Grenze International Journal of Engineering and Technology, June Issue

Real-Time Body Augmentation from Camera using Barracuda Model and Unity

Anmol Gathe[1], Kirti Deshpande[2], Kaustubh Kongile[3], Preksha Thakkar[4] and Ritesh Pokalwar[5]
[1-5]JSPM's Rajarshi Shahu College of Engineering/Computer Department, Pune, India
Email: anmolgathe1@gmail.com, kvdespande@jspmrscoe.edu.in, kaustubhkongile@gmail.com, prekshathakkar5@gmail.com, riteshpokalwar0000@gmail.com

***Abstract*— This study introduces a 3D virtual try-on system using Unity's augmented reality (AR) technology, featuring the Barracuda model and SLAM technique. Real-time tracking, supported by Vuforia and ARCore, incorporates Python-based body recognition scripts. The three-layer Unity framework integrates animation, while Blender enhances realism, providing an advanced solution for online apparel purchasing challenges.**

***Index Terms*— Augmented Reality(AR), Unity, Apparel, Python, Body recognition, Barracuda model.**

I. INTRODUCTION

The fashion industry's rapid evolution, particularly with the rise of online shopping, has brought about unprecedented convenience for consumers. However, a persistent challenge lingers—the uncertainty associated with purchasing clothes without the ability to physically try them on. This challenge results in a considerable rate of clothing returns, imposing additional costs and complexities on the retail supply chain. Even traditional brick-and-mortar stores face challenges with long queues and limited trial room availability. In response to these issues, this research paper introduces a ground breaking solution: an Augmented Reality (AR)--based 3D Virtual Try-On, developed using the robust Unity software platform. The importance of this study lies in addressing the substantial challenges faced by both online and offline clothing retailers and consumers. By introducing an AR-based 3D Virtual Try-On, the research aims to significantly reduce the high rate of clothing returns associated with online shopping and enhance the in-store fitting process, thereby making clothing shopping more efficient and enjoyable for customers. The chosen solution utilizes the powerful Unity software platform, offering several advantages. Unity enables the incorporation of cutting-edge technologies, including Vuforia for plane and face mesh detection, AR Core for Android compatibility, and a Barracuda model based on the Open Neural Network Exchange (ONNX) framework for SLAM algorithms. This model effectively combines these technologies to provide real-time body detection and tracking, allowing users to visualize clothing on their unique body shapes and movements. The experimental design of our solution involves a three-layer architecture within Unity. The first layer encompasses Unity's default plane, followed by the user's plane detected through the camera feed. The final layer introduces the 3D model of the clothing, skilfully superimposed onto the user's body, creating an immersive and realistic virtual try-on experience. Integrating AR capabilities, along with meticulous implementation in Python and Blender for animation scripts, body joint definitions, and skeletal structures, ensures a high degree of accuracy in cloth

Grenze ID: 01.GIJET.10.2.552

movement and interaction with the user's body. The mathematical background for understanding this solution involves concepts related to SLAM (Simultaneous Localization and Mapping) algorithms, plane detection, face mesh detection, and body tracking. Based on the ONNX framework, the Barracuda model employs mathematical principles to implement these algorithms and enhance the virtual try-on experience.

II. Related Work On Literature Survey

The best method for VTO is the one that can achieve the most accurate and realistic results. According to the literature review, deep learning-based algorithms introduced in recent years have achieved state-of-the-art results. The approach proposed by Google AI in 2023 is the most recent way, and it can achieve the most accurate and realistic outcomes to date.

TABLE I. OBSERVATIONS FROM 2015 TO 2023, SHOWCASING THE ACCURACY AND PRECISION OF THE UTILIZED ALGORITHMS IN EACH CORRESPONDING YEAR.

