MAT JOURNALS
Journal of Information Technology and Sciences
e-ISSN: 2581-849X
Volume 5 Issue 2

Kinect based Telerehabilitation for Virtual Therapy

Ramkrishna Vadali[1], Aditya Bodhankar[2], Priyanka Jagadale[2], Snehal Pethkar[2]
[1]Professor, [2]UG Students
[1,2]Department of IT, Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India
Email: *priyajagadalepj@gmail.com*
DOI:

Abstract
Virtual reality technology is currently widely applied in therapeutic rehabilitation therapy. The ability to trace joint positions for Microsoft Kinect might be helpful for rehabilitation, both in clinical setting and at home. Currently, most systems developed for virtual rehabilitation and motor coaching need quite advanced and high-priced hardware and may be used only in clinical settings. Now, a cheap rehabilitation game coaching system has been developed for patients with movement disorders, it is suitable for home use under the distant supervision of a therapist. This research explores the potential and therefore the limitations of the Kinect in the application of e-rehabilitation.We evaluated the tools that could be accustomed help promote physical rehabilitation reception by reducing the frequency of hospital visits, leading to the reduction of care value. A system should be developed such as it is useful in Telerehabilitation. As we see from the result, it indicated a slightly positive outcome for the patients after got involved in the treatment.

Keywords: *Kinect, Microsoft, telerehabilitation, therapy, segmentation*

INTRODUCTION

Virtual reality (VR) offers a possible solution, which allows the user to directly interact with a computer-simulated environment. It can lead to new and exciting discoveries in these areas which impact upon our day to day lives. Modern input devices have been massively influenced by VR and may become the corner stone of further virtual reality developments[2]. Nintendo Wii and Microsoft Xbox Kinect are good examples. The former uses a controller, which can be latched to the hand, to make movement becomes a form of input, while the later uses a camera to track a user movements, which are then reflected in exercise[3]. Our society is aging rapidly. Currently, a little less than 8% of the world's population is 65 or older and this percentage is expected to reach 16% by 2050. The growth in the elderly population is more accentuated in developed countries where life expectancy continues to rise. As a result, the number of patients with motor function disorders can drastically increase while the ability to care for them will be limited by public expenditure and human resources. Thus, there is high demand for computer-aided tools which support in-home rehabilitation.

LITERATURE SURVEY

1. Paper name: Communication interface for deaf-mute people using Microsoft Kinect.
Author: 2016 International conference on automatic control and dynamic optimization techniques.[11] (icacdot)
Nowadays, there are many people who are suffer from hearing disabilities. Many people are having disability in hearing. Therefore, the interaction or communication with these people is major difficulty for normal people. Therefore this research developed the communication interface between deaf-mute people and normal people who are able to hear. For this purpose MicrosoftKinect is used.
2. Paper name: Hand detecting and positioning based on depth image of kinect sensor.

Author: International journal of information and electronics engineering, vol. 4, no. 3, May 2014; [9]this research developed the technique for hand detection and positioning and identifying the gestures. For this technique to be implemented python programming environment is used. The kinect sdk is used for the development, which tracks the human body i.e gesture identification. Then depth image is extracted from the gestures.

3.Paper name: A Kinect sensor based windows control interface.
Author: International journal of control and automation[10]many new concepts have recently been developed for human–computer interfaces. In this research an intelligent interface is developed for windows operating system. For this purpose Microsoft kinect is used. Various kinect functionalities are used for this. Kinect RGB camera is used for gesture tracking. Kinect also have an audio interface.

4.Paper name: robust part-based hand gesture recognition using kinect sensor.
Author: Zhou ren, junsong Yuan, Jingiingmeng EEE transactions on multimedia, vol. 15, no. 5, august 2013 [12]

Kinect provide many functionalities for the applications developed for the human computer interaction. There are many languages which are supported by kinect for the development. Initially, when kinect is developed, it is used for gaming purpose only, but after adding functionalities it is used in many healthcare applications.

