Virtual Cloth Try-On Using Augmented Reality - Marker Based Approach

Prajakta Joglekar
Dr. Vishwanath Karad MIT World Peace University
Vinaya Gohokar (vinaya.gohokar@mitwpu.edu.in)
Dr. Vishwanath Karad MIT World Peace University

Research Article

Keywords: Feature, Augmented Reality, Rendering, descriptors, Marker, Homography

Posted Date: August 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1918747/v1

Abstract

The Virtual cloth Try-on is one of the biggest inventions took place in fashion industry which contributes to enhance user experience by allowing them to try out garments virtually without wearing it. Researchers are working on several technologies right from Image processing to Augmented reality (AR) to develop and deploy a stable, sustainable platform to enable virtual cloth try-on experience. In proposed system Augmented reality technique is used to create the same experience. With the help of OpenCV and Python programming, Marker based AR is designed for four types of 2D image garment datasets such as Dress, Top, Jeans and Skirt.

1 Introduction

Technology has been growing rapidly in many industries in recent years, remarkably in the clothing sector, which attempts to fulfill customer needs and expectations. One of these requirements is that garments be tried on before being purchased. Retail e-commerce sales worldwide hit 4.28 trillion dollars in 2020, with e-commerce revenues expected to reach 5.4 trillion dollars in 2022. When it comes to fashion, one of the most significant offline experiences that on-line buyers overlook is the changing room, where a garment item can be tried on. Additionally, Pandemic time has changed the way of living. From attending virtual work meetings to buying favorite apparel online everyone is getting more inclined towards taking outmost advantages of technologies. Even though online shopping is not a big thing in today's culture, there are some customers who still opt for visiting a store, trying out an outfit and then buy it. Currently, most of the shopping malls, apparel shops are having trial rooms for customers to try out apparels before buying it. In pandemic time these trial rooms are restricted for customers to maintain no contact policy and hygiene. Trying out cloths in these trial rooms could sometime hamper users' privacy and lastly, providing trial rooms is difficult task for small shop keepers as far as space is concerned. Proposed idea can tackle these issues. This system allows user to try clothes virtually without physically wearing it. In this, real time human body detection and tracking to get user's body image and Augmented Reality (AR) techniques to super impose cloth image on previously obtained user's body image is used. Augmented Reality is one of the prominent technologies comes under big Industry 4.0 umbrella. These Reality techniques escalates user experience in variety of fields by developing virtual world as per required application. In Augmented reality, by superimposing visual phenomena such as image, video or text on the real time object, virtual world can be created for user through which user can perceive the experience without physically building the environment.

From past decades researchers are working to achieve stable platform for virtual try-on for Apparel market as per consumer requirement. Some of those major research milestones are discussed in related research. In system overview, broad comparison between Marker based and Marker less AR approach is specified which will give clarity behind choosing marker-based AR approach as a baseline for the proposed system. Next, workflow of proposed system including intermediate stages like Video frame capture, recognizing image target surface using feature detection, feature descriptors and feature matching and 2D Image Rendering on target image surface is discussed with the help of block diagram.

At the end of this chapter results are shown. Lastly, issue and challenges faced while developing this system are discussed.

2 Related Research

Researchers used various techniques to develop Virtual try-on systems such as Image Segmentation, Image warping and in Augmented reality also, using 2D images, 3D images, Markers, and various S DKs the research has been done. Image Segmentation in which developer can take user image as an input, divide it into require body shape on which cloth image can be added. For this segmentation task a parser, an element of a preprocessor that divides into little components. Also, research has been done in parser-free algorithms where developer takes pretrained decomposed data as an input image and then Garment image can be added to desired body part [2][3][4]. In image warping, Developer can use perspective transformation, wrap transformation to add image on the user's required body part. All though all these techniques use Image processing as baseline, advancement can be done with the use of deep learning and neural network [5]. Many researchers use Augmented Reality technique as a baseline to develop virtual try-on system. AR can be classified based on Image type which being used for AR such as 2D, 3D images. In 2D Image type again Marker-based and Marker-less are two techniques which has been discussed. In next chapter methodology behind all these techniques is discussed in detail [6][7].

