

How Will It Drape Like? Capturing Fabric Mechanics from Depth Images Supplementary Material

Carlos Rodriguez-Pardo^{1,2,3} , Melania Prieto-Martin² , Dan Casas² , Elena Garces^{1,2} 

¹SEDDI, Madrid, Spain

²Universidad Rey Juan Carlos, Madrid, Spain

³Universidad Carlos III de Madrid, Spain

1. Overview

We divide this supplementary material in different sections, as follows:

- On Section 2, we provide additional implementation details on our model design and training configuration.
- On Section 3, we provide additional information on the procedure we followed on our user study.
- On Section 4, we show the results of our ablation study on our distance metric.
- On Figure 8, we provide additional details on our user study, particularly the demographics of the test.
- On Figure 1, we show the correlation between distances on the parameter space, and on our tSTE embedding, as well as our Drape Similarity Metric.
- On Figure 2, we show some examples of our simulation-space data augmentation policy, aimed at recreating possible misalignments that may be present on the real material capture setup.
- On Figure 3, we show the top-K classification accuracy of the rankings computed on the parameter space, and through our drape similarity; compared to the tSTE embedding rankings.
- On Figures 4 to 7, we show the relative similarity rankings computed through the tSTE embedding of our user study, compared to the ground truth and predicted parameter distances, as well as the drape similarity computed using the ground truth and predicted parameters.
- On Figures 10 to 14, we show the specific characteristics of each material in our test set, including their composition, structure and a visualization of their texture on both sides of the material (a 1×1 cm image).
- On Figures 15 to 24, we show per-material Neural Saliency Maps, for every material in our test set, divided by scene and target parameter.
- On Figure 35, we show comparisons of real and simulated garments, for two different fabrics.
- On Figures 25 to 34, we show 10 simulations done using the Ground Truth parameters and the parameters predicted by our model. It can be seen that the simulations are highly non-deterministic.
- On Figures 36 to 45, we show the relative orderings obtained by different methods, from left to right: The average per-parameter distance, the distance captured by the embedding of our user study, and the perceptual distance, obtained using the ground truth simulations, and the simulations done using the parameters estimated by our model.

2. Implementation Details

Training. We train our model using PyTorch [PGM*19] as our learning framework and Kornia [RMP*20] for image-processing operations, including normalization and data augmentation. The model is optimized using AdamW [LH17] for a total of 50 epochs, with a learning rate of $\alpha = 0.0003$, which is halved every 10 iterations. All other optimizer hyperparameters followed the default AdamW configuration. We use mixed precision training [MNA*18], a batch size of 256, and a weight decay of 0.000002. The input images are of 180×180 pixels, and we use a Mean Squared Error (MSE, ℓ_2) as our loss function for our regression problem. Model size design and hyperparameter selection were conducted using Bayesian hyperparameter tuning [Bie20]. Training takes approximately one hour on a NVIDIA 2080 RTX GPU.

Network Design. Our feature extractor is a pre-trained ResNet-18 [HZRS16] (TorchVision pretrained weights checkpoint `weights='IMAGENET1K_V1'`), followed by a Self-Attention layer [ZGMO19] with 64 attention heads. The output of this layer is max and average-pooled, and a final multi-layer perceptron (MLP) with 4 hidden layers with 512 hidden units each, receives the pooled features and the density value, and outputs the 6 predicted mechanical parameters. The MLP layers use ReLU [MHN13] non-linearities, followed by Layer Normalization [BKH16], and a Dropout [SHK*14] rate of 0.5. Both the feature extractor and the MLP are trained simultaneously.

Data Augmentation. We use the following policy, in order: First, we randomly rescale the input images using (0.5, 1.5) rescale ranges. We then randomly erase [ZZK*20] parts of the input images ($p = 0.2$), randomly rotate them ($p = 0.2, angle = \pm 10^\circ$), apply a random perspective change ($p = 0.5, scale = 0.5$), and a random thin plate spline warp ($p = 0.2, scale = 0.5$). We then crop the center area, with a resolution of 180×180 pixels. To those images, we randomly change their contrast and brightness, on the (0.5, 1.5) ranges. We then apply a final set of transformations, including image equalization, random horizontal flip ($p = 0.5$), random Gaussian noise ($p = 0.5, \mu = 0, \sigma = 0.02$), random posterization ($p = 0.5$), random sharpening ($p = 0.5$), and random Gaussian blur ($p = 0.5$, kernel size of 3×3). Please refer to the Kornia documentation [RMP*20] for specific implementation details of each of these transformations and the specific meaning of the mentioned parameters.

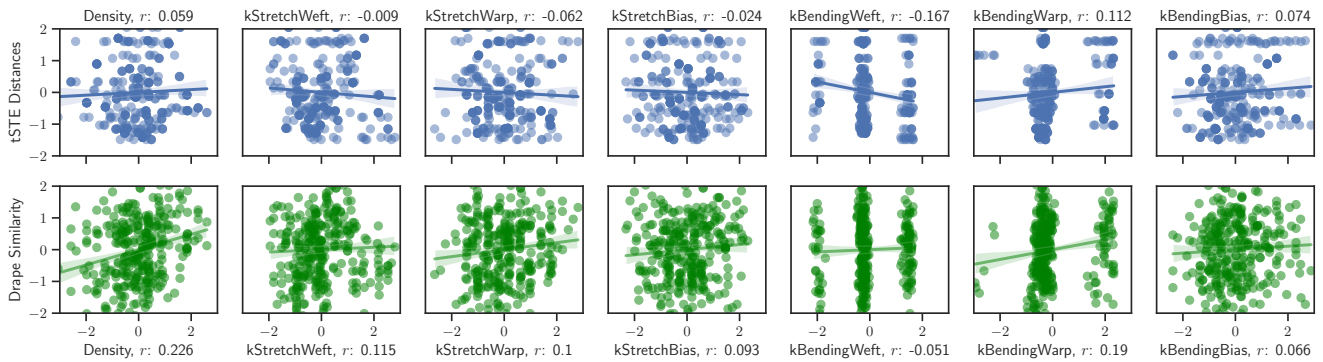


Figure 1: Relationship between the distance obtained by our user study, our distance metric and the difference in each parameter. On the top row, we plot the difference between the value of a parameter against the average distance between fabrics we obtained through our user study. On the bottom, we show the same parameter distances, now against the distance predicted through our perceptual metric. As shown, no mechanical parameter dominates the distance predicted by either the user study or our metric. The perception of mechanical properties can be understood as a highly non-linear phenomenon in which many parameters interact in complex ways. We use z-scores to help visualization.

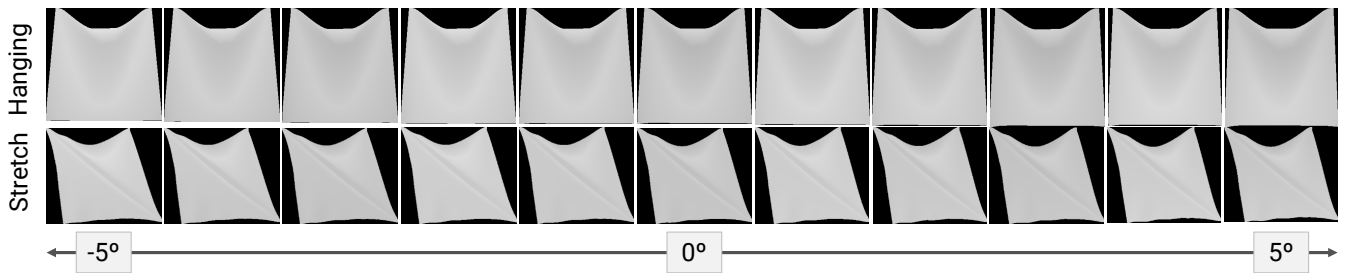


Figure 2: To recreate possible misalignments that may be present on the real images, we propose a simulation-space data augmentation policy, in which we simulate the drapes from different rest positions, on the $[-5, 5]$ degree ranges. For every scene, we generate 11 different simulations uniformly in this range, effectively creating a dataset of images for a single material.

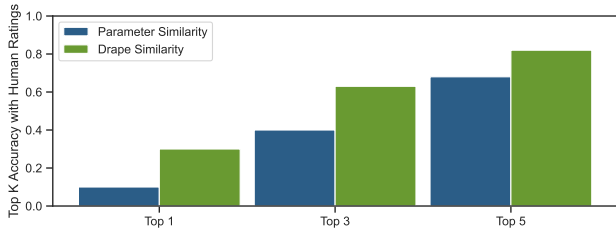


Figure 3: Ranking classification accuracy on different ways of comparing materials. Comparing materials simply by measuring their distances on parameter space yields inaccurate similarity rankings, as the parameter space is not orthogonal and cannot be linearly correlated with human perception of material similarity. Our Drape similarity metric provides us with a way of assessing the similarity between materials on a more perceptual-aware space, yielding more accurate rankings.

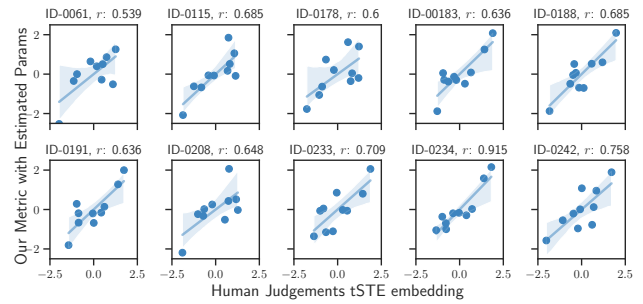


Figure 6: Correlation between the ordering provided by the tSTE embedding from our user study (x-axis) and our distance metric, computed using the estimated parameters (y-axis). We plot z-scores instead of the raw distances to help visualization.

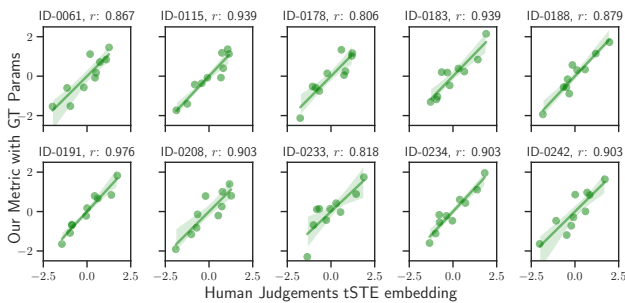


Figure 4: Correlation between the ordering provided by the Human Judgments (x-axis) and our Drape Similarity Metric (y-axis). We plot z-scores instead of the raw distances to help visualization.

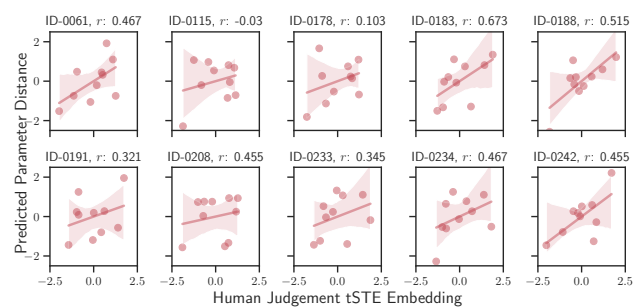


Figure 7: Correlation between the ordering provided by the tSTE embedding from our user study (x-axis) and the distance computed on the parameter space, for the model predictions (y-axis). We plot z-scores instead of the raw distances to help visualization.

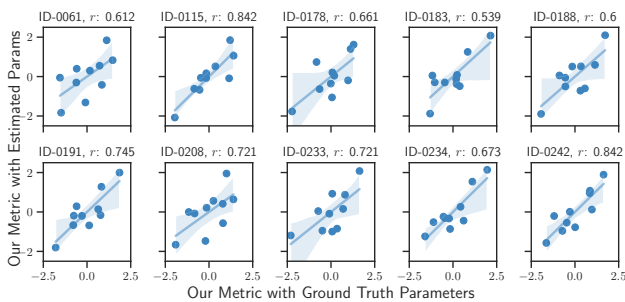


Figure 5: Correlation between the ordering provided by our similarity metric with the Ground Truth (GT) parameters (x-axis) and the estimated ones (y-axis). We plot z-scores instead of the raw distances to help visualization.

3. User Study: Procedure and Stimuli

Stimuli. We use ten physical fabric samples of our real dataset, which are representative of different types of fabric structures, thickness, densities, and mechanical properties. We leverage 25×25 cm samples, which is an adequate size for the manipulation of the materials.

Participants. A total of 30 volunteers took part in the test. The participants had different backgrounds and diverse levels of expertise on physical manipulation of textiles. A study on their demographics can be found in the Supplementary Material. Figure 8.

Procedure. Following previous work on human perception [ZIE*18, GAGH14], we design a Two-Alternative Forced Choice (2AFC) user study in which participants are presented with triplets of physical fabric samples. We performed the experiment in a lab-controlled environment because the fabrics samples had to be physically manipulated. Participants were asked to select which of two fabrics is most similar to a reference fabric. They were suggested to manipulate the samples (stretching, bending them, etc.) and to focus only on mechanical similarity, thus ignoring other factors such as optical properties (color, specularity, transparency) or irrelevant tactile feedback (softness, for instance). Material reflectance was thus ignored by participants. Volunteers were not instructed to leverage fabric motion to make their decisions but were free to dynamically manipulate the samples, which some participants did. Following recommendations in user study design for graphics [BHH*22, DSKP22], we informed the users that there are no right or wrong answers, allowed them to use their own criteria for answering the questions, and self-identify their demographics and levels of expertise. The volunteer demographics can be found on 8.

Each participant rated 20 triplets that were pseudo-randomly sampled following some rules: each of the 10 fabrics in our test set is shown as reference at least twice to each participant, the order of the triplets is random, each triplet was evaluated by at least 2 participants. Each test took between 15 to 30 minutes, depending on the participant, and was followed by an open discussion, to better understand which criteria each participant used to make their decisions. We started each quiz with a brief interview to register the self-reported participants demographics [DSKP22], we continued it with the 20 perceptual comparisons, and finished it with an open discussion so as to better understand which factors each participant used to make their decisions.