EXISTING METHODS

In vision-based Shining Path recognition, the key factor is the accurate and fast hand tracking and segmentation. It uses image processing algorithms to detect and track hand signs moreover as facial expressions of the signer, which is less complicated to the signer without sporting gloves[5]. However, it is very difficult for the complex backgrounds and also there is accuracy problem related to image processing algorithms[9].

PROPOSED METHODS

Kinect sensor is used on large scales, depth camera. It is device that can view in 3Dimensions. Microsoft made the device principally as a game controller and is operated by the company’s own software package for analyzing its 3D data, including proprietary computation for, scene analysis, feature trailing, motion tracking, gesture recognition and skeletal trailing[4]. The effectiveness of work is done in matlab and visual studio and verified exploitation Kinect. Kinect using Matlab may be a smart answer for converting sign language to speech communication and this is the idea of this paper which is delineated here in details.

RELATED WORK

OVERVIEW OF MODEL:

Definitions and data formats

We will begin by introducing some terms to be used in the rest of this report. A “frame” is a single time slice. Each frame contains the locations and joint angles of all 20 joints. Each joint is represented as a 3-dimensional point; in addition, each joint has a parent joint, which is intuitively defined. A “gesture” or “input gesture” is a non-empty sequence of frames. We call the output of the recognition engine a “result”.

Recognizing a gesture

The process of recognizing a gesture involves several steps. First, we train the various models used in the program with an appropriate set of training data[1]. There are two trained models: one, the “neutral stance” model, recognizes whether single frames represent a “neutral stance”. The second, the “recognizer”, takes input gestures (sequences of frames as defined above) and produces a result.

During real-time operation, the engine receives 30 raw frames per second from the motion tracker. Raw frames are pre-processed by storing the absolute position of the neck joint, and then recomputing the positions of the other joints relative to the neck joint (joint angles are also computed at this step) as shown in figure 1:(a) and figure2:(b). This is necessary, as the Kinect is able to track backward and forward motion, and our method needs to be translationindependent. The processed frames are sent to the "segmenter", which uses the "neutral stance" model to separate the input stream into logical segments. Some of these segments are then sent to the "recognizer" as input gestures[3].

The "Neutral Stance" Model

The neutral stance model represents a stance where the user is idle and not in the middle of performing a gesture. The neutral stance model is fairly simplistic: we model the mean and variance of the normalized positions of all the joints from a customized set of parents. The selections were made in this model to fit closely with expectations about the human body (all the right arm joints, for example, were parented to the right shoulder joint), as well as to minimize model error from differing body shapes, height, etc[6].

DEVELOPMENT AND IMPLEMENTATION:

Segmentation Algorithm

The segmentation algorithm uses a more fixed heuristic, and learns only the neutral stance feature.

By using more "distant" parents, we cut down significantly on sampling noise (for example, in the hand joints); and by choosing parents carefully, we can maintain sufficient body-shape independence. Currently, our model parents all the joints on the limbs to their respective "connecting" joints on the main body elbow, right wrist, and right palm are parented to the right shoulder). This approach segments fairly effectively - we were able to achieve 60-70% accuracy (subjectively) on our testing sequences.Initially an image is selected and then converted intoGray level image of size 256x256. Process:

- The original and resized image is displayed.
- The value of the concerned parameters, if any, isselected.
- The result of each algorithm is displayed in theseparate figure window.
- The segmented image is displayed at a particularposition according to the range of the selected valueof the parameter.
- Every output image in that figure window containsthe value of parameters and the duration of process.

Initially an image is selected and then converted into
Gray level image of size 256x256.

- The original and resized image is displayed.
- The value of the concerned parameters, if any, isselected.
- The result of each algorithm is displayed in theseparate figure window.
- The segmented image is displayed at a particular
- Position according to the range of the selected valueof the parameter.
- Every output image in that figure window containsthe value of parameters and the duration of process.Initially an image is selected and then converted intoGray level image of size 256x256.
- The original and resized image is displayed.
- The value of the concerned parameters, if any, isselected.
- The result of each algorithm is displayed in theseparate figure window.
- The segmented image is displayed at a particularposition according to the range of the selected valueof the parameter.

