Explicit Human Computer Interaction based virtual system for dynamic clothing and fitting simulation.

Mohammed Ali Chherawalla[1], Prashant Bhatt[2], Karan Gadiya [3], Eshaan Thakkar[4]

[1] *Student, Computer Engineering, Vishwakarma Institute of Information Technology, Maharashtra, India*
[2] *Student, Computer Engineering, Vishwakarma Institute of Information Technology, Maharashtra, India*
[3] *Student, Computer Engineering, Vishwakarma Institute of Information Technology, Maharashtra, India*
[4] *Student, Computer Engineering, Vishwakarma Institute of Information Technology, Maharashtra, India*

ABSTRACT

Small boutiques have a space constraint and hence have a single changing room which creates a bottleneck. People avoid visiting such boutiques since the entire process takes a very long time. This system uses an Android Phone, Leap Motion Controller, Google Cardboard and a mirror that aims at eliminating the space requirements and cumbersome process of physically trying on clothes. This system aims to overcome these challenges by showing the user how clothing articles will look on him or her. The user wears the headmounted gear and stands in front of the mirror, the application captures the user's image as it is seen in the mirror. The catalogue of the mall or boutique containing all the clothing articles available is present in the application, the user selects the clothing article he or she would like to try on. The selected clothing article is then rendered on this image of the user and the user is able to understand how that item would appear on him or her. The interaction with the system is done entirely using gestures to make the interaction more natural and to imbibe embedded hand gesture control to create a completely virtual experience with easy to understand working.

Keyword : - *Virtual Reality, Gesture Recognition, Leap Motion, Google Cardboard, Image Processing, Computer Vision*

1. Introduction

In the current trend of online shopping and ecommerce, there are still a large amount of people who prefer going to shopping malls to see the clothes that they want, as not everyone prefers shopping on a computer. As we know, a mall is a place where everyone, no matter what category they belong to, go
for shopping. This product will benefit each and everyone in the same manner. Our basic goal concerns the mass population and that is based on an unbiased concept. Installing this system in such a mall might just attract the attention of shoppers looking for a less tedious shopping experience. However, in a country like India, even after the sprouting of a large amount of malls,

we do face a lot of crowd and lines when it comes to trying out new clothes at malls. A major reason for this is the amount of the time
that goes in an individual trying out various clothes before selecting a final option. Our system will eliminate this tedious process, when it comes to seeing the appearance of the dress on the body of an individual. Malls that are popular and are functioning properly will make the most use of this system and the crowd waiting for changing rooms will reduce. Subsequently, the space required to build numerous changing rooms will also be reduced, leaving more space for shopping purposes. The user wears the headmounted gear and stands in front of the mirror, the application captures the user's image as it is seen in the mirror and then application eliminates the background with the help of background elimination techniques. This captured image is then rendered upon with the 3D vector image of the clothes available at the particular mall and displayed to the user.

Image processing is the mathematical treatment of an image and manipulation of the image according to the requirements of the user. Image processing refers to Digital image processing but optical and analog image processing is also possible. Image processing is closely related to computer graphics and computer vision. In computer graphics images are manually made from physical world object without acquiring them from some device like camera. Computer vision on the other hand is high level image processing and in this an image is acquired using a camera and various techniques are then applied to manipulate that image. In our system we will be using the computer vision part of image processing. We have a camera that will acquire the image from a mirror and then using image processing algorithms modifications like background subtraction, background illumination would then be applied to create a final image.

Virtual reality (VR) is an immersive multimedia or computer simulated life. It replicates an environment that will give the user an illusion of physical presence at real world places or an imagined world. VR concepts have been widely used in developing gaming applications. Nowadays some advance system use tactile information known as the force feedback along with VR to create more lifelike applications. Providing the user a completely hands free experience we will be using this concept. The user will select an item and then the item would be rendered on him so that he/she gets an idea as to how that clothing would look on him/her.

Computer vision is a field that includes methods for acquiring, processing, analyzing and understanding an image basically manipulating a high dimensional data from the real world and generating useful information from that. In more technical terms computer vision aims at creating artificial systems that extract valuable information from images. The image data could be in any for views form multiple cameras, video sequence or a multidimensional data.

