Tuijin Jishu /Journal of Propulsion Technology
ISSN:1001-4055
Vol. 44 No. 04 (2023)

DressWiz: Virtual Try-on for Dresses using Computer Vision and 3D Modeling

D. I. De Silva[1], M. P. Gunathilake[2], K. O. P. S.Helapalla[3], T. D. R.Mallawaarachchi[4], N. M. N. N.Narasinghe[5], J. R.Munasinghe[6]

[1] [2] [3] [4] [5] [6]Departmentof Computer Science and Software Engineering, Sri Lanka Institute of InformationTechnology, Malabe, Sri Lanka.

Abstract:-The apparel industry currently plays a significant role in the national economy as a dynamic and impactful force. In response to the persistent demand for clothing, physical and online shopping methods have recently emerged to cater to diverse customer needs. During the recent pandemic season online shopping notably became popular. However online shopping presents several drawbacks. One of the significant drawbacks is fitting issues like size and color mismatches when purchasing clothes online. In a technology-driven era, introducing a digital platform to address these fitting issues in online shopping offers an innovative solution to the traditional approach of fitting clothes. The application aims to provide an opportunity for clients to try clothes on an online platform which reduces time wastage, clothing mismatches, and less user satisfaction. A 3D model is used in the application with skin color detection and suggesting matching cloth colors to the user considering the skin tone to showcase the virtual dressing experiences.

Keywords: *virtual dressing room, 3D model, computer vision, virtual fitting, skin tone detection.*

1. Introduction

Fashion is a rapidly changing event with rapid and remarkable growth which has evolved into a dynamic and influential industry. With each passing season, emerging trends, styles, creativity, technological innovations, and rapid changes in consumer preferences, the apparel industry has been undergoing a profound transformation. Due to the never-ending demand and strong preference of the public for clothing, both physical and online shopping emerged as solutions to fulfill customer needs. During the recent pandemic season online shopping notably became popular. The phenomenon caused the limited accessibility for physical clothing stores while encountering safety issues. As a result, the usage of online shopping platforms highlighted a sudden upsurge. However online shopping consists of several issues when considering fitting and matching clothes. To attract them to a technologically immersed world and to make the fashion-trying process convenient for the customer, a digital platform can offer a novel approach to the old-fashioned try-on method [1].

When visiting a clothing store physically, customers have access to a real fit-on experience to select the best matching clothes by selecting preferred colors with the best measurements. They can also have assistance from a salesperson to make decisions considering their body measurements, shape, and skin tone throughout the process of shopping from the beginning to the end. But when using an online platform, some people who lack a fashion sense and who are not able to choose proper clothing items for themselves will face immense difficulties. These obstacles include mainly dissatisfaction with the measurement selections and difficulty in choosing matching colors. It will make the customer dissatisfied with the whole business platform and will undoubtedly have a negative impact on the company's reputation and its standing in the industry.

Most of the online shopping platforms do not provide the facility for an online fit-on experience. Furthermore, most online platforms that facilitate this experience do not have fitting features that are able to satisfy the customer. Due to these reasons, customers think twice before purchasing clothes online [2]. To overcome these challenges a reliable virtual dressing room application that is expected to have high user satisfaction is presented in the following sections of the research paper after utilizing upcoming technology aspects.

The web application addresses many aspects of an online clothing store to overcome the inability to facilitate a better user experience in the fit-on process. In the first stage of the application, the face of the user will be detected using "Tiny Face Detector" [3] which is a pre-trained machine learning algorithm. Then an array of pixels detected by the face area will be sent through an implementation of K-means clustering algorithm to cluster into three slightly different color variations [4] present in the skin and then the three colors will be averaged to a standard color palette. Then in the second stage, "PoseNet" and "BodyPix" machine learning models [5] are used to measure body measurements like shoulders and hip joints utilizing computer vision. A web canvas drawn over the camera feed simultaneously displays detected points and other annotations. Then in the third stage, a customized 3D model will be rendered using the "Three.js" JavaScript library by utilizing the body measurements and the skin tone captured during the early stages of the process. In the final stage, the application allows users to select preferred colors of dresses in different sizes to check whether they match their body measurements and skin tone. To achieve this objective, available clothing colors in the clothing store are filtered by comparing them with a pre-defined set of colors that match different skin tones. The application is built using JavaScript utilizing the ReactJS library.

