BLENDED INTERACTION FOR AUGMENTED LEARNING-AN ASSISTIVE TOOL FOR COGNITIVE DISABILITY

Suman Deb[1], Subir Saha[1] and Paritosh Bhattacharya[2]
[1]*Computer Science and Engineering Department*
[2]*Mathematics Department*
NIT Agartala, India

ABSTRACT

The fundamental right of education and employment for cognitive disable person is recognized in very recent times. With the advent of technology there have been many interactive way to combat with the challenges of disability and provide supplement training for enabling the challenged person with the skill of employability, so that they can individually lead the normal life without depending on others. In this work we try to identify the problem faced by cognitively week people and find appropriate solution by introducing computer aided interaction. We proposed a platform to blend audio and depth vision for augmented learning. The primary goal of this work is to create a play way teaching environment for individuals with learning disability so that the enhanced method help them for elementary education and provide self-confidence and motivation and encourage for continuous learning. While implementing, the system has shown significant improvement on learning by a range of people including learning challenged and normal kinder garden kids. The augmented method proposed here can be added advantage with the concept of visual teaching, live experience and reactive and pro-active mechanism rather than reading static content from book. This will also involve the physical movement with precise control that can lead to motivated control of limbs by any neuro motor disability effected person.

KEYWORDS

Depth Vision, Hand Joint Recognition, Voice Recognition, Cognitive Map.

1. INTRODUCTION

A cognitive disable child, besides processing a low I.Q, demonstrates impaired or deficient adaptive behaviour originating from conception and continuing into maturity. They fail to learn what average children learn and find difficult to detect an absurdity in a logical statement. They are also limited with respect to imagination. They remain behind their classmates and become bitter and hostile towards them and others. Repeated failures deprive them of confidence and lack the motivation to learn. In a highly technological age, where talent are very much needed, dissipation of human resources is great problem .We can hardly ignore the gravity of the problem and the care education and training of these children are of great importance and significance. Main deficits that people with cognitive disabilities demonstrate are:

A. Memory: Memory refers to the ability to be able to recall what has been learned over time. Meaningful information is typically moved from sensory memory (stored for seconds) to working memory (stored minutes) and then stored in long term memory. People with cognitive disabilities have difficulty with one of these types of memory [1].

B. Problem-solving and attention : People with cognitive disabilities often have difficulty problem solving. One difficult problem arise, such as learning new material in class, they can typically become frustrated and have difficulty expressing their frustration and have difficulty focusing on the task [1].

If we teach the children who faces some kind of learning disability, in more interactively and visualize the teaching content as natural environment rather static content, they can find interest in learning and develop feelings of power to overcome difficulties and accumulate a sense of self satisfaction. So far many interactive systems have been developed but very few interactive systems were developed to educate them based on what has been taught in the class room. Interactive learning provides keen interest in learning and makes learning fun therefore they can remember things more easily. The main focus of this paper is to develop an interactive learning platform for assisting them to learn more quickly through playing and visualization. Microsoft Kinect sensor is used for making the system. It contains one IR sensor and IR emitter to process depth data. It will return depth data of a point with 16 bit gray scale format with a viewable range of 43 degrees vertical and 57 degrees horizontal [2]. The depth data contain the distance between the device and the object in front of the device. For example, if a pixel coordinate is 200 x 300, the depth data for that pixel point contains distance in millimeters from the Kinect device .figure [1] shows the depth data processing of the Kinect device.

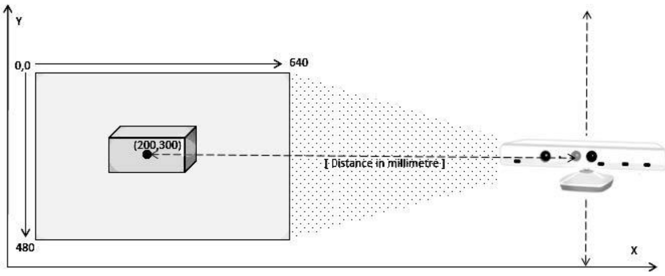


Figure 1. Depth Data Processing [2] of Kinect Sensor

In this proposed prototype Kinect Microphone Array and Kinect speech recognition engine are also used to recognize the speech. Advantages of using Kinect sensor is that it provide us 20 point human skeleton [2] through which gesture can be easily recognized. Moreover the Kinect microphone array have the capability to Identify the source direction of the incoming sound and helps to recognize human speech very clearly by focusing only in a particular direction and cancelling noises in the environment. Figure [2] shows the 20 point skeleton joint of human body return by the kinect sensor.

