A Hybrid Photorealistic Architecture Based on Generating Facial Features and Body Reshaping for Virtual Try-on Applications

Tran Van Duc[1], Pham Quang Tien[1], Hoang Duc Minh Trieu[1], Nguyen Thi Ngoc Anh[2], Nguyen Tien Dat[1,✉]

[1]Viettel High Technology Industries Corporation, Hanoi, Vietnam
[2]VNU University of Engineering and Technology, Hanoi, Vietnam

anhnguyenngoc@vnu.edu.vn, datnt65@viettel.com.vn[✉]

Abstract

Online shopping using virtual try-on technology is becoming popular and widely used for digital transformation because of sustainably sourced materials and enhancing customers' experience. For practical applicability, the process is required for two main factors: (1) accuracy and reliability, and (2) the processing time. To meet the above requirements, we propose a state-of-the-art technique for generating a user's visualization of model costumes using only a single user portrait and basic anthropometrics. To start, this research would summarize different methods of most virtual try-on clothes approaches, including (1) Interactive simulation between the 3D models, and (2) 2D Photorealistic Generation. In spite of successfully creating the visualization and feasibility, these approaches have to face issues of their efficiency and performance. Furthermore, the complexity of input requirements and the users' experiments are leading to difficulties in practical application and future scalability. In this regard, our study combines (1) a head-swapping technique using a face alignment model for determining, segmenting, and swapping heads with only a pair of a source and a target image as inputs (2) a photorealistic body reshape pipeline for direct resizing user visualization, and (3) an adaptive skin color models for changing user's skin, which ensures remaining the face structure and natural. The proposed technique was evaluated quantitatively and qualitatively using three types of datasets which include: (1) VoxCeleb2, (2) Datasets from Viettel collection, and (3) Users Testing to demonstrate its feasibility and efficiency when used in real-world applications.

1 Introduction

Due to the global impact of Coronavirus disease (Covid-19) and environmental pollution, the fashion industry is changing quickly to adopt the requirement in driving consumers and sustainable fashion. Offline retailers and department stores had already seen massive declines in sales and met huge pressure in inventory during the lockdown period. As a result, Online purchase is widely booming and numerous apparel companies are leveraging digital transformation to boost their sales. However, it is not easy for consumers to see and feel their actual appearance if they can not physically try products. Generally, there are two main approaches to creating a reality and immersive shopping experience: (1) Using 3D Interactive Simulation, (2) Using 2D Photorealistic Generation. The first method produces 3D models with collisions in virtual space. It could generate an acceptable shape and interact with stable performance, but it requires the preparation for input requirements and the ability to scalable both in horizontal and vertical

directions. For the second approach, the speed and input requirement are the main strong points, but, this method meets a big issue in controlling outcomes, which prevents it from being widely adopted around the world. In general, most methods face three main challenges: (1) The complexity to process and manage the input resource data, (2) the accuracy and reality of virtual try-on with clothes, and (3) Maintaining the system's high stability and reliability.

To address these issues, two start-up companies: Zeekit[1] and 3D Look[2] have been developed with their patented technology. Zeekit was founded by Yael Vizel, Nir Appleboim, and Alon Kristal in 2014. Zeekit divides a person's image into thousands of parts using its patented real-time image processing technology. A similar process is applied to clothing, and the analogous points from the two simulations are re-mapped to create the final simulation. The final product depicts a virtual representation of a person wearing clothing

---

[1]https://zeekit.me/
[2]https://3dlook.ai/

that takes into account their body type, fit, and fabric. Meanwhile, the body shape is changed by 3D look algorithms' creation of a geometrically accurate virtual try-on. The segmentation map of the subject is altered by the semantic generation module, which then warps the clothing mask to cover the appropriate portion of the body.

The main contributions of this research are listed below: (1) Proposing a unique and fully automatic architecture system applied for virtual try-on clothes that combines facial feature models with body reshaping, (2) A state-of-the-art technique is presented with high accuracy, stable and ease to integrate and deploy in a cloud server. This research is divided into four main parts: Section 2 summarizes previous works on virtual try-on clothes; the proposed method is provided in Section 3; Section 4 contains a detailed description of the dataset used for evaluating and results in quantitative and qualitative to prove the method's efficiency; section 5 concludes the paper and opens future works.

2 Related work

Recent research has shown a significant improvement in reconstructing 3D human models based on multiple approaches. In this part, two main methods including (1) Interactive simulation between the 3D models, and (2) 2D Photorealistic Generation are introduced and analyzed in terms of their upsides and downsides.

2.1 3D interactive simulation

2.1.1 Physics-based simulation

A physics-based simulation of clothing basically involves three main processes: Calculation of internal forces, collision detection, and collision reaction forces; the total simulation cost is the result of combining the effects of the three processes. Bender et al.[1] provides a survey to analyze the trade-off between accuracy and cost simulation. Approximation methods for video game applications are reasonable, but they cannot transfer the realistic material behavior required for virtual fitting. Another approach by Fratarcangeli et al. [6] is to improve the efficiency of computing forces and collision constraints using GPU-based while maintaining the simulation accuracy. These approaches provide very well-controllable results with a physics-based model. However, the high computational cost and the performance are big fences for application and scalability.

2.1.2 Network-based

Li et al. [8] introduces a mesh-based network, named N-cloth. The main purpose of the method is to predict the 3D Cloth Deformation between cloth mesh and the target obstacle mesh for various scenes. The main advantages are an end-to-end neural network, high performance (30-45 fps with up to 100.000 triangles), and acceptable accuracy. This work inherits, and extends the encoder-decoder architecture [11] and creates a fusion network to deal with the deformation mesh from the input clothes and multi-type of obstacle meshes for improving the overall accuracy. However, preprocessing training data using a physics-based simulator is a time-consuming procedure. Besides, the training process must be re-done if the mesh topology of clothes is changed.

Santesteban et al. [14] proposes a learning-based clothing animation with two major steps: (1) Preprocess Data: using a physics-based simulation to generate multiple animated human bodies wearing the same garment, (2) Process: Combining a global fit to a fixed body shape with local garment wrinkles. This research provides an efficient way to generate cloth simulation with high speed but struggles for processing input data and remaining completeness of clothes.

2.2 2D Photorealistic Generation

Xiang et al. [20] presents a photorealistic and animated clothing approach by using neural rendering to physically-simulated garments. There are three modules in this method: (1) Base body avatar: predicts body geometry and texture, (2) Cloth simulator: simulates and creates clothing deformation from body geometry, (3) Clothing appearance model: predicts photorealistic clothing texture. This approach is much better than prior work, however, there are still issues in dealing with: (1) data process: The clothing appearance model requires real captured data with the registered body, and clothing geometry (2) the clothes with losing form or multi-layer clothing.

Yang et al. [22] proposes a method named: Adaptive Content Generating and Preserving Network (ACGPN). It generates a photorealistic try-on composing of three main parts: (1) Semantic Generation: separate target clothing region and segments the body parts, (2) Clothes Warping: warps the target clothing image to predicted segmentation layout, (3) Content Fusion: Renders the results by utilizing the target clothing image, original clothing mask and body part image. Outcomes are acceptable, but ACGPN has trouble with various kinds of human poses and handles the texture of clothing input.

Another approach is proposed by Sarkar et al. [16], called: StylePoseGan for generating a photorealistic multi-view of users from a single input image with explicit control over pose and per-body-part appearance. This method, which is basically an end-to-end model trained with image reconstruction loss and adversarial loss, applies multi-purpose such as garment transfer, head swap, and image interpolation. Due to its architecture, this method has a difficulty in reconstructing images/text in the clothes.

