Scan-over-Clothes (SOC): Improved Body Measurement Accuracy When Scanning Loose-Clothed Subjects

Matthew S. GILMER, Steven C. HAUSER David BRUNER
Size Stream LLC, Cary NC, USA.

https://doi.org/10.15221/25.04

Abstract

At Size Stream we continuously strive to enhance the accuracy of our 3D body scanning technology while minimizing user friction. We have previously employed a model that detects loose clothing during scanning. In the cases of positive detection, we have prompted the user with a suggestion to change into tighter-fitting attire. We now introduce a method that effectively compensates for loose clothes for our avatar generation and measurement predictions. For scans of subjects wearing clothing that is marginally loose and only partially covering the body, this compensation results in an accuracy degradation of less than 10% compared to scans of those same subjects wearing ideal scanning attire. This improvement was achieved through two primary developments. First, we developed a `Body-Plus-Clothes (BPC)` segmenter which, in addition to separating the subject from the background, distinguishes the clothing from the bare body. Aside from the added capability to segment clothing, this model also provides a substantial accuracy improvement in separating the subject from the background when compared to our previous segmenter (`S2023`). The `BPC` segmenter was trained using real images and supervised to manually-corrected output from an open-source segmentation model. Second, leveraging the BPC segmenter, we developed a `Scan-Over-Clothes (SOC)` Body Measurement Model (BMM) specifically designed to adjust for the presence of loose clothing during body reconstruction and measurement estimation. The `SOC` model was trained using a combination of `BPC` segmentation of real images and internally-generated synthetic data. This novel approach results in a solution that substantially enhances measurement accuracy for scans involving loose clothing, raising the possibility for reliable Ready-To-Wear (RTW) sizing with relaxed scan wear requirements.

Keywords: 3D body scanning, machine learning, mobile scanning, segmentation, synthetic data

1. Introduction

We at Size Stream have long been providing Made-To-Measure (MTM) clothing to our customers. When we started, we required measurements from one of our booth scanners. The final version of this was the `SS20`. Beginning in 2020, we shifted our focus to mobile measurement solutions to make the scanning process more accessible. After years of improving the measurement quality in our mobile scanning, we achieved a 7% remake rate for MTM clothing. Now we are exploring another avenue towards making the scan process easier. In this paper we describe our first steps towards relaxing our scan wear requirements. We introduce two key innovations: the `Body-Plus-Clothes (BPC)` segmenter, capable of distinguishing a subject's body from their worn clothing with high fidelity, and the `Scan-Over-Clothes (SOC)` Body Measurement Model (BMM), trained to compensate for loose garments. As a bonus, we find that the `BPC` segmenter significantly outperforms our previous segmenter (`S2023`) in the task of separating the subject from the background. While we have yet to put the `SOC` BMM into practice, we began upgrading our mobile apps to use the `BPC` segmenter for subject segmentation in June of this year.

This paper is organized as follows: in Section 2, we describe the methods by which we developed these key innovations beginning with the `BPC` segmenter (Section 2.1) and ending with the `SOC` BMM (Section 2.2). In Section 3 we present the results following the same structure as Section 2. Finally, we summarize our findings in Section 4.

2. Methods

2.1. Body-Plus-Clothes (BPC) Segmenter

Here we describe the development of the `BPC` segmenter that we first deployed in June. It is currently only being used for the task of separating the subject from the background, but it also separately segments clothing worn by the subject.

2.1.1. Preparation of Training Data

We trained the `BPC` segmenter on tens of thousands of images collected from our body-scanning applications. We determined the segmentation labels for these images by first batch processing the images using an open-source segmentation engine. The raw batch process output was initially not suitable for training with, as articles of clothing were often missed (partially or entirely) and/or background objects would be segmented.

