PG23

Daejeon, Korea

Hand Shadow Art: A Differentiable Rendering Perspective

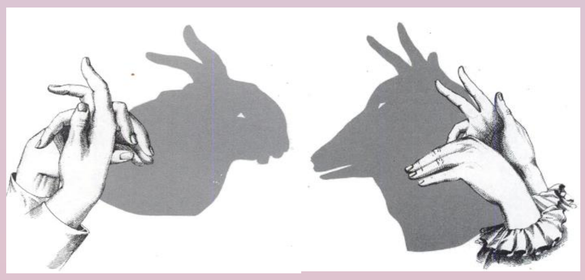
Aalok Gangopadhyay, Prajwal Singh, Ashish Tiwari, and Shanmuganathan Raman

Computer Vision, Imaging, and Graphics Lab

Indian Institute of Technology Gandhinagar

PROBLEM

- Shadow art is an exciting form of sculptural art that produces captivating artistic effects through the 2D shadows cast by 3D shapes.
- Hand shadows, also known as shadow puppetry or shadowgraphy, involve creating various shapes and figures using your hands and fingers to cast meaningful shadows on a wall.
- In this work, we propose a differentiable rendering-based approach to deform hand models such that they cast a shadow consistent with a desired target image and the associated lighting configuration.
- We show the results of shadows cast by a pair of two hands and the interpolation of hand poses between two desired shadow images.



The bunny and the deer as hand shadows (*Fun with Hand Shadows, Frank Jacobs*)

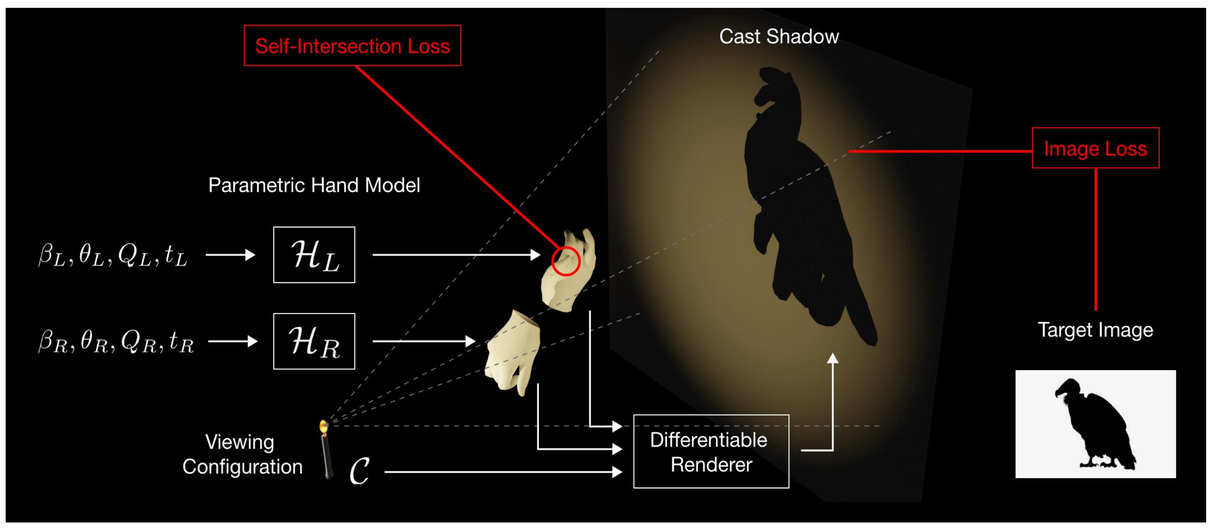
RELATED WORK

- A few recent methods have addressed shadow art using either optimization [2] or differentiable rendering [3]. However, none of them have explicitly addressed hand shadow art.
- Mitra *et al.* [2] described shadow art more formally by introducing a voxel-based optimization framework to recover the 3D shape from arbitrary input (shadow) images by deforming them and handling inherent image inconsistencies.
- Sadekar *et al.* [3] demonstrated the potential of differentiable rendering (mesh and voxel based) in generating 3D shadow sculptures all from arbitrary shadow images without any explicit input image deformation.

OVERVIEW

- We explicitly focus on developing a differential rendering based optimization framework to create hand shadow art by deforming the hand mesh models [3] and show the dynamics involved in interpolating between two shadow images.
- The hand model parameters are optimized by minimizing the image rendering loss and the mesh intersection loss. A model capture setup is shown below with camera moved from left to right.