Results: Agreement 30 participants, with diverse levels of expertise with fabrics and computer simulations, volunteered for this user study. Given the same triplet, we observe an average of 86.68% agreement between our participants across all experiments and materials, suggesting that there is a perceptual understanding of fabrics mechanics that humans share. We did not observe any significant differences in agreement depending on the volunteer demographics, or level of expertise on fabric handling or simulation. This suggests that the perception of similarity on the mechanical properties of textiles may be relatively universal. When prompted, participants generally admitted that some triplets were easier to evaluate than others, either because neither candidate fabrics were similar to

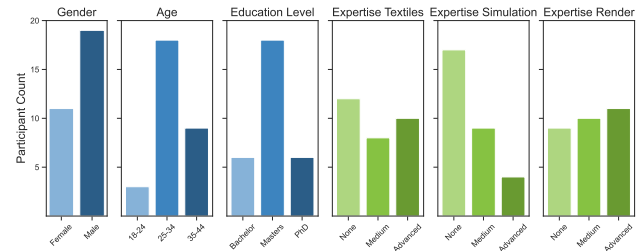


Figure 8: Demographics of the participants of our perceptual study. Following recommendations on study for computer graphics [DSKP22, BHH*22], we ask for capture the self-identified gender, age and education level of each participant. We also ask them to identify their level of expertise with handling physical textile materials, as well as their expertise on computer simulation and rendering of fabrics.

the reference or because both were very similar and thus difficult to make a definitive decision. Participants tended to evaluate elasticity and the shape of the fabric when it bends. When in doubt about those factors, a standard procedure was to assess similarity using the perceived weight of the material.

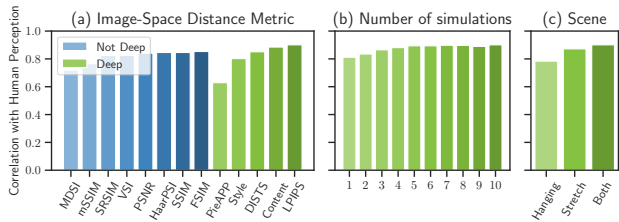


Figure 9: Ablation study on different factors of our similarity metric. Y-axis is Spearman correlation. (a) Results for several image-based metrics *IM* (we use the implementations provided in PIQ [KZP19]). (b) Number of simulations needed to account for the non-determinism (optimal choice is 5). (c) Impact of the scene.

4. Ablation Study: Similarity Metric Parameters

Here, we study the factors that might impact the performance of the metric: the image-based similarity metric (*IM*), the number of simulations necessary to account for the non-determinism of the simulation (*N*), and the scene (*hanging* or *stretch*).

Choice of *IM* The image-based metric should be able to capture the subtle differences in wrinkles, folds, and overall shape of the simulated drape. While previous work has learned this metric from data [GAGH14, LMS*19], we proposed a simpler approach using existing image-based similarity metrics. Figure 9 (a) shows results for several deep and non-deep methods. For deep learning-based metrics, it is particularly relevant that the metric can account for local and positional differences. A *Content loss* or the *LPIPS* shows stronger performance than translation-invariant losses (e.g. *Style Loss* or *DISTS*).

Choice of *N* We study how many simulations that are needed to account for the non-determinism of the simulation. Figure 9 (b) shows that $N \geq 5$ is enough to obtain accurate results.

Choice of scene We study which scene is best for estimating drape similarity in Figure 9 (c). The *stretch* scene generally provides more information than the *hanging* scene, while using both achieves the best agreement overall.

References

[BHH*22] BYLINSKII Z., HERMAN L., HERTZMANN A., HUTKA S., ZHANG Y.: Towards better user studies in computer graphics and vision. *arXiv preprint arXiv:2206.11461* (2022). 5

[Bie20] BIEWALD L.: Experiment tracking with weights and biases, 2020. Software available from wandb.com. URL: <https://www.wandb.com/>. 2

[BKH16] BA J. L., KIROS J. R., HINTON G. E.: Layer normalization. *arXiv preprint arXiv:1607.06450* (2016). 2

[DSKP22] DODIK A., SELLÁN S., KIM T., PHILLIPS A.: Sex and gender in the computer graphics research literature. *arXiv preprint arXiv:2206.00480* (2022). 5

[GAGH14] GARCES E., AGARWALA A., GUTIERREZ D., HERTZMANN A.: A similarity measure for illustration style. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 1–9. 5, 6

[HZRS16] HE K., ZHANG X., REN S., SUN J.: Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 770–778. 2

[KZP19] KASTRYULIN S., ZAKIROV D., PROKOPENKO D.: PyTorch Image Quality: Metrics and measure for image quality assessment, 2019. Open-source software available at <https://github.com/photosynthesis-team/piq>. URL: <https://github.com/photosynthesis-team/piq>. 6

[LH17] LOSHCHILOV I., HUTTER F.: Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101* (2017). 2

[LMS*19] LAGUNAS M., MALPICA S., SERRANO A., GARCES E., GUTIERREZ D., MASIA B.: A similarity measure for material appearance. *ACM Transactions on Graphics (TOG)* 38, 4 (2019), 1–12. 6

[MHN13] MAAS A. L., HANNUN A. Y., NG A. Y.: Rectifier nonlinearities improve neural network acoustic models. In *Proceedings on the International Conference on Machine Learning* (2013), vol. 30, Citeseer, p. 3. 2

[MNA*18] MICIKEVICIUS P., NARANG S., ALBEN J., DIAMOS G., ELSSEN E., GARCIA D., GINSBURG B., HOUSTON M., KUCHARIEV O., VENKATESH G., ET AL.: Mixed precision training. In *International Conference on Learning Representations* (2018). 2

[PGM*19] PASZKE A., GROSS S., MASSA F., LERER A., BRADBURY J., CHANAN G., KILLEEN T., LIN Z., GIMELSHEIN N., ANTIGA L., ET AL.: Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* 32 (2019), 8026–8037. 2

[RMP*20] RIBA E., MISHKIN D., PONSÁ D., RUBLEE E., BRADSKI G.: Kornia: an Open Source Differentiable Computer Vision Library for PyTorch. In *Winter Conference on Applications of Computer Vision* (2020). URL: <https://arxiv.org/pdf/1910.02190.pdf>. 2

[SHK*14] SRIVASTAVA N., HINTON G., KRIZHEVSKY A., SUTSKEVER I., SALAKHUTDINOV R.: Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* 15, 1 (2014), 1929–1958. 2

[ZGMO19] ZHANG H., GOODFELLOW I., METAXAS D., ODENA A.: Self-attention generative adversarial networks. In *International conference on machine learning* (2019), PMLR, pp. 7354–7363. 2

[ZIE*18] ZHANG R., ISOLA P., EFROS A. A., SHECHTMAN E., WANG O.: The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 586–595. 5

[ZZK*20] ZHONG Z., ZHENG L., KANG G., LI S., YANG Y.: Random erasing data augmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)* (2020). 2

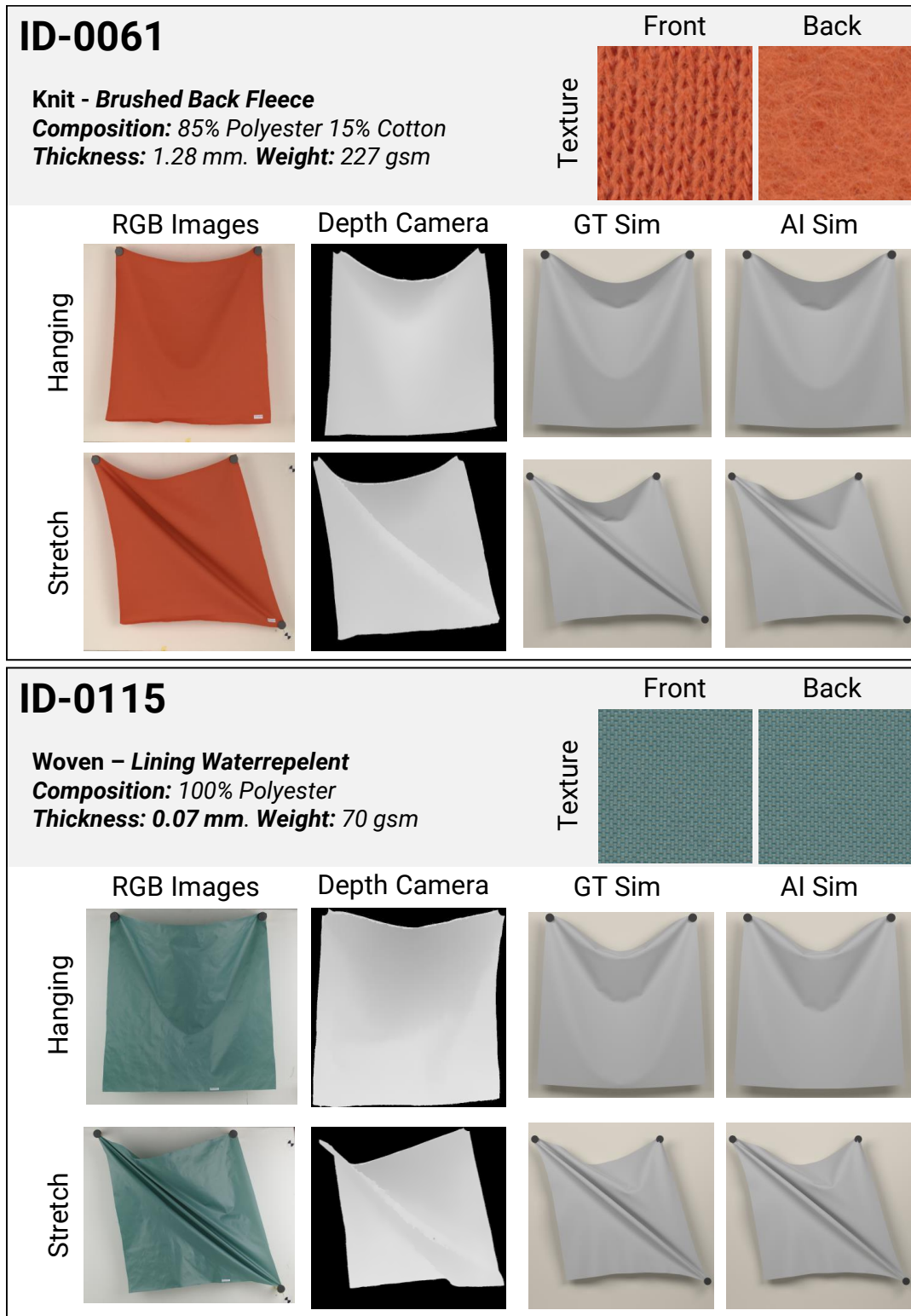


Figure 10: Characteristics of the materials ID-0061 and ID-0115 of our test set.

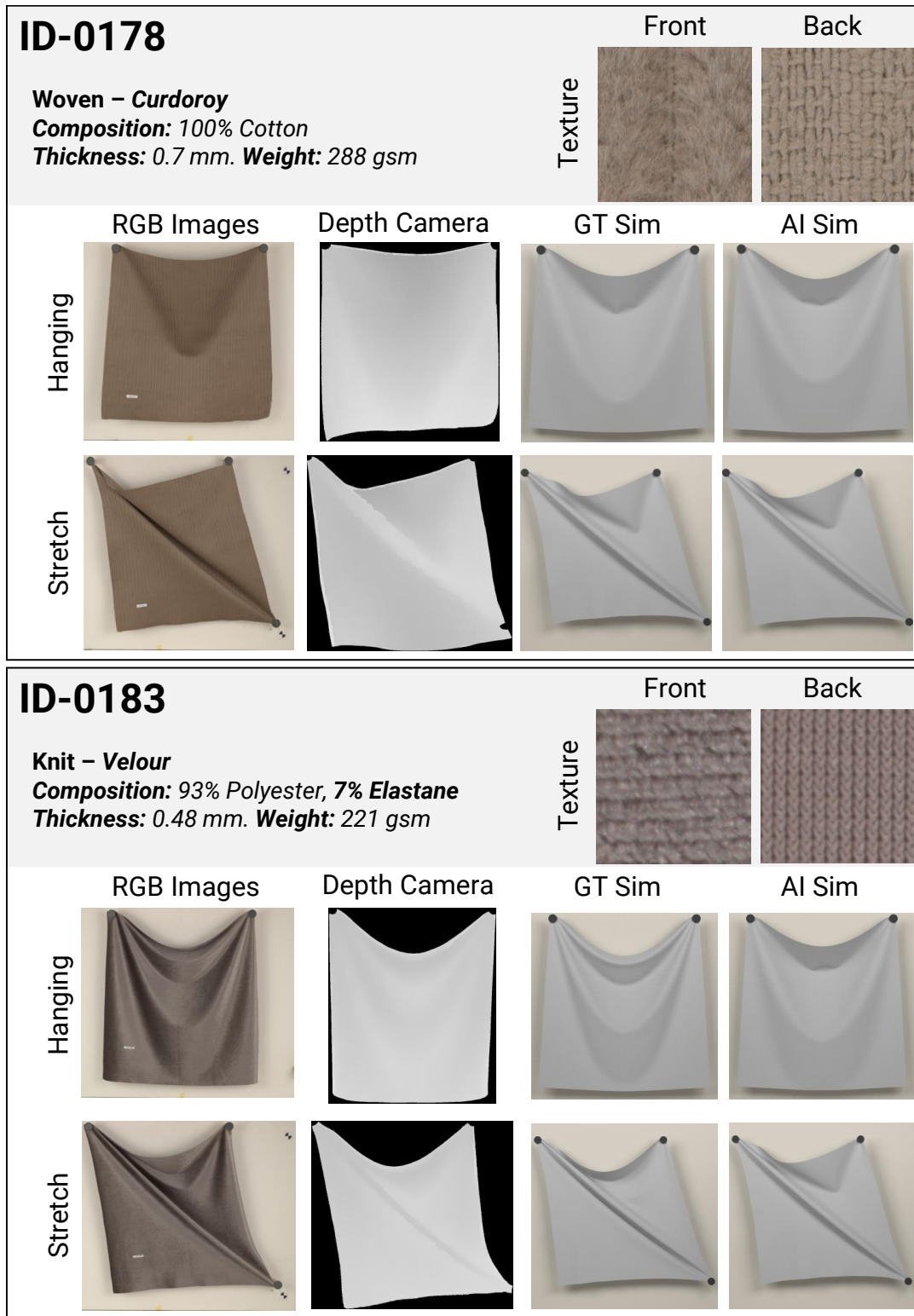


Figure 11: Characteristics of the materials ID-0178 and ID-0183 of our test set.

