# RESPONSE TO THE REVIEWERS’ COMMENTS

**Real-time Virtual Fitting with Body Measurement and Motion Smoothing**

**Umut Gültepe, Uğur Güdükbay**

**(submitted to Computer-Aided Design)**

We would like to thank the anonymous reviewers for their comments and suggestions. We had a native speaker proofread the manuscript and rewrite various sentences in the manuscript to improve the presentation. We also included a new section describing how we handle the footskating problem, together with the related references. The following is a log of changes we made to address to the issues raised in the reviews.

# REVIEWER # 1

**Reviewer’s Comments:**

The virtual fitting room framework operating a depth sensor for realistic fitting experience, customized motion filters, size adjustments and physical simulation are on top in this context. The presentation part about state of the art in chapter 'Previous Work' is good. It is a question whether TOF technology is more perspective than depth sensors. The presented virtual dressing framework is well elaborated. The firmware phase defining skeleton form and joint positions based on Kinect is convincing and upon the experiments it works well. The presentation of the mathematical and algorithmical background of depth map filtering, body dimensions and collision spheres and temporal optimization are very expressive. The motion smoothing and position filtering primarily serve for the motion simulation. The presented method of bone splitting is a good idea. The presented experiment results prove the usability and perspectives of this system. I felicitate authors on their work, on the development of this powerful system. There are problems with the title. No woman would buy a dress after this particular virtual try on. I suggest using quotation marks in the title "virtual fitting". I think that a Kinect-like depth sensor is good for more than defining height and width body parameters or collision sphere radii and in the future the different material properties can be used in physic simulations.

**Authors’ Answer:**

*We thank to the referee for his constructive comments. We modified the title as suggested by the reviewer as “Real-time “Virtual Fitting” with Body Measurement and Motion Smoothing.” We also agree with the reviewer that Kinect-like depth sensors provides more than width and height body parameters and collision sphere radii, which can be used in physically-based simulations for such applications. As a future study, we mentioned in the Conclusion that with the advances in depth sensing technology, our framework could be extended to provide a more customized fitting experience.*

# REVIEWER # 2

**Reviewer’s Comments:**

This paper describes a virtual fitting room application using Kinect. Given the depth map, they estimate joint centers and apply filters and constraints to obtain smooth joint orientations. An additional process of bone splitting (sort of addition of pseudo bones) has been made in order to obtain realistic skin deformation. Existing physics engine (PhysX by Nvidia) has been adopted for the garment simulation. It presents an interesting system, but the technical contribution devoted to each component is rather insignificant. For instance, the idea of adding pseudo joints (which the authors claim as novel) has been introduced in the following reference:

"Building Efficient, Accurate Character Skins from Examples", Alex Mohr and Michael Gleicher, SIGGRAPH 2003.

**Authors’ Answer:**

*We thank to the referee for his constructive comments. We would like to note that the proposed system employs numerous motion filtering and animation techniques, which are not available in the commercially available virtual fitting room suites with time of flight cameras.*

*We investigated the work by Mohr and Michael, which indeed proposes a similar approach to ours.* ***We removed our claim that the idea of adding pseudo joints is novel by citing the suggested reference.*** *Although both approaches provide the similar functionalities, the starting point and application areas are different, as they aim to solve a design stage problem through extrapolation of example postures whereas we aim to solve an animation stage problem through interpolation. We discuss the similarities and differences with this reference in the related section.*

**Reviewer’s Comments:**

I have some technical questions, too:

- At the end of Section 3, you say the rotation of the user is transferred to the avatar naturally "with proper weights": How are the weights determined? Do you mean per-vertex joint influence weight?

**Authors’ Answer:**

*The weights we mentioned were per-vertex joint influence weights as the reviewer predicted correctly. The rotation of the user's arm is transferred to the virtual avatar naturally without any artifacts by the use of proper weights; these weights are derived from the joint-vertex influence weights automatically assigned to a single bone and distributed linearly between the elbow and wrist joints. We modified the related part at the end of Section 3 as described here with a reference to the Blender Software Suit.*

**Reviewer’s Comments:**

- How are the clothes resized according to the body parameters that are extracted? Although in abstract you mentioned that "The proposed scaling method ... determines a specific apparel size", the way how you determine a garment size is not described in the manuscript.

