Light field intrinsics with a deep encoder-decoder network

CVPR Submission #2351 - supplementary material

Abstract

In this supplementary material, we show quantitative and qualitative results of our method on several synthetic scenes. To demonstrate the large variety of data our method works on, we evaluate on real world Lytro light fields, the HCI benchmark [3], the light field benchmark [1], datasets from the Stanford archive [2], and our own synthetic datasets created for this submission. We present results for both seen and unseen light fields. Since ground truth is only available for our synthetic scenes and [1], we show quantitative evaluations only for those data sets.

For the diffuse and specular component, we compute local mean-squared error (LMSE), global mean squared error (GMSE) and structural similarity index (SSIM). For LMSE and GMSE, smaller values are better, while for SSIM, a larger value is better. Accuracy of depth maps is assessed with (MSE). For better visualization and consistency with the main paper, all metrics are multiplied by 100. For visual evaluation, we show center views of diffuse and specular components, estimated disparity, and center view of reconstructed light field. On the last two pages, we also link video files to demonstrate angular consistency. Since we use a crosshair-shaped subset of 17 views, the video shows only 17 frames corresponding to the vertical and horizontal directions in the video files.

Note: text appearing in red is a link and can be used to navigate the document. Text appearing in blue is a link to a video file.

References

[1] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke. A dataset and evaluation methodology for depth estimation on 4D light fields. In *Asian Conf. on Computer Vision*, 2016.

[2] V. Vaish and A. Adams. The (New) Stanford Light Field Archive. `http://lightfield.stanford.edu`, 2008.

[3] S. Wanner, S. Meister, and B. Goldluecke. Datasets and benchmarks for densely sampled 4D light fields. In *Vision, Modelling and Visualization (VMV)*, 2013.

Contents

Light field *First training example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

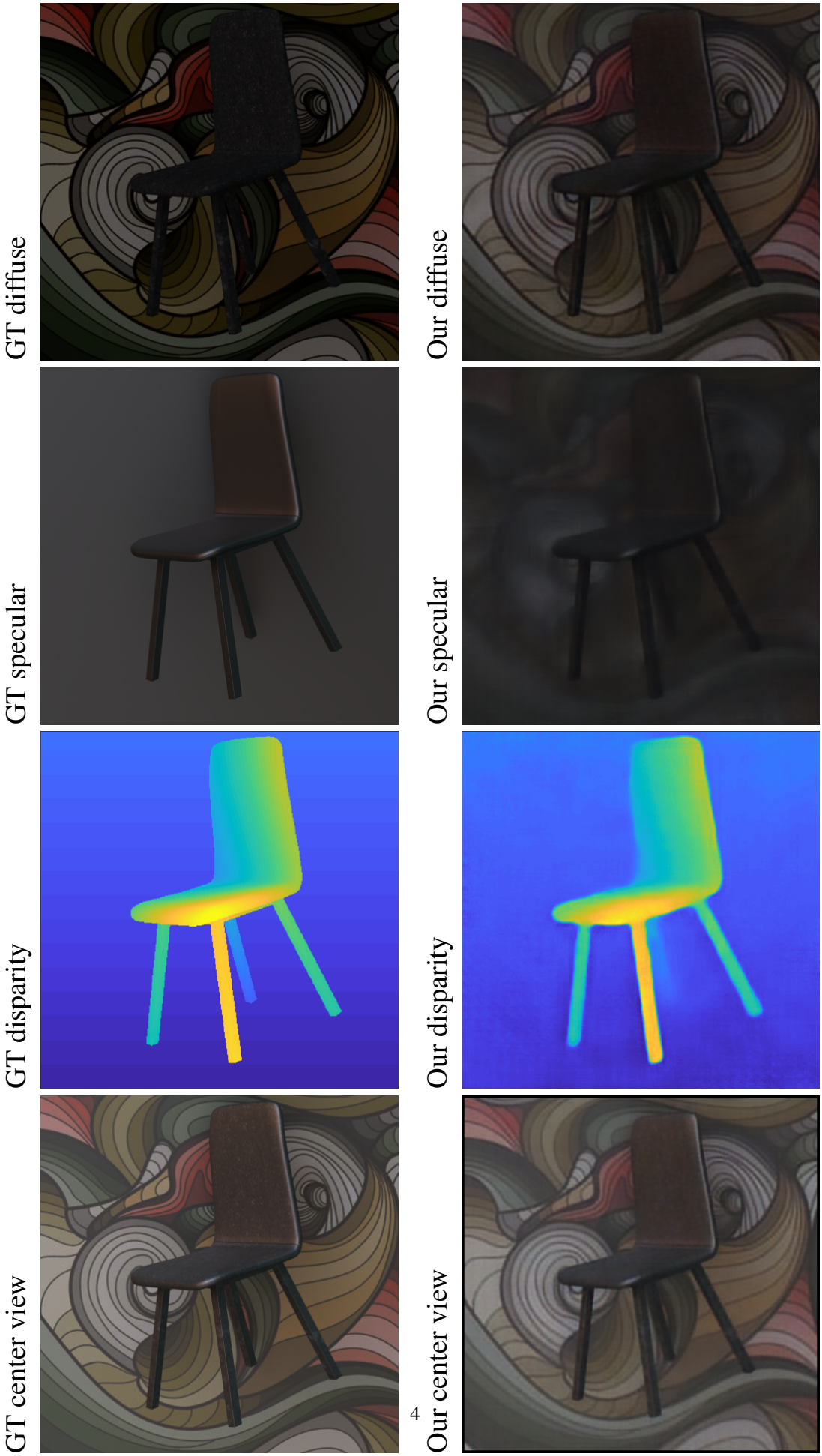
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-2.55 \;\; 1.44]$