The rest of the paper discusses the research as follows: Section II demonstrates the literature review of the research area. Section III follows the methodology. Section IV discusses the obtained achievements and test results.

2. Literature Review

A survey was conducted to collect user opinions on common online shopping-related issues. The predominant concern highlighted among the respondents was fitting-related issues in online clothing platforms. Most online shopping stores do not provide a real-time fit-on experience for users. The lack of ability to try clothes before purchasing them has the potential to lead to different problems like size and color mismatches which need to be solved by exchanging the purchased items since they do not have a clear idea about sizes and color combinations when exploring online clothing stores. The primary consideration for most of the users in garment shopping is typically the color. They consider whether the color of the garment matches their skin tone. Colors of the garment items are really important because it is the first impression when looking at someone. Selecting matching colors in online shopping is really hard for some people. The survey results show that the majority of people started using online clothing platforms after the Covid-19 pandemic era. As it is easy and safe for them during the pandemic season. In a physical clothing store, it is possible to select colors easily as users can seek assistance as well as try garments in real-time. Due to the unavailability of fitting rooms in online clothing stores and less assistance in the shopping process, people face many difficulties. So, there is no doubt that online shopping platforms often lack the same level of customer service compared to the traditional way of shopping when analyzing the results.

Through the past years, many researches were conducted to address the aforementioned issues and the final solution was to introduce a Virtual Dressing Room as a software solution to offer users a convenient customer service in the apparel industry. All the research was conducted for the purpose of building a virtual dressing room solution using innovative methods with different technologies. There are two different virtual fitting methods 3D and 2D. In 3D model-based virtual fitting, the approach uses 3D models of the human body and garments to simulate how the garment would fit on the body. This approach can be very accurate and at the same time, it can be computationally expensive and require special hardware and software. In 2D image-based virtual fitting, 2D images of the user and the garment are used to simulate how the garment would fit on the body. This approach is less accurate than a 3D model-based virtual fitting, but it is less computationally expensive and can be implemented on a wide range of devices as well. In some 2D model-based research, 2D projections of multiple shapes and poses generated from 3D body models are used for a realistic experience [6].

In 3D virtual fitting, simulations are done using physical characteristics and parameters for both human and garment models. The Kinect sensor is used to revolutionize the user experience by enabling depth sensing and skeletal tracking methods to acquire 3D geometric objects from the real world. It tracks body dimensions and movements in real time [7]. Foreground extraction and contour extraction is another vital part of 3D modeling. There are two main image processing techniques used in computer vision and image analysis. The techniques are often used in object recognition and image segmentation in computer vision applications mainly to isolate objects from the background in a captured video frame. Techniques vary, from simple thresholding to advanced algorithms like grab-cut and contour extraction. In 2D fitting, contour extraction is crucial to pinpoint body outlines. This uses methods like thresholding, grab-cut, and background subtraction [8].

Having a strong contrast and optimal lighting settings are two critical factors in the realm of 2D image-based virtual fitting. To ensure the precision of virtual fitting, both the customer's images and the clothing items necessitate high contrast and well-balanced lighting conditions [9].

Facecake's Swivel is an existing 3D virtual dressing room solution that applies clothes to the user image in real time. It uses non-flexible 2D models of clothes and does not accurately follow a person's curves and movements [17]. LazyLazy is another platform for trying clothes virtually using real-time customer images. In here also clothes are depicted as images in the center of the screen. Customers can adjust the size of clothing item images to achieve a closer fit based on their body size and distance to the camera.

In [10], the focus was on garment modeling, which involves the creation of virtual clothing designs using standardized body measurements. The process began by generating 2D garment patterns using splines. These patterns were then wrapped around a virtual human body to establish the initial garment shape. To make the virtual clothing behave realistically, a simulation was performed by applying physical parameters derived from actual fabric characteristics. Once the garment was created, a real-time web-based platform was established for seamless interaction and presentation of the virtual clothing designs.