3 System Overview

3.1 Marker Based Vs Marker Less Approach

For proposed system, 2D image is used as an input data for AR with the help of Marker based approach. Basically, in 2D image-based AR technique there are two approaches are invented namely, Marker-based approach and Marker-less approach. Marker is data input that might be in terms of an image or a RFID tag or a QR/Barcode which helps developer to focus on which particular area in plane, the image has to be augmented. The marker-based AR technique involves positioning an image target next to the camera and adding it with a new visual model. Feature identification, feature description, and feature matching, homography, projection, and placing a object into the scene are

Table 1. Comparison between Marker-based and Marker-less approach [1]

Parameters		Marker-based approach	Marker-less approach
Methods	Relative Position/ angle	Depends on markers	Depends on localization technology and gyroscope
	SDK	Commonly used	Rarely used
Position accuracy	High/low	Relatively higher	Relatively lower
	Influence factors	Brightness	Localization technology
Stability	High/low	Relatively lower	Relatively higher
	Influence factors	Marker clarity	Geographical co-ordinates
Licensing		Not required	Required to access exclusive Libraries
Device support	Computer	Supported	Usually not supported
	Mobile	Supported	Supported

all data whereas Marker-less approach uses techniques such as SLAM (Simultaneous Location and Mapping), Multiplayer rendering, co-ordinate tracking. Marker-less Approach is nothing but creation of an App with the help of AR development SDKs where geographical data fetched from sensors like gyroscope, accelerometer, magnetometer, GPS is used for detection of targeted co-ordinates Comparison between these two approaches is given in Table. 1

Based on the given comparison developer can select which approach can be efficient for the specific application. In virtual try-on position accuracy is necessary to impose garment image on user's body. Also using licensed SDKs may affect future development of this project due to environment incompatibility. To avoid this dependency, Marker-based approach is used to develop proposed system. For implementation, OpenCV and Python programming language is used and Jupyter notebook is used for execution. Workflow is discussed next.

3.2 Workflow

As shown in Fig. 1 proposed system works in 3 different stages:

A. Video frame capture:

Since we are using marker-based approach for image rendering, capturing accurate real time video frame of user is necessary as algorithm will recognize marker (user image) from captured frame. Device camera can be enable using OpenCV algorithm and frame is captured as shown in Fig 2 (A). If some error occurs while enabling the camera input, system will check for errors and restart.

B. Recognizing image target surface using feature detection, feature descriptors and feature matching

To augment required garment image on user body frame, system needs to capture marker from live video frame, which can be find out using feature extraction and feature matching technique. As shown in Fig. 2, there will be complete user body image priorly saved in system. once system enables camera and captures video frame, system will try to detect user body and will try to match feature points similar to priorly saved user body image. Here, user body image will work as marker for rendering the image, discussed in next step. In computer vision feature is referred as piece of information captured from visual data for processing or representing characteristics. These features can be key point of change in intensity, gradient, or grayscale. While detecting features each of these parameters is set up to a certain threshold value and if the captured value goes beyond threshold level the feature is detected. There are several renowned algorithms developed for feature detection. In proposed system ORB feature detection algorithm is used. The letters ORB stands for Oriented FAST and Rotated BRIEF. It is faster and less expensive to compute. Unlike SIFT and SURF, it is not patented. ORB uses the BRIEF description and a revised version of the FAST key point detector. The features of FAST are not scale- or rotation-invariant. As a result, ORB uses a multi-scale pyramid to achieve scale invariance. A multiscale pyramid is constructed of many layers, including one that contains a downscaled version of the image from the preceding layer. ORB recognizes characteristics at all layers and scales. The shift in intensities around each key point determines its orientation (left or right). As a result, ORB is a rotation invariant too. Table 2 shows feature points detected using ORB feature detection. In Table 2 generic results of feature detection using ORB algorithm is shown. If results of feature detection techniques discussed above is compared, one can clearly rely on ORB feature detection technique. Since Harris corner detection and Shi-Tomasi gives insufficient feature points for further processing (Feature matching and Image Augmentation) which can affect accuracy on the other hand FAST feature detection algorithm provides additional feature points including many false one which may hamper accuracy.

