Augmented Reality by using Hand Gesture

Jyoti Gupta
Computer Department
Dr. D.Y. Patil COE Pune,
India

Amruta Bankar
Computer Department
Dr. D.Y. Patil COE Pune,
India

Mrunmayi Warankar
Computer Department
Dr. D.Y. Patil COE Pune,
India

Anisha Shelke
Computer Department
Dr. D.Y. Patil COE Pune,
India

Abstract — **In this paper we have discussed a human-computer interaction interface (HCI) based on Hand gesture recognition. It is a challenging problem in its general form. We consider a set of manual commands and a reasonably structured environment, and develop a simple, yet effective, procedure for gesture recognition. Our approach contains steps for segmenting the hand region, locating the fingers, and finally classifying the gesture. The algorithm is invariant to translation, rotation, and scale of the hand. We demonstrate the effectiveness of the technique on real imagery.**

However, technology is progressing. In particular, Augmented Reality (AR), an emerging Human-Computer Interaction technology, which aims to mix or overlap computer generated 2D or 3D virtual objects and other feedback with real world scenes, shows great potential for enhancing e-commerce systems. Unlike VR, which replaces the physical world, AR enhances physical reality by integrating virtual objects into the physical world. The virtual object becomes, in a sense, an equal part of the natural environment. This chapter presents a new type of e-commerce system, AR e-commerce, which visually brings virtual products into real physical environments for user interaction. The new approach gives customers a chance to "try" a product at home or in another use environment. The chapter presents development of a prototype AR e-commerce system and a user study of the developed prototype. Experiment results and data both validate the new Human-Computer Interaction .AR e-commerce system and provide suggestions for improvement. Overall results of the study show that the AR e-commerce system can help customers make better purchasing decisions.

Keywords—Human–Computer Interaction(HCI), Gesture Recognition, Augmented Reality(AR)

I. INTRODUCTION

Augmented Reality is a technology in which we can mix computer-generated virtual objects with real world objects. Unlike VR, which experientially replaces the physical world, AR enhances physical reality by mixing virtual objects into a physical world. Generated virtual objects become an equal part of the natural environment. There are two types of AR:1. Optical see-though in which AR[5] uses a semi-transparent screen by using that screen computer generated objects can be displayed; users, can simultaneously see the computer generated images and the natural background environment and, thus, see an integrated AR scene. 2. Video see-through in which AR uses cameras to capture the live pictures as a video stream. For each viewed image frame, a captured video image frame is processed and computer generated virtual objects are added. One advantage of video see-through AR is that the

mixed scene can then be displayed on different devices. Between the two prominent AR methods, video-based AR has attracted the most attention from researchers.

From study of Augmented Reality we are developing an application in which we can navigate a virtual 3D model using hand gesture. In this the user interact with webcam and different virtual symbols on screen are provided for navigating 3D object, user will select a particular button from screen by placing his hand in front of that particular virtual key. We use Image Processing algorithms like Grayscaling [9], Blurring [6], Thresholding [7], HSV model [15], Blob Detection [8] for recognition of hand. After that user will get expected output. Like, if user had select zoomin button, then the 3D object will get zoomed. Our application can be implemented in Malls, Colleges, Museum, and Showroom.

II. AUGMENTED REALITY

A. Applications of Augmented Reality

AR represents the modern society's social-technological development. AR applications are being created by independent groups and organizations all over the world for use within many different fields. The goal of AR is to integrate 3D virtual objects as tools into real environment to get a realistic feelings of virtual objects.AR technologies can be designed to interact through many sensory channels (e.g. auditory, visual, olfactory, and haptic).

B. Gesture Recognition

Nowadays, gestures still are used by many people and natural interaction way for deaf people. In recent years, the gesture control technique has become a new developmental trend for many human-based electronics products, such as computers, televisions, and games. By using gestures recognizing technique people can control these products more naturally, conveniently. This technique also can be a well substitutive human-machine interaction way for some special people, such as deaf, dumb or physically disabled people, drivers, workers, even game players.

