 2011 18th IEEE International Conference on.* 2011. IEEE.

[45] Li, P., J. Ma, and S. Gao. *Actions in still web images: Visualization, detection and retrieval.* in *International Conference on Web-Age Information Management.* 2011. Springer.

[46] Yao, B., et al. *Human action recognition by learning bases of action attributes and parts.* in *Computer Vision (ICCV), 2011 IEEE International Conference on.* 2011. IEEE.

[47] Delaitre, V., J. Sivic, and I. Laptev. *Learning person-object interactions for action recognition in still images.* in *Advances in neural information processing systems.* 2011.

[48] Li, P. and J. Ma. *What is happening in a still picture?* in *Pattern Recognition (ACPR), 2011 First Asian Conference on.* 2011. IEEE.

[49] Thurau, C. and V. Hlaváč. *Pose primitive based human action recognition in videos or still images.* in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on.* 2008. IEEE.

[50] Dalal, N. and B. Triggs. *Histograms of oriented gradients for human detection.* in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on.* 2005. IEEE.

[51] Gupta, A., A. Kembhavi, and L.S. Davis, *Observing human-object interactions: Using spatial and functional compatibility for recognition.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009. **31**(10): p. 1775-1789.

[52] Desai, C. and D. Ramanan. *Detecting actions, poses, and objects with relational phraselets.* in *European Conference on Computer Vision.* 2012. Springer.

[53] Raja, K., et al. *Joint pose estimation and action recognition in image graphs.* in *Image Processing (ICIP), 2011 18th IEEE International Conference on.* 2011. IEEE.

[54] Yao, B. and L. Fei-Fei, *Recognizing human-object interactions in still images by modeling the mutual context of objects and human poses.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012. **34**(9): p. 1691-1703.

[55] Sener, F., C. Bas, and N. Ikizler-Cinbis. *On recognizing actions in still images via multiple features.* in *European Conference on Computer Vision.* 2012. Springer.

[56] Belongie, S., G. Mori, and J. Malik, *Matching with shape contexts,* in *Statistics and Analysis of Shapes.* 2006, Springer. p. 81-105.

[57] Wang, Y., et al. *Unsupervised discovery of action classes.* in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on.* 2006. IEEE.

[58] Yao, B. and L. Fei-Fei. *Modeling mutual context of object and human pose in human-object interaction activities.* in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on.* 2010. IEEE.

[59] Oliva, A. and A. Torralba, *Modeling the shape of the scene: A holistic representation of the spatial envelope.* International journal of computer vision, 2001. **42**(3): p. 145-175.

[60] Prest, A., C. Schmid, and V. Ferrari, *Weakly supervised learning of interactions between humans and objects.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012. **34**(3): p. 601-614.

[61] Bay, H., et al., *Speeded-up robust features (SURF).* Computer vision and image understanding, 2008. **110**(3): p. 346-359.

[62] Ikizler, N., et al. *Recognizing actions from still images.* in *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on.* 2008. IEEE.

[63] Ahmed, M., N. Al-Jawad, and A.T. Sabir. *Gait recognition based on Kinect sensor.* in *Real-Time Image and Video Processing 2014.* 2014. International Society for Optics and Photonics.

[64] Andersson, V.O. and R.M. de Araújo. *Person Identification Using Anthropometric and Gait Data from Kinect Sensor.* in *AAAI.* 2015.

[65] Ball, A., et al. *Unsupervised clustering of people from 'skeleton' data.* in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction.* 2012. ACM.

[66] Dikovski, B., G. Madjarov, and D. Gjorgjevikj. *Evaluation of different feature sets for gait recognition using skeletal data from Kinect.* in *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on.* 2014. IEEE.

[67] Kwolek, B., et al. *3D gait recognition using spatio-temporal motion descriptors.* in *Asian Conference on Intelligent Information and Database Systems.* 2014. Springer.

[68] Preis, J., et al. *Gait recognition with kinect.* in *1st international workshop on kinect in pervasive computing.* 2012. New Castle, UK.

[69] Farooq, A., A. Jalal, and S. Kamal, *Dense RGB-D Map-Based Human Tracking and Activity Recognition using Skin Joints Features and Self-Organizing Map.* KSII Transactions on Internet & Information Systems, 2015. **9**(5).

[70] Yoshimoto, H., N. Date, and S. Yonemoto. *Vision-based real-time motion capture system using multiple cameras.* in *Multisensor Fusion and Integration for Intelligent Systems, MFI2003. Proceedings of IEEE International Conference on.* 2003. IEEE.