Scene description: The data set contains a glossy object with complicated geometry, and a mixture of planar and curved surfaces. The foreground object has almost no structure. This light field was used for training the network.

Light field preview


GT diffuse
GT specular
GT disparity
GT center view
Our diffuse
Our specular
Our disparity
Our center view

- **Diffuse**

  LMSE = 0.2

  GMSE = 0.24

  SSIM = 67.58

- **Specular**

  LMSE = 0.05

  GMSE = 0.08

  SSIM = 84.26

- **Disparity**

  MSE = 3.92

Light field *Second training example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

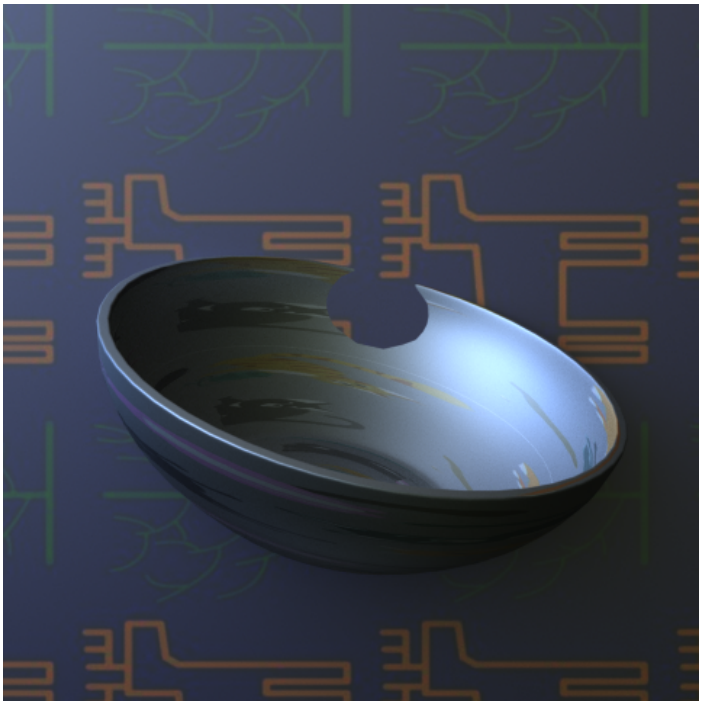
Size: 9×9 views

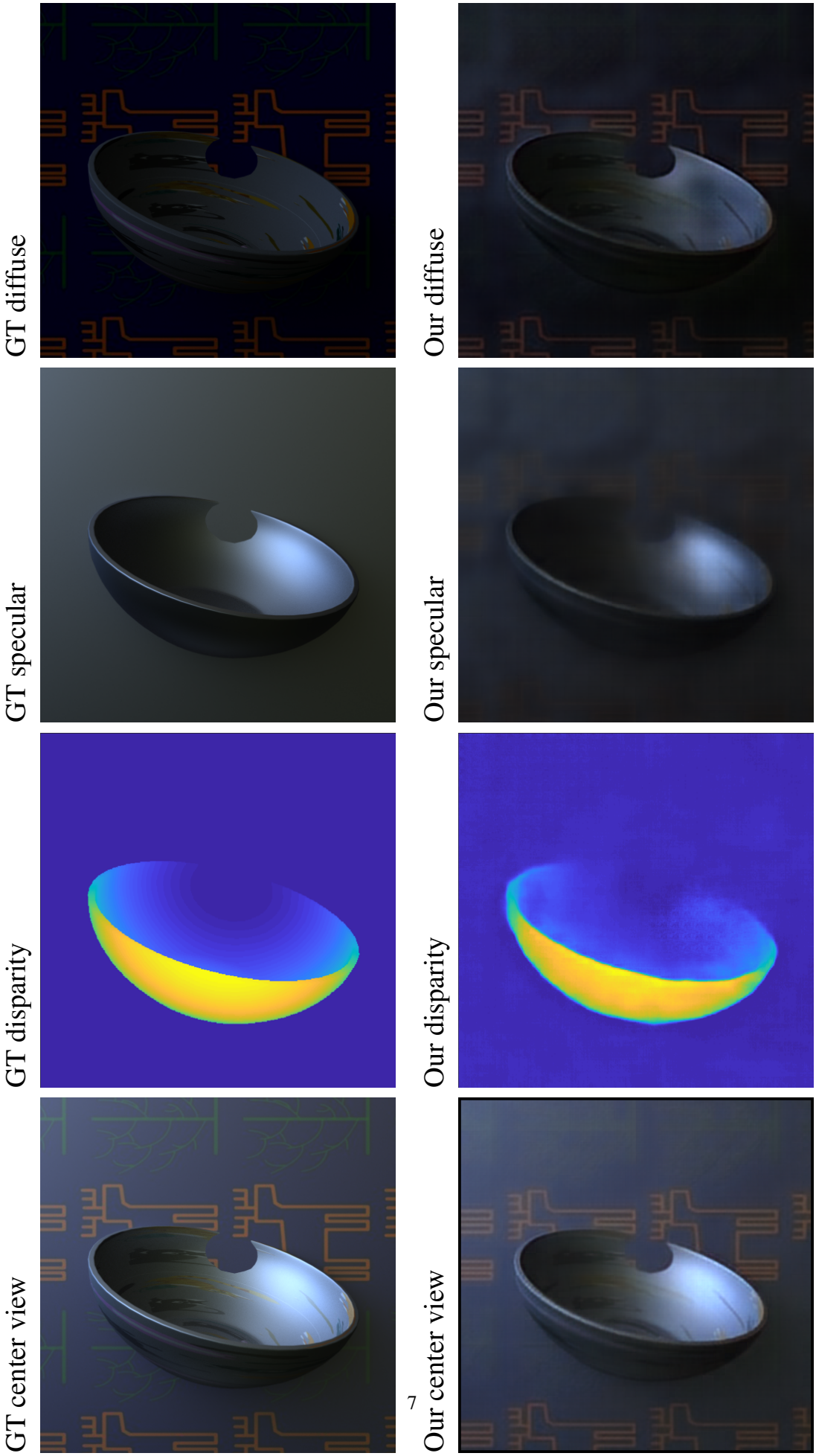