After moving forward with ORB algorithm passable feature points are detected which will be feed to Feature Matching algorithm discussed next. After extracting features from the reference image next task is to match those features with input image (marker). These feature points will help us compare and detect objects in photos in image processing. consider Fig 2 (B). The goal is to extract key points from the left-hand image, which we refer to as the target image. For this, feature matching techniques can be used to identify all of the matches in the image on the right. There are several feature matching techniques are available. In proposed system FLANN (Fast Library for Approximate Nearest Neighbors) based feature matching is used. Fast Library for Approximate Nearest Neighbors (FLANN) is an acronym for Fast Library for Approximate Nearest Neighbors. It includes a series of techniques that are optimized for rapid nearest neighbor searches in huge datasets and high-dimensional characteristics. For large datasets, it is faster than Brute Force Matcher. FLANN based feature matcher allows two sets of arguments that specify the algorithm to be used, as well as its associated parameters and so on.

The first is Index, which displays the variables of the algorithm in use, and the second is Search. It defines the number of times the index's trees to be explored recursively. Higher values provide more precision, but they also take longer time for execution. The Hamming distance is used to match binary descriptors. For binary strings, this distance is comparable to counting the range of unique elements (population count after an XOR operation):

$$d_{hamming}(a,b) = \sum_{i=0}^{n-1} a_i \oplus b_i \qquad (1)$$

where, a and b are two targeted key points.

When the distance ratio dhamming as shown in Eq.(1) between the two closest matches of a selected key point is less than a threshold, it is a good match.

From the shown output image, one can easily choose FLANN-

based feature matcher as it is giving sufficient matched feature points from real-time image. After finding matched feature points those matched can be fed to image augmentation algorithm, discussed next.

C. 2D Image Rendering on target image surface:

For image augmentation researchers have been using Homography matrix from past many years. Homography is known as the mapping between two planar projections of an image.

In a nutshell, planar homography is the relationship between the transformations of two plan4es (up to a scale factor):

$$s\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = H\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

As it is approximated up to a scale, the homography matrix is a 3x3 matrix with 8 DoF (degrees of freedom) [8].

h33=1 or

h211+h212+h213+h221+h222+h223+h231+h232+h233=1. (2)

Two homographs must be employed if two points are not on the identical plane. Similarly, one must utilize n number of homographs for n planes. Using homography matrix, perspective transform can be applied on the image to wrap outfit on human body. For specific set of garments dataset, correct size of marker should be used for example if we want to try-on Top or T-shirt on woman’s body marker of woman’s upper body part should be used. Similarly, if we want to try-on Skirt or Jeans on woman’s body, we must use marker of woman's lower body part. Cloth image will get rendered accurately only when correct marker is used. Table 3 shows how efficiency and error in image rendering changes with the use of incorrect and correct marker.

Using correct marker for various dataset will give virtual try-on results. If you see in column 3 of Table 3, there are some virtual try-on which are giving incorrect results since incorrect marker is used for specified garment.

Results are shown in Table. 7, Table. 8 and Table. 9

4 Issues And Challenges

Despite the fact that smartphones are becoming more powerful, they are still not very compatible in terms of processing speed when compared to a PC environment. As a result, using some high-end complex terms in this system is difficult because it will damage speed and accuracy. Second, because Augmented Reality is still a relatively new technology, development resources, documentation, and tools are limited.

