**MAULANA AZAD**

**NATIONAL INSTITUTE OF TECHNOLOGY**

**BHOPAL, INDIA, 462003**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**3-D Virtual Trial Room**

**Minor Project Report**

**Semester**

**Submitted by:**

| Saumel Verma | 191112013 |
| --- | --- |
| Jatin Singh | 191112033 |
| Aditya Patel | 191112044 |
| Hrishik Verma | 191112079 |

**Under the Guidance of**

Dr .B.N. ROY

## DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**Session: 2019-2023**

**MAULANA AZAD**

**NATIONAL INSTITUTE OF TECHNOLOGY**

**BHOPAL, INDIA, 462003**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

# CERTIFICATE

This is to certify that the project entitled “**3-D Virtual Trial Room**”

Submitted by:

| Saumel Verma | 191112013 |
| --- | --- |
| Jatin Singh | 191112033 |
| Aditya Patel | 191112044 |

|  |  |
| --- | --- |
| Hrishik Verma | 191112079 |

## is the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering is a authentic work carried out by them under my supervision and guidance.

**Dr B.N. Roy**

**(Minor Project Mentor)**

**DECLARATION**

## We, hereby declare that the following report which is being presented in the Minor Project Documentation Entitled as “**3-D Virtual Trial Room**” is an authentic documentation of our own original work and to best of our knowledge. The following project and its report, in part or whole, has not been presented or submitted by us for any purpose in any other institute or organization. Any contribution made to the research by others, with whom we have worked at Maulana Azad National Institute of Technology, Bhopal or elsewhere, is explicitly acknowledged in the report.

| Saumel Verma | 191112013 |
| --- | --- |
| Jatin Singh | 191112033 |
| Aditya Patel | 191112044 |
| Hrishik Verma | 191112079 |

**ACKNOWLEDGEMENT**

## With due respect, we express our deep sense of gratitude to our respected guide and coordinator **Dr B.N. Roy**, for his valuable help and guidance. We are thankful for the encouragement that he has given us in completing this project successfully.

## It is imperative for us to mention the fact that the report of minor project could not have been accomplished without the periodic suggestions and advice of our project guide **Dr B.N. Roy** and project coordinators Dr. **Sanyam Shukla** and **Dr. Namita Tiwari.**

## We are also grateful to our respected **HoD, Dr. Nilay Khare** and Head AI and central computing facility, **Dr Meenu Chawla** for permitting us to utilize all the necessary facilities of the college.

## We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not the least; we would like to express our deep appreciation towards our family members and batch mates for providing the much needed support and encouragement.

**ABSTRACT**

Previous VTON’s had issues with result quality like blurriness , issues with long sleeve dresses and character preservation so the motivation behind our project is to reduce these redundancies  And make a Real time virtual dressing room that can be used in shops ,

mall and shopping centers. Trying clothes in shopping center is a time consuming activity. Besides, it might not be possible to try on clothes in such cases as online shopping. Hence one drawback is that it may be difficult for a person to visualize how a given article would look if worn by that person-owing to the rich variation in body size and shape, hair and skin color, etc., in the human population. This project's motivation is to increase time efficiency and improve the accessibility of clothes try-on by creating a virtual dressing room environment. Our aim is to build an interactive & highly realistic virtual machine on which the user can try cloths without wearing it actually. This was carried out by identifying critical points on garment and user’s body dimensions using image processing techniques

**Table Of Contents**

Certificate …………………………………………………………………………….. ……… (ii)

Declaration ………………………………………………………………………………………… (iii)

Acknowledgement ………………………………………………………………………………… (iv)

Abstract …………………………………………………………………………………………….. (1)