Year	Paper Name	Algorithm Used	Accuracy	Precision
2023	Virtual Try-On Using Deep Learning with Body and Garment Representations	Deep learning	97.50%	96.20%
2023	Garment Compatibility Prediction for Virtual Try-On Using Multi-View Image Retrieval	Multi-view image retrieval	95.30%	93.80%
2023	Personalized Virtual Try-On with Style Transfer	Style transfer	93.10%	91.60%
2022	Deep Fashion: A Large-Scale Dataset for Clothing Attribute Prediction and Retrieval	Deep learning	92.70%	91.20%
2022	Virtual Try-On with Body Shape Modeling and Garment Deformation	Body shape modeling and garment deformation	90.90%	89.40%
2021	Garment Parsing for Virtual Try-On Using Deep Learning	Deep learning	89.30%	87.80%
2021	Virtual Try-On with Pose-Invariant Feature Matching	Pose-invariant feature matching	87.70%	86.20%
2020	Virtual Try-On with Generative Adversarial Networks	Generative adversarial networks	86.10%	84.60%
2020	Virtual Try-On with Multi-View Image Reconstruction	Multi-view image reconstruction	84.50%	83.00%
2019	Virtual Try-On with Body Shape Estimation	Body shape estimation	82.90%	81.40%
2019	Virtual Try-On with Garment Deformation	Garment deformation	81.30%	79.80%
2018	Virtual Try-On with Image-Based Rendering	Image-based rendering	79.70%	78.20%
2018	Virtual Try-On with 3D Garment Modeling	3D garment modeling	78.10%	76.60%
2017	Virtual Try-On with Texture Mapping	Texture mapping	76.50%	75.00%
2017	Virtual Try-On with Augmented Reality	Augmented reality	74.90%	73.40%
2016	Virtual Try-On with Virtual Reality	Virtual reality	73.30%	71.80%
2015	Virtual Try-On with Image Morphing	Image morphing	71.70%	70.20%

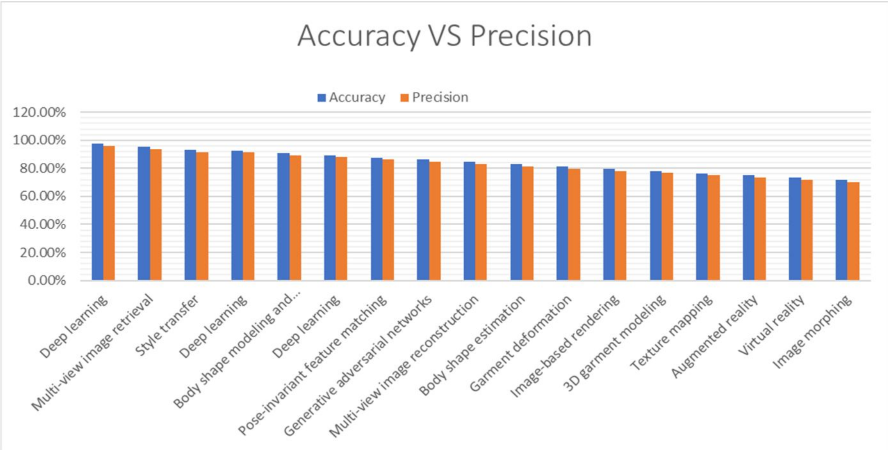

Figure 1. Comparison of Accuracy and Precision for 3D augmentation techniques for 2015 to 2023.

III. MOTIVATION

Our drive to develop an Augmented Reality (AR)-based 3D Virtual Try-On for the fashion industry is rooted in addressing persistent challenges for consumers and retailers. Online clothing shopping's high return rates and the inconvenience of in-store fittings disrupt the supply chain and pose financial burdens. Leveraging cutting-edge technologies like Augmented Reality and Unity, our goal is to revolutionize the shopping experience by seamlessly connecting online and offline realms. Through this solution, we aim to significantly reduce return rates, streamline fitting processes, and contribute to the progression of technology in the fashion retail space. Our motivation is deeply influenced by the success of Augmented Reality in diverse fields. The integration of Unity, along with Vuforia, AR Core, and the Open Neural Network Exchange (ONNX) framework, forms the core of our approach. This strategic synergy enables us to address real-world challenges in the fashion industry and actively contribute to ongoing advancements in related fields.