Resolution: 512×512 pixels

Disparity range: $[-2.04 \;\; 2.1]$

Scene description: The data set has an object with soft reflection. The background contains some texture with clear pattern. This data set is used for training the network.

Light field preview




GT diffuse
GT specular
GT disparity
GT center view
Our diffuse
Our specular
Our disparity
Our center view

- **Diffuse**

  LMSE = 0.11

  GMSE = 0.14

  SSIM = 56.16

- **Specular**

  LMSE = 0.12

  GMSE = 0.17

  SSIM = 83.23

- **Disparity**

  MSE = 3.45

Light field *First test example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

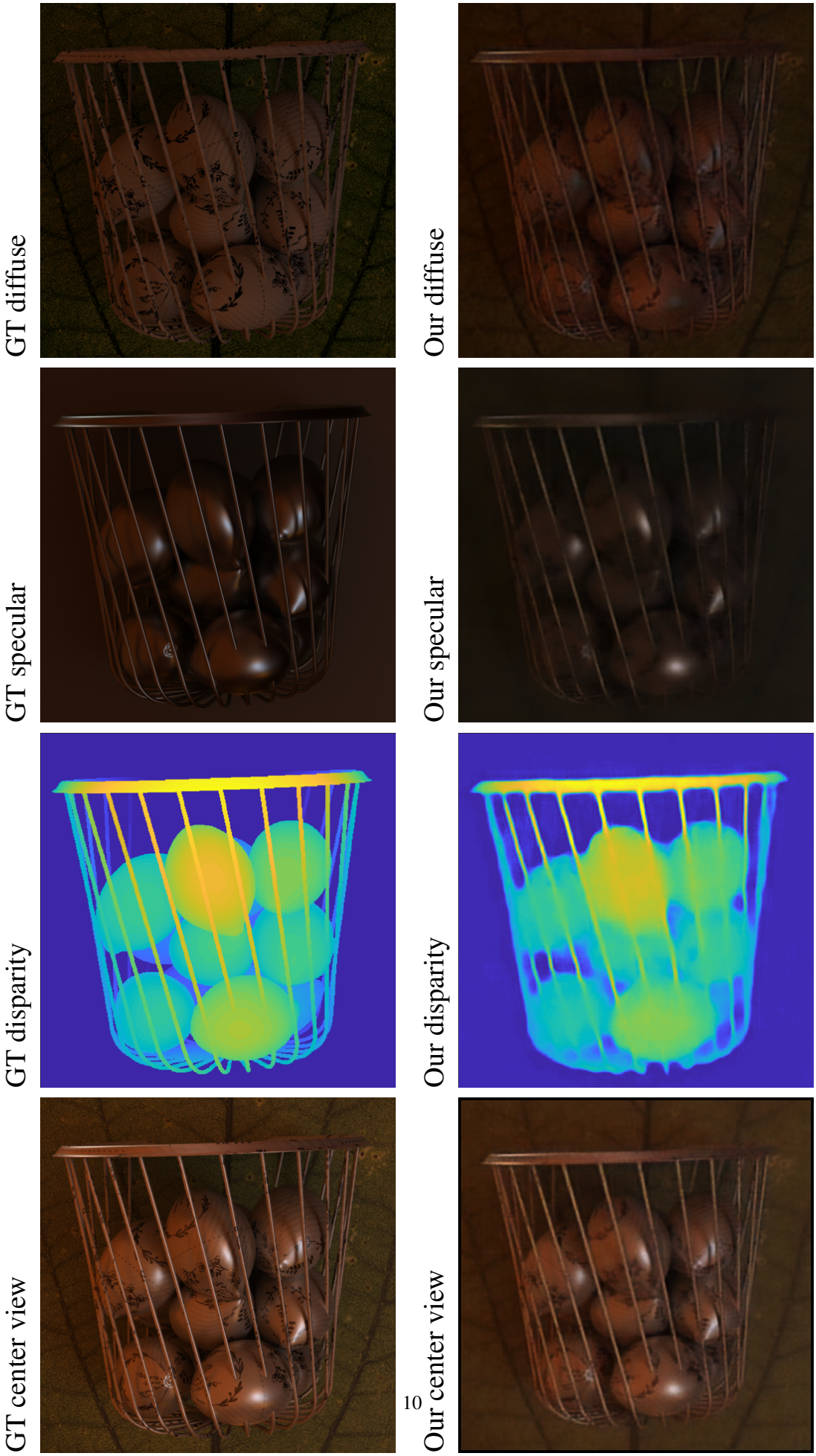
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-2.04 \;\; 2.86]$