DeepFaceLab (DFL) [10] focuses on photorealistic face-swapping using the fundamentals of GAN and StylGan. It would consist of three main phases: (1) Extraction: use face detection, alignment, and segmentation to extract a face from given data (2) Training:
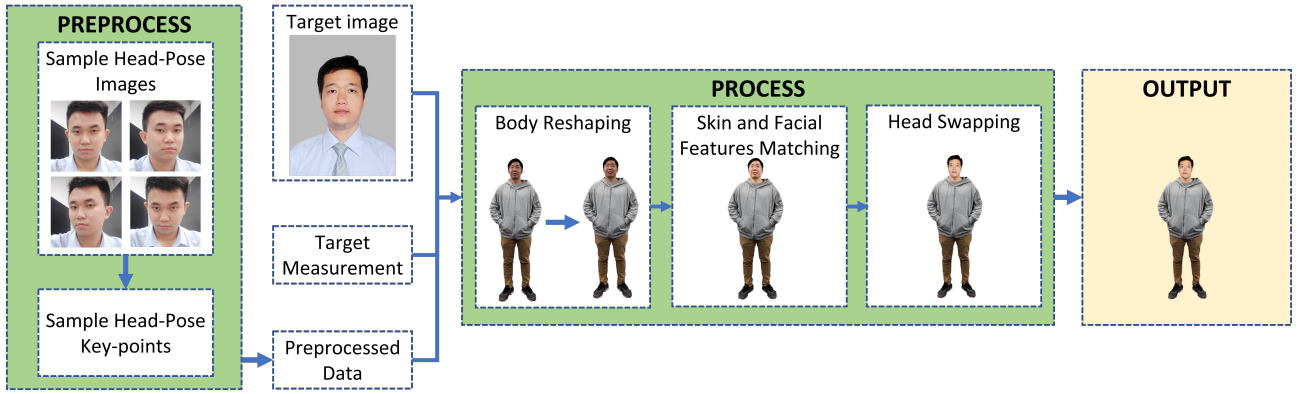
Figure 1: System diagram of the proposed method.

A structure named LIAE is proposed to achieve the generalization of src and dst (abbreviation of source and destination), through the shared Encoder and Inter and inherit fully information such as lighting (3) Conversion: An Encoder is used to blend and sharp for face generation and re-alignment. These results are quite well in the wild; however, only face regions are swapped and hairstyle, skin tones, and face shapes do not remain.

Shu *et al.* [17] presents a method named: Head Swapper (HeSer) with the main contribution is to produce photorealistic head swapping on different scenes using two parts: (1) Head2Head aligner: is used to migrate expression and position from target to source head, (2) Head2Scene Blender: is used to custom and edit the facial skin color and background texture unmatches. Qualitative and quantitative results prove its efficient approach, but it only focuses on head-swapping and does not mention body re-shape. Furthermore, there are several issues during processing with long-hair users and huge demand for input data.

3   Methodology

3.1   Overview

The system diagram of the proposed method is illustrated in Fig. 1. Our framework begins with data pre-processing, which produces two database: Sample Head-Pose Images and Sample Head-Pose Key-points. In the main process, the target image and measurement are operated with the pre-processed data through three main modules: body reshaping, skin and facial feature matching, and head swapping. While the body reshaping module resize the source image body to fit the target measurement, the skin and facial feature matching module take the target image facial properties to apply to the source image. Finally, after the source body and facial features are visually suitable with the target, the head swapping module is executed to place the target head on the right acceptable of source body's neck.
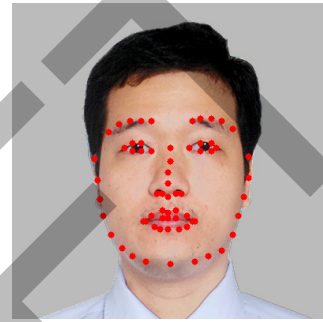

Figure 2: 68 face landmarks of the Sample Head-Pose Key points.

3.2   Pre-Process

The preprocess module generates a database for changing the head pose algorithm in the head-swap module: Sample Head-Pose Images and Sample Head-Pose Key-points. Sample Head-Pose Images contain multi-head images of one person that cover almost the natural position of the head. To ensure the reality of the change head pose process, the rotation angle of the head in three axes is limited in range $[-45°, 45°]$. And for the sake of simplicity, the rotation angle is uniformly sampled in a limited range with a $5°$ division. The Sample Head-Pose Key-points are 68 face landmarks extracted from Sample Head-Pose Images as shown on Fig. 2.

The landmarks are normalized based on the area of the human face. In order to reduce facial distortion in photos captured by short focal length lens cameras, Sample Head-Pose Images are utilized to change the users' faces from "near pose" to "far pose". Sample Head-Pose Images are split into two sets: near and far depending on the capture distance from the camera to the user's face, which is defined as 30cm and 90cm. For each "near" image, there is a head-angle corresponding "far" image and the image pair having the closest head angle to the user's face is used for the pose-changing task with DaGAN [7] implemented.
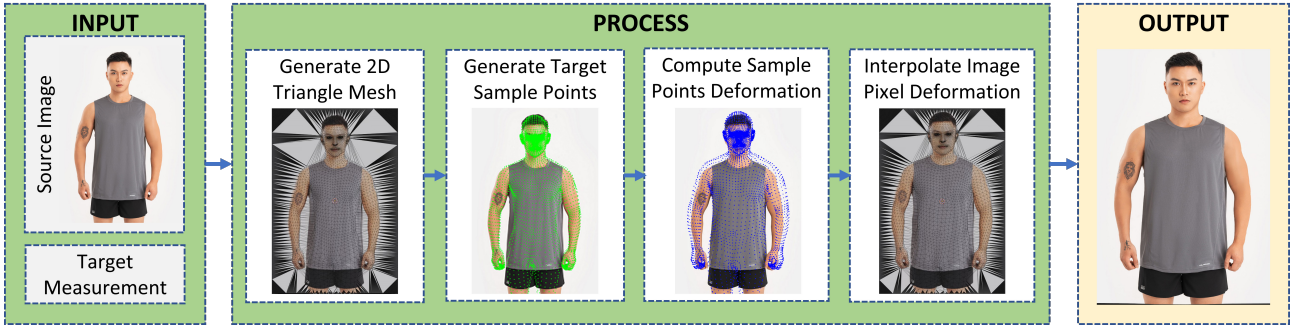
Figure 3: System diagram of the Body Reshape module.

3.3 Body Reshape

3.3.1 Generate 2D Triangle Mesh

Firstly, a 3D human parametric model is fitted to human images. The fitting process follows SMPLify-X by Pavlakos *et al.*[9]. The following is a short introduction, and more details are described in [9]. SMPL-X model is fitted to the human body in an image using 2D key points and an iterative optimization approach. The total objective is:

$$
\begin{aligned}
E(\theta, \beta, \psi) = {} & E_{\mathrm{J}}\left(\theta, \beta, K, J_{\mathrm{est}}\right) + \lambda_{\theta_b} E_{\theta_b} + \lambda_{\theta_f} E_{\theta_f} \\
& + \lambda_{m_h} E_{m_h} + \lambda_\alpha E_\alpha + \lambda_\beta E_\beta + \lambda_\varepsilon E_\varepsilon \quad (1) \\
& + \lambda_{\mathrm{weight}} E_{\mathrm{weight}} + \lambda_{\mathrm{height}} E_{\mathrm{height}}
\end{aligned}
$$

Where:

- $\lambda_{\theta_b}, \lambda_{\theta_f}, \lambda_{m_h}, \lambda_\alpha, \lambda_\beta, \lambda_{\mathrm{weight}}, \lambda_{\mathrm{height}}$ are scalar weights.

- θ, β, ψ are the pose, shape, and facial expressions of SMPL-X model.

- θ_b, θ_f, m_h are the pose vectors for the body, face, and two hands, respectively.

- K, J_{est} are the camera parameters and 2D detected key-points.

- E_J penalty 2D euclidean distance between projected SMPL-X joints and $J_{est}.E_{m_h}, E_{\theta_f}, E_\beta, E_\varepsilon$ are simple L2 priors for the hand pose, facial pose, body shape, and facial expressions, penalizing deviation from the neutral state.

- E_α is a simple prior penalizing extreme bending only for elbows and knees. E_{θ_b} is a VAE-based body pose prior. The interpenetration error term is removed due to making the fitting process very slow and having a little contribution to fitting performance.

- $E_{\mathrm{weight}}, E_{\mathrm{height}}$ are added to the total objective function to penalize the difference between the predicted and real weight, and height of a human.

When the optimization process is done, the 3D human model with optimized parameters is projected to 2D sample source points. To ensure the smoothness of the image warping process, the 2D sample source points only take visible points which are detected by using Ray Casting [12] method. Then, a 2D triangle mesh is attached to the image to control image warping. The 2D triangle mesh is constructed based on 2D sample source points and the image bounding box using Delaunay triangulation [13].

3.3.2 Generate target sample points

A simple model is constructed to predict SMPL-X shape parameters from height and weight in this step. The dataset for training is generated by registered SMPL-X model to Vietnam human scan data (Viettel dataset). This dataset includes more than 900 samples with different gender, occupation, region, pose and body mass index to ensure the diversity of the population in Vietnam.