**Authors’ Answer:**

*The clothes are resized according to the extracted body parameters as follows: the garment sizes we referred to would be the standard cloth sizes, from XS to XXL. Although there are standardized cloth sizes through Europe, Asia, USA and the rest for the world, there is not a strict regulation; hence, the exact size ranges would be determined according to the specific company and apparel piece. We use a generic chart provided by a sports goods manufacturer.*

*In order to clarify the resizing process, we elaborate on how we determine the garment size in the related section on body measurement and we modify the abstract by adding the phrase “… determines a standardized apparel size according to user’s measurements” in the revised version of the article. We also by refer to the Sports Goods Manufacture web site that provided the generic chart we used for the resizing process.*

**Reviewer’s Comments:**

A minor comment:

Section 3.2.3: A more appropriate title would be "Temporal averaging" (instead of "Temporal optimization")

**Authors’ Answer:**

*We replace the title of Section 3.2.3 as “Temporal Averaging” as suggested.*

# REVIEWER # 3

**Reviewer’s Comments:**

The paper presents a visual fitting room framework by using depth sensor data, which has applications for apparel shops and online stores. The proposed framework is real time with one second preprocessing time. It mainly includes the techniques of physics engine, body measurements, and motion smoothing. The cloth simulation in the framework is handled by PhysX's cloth simulation engine. The body parameters (Body height, width, and radii) are captured for displaying the cloth onto a virtual avatar reflecting the body characteristics of the user. Motion filters avoid unnatural results by smoothing the depth data and estimate the self-occluded parts. The quality of body parameters and motion capture affects the accuracy of fitting and animation results.

The paper involves many aspects and provides a complete cloth fitting framework. The algorithms proposed are explained in detail with pseudo-code provided.

**Authors’ Answer:**

*We thank to the referee for his constructive comments.*

**Reviewer’s Comments:**

Nevertheless, some aspects of this paper should be improved:

1. This paper lacks the comparisons with the prior work. Both theoretically and experimentally, this paper does not clarify the advantage of the proposed method in comparison with the state-of-art. The key difference between the proposed approach and existing cloth fitting methods introduced in Section 2 is not clear. It is hard to judge either objectively or subjectively the effectiveness of the proposed approach. Table 3 alone is not sufficient. More visual and animation comparisons are expected.

**Authors’ Answer:**

*We described the distinction between our approach and previous works in the beginning of chapter 3, including the goal and the details of the proposal. We provided the frame rates of simulations for two different apparel meshes and explain the reasons for the patterns in the graphs in the Experiments section. We also provide a figure showing the corrected foot displacements.*

*As most of the previous works in this field focus on content creation rather than simulation, the availability of reference data for comparison is low. We were able to find a similar study and included the comparison in Table 3.*

2. The limitations of the proposed method are not mentioned in the paper. In my opinion, one limitation is that the cloth industry is more concerning about the customization. However, the standard-sized apparel fitting without customization highly limits the applied scope of the proposed method.

**Authors’ Answer:**

*The limitations are now discussed in the conclusion section; the most important ones being the customization and small simulation area problems, together with the possible ways to improve the framework to overcome these limitations.*

**Reviewer’s Comments:**

3. Parameters of rendering and detail of meshes should be provided.

**Authors’ Answer:**

*We include a FPS graph in various stages of the simulation in experiments chapter. We also added screenshots of different apparels along with the detail level of all meshes used in the system (in vertex and triangle count) which can be found in Table 6.*

**Reviewer’s Comments:**

4. In Section 3.2.1, it would be necessary to provide experimental results to support the efficiency of the depth map filtering method.

**Authors’ Answer:**

*We provide a graph and a screenshot about the experimental results of depth map fıltering method and interpret the results in a new paragraph in Section 3 to support the proposal’s efficiency.*

**Reviewer’s Comments:**

5. The writing needs revision in some places.

-Eq(1): The symbol 'i' ,'n', and 'd' in the equation are not explicitly explained in the text.

-The caption of Figure 5 should be rephrased. "row" -> "raw" ". for right the humerus roll angle" ?

-Section 3.2.2: "effect others in the estimation process and vice versa" -> "affect others in the estimation process and vice versa"

**Authors’ Answer:**

*We applied all the revisions suggested for improving the presentation. We also had a native speaker proofread the manuscript and rewrite various sentences in the manuscript to improve the presentation.*