5 Future Scope

In future, the same system can be used for various body sizes from XXS to XXL. By changing some parameters in feature detection algorithm same method can be used for Men garment virtual try-on. Techniques like deep learning with blend of AR can be used for non-apparel try-on like makeup, jewelry, handbag, etc.

6 Conclusion

Introducing Augmented reality system for trying on virtual 2D clothes that allows users to see themselves wearing virtual clothes while looking at the mirror display without taking off their actual clothes. Users can select and try on a variety of virtual clothing. The system physically simulates the selected virtual clothing on the user's body in real time, allowing the user to see the virtual clothing sitting in the mirror image from different angles while on the move. Here new methodology for performing virtual try-ons based on images of people is presented. Image rendering technology is used to map clothing images to individual body images from simple inputs such as photographs, and automatically implements customized virtual fitting of clothing by humans. It has been shown that the number of feature points is the most important factor in the image distortion algorithm to achieve an acceptable try-on effect. The optimum number of feature points was obtained through experiments. From the application of virtual fitting automation to humans, the results show that the proposed approach can effectively perform virtual fitting in real time using images of humans. Compared to other virtual fitting techniques, our approach is not only feasible and practical, but also simple and inexpensive. The successful application of image warping not only provides a new way of image processing, but also greatly helps the fashion industry in improving customer experimentation.

References

1. Chao Lin, Zhao Li, Sheng Zhou, Shichang Hu, Jialun Zhang, Linhao Luo, Yuan He (2022) RMGN: A Regional Mask Guided Network for Parser-free Virtual Try-on
2. Cheng JCP, Chen K, Chen W (2017) Comparison of Marker-Based and Markerless AR: A Case Study of An Indoor Decoration System. In: Lean and Computing in Construction Congress - Volume 1: Proceedings of the Joint Conference on Computing in Construction. Heriot-Watt University, Heraklion, Crete, Greece, pp 483–490
3. Ge Y, Song Y, Zhang R, Ge C, Liu W, Luo P (2021) Parser-Free Virtual Try-on via Distilling Appearance Flows. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Nashville, TN, USA, pp 8481–8489
4. Grubert J, Langlotz T, Zollmann S, Regenbrecht H (2017) Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. IEEE Trans Visual Comput Graphics 23(6):1706–1724. https://doi.org/10.1109/TVCG.2016.2543720
5. Han X, Wu Z, Wu Z, Yu R, Davis LS (2018) VITON: An Image-Based Virtual Try-on Network. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, Salt Lake City, UT, pp 7543–7552
6. Liang X, Gong K, Shen X, Lin L (2019) Look into Person: Joint Body Parsing & Pose Estimation Network and a New Benchmark. IEEE Trans Pattern Anal Mach Intell 41(4):871–885. https://doi.org/10.1109/TPAMI.2018.2820063
7. San Günes, Okan Sanli, Övgü Öztürk Ergün (2015) Augmented Reality Tool for Markerless Virtual Try-on around Human Arm. In: 2015 IEEE International Symposium on Mixed and Augmented Reality - Media, Art, Social Science, Humanities and Design. IEEE
8. Sekhavat YA (2017) Privacy Preserving Cloth Try-On Using Mobile Augmented Reality. IEEE Trans Multimedia 19(5):1041–1049. https://doi.org/10.1109/TMM.2016.2639380
9. Sen H, Yi-Zhe S, Tao X Style-Based Global Appearance Flow for Virtual Try-On
10. Zeng X, Ding Y, Shao S Applying Image Warping Technique to Implement Real-Time Virtual Try-on Based on Person's 2D Image. Information Science and Engineering, International Symposium on
11. OpenCV: Basic concepts of the homography explained with code. https://docs.opencv.org/4.x/d9/dab/tutorial_homography.html?msclkid=7aa4545dbe3c11ec81daa60f19f020a1. Accessed 16 Jun 2022

Tables

Tables 2-6 are available in the supplementary files section. Tables 7-9 not available with this version.