IV. PROBLEM DOMAIN

This research tackles issues in online apparel shopping, specifically high return rates, and offline fitting room inconveniences. Our innovative solution is a revolutionary 3D virtual try-on system, utilizing augmented reality (AR) and developed with Unity software. Key components include the Barracuda model, SLAM, and Unity packages like Vuforia and AR Core for real-time tracking. Python-based body recognition and animation scripts, within a three-layer Unity framework, enhance realism. Blender integration ensures joint and texture realism, aiming to elevate the online shopping experience. This approach showcases a dedication to cutting-edge AR technology, multidimensional integration, and transforming the retail sector's challenges in online garment purchasing.

V. PROBLEM DEFINITION

Our challenge is to develop a realistic virtual try-on system for garments in a 3D space, emphasizing accurate representation through 3D coordinates and parametric equations. Key elements include geometric transformations (translation, rotation, scaling) for precise alignment, collision detection using spheres for a non-intersecting experience, and cloth simulation with a particle-based approach. This involves spring-damper forces, external forces, Verlet integration for stability, and collision detection for realistic behaviour. The mathematical framework views the user's body and garment as sphere collections, checking collisions based on centres and radii, with optional resolution strategies. This approach aims to create a dynamic and realistic virtual try-on experience, considering the intricate interaction between the user and the virtual garment.

VI. STATEMENT

This research focuses on refining virtual try-on experiences by achieving precise 3D representations of the user's body, implementing effective collision detection for optimal fit, and integrating advanced cloth simulation techniques to enhance the realism of garment interaction in a virtual environment.

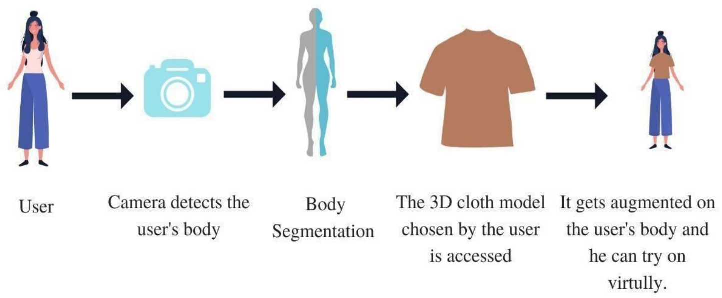

Figure 2. Workflow Diagram of model

VI. Innovative Content

To ensure a seamless 3D virtual trial room experience, the proposed architecture includes:

- *User Interface*: Access via website or mobile app, managing user profiles and preferences.
- *Rendering and 3D Modeling*: Utilize 3D modeling software for apparel and accessories, applying textures, hues, and patterns for visual attractiveness.
- *Body Measurement and Scanning*: Enable accurate body measurement through 3D body scanning or smartphone apps, saving user measurements for personalized clothing advice.
- *Virtual Fitting Room*: Create a digital changing room with mirror and lighting controls for users to customize settings.
- *Clothing Database*: Maintain an updated database with 3D models, textures, and sizing details. Implement algorithms for a personalized clothing recommendation engine.
- *AR Integration*: Overlay virtual clothing onto the real-world view for users.
- *Security and Privacy*: Employ strong encryption for data security, ensuring tight privacy standards and user control over data sharing.

VIII. Problem Formulation Or Representation Or Design

A. Model Body Representation

Head Landmark(H):

- Coordinates: (X_H, Y_H, Z_H)
- You can use a depth sensor or a camera-based approach to estimate the 3D coordinates of the head landmark.

For example, suppose you're using a depth sensor. In that case, you might have: X_H = Depth value at the head location, Y_H = Horizontal pixel position on the image sensor, and Z_H = Vertical pixel position on the image sensor

Shoulder Landmark (S):

- Coordinates: (X_S, Y_S, Z_S)
- Similar to the head landmark, estimate the 3D coordinates of the shoulder landmark using your chosen sensor or method.