In this the user interact with webcam and different virtual symbols on screen are provided for navigating 3D object, user will select a particular button from screen by placing his hand in front of that particular virtual key. We use Image Processing algorithms like Grayscaling, Blurring,

Thresholding, HSV model, Blob Detection for recognition of hand.

Figure.1.Input Image

Figure.2.Virtual Buttons on Screen

The different virtual buttons which present on screen are move-up, move-down, move-left, move-right, zoom-in, zoom-out.Fig.2 shows different different buttons on screen.

Many researches indicate that the gesture control will become the new trend of HMI. In daily life, besides home life and work environment, in vehicle driving we can also use gesture recognizing technique.

III. SYSTEM WORKFLOW

At first, the input of the hand is accepted by the camera, the frame extraction is done on the accepted picture by removing the unnecessary background from the input image. To detect the exact location of the hand certain image processing algorithms are performed on the image of the hand to recognize the gesture. This is done by performing the various image processing algorithms like blurring, thresholding, RGB to HSV and blob detection.
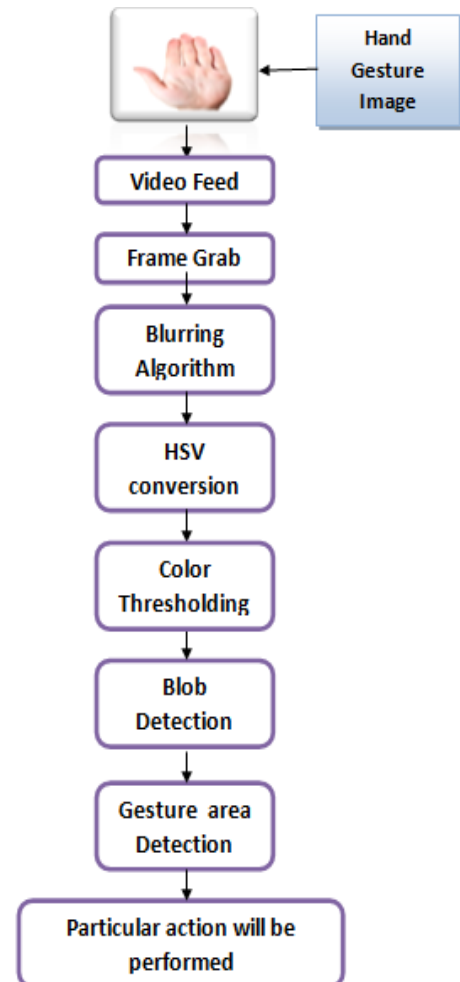

Figure.1 Workflow

IV. IMAGE PROCESSING

Image processing is any form of signal processing. For image processing input is an image and output is also an image such as a photograph or video frame or a set of parameters related to the image. In image-processing techniques we treat the image as a two-dimensional signal and we apply signal-*processing* techniques to it.

Augmented reality (AR) is a kind of created environment that allows the users to manipulate 3D virtual models. In this system, we are going to perform adaptive navigations on the presented model using the hand gestures. The system displays the sign icons on the screen, at both the boundaries. The user just has to take up his hands towards the icon he wants to bring into action. The webcam installed in the system detects the hand using the hand detecting algorithms.

A. Grayscaling

An image is collection of same type square pixels arranged in horizontal and vertical manner. In images each pixel has intensity varying from 0(black) to 255(white).

A grayscale image is normally seems like an black and white image, but as name indicate it include many shades of

grey. Each pixel is of 24 bit having 8 bit of Red, Green, Blue each.

For making grayscale image, we calculate the average value of RGB and assign average value to each of them. When all three primary color values are the same, the result is grayscale[9]. For example, if all three primary colors are 0 percent, the result is black. If all three primary colors are 100 percent (the maximum value), the result is white.

