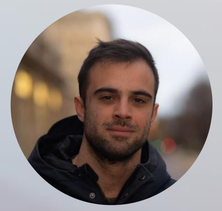
Applications Of Physically Based Differentiable Rendering In Vision Science

Benjamin Beilharz, Joshua Martin & Thomas S. A. Wallis
Perception Lab – Technical University of Darmstadt


Perception Lab

Centre for Cognitive Science

Funded by the European Union

European Research Council
Established by the European Commission

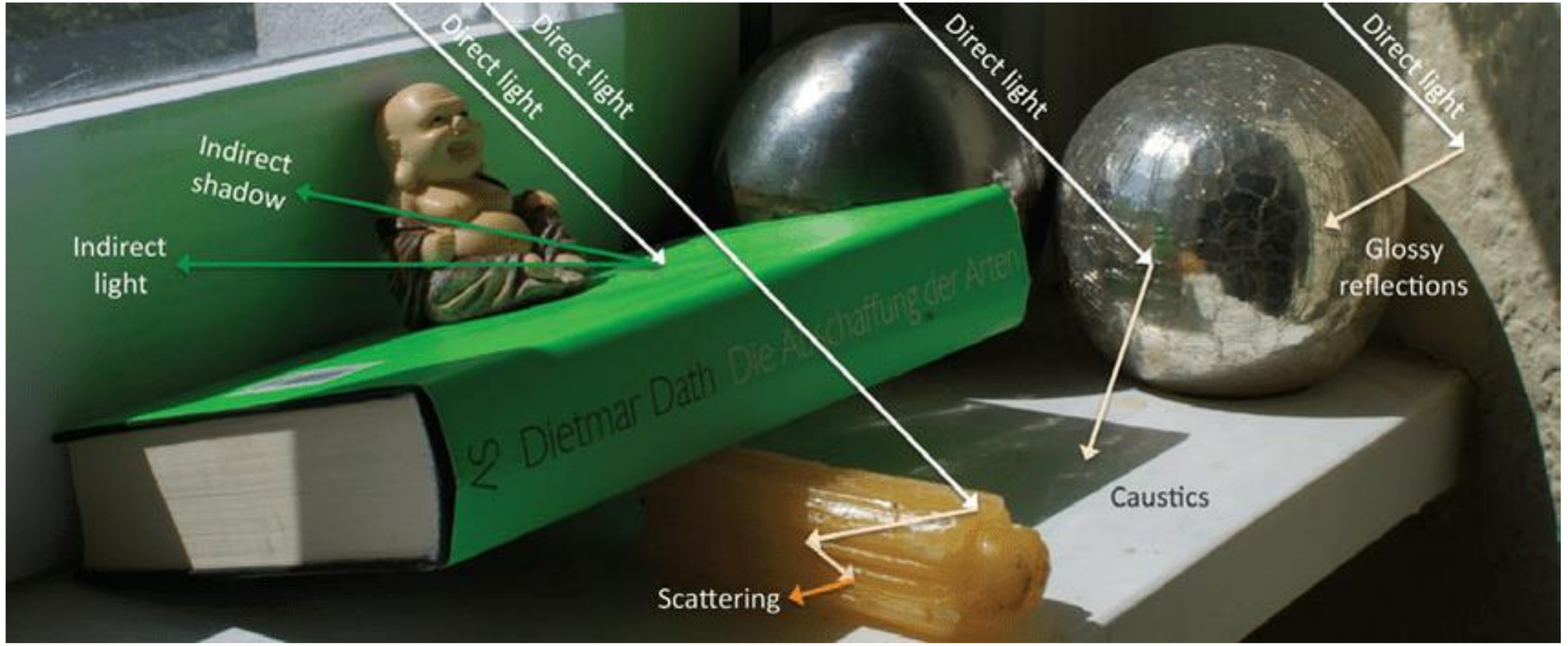
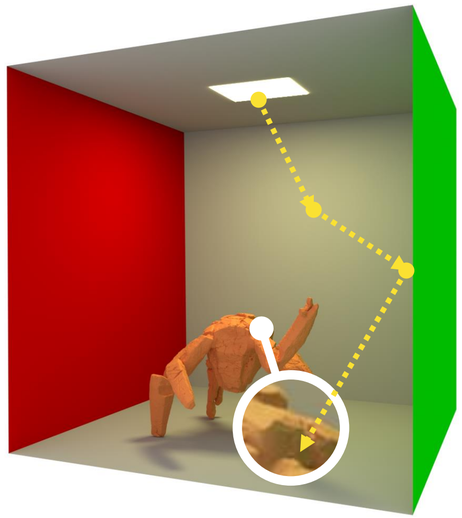
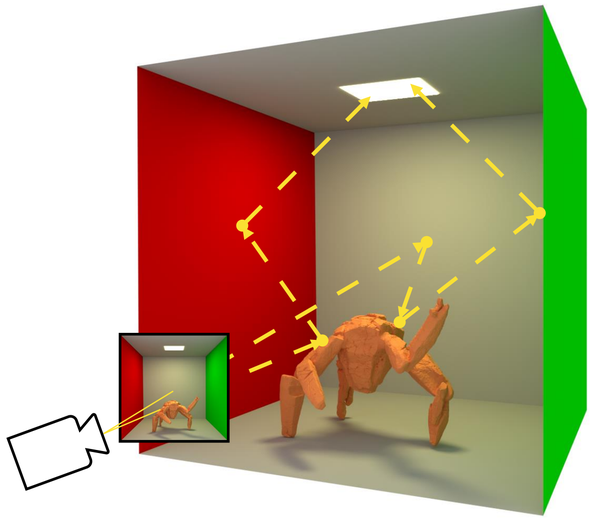