The SMPL-X model has two important properties. Firstly, body vertices of SMPL-X have a linear relation with shape space because SMPL-X model is constructed using Principal Component Analysis (PCA). Secondly, the SMPL-X model explores the relationship between the weight and volume of 900 registered sample data, and the linear relation between them is shown in Fig. 4: the more heavy the body is, the more massive the body volume has. This linear relation between weight and volume, combined with the linear relation of body vertices of SMPL-X and shape space come to a conclusion that the relationship between height, weight, and shape space of SMPL-X model is assumed to be linear. Therefore, a simple linear regressor is utilized for the training process. The 2D target sample points are obtained by using projections like 2D source sample points.

3.3.3 Compute sample points deformation

The deformations of sample points are represented as affine transformation matrices obtained by optimizing the objective function below:

$$
E_A = \omega_p E_p\left(J_S, J_T\right) + \omega_s E_S \quad (2)
$$

Where:

- ω_p, ω_s are scalar weights.

- $E_p\left(J_m, J_u\right) = \|J_T - J_s\|_2$ penalty 2D euclidean distance between 2D sample target points J_T and source points $J_S.E_S = \sum_i \sum_j \|A_i - A_j\|$ the term
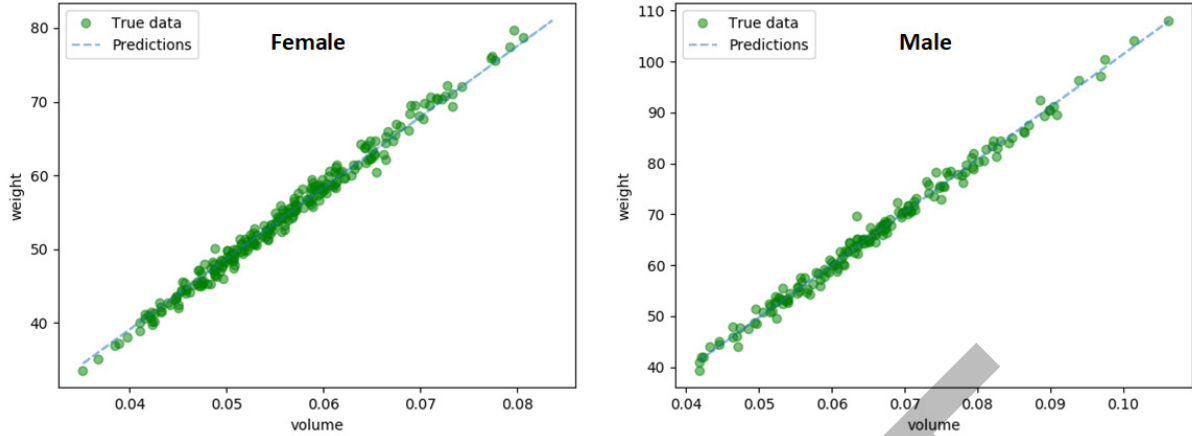
Figure 4: The linear relationship between weight and volume.

enforces smooth surface deformation of 2D triangle mesh.

- A_i, A_j are affine transformation matrices of two vertices on an edge of triangles.

The final result after applying image pixels deformation is shown in Fig. 5. For convenience, the image pixel deformations are pre-computed for a sample height, and weight. And the body in the image can be changed in real-time by scaling the pixel deformations via target height, and weight.

3.4 Head Swapping

The diagram of the HeadSwap module is illustrated in Fig. 6. The preprocess data, combined with source and target images, is utilized as the input data and processed in the HeadSwap module. The input data is processed in five stages: (1) Key point Extraction, (2) Best-frame Finding, (3) Animated Head, (4) Head Removal, and (5) Facial Landmark Transformation.

Two sets of 68 facial key points corresponding to preprocess data are collected from the source image and target image by implementing 2D Face Alignment [2] in the key point Extraction stage. In total, there are three key point sets belonging to the source image, target image, and preprocess data.

Sample Head-Pose Images are generated from multiple face images with the same identity, various poses, and expressions with the aim of representing the majority of human facial poses. The more diverse the Sample Head-Pose Images are, the more accurate the reference key points detection task is. In the next stage, these three key point sets are normalized by pupillary distance. Two key point sets from the source image and target image are compared in sequence with Sample Head-Pose key points to find the most identical frame in Sample Head-Pose Images.

This technique performs improvement in creating animated faces compared to other methods [7][3][24]. Animated Head images are created with only one
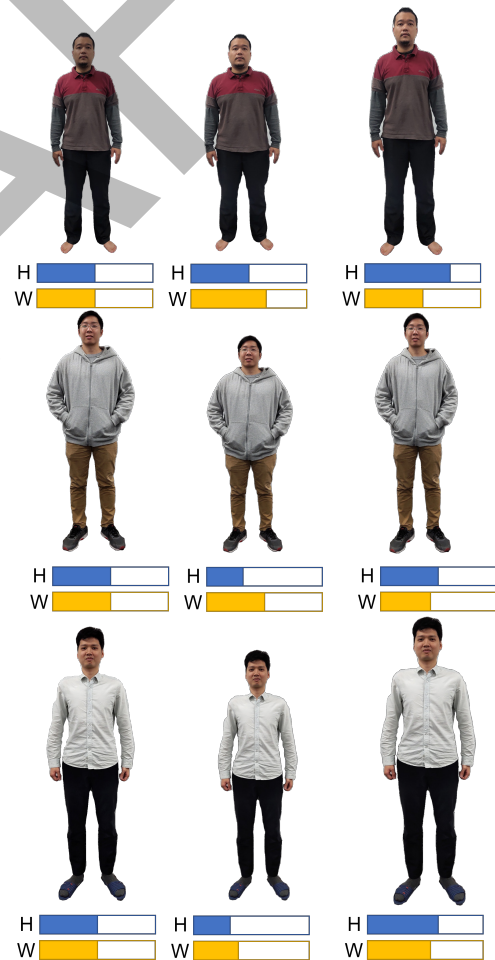

Figure 5: Results of body reshaping process with H and W indicate height and weight, respectively.
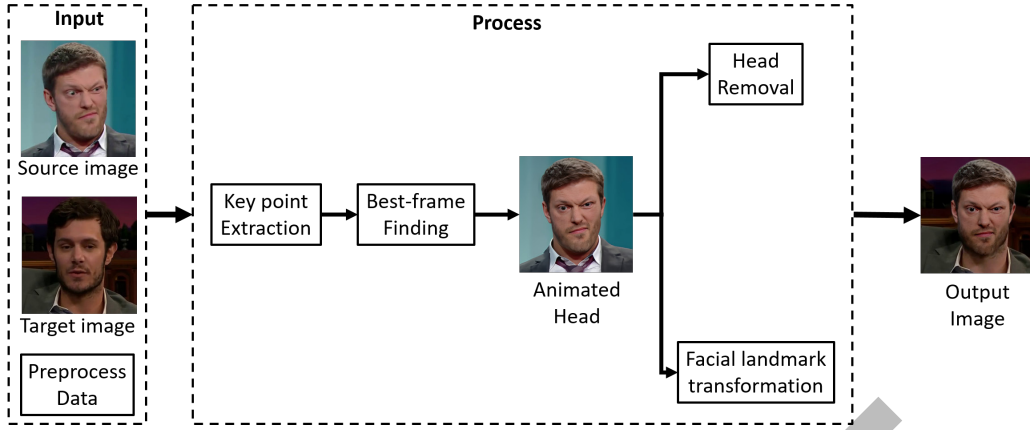
Figure 6: Diagram of the HeadSwap module.

source image and one destination image without the need for a source video or a few shot images and a target image. Therefore, the processing time is shortened with fewer images or frames.

DaGan [7] is integrated into the Animated Head stage with the inputs including two sets of key points and the source image and the output is the animated source head with the angles of the target head. The Animated Image and target image are fed into a segmentation network to detect the head region in each image. The head region in the target image is removed and filled with colors of the surrounding region to avoid blank space before being replaced with the head region in the source image.

In the facial landmark transformation stage, the transformation matrix is calculated from the key points obtained in the key point Extraction stage and the image distortion algorithm is utilized to transform the image of the user's face to match the model's face in the image. The matrix is calculated based on the spatial matrix transformations as follows:

$$M = \begin{bmatrix} \text{scale.} \ R & c_u^T - \text{scale.} \ R.c_m^T \\ 0 & 1 \end{bmatrix} \quad (3)$$

$$R = \left(U.(\text{V})^T\right)^T \quad (4)$$

$$U, \Sigma, (\text{V})^T = \text{SVD}\left((\text{lmk}_m)^T . \text{lmk}_u\right) \quad (5)$$

$$c_u = \text{mean}\,(\text{lmk}_u) \quad (6)$$

$$c_m = \text{mean}\,(lmk_m) \quad (7)$$

$$\|lmk_m\| = \frac{\text{lm}\,k_m - c_m}{\text{std}\,(lmk_m - c_m)} \quad (8)$$

$$\|lmk_u\| = \frac{\text{lm}\,k_u - c_u}{\text{std}\,(lmk_u - c_u)} \quad (9)$$

$$\text{scale} \ = \frac{\text{std}\,(lmk_u - c_u)}{\text{std}\,(lmk_m - c_m)} \quad (10)$$

Where:

- R is the rotation matrix.
- lmk_u and lmk_m are the key point sets of the user and model, respectively.