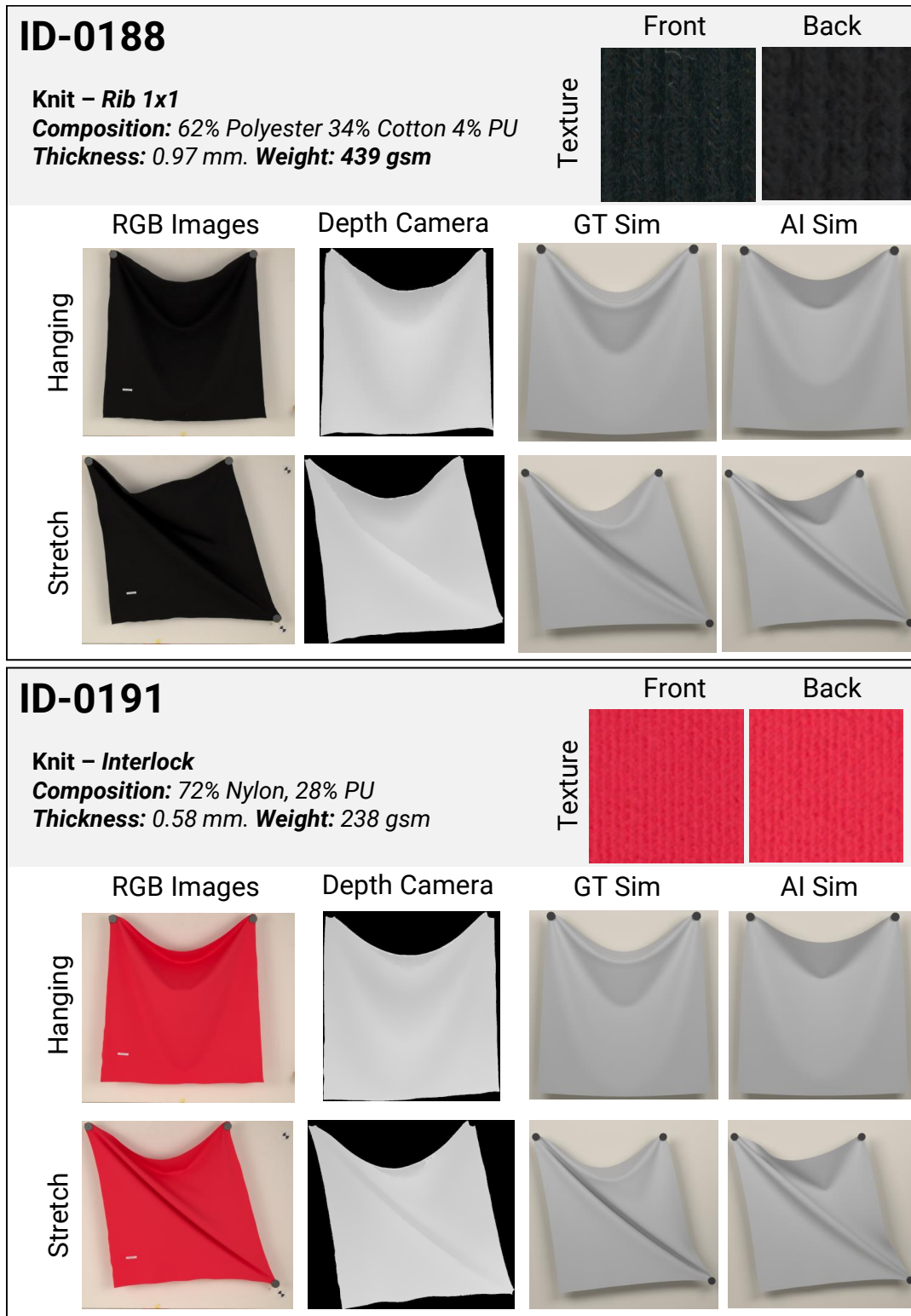


Figure 12: Characteristics of the materials ID-0188 and ID-0191 of our test set.

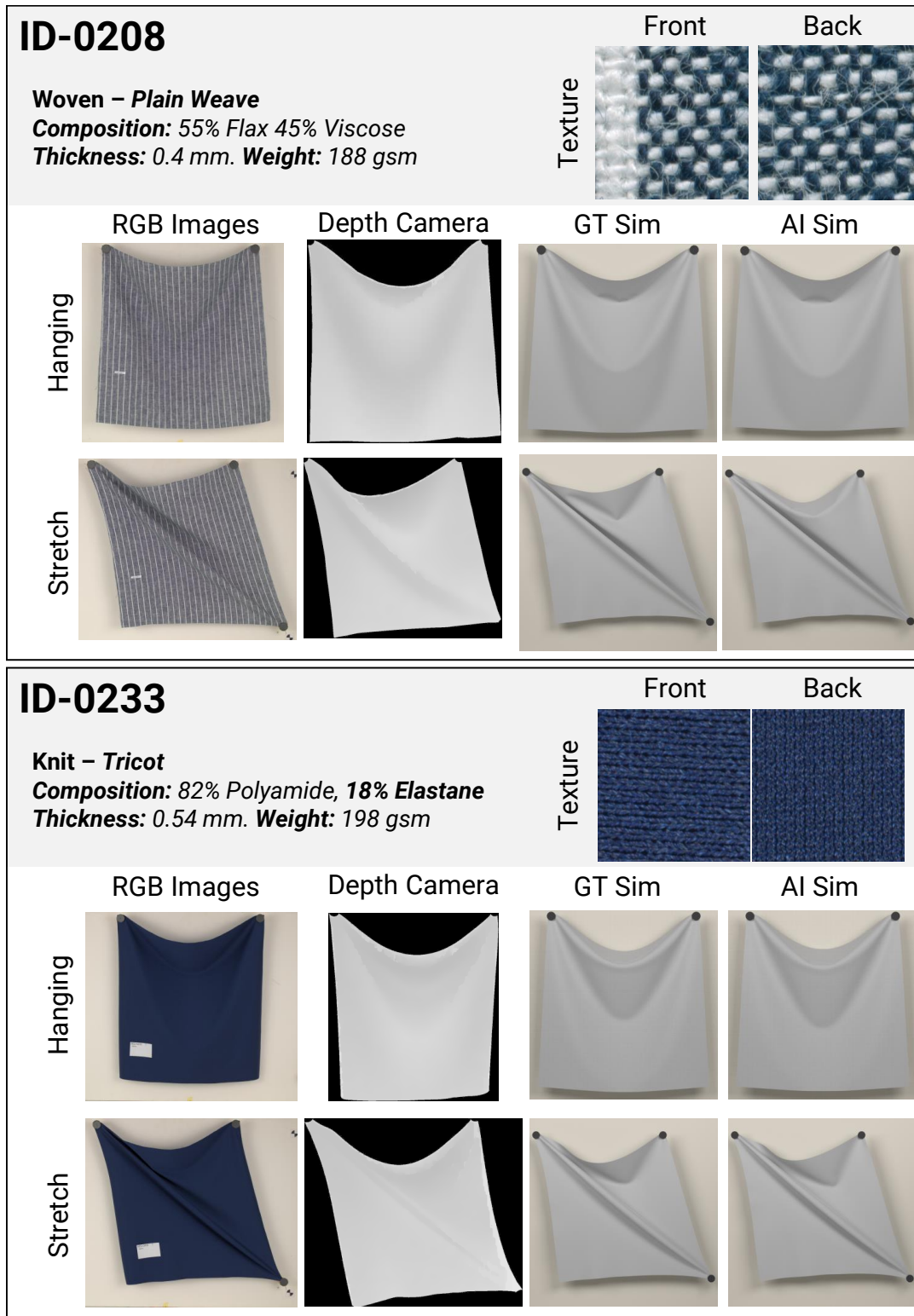


Figure 13: Characteristics of the materials ID-0208 and ID-0233 of our test set.

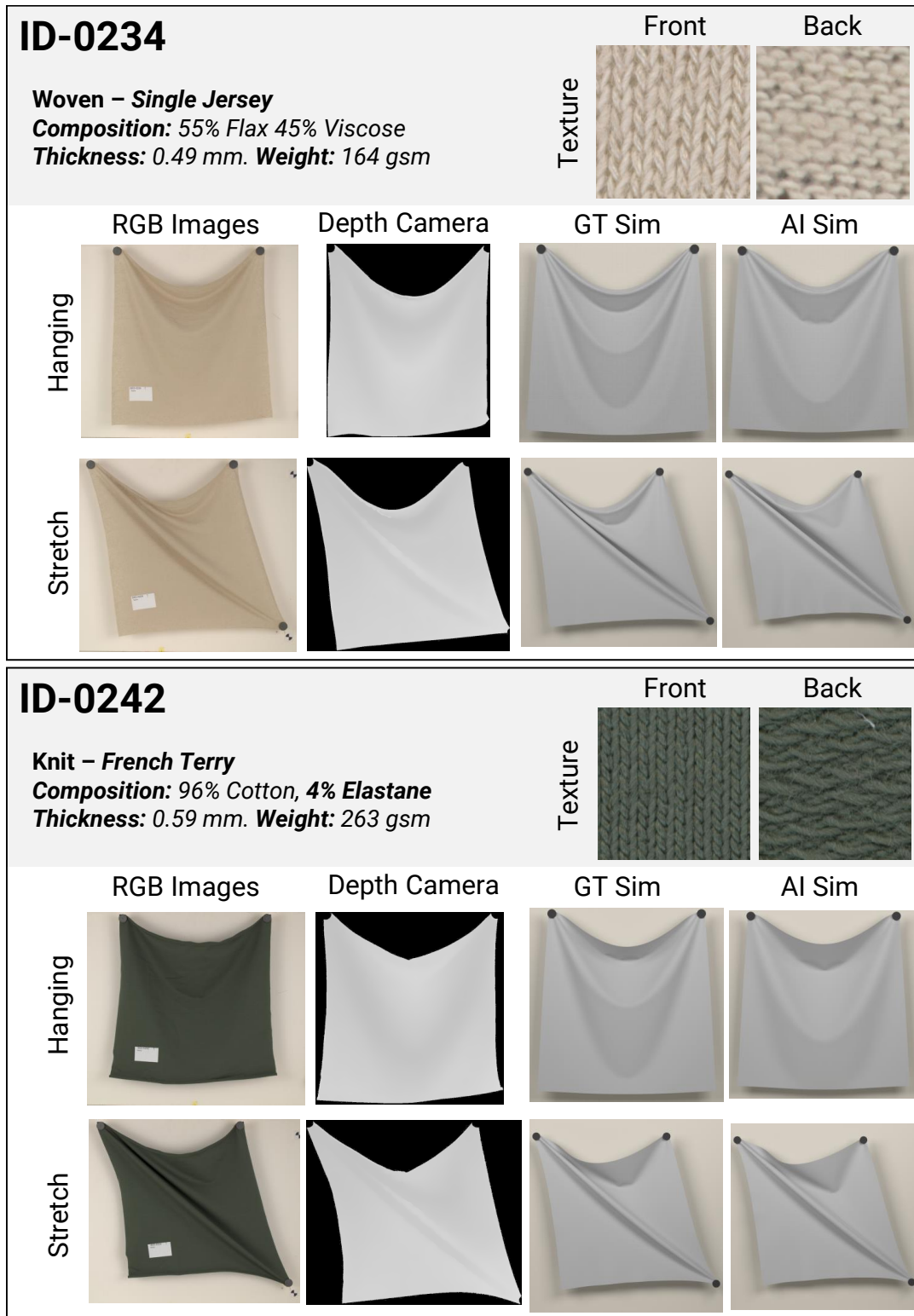


Figure 14: Characteristics of the materials ID-0234 and ID-0242 of our test set.

ID-0061

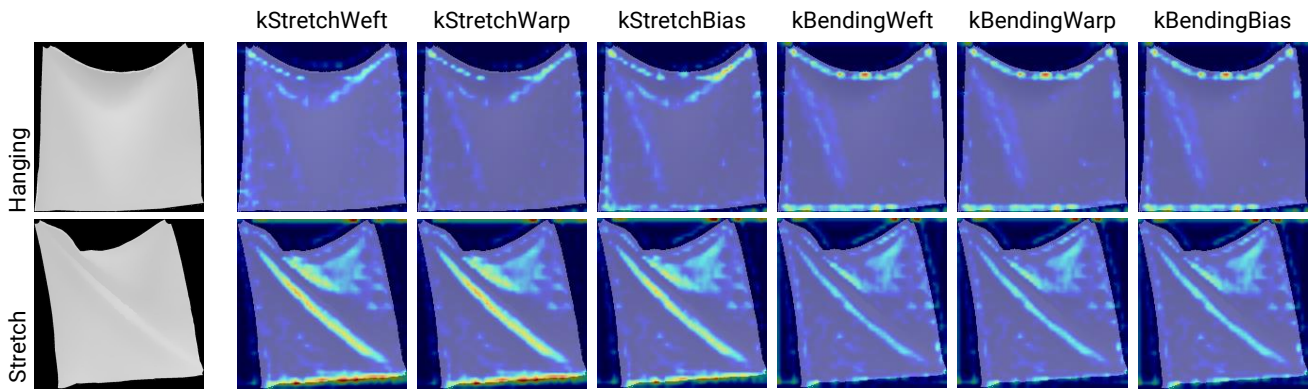


Figure 15: Neural saliency maps for the fabric ID-0061 of our test set, separated by scene and parameter.

