# Supplemental Material
# Factored Occlusion: Single Spatial Light Modulator Occlusion-capable Optical See-through Augmented Reality Display

Brooke Krajancich*, Nitish Padmanaban*, Gordon Wetzstein

## 1 Supplemental Results

In this document, we show additional and extended simulated and captured results in support of those shown in the primary document. Figures S1–S5 show the full set of captures for Figures 1, 5, 6 and 8, respectively, of the primary text. Figures S6–S8 present additional scenes not shown in the primary text.

• *Brooke Krajancich, Nitish Padmanaban and Gordon Wetzstein are with Stanford University. E-mail: {brookek | nit | gordonwz}@stanford.edu. *indicates equal contribution.*

*Manuscript received xx xxx. 201x; accepted xx xxx. 201x. Date of Publication xx xxx. 201x; date of current version xx xxx. 201x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.201x.xxxxxxx*

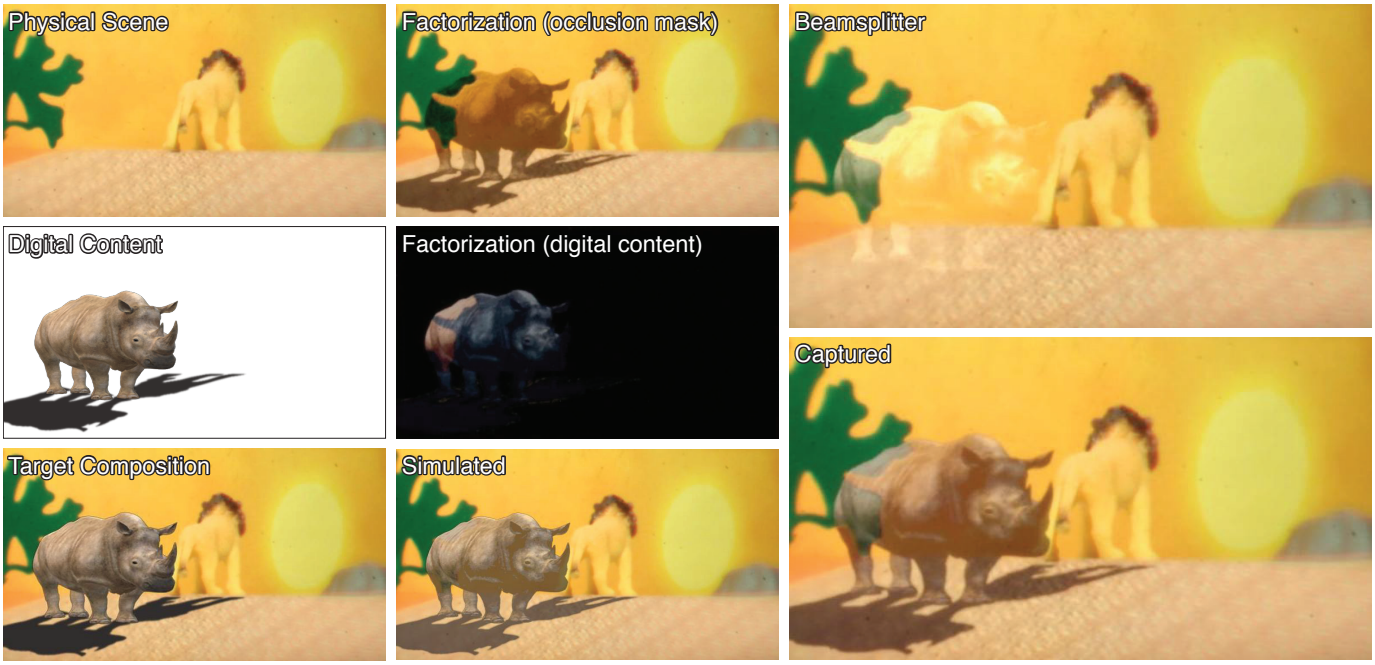


Fig. S.1. *Rhino* scene. We combine a physical scene (left, top) and digital image with plausible shadows (left, middle) to create a target composition (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity, light blocking, and ability to include realistic lighting effects over a conventional beamsplitter approach (right, top).