Scene description: This data is used in the main paper. With non-trivial illumination, the scene exhibit lots of specularity. The data set has never been shown to the network, and is used for testing.

Light field preview


GT diffuse
Our diffuse
GT specular
Our specular
GT disparity
Our disparity
GT center view
Our center view

- **Diffuse**

  LMSE = 0.11

  GMSE = 0.13

  SSIM = 84.25

- **Specular**

  LMSE = 0.14

  GMSE = 0.15

  SSIM = 77.85

- **Disparity**

  MSE = 18.42

Light field *Second test example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

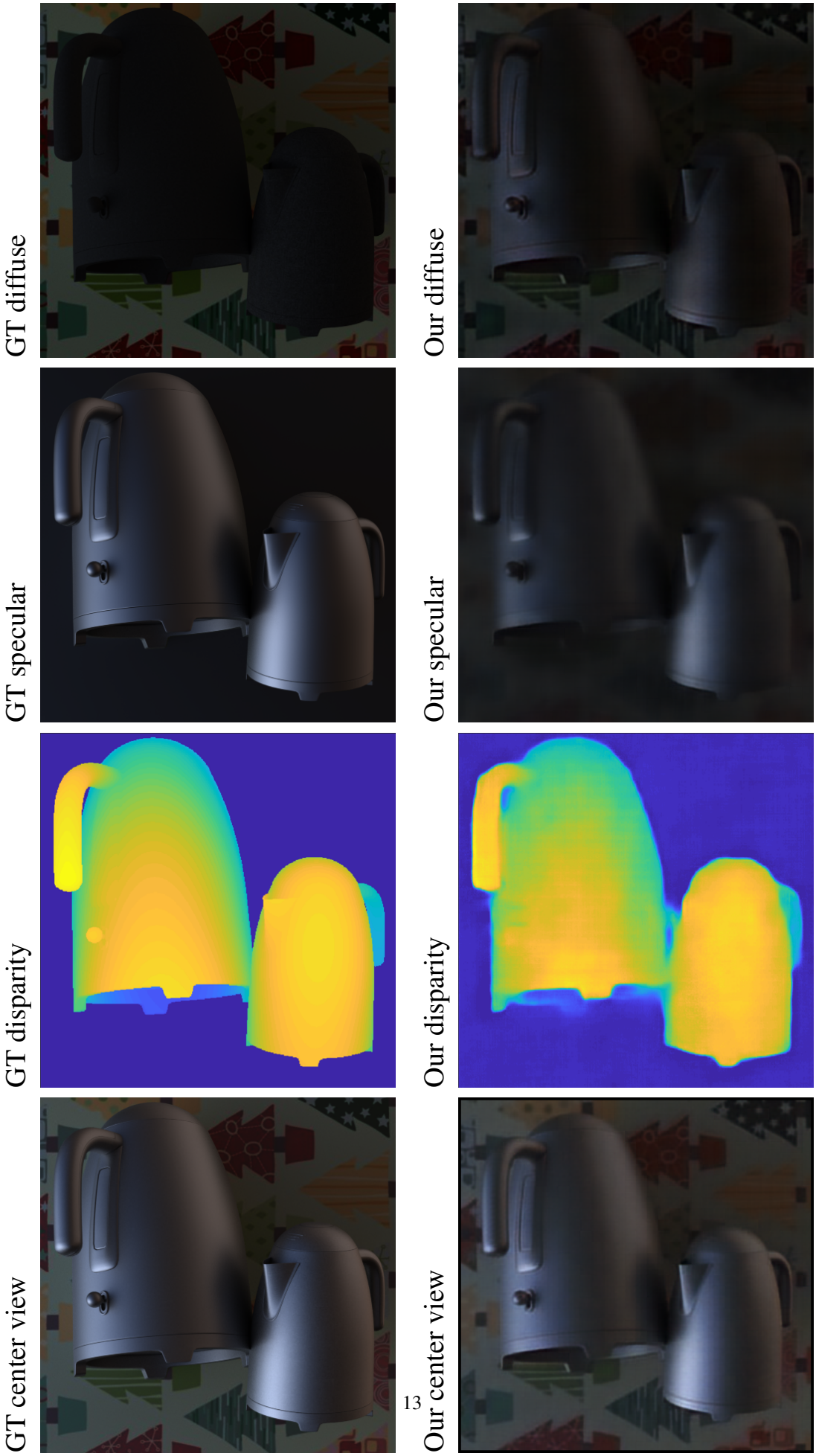
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-2.04 \;\; 1.76]$