In conclusion, while online clothing shopping offers undeniable advantages in a technology-driven era, several obstacles like the lack of ability to replicate the in-store fitting experiences continue to hinder the online shopping experience. From sizing issues and color mismatches to accurate representation, the virtual dressing room concept explored by researchers has the main goal of bridging the gap between the user-friendly and convenient environment of online shopping and the personalized experience of physical stores, enhancing customer satisfaction in the online fashion industry as well as the efficiency in online shopping platforms by attracting customers.

3. Methodology

Before starting the application, the camera should be properly positioned in a place where there is no harsh lighting directly aimed at it. For the application to function properly, the background should be plain and free from distractions (i.e., human figures other than the user).

Fig. 1. shows a high-level system overview diagram. Each of the major components has been explained below.

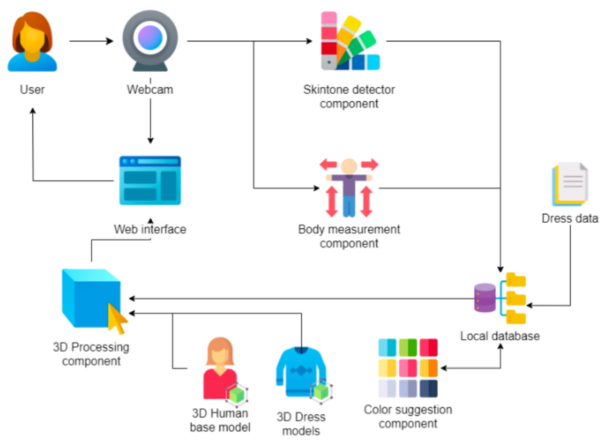


Fig. 1. System overview diagram

3.1 Identifying the skin color

The application consists of a module that identifies the prominent skin color of the user and suggests matching dress colors out of available colors for a particular dress. As the first step of the process, the skin color should be identified using the device camera. This has been achieved with the use of computer vision along with the K-means clustering algorithm.

First, the user is asked to appear in front of the camera. Then the application looks for a face in the camera feed as shown in Fig. 2. Once a face is detected, the pixels from the detected face area are extracted and returned as an array. From this, a "skin score" is calculated and the skin pixels are separated [11]. This array contains data of each pixel in the detected area with red, green, and blue values. This array of pixels is later sent through the K-means clustering algorithm to extract the most prominent color, which in turn is used as the skin color of the user.

The face detection is achieved by a pre-trained machine learning model called "Tiny face detector" a lightweight and high-speed computer vision algorithm model tailored for devices with limited computational resources. It is designed to identify faces in different aspect ratios in image or video frames. Further, this model provides the coordinates of the points where the face is detected in the camera feed. This feature is leveraged to isolate the skin area of the face rather than using the entire video frames to find the prominent color, which might be inaccurate when a part of a dress or the background colors are visible on the camera feed [12] [3].

An earlier solution for this problem was to have a region of interest in the center area of the camera feed and ask the user to hold their hand or any other area that has exposed skin. Another alternative idea was to use a machine learning model that is trained to detect and separate human skin from an image.

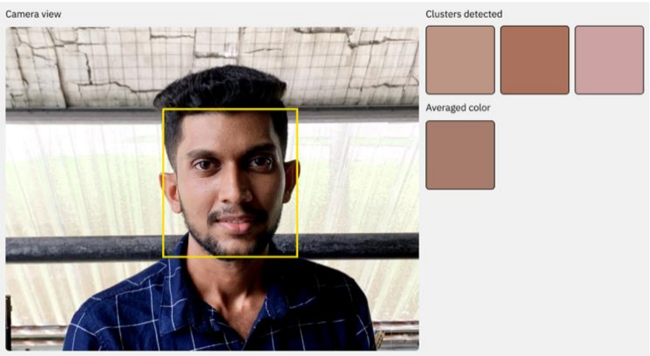


Fig. 2. Detecting face and skin tone

To identify the prominent skin color of the user, several methods and techniques are combined. Initially, the pixel area extracted by the previous component will be processed using the K-means clustering algorithm to obtain three clusters of colors. This function will look for three slightly different variations of skin colors and will cluster them according to common skin tones. Then these three colors will be averaged to one simple color tone, which is the standard color palette used for skin colors [13].