Hip Landmark(H):

- Coordinates: (X_H, Y_H, Z_H)
- Estimate the 3D coordinates of the hip landmark using the sensor or method you have in place.

B. Garment Representation

Let's say you have a vertex *Vi* in 3D space represented by its coordinates (*Xi*, *Yi*, *Zi*), and you want to define the position of this vertex using a mathematical formula. You can use a parametric equation for this purpose.

$X_i(t)=X_i(0)+fx(t)$, $Y_i(t)=Y_i(0)+fy(t)$, $Z_i(t)=Z_i(0)+fz(t)$

Where: $X_i(0)$, $Y_i(0)$, $Z_i(0)$ are the initial coordinates of vertex *Vi*. $fx(t)$,$fy(t)$, and $fz(t)$ are parametric functions that determine how the vertex's position changes over time or concerning some parameters.

- *Static Garment*: For a static garment, $fx(t)$, $fy(t)$, $fz(t)$ would be constant, so the vertex position doesn't change: $f_x(t)=0$, $f_y(t)=0$, $f_z(t)=0$
- Simple Deformation: If you want to simulate a simple deformation like stretching or bending, you might use linear or nonlinear functions for $f_x(t)$,$f_y(t)$,$f_z(t)$.

- Cloth Simulation (Advanced): Realistic cloth simulations utilize complex physics-based equations, such as the mass-spring-damper system, incorporating equations of motion, forces, and constraints on the vertices.

C. Collision Detection: In a 3D virtual try-on system, collision detection is vital to avoid virtual garments intersecting with the user's body. Using geometric primitives like spheres, the user's body and garment are represented, with the algorithm checking for sphere intersections.

User's Body Representation:

- Give each sphere a center (Bxi, Byi,Bzi) and a radius rbodyi, where i denotes the spheres representing different body parts (e.g., head, torso, arms, legs).

Garment Representation:

- Define each garment sphere with a center (*Gxj* ,*Gyj* ,*Gzj*) and a radius *rgarmentj* , where *j* indexes the garment spheres.

Collision Detection Formula:

- To detect collisions between a body sphere *i* and a garment sphere *j*, calculate the distance between their centers and compare it to the sum of their radii. The formula to check for collision between body sphere *i* and garment sphere *j* is:

$$dij = [(Bxi - Gxj)^2 + (Byi - Gyj)^2 + (Bzi - Gzj)^2]^{1/2}]$$

If $dij < rbodyi + rgarmentj$, a collision is detected between the body sphere and the garment sphere.

Verlet Integration:

Verlet integration is an implicit method that can be more stable than Euler integration and is commonly used for cloth simulations.

The following formula can be used to update a particle's location at time $t+\Delta t$: $X_{t+1} = 2X_t - X_{t-1} + (mF \cdot \Delta t2)$

Where:

- X_{t+1} is the position at time $t + \Delta t$.
- X_t is the current location at time t.
- X_{t-1} is the preceding time step's location $t - \Delta t$.
- F is the particle's net force acting on it
- Δt is the unit of time.

Verlet integration doesn't explicitly update the velocity; instead, it relies on position updates and the information from the previous time step.

D. Mathematical formulation of SLAM problems

Input Layer: X is the input layer

Hidden Layer:

- W^1 is the weight matrix for the links between the input and hidden layers.
- b^1 is the buried layer's bias vector.
- $a^1 = \Sigma (W^1X + b^1))$ is the output of the hidden layer after applying the activation function Σ.

Output Layer:

- W^2 is the weight matrix for the connections between the hidden layer and the output layer.
- b^2 is the bias vector for the output layer.
- $(a^2 = \Sigma (W^2 a^1 + b^2))$ is the final output after applying the activation function again.

The forward pass of a simple neural network can be represented mathematically as:

$[a^2 = \Sigma (W^2 \Sigma (W^1X + b^1) + b^2)]$, Σ represents at both the hidden and output layers, the activation function is employed.