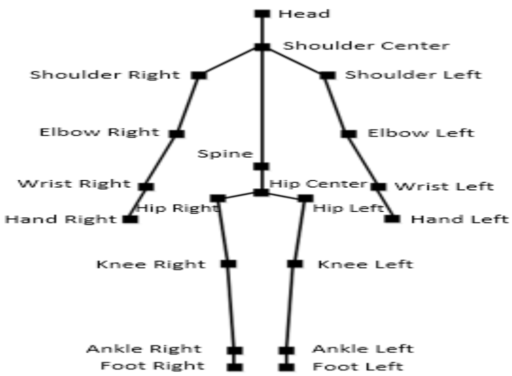


Figure 2. 20 point Skeleton joint [2] of human body

2. EXISTING WORK

Interactive building blocks [5] were developed for the children to learn the concept of geometric structure and shape. First time the system deployed a picture and gives some instruction to build the object. A pattern recognition algorithm is used to compare the children assemble object with the displayed picture .Virtual Laboratory[6] has been developed by using kinect Unity 3D and gesture classification algorithm which is mainly used to assist students to gain interest on particular subject. In paper [7] author developed a useful interactive tool for learning math using blender and Kinect to make learning more visual animated and lively. In paper [8] author developed a e-learning system where a student in remote location effectively interact with the professor by hand gesture. Therefore professor can pay equal attention to both the remote and local student. In paper [9] author proposed an interactive learning platform which combine the full body motion sensing a virtual environment. Therefore student can enter into the virtual environment and have interaction with the object and the virtual character gives some response to this student.

3. SYSTEM ARCHITECTURE

3.1 Hardware and Software System

Kinect was launched on 4 November 2010 by Microsoft and built specifically for the Xbox 360 gaming console. It enables the user to interact directly with the Xbox, allowing the user to perform touch-free operation. The key components of Kinect are (1) Multi-array microphone, (2) IR laser emitter, (3) IR camera, (4) Motorized tilt, (5) USB cable and (6) RGB camera .Below figure [2] shows the key component of the Kinect sensor. Microsoft visual studio 2010, C#, kinect for windows SDK and in back end Microsoft Sql server 2008 is used to develop the system.

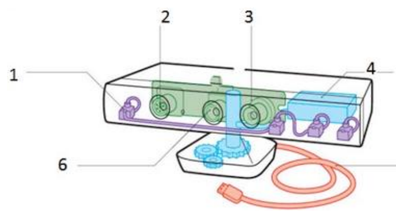


Figure 3. Block diagram of Microsoft Kinect Sensor [3]

3.2 Proposed System

In this work we created a platform where the sensor take the depth information of the object and send it to the process box. This process box performs depth identification ,depth analysis and depth derivation , after analyzing the depth data the computer create their own vocabulary and make a mapping of that vocabulary and store it for future use . Over view of our proposed system are shown in figure [4].

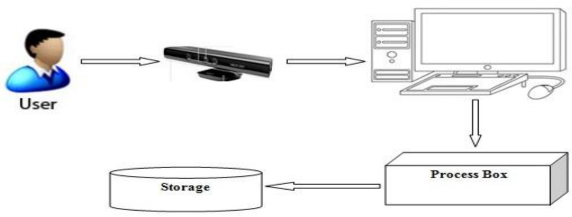


Figure 4. System Architecture of the Proposed System

We divide the system into two phases.

A. Encoding phase

B. Recall phase.

A. Encoding Phase

In this phase when a child comes in front of the Kinect sensor it will first extract the skeleton by processing the depth information and identifies the hand joint. Through this it will recognize the hand portion of the children, after that when the children moves his/her hand in front of the sensor it will map the hand point and draw the object into the computer screen according to the hand movement. After drawing the object he/she assign any pre defined name of this drawing object. The Kinect sensor first recognize the assign name using the Kinect speech recognition engine and mapped the drawing object with given name and store the mapped information into the database. This step was performed repeatedly for acquire accuracy in decision making. Work flow diagram of encoding phase is shown in figure [5].

ISBN: 978-989-8533-43-2 © 2015

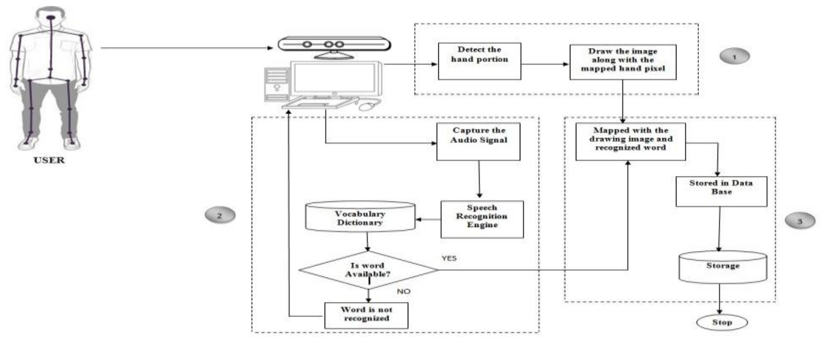


Figure 5. Workflow diagram of encoding phase.