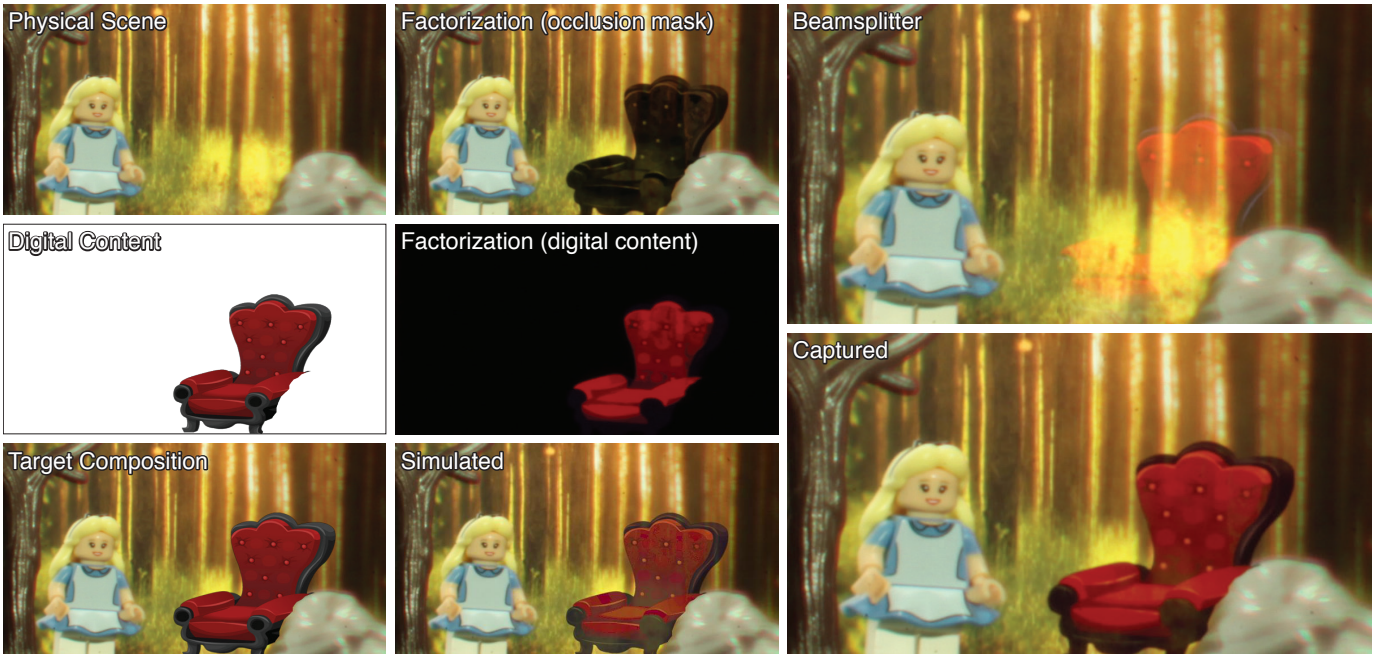


Fig. S.2. *Chair* scene. We combine a physical scene (left, top) and digital image (left, middle) to create a target scene (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity and light-blocking ability over a conventional beamsplitter approach (right, top).