Scene description: This data set has two objects and is also used in the main paper for the comparisons with other methods, and reproduced here for better image resolution. This is an unseen light field that is used for testing.

Light field preview


GT diffuse
GT specular
GT disparity
GT center view
Our diffuse
Our specular
Our disparity
Our center view

- **Diffuse**

  LMSE = 0.1

  GMSE = 0.26

  SSIM = 75.03

- **Specular**

  LMSE = 0.13

  GMSE = 0.28

  SSIM = 82.9

- **Disparity**

  MSE = 5.87

Light field *Third test example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

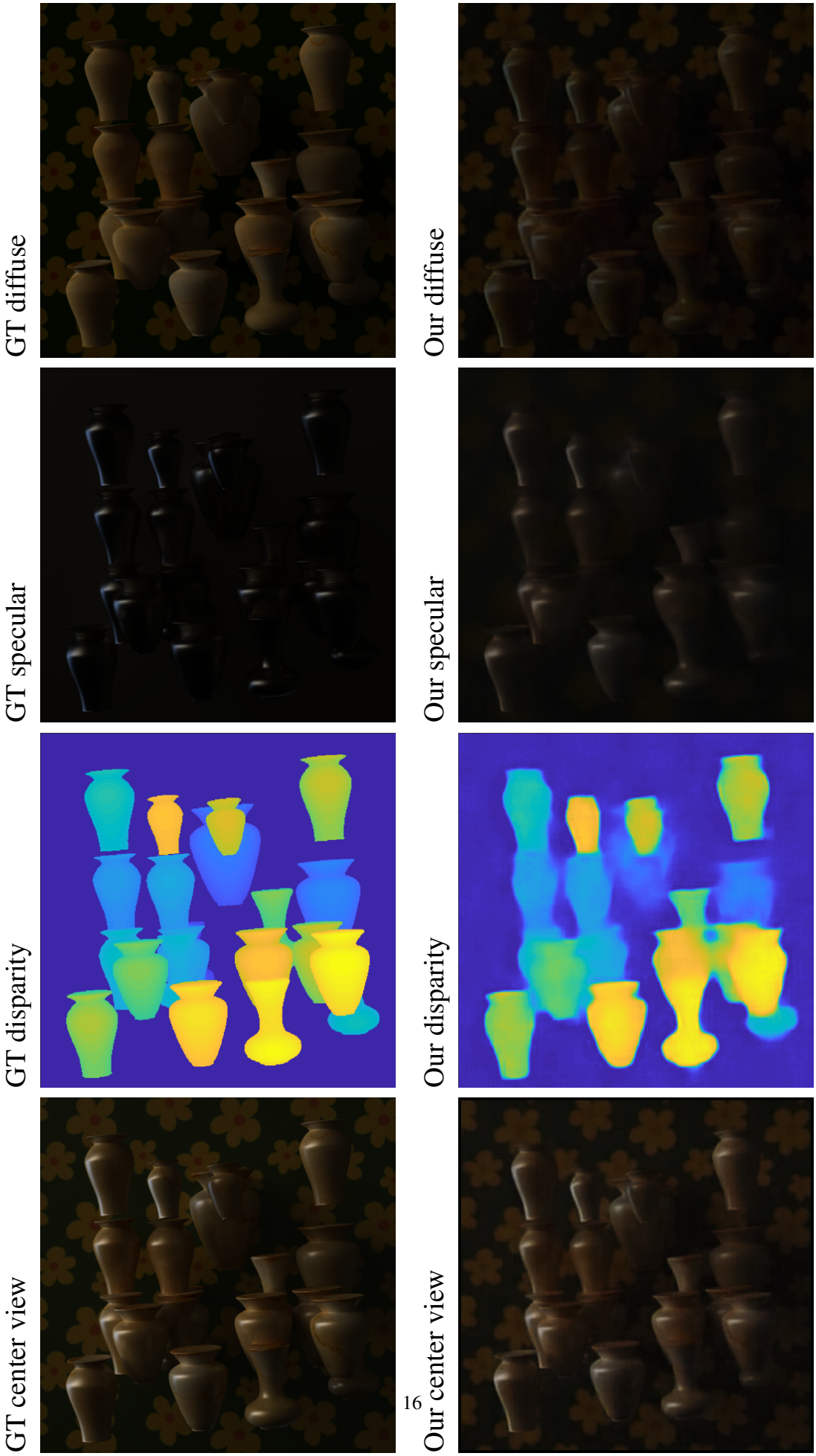
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-2.04 \;\; 1.99]$

