Design and Implementation of Interactive Augmented Trial Room

[1]Dr. N.Pughazendi, [2]G.Madankumar, [3]R.Rajkumar, [4]R.Ramsuraj
[1]Professor, [2,3,4]UG Scholar,
Department of computer science and engineering
Panimalar Engineering College

Abstract - This paper gives user friendly visual interface which auto-detect the human face and tries to merge the chosen accessories (either jewelry or eye-glasses) on them using a webcam as an input device and displays it to the screen based on Augmented Reality [AR]. With this, a lot time is saved to choose the accessories in a virtual display. To achieve this we use HAAR algorithm which takes the responsibility to detect the face thereby merging the accessory. Here the accessories are merged using the joints and position of the coordinates. Thus by doing so, the accessories are automatically positioned to the detected human face using an affine transformation. In addition, our proposed paper also detects the red pixels on the user's finger tip to change the accessories based on the gesture automatically. Thus this makes an user-friendly virtual trial room application instead of a real-time trail room.

Key words: Augmented reality, Merging accessories, fitting system, face pose and scale.

I INTRODUCTION

Augmented reality (AR) is an emerging form of experience in which the real world (RW) is enhanced by computer-generated content which is tied to specific locations and/or activities. In simple terms, AR allows digital content to be seamlessly overlaid and mixed into our perceptions of the real world. In addition to the 2D and 3D objects which many may expect, digital assets such as audio and video files, textual information, and even olfactory or tactile information can be incorporated into users' perceptions of the real world. Collectively, these augmentations can serve to aid and enhance individuals' knowledge and understanding of what is going on around them. Rather than seeming out of place, the digital markups inherent in AR lets users perceive the real world, along with 'added' data, as a single, seamless environment.

Augmented trial room is a digital interactive platform that helps the shoppers try out glasses and jewelries quickly and seamlessly. The augmented reality program captures the image of the customer and selects the desired accessory in front of personal computer which allows the customer to see how they look in it.

Today, mobile commerce and online sales are increasing at a rapid rate. Sales on mobile devices increased to 9.8 percent from 3.2 percent year over year. Nevertheless, It is estimated that majority of the consumers don't buy jewelry's online because they don't want to take any risk . This brings an additional financial burden to retail companies. Therefore, the objective of this work is to develop a virtual fitting room (VFR) application that can run on any mobile device that has a camera and network connection. This VFR application can enhance the way customers shop online and help them to choose the correct type of jewelry's item.

The inflection point for online fashion shopping will happen if your house can become the trial room. In e-commerce categories like apparel, personal accessories and jewelry, customers still like to touch, feel and try before purchase. Even though many e-tailers have lenient and efficient return policies, and some have virtual trial rooms that enable buyers to check out how products look on them via an interactive webcam, these are still not options that customers are comfortable with.

Though science fiction long ago introduced the concept of AR into our collective consciousness, until very recently, many might have labeled it a feature of our distant future. Now, however, we are riding the crest of a technological wave. AR is on the verge of becoming a household term, and perhaps, for many, an inseparable part of everyday life.

What is Augmented Reality?

Augmented reality refers to a wide spectrum of technologies that project computer generated materials, such as text, images, and video, onto users' perceptions of the real world. Initially, researchers defined AR in terms of specific facilitating devices, such as head mounted displays (HMDs). However, arguing that such definitions were too simplistic for an evolving and expanding field, Azuma and other researchers defined the implementation of AR by three characteristics: (a) the combination of real-world and virtual elements, (b) which are interactive in real-time, and which (c) are registered in the display of virtual objects or information is intrinsically tied to real-world loci and orientation. Similarly, Höllerer and Feiner define AR systems as those which combine "real and computer-generated information in a real environment, interactively and in real time, and virtual objects with physical ones". Ludwig and Reimann define AR as "human-computer-interaction, which adds virtual objects to real senses that are provided by a video camera in real time" . Zhou, Duh, and Billinghurst simply define AR as technology "which allows computer generated virtual imagery to exactly overlay physical objects in real time".