Image courtesy of Tobias Ritschel

Physically Based Rendering

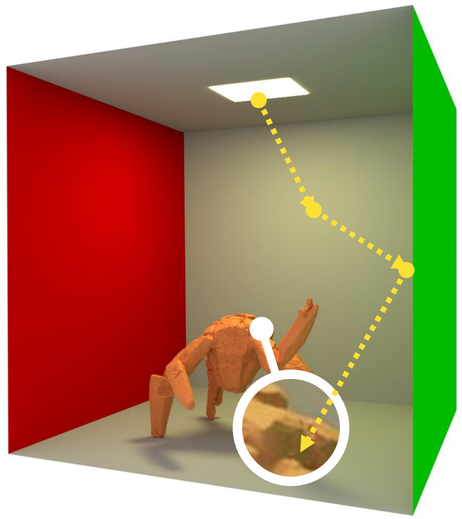


Physically based

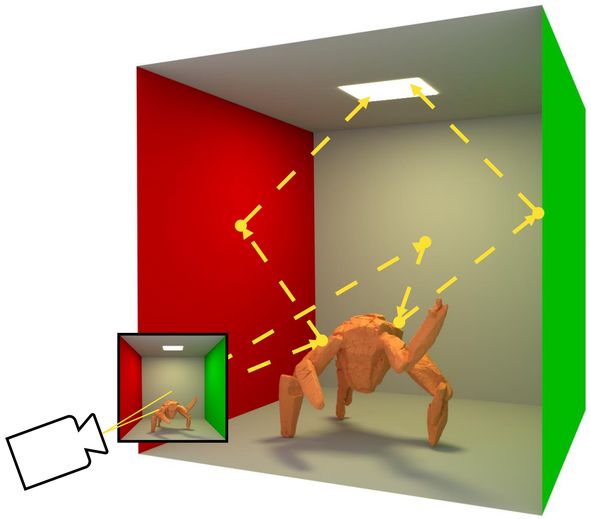


Interpretable

Physically Based Differentiable Rendering



Physically based



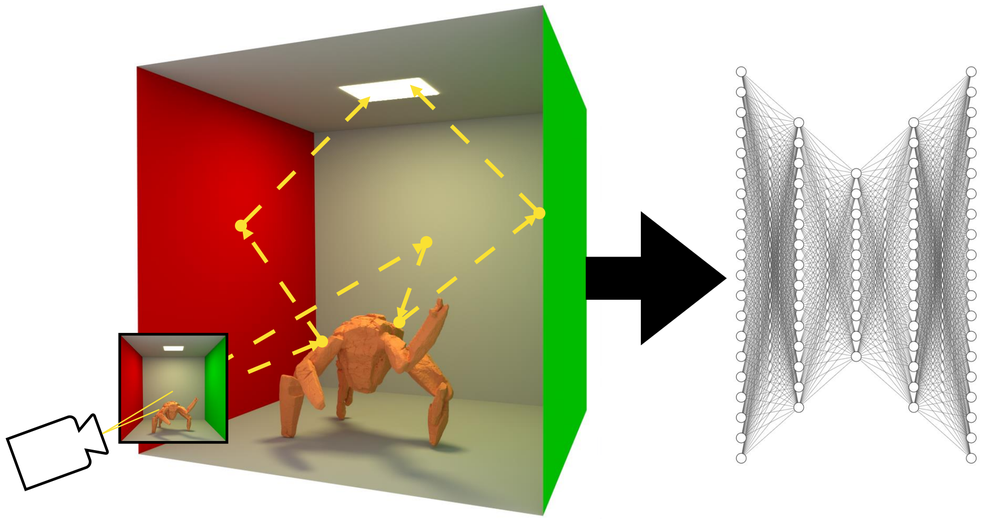
Interpretable

Differentiable

Differentiable Rendering (DR) grounds models in physical parameters.

Encoder

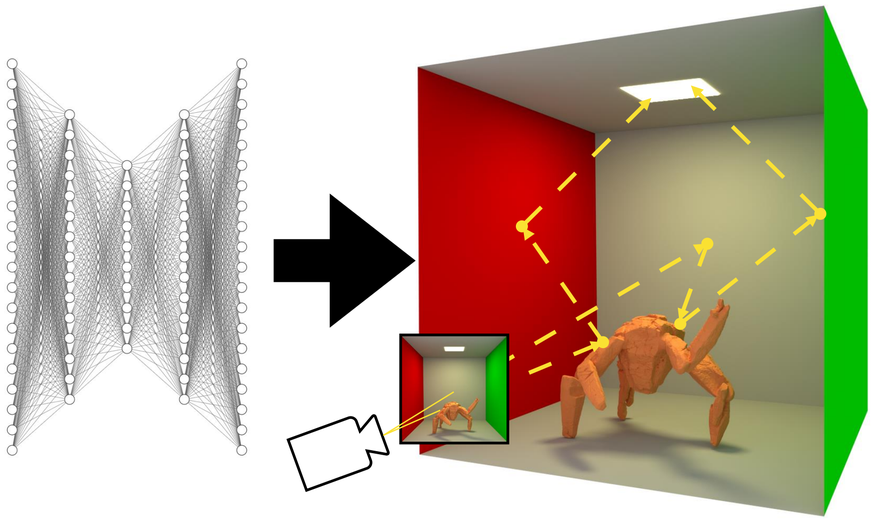
Decoder



DR as the *encoder* can ground the *decoder* in physical parameters. Full control over stimulus generation. End-to-end differentiable.

Encoder

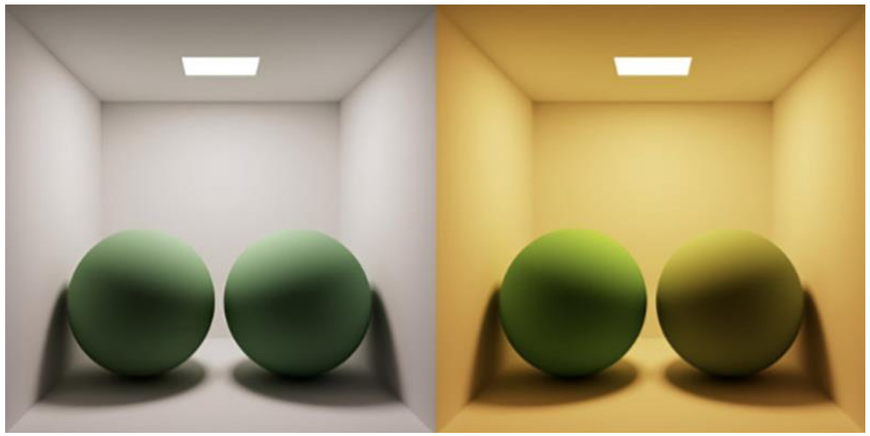
Decoder



The *Encoder* maps modality to physical parameters. DR as the *decoder* grounds the modality in verifiable physical parameters. End-to-end differentiable.

MRD: Using Physically Based Differentiable Rendering to Probe Vision Models for 3D Scene Understanding (accepted at JoV; in press)

Metamers



Colour metamer

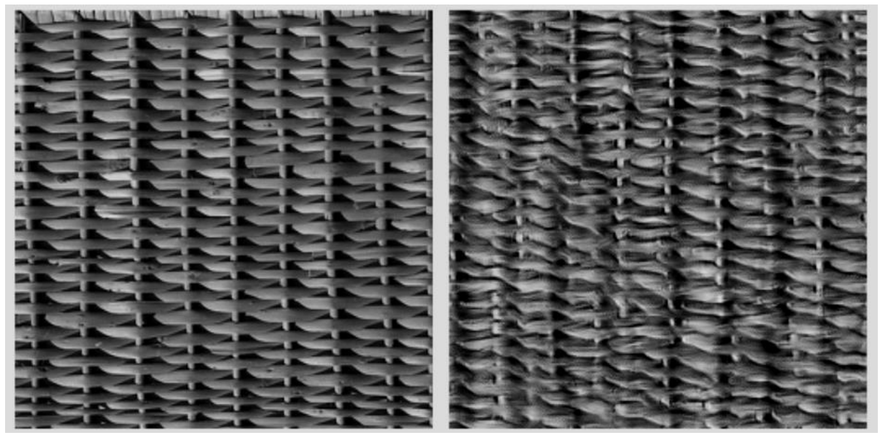
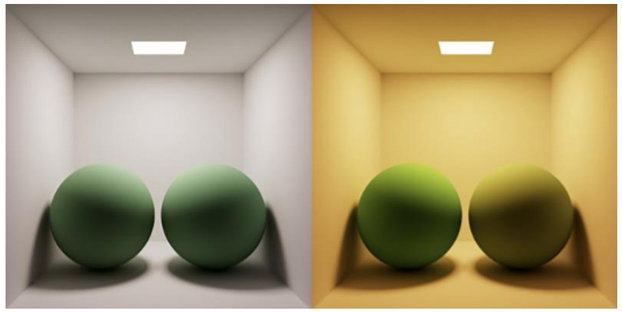


Image statistics metamer (Portilla & Simoncelli 2000, Freeman & Simoncelli 2011)

Neural Network metamer
(Feather et al. 2023)