Calculation for Greyscaling

Consider a pixel from image and calculate its red, green and blue values. Here col represents the color from which RGB has to be extracted.

$$B = col \,\&\&\, 0XFF \qquad (1)$$
$$G = (col>>8) \,\&\&\, 0XFF \qquad (2)$$
$$R = (col>>16) \,\&\&\, 0XFF \qquad (3)$$

Now, calculate the average value of these three values

$$Avg = R+G+B/3 \qquad (4)$$

Now, assign this average value to R, G, B.

$$R=G=B=Avg \qquad (5)$$

Now combine the pixel
$$col = ((R<<16)|(G<<8)|B) \qquad (6)$$

B. Thresholding

Image segmentation is also done by Thresholding. We create binary images from a grayscale image, by using thresholding. In thresholding process, we randomly assume a threshold value from range 0-255.If pixel of image is having value less than threshold value then we assigning it value to "0" that means black else if pixel value is greater than threshold value then assigning it "1" that means white. Finally, a binary image is created by coloring each pixel white or black, depending on a pixel's labels.
Fig.1 shows original image and figure.2 shows threshold image.

Figure.1   Figure.2

Calculation for Thresholding

Select a threshold value from range 0-255

If Avg<Th
    R=G=B=0 (black)
else Avg>Th
R=G=B=1 (white)

C. Blurring

Blurring algorithm[6] is actually used to reduce noise and sharpness from image. In our project we are taking the input from webcam so for easier and faster detection of hand gesture ,we are reducing noise and sharpness of input image by applying blurring algorithm .There are two types of blurring 1) color blurring 2) Grayscale blurring.In our system we are using colorblurring .
Fig.1 shows original image and Fig.2 shows blur and grayscale image

Figure.1   Figure.2

Calculation of blurring

For 3X3 frame consider a pixel and surrounding 9 pixels, calculate RGB value of each pixel. Now take three constant R_{sum}, G_{sum}, B_{sum} whose initial value is zero.

$$R_{sum}= R_{sum} + R.$$
$$G_{sum}= G_{sum}+ G.$$
$$B_{sum} = B_{sum} + B.$$

Now,

$$R = R_{sum} / 9.$$
$$G = G_{sum} /9.$$
$$B = B_{sum} / 9.$$

D. Color Models

The purpose of a color model is to facilitate the specification of color. In our project the input image which we are taking from webcam is in RGB format. The color is determined using RGB color model after taking input from webcam. Once the webcam has read these values, they are converted to HSV values. For determining the location of hand gesture we are using HSV values in code. Each pixel from input image is matched with predefined color threshold value.

In HSV model, H is the Hue which represents the color type, S is the Saturation which represents vibrancy of the color and V is the Value which represents the brightness of the color. Hue can be described in terms of an angle on the circle. The hue value is normalized to a range from 0 to 255, with 0 being red. Saturation value range is from 0 to 255. The lower the saturation value, the more gray is present in the color. Value ranges from 0 to 255, with 0 being completely dark and 255 being fully bright. White has an HSV value of 0-255, 0-255, 255. Black has an HSV value of 0-255, 0-255, 0.

Figure.1 Original image   Figure.2 Hue

Figure.3 Saturation   Figure.4 Value

Calculation for RGB to HSV

Calculate R, G,B values of pixel then assign, maximum value among R,G,B to rgb_{max}, and minimum value to rgb_{min}. Initially h, s, v represents consider hue, saturation and value respectively.

$V = rgb_{max}$ or rgb_{min}

If v=0 , then

$$s=0 , h=0$$

else

$$S= 255 * rgb_{max}\text{-}rgb_{min}/V \qquad (1)$$

If s=0 h=0,

If $rgb_{max}=R$
then

$$H= 0+ 43*(G\text{-}B)/rgb_{max} - rgb_{min} \qquad (2)$$

If $rgb_{max} = G$

$$H= 85 + 43*(B\text{-}R)/rgb_{max} - rgb_{min} \qquad (3)$$

If $rgb_{max} = B$

$$H=171 + 43*(R\text{-}G)/ rgb_{max} - rgb_{min} \qquad (4)$$

And finally from Equation (1), (2), (3) and (4) we get values of hue , saturation and value as H,S,V.

E. Blob Detection

Blob detection refers to the module which aims at detecting the points or regions in the image that differ in properties like brightness, intensity of color compared to the surrounding. Blobs are group of pixels. Each pixel of blob belongs to predefined range. We scan each and every pixel from the input image from top to bottom; if that pixel belongs to particular range then we consider it as a single blob.