ID-0115

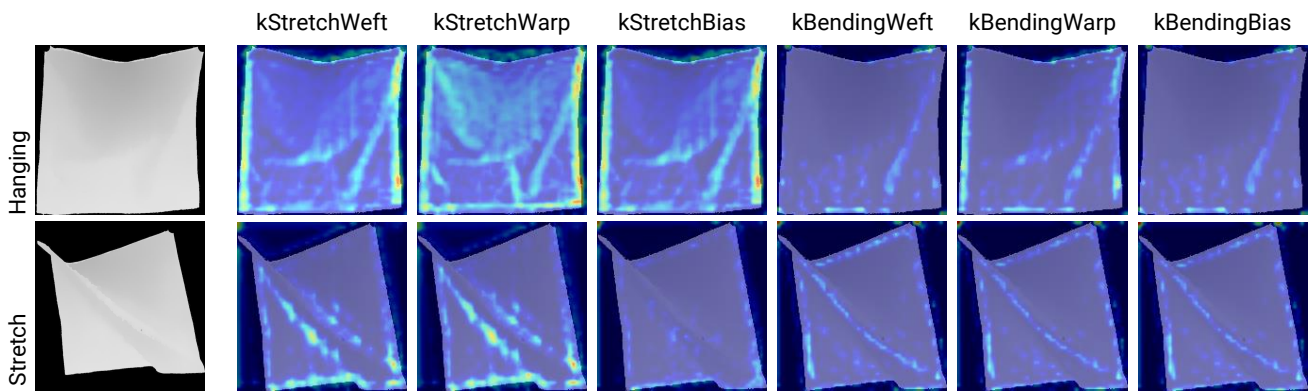


Figure 16: Neural saliency maps for the fabric ID-0115 of our test set, separated by scene and parameter.

ID-0178

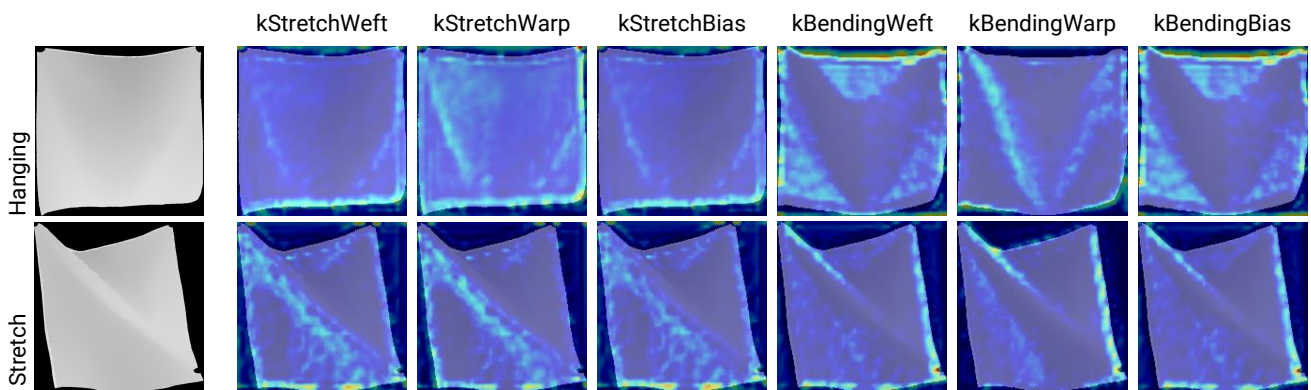


Figure 17: Neural saliency maps for the fabric ID-0178 of our test set, separated by scene and parameter.

ID-0183

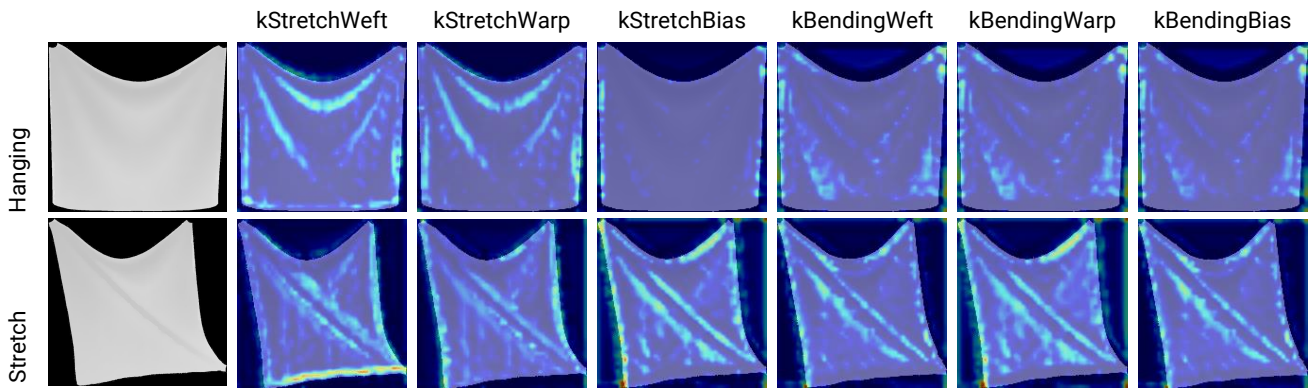


Figure 18: Neural saliency maps for the fabric ID-0183 of our test set, separated by scene and parameter.

ID-0188

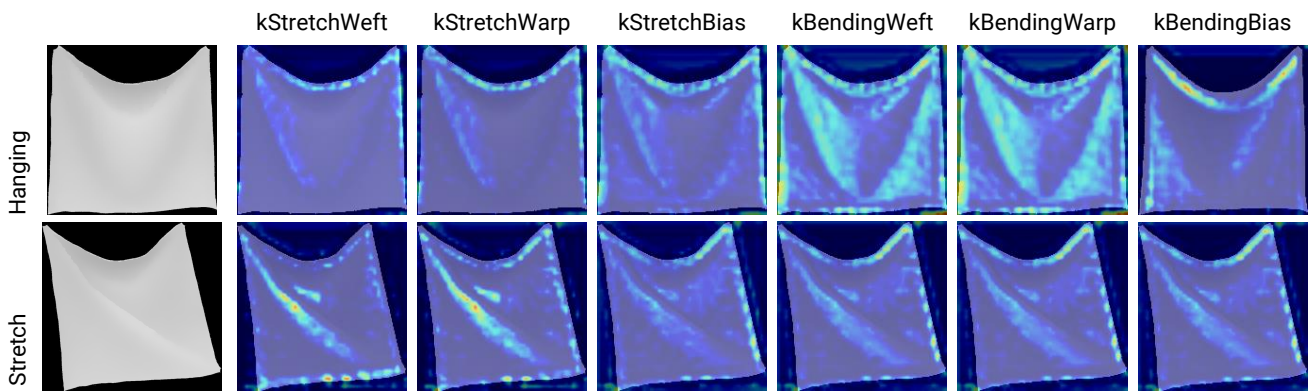


Figure 19: Neural saliency maps for the fabric ID-0188 of our test set, separated by scene and parameter.

ID-0191

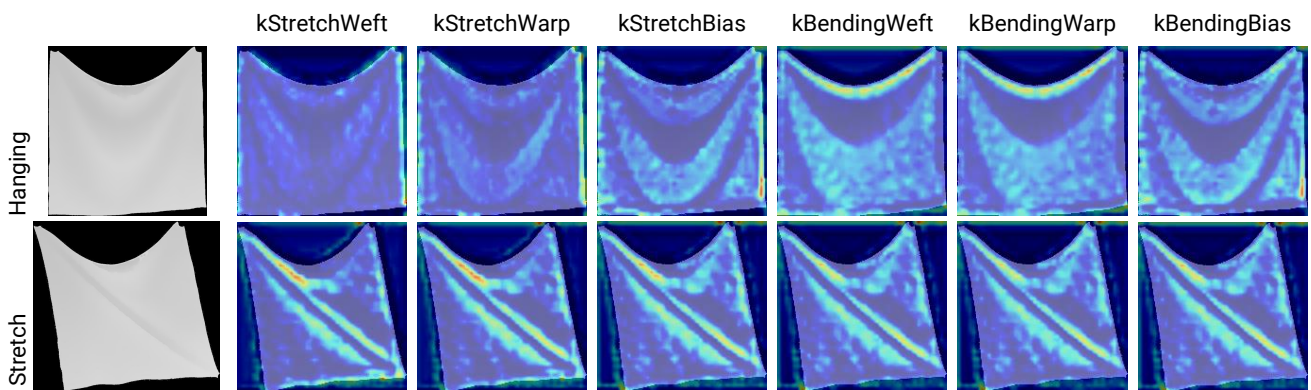


Figure 20: Neural saliency maps for the fabric ID-0191 of our test set, separated by scene and parameter.

ID-0208

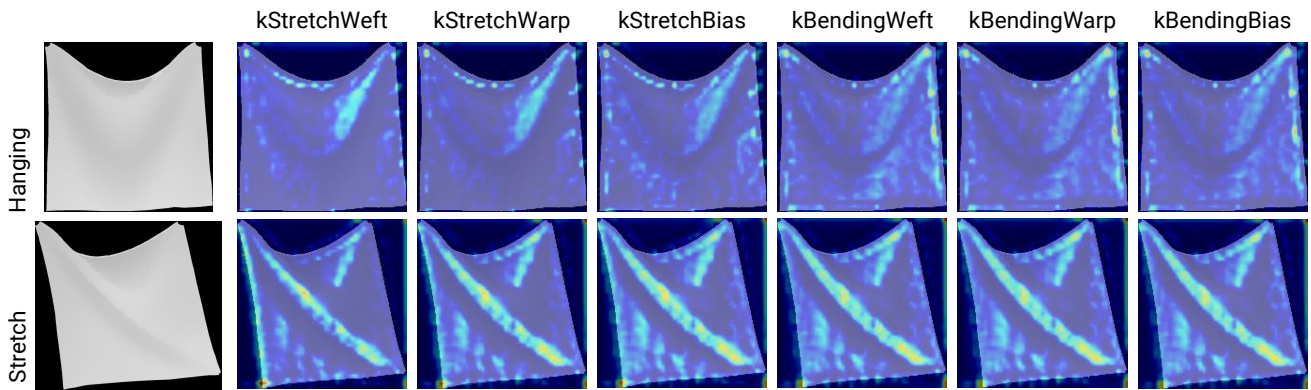


Figure 21: Neural saliency maps for the fabric ID-0208 of our test set, separated by scene and parameter.

ID-0233

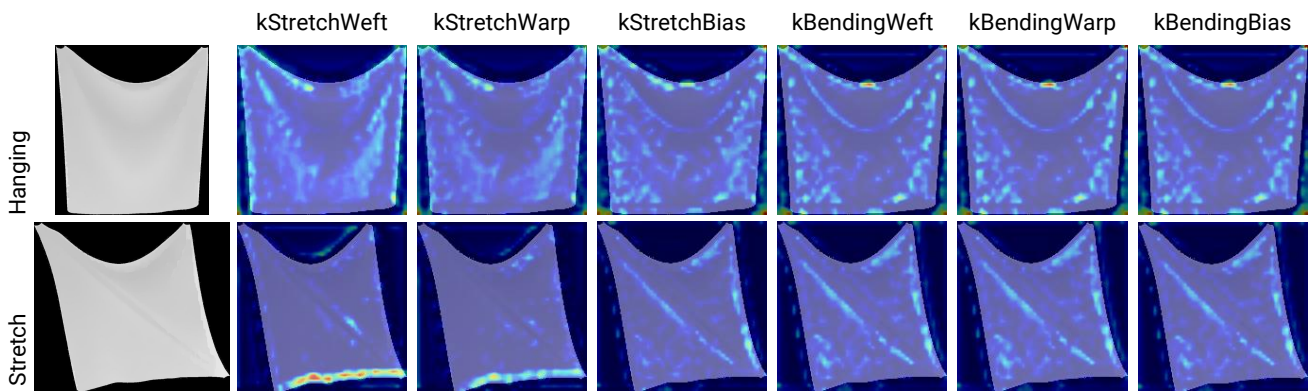


Figure 22: Neural saliency maps for the fabric ID-0233 of our test set, separated by scene and parameter.