Skin colors, when considered within a set of people, differ vastly due to slight changes in complexion. A skin tone palette called Monk Skin Tones (MST) is used here (Fig. 3). The common mathematical formula Euclidean Distance (1) is used to determine the color that is closest to an existing color in the MST palette [14].

$$d = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2}$$

(1)

The colors are separated into three channels Red, Green, and Blue (RGB) with each channel ranging from values 0-255. These channels are compared with the rest of the colors and averaged. Once this is done, the color gets saved in the browser's local storage to be used in the later steps.

Fig. 3. Monk Skin Tones (MST) palette

To suggest the matching dress colors to a certain skin tone, a matching algorithm is written using fuzzy logic with the help of a dataset of potential dress colors.

A more efficient method would be to use a machine learning model specifically trained to identify whether a color matches when given the two colors. A model of such scope needs to be trained with a large set of data which needs to be gathered by doing user research on matching colors for skin color tones.

3.2 Detecting body measurements

For the 3D model of the user to be rendered accurately, several measurements should be taken only using the device camera, without special distance measuring sensors (i.e., Kinect imaging sensor, LiDAR sensor). To achieve this, a methodology was developed with the help of several machine learning models that run on top of the JavaScript library TensorFlow.js, an adaptation of the popular machine learning platform to run on web browsers. The machine learning models used are "PoseNet" [5] and "BodyPix" [15]. The body measuring component consists of a camera feed and a web canvas drawn over the feed to simultaneously display any point/annotation necessary. For ease, the full process will be divided into separate sections and explained.

The user is asked to stand within a fixed distance from the camera. This value is pre-calibrated so that the cartesian coordinates can be translated using a scale to obtain the actual distance between two points.

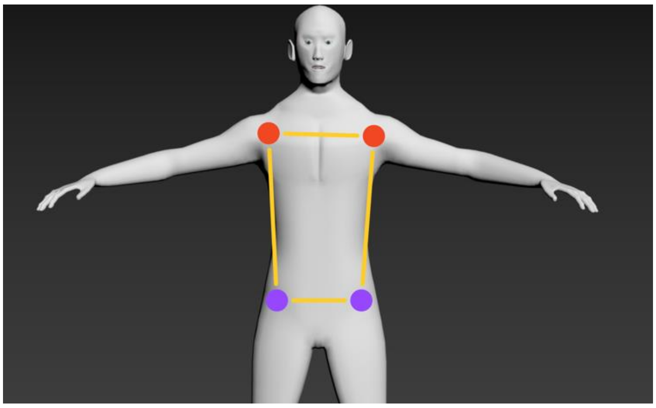


Fig. 4. Shoulder and hip joints detected by PoseNet

3.2.1 Detecting shoulder and hip joints

The camera feed is sent to be processed through the machine learning model "PoseNet" [5] to return the coordinates of the joints in the body that are visible in the camera feed as shown in Fig. 4, and the necessary data are filtered and stored as an object. Later, these data are displayed on the web canvas directly over the camera feed, such that the points are visible on the exact positions of the image. In this application, the displayed points are the shoulder joints and the hip joints.

These points can be used to obtain the distance between the joints, but that does not accurately reflect the shoulder and hip length of the user since the pose estimation only provides the coordinates of the bone joints without the width/thickness of the muscles. There should be a way to add muscle thickness to this measure. For this, a different approach was used.

The straight lines passing through each of these sets of coordinates are obtained using the straight-line equation (2). These straight lines are then displayed on the canvas using inbuilt functions. For example, the straight line passing through the shoulder points of the figure will be displayed on the canvas. The same goes for the hip joints.

$$y = mx + b$$

(2)

The machine learning model "body-pix" is used to perform human body segmentation [15] on the images passed in. This model works with TensorFlow and returns a masked area in which the different segments exist. (i.e., face, left and right arms, etc.), which is then displayed over the webcam video feed via a JS canvas. The returned object is an array of pixels where the color value and the opacity values are included.