B. Recall Phase

In this phase when a child says any word in front of the sensor, it starts its speech recognition engine for recognising the word. If the word is successfully recognised then the system will try to find the mapping between the word and the image from the database. If mapping is found then the image corresponding to the recognized word is displayed in the screen. Below figure [6] shows the workflow diagram of the recall model.

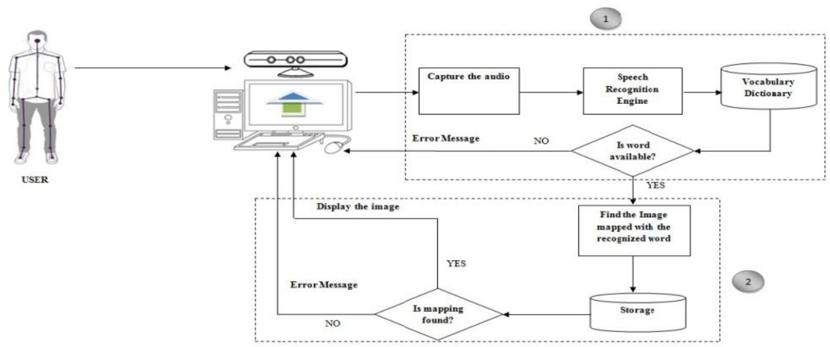


Figure 6. Workflow diagram of recall model

3.3 Algorithm Used

3.3.1 Algorithm for Hand Portion Recognition

Position of both hand for making interactive learning are acquired after skeleton analysis of depth image data return from the Kinect sensor. Following algorithmic step was performed to detecting Hand portion.

***Step 1**: Capture the x,y,z coordinate of all 20 skeleton Joint*

***Step 2**: Analyse the x,y coordinate of all joint*

***Step 3**: Compare the coordinate as*

```
If(x,y coordinates around hand) {
  If( x,y coordinates around Hand_left) {
     Left hand is recognized.
     }
  Else {
         Right Hand is recognized.
     }
     Else {
         Hand is not recognized go to Step 2.
  }
```

***Step 4**: stop*

3.3.2 Algorithm for Speech Recognition

Recognizing the Speech we are used Kinect Speech recognition Engine. The speech recognition engine consists of the following two major modules [2]:

- Acoustic model
- Language model

Each one of the modules has a sole responsibility for recognizing speech. The following is the list of operations performed for recognizing the speech:

Step 1*: Microphones capture the audio stream convert the analog audio data into a digital wave.*

Step 2*: The audio sound signals are sent to the speech recognition engine to recognize the audio.*

Step 3*: The acoustic model of the speech recognition engine analyzes the audio and converts the sound into a number of basic speech elements; we call them phonemes.*

Step 4*: The language model analyzes the content of the speech and tries to match the word by combining the phonemes within an inbuilt digital dictionary as.*

```
If (the word exists in the dictionary) {
        Recognize the word
} Else {
        Word is not recognized.
}
```

Step 5*: Stop.*

4. MACHINE LEARNING USING COGNITIVE MAP

The word cognitive map refers to a process of mental activities which abstract the information from a real world scenario and encoded into the memory for automatic recall [4]. It represents the cause – effect relationships of event in knowledge. Modelling of cognitive map using fuzzy logic is called fuzzy cognitive map. It has been recently been introduced in the field of machine intelligence coined by Barl kosko [4].It is a graph structure capable of encoding knowledge. In this work we used Pal and Konar's Fuzzy cognitive model (FCM). In this approach they represent cognitive map as an associative structure consisting of nodes and directed arcs where the nodes carry the fuzzy belief and the arcs or edge carry the connectivity strength [4].Here we represent each recognized word and the image associated to the word as a node and the edge between the nodes represent the connectivity strength. Below figure [7] shows the cognitive map of the proposed system.

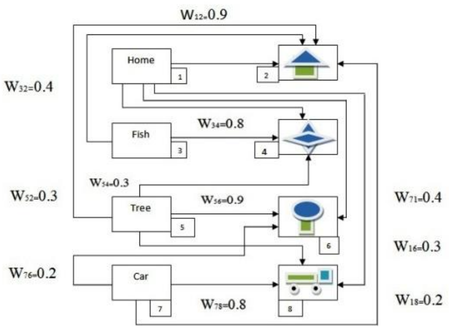


Figure 7. Cognitive map of our proposed system.