ID-0234

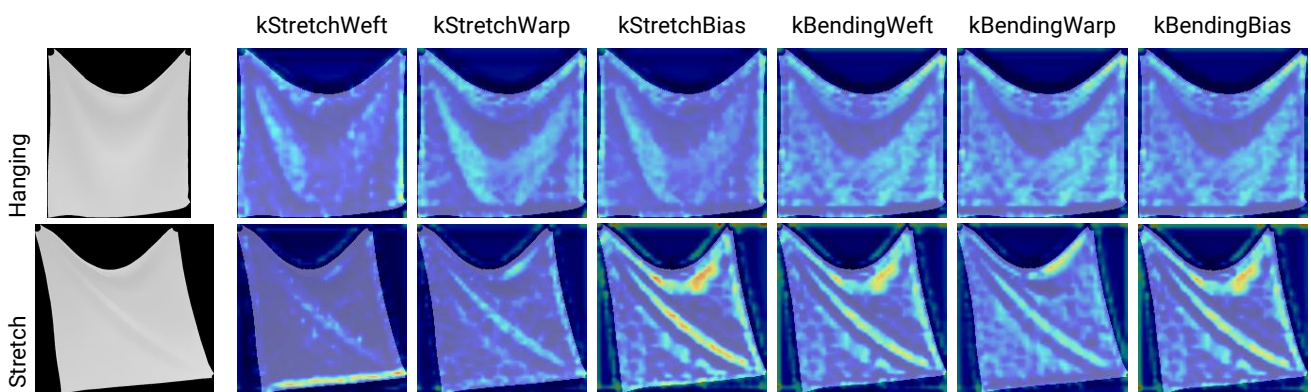


Figure 23: Neural saliency maps for the fabric ID-0234 of our test set, separated by scene and parameter.

ID-0242

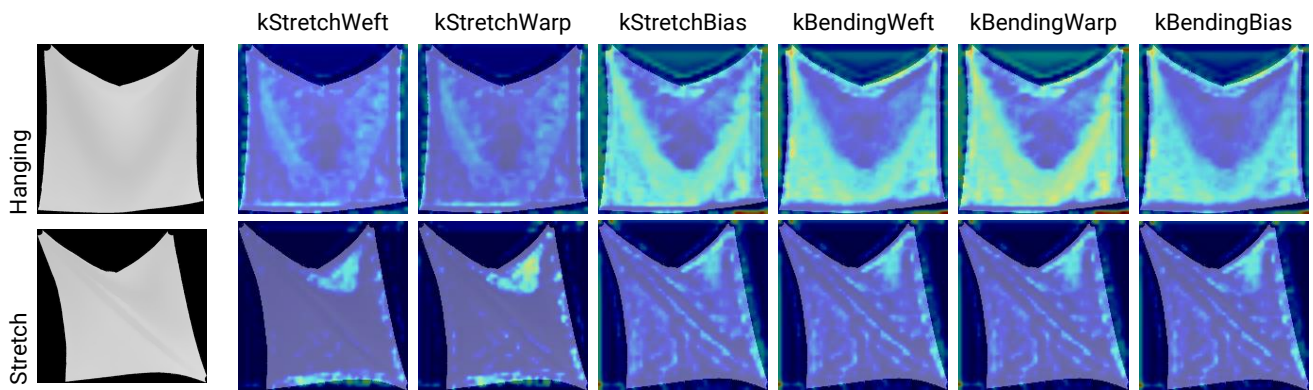


Figure 24: Neural saliency maps for the fabric ID-0242 of our test set, separated by scene and parameter.

ID-0061

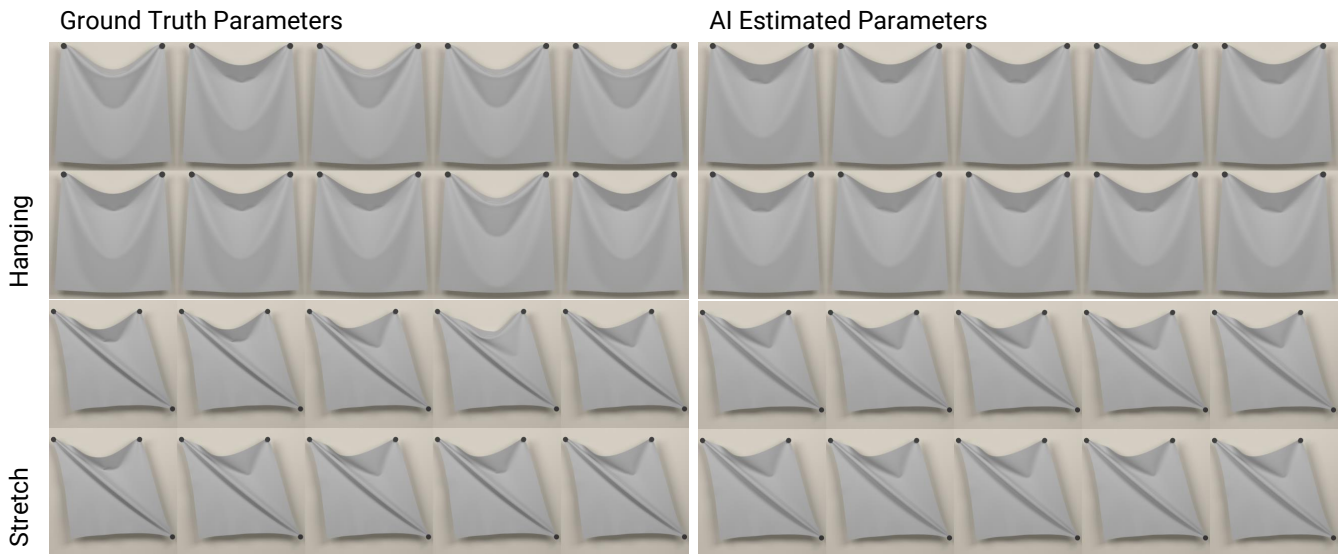


Figure 25: Different simulations for the fabric ID-0061 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0115

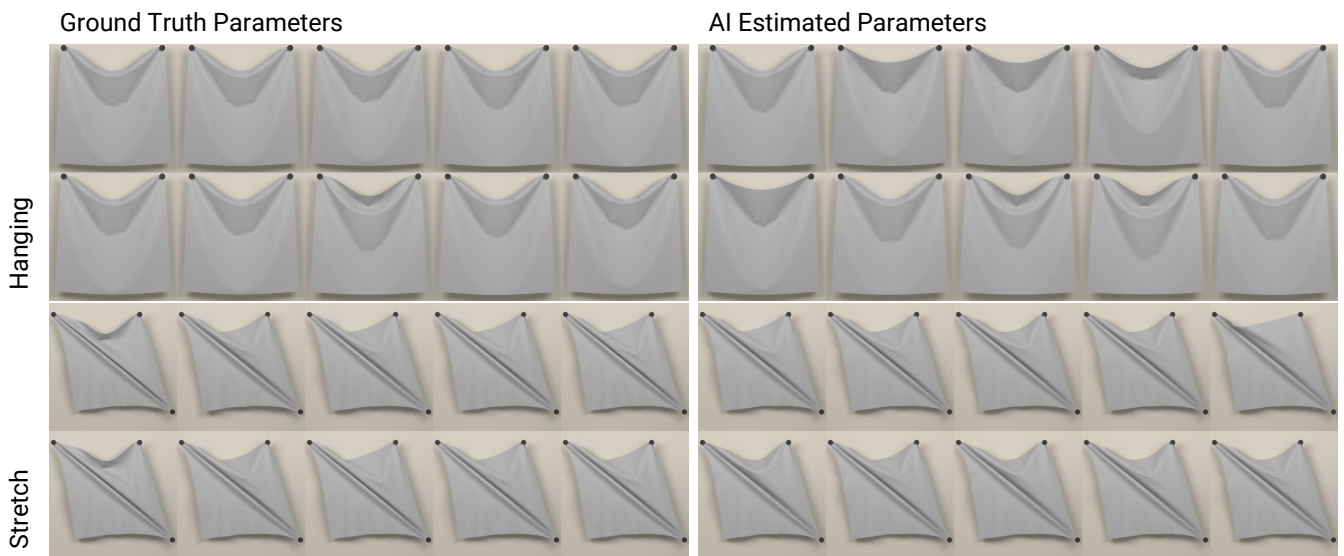


Figure 26: Different simulations for the fabric ID-0115 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0178

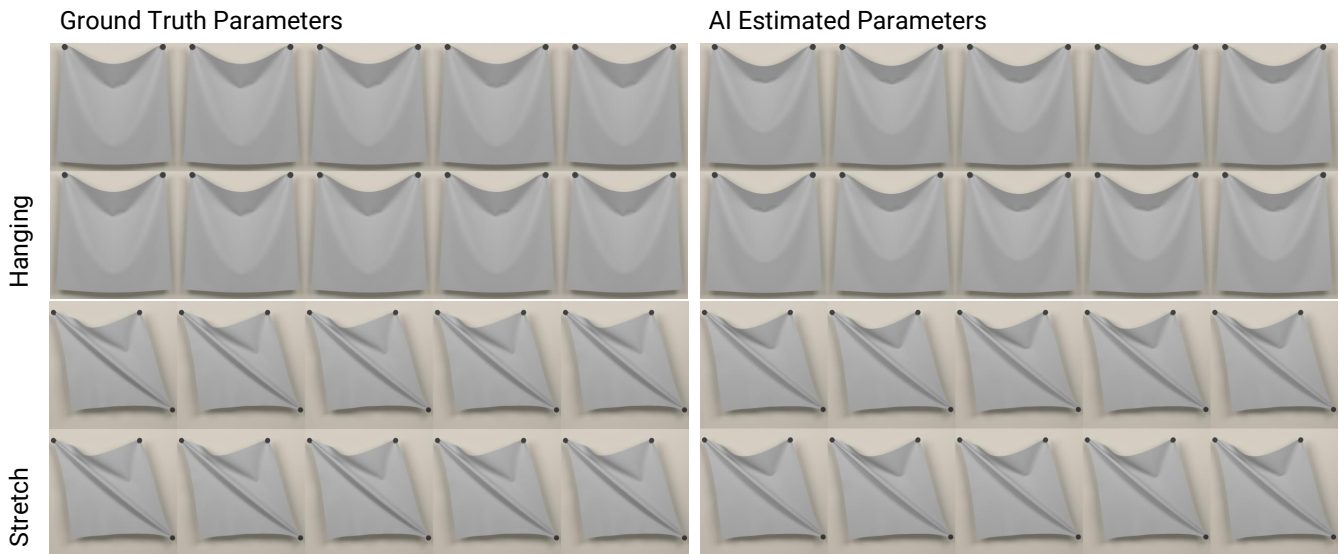


Figure 27: Different simulations for the fabric ID-0178 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0183

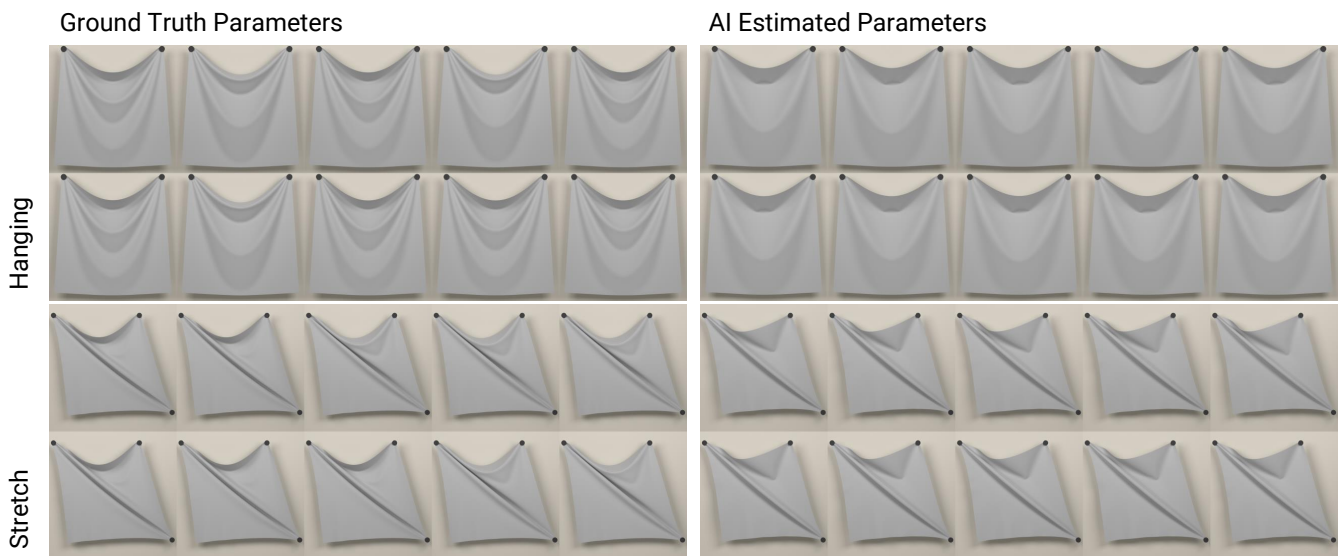


Figure 28: Different simulations for the fabric ID-0183 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0188

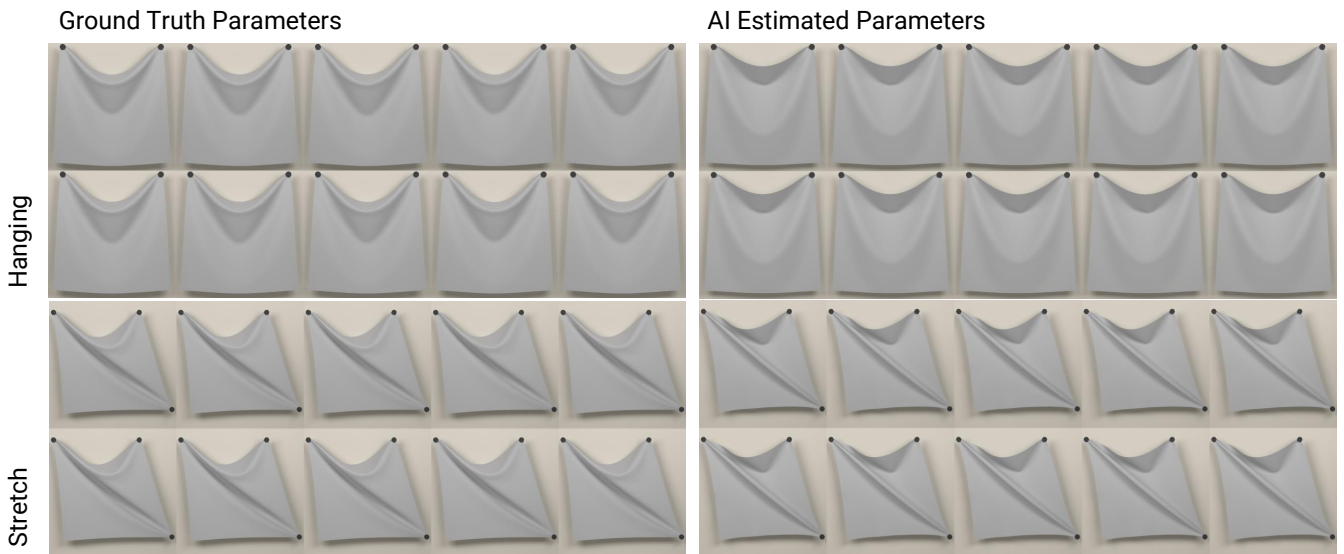


Figure 29: Different simulations for the fabric ID-0188 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0191