- $mean$ is the average value of the dataset.
- std is the standard deviation value.
- SVD is the Singular Value Decomposition (SVD), which is a factorization of a matrix into orthogonal matrices and non-square diagonal matrices.

3.5 Skin and Facial Features Matching

In terms of output image natural enhancement, a skin-changing technique is implemented in the pipeline to make models' skin correspond to users' skin. The skin-changing technique is split into 2 phases: skin tone detection and skin changing as Fig. 7 and Fig. 8 show, respectively.

In the first phase, the human skin tone is determined by three color spaces: RGB (for red, green, and blue), HSV (for hue, saturation, and value), and YCRCB. The RGB color space is utilized for separating images into three channels: red, green, and blue. Three threshold values, which represent the human skin color range in different color spaces, are defined in the H channel of HSV space and two channels of YCRCB space to create three masks: one for HSV and two for YCRCB.

After that, these three masks are merged by "AND" logical operators and applied to the input image. The region of the input image covered in the merged mask is considered as "skin region". In "skin region", the human skin tone in RGB is collected by dividing respectively three total values of red, blue, and green by the total number of pixels. In the proposed method, both user and model skin tones are calculated for the next phase: skin changing.

In human faces, there are some non-skin regions such as eyebrows, eyes, lips, glasses, etc. When applying skin changes to these regions, there is a high risk for the output images to become unnatural. Therefore, a preprocessing stage is added before the color-changing stage to filter out these non-skin regions by using lower and upper thresholds of the human skin color range in the skin-changing phase. After the preprocessing stage, the remaining skin regions are altered to match
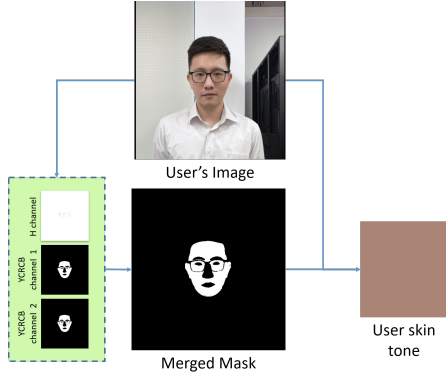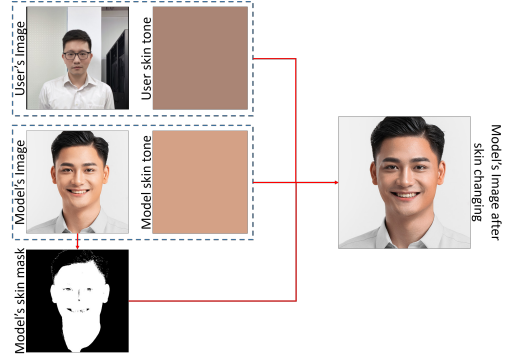
Figure 7: Tone detection phase.

Figure 8: Skin changing phase.

the models' skin. Firstly, the disparity between two skin tones is expressed in three ways as follows:

$$diff_1 = \frac{c_{ref}}{c_{ori}} \qquad (11)$$

$$diff_2 = \frac{255 - c_{ref}}{255 - c_{ori}} \qquad (12)$$

$$diff_3 = \frac{255(c_{ref} - c_{ori})}{255 - c_{ori}} \qquad (13)$$

While $diff_1$, $diff_2$, and $diff_3$ are three values to define the difference between two skin tones, c_{ref} and c_{ori} refer to the reference and original color, respectively. After that, the result image i_{out} is created by comparing with c_{ori} and processing the original image i_{in} with three values $diff_1$, $diff_2$, and $diff_3$:

$$i_{out} = \begin{cases} i_{in} \times diff_1, \text{ if } i_{in} < c_{ori} \\ i_{in} \times diff_2 + diff_3, \text{ if } i_{in} \geq c_{ori} \end{cases} \qquad (14)$$

Notice that these formulas above are applied in a single channel only. Therefore, the input needs to be split into three channels: red, green, and blue to calculate each channel's i_{out} and then compose three values to get the complete output image.

Due to differences in light environment between model and user pictures such as intensity, directions, etc, inspired by Exchanging Latent Encodings with GAN for Transferring Multiple Face Attributes (EleGANt) [21], there is an additional makeup transfer module after skin changing phase to make the user image match the light condition of the model image. In the makeup transfer module, the skin-changed model image and the user image are defined as reference and source, respectively. Thanks to the makeup transfer module, dark and light regions in the model face are estimated and applied to the user's face to create the *light effect* in the user's face equivalent to the model body.

4 Results

This section would divide the results into three separate parts: (1) Evaluate the quantitative in Head Swap,

(2) Evaluate the quantitative in Reshaping module, (3) Evaluate the quantitative and qualitative in the whole process.

4.1 Datasets and hardware specification

To evaluate our proposed method, four datasets: two self-collected and two sub-datasets from VoxCeleb2 [5] are utilized. The first self-collected dataset is Vietnam adult scan data (Viettel dataset), which contains scan data of 156 males and 250 females aged from 18 to 60. Viettel dataset is used for body reshape error ratio evaluation. The second dataset includes portraits and full-body images of five men and ten women for qualitative assessment tasks. The other two sub-datasets are extracted from VoxCeleb2 [5] dataset for quantitative result collection tasks. VoxCelebs [5] is a collection of short interview videos (which last less than five minutes) of celebrities around the world.

The first sub-dataset is accumulated with 13079 pairs of corresponding source and destination images in VoxCeleb2 [5]. These image pairs are used for calculating the Structural Similarity Index measure (SSIM), the Learned Perceptual Image Patch Similarity (LPIPS), and Peak Signal to Noise Ratio (PSNR) of the output from our proposed method. Due to the fact that the Pose Reconstruction Error (E_p) is only evaluated on the same person, another sub-dataset is collected from these 13079 image pairs by sampling source and destination images on each pair of frames of the same video to create 228 new image pairs.

Our proposed method is developed and assessed using an Internet-connected workstation with an Intel Xeon W-2245 CPU, 64GB of memory, and an Nvidia Quadro RTX 8000 GPU running Windows 10 Pro for Workstations with Pytorch, OpenCV, Numpy, and Pycharm installed.

4.2 Head Swap

Our proposed method is compared to the previous state-of-the-art models which included: (1) HeSer[17], (2) FOMM[18], (3) LPD[3], (4) Siarohin *et al.*[19]. Four popular metrics [15][3] are used to make the evaluation:

Table 1: Quantitative comparison on Head Swapping

No.	Method	EP↓	SSIM↑	LPIPS↓	PSNR↑
1	FOMM[18]	0.275	0.76	0.18	30.92
2	LPD[3]	0.063	0.52	0.50	28.84
3	Siarohin *et al.*[19]	0.137	0.73	0.20	30.01
4	HeSer[17]	0.026	0.77	0.19	31.33
5	**Ours**	**0.022**	**0.90**	**0.09**	**34.82**

(1) Pose reconstruction Error (E_p): represents the difference head pose between synthesized and source image using the facial key points, (2) Structural similarity index measure (SSIM): Estimate the perceived quality of images such as luminance masking, contrast masking, hand contrast (3) Peak Signal-to-Noise Ratio (PSNR): Measure the quality of reconstruction of loss image compression codecs (4) Learned Perceptual Image Patch Similarity (LPIPS): Measure the Semantic perceptual similarity between 2 images via the AlexNet [23].

The table gives information about the quantitative results of every metric of five different methods. Overall, our approach outperforms other methods on every metric. As can be seen from Table 1, Pose Error and Structure similarity index are 0.022 and 0.9, better than approximately 15% and 17% respectively compared to "HerSer" model. The biggest difference comes from LPIPS, the best result before is around 0.18 [18], while ours is 0.094, nearly 48% higher.