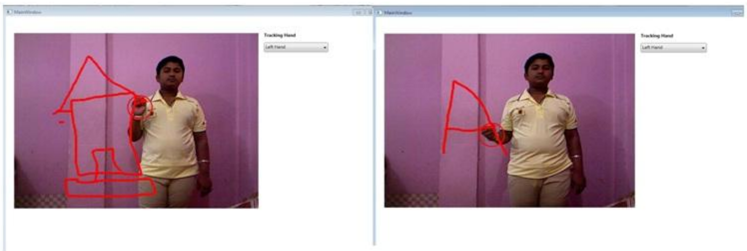
It follows the encoding and recall cycle to build the cognitive map. Process of creating fuzzy cognitive map of our system is as follows.

Let ni, nj be the fuzzy belief of node Ni and Nj.

***Step1**: Represent each recognised word and drawing object as node.*
***Step 2**: follow the Hebbian learning to find self mortality of Wij as:*
$\partial W_{ij}(t) = \alpha\, W_{ij}(t) + S(n_i(t)).S(n_j(t))$. *Where*
$S(n_k) = 1/[\exp(-n_k)]$ for $k = \{i, j\}$. α *represent the mortality rate*
***Step 3**: determine the recall model as*
$n_i(t+1) = \mathrm{Max}\,[n_i(t), \mathrm{Max}_k\{(n_k \ \mathrm{Min}\ w_{ki})\}]$
***Step 4**: repeat step 2 and 3 (Encode-Recall Cycle) until steady state condition is reached.*
***Step 5**: Stop*

5. EXPERIMENT AND RESULT.

For measuring the accuracy, we test the proposed system several times and found that the performance of encoding and recall cycle is satisfactory. Following figure shows some of screen shot while testing the system.



1 **2**

Figure 8. Screen shot while testing encoding phase of the system

Figure 9. Screen shot while testing recall phase of the system

The results in Tables show that in the majority of the attempts are successfully recognized and the corresponding object are displayed in the screen. This system is tested on different noisy environment and it will produce the accurate result irrespective of any noise. The overall accuracy of the system is about 85%.

Word	Number of attempts	success	failure	Accuracy (%)
Home	50	43	7	86
apple	50	41	9	82
orange	50	42	8	84
car	50	40	10	80
tree	50	38	12	76
fish	50	45	5	90
Sun	50	43	7	86
Moon	50	44	6	88
			Average Accuracy	84

Alphabets	Number of attempts	success	failure	Accuracy (%)
A	50	42	8	84
B	50	41	9	82
C	50	43	7	86
D	50	45	5	90
E	50	46	4	92
			Average Accuracy	86.8

Numbers	Number of attempts	success	failure	Accuracy (%)
1	50	47	3	94
2	50	45	5	90
3	50	41	9	82
4	50	44	6	88
5	50	42	8	84
			Average Accuracy	87.6

Operation	Number of attempts	success	failure	Accuracy (%)
Plus	50	46	4	92
Minus	50	42	8	84
Multiply	50	41	9	82
Divide	50	43	7	86
			Average Accuracy	86

Figure 10. Accuracy matrix of different test case.

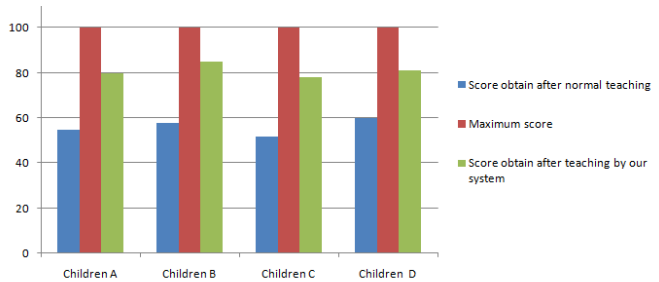
This system is tested on four children who suffer from this disability. First of all we taught them in traditional classroom system based on text book and take a test and the score obtain by each individual are recorded. After few days we taught them the same content using our system and again arrange a test. Significant improvement was found between two score card. All the children secured better score than the previous one and the encouragement of learning among the children is remarkable. They find more interest and motivation to learn and remember things more quickly. Figure [11] shows some experimental result of our proposed system.

Participating Children	Score Secured
Children A	55
Children B	58
Children C	52
Children D	60

[a]

Participating Children	Score Secured
Children A	80
Children B	85
Children C	78
Children D	81

[b]



[c]

Figure 11. [a] Score obtain before after normal teaching [b] Score obtain after teaching by proposed prototype [c]Graphical representation of the experiment result.