Augmented Reality and Virtual Reality

While considering the array of developing technologies, all seeking to modify, augment, interface with, or even replace our perceptions of reality, Milgram and Kishino sought to clarify the work being done by defining four types of environments. First is the real world, or the real environment, which we are all familiar with. On the opposite end of the scale are virtual worlds, or virtual environments (often previously labeled virtual reality), in which all information perceived by the user is computer-generated and completely unrelated to real world locations, objects, or activities. Between these two extremes exist, at least conceptually, two types of augmented environments: Augmented Reality (AR) which takes the real world and real environments as its backdrop and inserts computer-generated content, and augmented virtuality, in which a computer-generated world serves as the backdrop while real-world data is blended in and superim-posed.

III NEED FOR PROPOSED SYSTEM

The already existing system has the body scanning implemented using Microsoft's Kinect and Asus' Xtion devices. A VFR implementation by JCPteen gets an image of the user and using adobe flash player displays the clothing items. At the beginning, it shows a shadow on the screen where users have to fit themselves and after that the cloth is displayed. It is based on the augmented reality concept. The VFR doesn't consider the proportions of the user, only shows how it looks as a fixed template. Similarly, Swivel is labeled as a Try-On system that let users to see how clothes and accessories look on them in real-time.

The major drawback of previously existing system is that the initial cost spent for constructing the system and the maintenance cost of it. So in order to overcome these drawbacks we are coming up with a low cost system which yields out the same result.

Description of proposed system:

We get the frame continuously from the camera and it undergoes to the process. Then it is converted into grayscale images which makes it easier to eliminate the background pixels and other unwanted frames. Using HAAR cascade classifier we detect the face from the capturing frame. Once if the face is detected, then it will show the jewel on the neck of the user. Users are able to select a jewel and they can adjust the position of the jewel. In order to change the jewel, next and previous image of the jewel is selected using the red cap (red pixel) moving to the specified region.

The Raspberry pi

The Raspberry Pi is a low cost, **credit-card sized computer** that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. It's capable of doing everything you'd expect a desktop computer to do, from browsing the internet and playing high-definition video, to making spreadsheets, word-processing, and playing games.

What's more, the Raspberry Pi has the ability to interact with the outside world, and has been used in a wide array of digital maker projects, from music machines and parent detectors to weather stations and tweeting birdhouses with infra-red cameras. We want to see the Raspberry Pi being used by kids all over the world to learn to program and understand how computers work.

Raspberry Pi Technical Features :

The Raspberry Pi features:

- The Broadcom BCM2835 ARM11700Mhz 'System on Chip' Processor (Similar performance to a 300MHz Pentium 2 Processor).
- 512Mb SDRAM
- The free, versatile, and highly developer friendly Debian GNU/Linux Operating System
- 2 x USB Ports
- HDMI Video Output
- RCA Video Output
- 3.5mm Audio Output Jack
- 10/100Mb Ethernet Port for Internet Access
- 5V Micro USB Power Input Jack
- SD, MMC, SDIO Flash Memory Card Slot
- 26-pin 2.54mm Header Expansion Slot (Which allow for peripherals and expansion boards.

IV HAAR ALGORITHM

The human face poses even more problems than other objects since the human face is a dynamic object that comes in many forms and colors. However, facial detection and tracking provides many benefits. Facial recognition is not possible if the face is not isolated from the background. Human Computer Interaction (HCI) could greatly be improved by using emotion, pose, and gesture recognition, all of which require face and facial feature detection and tracking.