4.3 Reshaping

To identify the accuracy of the reshaping module, a measurement for quantitative via three ratios in the leg, shoulder, and arm is conducted. OpenPose[4] is used to detect 25 key points and formulas to calculate the ratios are illustrated as follows:

$$leg_{ratio} = \frac{b}{c} \tag{15}$$

$$shoulder_{ratio} = \frac{a2}{a1 + a2 + a3} \tag{16}$$

$$arm_{ratio} = \frac{a1 + a2 + a3}{c} \tag{17}$$

when $a1$, $a2$, $a3$, b, and c are the body part lengths shown in Fig. 9.

The boxplots in Fig. 10 and Fig. 11 and Table 3 illustrate the error ratio (%) by males and females which includes leg, shoulder, and arm. Overall, the mean error in the whole ratio is around 5% and maintained the same level for 406 samples. However, the maximum errors of leg ratios are approximately 18% both in males and females. As can be seen from Fig. 10 and Fig. 11, the number of errors staying below 8% is much larger than the values above.

4.4 Runtime and performance analysis

The average processing time of the head-swapping task over 13079 image pairs is 4 seconds, and the response

Table 2: Qualitative results of the proposed method
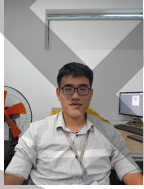
Source Image	Target Image	Result Image
H: 1.70m W: 66kg BMI: 22.84	H: 1.68m W: 67kg BMI: 23.74	
H: 1.70m W: 85kg BMI: 29.41	H: 1.68m W: 67kg BMI: 23.74	
H: 1.70m W: 72kg BMI: 24.91	H: 1.68m W: 73kg BMI: 25.86	
H: 1.45m W: 55kg BMI: 26.20	H: 1.62m W: 50kg BMI: 19.05	
H: 1.50m W: 40kg BMI: 17.80	H: 1.62m W: 50kg BMI: 19.05	
H: 1.60m W: 48kg BMI: 18.80	H: 1.62m W: 50kg BMI: 19.05	

Table 3: Body error ratio of males and females

Gender	Leg error ratio (%)	Shoulder error ratio (%)	Arm error ratio (%)	Mean (%)
Male	5.96	4.72	4.80	5.16
Female	5.68	4.77	4.75	5.07
Mean (%)	5.82	4.75	4.78	5.12
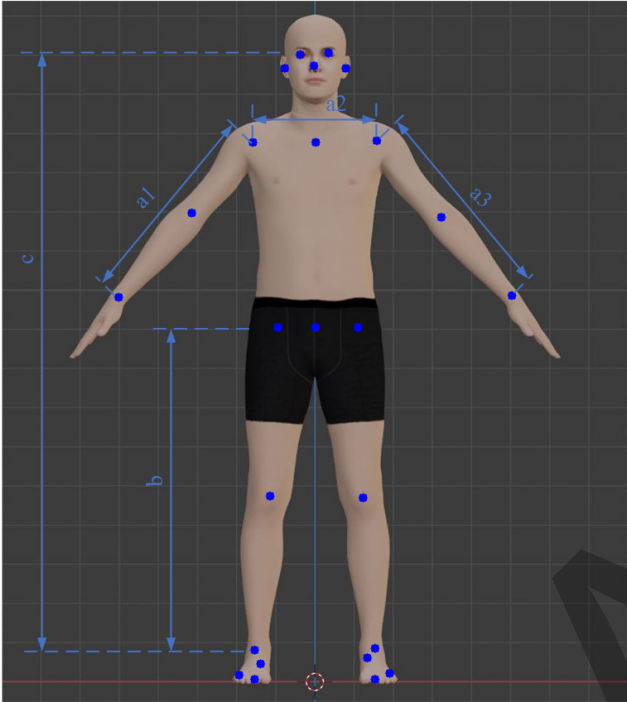
Figure 9: Body part length used in body reshaping assessment.

Figure 10: Male body error ratio.
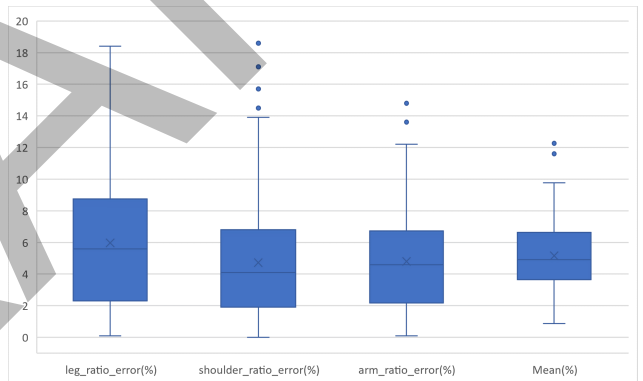
Figure 11: Female body error ratio.

time of the body-reshape module on the Viettel dataset is 0.021 seconds, which almost achieves real-time speed. Table 2 indicates the qualitative outcomes on different body types both male and female. More information and details are put in the Appendix B. Overall, the images after processing express high fidelity to capture the user's characteristics with various body types while maintaining the feature of target images.

Three body types of source and target are chosen to verify the efficiency of the algorithm with a multi-class of BMI: (1) Overweight: $BMI > 25$, (2) Normal weight: $18.5 < BMI < 25$, (3) Underweight: $BMI < 18.5$. The area in the neck and facial skin color are generated naturally and consistent with the whole body. Furthermore, our algorithm could handle and perform very well on long hair cases. The accuracy of segmentation and head swapping combined with skin color present the best results with high precision. Besides, the features of target images such as belly fat/thin or muscular chest,... are kept and performed nicely while changing the outlook to fit based on the measurement of users through the reshaping model.

5 Discussions and future work

5.1 Discussions

In this research, we present a unique photorealistic framework for virtual try-on applications that combines: (1) Face swapping, (2) Body shape Reconstruction, and (3) Skin adaptive changing. Our approach achieves high-fidelity in quantitative results comparing others methods and ensures the visual appearance of users on clothing after try-on processing. The average runtime is around 5 seconds on a single thread and there is no specific requirement in preparing data for processing with an image size of 1280 x 720 when applying a new type of clothing. Besides that, the ease of package modules brings our work a huge potential to apply and integrate into a complete mobile application and system. Our research is the first all-in-one pho-

torealistic method that can handle complex clothing details, facial features, skin tone, and different body shapes with the quickest runtime and simple requirements.

5.2 Limitations

Our approach is sensitive to input the user's portrait image. Selfie photos from users could be an issue if the distance is too close (less than 25cm from the face to camera), the output when swapping faces could be unsuitable with the target and cause an in-efficient visual. Furthermore, the size of the package and hardware requirements are not completely optimized for low-end or mobile devices. Besides, the extreme body shapes are not validated and verified with our solution.

5.3 Future work

Our current method should be to improve the head swapping module by using two approaches: (1) re-process the input data, (2) custom the layers, and (3) re-organize the processing stream to combine modules and reduce processing time. An optimization framework with more lightweight models is being developed for mobile devices like smartphones or low-end computers. A deeper investigation into various body types must be present in the future direction. It would be interesting to add a feature that could reconstruct and change body shape in real-time in the outcome. Finally, our approach needs to integrate with various applications corresponding to virtual try-on and evaluate the performance of end-users.

A Qualitative results of the proposed method

B Body error ratio

References

[1] BENDER, J., MÜLLER, M., OTADUY, M. A., TESCHNER, M., AND MACKLIN, M. A survey on position-based simulation methods in computer graphics. In *Computer graphics forum* (2014), vol. 33, Wiley Online Library, pp. 228–251.

[2] BULAT, A., AND TZIMIROPOULOS, G. How far are we from solving the 2d & 3d face alignment problem? (and a dataset of 230,000 3d facial landmarks). In *International Conference on Computer Vision* (2017).

[3] BURKOV, E., PASECHNIK, I., GRIGOREV, A., AND LEMPITSKY, V. Neural head reenactment with latent pose descriptors. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2020).

[4] CAO, Z., HIDALGO MARTINEZ, G., SIMON, T., WEI, S., AND SHEIKH, Y. A. Openpose: Real-time multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019).

[5] CHUNG, J. S., NAGRANI, A., AND ZISSERMAN, A. Voxceleb2: Deep speaker recognition. In *INTERSPEECH* (2018).

[6] FRATARCANGELI, M., TIBALDO, V., AND PELLACINI, F. Vivace: A practical gauss-seidel method for stable soft body dynamics. *ACM Transactions on Graphics (TOG) 35*, 6 (2016), 1–9.

[7] HONG, F.-T., ZHANG, L., SHEN, L., AND XU, D. Depth-aware generative adversarial network for talking head video generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 3397–3406.