2. Literature Survey

Among a variety of gestures, hand gesture is the most expressive and the most frequently used. Human gestures constitute a space of motion expressed by the body, face, and,or hands. The proposed technique overcomes background complexity and distance from camera which is up to 2.5 meters. Experimental results show that the proposed technique can recognise 12 gestures with rate above 94 percent [1]. Recent technologies in vision sensors are capable of capturing 3D finger positions and movements. The positions of fingers are precisely captured by a computer vision device. By tracking the moving patterns of fingers, we can then recognize users' intended control commands or input information. Experiments show promising recognition performance and speed. This paper presents a new type of user input for writing: given finger data from the Leap Motion controller, identify characters and words that are written in the air. This problem is novel because no pen up or pen down gesture exists that determines the beginning and end of data. Rather, characters must be recognized in real time[4].

Since edge detection is in the forefront of image processing for object detection, it is crucial to have a good understanding of edge detection algorithms. The commonly used edge detection algorithms are namely ISEF, Canny, MarrHildreth, Sobel, Kirsch, Lapla1 and Lapla2. A representative group containing the above seven algorithms are the implemented in C++ and compared subjectively, using 30 images out of 100 images. The Classification is based on the behavioral study of these edges with respect to differentiation operators like Gradient edge Detectors

(first derivative or classical), Zero crossing (second derivative), Laplacian of Gaussian (LoG), Gaussian edge detectors, and Coloured edge detectors[12].

Another technique of detecting object hidden in the background image is put forward by the authors. Camouflage detection method or Decamouflaging method is basically used to detect foreground image hidden in the background image. Authors have presented various camouflage detection methods for different applications and areas[13].

In [14], the authors have elaborated the use of Canny and Sobel edge detection algorithms for edge detection purpose. These algorithms are used to identify and locate sharp discontinuities in an image. Besides edge detection, its other uses are object recognition, target tracking, segmentation, data compression, image reconstruction and so on. They have also shown how Canny edge detection algorithm works better than Sobel edge detection algorithm's with the help of various experiments. In [15], various techniques for detecting moving objects using background cancellation are described. These techniques will prove useful in our system wherein we would have to imbibe the vector image of the selected cloth on the person even when the person moves. So this will make the system more flexible and people will have a better vision of how the dress would look on them. Larry Venetsky, Ross Boczar, and Robert LeeOwn depict background subtraction for image enhancement in their paper by improving the blob(connected component) content in the image. Quality blob content is needed to recognize and track objects. Three methods for background optimization were examined with the goal of finding a more robust denoising architecture consisting of the fusion of several background synthesis techniquesgenetic algorithm, an analytical solution based in convex optimization, and a related CVX solver toolbox implementation. These algorithms were used for the application of extracting foreground objects from aircraft carrier deck images. These objects were typically aircraft, related support equipment, or personnel. The incoming image frames were composed of multiple subimages ('cells') with sometimes very large differences in pixel intensities. This is because each cell was produced by a different camera under varying ambient lighting conditions as well as varying degrees of wear and tear from deployment [16].

Various techniques for detecting human faces in digital colour images are presented in [17]. Ones of these techniques uses a skin filter which relies on colour and texture information, after which face detection is performed on grayscale images containing only the detected skin areas. The process for detection of faces was based on a twostep approach. First, the image is filtered so that only regions likely to contain human skin are marked. This filter was designed using basic mathematical and image processing functions in MATLAB. The second stage involves taking the marked skin regions and removing the darkest and brightest regions from the map[17].

Authors in [19], have shown how imagebased face detection and recognition is done without the coordination of the user.Ada Boost classifier is used with Haar and Local Binary Pattern features whereas Support Vector Machine(SVM) classifier is used with Histogram of Oriented Gradients (HOG) features for face detection evaluation. They have shown the results of 5 different data sets and their respective results are demonstrated.

A background cut as well as a high quality and real time foreground layer extraction algorithm, from a single video sequence with a moving foreground object and stationary background, combines background subtraction, colour and contrast cues to extract a foreground layer accurately and efficiently. The key idea in this entire plan is background contrast attenuation, which adaptively attenuates the contrasts in the background while preserving the contrasts across fore-ground/background boundaries. This algorithm also builds upon a key observation that the contrast (or more precisely, colour image gradient) in the background is dissimilar to the contrast across foreground or background boundaries in most cases. Using background cut, the layer extraction errors caused by background clutter can be substantially reduced [20].

eShakti.com, Staten Island Mall, New York: The Staten Island Mall, the first one in New York City hosts the "virtual fitting room" via the ecommerce shopping website eShakti.com. This idea has been made possible with the help of 3D photography. In this approach an avatar of the person is created on the computer screen with precise measurement of body size, shape and posture which can then be used to customdesign clothing selected from an online catalogue. After the avatar is created whenever the customer shops at eShakti.com the avatar is made available to them automatically. This allows the customers to get madetofit clothing articles for themselves.