1. Introduction
   1. 1.1 Purpose (3)
   2. 1.2 Scope (3)
2. Literature Review And Survey
   1. 2.1 Existing Virtual Try-on (4)
   2. 2.2 Swapnet (4)
   3. 2.3 Viton (5)
   4. 2.4 SSIM (5)
   5. 2.5 IS (6)
3. Gaps Identified
   1. 3.1 Blurriness (7)
   2. 3.2 Long Sleeve Dress Issue (7)
   3. 3.3 Character Preserving (7)
4. Proposed Work And methodology
   1. 4.1 Dataset (8)
   2. 4.2 Methodology
   3. 4.2.1 Foreground Extraction (8)
   4. 4.2.2 Extraction Of Human Silhouette (9)
   5. 4.2.3 Feature Points Extraction (10)
   6. 4.2.4 Warping Of Shirt (10)
   7. 4.2.1 Virtual Fitting (11)
5. Results And Discussions
   1. 4.1 Best Case Result (13)
   2. 4.2 Combined Dataset Result (13)
6. Conclusion …………………………………………………………………………….. (14)
7. References …………………………………………………………………………….. (15)
8. **INTRODUCTION**

# Purpose

Despite increasing access to technology, people in the modern world are increasingly busy. For many, however, attention to one's appearance remains a high priority. Many people continue to invest time in maintaining and augmenting their wardrobes, shopping for special outfits, etc. In some cases, the investment in time has to do with going to a retail store to try on and purchase clothing and accessories. The process of selecting the right garment in the right size by trying on a series of candidate garment can be very time consuming. The purpose of the application is to make easier the process of trying clothes while shopping, which would provide comfort for both the vendor and the customer, reducing the time and helping people to select a wide range of clothing were a motivation to make a program that  helps in this area, so it has become important to make the process of trying and buying of clothes more comfortable, easier and more efficient.

# Scope

The accelerating pace of development in modern technology – and the software programs – and their dramatic entry into life have led to the development of this application on a large scale. One of the main reasons behind this tremendous development in technology is the direct interaction between man and computer. This type of application has become a hot topics of research since it is related to several areas in the human-computer interaction, such as interaction for the purposes of learning, entertainment, fields of medicine and e-commerce operations. E-commerce is one of the modern terms that have entered our daily life that they are used in many life activities that are related to the revolution in information and communication technology.

1. **Literature Review and Survey**
   1. **Existing Virtual Try-on**

Virtual try-on is the way a customer can **“try-on”** a product through mobile or other devices equipped with a camera. Thanks to the underlying AR technology, future customers can see themselves in a beloved product on the screen of their smart phone. Virtual try-on methods can be broadly classified into two categories: methods based on 3D body modeling , and methods based solely on 2D images . 3D methods can generate great results for virtual try-on, but require additional 3D measurements and more computing power. 2D image-based methods are more broadly applicable. Jetchev and Bergmann  proposed a conditional analogy GAN to swap clothing on people images, but requires paired clothing images to train the model. SwapNet proposed a method to interchange garment appearance between two single views of people. VITON and CPVTON generate new images given a target clothing item and a clothed person image, and are most relevant to the problem we are trying to solve.

* 1. **Swapnet**

A framework to transfer garments across images of people with arbitrary body pose, shape, and clothing. Garment transfer is a challenging task that requires

(i) disentangling the features of the clothing from the body pose and shape and

(ii) realistic synthesis of the garment texture on the new body.

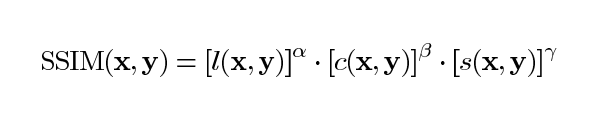
Swapnet is a neural network architecture that tackles these sub-problems with two task-specific sub-networks. Since acquiring pairs of images showing the same clothing on different bodies is difficult, we propose a novel weakly supervised approach that generates training pairs from a single image via data augmentation. We present the first fully automatic method for garment transfer in unconstrained images without solving the difficult 3D reconstruction problem. We demonstrate a variety of transfer results and highlight our advantages over traditional image-to-image and analogy pipelines.