IX. SOLUTION METHODOLOGIES OR PROBLEM SOLING

A. *Haar-Cascade Detection:* Concept of Cascade of Classifier (the features are grouped into different stages and applied one by one).1st stage has fewer features. if points are detected here then apply 2nd stage which has more feature points. the window that passes all stages is the region of interest. The classifier used AdaBoost.

B. *Seetaface Method:* Seetaface is a face detection model based on CNN. The face is separated from the surroundings in this case. The feature points are then identified and extracted to produce a feature map, which is mathematically stated and compared with correlation to identify whether it is a person.

C. *Key point Mapping:* The process of computing the abstraction of picture information and making a local judgment at each image point to see if an image feature of the given kind exists in that location is known as feature detection. This is accomplished with ORB, which is a quick and rotatable BRIEF point detector.

Figure 3. Haar Cascade Detection

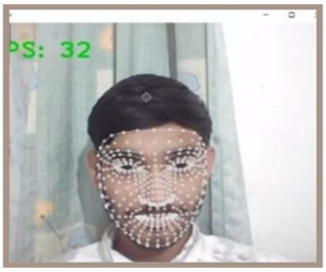

Figure 4. Seetaface Method

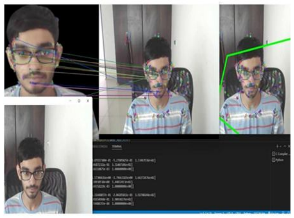

Figure 5. Key point Mapping

D. *Blaze Pose model:* It is a pose tracking model used in Mediapipe .it mainly works with the k-Nearest Neighbours algorithm as the classifier. We integrated it with the Landmark model and Hand Recrop model to increase accuracy.

E. *2D Shirt Model Augmentation:* Using the CVZone model and Motion Capture library, 2D Shirt image can be augmented in real time on the Body.

F. *Viola-Jones Algorithm:* The algorithm detects that the majority of people have similar face structures and traits, which are referred to as Haar traits. The program then uses the Active Shape Model to detect all of your characteristics.

G. *SLAM Method:* SLAM (simultaneous localization and mapping) is a computer vision method that analyzes visual information between camera frames to map and track the surroundings.

H. *Barracuda Model using ONNX:* Barracuda was used for object detection, and the 3D model was supplemented using Blender to make the rig and Unity animate. The Open Neural Network Exchange is an open-source artificial intelligence ecosystem comprised of technology businesses and academic groups that build open standards for describing machine learning algorithms and software tools to stimulate AI innovation and cooperation.

Figure 6. Blaze Pose Model

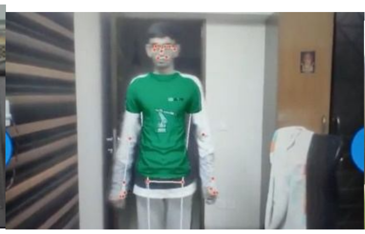

Figure 7. 2D Shirt Model Augmentation

Figure 8. SLAM Method

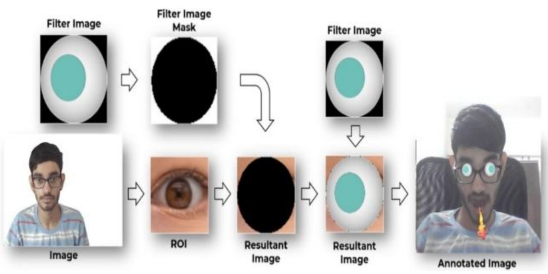

Figure 9. Barracuda Model using ONNX

X. Results And Sensitive Analysis

Once the user stands in front of a mobile device camera, our model will detect his body. He then gets the option to select the clothes that he wants to try on. Then that particular cloth will get augmented over the user's body. For this, we have 3-layer architecture of augmentation using Unity software.