[71] Zhang, X., et al., *A framework for hand gesture recognition based on accelerometer and EMG sensors.* IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2011. **41**(6): p. 1064-1076.

[72] Kammer, D., et al. *Towards a formalization of multi-touch gestures.* in *ACM International Conference on Interactive Tabletops and Surfaces.* 2010. ACM.

[73] Hoggan, E., et al. *Multi-touch rotation gestures: Performance and ergonomics.* in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 2013. ACM.

[74] Zabulis, X., H. Baltzakis, and A.A. Argyros, *Vision-Based Hand Gesture Recognition for Human-Computer Interaction.* The universal access handbook, 2009. **34**: p. 30.

[75] Buehler, P., et al., *Upper body detection and tracking in extended signing sequences.* International journal of computer vision, 2011. **95**(2): p. 180.

[76] Rautaray, S.S. and A. Agrawal, *Vision based hand gesture recognition for human computer interaction: a survey.* Artificial Intelligence Review, 2015. **43**(1): p. 1-54.

[77] Ren, Z., J. Yuan, and Z. Zhang. *Robust hand gesture recognition based on finger-earth mover's distance with a commodity depth camera.* in *Proceedings of the 19th ACM international conference on Multimedia.* 2011. ACM.

[78] Liou, W.-G., C.-Y. Hsieh, and W.-Y. Lin. *Trajectory-based sign language recognition using Discriminant Analysis in higher-dimensional feature space.* in *Multimedia and Expo (ICME), 2011 IEEE International Conference on.* 2011. IEEE.

[79] Tran, C. and M.M. Trivedi, *3-D posture and gesture recognition for interactivity in smart spaces.* IEEE Transactions on Industrial Informatics, 2012. **8**(1): p. 178-187.

[80] Alon, J., et al., *A unified framework for gesture recognition and spatiotemporal gesture segmentation.* IEEE transactions on pattern

analysis and machine intelligence, 2009. **31**(9): p. 1685-1699.

[81]	Sarkar, S., et al. *Segmentation-robust representations, matching, and modeling for sign language*. in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*. 2011. IEEE.

[82]	Lee, B. and J. Chun. *Interactive manipulation of augmented objects in marker-less ar using vision-based hand interaction*. in *Information Technology: New Generations (ITNG), 2010 Seventh International Conference on*. 2010. IEEE.

[83]	Störring, M., et al. *Computer vision-based gesture recognition for an augmented reality interface*. in *4th IASTED international conference on visualization, imaging, and image processing*. 2004.

[84]	Yang, M.-T., W.-C. Liao, and Y.-C. Shih. *VECAR: Virtual English classroom with markerless augmented reality and intuitive gesture interaction*. in *Advanced Learning Technologies (ICALT), 2013 IEEE 13th International Conference on*. 2013. IEEE.

[85]	Lee, M., R. Green, and M. Billinghurst. *3D natural hand interaction for AR applications*. in *Image and Vision Computing New Zealand, 2008. IVCNZ 2008. 23rd International Conference*. 2008. IEEE.

[86]	Hassanpour, R., S. Wong, and A. Shahbahrami. *Visionbased hand gesture recognition for human computer interaction: A review*. in *IADIS international conference interfaces and human computer interaction*. 2008. Published.

[87]	Leap, C., *API--https://developer. leapmotion. com/documentation/cpp/api.* Leap_Classes. html.

[88]	Guna, J., et al., *An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking.* Sensors, 2014. **14**(2): p. 3702-3720.

[89]	Kean, S., J.C. Hall, and P. Perry, *Microsoft's kinect sdk.* Meet the Kinect, 2011: p. 151-173.

[90]	Buchmann, V., et al. *FingARtips: gesture based direct manipulation in Augmented Reality*. in *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia*. 2004. ACM.

[91]	Park, H., H.-K. Jung, and S.-J. Park, *Tangible AR interaction based on fingertip touch using small-sized nonsquare markers.* Journal of Computational Design and Engineering, 2014. **1**(4): p. 289-297.

[92]	Baldauf, M., et al. *Markerless visual fingertip detection for natural mobile device interaction*. in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*. 2011. ACM.

[93]	Lee, T. and T. Hollerer. *Handy AR: Markerless inspection of augmented reality objects using fingertip tracking*. in *Wearable Computers, 2007 11th IEEE International Symposium on*. 2007. IEEE.

[94]	Wachs, J.P., et al., *Vision-based hand-gesture applications.* Communications of the ACM, 2011. **54**(2): p. 60-71.