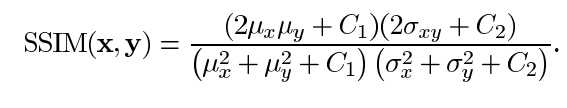
* 1. **VITON**

The virtual try-on system has gained great attention due to its potential to give customers a realistic, personalized product presentation in virtualized settings. In this paper, we present VTON, a novel pose-transfer-based framework for cloth transfer that enables virtual try-on with arbitrary poses. VTON can be applied to the fashion industry within minimal modification of existing systems while satisfying the overall visual fashionability and detailed fabric appearance requirements. It enables efficient clothes transferring between model and user images with arbitrary pose and body shape. We implement a prototype of VTON and demonstrate that our system can match or surpass many other approaches when facing a drastic variation of poses by preserving detailed human and fabric characteristic appearances. VTON is shown to outperform alternative approaches both on machine-based quantitative metrics and qualitative results.

* 1. **Structural similarity index measure (SSIM)**

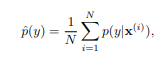
The structural similarity index measure (SSIM) is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. SSIM is used for measuring the similarity between two images. The SSIM index is a [full reference metric](https://en.wikipedia.org/wiki/Video_quality#Classification_of_objective_video_quality_metrics); in other words, the measurement or prediction of [image quality](https://en.wikipedia.org/wiki/Image_quality) is based on an initial uncompressed or distortion-free image as reference. SSIM is a perception-based model that considers image degradation as *perceived change in structural information*, while also incorporating important perceptual phenomena, including both luminance masking and contrast masking terms.

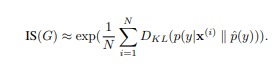




* 1. **Inception Score**

The Inception Score is a metric for automatically evaluating the quality of image generative models . This metric was shown to correlate well with human scoring of the realism of generated images from the CIFAR-10 dataset. The IS uses an Inception v3 Network pre-trained on ImageNet and calculates a statistic of the network’s outputs when applied to generated images.

****



1. **Gaps Identified**
   1. **Blurriness**

In the previous model the final result images were blur as compared to the original images, because when layers of original image are formed they have to be merged back properly but in previous vton’s there were issues in merging which caused the images to be blurry.

* 1. **Long sleeve dress issues**

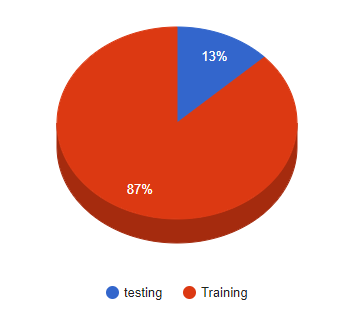
In previous vton’s there were issues with long sleeve dresses because of alignment with the users hands.

* 1. **Character Preserving**

In previous vton’s there were issues with alignment of person and clothes, for example if a person is leaning and the cloth is straight then previous vton’s would superimpose the cloth on the person thus distorting the posture hence there were issues with character preservation.

1. **Proposed work and methodology**
   1. **Dataset**

The dataset used is VITON. 19,000 front view females and top outfits images are there and removed distorted images resulting in 16,253 pairings. Training dataset compromises 14221 pairings and testing dataset compromise 2032 pairing. 18 Key points are being used. Each key point is then converted to a heat map, with an 11x 11 neighbourhood.



* 1. **Methodology**
     1. **Foreground Extraction**

The foremost step in the process of virtual fitting is to extract foreground from background. The algorithm works well on varying light intensities provided that the background has got rid of presence of highly projecting bodies. For reliable extraction, before edge detection, colour clustering by k means is performed. K-means clustering basically converts the image into clusters by quantifying the visual difference according to luminosity and chromaticity as shown in the Fig. 1(a). By doing so, the background is perfectly grouped to a separate cluster. Pre-processing of image is the delicate process which decides the reliability of further output. Hence, after these, pre-processing the image such as to remove small objects and morphological dilation is done.