[8] LI, Y., ET AL. N-Cloth: Predicting 3D cloth deformation with mesh-based networks. *Computer Graphics Forum (Proceedings of Eurographics) 41*, 2 (May 2022), 547–558.

[9] PAVLAKOS, G., CHOUTAS, V., GHORBANI, N., BOLKART, T., OSMAN, A. A. A., TZIONAS, D., AND BLACK, M. J. Expressive body capture: 3d hands, face, and body from a single image. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* (2019).

[10] PEROV, I., ET AL. Deepfacelab: Integrated, flexible and extensible face-swapping framework. *arXiv preprint arXiv:2005.05535* (2020).

[11] RANZATO, M., HUANG, F. J., BOUREAU, Y.-L., AND LECUN, Y. Unsupervised learning of invariant feature hierarchies with applications to object recognition. In *2007 IEEE conference on computer vision and pattern recognition* (2007), IEEE, pp. 1–8.

[12] RAZAFINDRAZAKA, F. H. Delaunay triangulation algorithm and application to terrain generation. *International Institute for Software Technology, United Nations University, Macao* (2009).

[13] RIBEIRO, S., MAXIMO, A., BENTES, C., OLIVEIRA, A., AND FARIAS, R. Memory-aware and efficient ray-casting algorithm. In *XX Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2007)* (2007), IEEE, pp. 147–154.

[14] SANTESTEBAN, I., OTADUY, M. A., AND CASAS, D. Learning-based animation of clothing for virtual try-on. In *Computer Graphics Forum* (2019), vol. 38, Wiley Online Library, pp. 355–366.

[15] SARA, U., AKTER, M., AND UDDIN, M. S. Image quality assessment through fsim, ssim, mse and psnr—a comparative study. *Journal of Computer and Communications 7*, 3 (2019), 8–18.

[16] SARKAR, K., GOLYANIK, V., LIU, L., AND THEOBALT, C. Style and pose control for image synthesis of humans from a single monocular view. *arXiv preprint arXiv:2102.11263* (2021).

Table 4: Qualitative results of the proposed method

Source	Target	Result	Source	Target	Result	Source	Target	Result
H: 1.45m W: 55kg BMI: 26.20	H: 1.62m W: 50kg BMI: 19.05		H: 1.45m W: 55kg BMI: 26.20	H: 1.62m W: 50kg BMI: 19.05		H: 1.60m W: 48kg BMI: 18.80	H: 1.62m W: 50kg BMI: 19.05	
H: 1.60m W: 48kg BMI: 18.80	H: 1.62m W: 50kg BMI: 19.05		H: 1.50m W: 40kg BMI: 17.80	H: 1.62m W: 50kg BMI: 19.05		H: 1.50m W: 40kg BMI: 17.80	H: 1.62m W: 50kg BMI: 19.05	
H: 1.70m W: 85kg BMI: 29.41	H: 1.68m W: 67kg BMI: 23.74		H: 1.70m W: 85kg BMI: 29.41	H: 1.74m W: 65kg BMI: 21.5		H: 1.70m W: 85kg BMI: 29.41	H: 1.74m W: 65kg BMI: 21.5	
H: 1.70m W: 66kg BMI: 22.84	H: 1.68m W: 54kg BMI: 19.12		H: 1.70m W: 66kg BMI: 22.84	H: 1.68m W: 73kg BMI: 25.91		H: 1.70m W: 66kg BMI: 22.84	H: 1.68m W: 73kg BMI: 25.91	
H: 1.70m W: 72kg BMI: 24.9	H: 1.68m W: 54kg BMI: 19.12		H: 1.70m W: 72kg BMI: 24.9	H: 1.68m W: 54kg BMI: 25.91		H: 1.70m W: 72kg BMI: 24.9	H: 1.68m W: 67kg BMI: 23.74	

[17] Shu, C., et al. Few-shot head swapping in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 10789–10798.

[18] Siarohin, A., Lathuilière, S., Tulyakov, S., Ricci, E., and Sebe, N. First order motion model for image animation. In *Conference on Neural Information Processing Systems (NeurIPS)* (December 2019).

[19] Siarohin, A., Woodford, O., Ren, J., Chai, M., and Tulyakov, S. Motion representations for articulated animation. In *CVPR* (2021).

[20] Xiang, D., et al. Dressing avatars: Deep photorealistic appearance for physically simulated clothing. *ACM Transactions on Graphics (TOG) 41*, 6 (2022), 1–15.

[21] Xiao, T., Hong, J., and Ma, J. Elegant: Exchanging latent encodings with gan for transferring multiple face attributes. In *Proceedings of the European Conference on Computer Vision (ECCV)* (September 2018), pp. 172–187.

[22] Yang, H., Zhang, R., Guo, X., Liu, W., Zuo, W., and Luo, P. Towards photo-realistic virtual try-on by adaptively generating-preserving image content. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2020).

[23] Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR* (2018).

[24] Zhao, J., and Zhang, H. Thin-plate spline motion model for image animation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 3657–3666.

Table 5: Female body error ratio

No.	Sample	Leg (%)	Shoulder (%)	Arm (%)	Mean (%)	No.	Sample	Leg (%)	Shoulder (%)	Arm (%)	Mean (%)
1	F10	2.3	4.9	3.8	3.7	67	F161	5	14.6	8.3	9.3
2	F100	8.1	5.4	7.5	7.0	68	F162	6.5	0.8	7.6	5.0
3	F101	14.2	0.9	0.6	5.2	69	F163	10.1	1.9	3.8	5.3
4	F102	1.3	0.2	3.2	1.6	70	F164	6.9	0.2	6.3	4.5
5	F103	5.8	1.9	6.6	4.8	71	F165	6.1	4.5	7.2	5.9
6	F104	3.5	8	1.1	4.2	72	F166	2.4	6.6	7.4	5.5
7	F105	3.7	8	8	6.6	73	F167	0.3	3	8.3	3.9
8	F106	16.5	5.4	6.4	9.4	74	F168	18.4	4.2	4.7	9.1
9	F107	2.7	2	9.4	4.7	75	F169	4	7.3	4.1	5.1
10	F108	3.8	4.6	4.7	4.4	76	F170	1.3	10.8	5.8	6.0
11	F109	4.9	8.6	5.1	6.2	77	F171	0.9	3	0.6	1.5
12	F11	4.9	0.7	8	4.5	78	F172	0.8	4.1	1.1	2.0
13	F110	9.5	6.1	6.3	7.3	79	F173	2	0.6	3.2	1.9
14	F111	3.4	0.4	2.9	2.2	80	F174	1.3	0.9	0.6	0.9
15	F112	10.8	6.8	3.9	7.2	81	F175	5.2	10.2	6.7	7.4
16	F113	4.6	6.1	4.7	5.1	82	F176	5.3	2.8	2	3.4
17	F114	12.9	2.5	6.6	7.3	83	F177	6.3	7.4	4.9	6.2
18	F115	1.3	5	11.4	5.9	84	F178	0.2	2.5	4.6	2.4
19	F116	10	2.7	3.8	5.5	85	F179	1.3	5.4	1	2.6
20	F117	14.3	5.5	1.5	7.1	86	F18	2	3.9	5.7	3.9
21	F118	5.6	13.9	1.4	7.0	87	F180	0.8	10.5	1.2	4.2
22	F119	0.1	0.3	6.7	2.4	88	F181	4.6	10.9	0.7	5.4
23	F12	7.2	4.8	7.7	6.6	89	F182	12	1.3	2.5	5.3
24	F120	7.8	2.1	10.4	6.8	90	F183	10.7	4.4	5.8	7.0
25	F121	7.4	3.7	8.5	6.5	91	F184	8.9	5.8	0.6	5.1
26	F122	9.4	5	11.4	8.6	92	F185	13	2.3	6.6	7.3
27	F123	1.7	5.8	6.3	4.6	93	F186	8.2	5.8	5.9	6.6
28	F124	0.9	7	4.3	4.1	94	F187	10.2	6.5	1.3	6.0
29	F125	6.5	9.8	7.8	8.0	95	F188	5.9	5.9	1.7	4.5
30	F126	5.9	18.6	4.8	9.8	96	F189	18.1	13.6	5.1	12.3
31	F127	5.1	6.8	13.6	8.5	97	F19	6.8	2	3.5	4.1
32	F128	0.1	8.7	8	5.6	98	F190	12.4	5.8	2.1	6.8
33	F129	0.6	4.6	8.6	4.6	99	F191	0.8	7.9	0.8	3.2
34	F13	4	5	0.7	3.2	100	F192	4	2.9	3.1	3.3
35	F131	4.7	4.3	5.2	4.7	101	F193	6.8	3.5	4.1	4.8
36	F132	12	2.6	6.1	6.9	102	F194	7.2	0.5	0.6	2.8
37	F133	3.1	14.5	2.3	6.6	103	F195	6.8	7.4	1.1	5.1
38	F134	12	8.2	4.1	8.1	104	F196	13.8	6.2	2.3	7.4
39	F135	13	13.5	10.8	12.4	105	F197	7.6	15.7	2.4	8.6
40	F136	10.6	8.7	1.3	6.9	106	F198	8.5	3.8	0.8	4.4
41	F137	9.4	0.4	1.2	3.7	107	F199	2.1	6.8	3.6	4.2
42	F138	5.7	1.5	6.7	4.6	108	F20	5.5	6.6	5.3	5.8
43	F139	4.1	7.6	11.5	7.7	109	F200	11	0.5	3.7	5.1
44	F14	8.9	1.5	2.2	4.2	110	F201	7.4	11.6	3.8	7.6
45	F140	8.5	0.3	3.2	4.0	111	F202	7.9	7.2	6.7	7.3
46	F141	13.7	2.2	6.9	7.6	112	F203	0.1	5.7	2.3	2.7
47	F142	1.2	7.8	7.4	5.5	113	F204	8.5	1.4	8.8	6.2
48	F143	6.4	2	8	5.5	114	F205	0.2	2.2	2.3	1.6
49	F144	9.7	7.2	6.3	7.7	115	F206	5.4	5.4	5.9	5.6
50	F145	2.6	3.3	4.1	3.3	116	F207	0.1	10.3	5.7	5.4
51	F146	15.5	5.6	5.9	9.0	117	F208	0.6	3.9	4.6	3.0
52	F147	9.1	8.2	11.1	9.5	118	F209	5.7	2.1	6.1	4.6
53	F148	2.7	4.2	9	5.3	119	F21	5.6	3.4	6.1	5.0
54	F149	9.2	2.9	2.6	4.9	120	F210	2	7.5	2.1	3.9
55	F150	0.1	9.1	1.9	3.7	121	F211	11.3	3.2	0.8	5.1
56	F151	0.4	1.9	7.4	3.2	122	F212	1.8	6.8	0.8	3.1
57	F152	1	10.3	8.8	6.7	123	F213	3.1	1.9	4.4	3.1
58	F153	9.4	0	10.1	6.5	124	F214	8.4	5.4	4.4	6.1
59	F154	15.2	1.6	7.6	8.1	125	F215	11	7.7	4.3	7.7
60	F155	13.4	4.9	7.6	8.6	126	F216	4.1	3.8	1.3	3.1
61	F156	4.2	2.9	12.2	6.4	127	F217	0.9	4.3	3.1	2.8
62	F157	2.1	0.5	9.5	4.0	128	F218	7.2	2.8	6.9	5.6
63	F158	12.6	2.6	9.6	8.3	129	F219	3.7	5.7	5	4.8
64	F159	13.2	0.4	14.8	9.5	130	F22	2.3	5.8	0.2	2.8
65	F16	11.7	7.6	1.8	7.0	131	F220	6.9	6	3.2	5.4
66	F160	12.6	6.1	0.8	6.5	132	F221	3.4	7	4.4	4.9