The mask area is saved to a pixel array to be accessed for later calculations. This contains the color values of each pixel of the mask. The next part is finding the coordinates of the points that intersect with the edges of this mask. With this, the accurate distance between the shoulders and hips can be obtained rather than joints. To achieve this, first, the points that lie inside the mask which also belong to the straight line should be found. For this, each pixel of the pixel array needs to be iterated and checked if that point satisfies the equation of the relevant straight line. For example, the points that the straight line passing through the shoulder points that belong to the masked area can be retrieved with this. The image mask contains data in the form of hexadecimal values, and these are converted to RGB values to extract the alpha value which refers to the transparency. To check if a certain pixel is in the mask, the alpha value is checked. If the alpha value is greater than 0.5 it is considered as a point inside the mask (since the mask is also transparent with an opacity of 0.5). Once all the values inside the mask are retrieved, the farthest two points can be obtained with a simple mathematical formula. Which then gives the farthest two coordinates that lie on the masked area that also lies on the straight line. These two points can be used to determine the distance between the exact shoulders and the waist.

The height of the model will be measured by considering only the distance between the two lines joining the shoulder joints and hip joints (since only upper body height is required). For now, arm width will be calculated based on a generic constraint, but a possible way to do this would be to consider the midpoint of the line joining the shoulder joint and the elbow joint, drawing a line perpendicular to that, and obtaining the points that the line intersects with the mask obtained from segmentation [16].

Further, bust length, chest length, and abdomen area can be measured using a similar way. To obtain depth, a similar method can be implemented, and the user could be asked to provide a side view of the body to the camera feed.

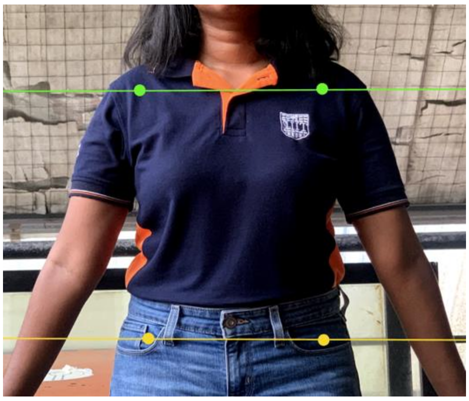


Fig. 5. Detected shoulder and hip joints

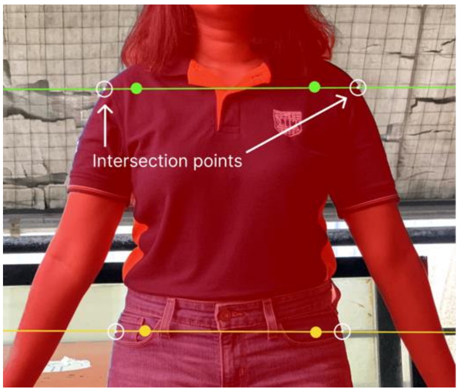


Fig. 6. Intersection points of straight lines joining shoulder & hip joints with the segmentation mask

3.3 Rendering the 3D model

To render the 3D model, a JavaScript 3D library "Three.js" is used. With this, 3D models can be rendered and manipulated easily on web browsers without having to rely on resource-heavy platforms like Unity, Blender or Unreal Engine.

This part of the application has several components that work together to show whether the selected dress fits the user or not. The 3D model chosen is a base model for now, and later other specific features can be added as improvements. The goal is to render a 3D model that resembles the user's body size, in order to check if a dress fits or not. For the dress, a 3D model is rendered over the base human body model such that the dress is wrapped around the body.

Once the measurements are obtained from the body measurement component, they are passed to the body rendering component to be displayed. As of now, the only accepted measurements are the shoulder length and waist length. With these measures, the 3D model will be adjusted accordingly as shown in Fig. 7.& Fig. 8. The 3D model is displayed in a particular scale such that one unit distance in pixels actually displays a real-world measuring unit such as centimeters.

To resize the base 3D model into the required size, a technique called vertex deformation is used. Since, 3D objects are built using vertices which are linked, such that it forms a mesh, and the areas between these vertices are filled with planar objects to achieve opaqueness. To adjust the size of such a part, using several vertices would not be sufficient since it would make the body rough. (When the position of a single vertex changes in the XYZ plane, that vertex specifically stands out.) To have a smooth effect, the vertex deformation technique [16] is used. Once a vertex is moved in the XYZ plane, the adjacent 3D mesh is deformed proportionally and smoothly.

