ISSN:0975-9646

Utsav Jambusaria et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (6) , 2014, 8119-8121

Finger Writing in Air using Kinect

Utsav Jambusaria[#1], Neerja Katwala[#2], Manali Kadam[#3], Harish Narula[#4]

[#1,2,3,4]Computers, D.J.Sanghvi College of Engineering
U-15 Bhakti Vedant Swamy Marg, Mumbai, 400056, India.

***Abstract*— The Microsoft Kinect released as an add-on accessory to the XBOX 360, has long gone beyond its originally intended purpose of an innovative gaming device. Its ability to capture colour and depth information to create a full 3 dimensional reconstruction of the human body in real-time has allowed avid developers to turn a person into a human remote. It has opened up an exciting avenue for Human-Computer Interaction (HCI) that promises to shake the very foundation of the way we interact with computerised devices. Our approach is aimed at introducing one such method, wherein our system correctly records and identifies finger drawn characters in the air. This has a sizeable potential to replace the traditional pen and paper method of teaching languages, as children are very likely to find this more interesting and interactive, and hence respond to it better than they do to conventional teaching methods.**

***Keywords*— Kinect, HCL, finger-writing, air, teaching**

I. INTRODUCTION

As computers get more and more integrated in our daily lives, HCI has become an integral area of research. And as the research intensifies, HCI is becoming increasingly closer to how people interact with each other. Gestures over clicks or button press, that is. HCI itself has come a long way. From the simple yet efficient keyboard and mouse, to joysticks and gamepads. From wired devices that limit our range of movement, to wireless ones with infrared support. However one thing remained a common constraint - all these require the use of an addition handheld hardware device.

The release of the Kinect is thus a very important breakthrough regarding this restraint. Using it, users can control devices using Gestures and body movements, without having to use any other hardware device. It offers them complete freedom of movement unlike any other experienced before.

In our application, users can use their fingers to draw alphabets or numerals in the air, which are then recorded and the forming pattern is input to a highly sophisticated handwriting recognition system, which gives the relevant output. This is done by first separating the hand, and thus the fingertip, from the rest of the body and a generally cluttered background. This in turn is based on the assumption that the human hand will be the closest part of the body to the Kinect, thus allowing for the depth sensors to identify and isolate it. Once identified, the trajectory of the hand and fingertip is then tracked and recorded, and it is this which is input to the handwriting recognition system.

We can thus see that for the success of this system in real-time usage, the following two criteria are absolutely imperative:

- It should provide an accurate conversion of the finger drawn input into the relevant character output, and in real-time.
- It should be able to give accurate character output even when introduced to different test cases which arise from different handwriting and the variability in the sizes of characters as drawn by different people in the air.

The greatest challenge we face is dealing with all these test cases, to ensure that this system works for every person who wishes to use it. If that is ensured, this can certainly be a very big breakthrough in the field of education as this hands-on, practical approach is sure to endear it to students.

II. HUMAN COMPUTER INTERACTION

HCI plays a significant role in most of the top, emerging technologies[4]. One of its applications is finger movement detection. Traditionally, a keyboard is used for any character input for computing technology. However, it has its limitations in size due to the size of our hands, precisely our fingers. To overcome this disadvantage, a video camera has been brought into use to detect the handwriting in air. The reason for the camera being used is that it can be miniaturized thus making the interface much smaller[5]. Similarly, in a Kinect, a video camera in the sensor enables the detection of the finger movement. This detection is thus helpful for the finger writing in the air. The characters drawn in the air are detected by the Kinect and processed later on the computer. The Kinect enables artless and effortless way which could be a significant for future development in HCI.

III. MICROSOFT KINECT

As we are aware, a Microsoft Kinect acts as a controller, replicating all our bodily movements occurring in real time within the game we are engaged in playing. The Kinect incorporates certain features which makes this innovation possible. It consists of advanced technical hardware such as cameras, sensors and four-microphone array that provide full-body 3D motion capture[2].

Fig. 1 (a) Microsoft Kinect Sensor

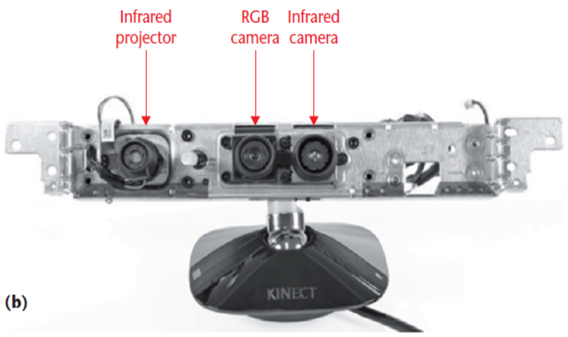

Fig. 1 (b) The components of the sensor.

A. *Depth Sensor*

- The depth sensor primarily consists of an IR projector and the IR camera.
- The IR camera is a CMOS(complimentary metal-oxide semiconductor) sensor, a commonly used technology for the designing of integrated circuits.
- The combination of these two helps the Kinect perceive the room in 3-D, regardless of proper illumination.

B. *RGB Camera*

- As the name suggests, RGB stands for Red, Green and Blue.
- The camera helps in detection of bodily movements and uses the above colour components for the same.
- It has 640 x 480-pixel resolution and runs at 30 FPS (frames per second)[3].

C. *Four-microphone array*

- Background noise is a very familiar yet a rampantly occurring disturbance.
- Voice control is an essential part of the Xbox 360.
- The microphone of the Kinect helps to distinguish between the voices of the users of the Kinect and the background noise in the room.
- This feature permits the users to use the voice control feature by even staying at some distance from the Kinect.