Rainbow Shopping Mall, Shenzhen, China: This mall has a virtual fitting room with 3D motion sensing which is used to try clothes, bags, shoes and other accessories by simple hand gestures

via virtual buttons right in the air. One can also snap a photo and share it with friends through Weibo and Wechat or can even add specific clothes into the cart for purchase.

3. Proposed System

Small boutiques have a space constraint and hence have a single changing room which creates a bottle neck. People avoid visiting such boutiques since the entire process takes a very long time. This system aims to overcome these challenges by showing the user how clothing articles will look on him or her.

In the current trend of online shopping and ecommerce, there are still a large amount of people who prefer going to shopping malls to see the clothes that they want, as not everyone prefers shopping on a computer. However, in a country like India, even after the sprouting of a large amount of malls, we do face a lot of crowd and lines when it comes to trying out new clothes at malls. A major reason for this is the amount of the time that goes in an individual trying out various clothes before selecting a final option. Our system will eliminate this tedious process, when it comes to seeing the appearance of the dress on the body of an individual.

This product is not specified or made to target a certain group of users. As we know, a mall is a place where everyone, no matter what category they belong to, go for shopping. This product will benefit each and everyone in the same manner. Our basic goal concerns the mass population and that is based on an unbiased concept. More than the user classes and types, we are more concerned about the areas where this system will be installed. Malls that are isolated and have a low customer count will not be able to benefit much from this system. However, on the other side, installing this system in such a mall might just attract the attention of shoppers looking for a less tedious shopping experience. Malls that are popular and functioning properly will make the most use of this system and the crowd waiting for changing rooms will reduce and subsequently, the space required to build numerous changing rooms will also be reduced, leaving more space for shopping purposes. Users that love shopping and are frequent visitors in malls to buy clothes will have a lot to reap out of this system as they will not have to keep on trying new clothes over and over again to see how they look on oneself.

Our system should mandatorily cover all the following objectives:

The system should be able to operate in real time via the Android mobile and with the help of the Leap Motion Controller will be able to react to the gestures of the people in real time too.

The transformation algorithms that will be used should be efficient enough to transform the image onto the body and move it according to the needs in real time or nearreal time The entire process of taking a pile of clothes to tryout in the changing room should be reduced and should in turn free up the area in the mall as well create more space for people to move around freely.

The user interface should be easy to understand and operate for the layman users as well. Basically, the main output of the entire system will be on an Android Mobile in the form of an Application.

Initially, we will have predefined 3D images of around 67 dresses that can be rotated and seen from all angles. Once that is done, we will be using the camera of an Android Mobile that is inserted in the Google Cardboard and a mirror in front of it. Google Cardboard is a virtual reality (VR) platform developed by Google for use with a foldout cardboard mount for a mobile phone.

In other words, it converts the 2D image of the mobile screen into a 3D image by applying barrel distortion. The Google Cardboard is a headgear which can be adjusted using the Velcro strap that is given. It has a slot for the mobile device which is right in front of our eyes. The 2D to 3D conversion takes place here. The customer will now have to wear the headset and stand in front of the mirror. Here, the high resolution camera on the mobile device will be facing the mirror and will detect the edges of the customers body. In order to ensure that the edge detection takes place depending on the customer's body size and not the clothes that the customer has already worn, we will have to ensure that the clothes are tucked in so that they can give a better perception of the body size. This will be dependent and will vary according to the size of each person, which in turn will be calculated by the edge detection algorithm. This information will be relayed to the Android Application. Once that is done, the 3D image of the selected dress will be embedded on the customer's body depending on the size that is selected. The dress will be in 3D and we can apply a rotation algorithm which, once the edge detection is done, will make the dress rotate as and when the customer rotates on an axis of 0 to180 (90 being

the initial position facing the mirror). Due to this, the customer will get the virtual feel of wearing the dress without having to go through the tedious process of waiting in line to try it out. The interfacing will be gesture based. Thus, if the customer wants to swipe through the catalogue of available dresses, he/she can do so by the flick of the wrist. They will be able to adjust the sizes by gesturing the system to zoom in and out. All this information, again, will be provided as an output to the mobile application, all in real time using the Leap Motion Controller to detect all these motions.