This method is applied on both the shoulder areas, and in the waist area, so that once the measurement is provided, it is passed into the deform vertices function and it will perform the deformation on the 3D model. The measurements obtained from the body measuring component will be converted to the scale on which the 3D model is rendered so that the model can be displayed accurately and proportionately [16].

Simultaneously, the skin color identified during the first step will be applied to the 3D model to provide a realistic look. To achieve a more realistic look, a patch from the skin could be extracted as a bitmap and then applied as a texture for the skin area of the 3D model. This provides a realistic representation of the skin. The texture can be applied using a method in Three.js which creates a material. If the skin texture bitmap is not provided it will only use the color.

The dresses have pre-defined sizes and for each of them, a separate model that is sized to a scale is included at the time of adding a dress. The system user interface allows the user to pick any of those and apply them to the human 3D model. The user can rotate or zoom the 3D model and see how the dress size matches the body as shown in Fig. 9.

The user can select various colors of the dress which have been configured previously and try different options on the 3D model. The user can see which colors match and do not match with the help of the color suggestion component. Once the user selects an available color for a dress, if it is a match, then a matching value (true or false) is shown by the side.

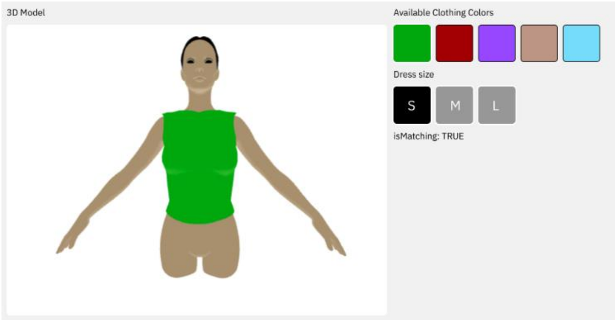


Fig. 7. View of the 3D model with the dress applied on

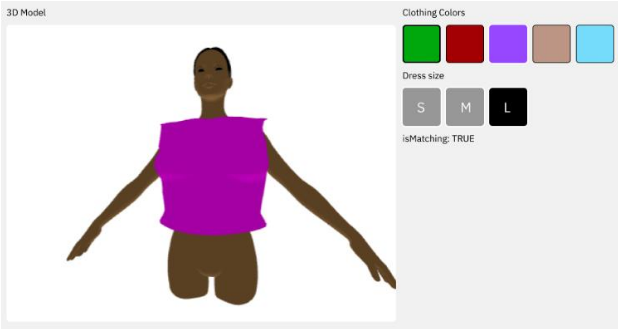


Fig. 8. Trying on different dress sizes

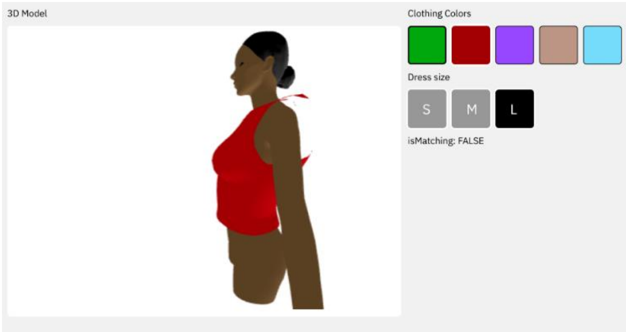


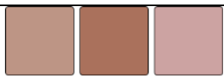
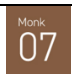
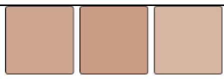
Fig. 9. Rotated view of the 3D model

4. Results and Conclusions

The following results and conclusions were made during the process of developing this application. These values and results can be utilized to understand if the application can be used to get accurate results.

The skin tone detection component first outputs three clusters identified from the extracted skin pixels. The Table 1 shows the skin patch, the identified clusters and the Monk Skin Tone matched to it.

Table1. Detected Three Clustersand Suggested Skin Tone

Skin patch	**Clusters**	**Suggested Monk Skin Tone**
		
		

The measurements obtained from the body measurement component rely on several factors like the quality of the camera, the lighting conditions, and distractions in the background. Table2 shows some results obtained from this component.