No.	Sample	Leg (%)	Shoulder (%)	Arm (%)	Mean (%)	No.	Sample	Leg (%)	Shoulder (%)	Arm (%)	Mean (%)
133	F222	2.3	6.2	6.3	4.9	192	F45	3.8	1.4	4	3.1
134	F223	8.2	17.1	3.4	9.6	193	F46	4.8	0.4	5.9	3.7
135	F224	3.7	6.5	3.5	4.6	194	F47	5.4	0.9	8.5	4.9
136	F225	1.9	1.4	6.6	3.3	195	F48	1.4	0.9	7	3.1
137	F226	1.1	0.6	5.3	2.3	196	F49	10.5	3.7	0.4	4.9
138	F227	8.5	0	4.3	4.3	197	F5	1.9	4.3	5	3.7
139	F228	0.3	4	4	2.8	198	F50	2.5	1.9	0.1	1.5
140	F229	1.5	2.8	6.4	3.6	199	F51	7.3	12.1	5.5	8.3
141	F23	9.7	1.4	0.2	3.8	200	F52	5.6	5.5	4.4	5.2
142	F230	0.7	6.5	1	2.7	201	F53	16.8	11.2	6.8	11.6
143	F231	6.4	2	4.6	4.3	202	F54	6.4	4.3	7.8	6.2
144	F232	0.5	5.8	3.4	3.2	203	F55	4.9	6.5	2.9	4.8
145	F233	2.8	1.6	0.6	1.7	204	F56	1	0.8	3.3	1.7
146	F234	7.5	5.4	11.8	8.2	205	F57	11	1.4	0.3	4.2
147	F235	4.6	1.1	4.5	3.4	206	F58	6.2	9.9	6.8	7.6
148	F236	6.1	2.7	0.6	3.1	207	F59	1.1	0.7	1.1	1.0
149	F237	5.1	3.9	6.7	5.2	208	F6	8.2	5	7.4	6.9
150	F238	1.7	0.2	6.6	2.8	209	F60	5.4	2.3	6.6	4.8
151	F239	1.1	2.8	5.8	3.2	210	F61	1.2	7.1	6.3	4.9
152	F24	2.6	1.8	5.2	3.2	211	F62	6.4	2.5	4.6	4.5
153	F240	11.2	0.7	2.5	4.8	212	F63	10.8	3.5	3.6	6.0
154	F241	4.1	2.1	0.1	2.1	213	F64	1.4	2.2	0.7	1.4
155	F242	2.7	1.1	5	2.9	214	F65	6.6	6.8	2.8	5.4
156	F243	6.6	7.8	0.1	4.8	215	F66	7.9	13.3	2.5	7.9
157	F244	3.5	10.4	6.8	6.9	216	F67	18	1.6	1.1	6.9
158	F245	10.1	4.1	1.1	5.1	217	F68	0.2	7.2	2.9	3.4
159	F246	9.1	2.5	6.8	6.1	218	F69	1.9	0.5	10.3	4.2
160	F247	2.3	2.4	7.9	4.2	219	F7	1.7	7.2	0.9	3.3
161	F248	10.6	0.1	3.3	4.7	220	F70	10.8	2.5	1.1	4.8
162	F249	1.3	1.3	0.4	1.0	221	F71	10.8	4.5	0.4	5.2
163	F25	6.5	3.7	5.8	5.3	222	F72	6.7	13.4	3.7	7.9
164	F250	9.8	0.3	7.9	6.0	223	F73	3.3	3.4	9.9	5.5
165	F252	8.1	1.7	7.3	5.7	224	F74	0.5	1.2	3.7	1.8
166	F253	10.7	6.1	7.5	8.1	225	F76	13	3.2	1	5.7
167	F254	5.6	8.2	5.9	6.6	226	F77	3.1	7.2	5.6	5.3
168	F255	10.4	8.4	8.9	9.2	227	F78	4.3	1.3	3	2.9
169	F256	10.3	11.5	1.5	7.8	228	F79	6.5	7.7	6.1	6.8
170	F257	12.8	3.3	1.4	5.8	229	F8	2.5	2.1	8.3	4.3
171	F258	2.3	1.8	5.3	3.1	230	F80	0.3	3.1	0.8	1.4
172	F259	9.5	10.4	6.5	8.8	231	F81	3	4.8	4.3	4.0
173	F26	3.9	9	3.3	5.4	232	F82	8.6	4.9	2.4	5.3
174	F27	5.8	1.1	1.5	2.8	233	F83	0.5	2	0.5	1.0
175	F28	1.6	4.4	1	2.3	234	F84	4.2	3.2	1.2	2.9
176	F29	3.7	4.1	6.7	4.8	235	F85	2.2	1	5.3	2.8
177	F3	0.4	4.8	7.7	4.3	236	F86	6.4	7.2	0.4	4.7
178	F30	7.8	3.2	2	4.3	237	F87	13.9	2.4	10.1	8.8
179	F31	6.3	4	0.4	3.6	238	F88	5.6	1.1	3.3	3.3
180	F32	8.7	1.1	1.5	3.8	239	F89	1.1	0.1	10.9	4.0
181	F34	5.3	4.2	1.2	3.6	240	F9	1.8	2.3	10.3	4.8
182	F35	6.8	1.2	7.5	5.2	241	F90	5.2	1.7	4.4	3.8
183	F36	4.2	6.6	1.8	4.2	242	F91	9.2	3.4	7	6.5
184	F37	6.6	12	6.4	8.3	243	F92	8.1	4.9	11.9	8.3
185	F38	6.2	5.2	3.8	5.1	244	F93	6.3	10.2	9.1	8.5
186	F39	3.1	1.2	3.2	2.5	245	F94	3.6	2.8	4.9	3.8
187	F4	2.7	9.3	6.7	6.2	246	F95	11	7.1	5	7.7
188	F41	13.4	4.4	9.2	9.0	247	F96	0.5	0	2.1	0.9
189	F42	1	2.8	4.6	2.8	248	F97	3.8	4.5	5	4.4
190	F43	6.8	1.8	3.2	3.9	249	F98	6.6	7.5	6.3	6.8
191	F44	6.7	1.7	6.3	4.9	250	F99	5	2.1	6.1	4.4