The next step was to review and correct (when needed) the results of the batch processes. We had previously developed a Graphical User Interface (GUI) for data annotation which we used to prepare training data for a pose detector and our real-time scan quality checking system (`CHECKER`), it will hereafter be referred to as `Annotation GUI (AGUI)`. We adapted `AGUI` to this new task in the following ways. First, we configured `AGUI` to display the RGB image with a single segmentation result overlaid (red for clothes or blue for body). The annotator can cycle through segmentation results for each image, choosing to discard them or not using the top panel of checkboxes, and then finally to the next image using keystrokes. Second, we added functionality for the annotator to draw a line over the image by clicking and dragging the mouse. If the annotator made a mistake, they could use the keyboard to clear the entire line or clear the last drawn point. The line drawn specifies a new blob in which the final point connects in a straight line to the initial point. The annotator could use keystrokes and this new blob to either add to or subtract from the displayed segmentation blob, with the display updating immediately.

2.1.2. Model Training

We chose a lightweight architecture suitable to the task and to mobile deployment. We trained the model in two steps. First, we trained the model using a publicly available fashion dataset, complete with ground-truth labels, containing tens of thousands of images. Then, we fine-tuned the model on our own dataset using standard image augmentation techniques (mirroring, translation, rotation, scaling, and color jitter), taking care to preserve the precise overlap of the body and clothing masks by performing the same augmentations on each (except in the case of color jitter).

2.2. Scan-Over-Clothes (SOC) Body Measurement Model (BMM)

We had previously distilled our BMM, `BMM2024`, onto a lighter weight architecture for the purpose of reducing processing time (`BMM2025`). This model took as input the front and side silhouettes along with the height, age, weight, and gender of the subject. Here, we used a similar strategy by training with the `BMM2024` measurement results as ground-truth but now leveraging the newly developed `BPC` segmenter. We modified the architecture to allow for the input of clothing masks in addition to subject silhouettes so that the model could distinguish between parts of the silhouettes that are clothed and parts of the silhouettes that are not.

2.2.1. Preparation of Training Data

We processed tens of thousands of RGB images collected from our body-scanning applications with the `BPC` segmenter. These images were from body scans where for each scan we had two images (a front view and a side view). However, we wanted to train only with ‘clean’ scans (in which subjects were wearing proper (skin-tight) scan wear) for which the `BMM2024` results accurately reflected the subject’s body measurements. To determine which scans were ‘clean’ and which were ‘loose’ instead, we used our previously developed loose clothes detector (a part of our real-time scan quality checking scheme, `CHECKER`). While we did not train with the ‘loose’ scans, we did use the clothing masks from their BPC segmentation results. Our data augmentation scheme in this case relies on randomly sampling from the clothing masks of the loose-clothed scans to be input along with ‘clean’ silhouettes.

2.2.2. Model Training

During training, the model always receives two silhouettes (front and side views) from one of the ‘clean’ scans. Some of the time the model receives the clothing masks of the very same ‘clean’ scan while the rest of the time it receives the clothing masks of a random loose-clothed scan. This ensures the model will learn to predict accurate measurements whether loose clothes are worn or not. Then, we apply standard image augmentation techniques as described in Section 2.1.2 (but without color jitter). After image augmentation, we took care to modify the silhouettes so that they too had in their foreground the same pixels as the accompanying clothing masks. This training scheme allows for an ever-changing stream of synthetic ‘clean’ and loose-clothed training examples.

3. Results

3.1. Body-Plus-Clothes (BPC) Segmenter

We found our new segmenter exceeded our expectations in being able to reliably segment out clothing from the body. Additionally, it showed a marked improvement over `S2023` in subject segmentation (separating the scanning subject from the background). Consequently, we are in the process of upgrading our scanning applications with the `BPC` segmenter, giving our measurement results a boost in both accuracy and precision. An example scan and its segmentation result is shown in Figure 1.