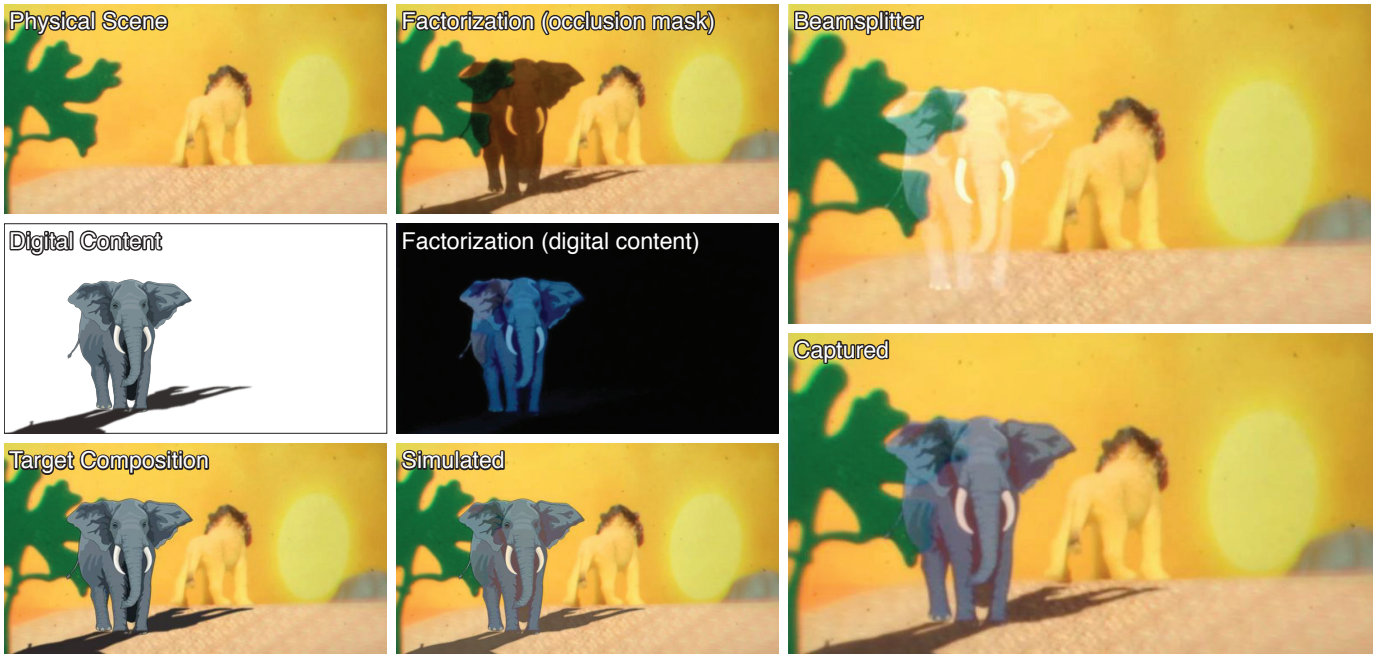


Fig. S.3. *Elephant* scene. We combine a physical scene (left, top) and digital image with plausible shadows (left, middle) to create a target composition (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity, light blocking, and ability to include realistic lighting effects over a conventional beamsplitter approach (right, top).