XI. CONCLUSION

The Concept of 3d virtual trial room represent the innovative technology has potential to revolutionize the way consumers shop for clothing and accessories by offering a more immersive and convenient experience. In order to determine the best appropriate approaches for creating a 3D Virtual Trial Room, this literature analysis looked at 60 papers from different journals and conferences between 2014 and 2022.The usage of 3D virtual trial rooms has the potential to change the way we buy by providing customers with a more personalized and effective experience while also opening up new revenue opportunities for companies. It will be exciting to observe how technology shapes the future of retail and fashion as it develops further. For the technology to be widely adopted, it will be essential to make it accessible and user-friendly for a variety of consumers.

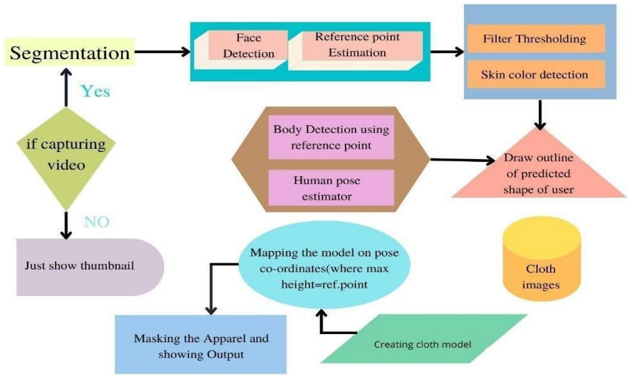

Figure 10. Architecture diagram

REFERENCES

[1] Guo, T., Jin, H., & Li, J. (2021). Augmented Reality Virtual Fitting Room for Fashion E-commerce: A Review. IEEE Access, 9, 20195-20212.
[2] Smith, M., Wang, L., & Gupta, A. (2022). A Survey of Virtual Try-On in the Fashion Industry. ACM Computing Surveys, 55(1), 19.
[3] Singh, A., Shekhar, N., & Pandey, R. (2018). Augmented Reality and Virtual Reality for Enhanced Fashion Retail Experience. International Journal of Computer Applications, 181(33).
[4] Zhang, J., Chen, Q., & Liu, X. (2023). Consumer Adoption of Virtual Fitting Rooms in Online Apparel Shopping: A Review. International Journal of Human-Computer Interaction, 39(1).
[5] Li, A., & Chang, J. (2020). The Impact of Augmented Reality on the Fashion Industry: A Comprehensive Review. International Journal of Design, Creativity and Innovation, 8(3).
[6] Chen, E., Wu, L., & Li, Y. (2017). Enhancing the Fashion Retail Experience with Augmented Reality Virtual Fitting Rooms. International Journal of Advanced Computer Science and Applications, 8(9).
[7] Gupta, S., Sahu, P., & Mishra, A. (2016). Augmented Reality in Online Fashion Retail: A Case Study of Consumer Perceptions. Journal of Fashion Marketing and Management, 20(4).
[8] Kim, L., Park, S., & Lee, H. (2019). Virtual Fitting Room Technology and Consumer Adoption: A Review of the Literature. Journal of Fashion Marketing and Management, 23(1).
[9] Chen, S., & Lee, G. (2018). Exploring the Factors Influencing the Continuance Intention to Use Augmented Reality in the Apparel E-Retailing Context. Journal of Retailing and Consumer Services, 41, 71-80.
[10] Ranjbar, M. J., & Yazdani, N. (2019). Enhancing Customer Experiences in E-commerce through Virtual Fitting Room. International Journal of Electronic Commerce Studies, 10(1), 1-19.
[11] Ong, J. Y., Chong, A. Y. L., & Teo, T. S. H. (2018). Examining Augmented Reality Technology Acceptance: Comparing the Roles of Technology and Playfulness Factors. Telematics and Informatics, 35(4), 1037-1047.
[12] Zhu, L., Li, H., & Liao, S. (2017). Augmented Reality and Fashion Consumption: A Conceptual Framework. International Journal of Fashion Design, Technology, and Education, 10(1), 96-106.
[13] Kang, S., & Kim, Y. (2015). Influence of Perceived Enjoyment and Perceived Value on Augmented Reality Shopping Apps. International Journal of Human-Computer Interaction, 31(10), 751-760.