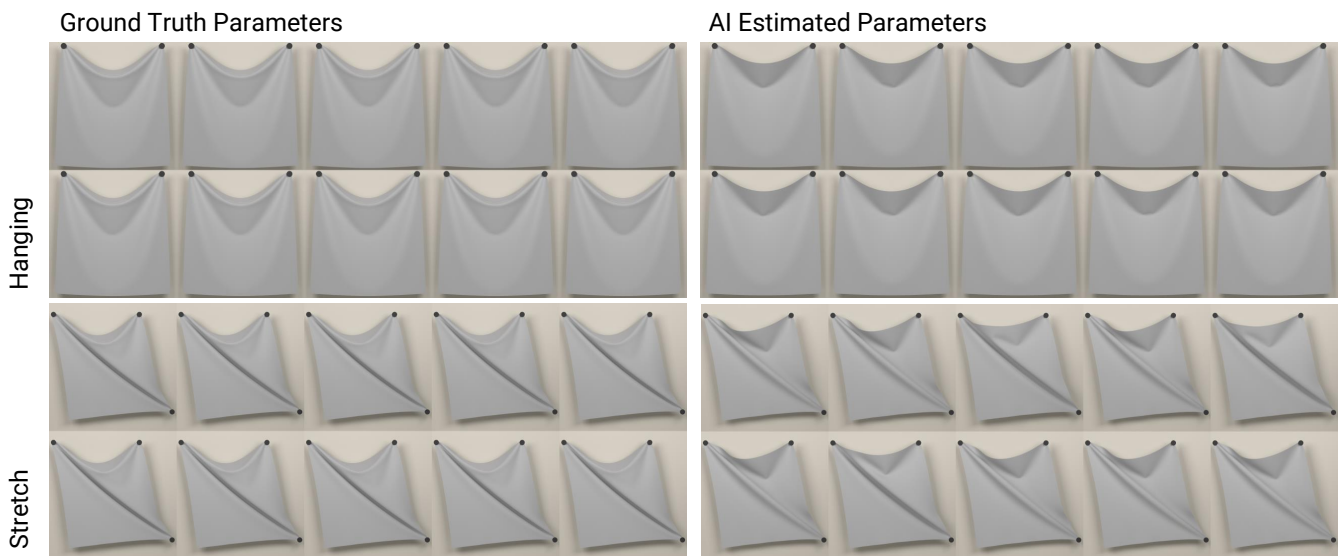


Figure 30: Different simulations for the fabric ID-0191 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0208

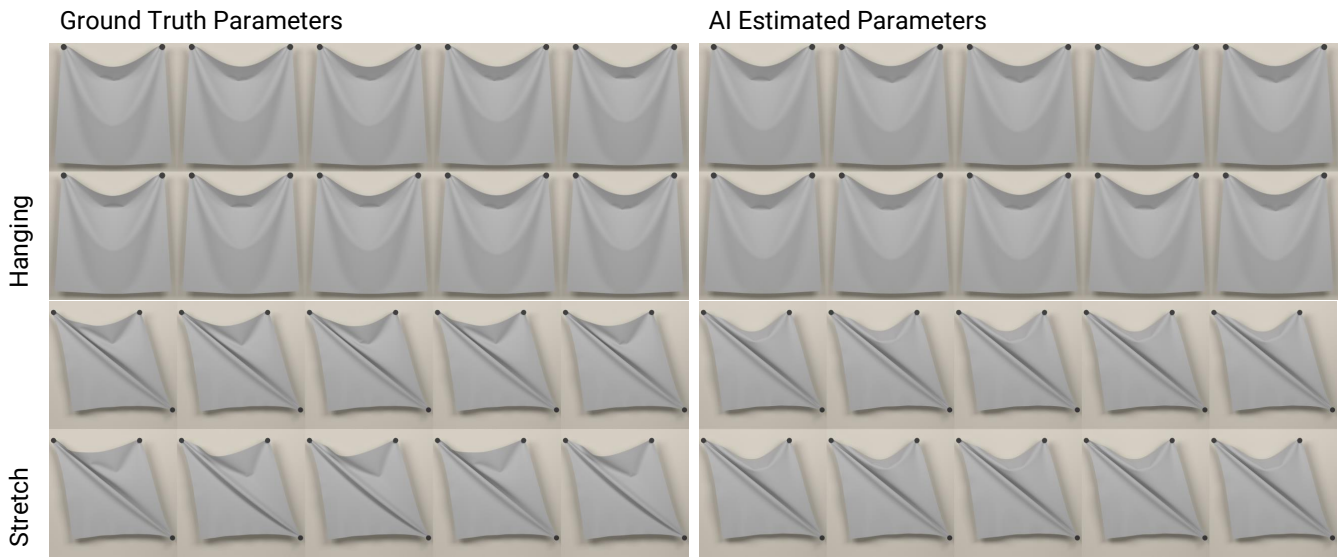


Figure 31: Different simulations for the fabric ID-0208 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0233

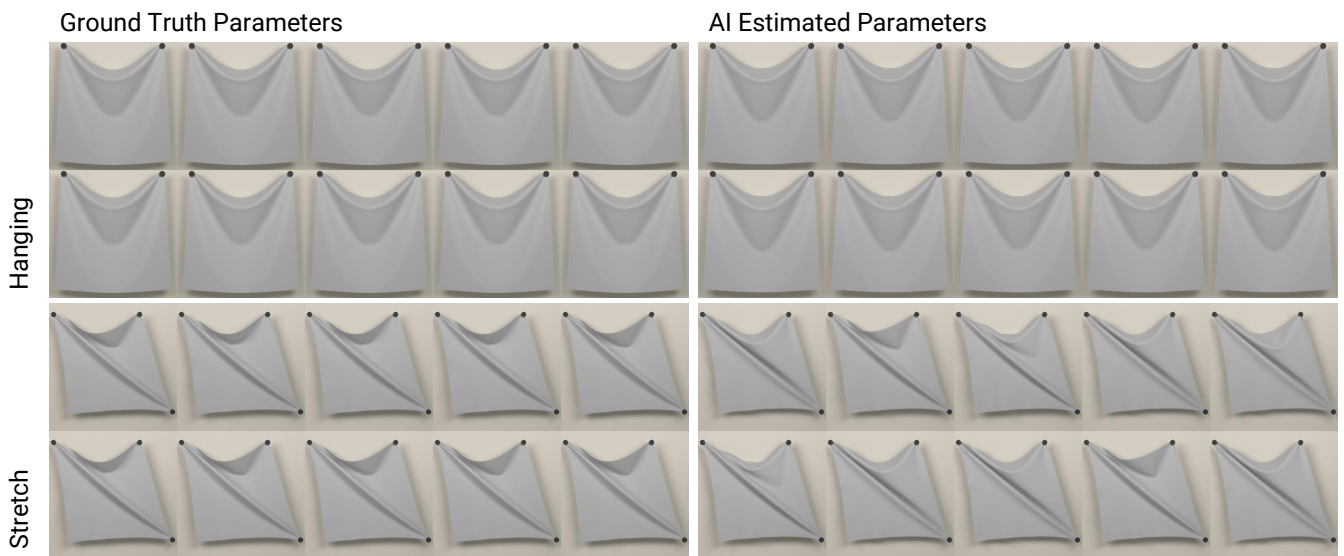


Figure 32: Different simulations for the fabric ID-0233 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0234

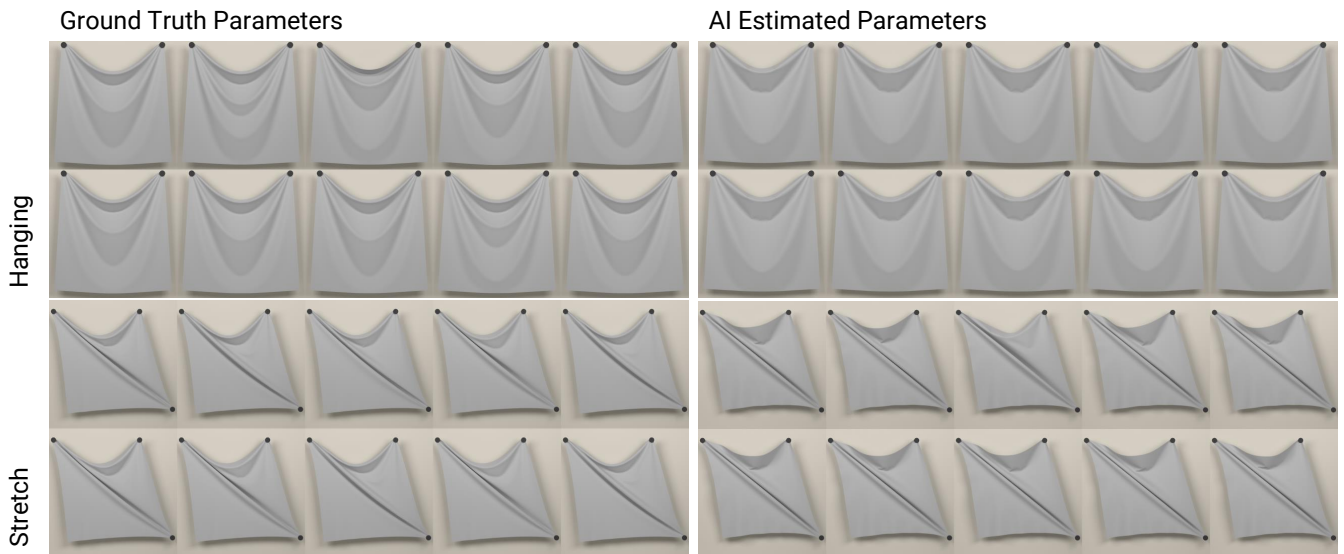


Figure 33: Different simulations for the fabric ID-0234 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).

ID-0242

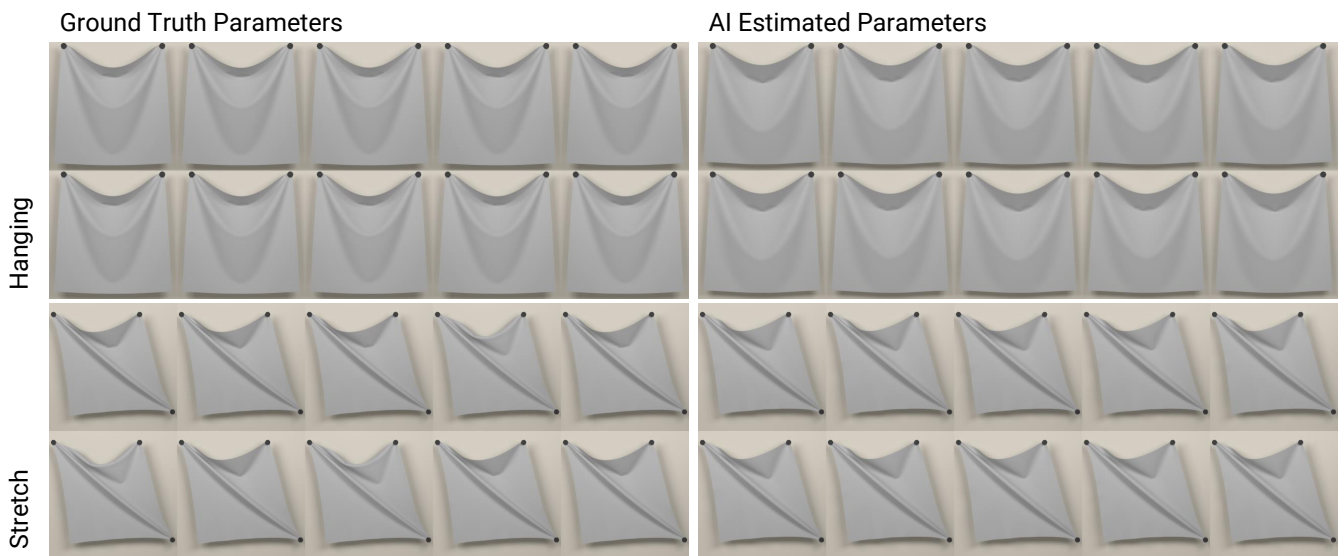


Figure 34: Different simulations for the fabric ID-0242 of our test set, for the Ground Truth parameter (left) and the parameters predicted by our model (right).



Figure 35: Comparisons of real garments on a mannequin, with virtual garments simulated using our engine and material model, for two different fabrics with different mechanical properties. On the left, we show a lightweight georgette, on the right, a thicker denim.