Metamers



Colour metamer

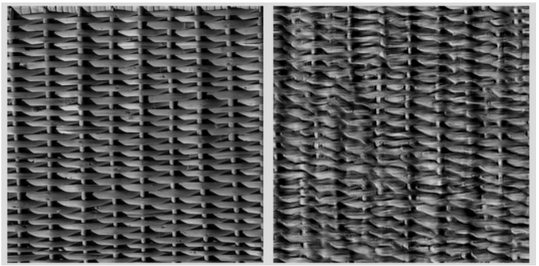
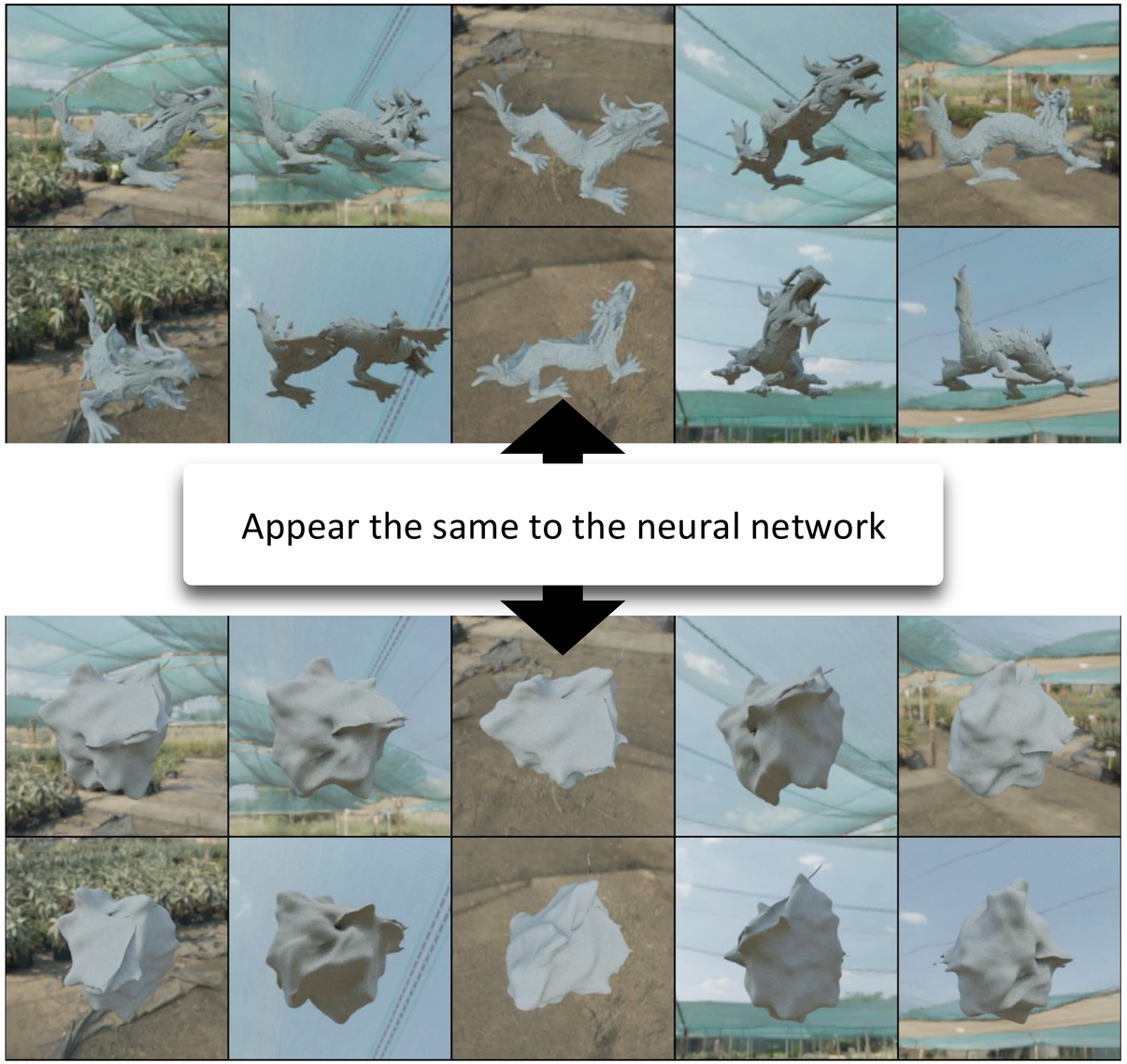


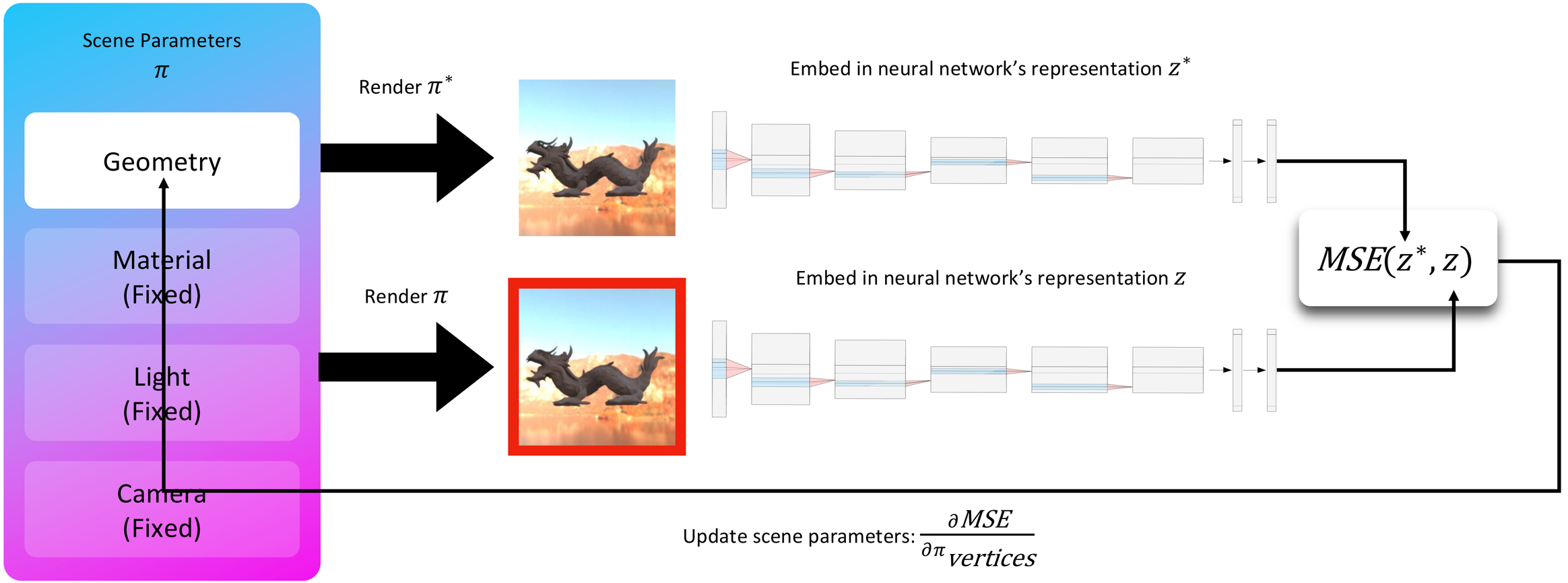
Image statistics metamer (Portilla & Simoncelli 2000, Freeman & Simoncelli 2011)

Neural Network metamer
(Feather et al. 2023)