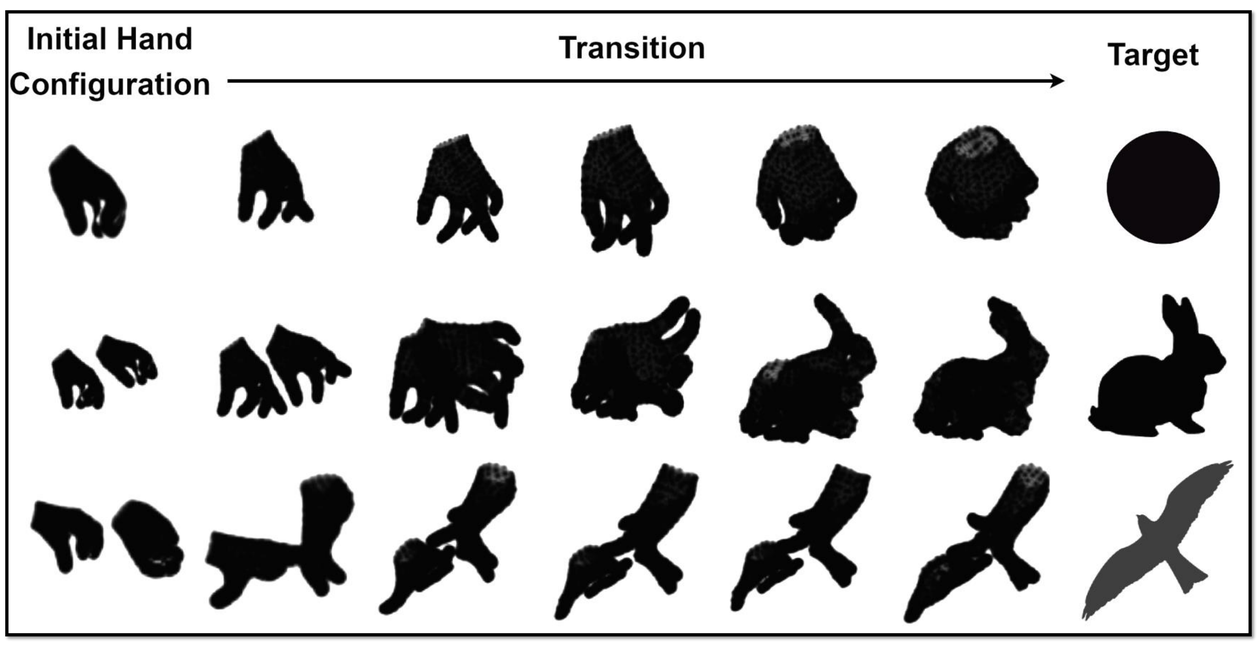
METHODOLOGY



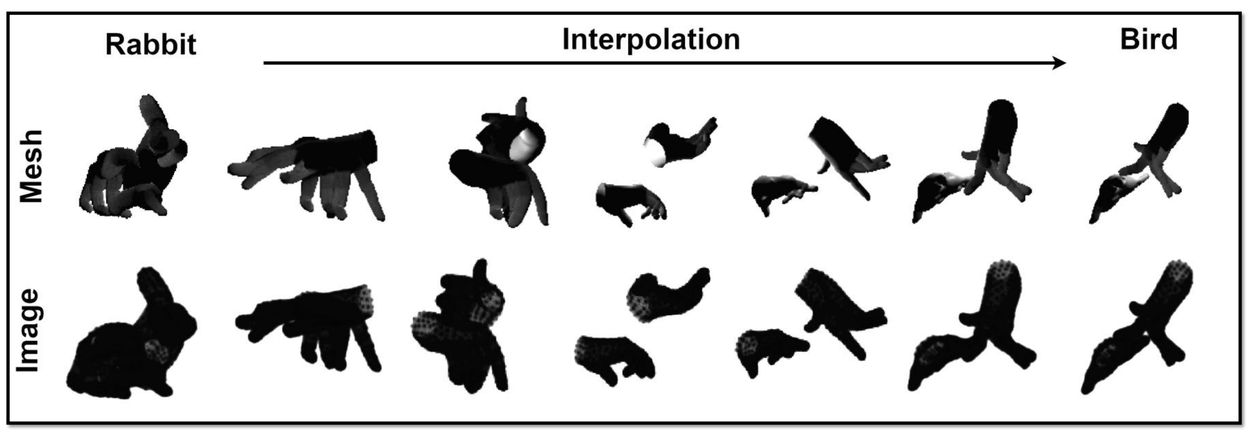
- $\mathcal{I} = [0,1]^{H \times W}$: space of all grayscale images defined on a grid of size $H \times W$.
- $I \in \mathcal{I}$ be a given target image.
- $\mathcal{M} = (\beta, \theta, Q, t)$ be the parametric hand model [3].
- β: Shape Parameter
- θ: Pose Parameter
- Q: Rotation Matrix
- t: translation vector
- $\mathcal{M}$: Hand mesh model

- $\mathcal{M}_L = \mathcal{H}_L(\beta_L, \theta_L, Q_L, t_L)$: Left hand model and $\mathcal{M}_R = \mathcal{H}_R(\beta_R, \theta_R, Q_R, t_R)$: Right hand model
- Consider $\mathcal{C}$ be the camera with fixed parameters placed at the origin. The lighting configuration with the camera $\mathcal{C}$ is called *viewing configuration* $(\mathcal{V})$.
- $\mathcal{R}(\mathcal{V}, \mathcal{M}_L, \mathcal{M}_R)$: Silhouette image rendered by arrangement of both hands as seen from the camera.
- Objective: Find the values of parameters $\theta_L, \theta_R, Q_L, Q_R, t_L$, and t_R that minimize
$$\left|\left|I - \mathcal{R}(\mathcal{V}, \mathcal{M}_L, \mathcal{M}_R)\right|\right|_2^2,$$
keeping β_L and β_R fixed.
- To avoid self-intersections and cross-intersections among a pair of hand meshes, we penalize such intersections according to pen loss [1, 5].
- We restrict the angular movement across 15 different joints (per hand) in the MANO hand models [4] to simulate realistic and plausible human hand movements.

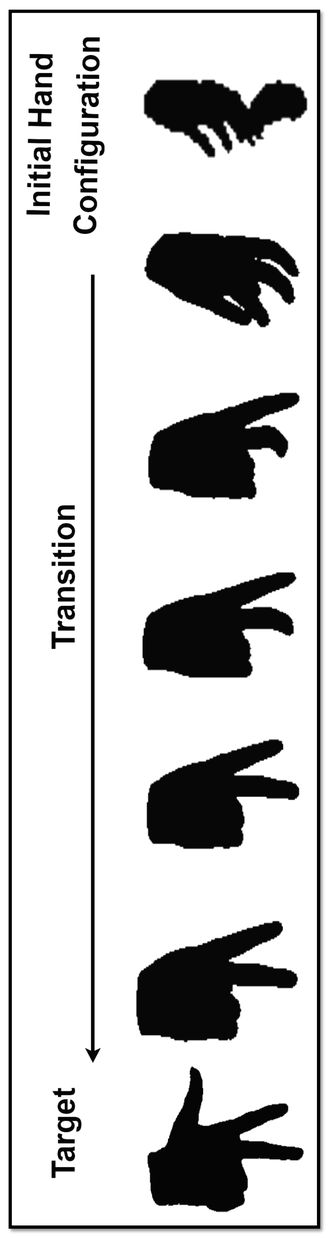
RESULTS



(a) Hand shadow art. We start with a random initial configuration of hand models (using one hand – row 1 and two hands – row 2 & 3) to reach the final pose, creating a shadow similar to the given target image when viewed from a fixed configuration.



(b) Interpolation. We interpolate between a pair of shadow images cast by a pair of hands. Given a pair of images I_A and I_B , we find a sequence of hand mesh models $(\mathcal{M}_L^t, \mathcal{M}_R^t)$ such that $\mathcal{R}(\mathcal{V}, \mathcal{M}_L^0, \mathcal{M}_R^0) = I_A$ and $\mathcal{R}(\mathcal{V}, \mathcal{M}_L^T, \mathcal{M}_R^T) = I_B$, where $t \in [0, T]$.



(c) Failure case. Inability to handle large and abrupt transitions.

REFERENCES

[1] Karras, Tero. "Maximizing parallelism in the construction of BVHs, octrees, and k-d trees." *Proceedings of the Fourth ACM SIGGRAPH/Eurographics conference on High-Performance Graphics*. 2012.

[2] Mitra, Niloy J., and Mark Pauly. "Shadow art." *ACM Transactions on Graphics* 28.CONF (2009): 156-1.

[3] Sadekar, Kaustubh, Ashish Tiwari, and Shanmuganathan Raman. "Shadow art revisited: a differentiable rendering based approach." *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2022.

[4] Romero, Javier, Dimitrios Tzionas, and Michael J. Black. "Embodied hands: Modeling and capturing hands and bodies together." *arXiv preprint arXiv:2201.02610* (2022).

[5] Tzionas, Dimitrios, et al. "Capturing hands in action using discriminative salient points and physics simulation." *International Journal of Computer Vision* 118 (2016): 172-193.

CVIG
IIT Gandhinagar