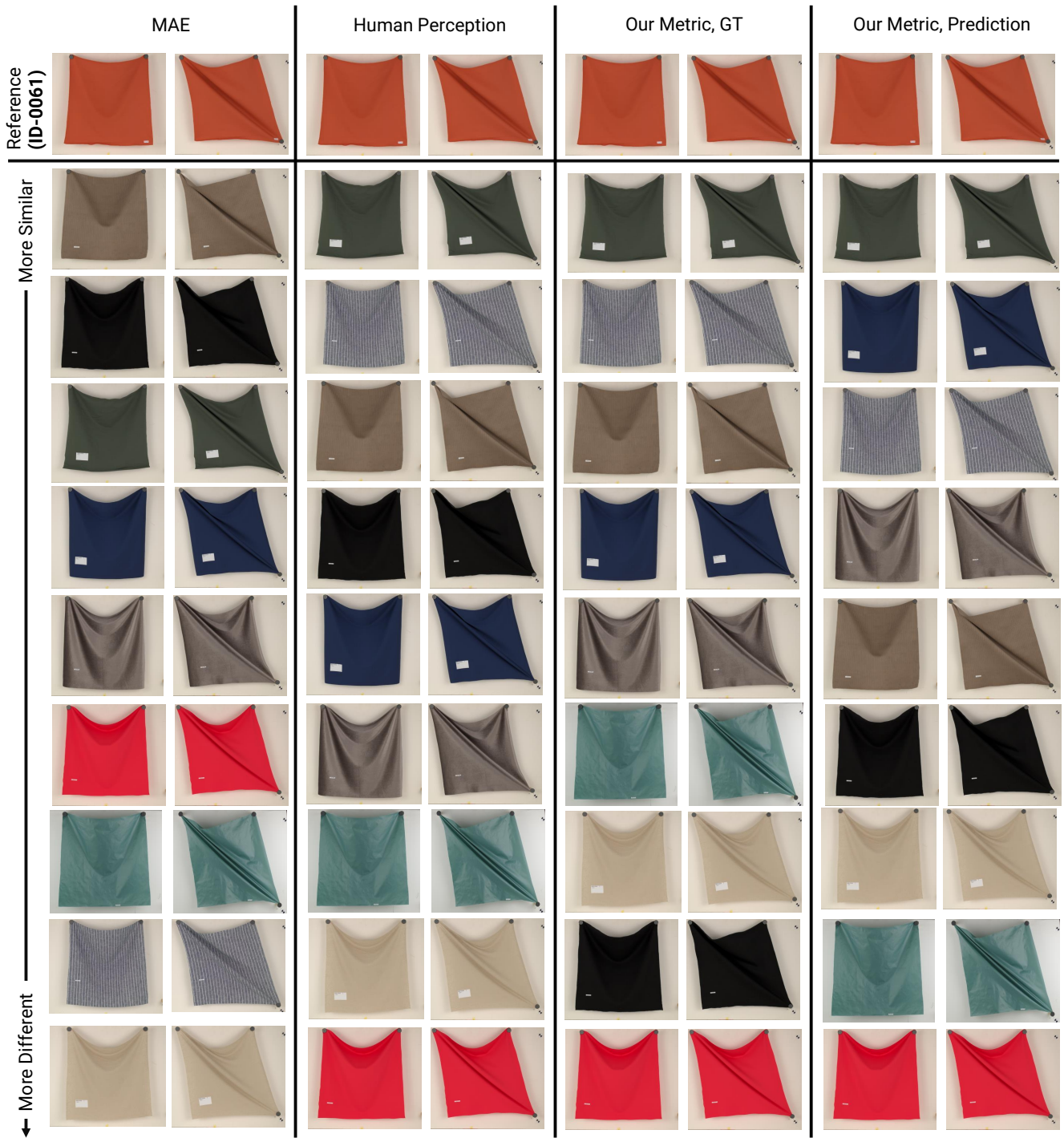


Figure 36: Relative orderings obtained for the fabric ID-0061 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).

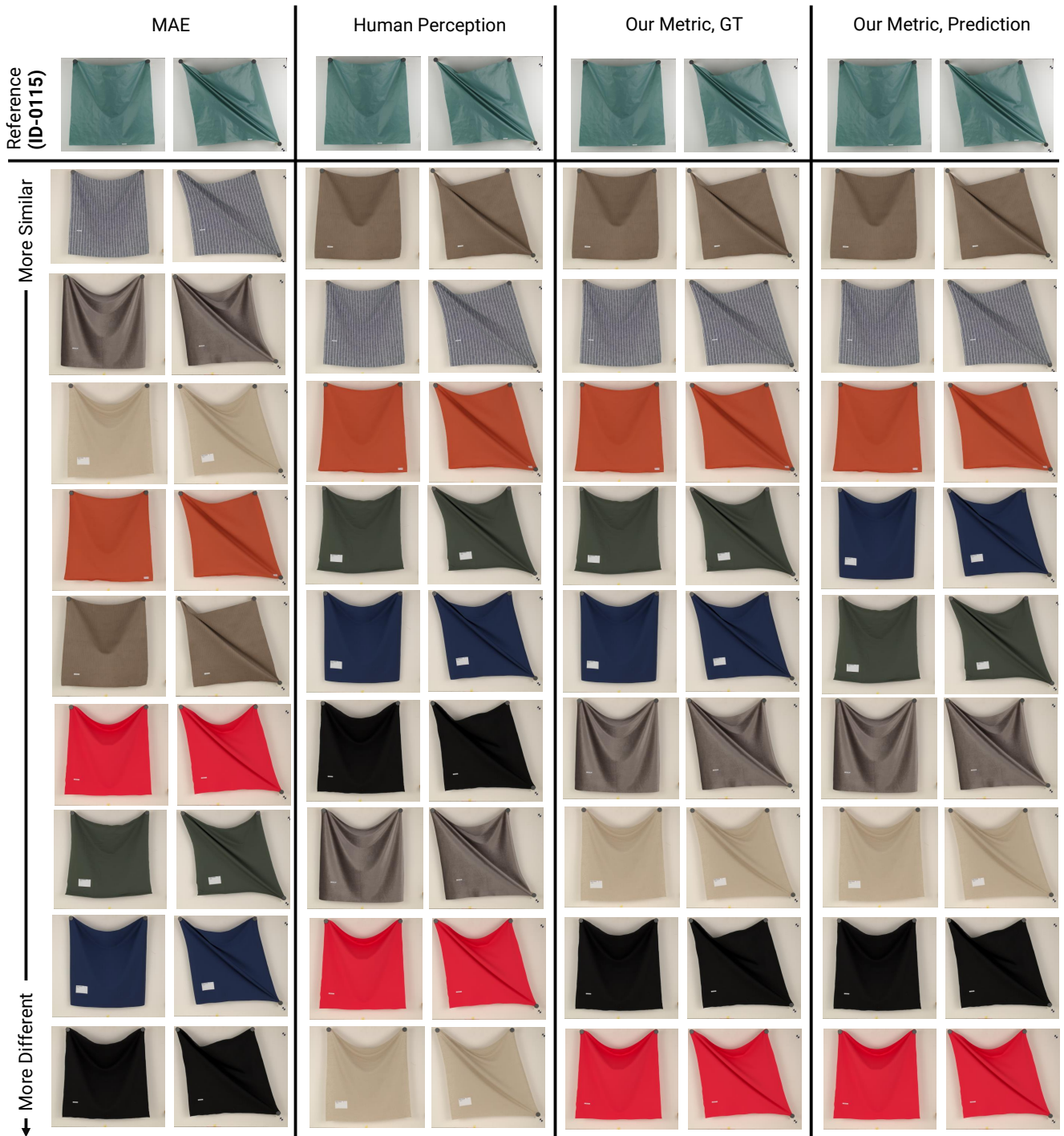


Figure 37: Relative orderings obtained for the fabric ID-0115 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).



Figure 38: Relative orderings obtained for the fabric ID-0178 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drapé Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drapé Similarity Metric evaluated using the Ground Truth parameters (rightmost column).

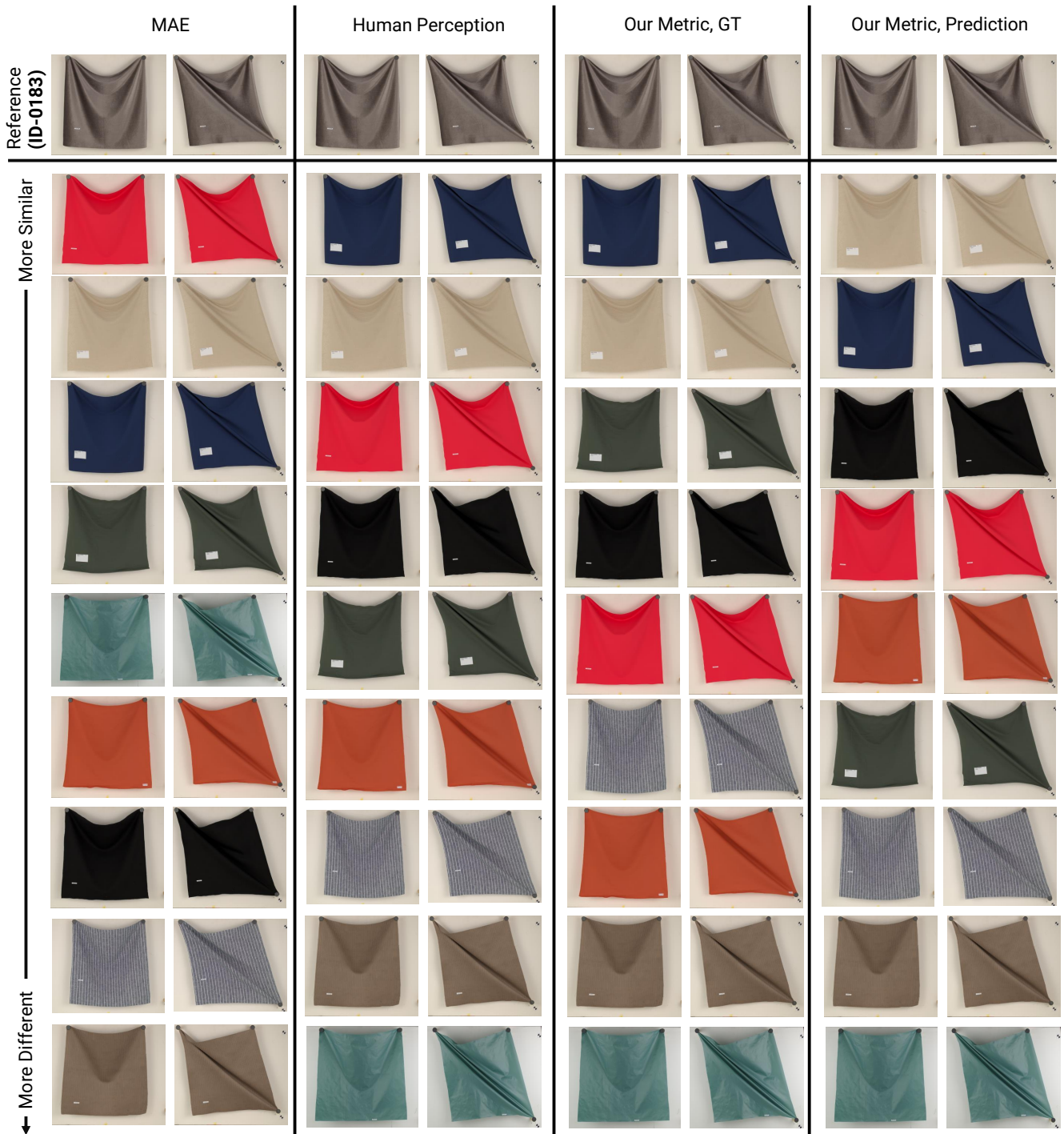


Figure 39: Relative orderings obtained for the fabric ID-0183 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).

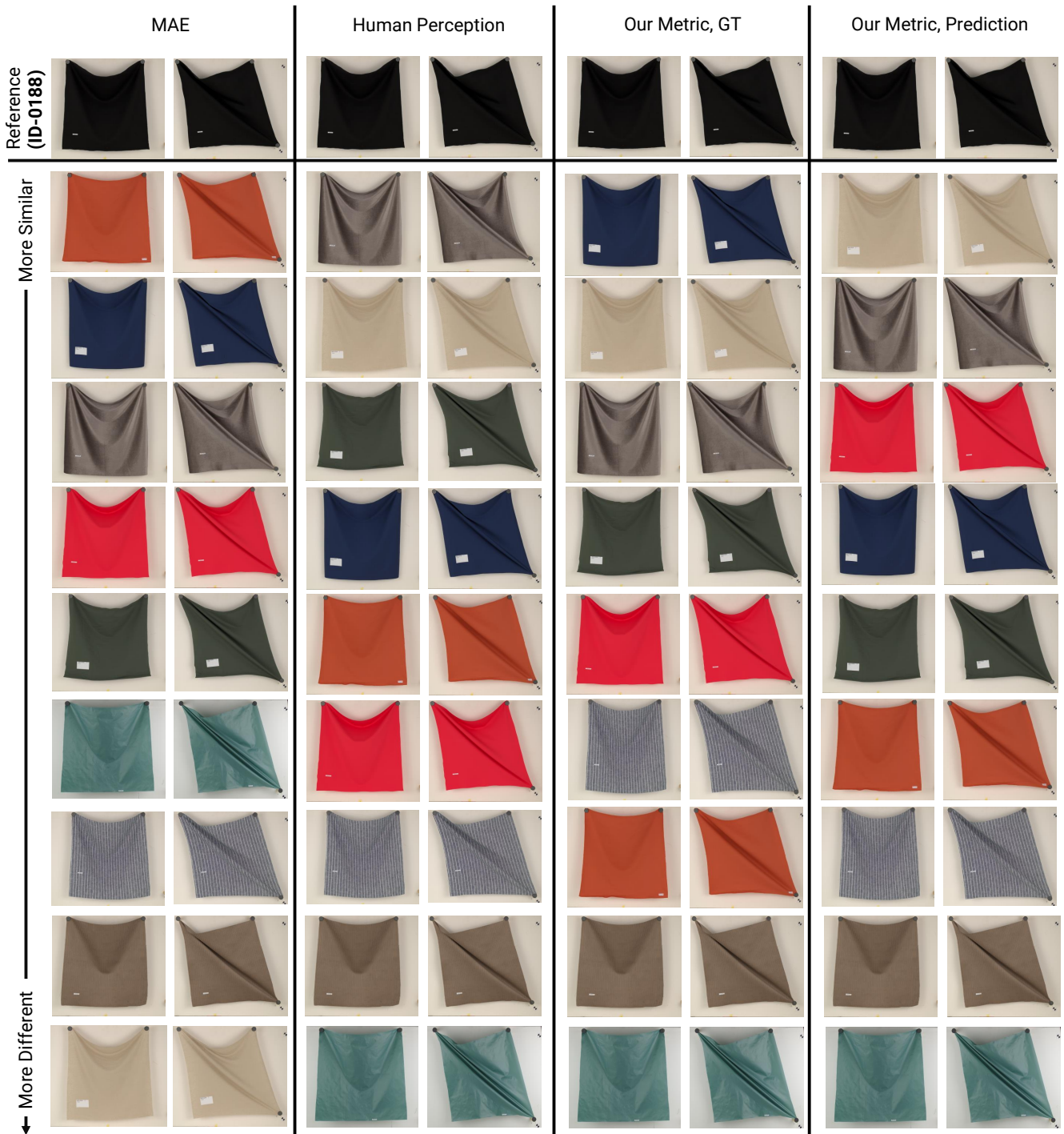


Figure 40: Relative orderings obtained for the fabric ID-0188 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).