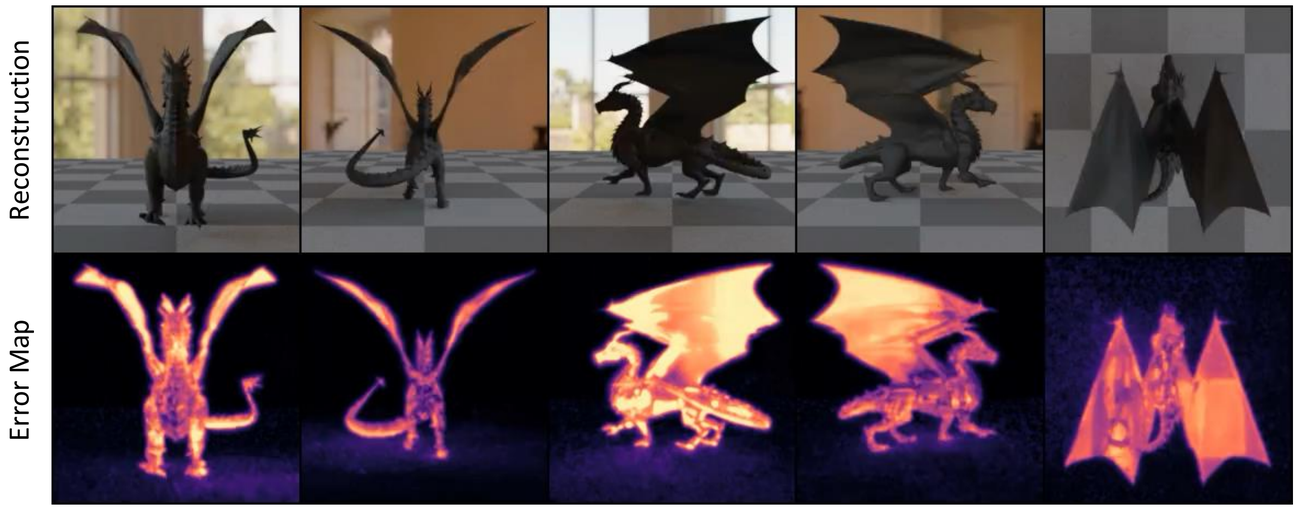


Reconstructing parameters from latent spaces

Geometry Example

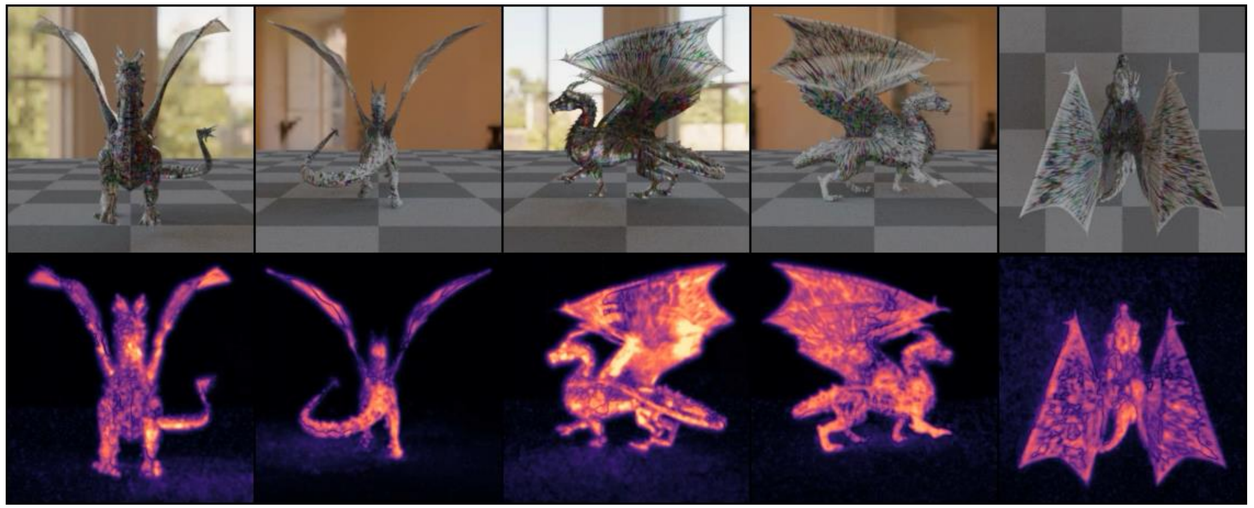


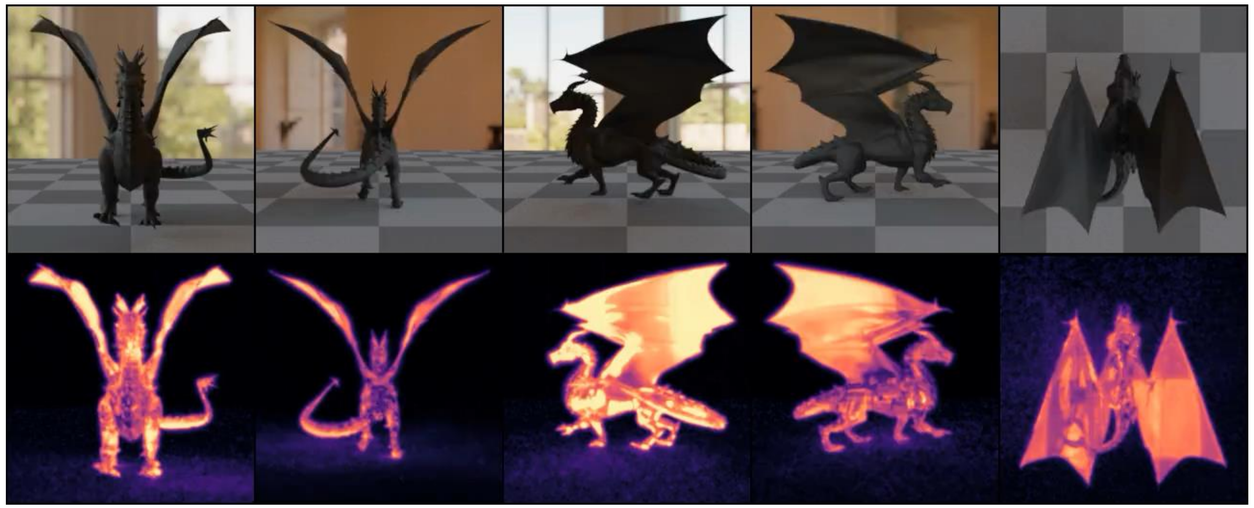
Results



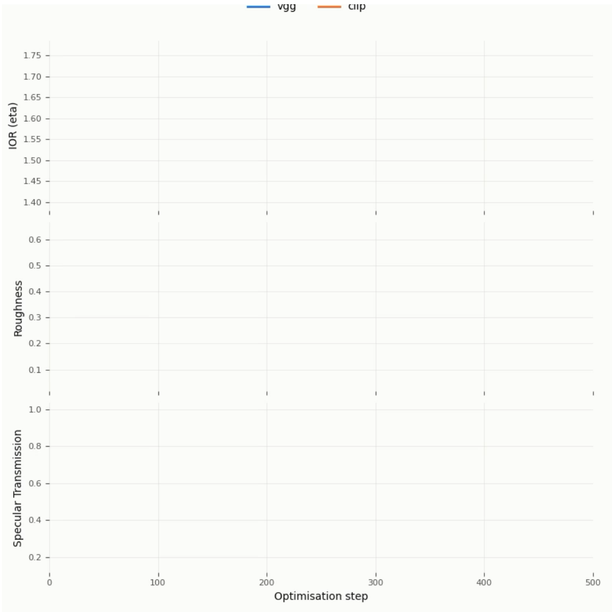
CLIP
Translucent

Results



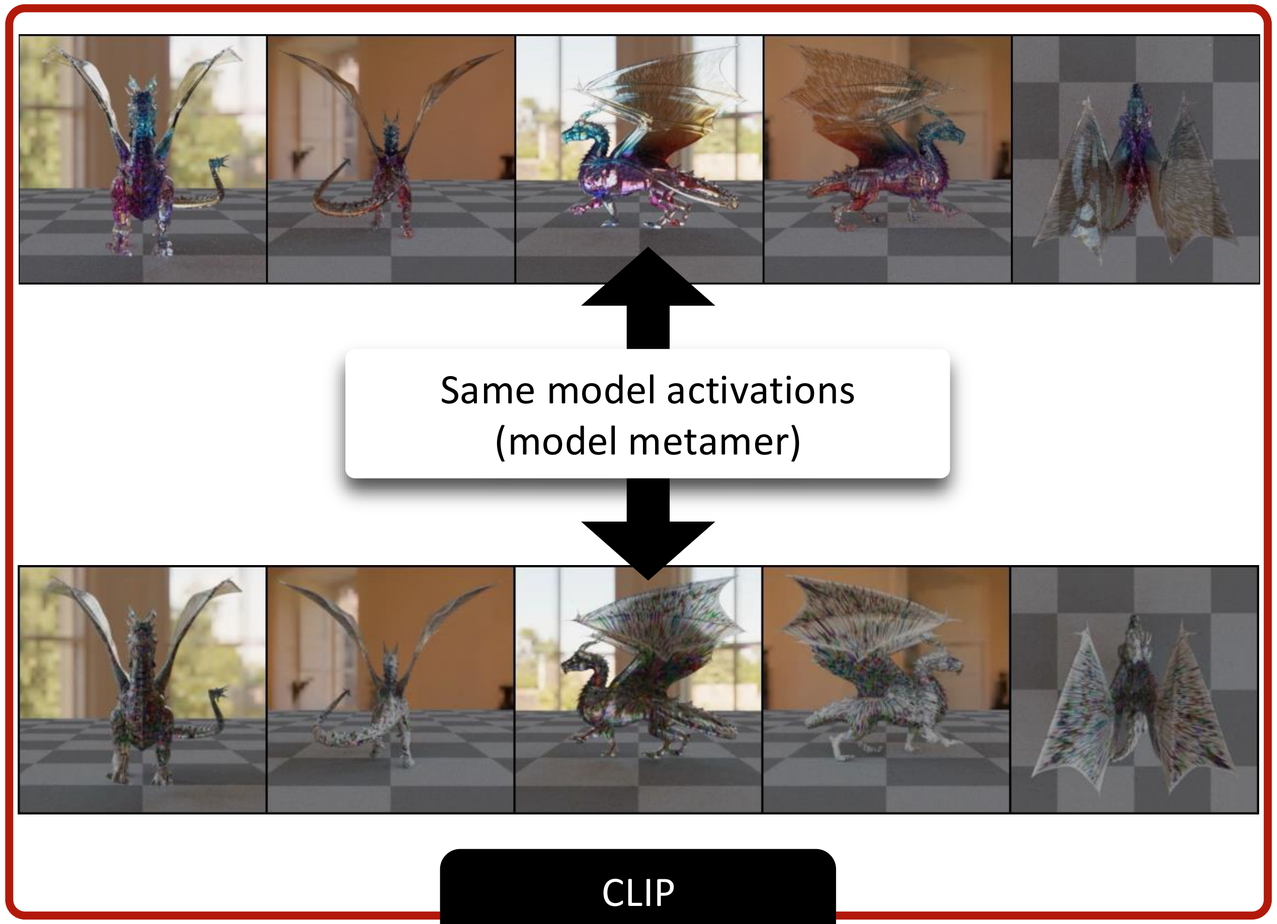