Although many different algorithms exist to perform face detection, each has its own weaknesses and strengths. Some use flesh tones, some use contours, and other are even more complex involving templates, neural networks, or filters. These algorithms suffer from the same problem; they are computationally expensive. An image is only a collection of color and/or light intensity values. Analyzing these pixels for face detection is time consuming and difficult to accomplish because of the wide variations of shape and pigmentation within a human face. Pixels often require reanalysis for scaling and precision. Viola and Jones devised an algorithm, called HAAR Classifiers, to rapidly detect any object, including human faces, using Ada Boost classifier cascades that are based on HAAR-like features and not pixels.

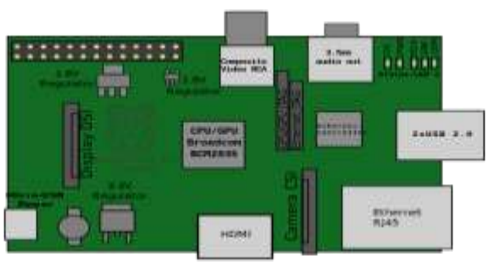

Fig. 4.2 Sectional view of components in the system

The architecture design shown in fig. 4.2 has an ARM11(BCM2835) processor which consists of power supply unit, SRAM(512MB) slot, Flash drive(8GB) ,2 x USB slots, HDMI port, stereo audio out, composite video RCA, Ethernet RJ45 and display OSI.

V **ARM VIRTUAL ADDRESSES (STANDARD LINUX KERNEL ONLY)**

As is standard practice, the standard BCM2835 Linux kernel provides a contiguous mapping over the whole of available RAM at the top of memory. The kernel is configured for a 1GB/3GB split between kernel and user-space memory. The split between ARM and GPU memory is selected by installing one of the supplied start*.elf files as start*.elf in the FAT32 boot partition of the SD card. The minimum amount of memory which can be given to the GPU is 32MB, but that will restrict the multimedia performance; for example, 32MB does not provide enough buffering for the GPU to do 1080p30 video decoding. Virtual addresses in kernel mode will range between 0xC0000000 and 0xEFFFFFFF. Virtual addresses in user mode (i.e. seen by processes running in ARM Linux) will range between 0x00000000 and 0xBFFFFFFF. Peripherals (at physical address 0x20000000 on) are mapped into the kernel virtual address space starting at address 0xF2000000. Thus a peripheral advertised here at bus address 0x7Ennnnnn is available in the ARM kernel at virtual address 0xF2nnnnnn.

ARM physical addresses

Physical addresses start at 0x00000000 for RAM.

• The ARM section of the RAM starts at 0x00000000.

• The VideoCore section of the RAM is mapped in only if the system is configured to support a memory mapped display (this is the common case).

The VideoCore MMU maps the ARM physical address space to the bus address space seen by VideoCore (and VideoCore peripherals). The bus addresses for RAM are set up to map onto the uncached1 bus address range on the VideoCore starting at 0xC0000000. Physical addresses range from 0x20000000 to 0x20FFFFFF for peripherals. The bus addresses for peripherals are set up to map onto the peripheral bus address range starting at 0x7E000000. Thus a peripheral advertised here at bus address 0x7Ennnnnn is available at physical address 0x20nnnnnn.

DMA Controller Registers

The DMA Controller is comprised of several identical DMA Channels depending upon the required configuration. Each individual DMA channel has an identical register map (although LITE channels have less functionality and hence less registers).

DMA Channel 0 is located at the address of 0x7E007000, Channel 1 at 0x7E007100, Channel 2 at 0x7E007200 and so on. Thus adjacent DMA Channels are offset by 0x100. DMA Channel 15 however, is physically removed from the other DMA Channels and so has a different address base of 0x7EE05000.

VI RESULTS AND DISCUSSION

The devices which are used in designing of the augmented trial room are Raspberry pi,web camera and a normal personal computer. Here the human face is detected using HAAR algorithm. The frame is got continuously from the camera and then it sent to the server side. Then detect the face of user, if the human face is detected then the jewel on neck or eye glass are merged accordingly and then send back to the users side display. In this process we use HAAR algorithm for detecting human face on frame.

For the present, while movements towards increased and improved distance education, both through real-world, AR, and VR interfaces are underway and seemingly speeding along, the incredible rapidity of technological change and development is a wave that educators have grown used to riding. While the world is undoubtedly changing, we will grow and adapt along with it. In fact, if we have our way, we will be the ones on the forefront, pushing forward with new innovations and improvements for teaching and learning. On a calmer note, most current educators will find that, while it is possible for them, as individuals, to create AR content using the tools mentioned earlier in this paper, truly user-friendly AR creation tools may still be just over the horizon. With that in mind, our task remains to keep examining what is available, and creatively utilizing it, to keep driving change forward

Fig. AR1 Positioning of the selected jewelry

Fig. AR2 Positioning of the selected eye-glass

The first step in facial feature detection is detecting the face. This requires analyzing the entire image. The second step is using the isolated face(s) to detect each feature. Since each the portion of the image used to detect a feature is much smaller than that of the whole image, detection of all three facial features takes less time on average than detecting the face itself. Using a 1.2GHz ARM11processor to analyze a 320 by 240 image, a frame rate of 3 frames per second was achieved. Since a frame rate of 5 frames per second was achieved in facial detection only by using a much faster processor, regionalization provides a tremendous increase in efficiency in facial feature detection.

The simple rectangular features of an image are calculated using an intermediate representation of an image, called the integral image. The integral image is an array containing the sums of the pixels' intensity values located directly to the left of a pixel and directly above the pixel at location (x,y) inclusive.

The Virtual Dressing Room using Augmented Reality would hence prove to make e-shopping for the customers more efficient and thus be beneficial to the e-shopping websites as well. By implementing each of the above steps accurately one can efficiently build the Virtual Dressing Room application which will prove to be a huge success and take online shopping altogether to a new level.

We conclude by saying that, the work done here is to create a Virtual Trial Room based on Augmented Reality. Here the user can visually feel how does he look when he wears the accessories and jewelries without touching the in the display screen. This VFR application can enhance the way customers shop online and help them to choose the correct item.

VII CONCLUSION AND FUTURE WORK

In many previous works the human has to adjust himself such that he fits into the given image displayed in the screen. This makes uncomfortable for the user, the size of the human may exceed the given image on the screen and the other checkbox method are also difficult in practical in which the user has to carry the checkbox to fix the points.

Thus our proposed system is more reliable than the previous existing systems. In this the human face is detected automatically so that the user may not fit in the given image. The face is detected automatically and then the given accessories are merged accordingly to the corresponding coordinates but the depth sensing is impossible in this.

One of the drawback of this system is that speed, the speed frame capture is low this has to be enhanced for better usage of this system. And in the proposed scheme only accessories like glasses and necklaces are merged with the human. In future more accessories like cloths, ear rings, caps and more things are to be added to make this system more useful in future.

Since there is an increase in online shopping this system could help for the better selection of the online products where the customer can visually feel how the given accessories suit him/her. These things are to be done in the future. These platforms are not only powerful decision tools for on-line shopping but also contribute to the fun factor for in-store shopping. In addition, the incorporation of social media in VTRs allows shoppers to obtain quick advice and feedback (e.g., Tweeter, Facebook) for a more pleasing and interactive shopping experience. Smartphone applications can also alert nearby stores to wishlisted items so that local shopping and deal finding can be usefully combined.

REFERENCES

[1] W. Broll, I. Lindt, J. Ohlenburg, I. Herbst, M. Wittkamper, and T. Novotny, "An infrastructure for realizing custom-tailored augmented reality user interfaces," *IEEE Trans. Visualiz. Comput.Graph., vol. 11*, no. 6, pp. 722–733, Nov.–Dec. 2005.

[2] D. Kalkofen, E. Mendez, and D. Schmalstieg, "Comprehensible visualization for augmented reality," *IEEE Trans. Visualiz. Comput.Graph.*,vol. 15, no. 2, pp. 193–204, Mar.–Apr. 2009.