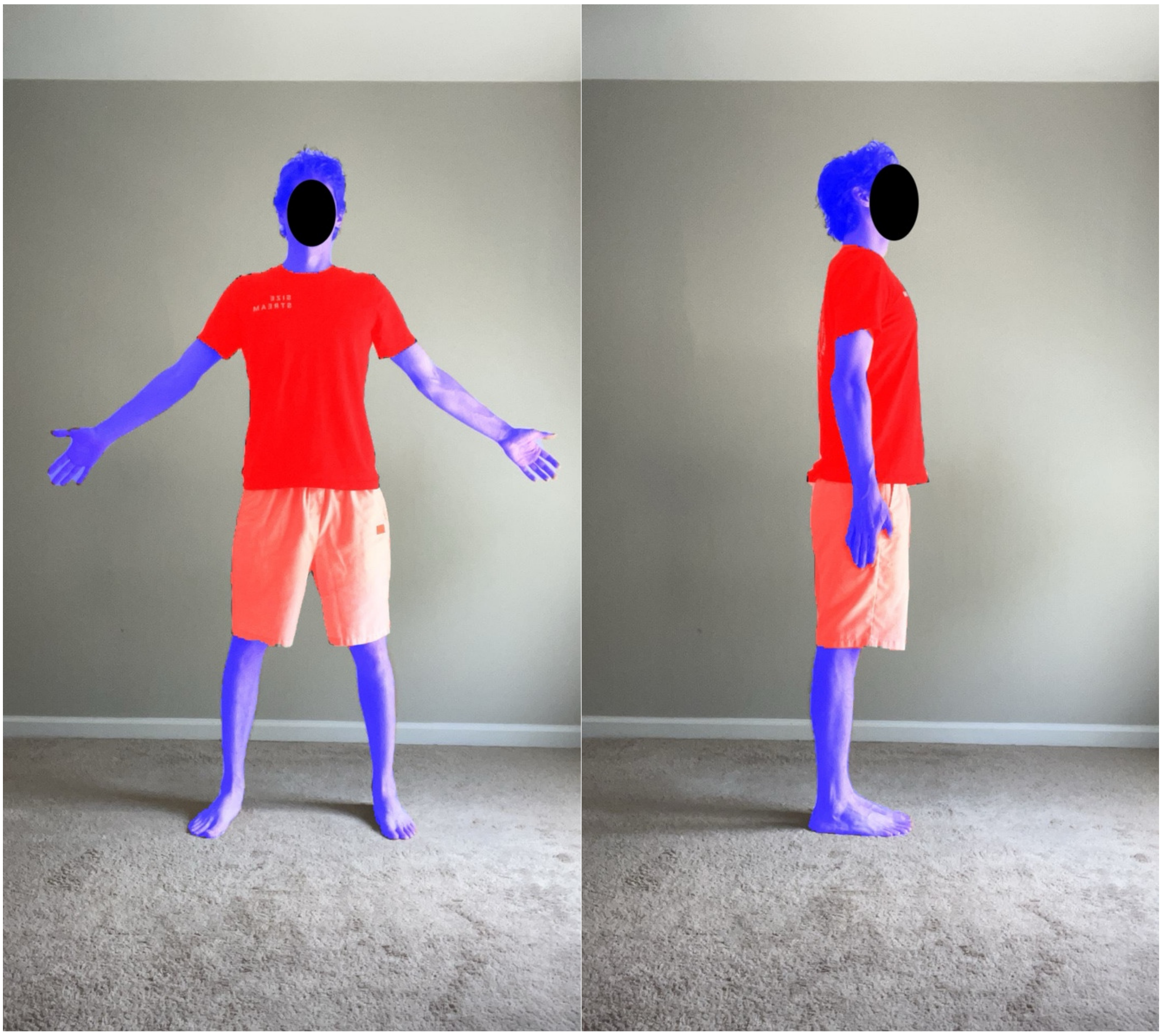

Fig. 1. Example scan with `BPC` segmentation masks represented (red: clothes, blue: body).

3.1.1. Accuracy of Subject Segmentation

The accuracy improvement in subject segmentation is reflected in the results of our internal ‘scan party’ datasets (hereafter referred to as Dataset 1) containing multiple scans of each of 30 male and 30 female subjects (see Table 1). Table 1 shows the Coefficient of Variation (CV) which expresses the variability measurements across multiple scans by the same subject taken one after the other, the dimensionless standard deviation (dimSTD), and the dimensionless Mean Average Error (dimMAE). The latter two are accuracy metrics calculated from the percent errors (relative to our booth scanner, the `SS20`) across subjects. Both metrics are averaged over all subjects and a set of 46 measurements. Additionally, we include in parentheses the percent change for the `BPC` numbers against those of the `S2023`. We find the largest improvement between `S2023` and the `BPC` segmenter is in the CV, but the data also show a slight accuracy improvement. The impact on our customers who scan at home is more difficult to quantify since we do not have `SS20` scans for them, however, it is likely stronger due to the prevalence of lighting issues and cluttered scan environments (absent in our 'scan party' data), which the `BPC` segmenter has improved resiliency against (see Figure 2). Additionally, the `BPC` segmenter perfectly segments limbs with tattoos that `S2023` could not and it routinely produces segmentations with sharper silhouette boundaries and less background noise. In a random sample of 100 scans (200 images), we found only two examples where there was a small portion of the background segmented as part of the subject and in each case the real boundary was difficult to see by eye due to background clutter.