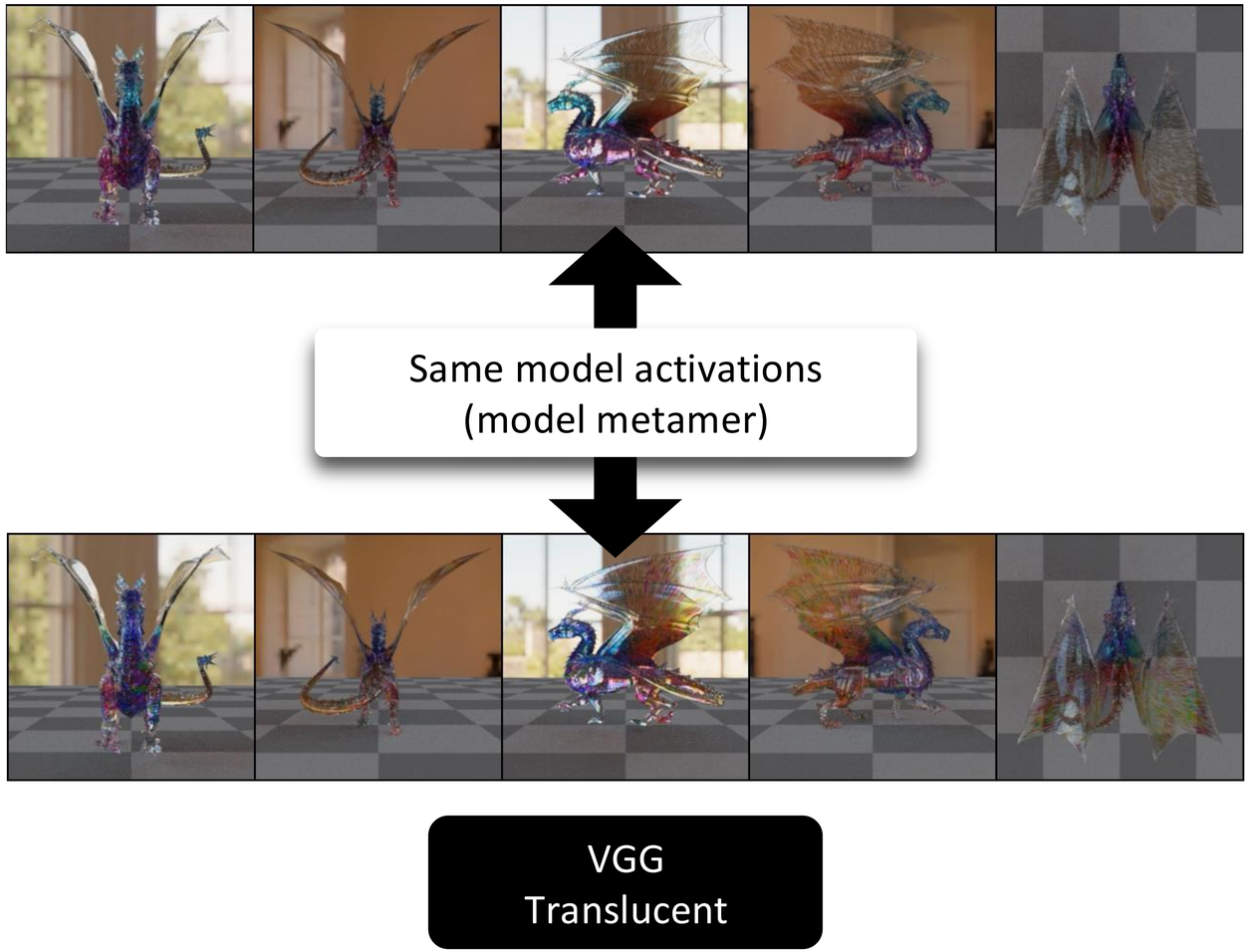
CLIP
Translucent



VGG
Translucent

Results

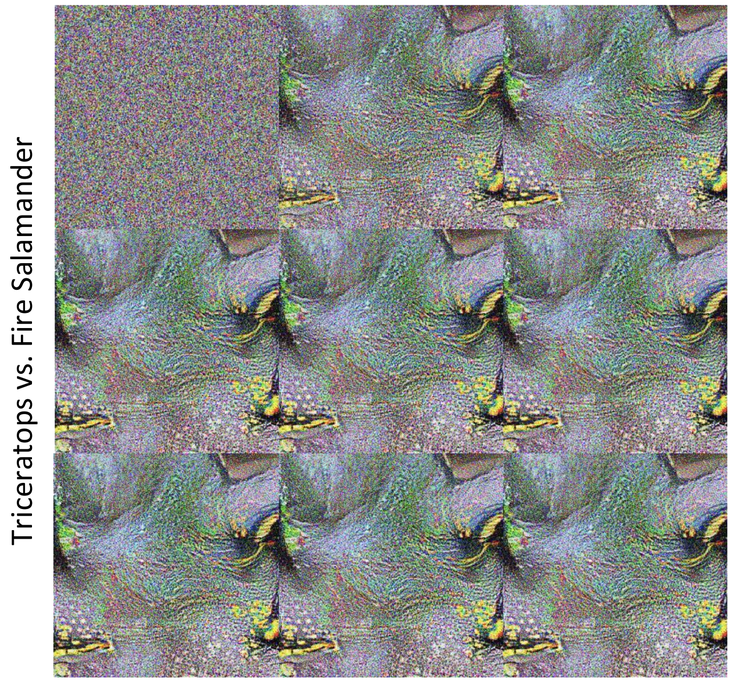

Same model activations
(model metamer)
CLIP
Translucent


Same model activations
(model metamer)
VGG
Translucent

Controversial Scenes
(work in progress)