Inorder to use the Google Cardboard we need to extend the CardboardActivity which implicitly calls the onNextFrame() method which is not viable for us as OpenCV implicitly calls the onCameraFrame() method on the availability of new frames.Hence we came up with our own algorithm to detect objects in the image render the cloth on both sides so as to give a 3D perception of the cloth on the person. OpenCV runsdirectly on the linux kernel by bypassing the Dalvik Virtual Machine thereby enabling faster response time.

The following images demonstrate the selected clothes being rendered on three subjects:

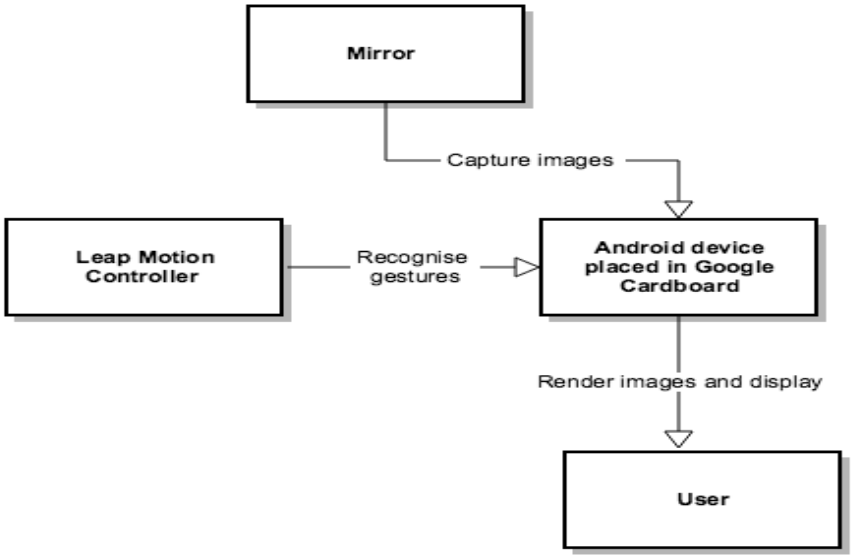
Fig-1: Test Subject 1

Fig-2: Test Subject 2

Fig-3: Test Subject 3

Frameworks used:
Android SDK: A software development kit (SDK or "devkit") is typically a set of software development tools that allows the creation of applications for a certain software package, software framework, hardware platform, computer system, video game console, operating system, or similar development platform.
Leap Motion Leap Motion Android SDK: The SDK required for the Leap Motion Controller
OpenCV:(Open Source Computer Vision) is a library of programming functions mainly aimed at realtime computer vision, originally developed by Intel. The library is crossplatform and free for use under the opensource BSD license.
NDK: The NDK is a tool set that allows you to implement parts of your app using nativecode languages such as C and C++. Typically, good use cases for the NDK are CPUintensive applications such as game engines, signal processing, and physics simulation.

4. System Architecture



The Fig-4: System Architecture

The given figure illustrates the flow of our entire system. The dimensions and posture of theperson standing in front of the mirror will be captured by the camera of the android application from the image of his or her reflection in the mirror. This captured image will be taken as input by the android application system; and along with it, the person's body dimensions like length from shoulder to shoulder, waist etc. For the system to gather appropriate coordinates the person will have to be properly dressed i.e. men will have to tuck their shirt in so that his captured image consists of the correct dimensions, or women will have to press their skirts close to their body to get appropriate dimensions. Once this is done, the person can choose over a variety of options available with the help of hand gestures. The entire user interface interaction will be gesturebased for further comfort. Once a cloth is selected by the person, it's vector image will be rendered on the image of the person that was initially captured by the camera and displayed. Thus the person will actually be able to visualize how the cloth looks on him or her.

The Google Cardboard: It is a head mounted gear that is used to convert 2D images to 3D images. There is a strap in which the user can keep his or her phone so that a 3D visualization of the mobile screen is created and displayed to the user. The strap is attached to a Velcro which makes the gear flexible for users with different head sizes to use it with comfort.