Table 6: Male body error ratio

No.	Sample	Leg (%)	Shoulder (%)	Arm (%)	Mean (%)	No.	Sample	Leg (%)	Shoulder (%)	Arm (%)	Mean (%)
1	M100	1.7	4.8	11.9	6.1	74	M27	1	0.2	8	3.1
2	M101	9.9	1.8	0.2	4.0	75	M28	6.5	6.5	11	8.0
3	M102	6.9	3.7	3.4	4.7	76	M29	1.5	1.1	6.2	2.9
4	M103	7.5	6.9	9	7.8	77	M30	3.3	2.5	6.1	4.0
5	M104	2	2.3	0.3	1.5	78	M31	4.4	2.3	7	4.6
6	M106	0.7	3.3	5.5	3.2	79	M32	4.6	6.9	1.1	4.2
7	M107	6.3	0.8	7.8	5.0	80	M33	3.7	1.9	3	2.9
8	M108	12.1	0.8	6.5	6.5	81	M34	2.9	3.4	2.4	2.9
9	M109	9.9	8.5	1.2	6.5	82	M35	1	4.3	2	2.4
10	M110	4.8	2.9	0.8	2.8	83	M36	0.4	7.6	10.2	6.1
11	M111	15	6	4.5	8.5	84	M37	10.1	3.4	2.9	5.5
12	M112	9.5	0.9	1.5	4.0	85	M38	1.9	8.6	7.8	6.1
13	M113	14.7	1.5	2.7	6.3	86	M39	3	5.3	0.1	2.8
14	M114	7.3	3.3	6.8	5.8	87	M40	2.1	2.6	4	2.9
15	M115	10.6	5.6	4.7	7.0	88	M41	5.7	5.6	2.1	4.5
16	M116	8.2	4.1	4.6	5.6	89	M42	0.6	10.5	1.9	4.3
17	M117	12.3	7.5	0.7	6.8	90	M43	3.4	2	5.6	3.7
18	M118	11.3	0.8	4.4	5.5	91	M44	4.8	12.8	7.1	8.2
19	M119	8.4	9.5	0.4	6.1	92	M45	10.2	2.4	9.9	7.5
20	M120	8.4	2.3	2.2	4.3	93	M46	9.3	4.6	1.7	5.2
21	M121	10.5	10.2	2.1	7.6	94	M47	2.7	3.7	1.6	2.7
22	M122	9.6	7.1	2.4	6.4	95	M48	16.4	2.5	4.4	7.8
23	M123	8.9	4.3	6.2	6.5	96	M49	0	2.9	4.8	2.6
24	M124	11.9	11.1	2.6	8.5	97	M50	2.6	0.9	3.4	2.3
25	M125	10.6	3.4	9.3	7.8	98	M51	9.5	2.3	2.7	4.8
26	M126	8.8	8.3	0.4	5.8	99	M52	4.1	0.9	7	4.0
27	M127	9.7	5.1	3.2	6.0	100	M53	2.3	2.3	8.3	4.3
28	M128	5.4	4.5	0.4	3.4	101	M54	0.2	7.8	8.6	5.5
29	M129	9	9.1	0.7	6.3	102	M55	1.2	7.2	6.4	4.9
30	M130	9	9.3	7.7	8.7	103	M56	1.6	5.4	3.6	3.5
31	M131	12.1	13.4	2.9	9.5	104	M57	7.8	3	6.5	5.8
32	M132	11.4	12.6	3.9	9.3	105	M58	8.3	4.4	4.5	5.7
33	M133	5.7	14.2	0.6	6.8	106	M59	0.3	4.1	8.3	4.2
34	M134	1.3	11.6	1.7	4.9	107	M60	3.3	1.3	1.7	2.1
35	M135	4.6	4	0.9	3.2	108	M61	3.9	3.4	3.6	3.6
36	M136	10.5	4.5	5	6.7	109	M62	0.2	7.7	4.1	4.0
37	M137	5	4.5	7.1	5.5	110	M63	2.5	4.7	4	3.7
38	M138	8.7	4.3	1.6	4.9	111	M64	1	0.4	3.7	1.7
39	M139	0.4	9.7	6.3	5.5	112	M65	14.4	2.9	10.4	9.2
40	M14	4.6	0.2	1.2	2.0	113	M66	11.3	0.7	0.9	4.3
41	M140	3.4	0.8	0.5	1.6	114	M67	2.4	11.3	1.3	5.0
42	M141	1.7	2.7	2.9	2.4	115	M68	1	0.9	0.1	0.7
43	M143	0.1	1.3	8.3	3.2	116	M69	17.1	5.5	3.6	8.7
44	M144	1.9	2.7	4.6	3.1	117	M70	5.5	12.9	2.2	6.9
45	M145	0.5	2.6	6.4	3.2	118	M71	5.1	1.9	6.4	4.5
46	M146	9.6	13.4	0	7.7	119	M72	1.1	0.7	5.2	2.3
47	M147	4.2	13.5	6.6	8.1	120	M73	0.7	11	0.5	4.1
48	M148	0.5	2	4.5	2.3	121	M74	12.4	10.4	0.5	7.8
49	M149	1.3	0.9	4.6	2.3	122	M75	10.1	3	7.9	7.0
50	M15	3.7	2	3.3	3.0	123	M76	4.6	2.9	8	5.2
51	M150	2.4	9.2	6.9	6.2	124	M77	6.2	14.7	10.5	10.5
52	M151	2.3	8.2	6.8	5.8	125	M78	4.7	4.3	2.5	3.8
53	M152	10.7	2.7	3.1	5.5	126	M79	7.2	14.5	11	10.9
54	M153	8.2	0.8	5.1	4.7	127	M80	6.1	1.6	8.1	5.3
55	M154	1.4	3.5	6.6	3.8	128	M81	2	8.6	6.2	5.6
56	M155	9.7	6.8	0.5	5.7	129	M82	0.9	4.7	8.2	4.6
57	M157	5.6	0.3	3.8	3.2	130	M83	6.4	0.3	11.3	6.0
58	M158	10.7	0.7	4.5	5.3	131	M84	6.7	4	8.5	6.4
59	M159	4.1	0.3	0.5	1.6	132	M85	1.5	1.4	10.2	4.4
60	M16	0.6	2.2	3.3	2.0	133	M86	8.8	0.3	6	5.0
61	M160	2.7	7.1	3.8	4.5	134	M87	5.7	9.9	4.3	6.6
62	M161	4	4.9	3.7	4.2	135	M88	12.1	2.9	9.2	8.1
63	M162	4.3	8.3	1	4.5	136	M89	10.4	0.7	10.5	7.2
64	M163	7.2	3.9	5.7	5.6	137	M90	3.9	14.3	4.7	7.6
65	M17	0.7	3.7	6.1	3.5	138	M91	7.8	1.2	0.6	3.2
66	M18	0.8	4.5	4.1	3.1	139	M92	2.7	2	8.1	4.3
67	M19	1.2	1.2	5.8	2.7	140	M93	10.9	8.3	13.7	11.0
68	M20	11.8	4	4.6	6.8	141	M94	6.9	1	7.5	5.1
69	M21	2.9	6.7	3.5	4.4	142	M95	5.2	5.8	8.4	6.5
70	M23	1.8	10.5	3.6	5.3	143	M96	14.7	7	4.9	8.9
71	M24	13.4	3.3	8.3	8.3	144	M97	6.9	0.7	8.4	5.3
72	M25	9.7	0.2	6.5	5.5	145	M98	2	0.7	8.1	3.6
73	M26	10	12.4	3.8	8.7	146	M99	0.4	0.4	8.1	3.0