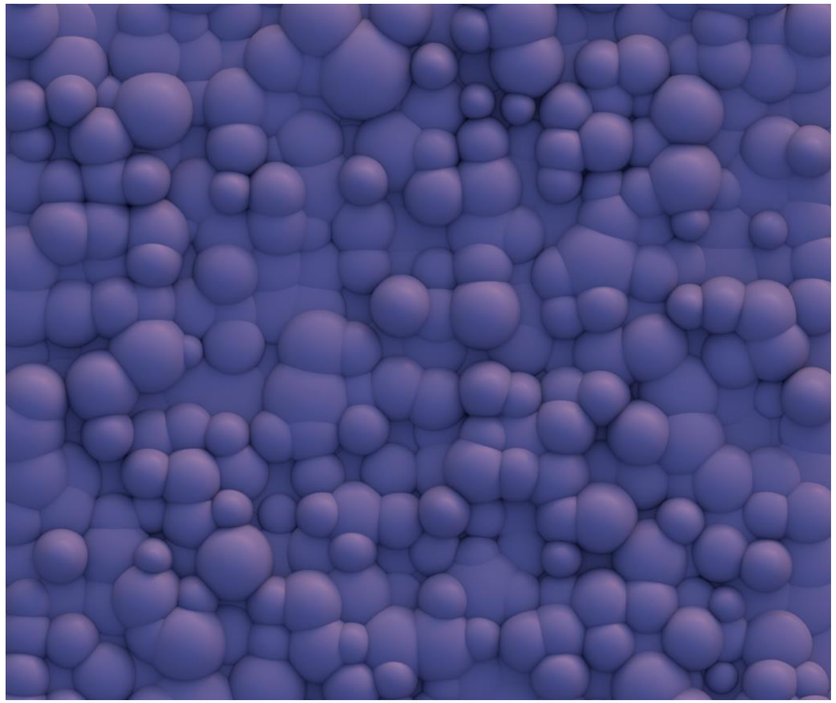
Controversial stimuli: Pitting neural networks against each other as models of human cognition

Tal Golan[a,1], Prashant C. Raju[b], and Nikolaus Kriegeskorte[a,c,d,e,1]

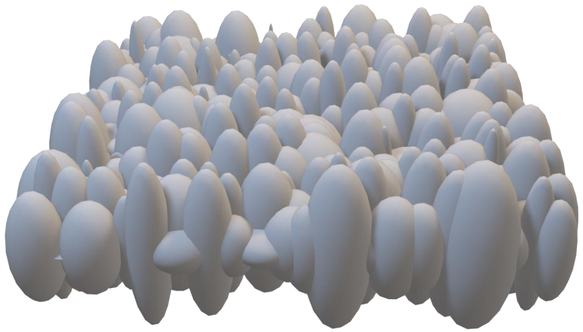


Unsupervised learning predicts human perception and misperception of gloss

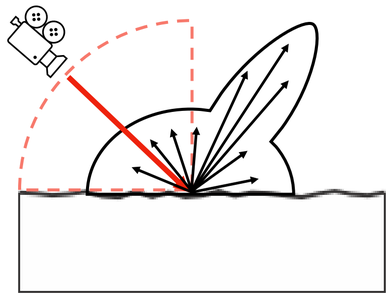
Katherine R. Storrs[1]✉, Barton L. Anderson[2] and Roland W. Fleming[1,3]



Glossiness has multiple factors



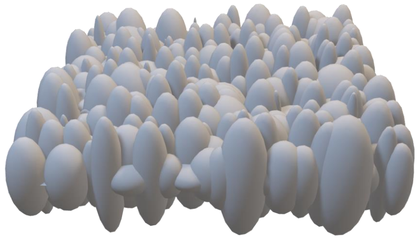
Surface Relief



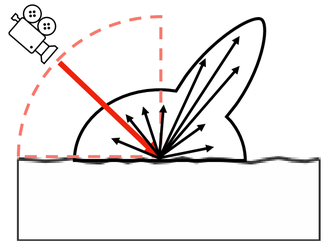
Material & Viewing Direction

Lighting

Dataset Generation

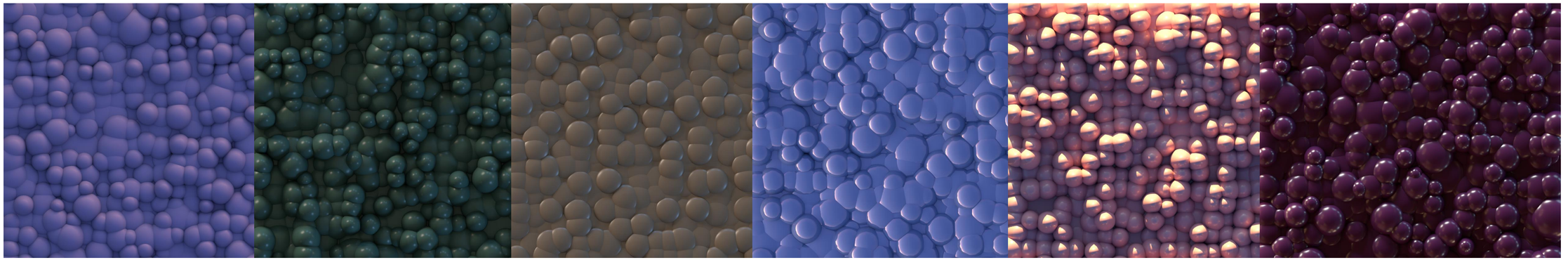


Surface Relief

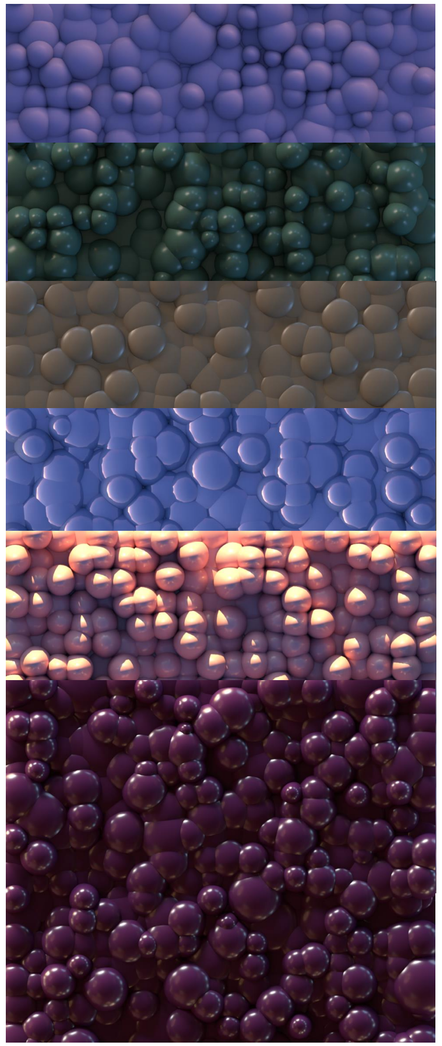


Material & Viewing Direction

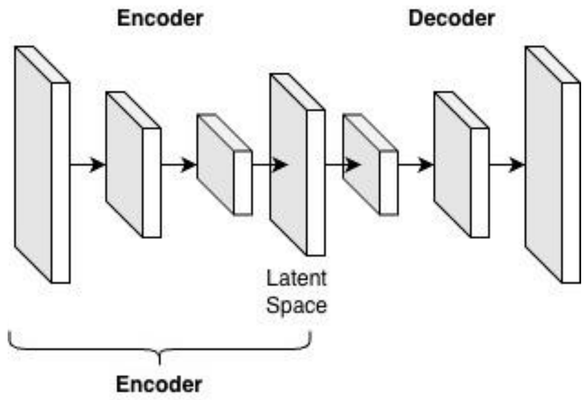
Lighting



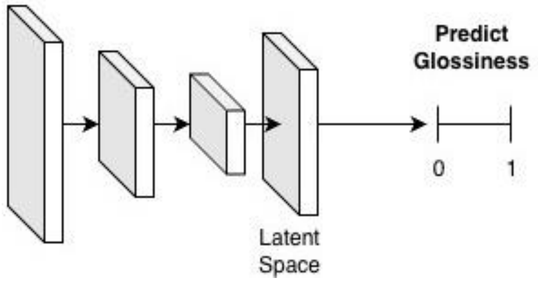
Learning Neural Networks on Glossiness



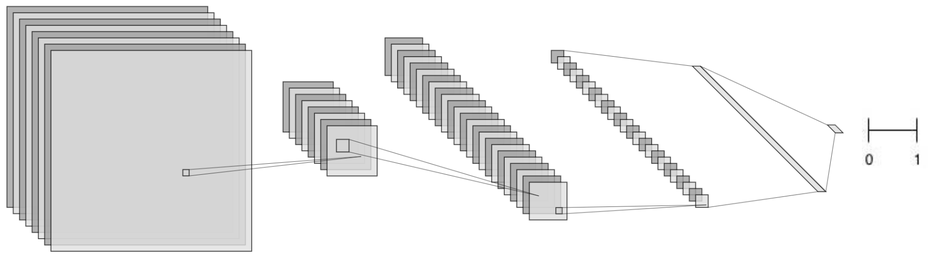
Unsupervised:
(PixelVAE)