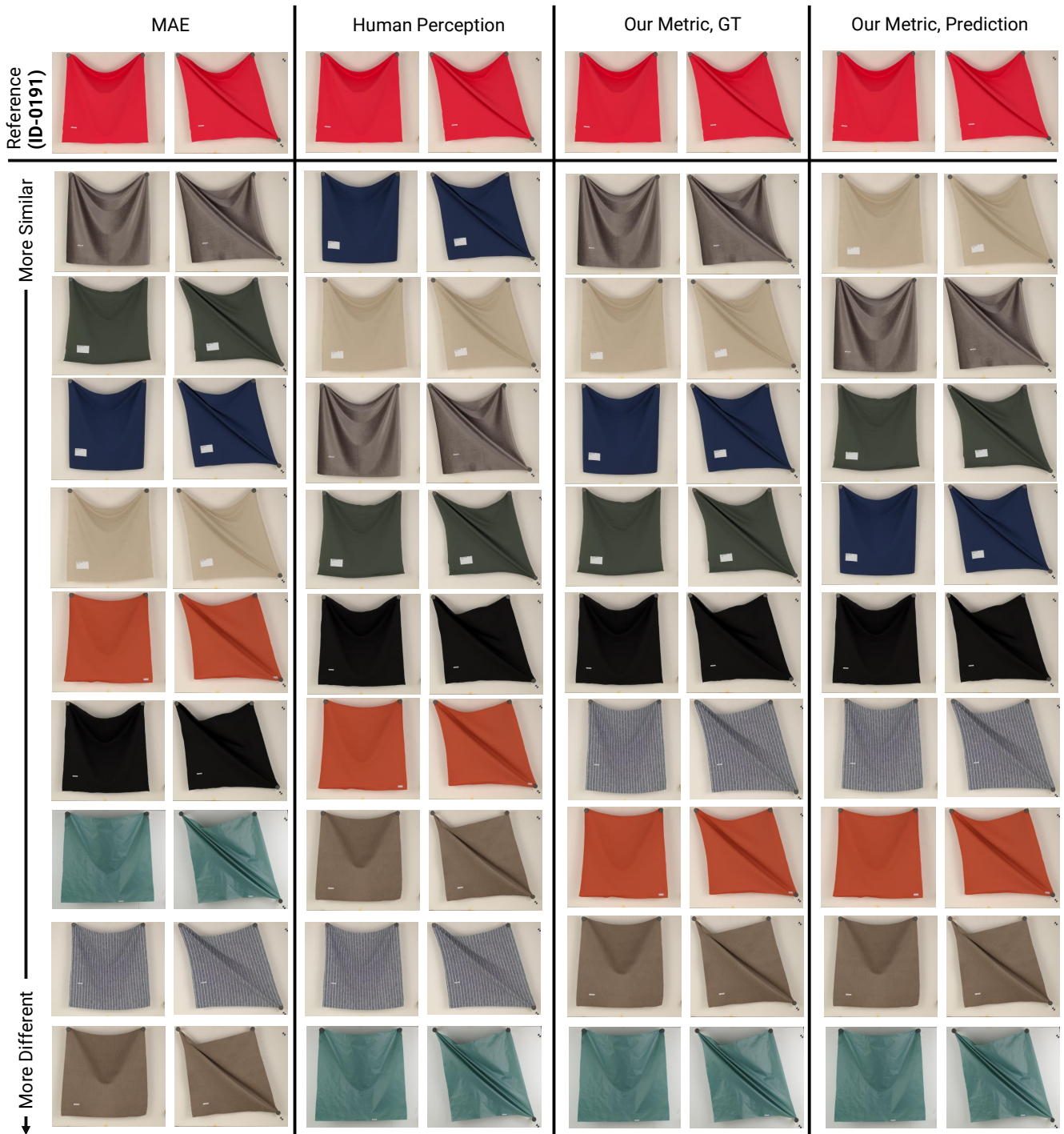


Figure 41: Relative orderings obtained for the fabric ID-0191 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).

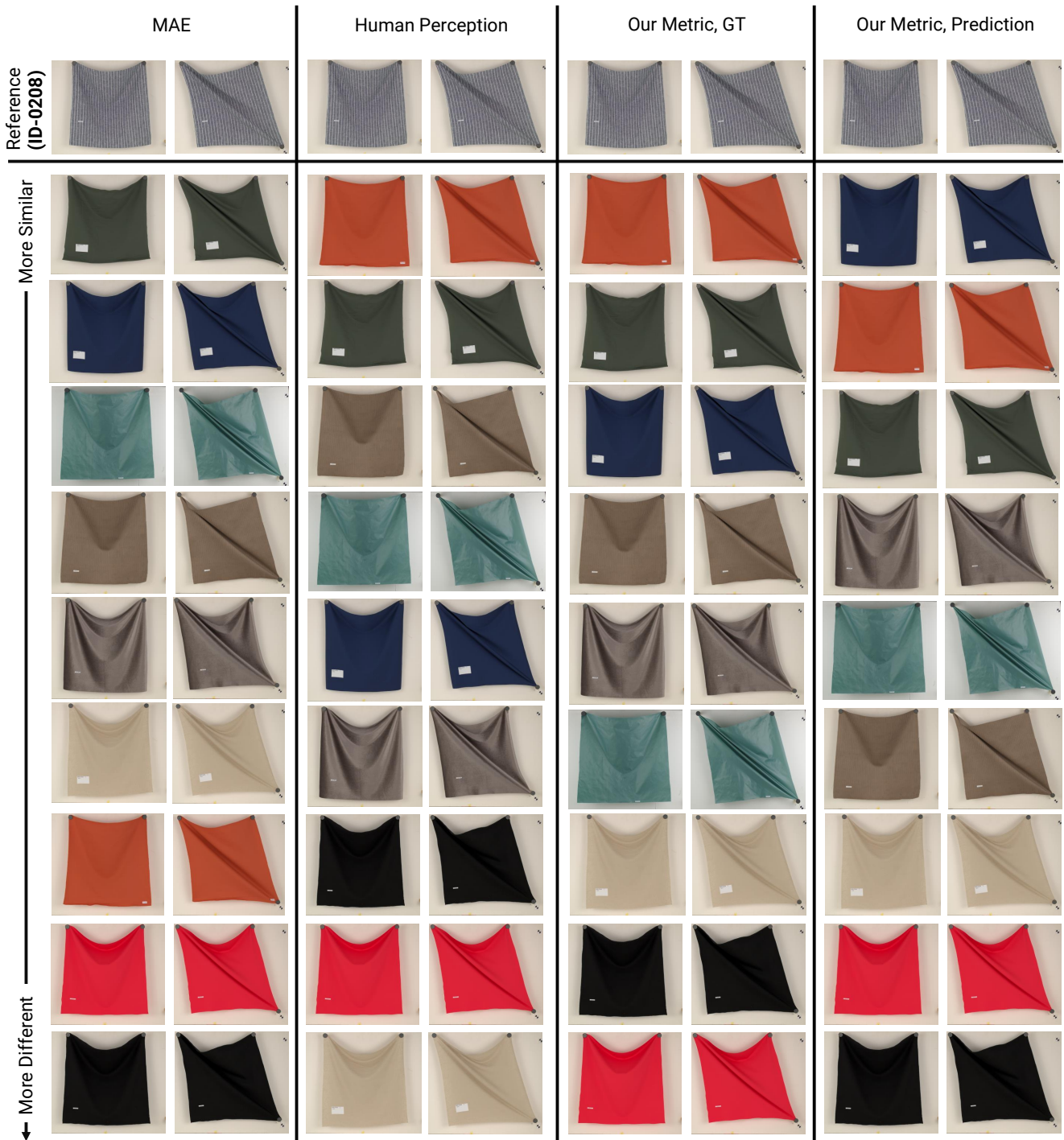


Figure 42: Relative orderings obtained for the fabric ID-0208 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).



Figure 43: Relative orderings obtained for the fabric ID-0233 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).

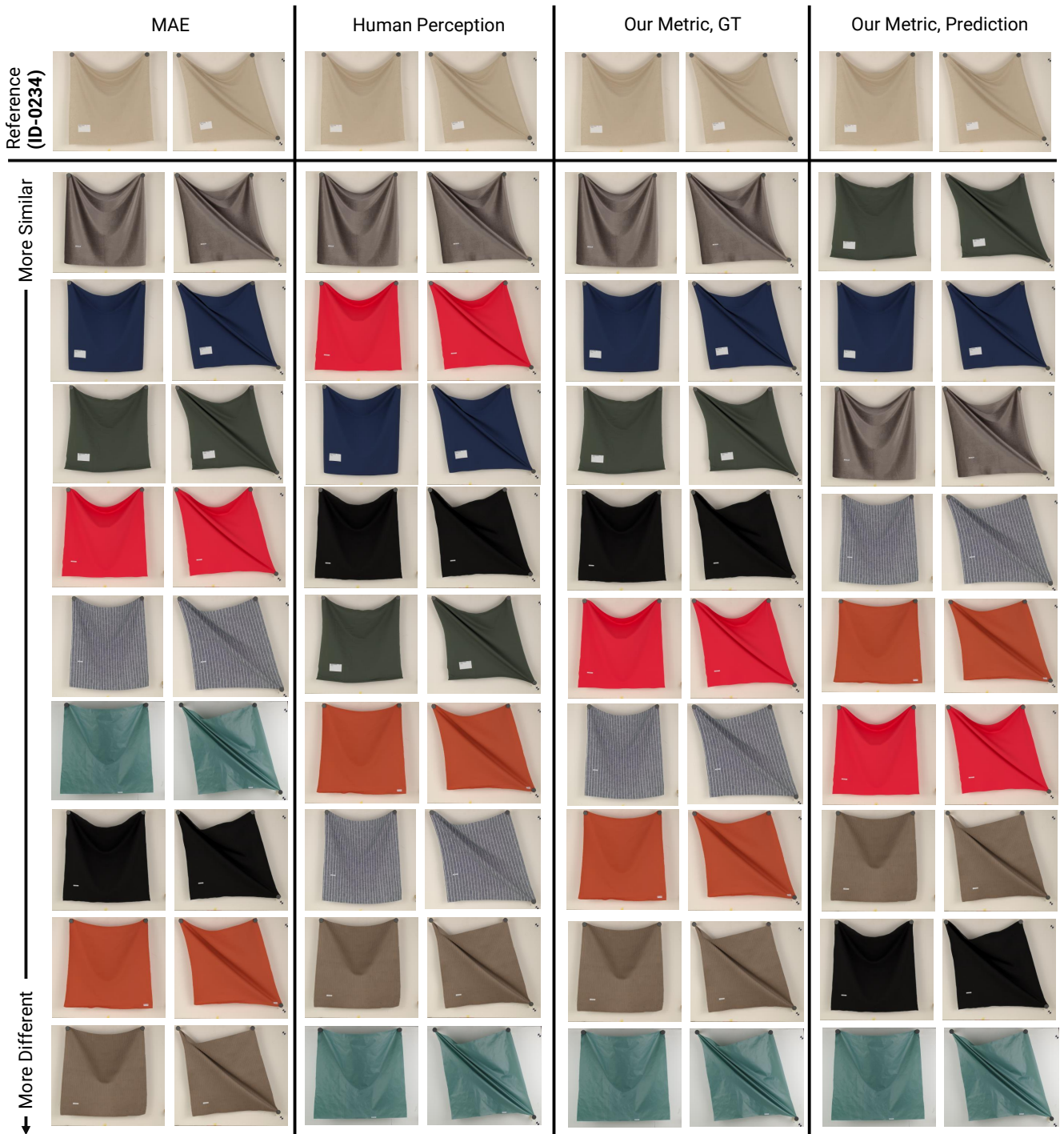


Figure 44: Relative orderings obtained for the fabric ID-0234 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).



Figure 45: Relative orderings obtained for the fabric ID-0242 of our test set, using the similarities provided by comparing the materials on the parameter space (leftmost column), the tSTE embedding (center-left column), our Drape Similarity Metric evaluated using the Ground Truth parameters (center-right column), and our Drape Similarity Metric evaluated using the Ground Truth parameters (rightmost column).