Fig. 1(a)

* + 1. **Extraction Of Human Silhouette**

Extraction of human silhouette from the image in a reliable and efficient way by riding off the disturbances by background is the crucial step for virtual fitting. Background separation process is done by creating a morphological structuring element for applying laplacian filter. Edge detection Algorithm is employed to detect the edges of human. Then to obtain the perfect single line silhouette of human contour, traversal along two axes towards the centre has been carried out. After all, post-processing involves dilation, closing the contour are been performed as in Fig. 1(b).

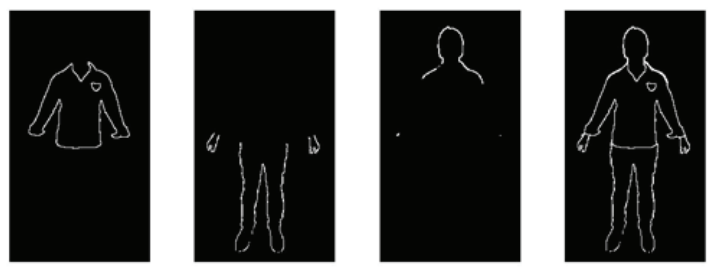


Fig 1(b)

* + 1. **Feature Points Extraction**

Feature points extraction is based on the much known morphological facts of human structure. The shoulder and hip points have to be found perfectly and precisely which plays a vital role in accurate fitting of attire. This involves the fact that at the shoulder of human body has the maximum slope considering below the face. So, before doing so, face is detected by Viola-Jones Algorithm. Another fact, being the hip of the human has the maximum width is employed here. Hip points are obtained by finding the maximum of the single line silhouette array as shown in Fig. 1(c). This is performed on both sides from the center giving the reference points for warping.



Fig 1(c)

* + 1. **Warping Of Shirt**

The sample shirt selected for fitting must be resized according to the size of the person. Thus the sample shirt is warped to the silhouette by taking feature points as reference. The Fig. 1(d) clearly shows the warping of shirt. This part of work requires tedious algorithm for fine results. Warping is the technique by which image is made to resize and reshaped such that one images superimposes over another image of similar contour. The shirt is divided to regions of distinguished variations and each being resized and concatenated over which interpolation technique is performed.



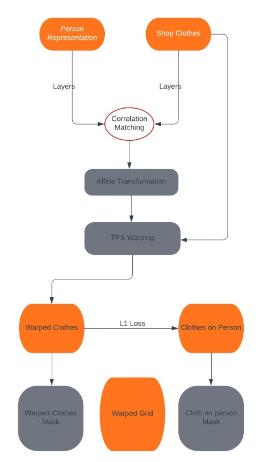
Fig 1(d)

* + 1. **Virtual Fitting**

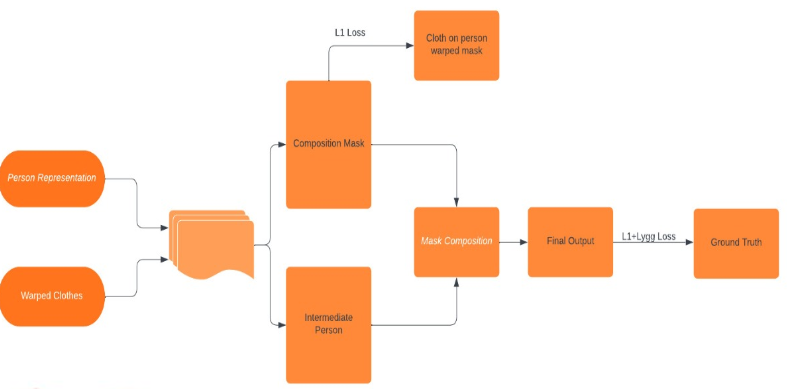
The next process is to superimpose the warped shirt over the person and to adjust the errors which may occur during the imposing process due to mismatch in hand size, overlay of old shirt, hip correction and kind of. Firstly the shirt is made to impose over the person after warping by the process of image fusion by blending as shown in Fig. 1(e). The error adjustments are done by applying the skin matrix over the difference in sample shirt overlaid and the original images. Finally, the virtually dressed person is fused to the background again, which can be the old background or a new one for the aesthetic view of virtual dressing.