- Every output image in that figure window containsthe value of parameters and the duration of process.

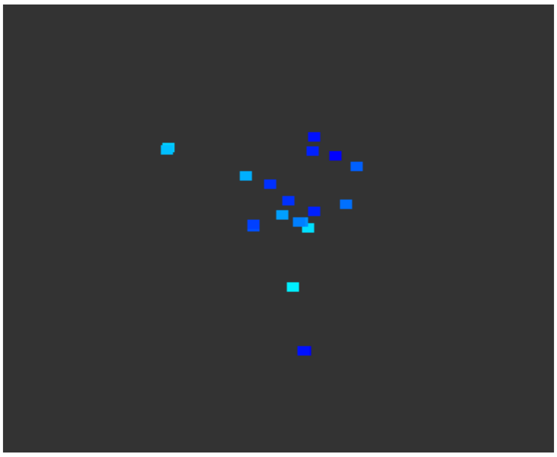

Figure 1: *(a) High kick gesture.*

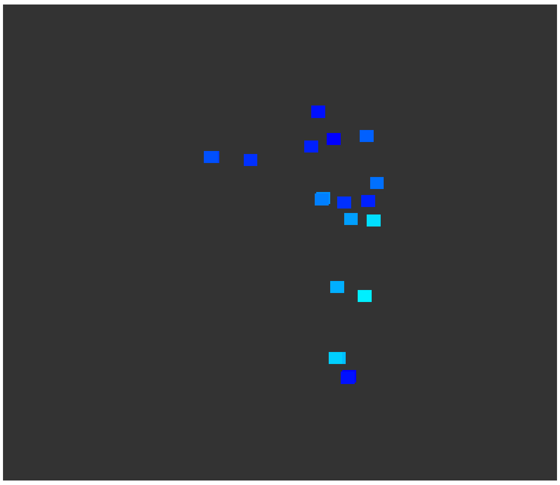

Figure 2: *(b) Punch gesture.*

SYSTEM ARCHITECTURE

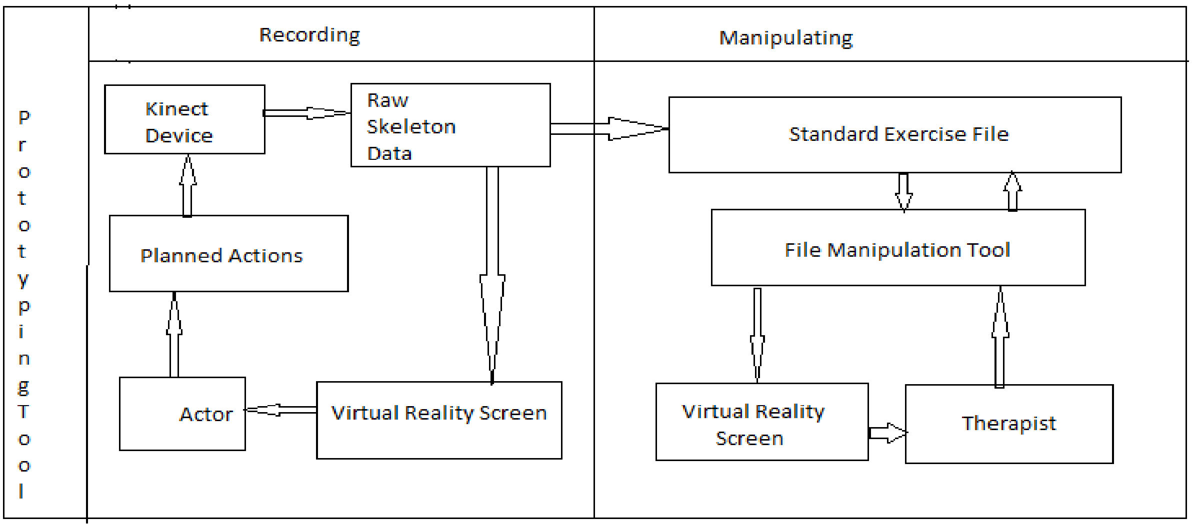

Figure 3: *System architecture.*

The overall system design consists of following modules (as shown in figure 3s):

- Kinect Device
- Raw skeleton data
- Planned Actions
- Actor
- Virtual reality screen
- Standard exercise file
- File manipulation tool
- Therapist

In this work,let's imagine that the human face and upper body are completely visible and do not have any occlusions. The sensor is stable and the motion of the subject along the camera axis is constrained to a range.