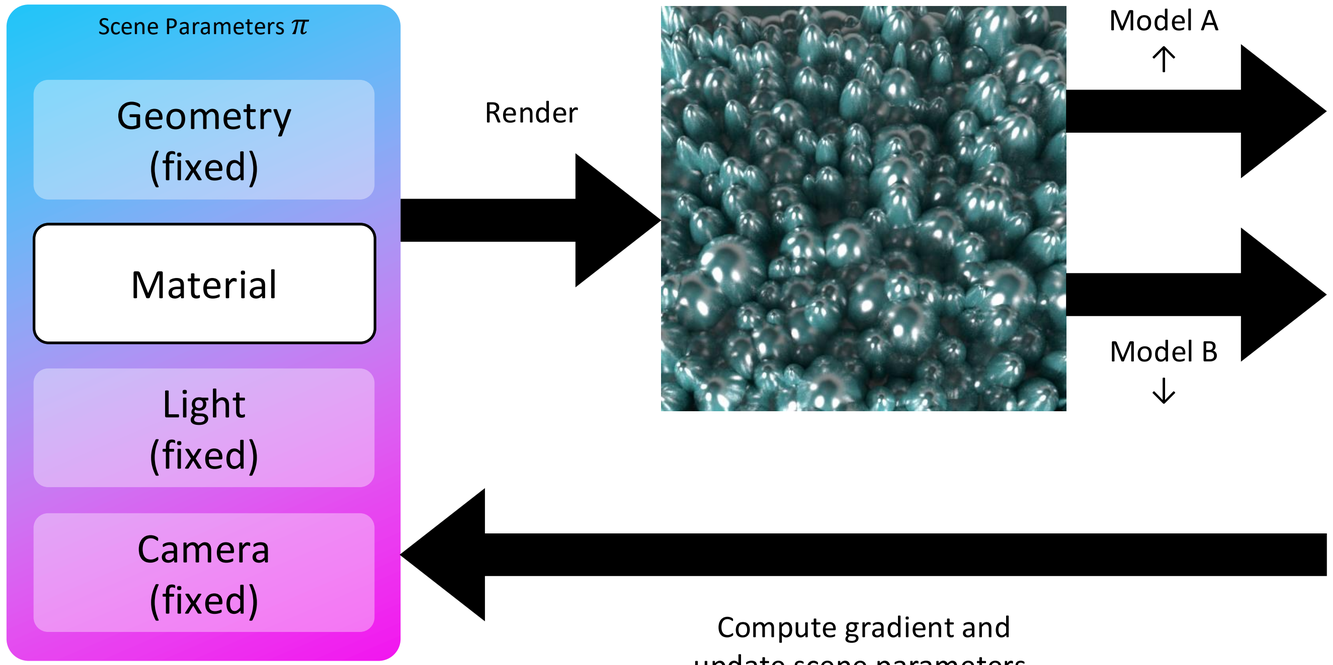
Fine-tuned unsupervised network on glossiness:
(PixelVAE + Regression)

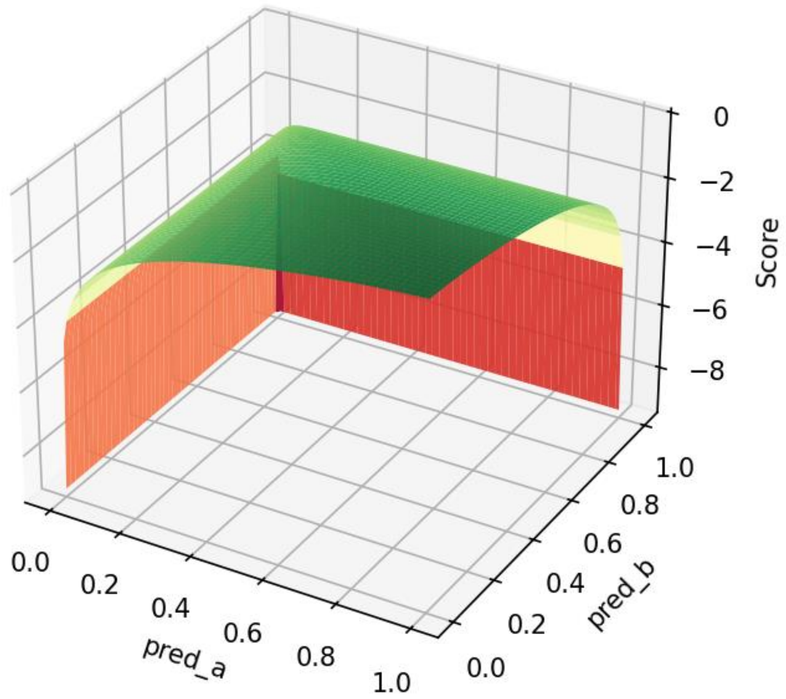


Supervised:
(ConvNext)



Controversial Scene Optimisation

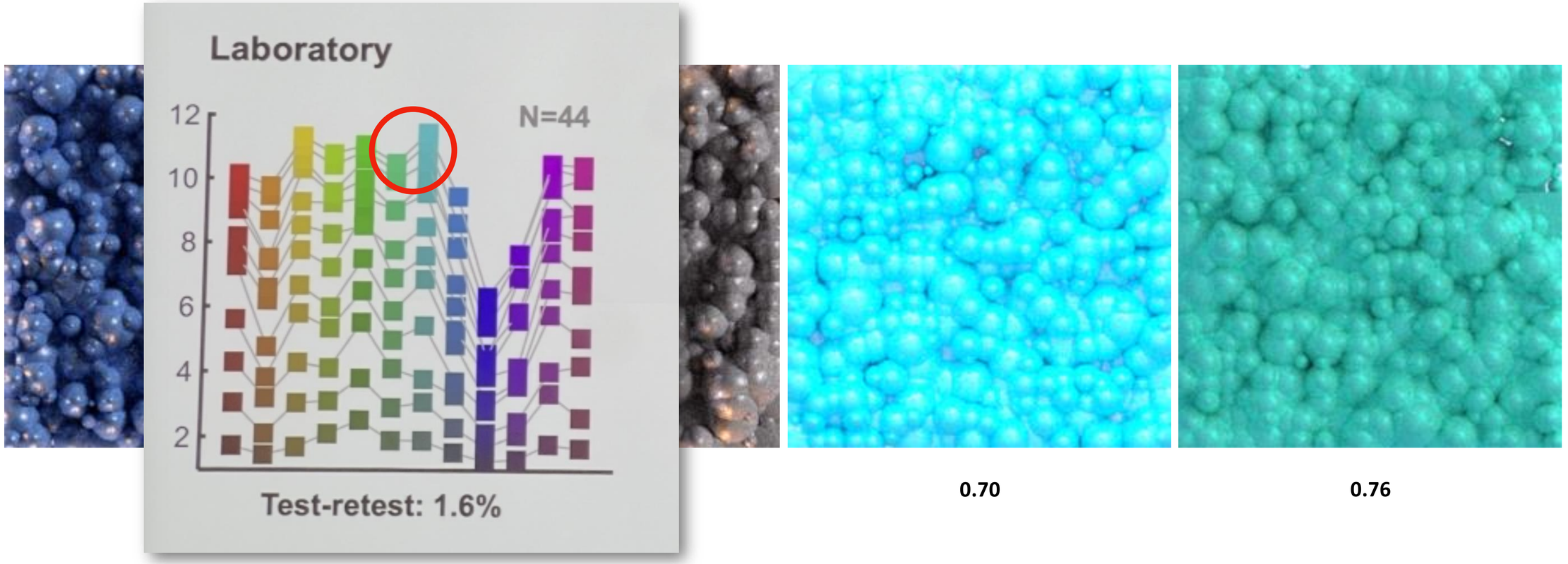




Compute unidirectional disagreement score
(Model A requires to predict high glossiness and Model B low glossiness)

Preliminary Results: Material Parameters

Fine-tuned model predicts high gloss, supervised model predicts low gloss.