Table2. Actual Values and Measured Values

Measurement	**Real measurement**	**Detected measurement**
Shoulders	35cm	32cm
Waist	40cm	38cm

Since the beginning of time, the apparel industry has played a major role all over the world. Traditional purchasing methods entailed physically going to stores to pick out and buy desired clothing. However, the busy lifestyles and time restraints of modern society have spurred a move to online purchasing platforms. Online activities became crucial during situations like the pandemic, which hastened this shift even more.

References

[1] Weerasinghe S.W.P.N.M., Rajapaksha R.M.D.D., Sathsara L.G.I., Gunasekara H.S.D.N, Dinuka R. Wijendra, Dilshan I. De Silva, "Virtual Dressing Room: Smart Approach to Select and Buy Clothes," in *Advancements in Computing (ICAC)*, Colombo, Sri Lanka, 2022.

[2] Edbert Junus, Clarissa Jocelyn, Daffa IrsyadAdilah and Alexander Agung Santoso Gunawan, "Application Fitting Room using Virtual Technology," in *IEEE Creative Communication and Innovative Technology (ICCIT)*, IEEE International Conference, 2022.

[3] Peiyun Hu, Deva Ramanan, "Finding Tiny Faces," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017.

[4] HongleiZhouf, Bo Zhang, Jipeng Zhai, "Clothing colour clustering method based on CC_K-Means algorithm," in *2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI)*, Kunming, China, 2022.

[5] Kosei Yamao, Ryosuke Kubota, "Development of Human Pose Recognition System by Using Raspberry Pi and PoseNet Model," in *2021 20th International Symposium on Communications and Information Technologies (ISCIT)*, Tottori, Japan, 2021.

[6] P. Guan, O. Freifeld, M. J. Black, "A 2D human body model dressed in eigen clothing," in *Proceedings of the 11th European Conference on Computer Vision*, Heraklion, Crete, Greece, 2010.

[7] J. Tong, J. Zhou, L. Liu, Z. Pan and H. Yan, "Scanning 3D full human bodies using kinects," *IEEE Transactions on Visualization and Computer Graphics (Proc. of IEEE Virtual Reality)*, vol. 18, pp. 643-650, 2012.

[8] Srinivasan K., Porkumaran K., Sainarayanan G., "Background Subtraction Techniques for Human body segmentation in Indoor video surveillance," *Journal of Scientific and Industrial Research*, vol. 73, pp. 342-345, 2014.

[9] Frédéric Cordier, WonSook Lee, HyeWon Seo, Nadia Magnenat-Thalmann, "From 2D Photos of Yourself to Virtual Try-on Dress on the Web," in *People and Computers XV - Interaction without Frontiers*, London, 2001.

[10] Dimitris Protopsaltou, Christiane Luible, Marlene Arevalo, Nadia Magnenat-Thalmann, "A body and Garment Creation Method for an Internet Based Virtual Fitting Room," in *Advances in Modelling, Animation and Rendering*, London, 2002.

[11] Pratheepan Yogarajah, Joan Condell, Joan Condell, Abbas Cheddad, Paul McKevitt, "A dynamic threshold approach for skin segmentation in color images," in *2010 IEEE International Conference on Image Processing*, Hong Kong, China, 2010.

[12] Cha Zhang, Zhengyou Zhang, "*A survey of recent advances in face detection*," Microsoft Research, Redmond, WA, 2010.

[13] Thilan Costa, Vincent Gaudet, Edward R. Vrscay, Zhou Wang, "Perceptual Colour Difference Uniformity in High Dynamic Range and Wide Colour Gamut," in *IEEE International Conference on Image Processing (ICIP)*, Abu Dhabi, United Arab Emirates, 2020.

[14] Ellis P. Monk, Jr, "*The Monk Skin Tone Scale*," SocArXiv, 2023.

[15] Zul Herri, Yulianta Siregar, Fahmi Fahmi, "Person Segmentation in Human Height Measurement Using Bodypix," in *6th International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)*, Medan, Indonesia, 2022.

[16] Ting Liu, LingZhi Li, XiWen Zhang, "Real-time 3D virtual dressing based on users' skeletons," in *4th International Conference on Systems and Informatics (ICSAI)*, Hangzhou, China, 2017.