Scene description: This data set has many glossy objects with similar shapes, which is challenging for the disparity estimation. This is an unseen light field that is used for testing.

Light field preview


GT diffuse
GT specular
GT disparity
GT center view
Our diffuse
Our specular
Our disparity
Our center view

- **Diffuse**

  LMSE = 0.05

  GMSE = 0.06

  SSIM = 82.02

- **Specular**

  LMSE = 0.04

  GMSE = 0.05

  SSIM = 75.64

- **Disparity**

  MSE = 11.26

Light field *Fourth test example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

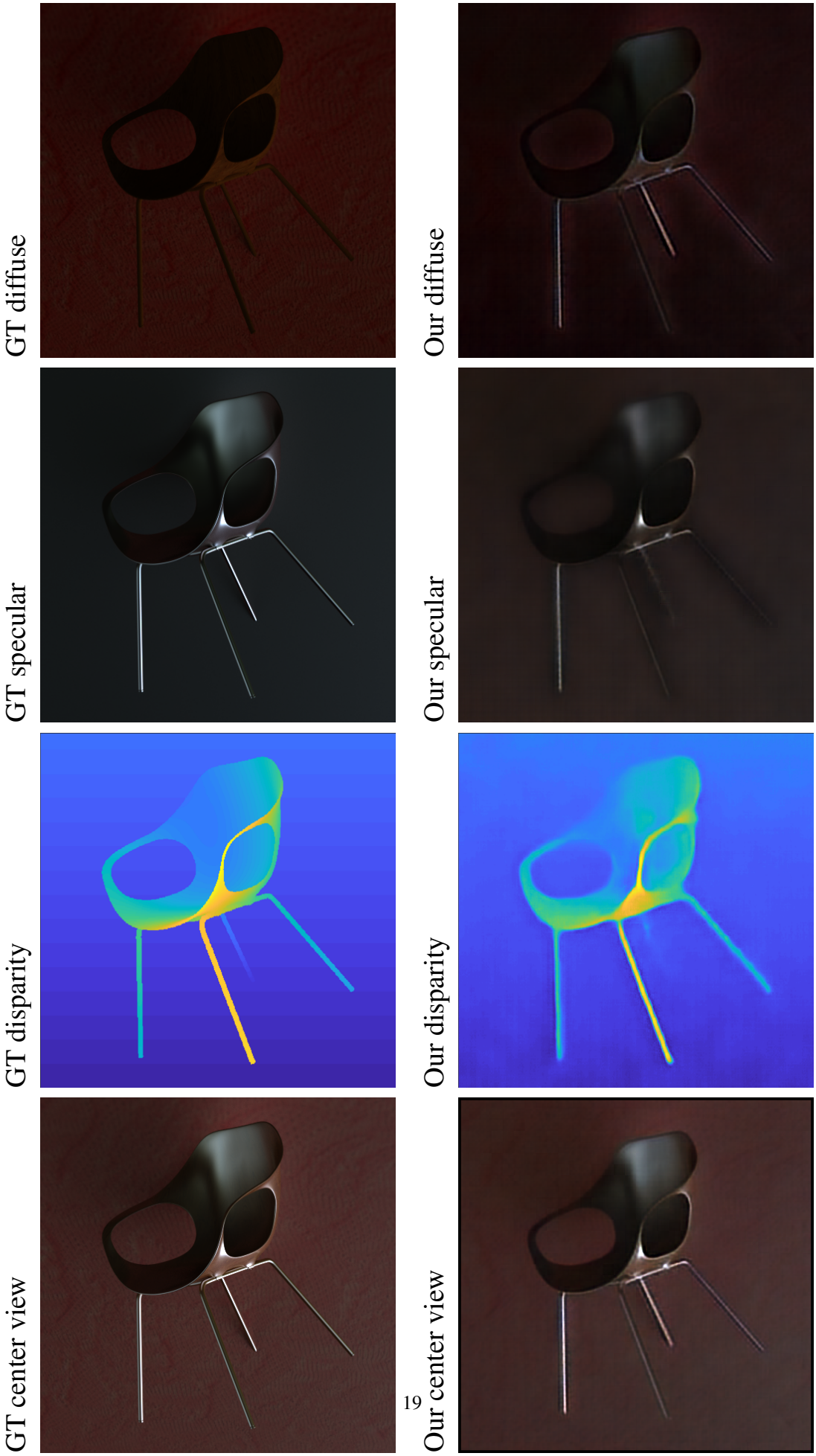
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-2.04 \;\; 1.95]$