Leap Motion Controller: It comes into the picture as soon as the system starts and is used as an interface for the users. The user can select a cloth, swipe through all the choices available, return to the main menu, or to see how the cloth looks on him or her, all with the help of hand gestures. The Leap Motion Controller makes this possible.

4. Description of Functions

4.1. Conversion of 2D to 3D

- Description and Priority: This feature will form the basic output on the Android mobile and will make the user feel as if the screen is 3 dimensional. The priority can be rated as 9 or 10 as if there is a slight chance that this part fails, the user will still be able to see the dress on the mobile screen.
- Stimulus or Response Sequences :The main output of the entire system will be on an Android Mobile in the form of an Application. Initially, we will have predefined 3D images of around 67 dresses that can be rotated and seen from all angles. Once that is done, we will be using the camera of an Android Mobile that is inserted in the Google Cardboard and a mirror in front of it. Google Cardboard is a virtual reality (VR) platform developed by Google for use with a foldout cardboard mount for a mobile phone. In other words, it converts the 2D image of the mobile screen into a 3D image by applying barrel distortion. The Google Cardboard is a headgear which can be adjusted using the Velcro strap that is given. It has a slot for the mobile device which is right in front of our eyes. The 2D to 3D conversion takes place here.
- Functional Requirements:

1) Mobile should be an Android Mobile with a SnapDragon 800 processor or faster.
2) The Cardboard should have clear lenses so that it need not be replaced from time to time.
3) A Velcro strap should be present at all times and should have a flexible size so that it can fit the head of everyone ranging from a child to an adult.

4.2. Image Capturing and Processing

- Description and Priority: Image Processing is the key feature that will run this system. This system feature will be the highest prioritized feature and will provide the basic input to the system.
- Stimulus or Response Sequences: The customer will now have to wear the headset and stand in front of the mirror. Here, the high resolution camera on the mobile device will be facing the mirror and will detect the edges of the customers body. In order to ensure that the edge detection takes place depending on the customers body size and not the clothes that the customer has already worn, we will have to ensure that the clothes are tuckedin so that they can give a better perception of the body size. This will be

dependent and will vary according to the size of each person, which in turn will be calculated by the edge detection algorithm. This information will be relayed to the Android Application.

● Functional Requirements:

1. Camera should be of a high resolution 5MP with 30 fps recording capabilities is preferred.
2. Person should not move initially while the edge detection takes place.

4.3. Image Transformation

● Description and Priority:

Image Transformation takes place after the edges of the person standing in front of the mirror have been detected. This is the part where the dress will be imprinted on the body of the person and he will get an idea of how it will look. Again, this part will have the highest priority along with the Image Processing.

● Stimulus or Response Sequences:

The 3D image of the selected dress will be embedded on the customer's body depending on the size that is selected. The dress will be in 3D and we can apply a rotation algorithm which, once the edge detection is done, will make the dress rotate as and when the customer rotates on an axis of 0 to180 (90 being the initial position facing the mirror). Due to this, the customer will get the virtual feel of wearing the dress without having to go through the tedious process of waiting in line to try it out. The interfacing will be gesture based.

● Functional Requirements:

1) Pre defined 3D dress models that can be scalable.
2) Minimally crowded background.

4.4. Gesture Control

● Description and Priority: The system will be controlled via gestures. This is the part where the user will interact with the system. Priority will be 8 or 10 as if the gestures dont work, we can interact with physical buttons or touch panels.

● Stimulus or Response Sequences:The interfacing will be gesture based. Thus, if the customer wants to swipe through the catalogue of available dresses, he or she can do so by the flick of the wrist.They will be able to adjust the sizes by gesturing the system to zoom in and out. All this information, again, will be provided as an output to the mobile application, all in real time using the Leap Motion Controller to detect all these motions.

● Functional Requirements:

1) Leap Motion Android SDK.
2) Leap Motion Controller

5. Conclusions

Gesture recognition is the future of computer evolution This paper is based on the system "Gesture based Utility Tool". Gesture recognition gives a new dimension to computer human interaction. It is the point of collision of the physical 5world and the digital world where real world gestures can be used to communicate with the computer system. In the days to come online shopping for clothes is probable to become the next big thing in the world of eshopping and fashion designing. This tool uses gesture recognition and allows hands free communication with the computer system. This system will boost the use of virtual reality in day to day life of the common man and can be enhanced to include more advanced and complex features as well. Creating such a system in a common place such as a mall will encourage companies and people to imbibe it in other sectors and places as well. Hence this system can be used to efficiently reduce the bottle neck caused by the problem of limited space in malls and also ease the process of trying on and buying clothes. Combining all the aspects of online shopping, the ever growing use of android applications, and the use of gesture based flexible and comfortable tools we aim to provide the most complete system to.
the shopcrazy users of the modern world.