Thus, with these features, the Kinect helps in the application of HCI, finger writing in air.

IV. Hand Segmentation

The most important step to implement this application is separating the hand from the rest of the captured image. Hand Detection can be performed using varied techniques and the technique used will directly affect the performance of the proposed model. One such method is the skin colour model. It helps to distinguish between skin and non-skin attributes from an image. However it was found that this method was not as efficient in scenarios where the hand was held in front of the face, when the brightness contrasts or in a chaotic environment.

The data given by Kinect is a 640 x 480 grayscale image that encodes the distance of the scene object surfaces from the Kinect's viewpoint [1]. Since the data provided is 3-D distance using depth for segmentation will solve most of the disadvantages that the skin model had. The key conjecture that the depth model uses is that from the Kinect's viewpoint, the hand will be nearest to the image sensor. The depth model can approximately identify the hand location but not correctly determine the exact shape with distinct edges.

The third technique is background model is used to differentiate between the background and the foreground part of the captured image. The various pixel values of the hand and non-hand region are stored in the database in this method.

In order to calculate the results efficiently, the results of the individual methods are fed as input to an artificial neural network (ANN). The network has nine outputs, three for each input that will help in determining the factor to be used for calculations. Various combinations of the inputs are taken in order to check the consistency of the outputs. The ANN is trained with the robust back propagation algorithm (RPROP) with a sigmoid function as the activation function.

The findings of the ANN are the confidence factors of the three models a, b, and c. The training network will help to decide which of the techniques is most compatible with the other two and that is used for the final design. It is quite clear that all models influence the result and no model can be the most consistent one. The sum of the three models gives the required result.

V. Fingertip Detection

The next step is detecting the fingertip and separating it from the hand so as to easily map its moments.

A dual-mode switching algorithm[1] can be used for the same. All diverse kinds of hand postures will be deliberated by this method and apply fingertip-detection tactics. The side mode indicates that the finger is pointing away from the camera. It is presumed that the finger is furthermost from the arm point within the area specified by the hand detection with respect to the depth. In the frontal mode however the exact reverse is true as the fingertip is the nearest point to the camera with respect to the depth and it is generally pointing towards the camera.

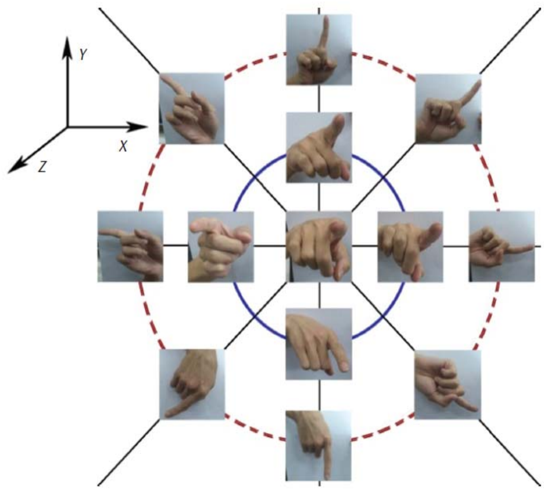

Fig. 2 Hand Postures which occur while writing

The mode(side or front) is decided based whether the furthermost point P from the hand should not be in the central part of the palm region. Taking basic human factors into consideration the angle between the hand and the finger (where they are pointing) should not be above a certain value. IF these criteria are fulfilled then the image belongs to the side mode. Otherwise, we switch to the frontal mode.

A black area has been observed around the fingertip in the depthimage because the infrared light is scattered bythe fingertip and depth values are incorrectlyset to 0. This is dealt with by using the inpainting techniqueto fill the hole using nearby pixels[1].

VI. Finger-Writing Character Recognition

The path followed by the finger is identified by using running simultaneous frames of fingertip positions.

A filter is then used to remove the noise which is produced due to weird hand movement and its wrongful classification. Our Character Classifier is able to distinguish between the 26 alphabets of the English language and all the characters of the Hindi language with the help of the model proposed above.

VII. Conclusions

The purpose of this paper is to make use of different hand detection and finger detection techniques to apply them in order to recognize the characters written. Using the sensors present in a Kinect we have created a classifier which is efficiently able to identify all characters of English and Hindi language. The ANN helps to make the classifier more efficient and help it to categorize a larger variety of hand postures.

Acknowledgment

We would like to thank our honourable principal Dr. Hari Vasudevan of D. J. Sanghvi College of Engineering and Dr. Narendra Shekokar, Head of Department of Computer Engineering, for giving us the facilities and providing us with a propitious environment for working in college. We would also like to thank S.V.K.M. for encouraging us in such co-curricular activities.

References

[1] X. Zhang, Z. Ye, LianwenJin, Z. Feng, S. Xu, A New Writing Experience: Finger Writing in the Air Using a Kinect Sensor, IEEE MultiMedia, vol. 20, no. 4, 2013, pp. 85-93.

[2] Z. Zhang, Microsoft Kinect Sensor and its Effect, IEEE MultiMedia, vol. 19, no. 2, 2012, pp. 4–10.

[3] (2010) Microsoft Kinect Working on Howstuffworks. [Online]. Available: http://electronics.howstuffworks.com/microsoft-kinect2.htm

[4] (2011) Human Interaction page on Wikipedia. [Online]. Available: http://en.wikipedia.org/wiki/Human%E2%80%93computer_interaction

[5] L. Jin et al., A Novel Vision Based Finger-Writing Character Recognition System, J. Circuits, Systems, and Computers (JCSC), vol. 16, no. 3, 2007, pp. 421–436.