High-Quality Face Capture Using Blendshape-Driven Muscles
Supplementary Material

A. Targeting 3D Geometry - Additional Results

We present additional comparisons between using blendshapes and simulations for targeting three-dimensional geometry in Figure 1. Our approach using muscle simulation results in facial expressions similar to that obtained via blendshapes, but also introduces physical properties such as volume preservation. Our results can be improved by further calibrating and refining the anatomical model. As seen in Figure 2, the resulting muscle activation weights are sparser and less overdialed than their blendshape counterparts. In particular, note how the muscle activations generally track the magnitude of the expression. This is especially evident in frame 2590 where the face is in a close to neutral pose; while the muscle activations are close to all 0, the blendshape weights are still dialed in heavily to match the expression. The overdialing of blendshape weights could be alleviated by increasing the L2 regularization of the weights; however, this will also cause the captured performance to become less representative of the original performance. Figure 3 shows that muscle activations result in anatomically and semantically meaningful information. Note that further calibration of the anatomical model will also lead to more accurate muscle activation weights.

B. Targeting RGB Images - Additional Results

We show additional results for targeting monocular RGB images in Figure 4. Furthermore, we show the resulting geometry and plates for the same frames from another camera perspective in Figure 5. The corresponding blendshape weights and muscle activations are shown in Figure 6. A visualization of the muscles' activations is shown in Figure 7. Currently, the muscle activations resulting from targeting RGB images do not permit as clean of an interpretation as those obtained when targeting geometry, although the incisivus labii superioris muscles tend to become activated in conjunction with expressions involving the mouth. However, we note that the general magnitude of the activations tends to match the magnitude of the expression. Future work calibrating the muscle model will improve semantic intepretability.

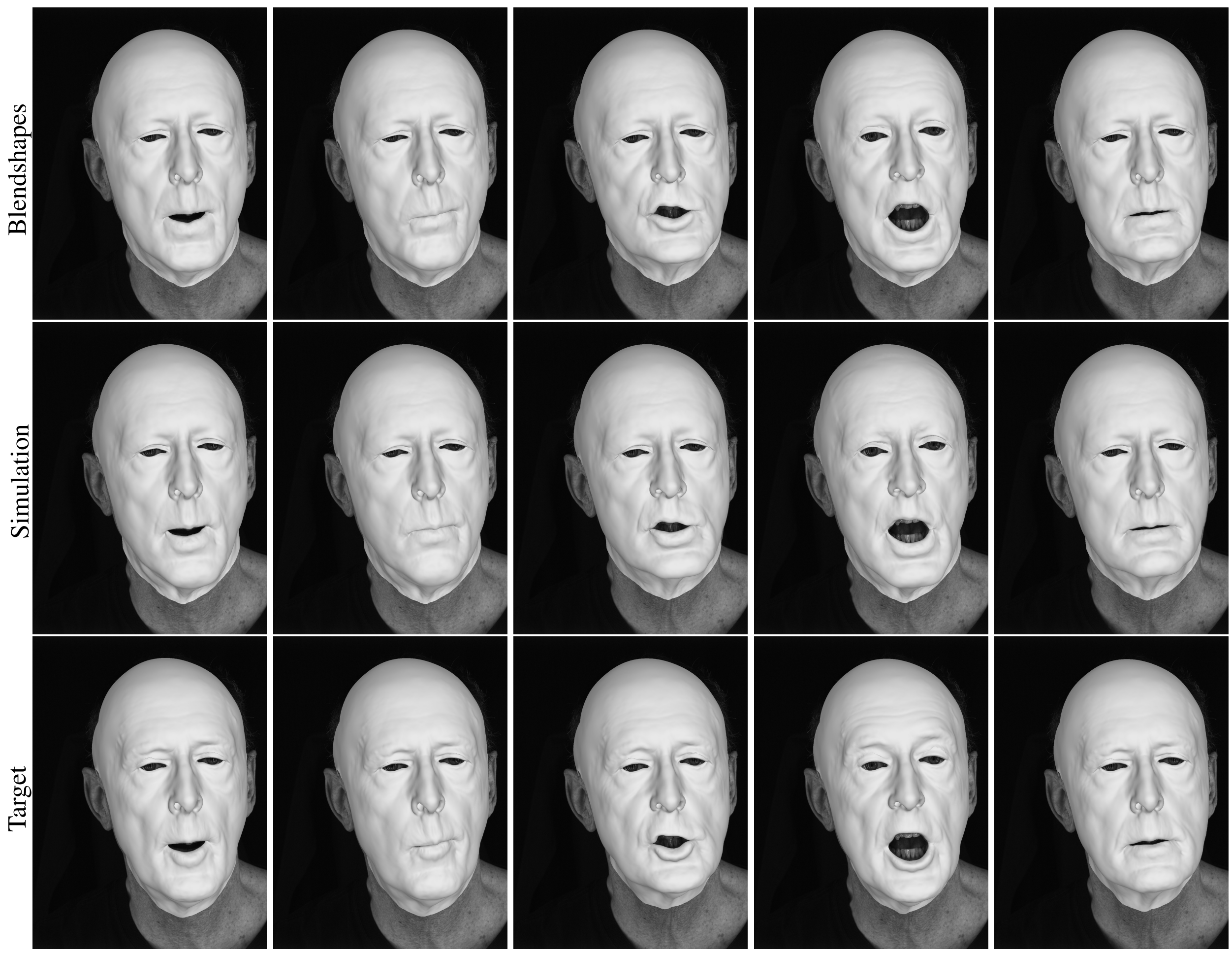
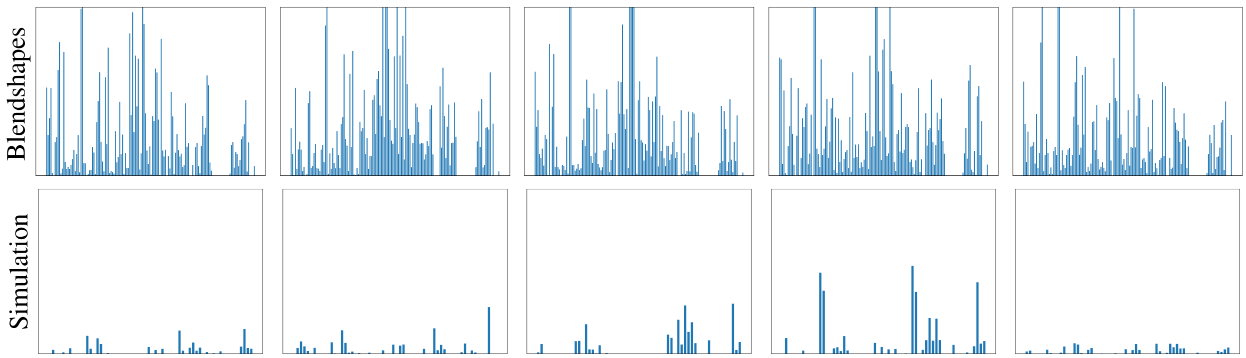
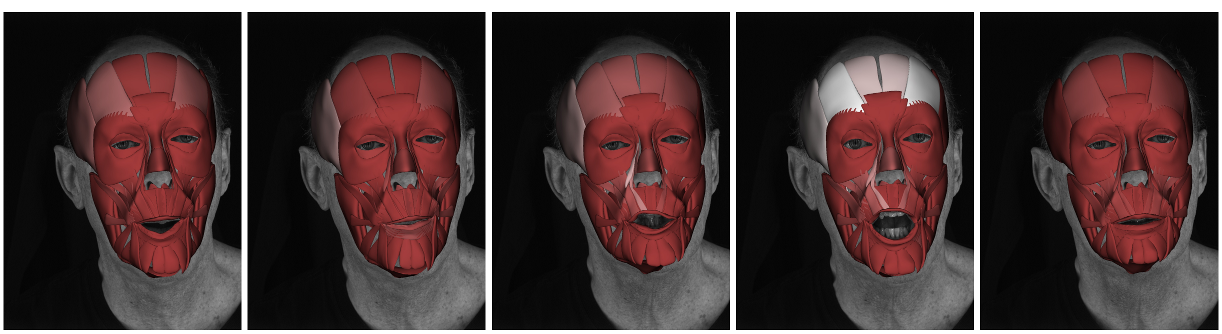


Figure 1. Additional comparisons when targeting geometry viewed from one of the original camera viewpoints.



(a) 2536 (b) 2540 (c) 2560 (d) 2573 (e) 2590

Figure 2. Additional comparisons between the resulting blendshape weights and muscle activations when targeting geometry.



(a) 2536 (b) 2540 (c) 2560 (d) 2573 (e) 2590

Figure 3. Muscle activations from Figure 2 visualized where activations greater than 0.5 are colored white and activations at 0 are colored red.

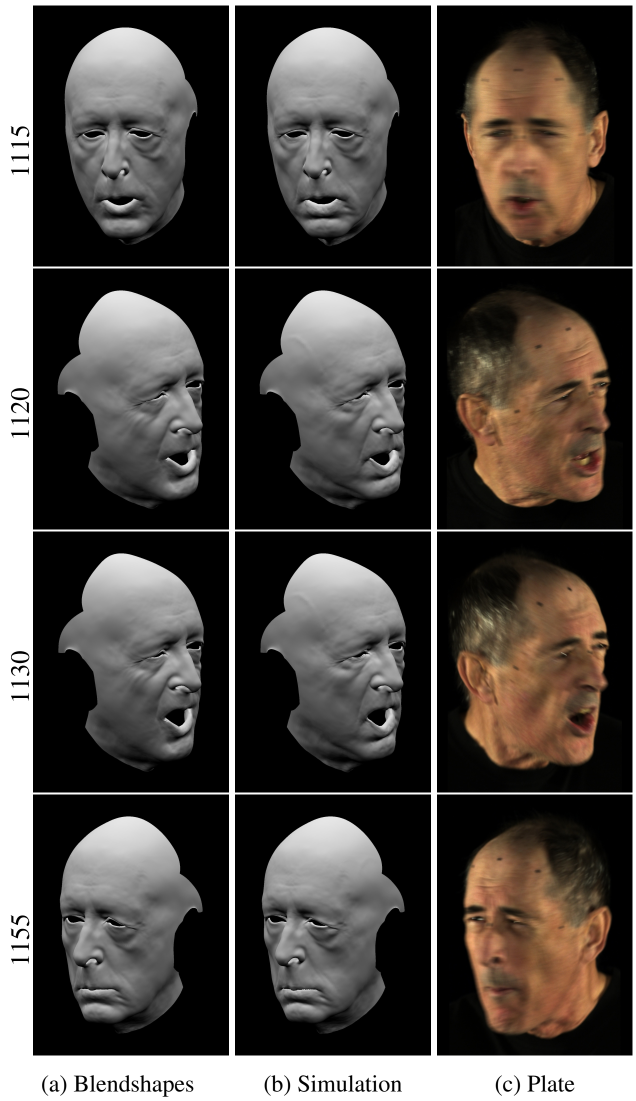


Figure 4. Targeting the monocular RGB image using shape-from-shading and rotoscope curves with blendshapes and simulation from the main camera's perspective.

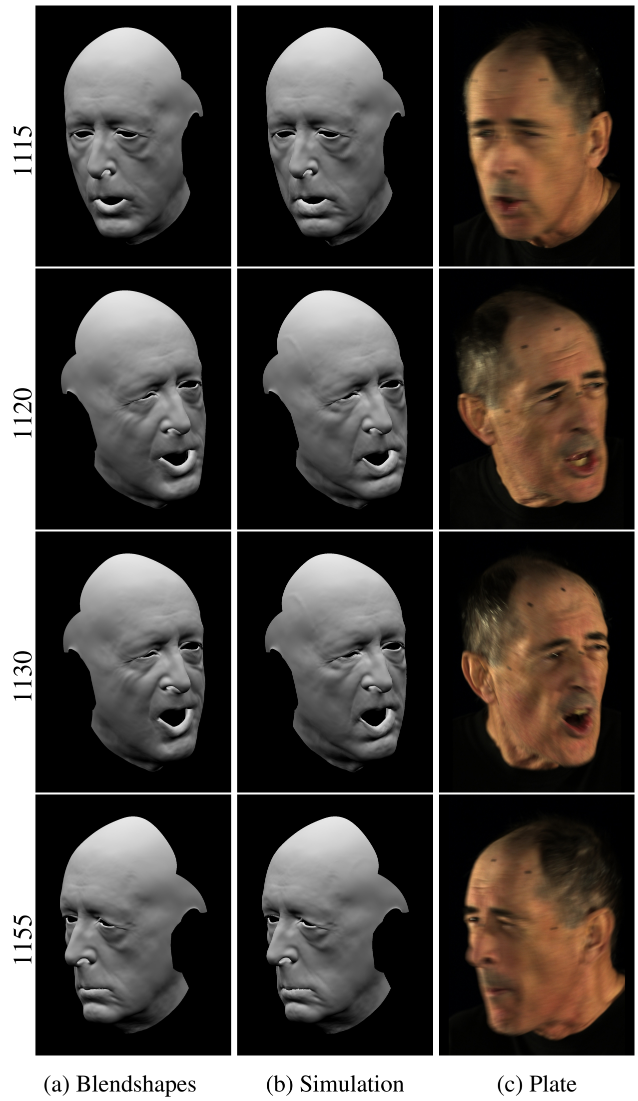
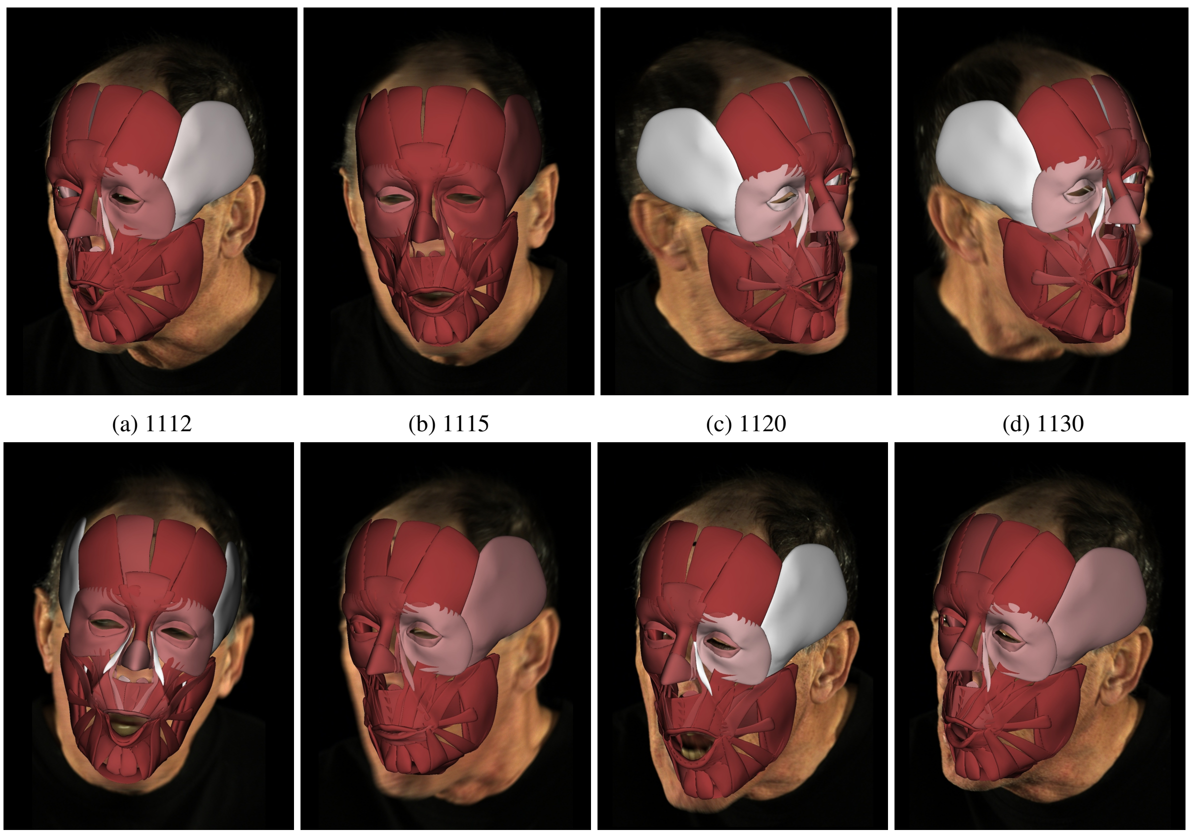


Figure 5. Targeting the monocular RGB image using shape-from-shading and rotoscope curves with blendshapes and simulation from an alternate camera's perspective.

Figure 6. Comparisons between the blendshape weights and muscle activations for all the monocular shape-from-shading results. The corresponding geometry for frames 1115, 1120, 1130, and 1155 are shown in Figures 4 and 5. The corresponding geometry for frames 1112, 1134, 1160, and 1170 are shown in the main paper.



(a) 1112 (b) 1115 (c) 1120 (d) 1130

(e) 1134 (f) 1155 (g) 1160 (h) 1170

Figure 7. Muscle activations from Figure 6 visualized where activations greater than 0.5 are colored white and activations at 0 are colored red.