Figures

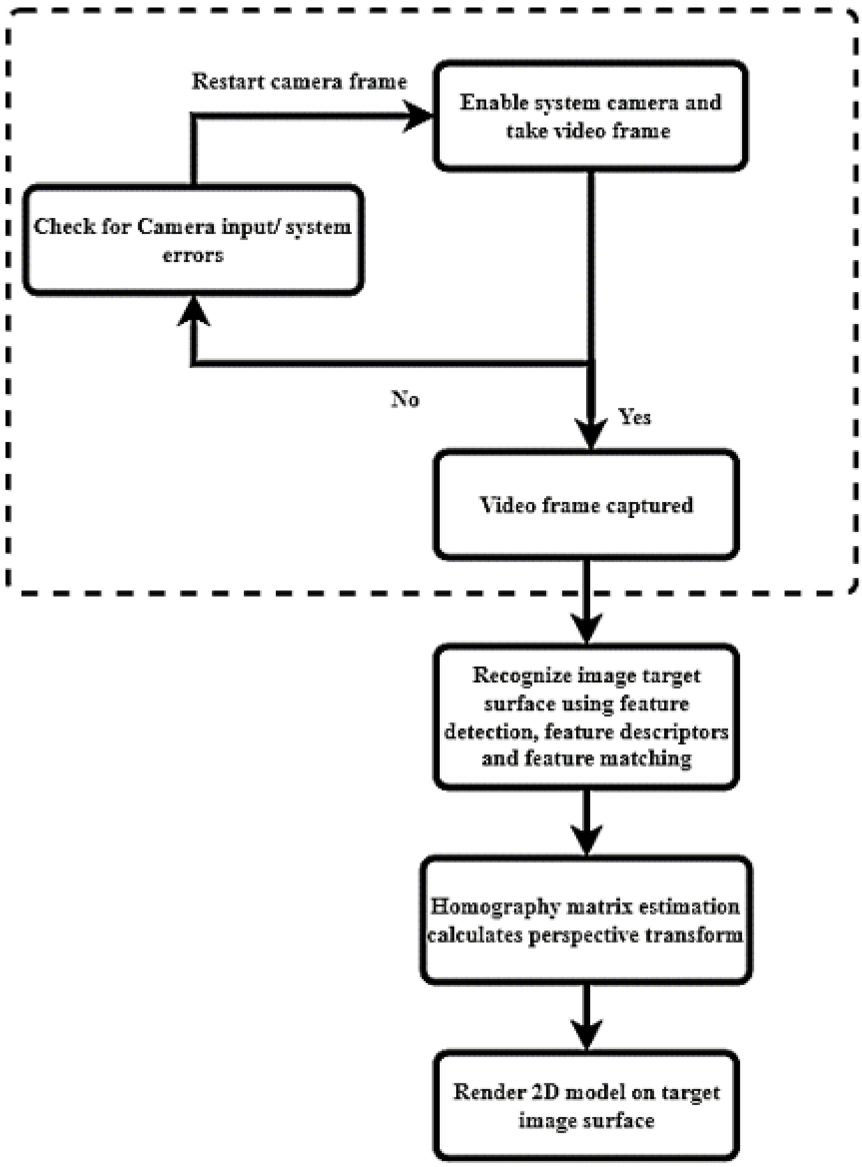

Figure 1

Workflow

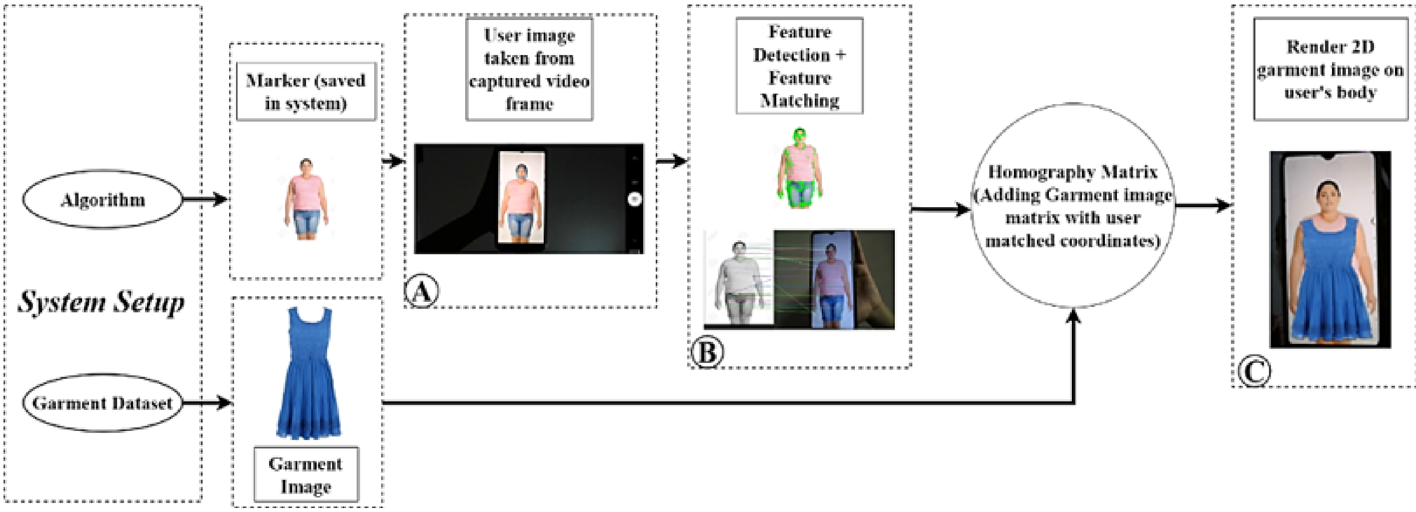

Figure 2

System Architecture

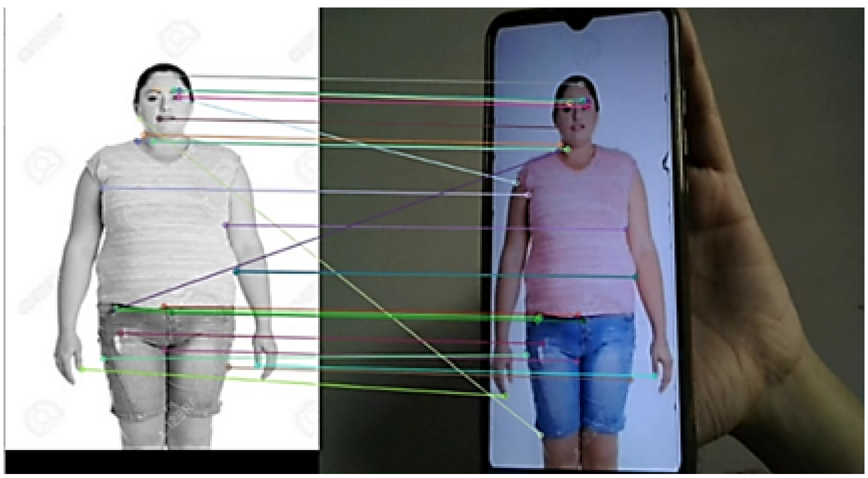