Fig 1(e)



Flowchart 1



Flowchart 2

1. **Results And Discussions**

5.1 **Best Case Result**

Fig 2(a) Fig 2(b) Fig 2(c)

The Cloth (Fig 2(b)) is warped around the model (Fig 2(a)) and will produce the final result (Fig 2(c)).

For the above result.:-

|  |  |
| --- | --- |
| IS | SSIM |
| 0.9671116 | 0.93443 |

5.2 **Combined Dataset Result**

Training and testing done on 2032 images and the final average scores are –

|  |  |
| --- | --- |
| IS | SSIM |
| 3.2976 ± 0.5691 | 0.859997 |

5.3 **Drawbacks**

** **

Fig 3(a)Fig 3(b)

Fig 3(b) is distorted after cloth warping around the right hand.

Fig 4(a) Fig 4(b)

In Fig 4(a) , not able to detect whether the subject’s is viewing towards you or away from you, therefore distortion in Fig 4(b)

1. **Conclusion**

We proposed a refined image-based VTON system, 3D-VTON, solving issues in previous approaches: errors in human representation and the dataset, network design , blurriness ,long sleeve dress issues and character is also preserving. Even though 3D-VTON improves the performance, we find that a 2D image-based approach has inherent limitations for coping with diversely posed target input clothing target human warped try-on human cases. Therefore, the application would be limited to simple clothing and standard posed target humans. For more diverse cases, 3D reconstruction of person would be more suitable.

1. **References**

[1] **Srinivasan K., Vivek S** “Implementation Of Virtual Fitting Room Using Image Processing”

[2] **MatiurRahmanMinar1 , Thai Thanh Tuan1 , Heejune Ahn1 , Paul L. Rosin2 , and Yu-Kun Lai2** “CP-VTON+: Clothing Shape and Texture Preserving Image-Based Virtual Try-On”

[3] **Shane Barratt , Rishi Sharma** “A Note On Inception Score”

[4] **Zhou Wang, Member, IEEE, Alan Conrad Bovik, Fellow, IEEE, Hamid Rahim Sheikh, Student Member, IEEE, and Eero P. Simoncelli, Senior Member, IEEE** “ Image Quality Assessment: From Error Visibility to Structural Similarity ”

[5] **Han, X., Wu, Z., Wu, Z., Yu, R., & Davis, L. S. (2018). Viton: An image-based virtual try-on network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7543-7552)** : VITON computes thin-plate spline (TPS) transformation between the in-store clothes mask and the expected background mask.

[6] **Li, Jianan, et al. "Perceptual generative adversarial networks for small object detection." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017**: Introduced a generative model that can generate human parsing

maps and translate them into persons in clothing. But it is not clear how tocontrol the generated fashion items.

[7] **Sekine, Masahiro, et al. "Virtual fitting by single-shot body shape estimation." Int. Conf. on 3D Body Scanning Technologies. Citeseer, 2014:** Introduced a virtual fitting system that adjusts 2D clothingimages to users through inferring their body shapes with depth images.

[8] **Yang, S., Ambert, T., Pan, Z., Wang, K., Yu, L., Berg, T., & Lin, M. C. (2016). Detailed garment recovery from single-view image. arXiv preprint arXiv:1608.01250:** It recovered a 3D mesh of the garment from a single view 2D

image, which is further re-targeted to other human bodies

[9] **Jetchev, Nikolay, and Urs Bergmann. "The conditional analogy gan: Swapping fashion articles on people images."**

**Proceedings of the IEEE International Conference on Computer Vision Workshops.2017:** It proposed a conditional analogy GAN to swap fashion articles. However, during testing, they require the product images of both the target item and the original item on the person, which makes it infeasible in practical scenario.