[95]	Chun, W.H. and T. Höllerer. *Real-time hand interaction for augmented reality on mobile phones*. in *Proceedings of the 2013 international conference on Intelligent user interfaces*. 2013. ACM.

[96]	Lee, J.Y., G.W. Rhee, and D.W. Seo, *Hand gesture-based tangible interactions for manipulating virtual objects in a mixed reality environment.* The International Journal of Advanced Manufacturing Technology, 2010. **51**(9-12): p. 1069-1082.

[97]	Lu, G., et al., *Immersive manipulation of virtual objects through glove-based hand gesture interaction.* Virtual Reality, 2012. **16**(3): p. 243-252.

[98]	Vatavu, R.-D., Ş.-G. Pentiuc, and C. Chaillou, *On natural gestures for interacting in virtual environments.* Advances in electrical and computer engineering, 2005. **24**(5): p. 72-79.

[99]	Bowman, D.A., D. Koller, and L.F. Hodges, *A methodology for the evaluation of travel techniques for immersive virtual environments.* Virtual reality, 1998. **3**(2): p. 120-131.

[100]	De Boeck, J., C. Raymaekers, and K. Coninx. *Are existing metaphors in virtual environments suitable for haptic interaction*. 2005. sl 2005.

[101]	Mitra, S. and T. Acharya, *Gesture recognition: A survey.* IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2007. **37**(3): p. 311-324.

[102]	Motion, L., *Leap motion developer portal*. 2014.

[103]	Dardas, N.H. and M. Alhaj, *Hand gesture interaction with a 3D virtual environment.* The International Journal of ACM JORDAN, 2011. **2**(3): p. 186-194.

[104]	Yang, J., Y. Xu, and C. Chen. *Gesture interface: Modeling and learning*. in *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*. 1994. IEEE.

[105]	Cooper, H., B. Holt, and R. Bowden, *Sign language recognition*, in *Visual Analysis of Humans*. 2011, Springer. p. 539-562.

[106]	Kim, J.-S., et al. *Study on intelligent autonomous navigation of Avatar using hand gesture recognition*. in *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*. 2000. IEEE.

[107]	Chen, Q., et al. *Accessing Learning Objects in Virtual Environment by Hand Gestures and Voice*. in *Proc. 3rd Annual Scientific Conference of LORNET Research Network (I2LOR-06), TELUQ-UQAM University in Montreal, Canada*. 2006.

[108]	Krum, D.M., et al. *Speech and gesture multimodal control of a whole earth 3d virtual environment*. in *Eurographics-IEEE Visualization Symposium 2002*. 2002. Citeseer.

[109]	Shao, L., *Hand movement and gesture recognition using Leap Motion Controller.*

[110]	Kerefeyn, S. and S. Maleshkov, *Manipulation of virtual objects through a LeapMotion optical sensor.* International Journal of Computer Science Issues (IJCSI), 2015. **12**(5): p. 52.

[111]	Khundam, C. *First person movement control with palm normal and hand gesture interaction in virtual reality*. in *Computer Science and Software Engineering (JCSSE), 2015 12th International Joint Conference on*. 2015. IEEE.

[112]	de Paiva Batista, R.M., *Navigating Virtual Reality Worlds with the Leap Motion Controller.* 2016.

[113]	Liang, H., et al., *Hand gesture-based interactive puppetry system to assist storytelling for children.* The Visual Computer, 2016: p. 1-15.

[114]	Nabiyouni, M., B. Laha, and D.A. Bowman. *Poster: Designing effective travel techniques with bare-hand interaction*. in *3D User Interfaces (3DUI), 2014 IEEE Symposium on*. 2014. IEEE.

[115]	Prazina, I. and S. Rizvić. *Natural interaction with small 3D objects in virtual environments*. in *Submitted to 20th Central European Seminar on Computer Graphics*. 2016.

[116]	Kammergruber, F., A. Ebner, and W.A. Günthner. *Navigation in virtual reality using Microsoft Kinect*. in *12th international conference on construction application of virtual reality, Taipei, Taiwan*. 2012.

[117]	Kim, J.-O., M. Kim, and K.-H. Yoo, *Real-time hand gesture-based interaction with objects in 3D virtual environments.* Int. J. Multimed. Ubiquitous Eng, 2013. **8**(6): p. 339-348.

[118]	Vultur, O.M., Ş.G. Pentiuc, and A. Ciupu. *Navigation system in a virtual environment by gestures*. in *Communications (COMM), 2012 9th International Conference on*. 2012. IEEE.