Gender	Segmenter	CV [%]	dimSTD [%]	dimMAE [%]
Male	BMM2025	0.68	4.13	3.28
	SOC	0.60 (-12%)	4.05 (-2%)	3.24 (-1%)
Female	BMM2025	0.63	3.56	3.34
	SOC	0.54 (-14%)	3.48 (-2%)	3.08 (-8%)

Table 1.

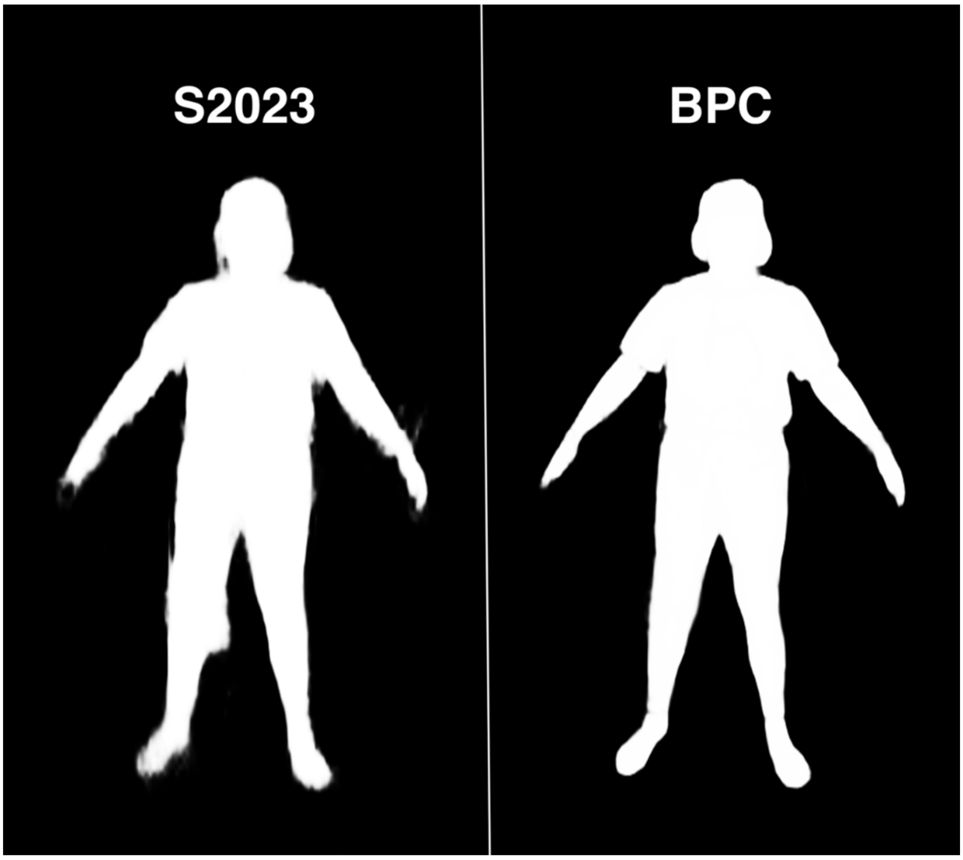

Fig. 2. Comparison of the `S2023` (left) and the `BPC` (right) segmentation results for an example front view image that had bad lighting and a cluttered background.

3.1.2. Accuracy of Clothing Segmentation

In the same sample of 100 scans (see Section 3.1.1) we found only two examples where a garment was partially missed, in both cases the missed region was counted as body but not clothes. The model even segmented relatively rare garments like hats. We concluded that the clothing segmentation of the `BPC` segmenter is reliable enough to be used as input for a SOC BMM.