The three functions in diagram are respectively corresponding to three modules: exercises prototyping tool, action evaluation module, data statistics and analysis module. 1) Prototyping tool: In this module the therapist only records the patient's skeleton data. An actor does the planned actions and Kinect device captures the poses and then generates the raw skeleton data. Then the raw skeleton data is processed and converted to standard file system. 2) Action evaluation module: this module captures the actions performed by the patient and if the patient performed action correctly, the therapist gives score to the patient according to the perfection of the exercise. At the same time, the Kinect device captures the poses processing procedure, the scores are given based on the comparison and generates the patient's skeleton data. 3) Data statistics and analysis module: This module gives the response to the patient about the correctness of exercise he/she is done. The patient's skeleton data is recorded for taking the feedback.

RESULTS

VISUALIZATION OF GESTURES:

We created visualizations of some gestures to help the reader visualize the input data. These visualizations also helped us during the development process. One graph of the joint components over a particular gesture is also included.

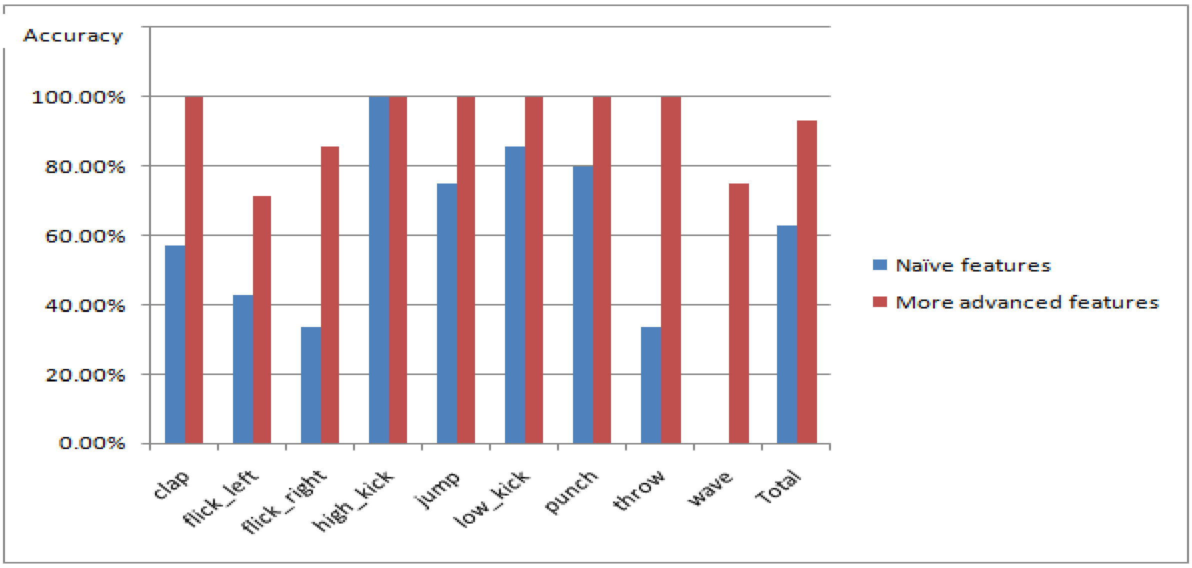

Figure 4: *Performance comparison.*

Single-gesture recognition

Using the L2-normalized logistic regression model, we achieved extremely good results in single-gesture recognition tests. Typically, cross-validation tests reported final feature set near or more than 90% accuracy. The output of a sample testing run can be seen in Table 1. We can immediately see the potential for this kind of model. We also include a plot which compares the results we obtain using our original set of naïve features and our final results using more holistic features.