Scene description: This data set shows an object with fine details and curved surfaces. This is an unseen light field that is used for testing.

Light field preview


GT diffuse
GT specular
GT disparity
GT center view
Our diffuse
Our specular
Our disparity
Our center view

- **Diffuse**

  LMSE = 0.19

  GMSE = 0.29

  SSIM = 68.57

- **Specular**

  LMSE = 0.3

  GMSE = 0.37

  SSIM = 51.52

- **Disparity**

  MSE = 5.82

Light field *Fifth test example*

Type: Synthetic data set generated with Blender using Cycles rendering engine

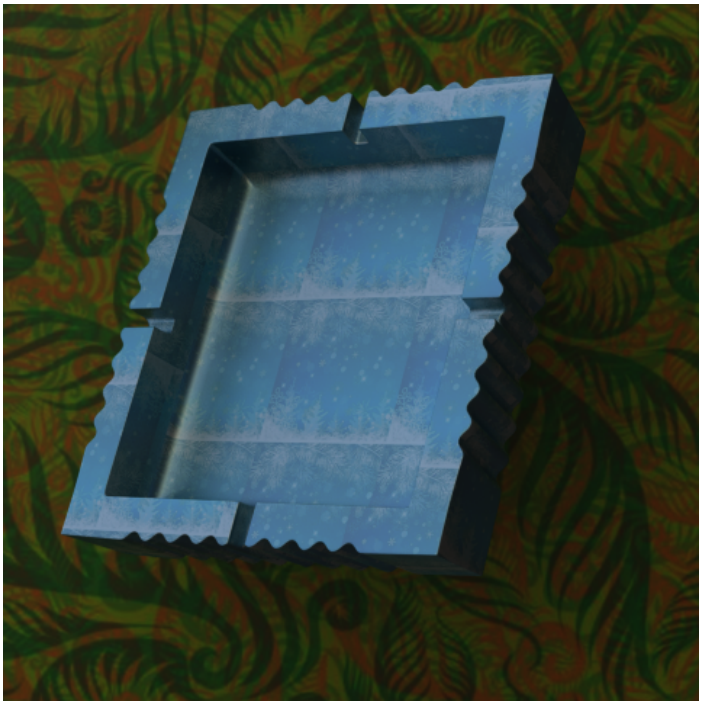
Size: 9×9 views

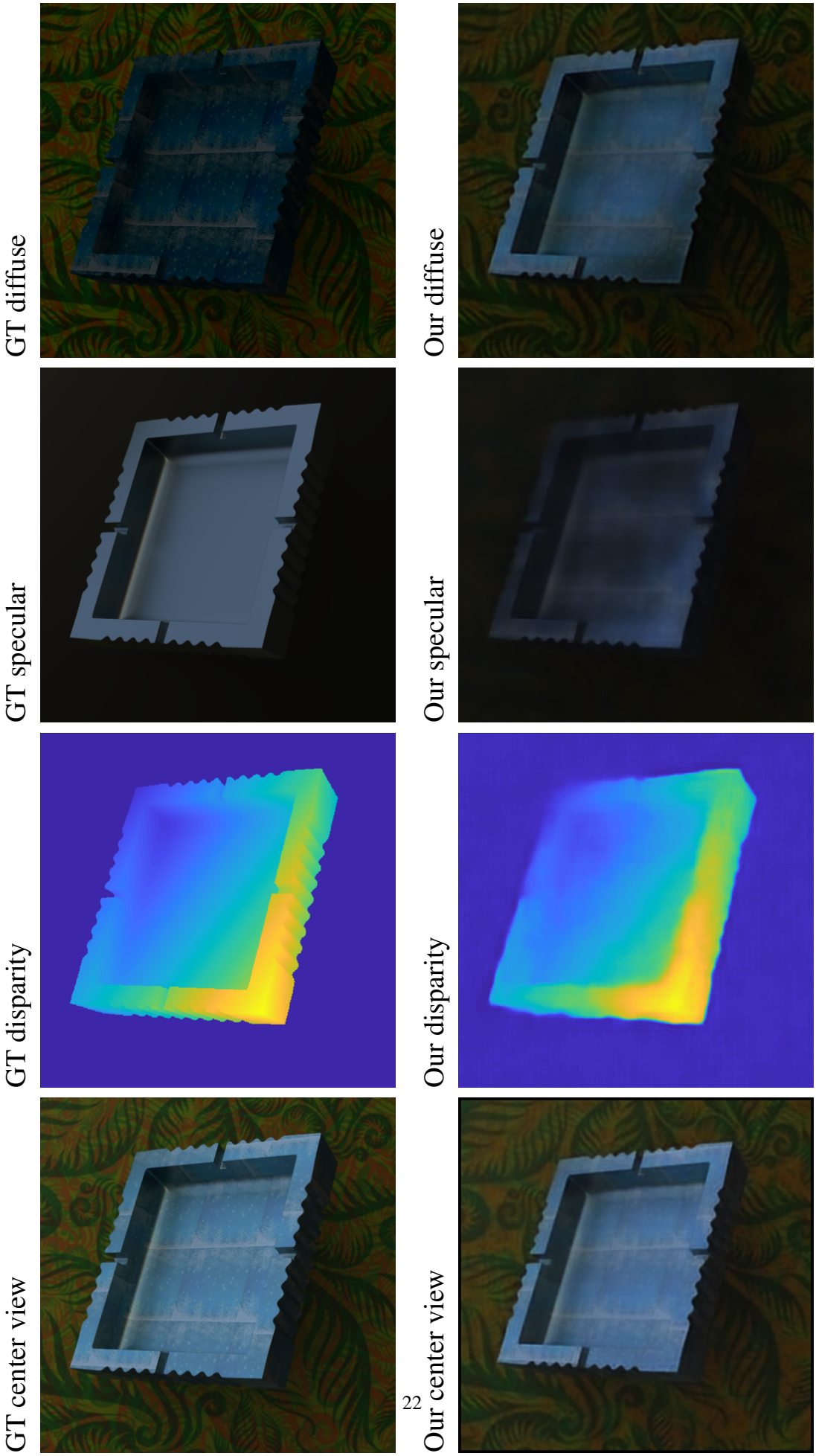
Resolution: 512×512 pixels