3.2. Scan-Over-Clothes (SOC) Body Measurement Model (BMM)

We evaluate our `SOC` BMM on two different datasets. The first is `Dataset 1` (introduced in Section 3.1.1). This dataset has only 'clean' scans, but we include the evaluation here since we trained the model to be able to handle both 'clean' and 'loose' scans. The second is `Dataset 2` which contains multiple scans of each of 8 male and 7 female subjects. This dataset is half 'clean' scans and half 'loose' scans. For the 'loose' scans, the subjects were allowed to choose their clothing which ended up being a mix of shorts, slacks, jeans, sweatpants, t-shirts, and sweaters.

3.2.1. Evaluation on Dataset 1

The results for this evaluation are shown in Table 2. We also include the percent changes of all metrics for the `SOC` BMM against those of `BMM2025` (negative in green and positive in red). We find that for the `SOC` BMM the CV is slightly lower for the male subjects and slightly higher for the female subjects. We also find that the `SOC` model has dimSTD and dimMAE values that are slightly higher (with percent changes not exceeding 10%) for both male and female genders. Therefore, the `SOC` BMM is slightly less accurate on 'clean' scans than BMM2025.

Gender	BMM	CV [%]	dimSTD [%]	dimMAE [%]
Male	BMM2025	0.60	4.05	3.24
	SOC	0.57 (-5%)	4.34 (+7%)	3.42 (+6%)
Female	BMM2025	0.54	3.48	3.15
	SOC	0.57 (+6%)	3.83 (+10%)	3.31 (+5%)

Table 2.

3.2.2. Evaluation on Dataset 2

The results for this evaluation are shown in Table 3. Here the percent changes shown are with respect to the 'clean' scans run through `BMM2025`. For both male and female genders, the `SOC` BMM is only slightly less accurate than `BMM2025` on 'clean' scans. We also note the CV for female subjects on 'clean' scans is significantly lower than it is for `BMM2025`. These results, however, have less significance than the analogous results of 3.2.1 due to Dataset 2 having a quarter the number of subjects as `Dataset 1.`

When processing scans with `BMM2025` we see a strong decrease in accuracy for the 'loose' scans relative to the 'clean' scans of the same subjects. This is especially true for dimMAE which is more sensitive to biases than dimSTD. `BMM2025` is not equipped to account for loose clothes in its measurement predictions leading to overestimation of measurements. We also note that this accuracy degradation is stronger for the female subjects, and this is likely due to these subjects choosing looser clothing on average than the male subjects.

Finally, we describe the main result which is that when the `SOC` BMM is used on 'loose' scans instead of `BMM2025` all metrics are significantly reduced. The `SOC` BMM results still show a degradation in performance relative to the `BMM2025` results on the 'clean' scans (excepting the reduction in CV for male subjects) but this degradation is dwarfed by that of the results for using `BMM2025` on those same 'loose' scans. As a reminder the metrics in these tables have been averaged over a set of 46 measurements. The measurements averaged vary in their sensitivity to the 'loose' clothing, so we additionally include Figure 3 to highlight three circumference measurements that have a strong sensitivity. The biases, shown here as dimensionless mean differences (dimMD) introduced by 'loose' clothing are effectively eliminated when the `SOC` BMM is used. While we have not yet attempted to do RTW sizing with the `SOC` BMM, we are encouraged by the results presented here to expect it would prove successful. In the case of the male subjects, the accuracy degrades only 10% relative to using BMM2025 on 'clean' scans (a method already proven to work for MTM clothing).

Gender	'Clean' or 'Loose'	BMM	CV [%]	dimSTD [%]	dimMAE [%]
Male	'Clean'	BMM2025	0.49	3.62	3.20
		SOC	0.49 (+0%)	3.65 (+1%)	3.10 (+3%)
	'Loose'	BMM2025	0.58 (+18%)	5.20 (+44%)	5.60 (+75%)
		SOC	0.38 (-22%)	3.97 (+10%)	3.43 (+7%)
Female	'Clean'	BMM2025	0.52	3.82	3.34
		SOC	0.41 (-21%)	3.90 (+2%)	3.31 (-1%)
	'Loose'	BMM2025	0.78 (+51%)	5.66 (+48%)	7.21 (+116%)
		SOC	0.56 (+8%)	4.40 (+15%)	3.74 (+12%)

Table 3.