Figure 3

FLANN feature matching result 1

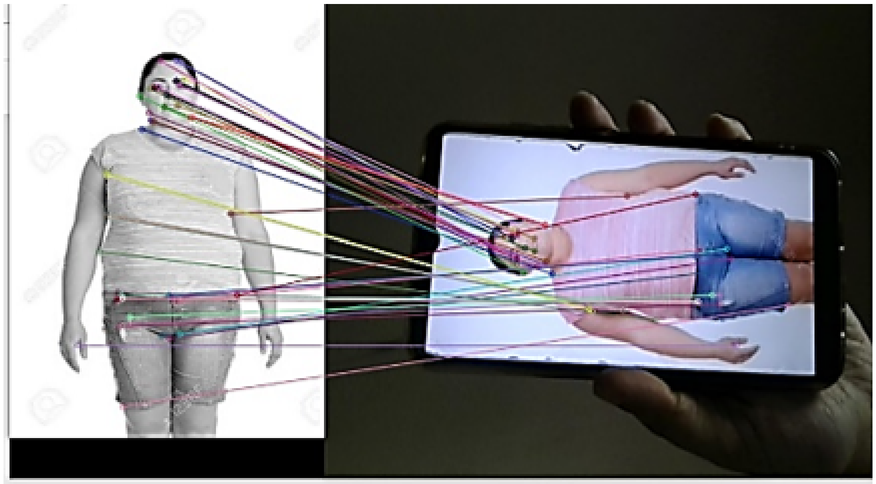

Figure 4

FLANN feature matching result 1- tilted

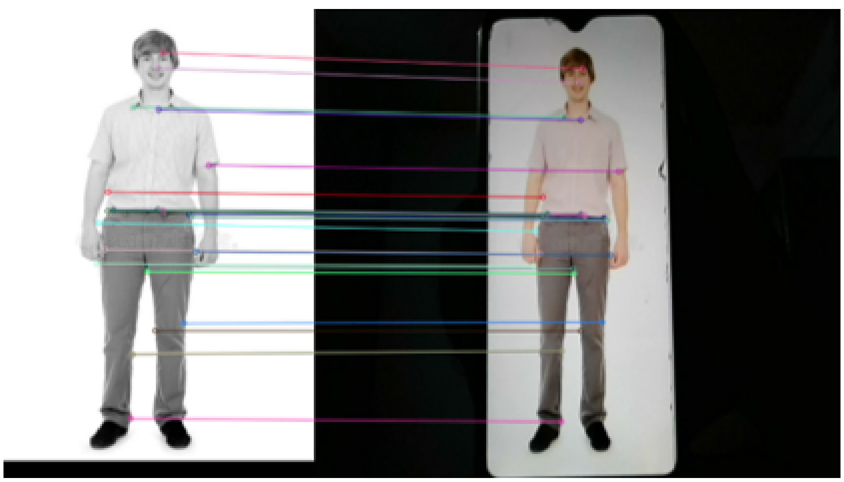

Figure 5