6. LIMITATIONS AND FUTURE WORK

In this proposed system some pre-defined word are stored in the vocabulary dictionary and only those word can be recognized by the Kinect sensor and corresponding image are drawn in the screen and mapped information will store in the database. These limitations can be overcome by making the vocabulary dictionary dynamic, therefore any word can be recognized by the system. More over this system can be used as a platform for device and appliance control at any place so that it will work as a assistive tool for aged and physically challenged people.

7. CONCLUSION

The main problem of the children with cognitive disability is the lack of interest and motivation to learn. Therefore a special type of care should be taken so that they can find themselves with higher self-confidence and stimulus environment for learning. The primary motive to achieve a mechanism for continues interest building towards learning has been significantly achieved. In this paper the proposed audio visual augmented interactive learning tool not only enhances the performance of the children effected by cognitive disability, but can also be used as significant support system towards elementary teaching to all level of children. The crucial accuracy of the system is 80-95% in achieving time versus learning outcome, which is a promising result and encourages the children to gain interest in continuous learning and make learning fun. This proposed first prototype is undergoing several real-life calibration like lighting condition, mobility, multilingual support etc. , refinement for individual proficiency and machine guided learning endeavored as a future goal of the work.

REFERENCES

[1] http://twbonline.pbworks.com/w/page/30889203/Inclusive%20Education.

[2] Abhijit Jana *"Kinect for Windows SDK Programming Guide Build motion-sensing applications with Microsoft's Kinect for Windows SDK quickly and easily"*.

[3] TraumaBot - 3D body reconstruction & recognition. Web resource link http://traumabot.blogspot.in/p/pcl-tutorial.html

[4] Amit Konar, 2005, *"Computational Intelligence ,Principles, Techniques and Applications".* Springer, India.

[5] Ke-Wei Chen et al, 2014. "*To design an interactive learning system for child by integrating blocks with Kinect".* In Global Engineering Education Conference (EDUCON), IEEE DOI: 10.1109/EDUCON.2014.6826061 , pp. 20 – 22.

[6] Yi-Zeng Hsieh. et al , 2013. *"To Develop the Virtual Physics Laboratory by Integrating Kinect with Gesture Classification Algorithm Signal-Image Technology & Internet-Based Systems(SITIS)",* International Conference on DOI: 10.1109/SITIS.2013.164. pp. 1017 – 1019.

[7] Thakkar, V. ; Shah, A. ; Thakkar, M. ; Joshi, A. ; Mendjoge, N. *"Learning Math Using Gesture "* Education and e-Learning Innovations (ICEELI), 2012 International Conference on DOI:10.1109/ ICEELI.2012.6360617 Publication Year: 2012 , pp. 1 – 3.

[8] Hariharan, B. ; Padmini, S. ; Gopalakrishnan, U. *"Gesture recognition using Kinect in a virtual classroom environment"* Digital Information and Communication Technology and it's Applications (DICTAP), 2014 Fourth International Conference on DOI: 10.1109/DICTAP.2014.6821668 Publication Year: 2014 , pp. 118 – 124.

[9] Tai-Yin Lin ; Chin-Feng Chen ; De-Yuan Huang ; Chi-Wen Huang ; Gwo-Dong Chen *"Using Resource of Classroom and Content of Textbook to Build Immersive Interactive Learning Playground"* Advanced Learning Technologies (ICALT), 2014 IEEE 14th International Conference on DOI: 10.1109/ ICALT.2014.78 Publication Year: 2014 , pp. 244 – 248.

[10] Hore, N.S. "*A Motion-Detection Biology-Based Learning Game for Children* " Potentials, IEEE Volume: 33 , Issue: 6 DOI: 10.1109/MPOT.2013.2295693 Publication Year: 2014 , pp. 31 – 36.

[11] Jaehong Lee ; Heon Gu ; Hyungchan Kim ; Jungmin Kim ; Hyoungrae Kim ; Hakil Kim *"Interactive manipulation of 3D objects using Kinect for visualization tools in education "* Control, Automation and Systems (ICCAS), 2013 13th International Conference on DOI: 10.1109/ICCAS.2013.6704175 Publication Year: 2013 , pp. 1220 – 1222.

[12] Fryia, G.D. ; Wachowiak-Smolikova, R. ; Wachowiak, M.P. *"Human-computer interface design in an e-Learning system for individuals with cognitive and learning disabilities "* Digital Information Management, 2009. ICDIM 2009. Fourth International Conference on DOI: 10.1109/ICDIM.2009.5356784 Publication Year: 2009 , pp. 1 – 6.