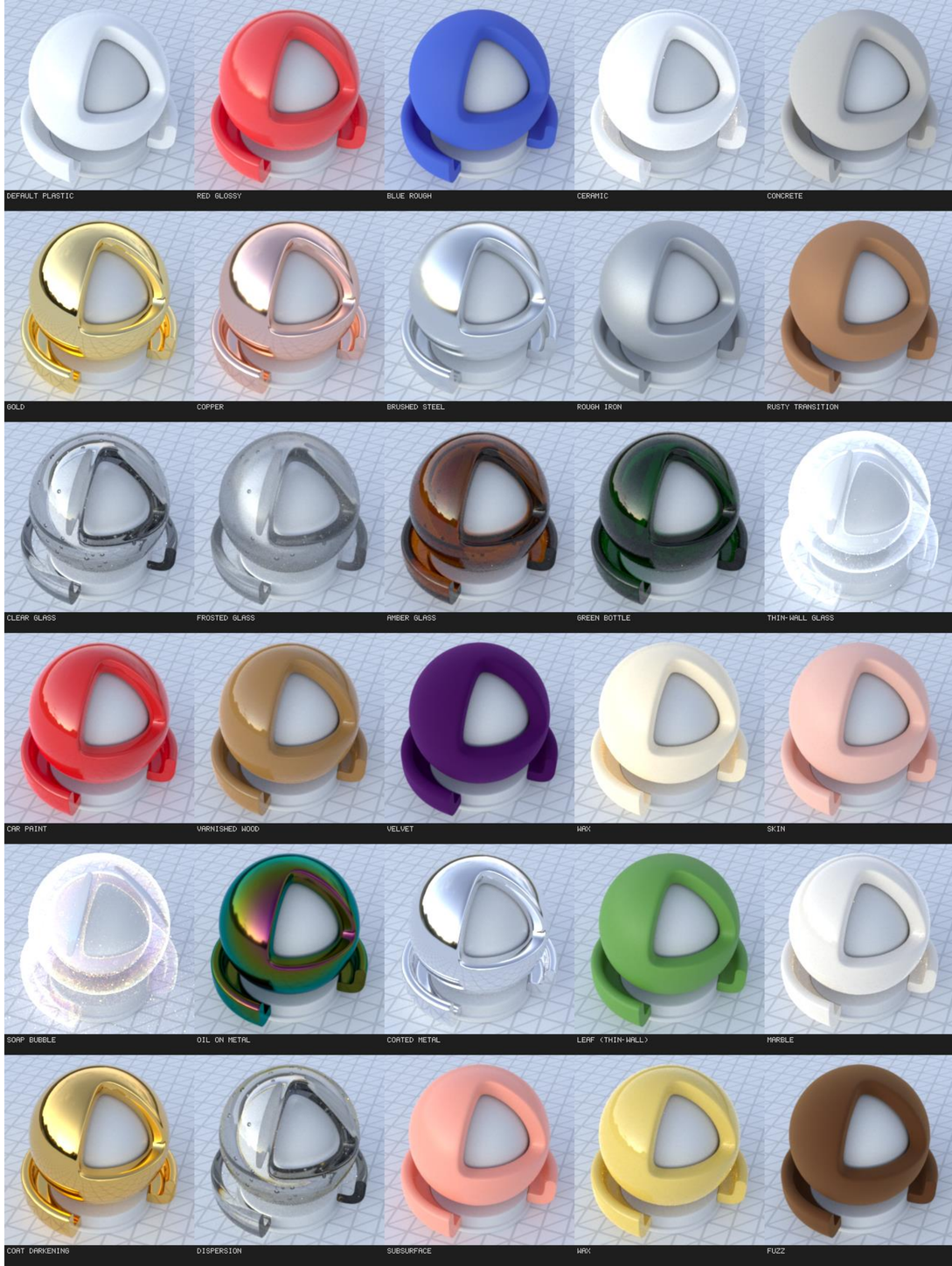


Psychophysics will follow.

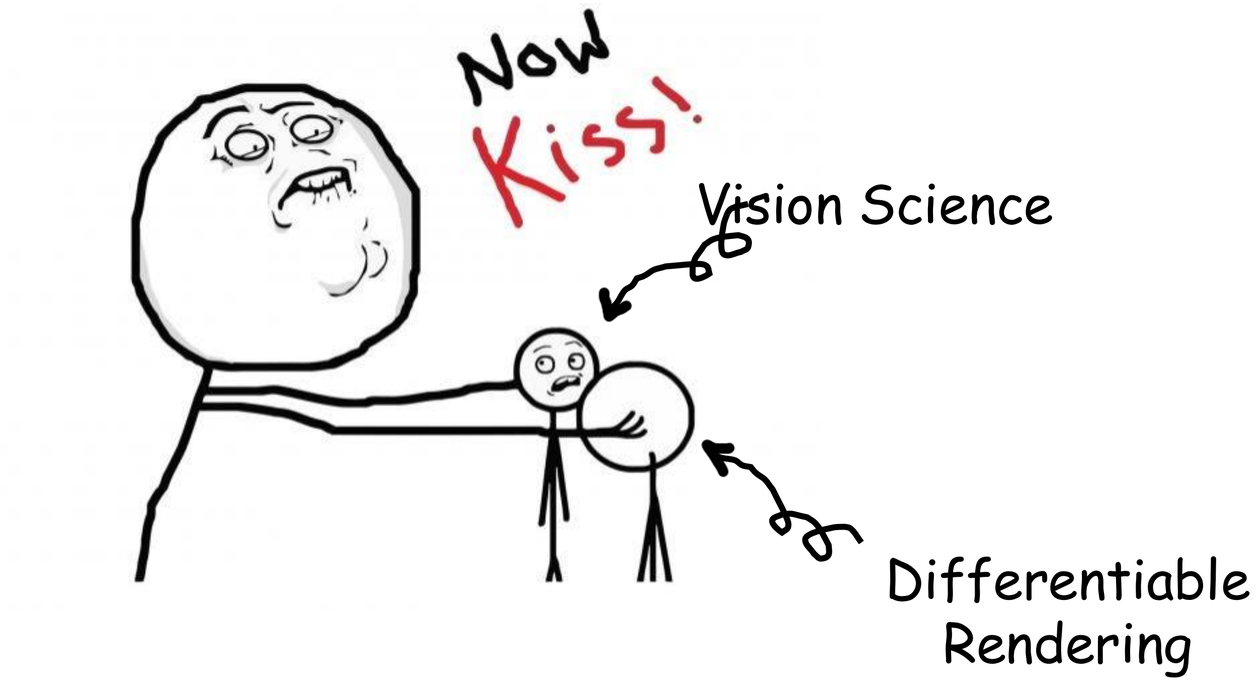
Higher score = higher model disagreement.

Verdict time!

You can always extend existing material models.



Differentiable rendering
grounds models of vision and stimuli
in an interpretable and optimisable physical space.



Thanks!

@engramm.ing

benjamin.beilharz@tu-darmstadt.de

MRD (Repo with paper link)

A primer on physically based differentiable rendering in vision science.