Real-time recognition

As mentioned above, we were able to achieve over 70%(subjective) accuracy in real-time recognition. (The resultswere measured by a human watching the animationand comparing his observations to the recognizedgestures.) Some gestures are missed and some singlegestures are "split", but for the most part, gestures canbe performed in real-time with fairly high recognitionrates.

***Table 1:** Sample run.*

Gesture Name	Correct	Total	%Correct
Clap	8	8	100%
Flick_left	5	7	71.4%
Flick_right	6	7	85.7%
High_kick	10	10	100%
jump	5	5	100%
Low_kick	8	8	100%
Punch	6	6	100%
Throw	3	3	100%
wave	3	4	75%
ToTal	**54**	**58**	**93.1%**

CONCLUSION

System will help for easier interaction and communication with impaired people. It acts as mediator between impaired user and ordinary user. They can easily convey the messages to each other by this system. While communicating, user can also add as many signs into the dictionary along with its corresponding meaning. The experimental results show that the system is working system for native Indian sign language recognition. This proposed system can be enhanced to recognize for continuous sentences.

REFERENCES

1. Ren, Zhou, JingjingMeng, Junsong Yuan (2011), "Depth camera based hand gesture recognition and its applications in human-computerinteraction",*IEEE InInformation, Communications and Signal Processing (ICICS) 8th International Conference on*, pp. 1−5.
2. W. Gao, G. Fang, D. Zhao, Y. Chen (2004), "A Chinese sign language recognition system based on SOFM/SRN/HMM", *Pattern Recognition*,Volume 37, Issue 12, pp. 2389–2402.
3. Hernandez-Rebollar, J. L., Kyriakopoulos, N., Lindeman, R. W. (May 2004,),"A new instrumented approach for translating American Sign Language into sound and text. In Automatic Face and Gesture Recognition", *2004. Proceedings. Sixth IEEE International Conference on*, pp. 547−552.
4. Huang, Xuedong D., YasuoAriki, Mervyn A. Jack (1990),"Hidden Markov models for speech recognition",*Edinburgh: Edinburgh university press*, Volume 2004,
5. Elina, Anton-Canalis, Luis, Elena Sanchez-Nielsen (2006),"Hand posture dataset creation for gesture recognition", *International Conference on Computer Vision Theory and Applications (VISAPP'06*
6. Bukhari, J., et al. (2015),"American sign language translation through sensory glove; signspeak",*Int. J. u-and e-Service, Science and Technology*,Volume 8,

7. D. Capilla, “Sign Language Translator using Microsoft KinectXBOX360”, *MSc Thesis*, Department of Electrical Engineering and ComputeScience Computer Vision Lab University of Tennessee, Knoxville,USA 2012.R. Nicole,
8. S. Lang, M. Block, R. Rojas (2012), “Sign Language Recognition Using Kinect”, *In Artificial Intelligence and Soft Computing*, (Eds) L.Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L. Zadeh, JZurada. Lecture Notes in Computer Science, Springer Berlin Heidelbergvol. 7267, pp. 394−402.
9. Chen, Z. H., Kim, J. T., Liang, J., Zhang, J.,Yuan, Y. B. (2014),“Real-Time Hand Gesture Recognition Using Finger Segmentation”,*The Scientific World Journal*,
10. Lee, Sang-Hyuk& Oh, Seung-Hyun. (2014). A Kinect Sensor based Windows Control Interface. International Journal of Control and Automation. 7. 113-124. 10.14257/ijca.2014.7.3.12.
11. Darryl O’Hare, William Hurst, David Tully, Abdennour El Rhalibi, *E-Learning and Games*, vol. 10345, pp. 182, 2017.
12. EEE transactions on multimedia, robust part-based hand gesture recognition using kinect vol. 15, no. 5, august 2013

Cite this article as: