PlatonicGAN

Escaping Plato's Cave: 3D Shapes From Adversarial Rendering

Philipp Henzler **Niloy J. Mitra** **Tobias Ritschel**

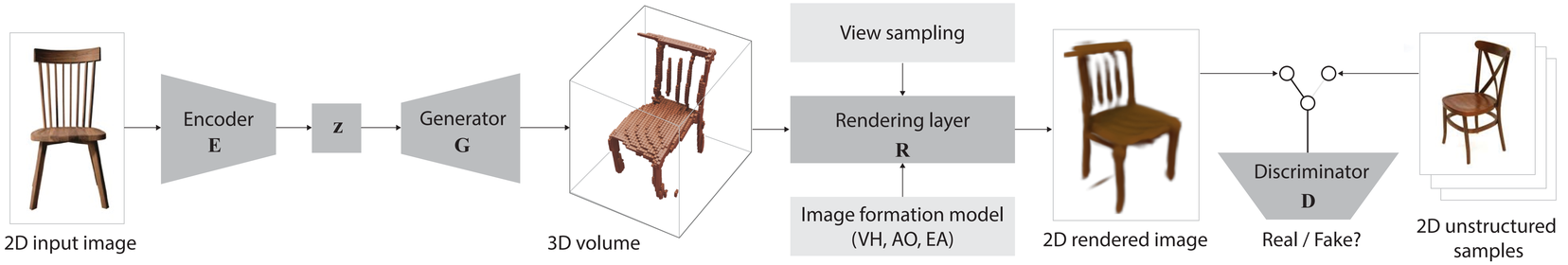
University College London

Motivation

Input: 2D image collection (different object, view, light, camera, etc.)

Output: Generative 3D model

Overview



Rendering layers

a) Visual Hull

$$\rho_{\mathrm{VH}}(\mathbf{v}) = 1 - \exp(\sum_i -v_i)$$

b) Absorption Only

$$\rho_{\mathrm{AO}}(\mathbf{v}) = 1 - \prod_i (1 - v_i)$$

c) Emission Absorption

$$\rho_{\mathrm{EA}}(\mathbf{v}) = \sum_{i=1}^{n_z} (1 - \prod_{j=1}^{i} (1 - v_{\mathrm{a},j}))$$

Algorithm

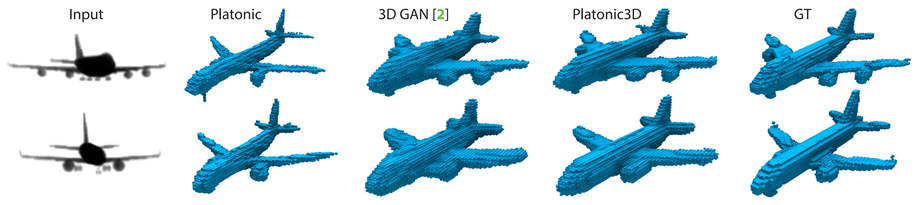
Algorithm: PLATONICGAN Reconstruction Update Step

$I_{\mathrm{Dat}} \leftarrow \mathrm{SAMPLEIMAGE}(p_{\mathrm{Dat}})$
$\omega \leftarrow \mathrm{SAMPLEVIEW}(p_{\mathrm{View}})$
$z \leftarrow \mathrm{E}(I_{\mathrm{Dat}})$
$v \leftarrow \mathrm{G}(z)$
$I_{\mathrm{View}} \leftarrow \mathrm{R}(\omega, v)$
$I_{\mathrm{Front}} \leftarrow \mathrm{R}(\omega_0, v)$
$c_{\mathrm{Dis}} \leftarrow \log D(I_{\mathrm{Dat}}) + \log(1 - D(I_{\mathrm{View}}))$
$c_{\mathrm{Gen}} \leftarrow \log(1 - D(I_{\mathrm{View}}))$
$c_{\mathrm{Rec}} \leftarrow \mathrm{L2}(I_{\mathrm{Dat}} - I_{\mathrm{Front}})$
$\Psi \leftarrow \mathrm{MAXMIZE}(c_{\mathrm{Dis}})$
$\Theta, \Phi \leftarrow \mathrm{MINIMIZE}(c_{\mathrm{Gen}} + \lambda c_{\mathrm{Rec}})$

Contribution

1. Adversarial training of 3D generator with 2D discriminator that operates exclusively on widely available unstructured collections of images, i.e. no relation among images is known.
2. Family of differentiable rendering layers.

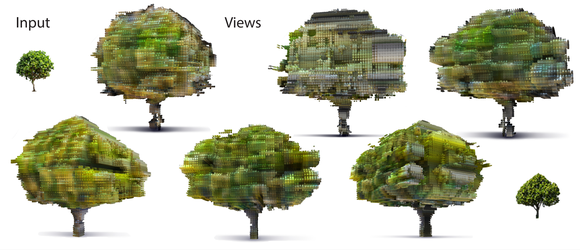
Comparison



Degree of supervision

Supvervision at training time	[1] Kanazawa et al.	[2] Kato et al.	[3] Eslami et al.	[4] Tulsiani et al.	[5] Tulsiani et al.	[6] PrGAN	[Ours]
Annotation-free	-	+	+	+	+	+	+
3D template-free	-	-	+	+	+	+	+
Unknown camera pose	+	-	-	-	+	+	+
No pre-defined camera poses	+	+	+	+	-	-	+
Only single view required	+	-	-	-	-	+	+
Color	+	+	+	+	-	-	+

Results



References

[1] Angjoo Kanazawa, Shubham Tulsiani, Alexei A. Efros, and Jitendra Malik. Learning category-specific mesh reconstruction from image collections. In ECCV, 2018.
[2] Hiroharu Kato, Yoshitaka Ushiku, and Tatsuya Harada. Neural 3D mesh renderer. In CVPR, 2018.
[3] SM Ali Eslami, Danilo Jimenez Rezende, Frederic Besse,Fabio Viola, Ari S Morcos, Marta Garnelo, Avraham Ruderman, Andrei A Rusu, Ivo Danihelka, Karol Gregor, et al. Neural scene representation and rendering. Science, 2018.
[4] Shubham Tulsiani, Tinghui Zhou, Alexei A Efros, and Ji-tendra Malik. Multi-view supervision for single-view reconstruction via differentiable ray consistency. In CVPR, 2017.
[5] Shubham Tulsiani, Alexei A Efros, and Jitendra Malik. Multi-view consistency as supervisory signal for learning shape andpose prediction. hhln CVPR, 2018.
[6] Matheus Gadelha, Subhransu Maji, and Rui Wang. 3d shapeinduction fromh 2d views of multiple objects. In 3DV, 2016
[7] Jiajun Wu, Chengkai Zhang, Tianfan Xue, Bill Freeman, andJosh Tenenbaum. Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling. In NIPS, 2016.

Acknowledgements

This work was supported by the ERCStarting Grant SmartGeometry, a GPU donation by NVIDIACorporation and a Google AR/VR Research Award.

Email: **p.henzler@cs.ucl.ac.uk**
Website **https://henzler.github.io/**