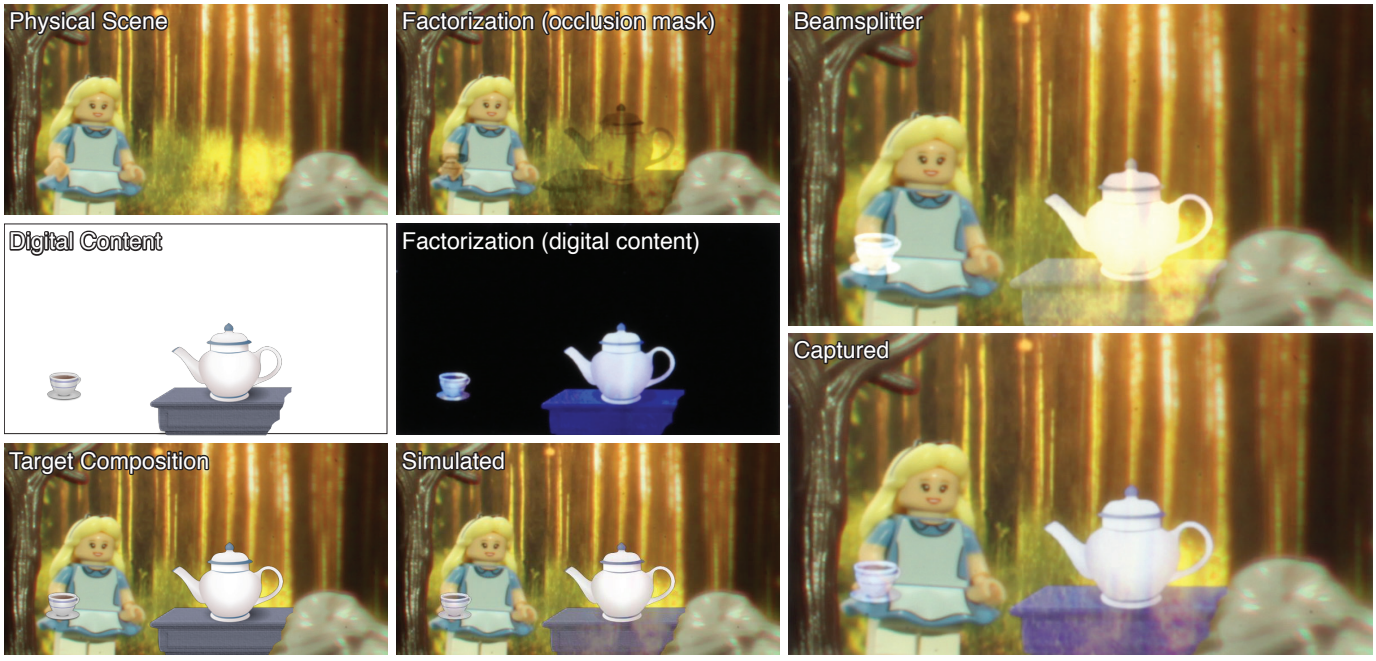


Fig. S.4. *Tea party* scene. We combine a physical scene (left, top) and digital image (left, middle) to create a target scene (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity and light-blocking ability over a conventional beamsplitter approach (right, top).


Fig. S.5. *Bird* scene. We combine a physical scene (a) and digital image (b) to create a target composition (c). Our optimization-based approach then factors this into occlusion mask (d) and digital content (e) components that combine as a best global approximation, as shown in simulation (f). This closely corresponds to the composition captured by our benchtop implementation (g), which shows significant improvement in color fidelity and light-blocking ability over a conventional beamsplitter approach (h). Since close-up viewing of the birds is likely to occur separately, we demonstrate that including a gaze-contingent weighting in our approach improves the rendering of each digital image. Simulating a viewer looking at the blue bird gives occlusion mask (i) and digital content (j) factorizations which combine to render the bird with a higher-intensity color and inclusion of more of the target's fine details, as demonstrated in both simulation (k) and capture (l). Similarly, applying gaze-contingent optimization to the red bird gives occlusion mask (m) and digital content (n) factorizations that combine to render the bird with a higher-intensity color and greater occlusion of the tree in the physical scene, both in simulation (o) and capture (p).

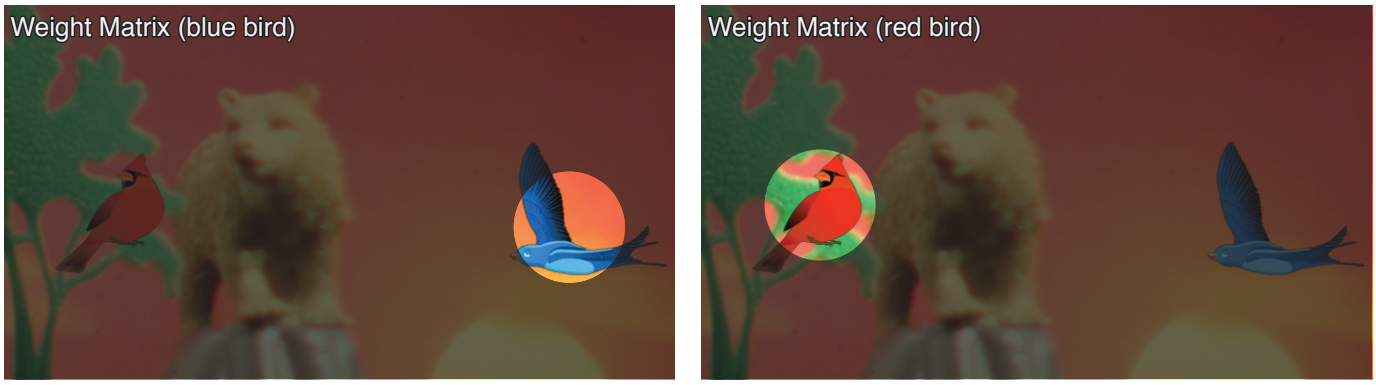


Fig. S.6. *Bird* scene weight matrix representation for blue bird (left) and red bird (right) gaze directions. Regions of faded image intensity indicate pixels given a weighting of 0.01, while the circular regions in full color indicate pixels given a weighting of 1. Each circular region has a diameter of 100 pixels, with the full image having pixel dimensions of 380 $\times$ 680.

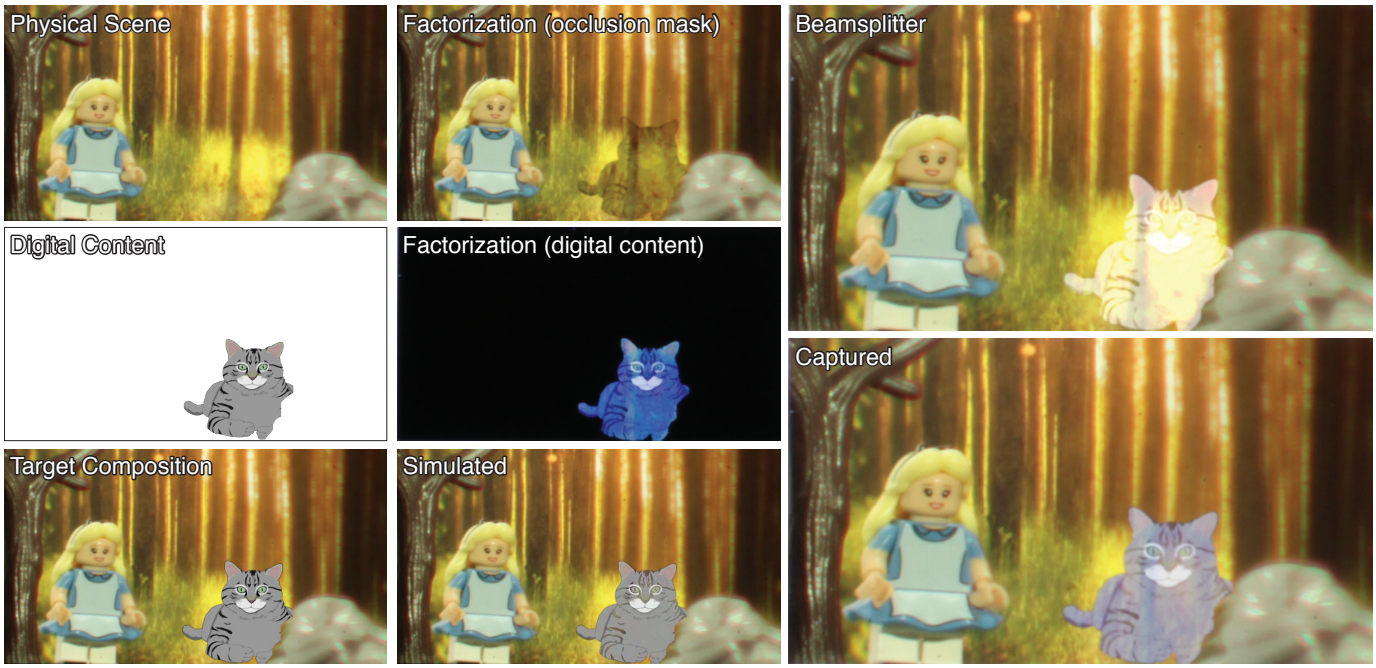


Fig. S.7. *Grey cat* scene. We combine a physical scene (left, top) and digital image (left, middle) to create a target composition (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity and light-blocking ability over a conventional beamsplitter approach (right, top).

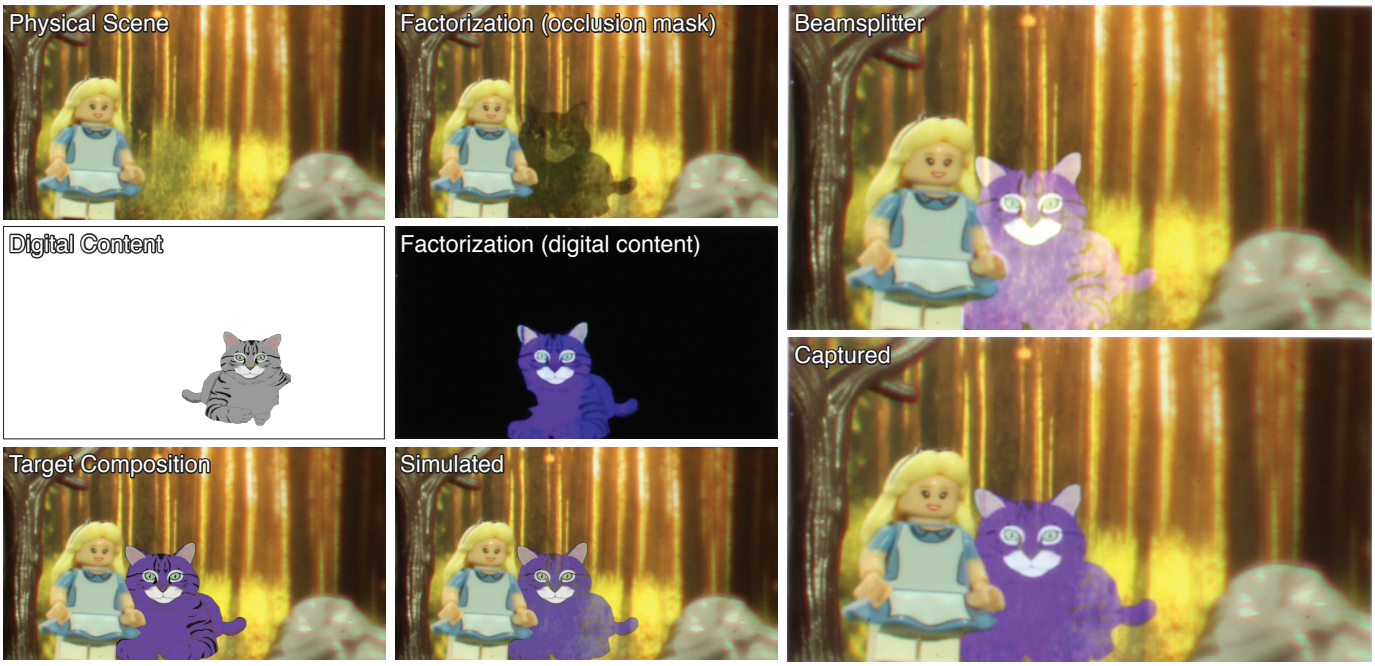


Fig. S.8. *Purple cat* scene. We combine a physical scene (left, top) and digital image (left, middle) to create a target composition (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity and light-blocking ability over a conventional beamsplitter approach (right, top).

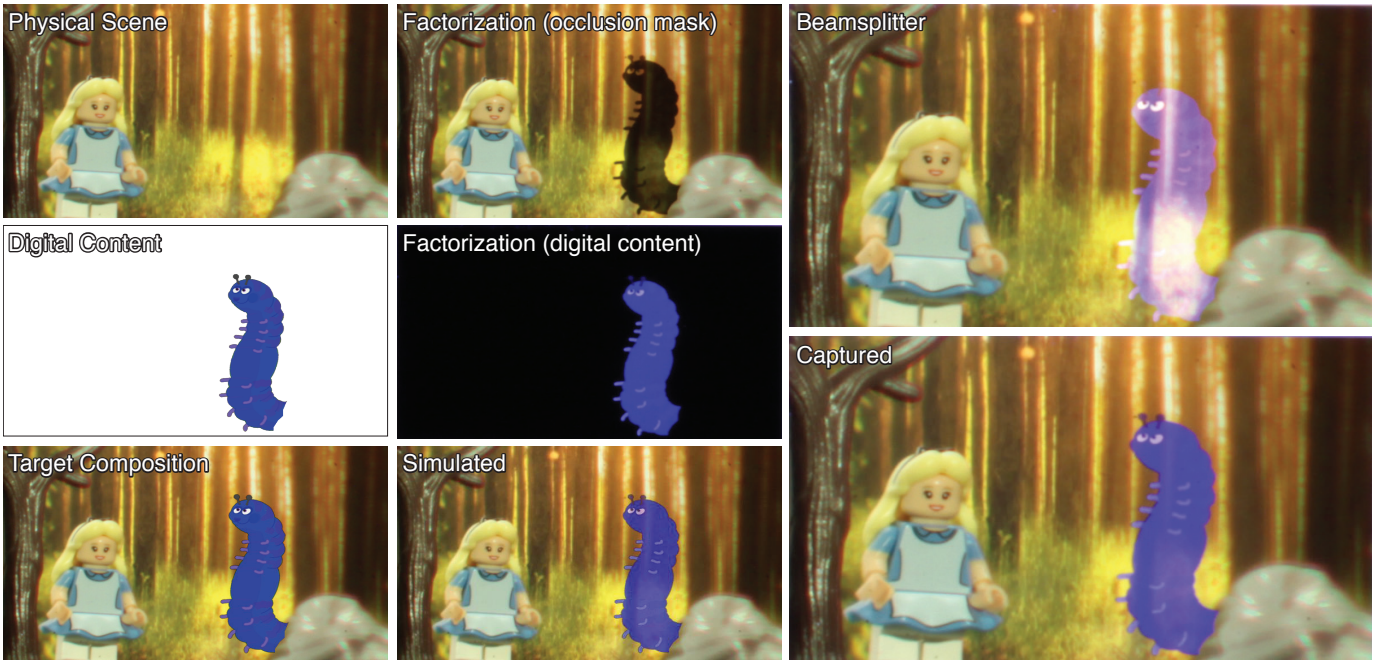


Fig. S.9. *Caterpillar* scene. We combine a physical scene (left, top) and digital image (left, middle) to create a target composition (left, bottom). Our optimization-based approach then factors this into occlusion mask (center, top) and digital content (center, middle) components that combine to form a best approximation, as shown in simulation (center, bottom). This corresponds closely to the composition captured by our benchtop implementation (right, bottom), which shows significant improvement in color fidelity and light-blocking ability over a conventional beamsplitter approach (right, top).

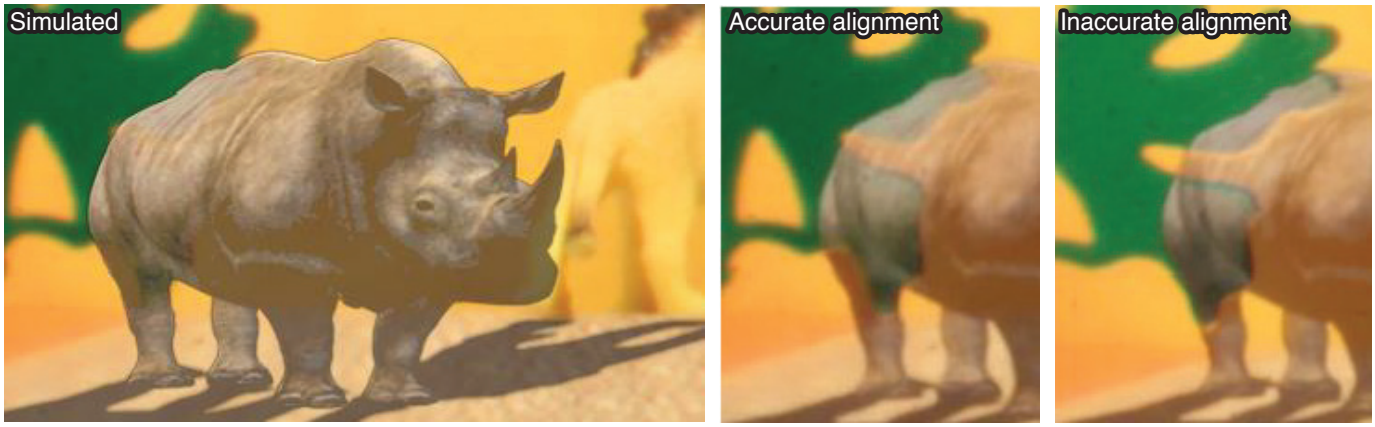


Fig. S.10. Incorrect alignment artifact. For a simulated target scene composition (left), demonstration of the captured result when the target composition and current physical scene are accurately (center) and inaccurately (right) aligned. Observe that this results in the edges of tree being poorly occluded and dark green artifacts in the Rhino. The Rhino's back leg also has a yellow spot, corresponding to where the optimization expected the tree to reach.