[3] S. Lee, J. Choi, and J. Park, "Interactive E-learning system using pattern recognition and augmented reality," *IEEE Trans. Consum. Electron.*,vol. 55, no. 2, pp. 883–890, May 2009.

[4] M. Sidhu and L. Kang, "Interactive augmented reality environments for engineering with open systems," in *Proc. ICOS, 2011, pp. 1–5.*

[5] A. Dias, M. Ordaz, and J. Lopez, "From 2D to augmented reality," in*Proc. CISTI*, 2011, pp. 1–7.

[6] Kyle Chard and Kris Bubendorfer and Simon Caton and Omer Rana, "Social Cloud Computing: A Vision for Socially Motivated Resource Sharing," IEEE Transactions on Services Computing, vol. 4, no. 4, 2011.

[7] Foster, Ian and Kesselman, Carl and Tuecke, Steven, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," International Journal of High Performance Computing Applications, vol. 15, no. 3, pp. 200–222, August 2001.

[8] E. Korpela, D. Werthimer, D. Anderson, J. Cobb, and M. Leboisky,"SETI@homemassively distributed computing for SETI," Computingin Science Engineering, vol. 3, no. 1, pp. 78–83, jan/feb 2001.

[9] A. Beberg, D. Ensign, G. Jayachandran, S. Khaliq, and V. Pande, "Folding@home: Lessons from eight years of volunteer distributed computing," in Parallel Distributed Processing. IPDPS 2009., May 2009, pp. 1–8.

[10] J. Howe, "The Rise of Crowdsourcing," Wired Magazinehttp://www.wired.com/wired/archive/14.06/crowds.html, June 2006.

[11] D. Anderson, "BOINC: a system for public-resource computing and storage," in Grid Computing. Fifth IEEE/ACM International Workshop on, Nov 2004, pp. 4–10.

[12] Top 500 Supercomputer Sites, "November 2010 | Top 500 Supercomputing Sites," http://www.top500.org/lists/2010/11, June 2011.

[13] F. Costa, L. Silva, and M. Dahlin, "Volunteer cloud computing: Mapreduceover the internet," in Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW).

[14] P. Cesar, J. Vierinen, and P. Vuorimaa, "Open graphical framework for interactive TV," in *Proc. ISMSE*, 2003, pp. 21–28.

[15] K. Kang, J. Kim, H. Lee, H. Chang, S. Yang, Y. Kim, *et al*., "Metadata broadcasting for personalized service: A practical solution," *ETRI J.*, vol. 26, no. 5, pp. 452–466, 2004.

[16] S. Kim, J. Cha, J. Kim, J. Ryu, S. Eom, N. Mahalik, *et al*., "A novel test-bed for immersive and interactive broadcasting production using augmented reality and haptics," *IEICE Trans. Inf. Syst*., vol. E89-D, no. 1, pp. 106–110, 2006.

[17] G. Holbling, T. Rabl, D. Coquil, and H. Kosch, "Interactive TV services on mobile devices," *IEEE Multimedia*, vol. 15, no. 2, pp. 72–76, Apr.–Jun. 2008.

[18] I. Lee, H. Lee, J. Hong, and J. Cha, "Interactive contents player for rich media service," *IEEE Trans. Consum. Electron.*, vol. 55, no. 1, pp. 112–118, Feb. 2009.

[19] C. Mart´ın, L. Garc´ıa, J. Men´endez, and G. Cisneros, "Access services based on MHP interactive applications," *IEEE Trans. Consum.Electron.*,vol. 56, no. 1, pp. 198–202, Feb. 2010

.[20] Y. Lim, K. Kim, and J. Jeong, "Hybrid interactive data service architecture over T-DMB with mobile network," *IEEE Trans. Consum.Electron.*,vol. 56, no. 4, pp. 2826–2833, Nov. 2010.