FLANN feature matching result 2

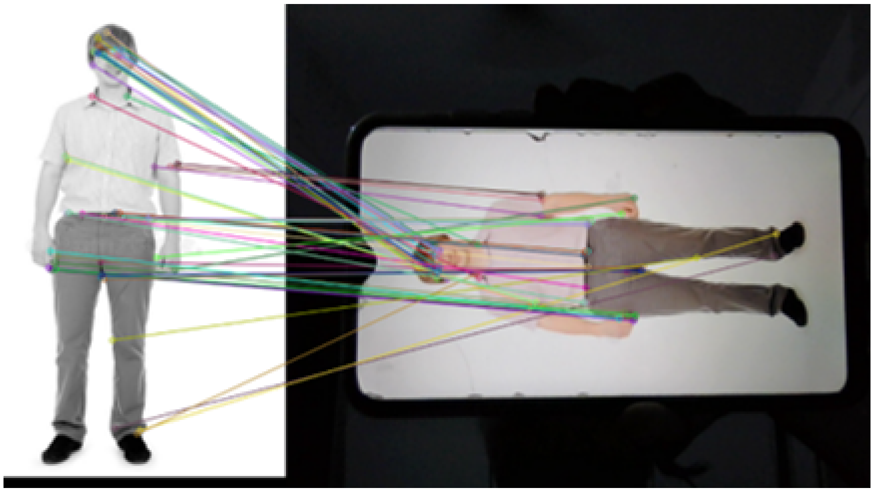

Figure 6

FLANN feature matching result 2- tilted

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tables.docx