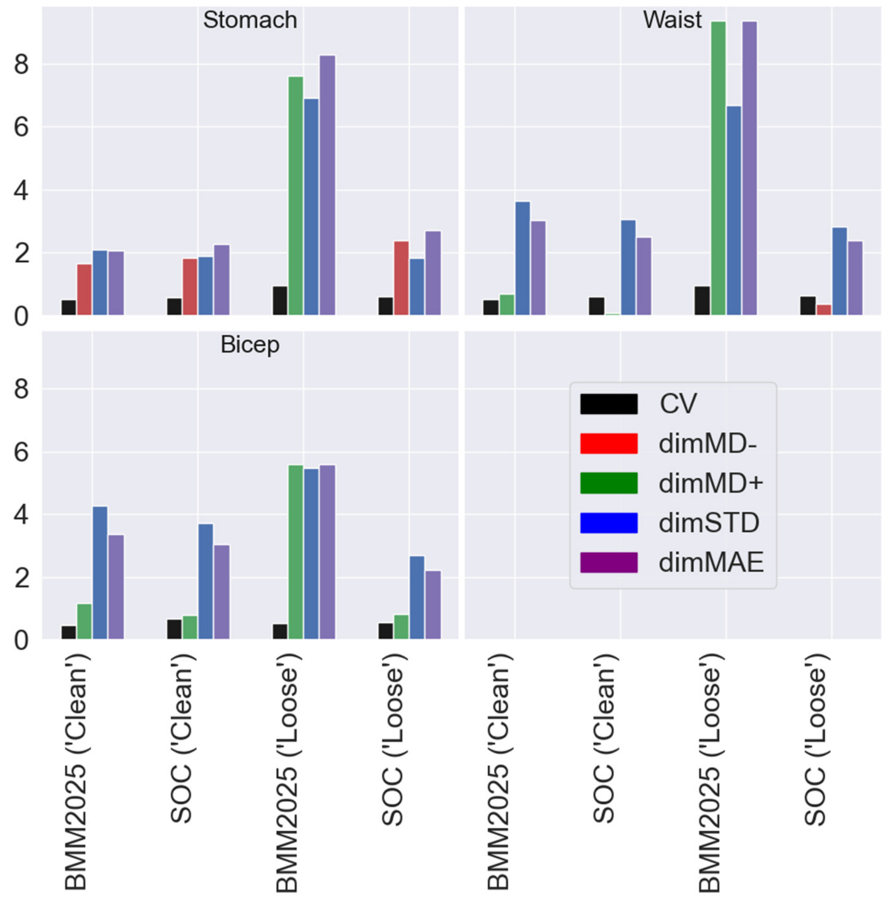

Fig. 3. Select circumferences (Stomach, Waist, and Bicep) and their associated dimensionless metrics for both sets of scans ('clean' and 'loose') and for both BMMs (`BMM2025` and `SOC`). Y-axis values are percentages.

4. Summary

We introduced a clothing-aware BMM (`SOC`) that greatly enhances accuracy compared to our current BMM (`BMM2025`) when subjects are scanned in loose clothing. The `SOC` BMM can additionally be used when subjects are wearing proper (skin-tight) scan wear with only a slight loss of accuracy (relative to `BMM2025`). This result was made possible by the development of the `BPC` segmenter whose ability to distinguish between a subject's body and their worn clothing provided the information necessary for the `SOC` BMM to learn to compensate for loose clothing.

As an additional bonus to the effort in developing the `BPC` segmenter, we have significantly improved our subject segmentation (previously handled by `S2023`). The subject segmentation with the `BPC` segmenter is more robust to cluttered scan environments and poor lighting and, when used with our current BMM (`BMM2025`), provides a significant improvement in precision and a modest improvement in accuracy to our reported measurements in ideal scanning conditions.

Future work may include an RTW sizing test. If we find we can provide reliable RTW sizing with relaxed clothing restrictions, which appears likely, we may explore further applications in the apparel and health spaces.