V. EXPERIMENTAL SET UP

In our experiment we require the webcam. The frame size of the webcam is 320*240. In this system low intensity webcam is preferred. The system is running on a computer with an Intel(R) Core(TM)i3-2370M CPU @ 2.40 GHz 2.40 GHz under Windows7.Installed memory (RAM) 2.00GB.

VI. FUTURE WORK

This system can be installed at the entrance of malls, colleges, museums and various exhibitions where the user can access the information about the place and get a complete view of it, prior to entering it. Sometimes, the system can also be used as similar to the maps to direct the place to reach the destination. Future work on this system includes the integration of speech recognition.

We can further add extra features in the system, so as it be used by the dumb and deaf people. Our hand gesture recognition can integrate with other application such as interactive game, smart home, auxiliary equipment and industrial control.

VII. CONCLUSION

In recent years, by using gesture control technique we can easily handle many human-based electronics products. This technique let people can control these products more naturally, instinctively and conveniently.

This paper proposes a gesture recognition scheme with well accuracy and efficiency. This scheme can be applied to be the human-machine interface so that user can handle many systems only with their hands. We can easily navigate a 3D model with hand gestures by providing different symbols of navigations on the screen. User has to take his hand in front of the symbol and the specified action will be performed. By using hand gesture in the system it has many characteristics including high learn ability, low mental load, high comfort, and intuitiveness.

REFERENCES

[1] Sheng-Yu Peng ,KanoksakWattanachote ,Hwei-Jen Lin ,Kuan-Ching Li, "A Real-Time Hand Gesture Recognition System for Daily Information Retrieval from Internet," *Fourth International Conference on Ubi-Media Computing*, 2011 IEEE.

[2] Ching-Hao Lai, "A Fast Gesture Recognition Scheme for Real-Time Human-Machine Interaction," *Conference on Technologies and Applications of Artificial Intelligence*, 2011

[3] Zhou Ren, Junsong Yuan, JingjingMeng, and Zhengyou Zhang, "Robust Part-Based Hand Gesture Recognition Using Kinect Sensor," IEEE transactions on multimedia, vol. 15, no. 5, august 2013.

[4] Chen-Chiung Hsieh and Dung-HuaLiou ,David Lee, "A Real Time Hand Gesture Recognition System Using Motion History Image, " *2nd International Conference on Signal Processing Systems (ICSPS),* 2010

[5] Ronald T. Azuma, "A Survey of Augmented Reality,"*In Presence: Teleoperators and Virtual Environments 6*, 4 (August 1997), 355-385

[6] http://en.wikipedia.org/w/index.php?oldid=436139840 *Contributors*: !melquiades, Alex:D, Army1987, AxelBoldt, Balizarde, Bawolff, BenFrantzDale, Cmdrjameson.

[7] http://en.wikipedia.org/w/index.php?oldid=451470153 Contributors: Anoko moonlight, Braksus, BrotherE, DARTH SIDIOUS 2, Dawoodmajoka.

[8] http://en.wikipedia.org/w/index.php?oldid=510104394 *Contributors*: 1ForTheMoney, Agent007bond, Casmith 789, Cbauckhage, Coupriec, Euchiasmus, Fjarlq

[9] http://heritage.stsci.edu/commonpages/infoindex/ourimages/color_comp.html

[10] " *Augmented Reality*" By Jason Separa and Gregory Kawano

[11] Yuen, S.; Yaoyuneyong, G.;& Johnson, E. (2011). "*Augmented reality: An overview and five directions for AR in education*". *Journal of Educational Technology Development and Exchange, 4*(1), 119-140.

[12] "*Converting from RGB to HSV*" BasketBall Robot,10-May-2005

[13] "*RGB and HSV colour models in colour identification of digital traumas*" Images Lidiya Georgieva, Tatyana Dimitrova, Nicola Angelov.

[14] Harpreet Kaur Saini, Onkar Chand," *Skin Segmentation Using RGB Color Model and Implementation of Switching Conditions*" International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 3, Issue 1, January -February 2013, pp.1781-1787