[119]	Phung, S.L., A. Bouzerdoum, and D. Chai. *A novel skin color model in ycbcr color space and its application to human face detection*. in *Image Processing. 2002. Proceedings. 2002 International Conference on*. 2002. IEEE.

[120]	Raees, M., et al., *Image based recognition of Pakistan sign language.* Journal of Engineering Research, 2016. **4**(1).

[121]	Piyathilaka, L. and S. Kodagoda. *Gaussian mixture based HMM for human daily activity recognition using 3D skeleton features*. in *Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on*. 2013. IEEE.

[122]	Jalal, A., S. Kamal, and D. Kim, *A depth video sensor-based life-logging human activity recognition system for elderly care in smart indoor environments.* Sensors, 2014. **14**(7): p. 11735-11759.

[123]	Jalal, A., et al., *Robust human activity recognition from depth video using spatiotemporal multi-fused features.* Pattern recognition, 2017. **61**: p. 295-308.

[124]	Jalal, A., S. Kamal, and D. Kim. *Shape and motion features approach for activity tracking and recognition from kinect video camera*. in *Advanced Information Networking and Applications Workshops (WAINA), 2015 IEEE 29th International Conference on*. 2015. IEEE.

[125]	Jalal, A., S. Kamal, and D. Kim. *Individual detection-tracking-recognition using depth activity images*. in *Ubiquitous Robots and Ambient Intelligence (URAI), 2015 12th International Conference on*. 2015. IEEE.

[126]	Wu, H., et al. *Human activity recognition based on the combined SVM&HMM*. in *Information and Automation (ICIA), 2014 IEEE International Conference on*. 2014. IEEE.

[127]	Jalal, A. and Y. Kim. *Dense depth maps-based human pose tracking*

and recognition in dynamic scenes using ridge data. in *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on*. 2014. IEEE.

[128] Pears, Alan (June 1998). "Chapter 7: Appliance technologies and scope for emission reduction". Strategic Study of Household Energy and Greenhouse Issues (PDF). Sustainable Solutions Pty Ltd. Department of Industry and Science, Commonwealth of Australia. p. 61. http://www.energyrating.gov.au/document/report-strategic-study-household-energy-and-greenhouse-issues. Retrieved on 26-06-2017.

[129] *Australian Greenhouse Office (May 2005). «Chapter 5: Assessing lighting savings».* Working Energy Resource and Training Kit: Lighting. https://web.archive.org/web/20070415151053/http://www.greenhouse.gov.au/lgmodules/wep/lights/training/training9.html. Retrieved on 26-06-2017.

[130] Australian Greenhouse Office. (May 2005). 'Chapter 5: Assessing lighting savings'. Working Energy Resource and Training Kit: Lighting. Retrieved 17-03-2017.

[131] "Low-Light Performance Calculator". http://www.scopecalc.com/ Retrieved on 03-06-2017.

[132] How to use a lux meter (Australian recommendation)' (PDF), April 2010. Sustainability Victoria. http://www.sustainability.vic.gov.au/ Retrieved on 09-05-2017.

[133] "Illumination. - 1926.56". Regulations (Standards - 29 CFR). Occupational Safety and Health Administration, US Dept. of Labor. www.osha.gov/pls/oshaweb/owadisp.show_document?p_table=STANDARDS&p_id=10630, Retrieved on 09-05-2017.

Mr. Rehman

He is currently pursuing Ph. D. degree in 3D Interaction from the University of Malakand, Pakistan. He received his M. Phil in Virtual and Augmented Reality and BS degree in Computer Science from University of Malakand in 2006 and 2016 respectively. His research interest includes 3D interaction, virtual and augmented reality, and machine learning.

Dr. Ullah

He is currently working as Associate Professor in the department of Computer Science & IT, University of Malakand, Pakistan. He received his PhD degree in Computer Science in specialty in Robotics from University of Evry Val d'Essonne, France in 2011. He also earned MS degree of Computer Science with specialty in Virtual Reality and Intelligent Systems from the same University in 2007. He received MSc degree of Computer Science and BSc with major in Mathematics and Computer Science from University of Peshawar in 2001 and 1998 respectively. Dr. Sehat Ullah has research interest in virtual and augmented reality systems, machine learning, image processing and haptic systems.

Mr. Raees

He is serving as assistant professor of computer science in the department of Higher Education, Khyber pakhtunkhwa, Pakistan. He received his M.Phil degree from University Of Malakand in 2015 and is currently pursuing Ph.D. in flexible 3D interaction from the same university. His research interest includes 3D interactions, computer vision and machine learning.