6. REFERENCES

[1] Mohamed Alsheakhali, Ahmed Skaik, Mohammed Aldahdouh, Mahmoud Alhelou, " Hand Gesture Recognition System " ,Computer Engineering Department, The Islamic University of Gaza Gaza Strip, Palestine, 2011.
[2] Tin Hninn Hninn Maung, "RealTime Hand Tracking and Gesture Recognition System Using Neural Network",
World Academy of Science,Engineering and Technology , Vol:3, 2009.
[3] A. Kurakin, Z. Zhang, Z. Liu :"A real time system for dynamic hand gesture recognition with a depth sensor", in 20th European Signal Processing Conference . 2012.
[4] Sharad Vikram, Lei Li , Stuart Russell, "Handwriting and Gestures in the Air, Recognizing on the Fly", ACM , 2013.
[5] "Design of gesture recognition system for dynamic user interface", in IEEE International Conference on Technology Enhanced Education (ICTEE), 2012.
[6] "A hand tracking framework using the 3D active tracking volume", in 1st International Conference on Emerging Trends and Applications in Computer Science (ICETACS) , 2013.
[7] "2D Human Gesture Tracking and Recognition by the Fusion of MEMS Inertial and Vision Sensors", IEEE Sensors Journal , vol. 14, no. 4.
[8] "Hand gesture recognition with depth images: A review", IEEE ROMAN , 2012.
[9] "Robust PartBased Hand Gesture Recognition Using Kinect Sensor, Multimedia", IEEE Transactions on vol. 15, no. 5.
[10] "A hidden Markov model based dynamic hand gesture recognition system using OpenCV", IEEE 3rd International on Advance Computing Conference (IACC) , 2013.
[11] "A Framework for RealTime Hand Gesture Recognition in Uncontrolled Environments with Partition Matrix Model Based on Hidden Conditional Random Fields, Systems, Man, and Cybernetics (SMC)", IEEE International Conference. 2013 .
[12] "A Classified and Comparative Study of Edge Detection Algorithms", Department of Computer Engineering Iran University of Science and Technology Narmak, Tehran 16844, IRAN mshar, mahfathy, tayefeh @iust.ac.ir
[13] Sujit K Singh, Chitra A Dhawale, Sanjay Mishra, "Survey of Object Detection Methods in Camouflaged Image", International Conference on Electronic Engineering and Computer Science , 2013.
[14] Dr. S. Vijayarani, Mrs. M. Vinupriya, "Performance Analysis of Canny and Sobel Edge DetectionAlgorithms in Image Mining", International Journal of Innovative Research in Computer and Communication Engineering , vol. 1, no. 8, October 2013.
[15] Aseema Mohanty, Sanjivani Shantaiya, "A Survey on Moving Object Detection using Background Subtraction Methods in Video", International Journal of Computer Applications (0975 8887) National Conference on Knowledge, Innovation in Technology and Engineering (NCKITE 2015)
[16] Larry Venetsky, Ross Boczar, Robert LeeOwn Advanced Technologies Branch, "Optimization of background subtraction for image enhancement", April 2013 DOI: 10.1117/12.2014319
[17] Ridhi Jindal, Anuj Gupta, Dr.Sonia Vatta, "Face Detection using Digital Image Processing", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 11, November 2013.
[18] Ebrahim Aghajari, Damayanti Gharpure, "Real Time VisionBased Hand Gesture Recognition for Robotic Application", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 3, March 2014.
[19] Faizan Ahmad, Aaima Najam and Zeeshan Ahmed, "Imagebased Face Detection and Recognition:State of the Art, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 1, November 2012.
[20] Jian Sun, Weiwei Zhang, Xiaoou Tang, HeungYeung Shu, "Background Cut", Microsoft Research Asia, Beijing, P. R. China.
[21] Staten Island Mall eShakti.com, Staten Island Mall, New York.
[22] Rainbow Shopping Mall
https://www.behance.net/gallery/24810775/3DVirtualFitting Roomin RainbowShoppingMall