Disparity range: $[-2.04 \;\; 1.84]$

Scene description: This data set contains an object with almost diffuse reflection and very soft specularity. This is an unseen light field that is used for testing.

Light field preview




GT diffuse
GT specular
GT disparity
GT center view
Our diffuse
Our specular
Our disparity
Our center view

- **Diffuse**

  LMSE = 0.11

  GMSE = 0.25

  SSIM = 77.28

- **Specular**

  LMSE = 0.18

  GMSE = 0.37

  SSIM = 84.5

- **Disparity**

  MSE = 3.54

Light field *Flowers*

Type: Real world light field obtained with Lytro Illum plenoptic camera

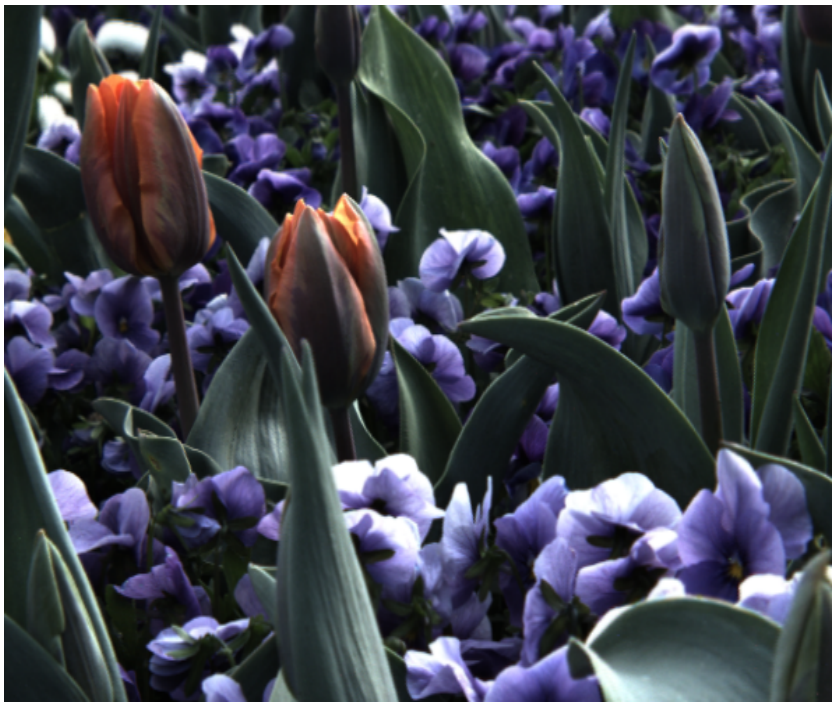
Size: 9×9 views

Resolution: 434×625 pixels

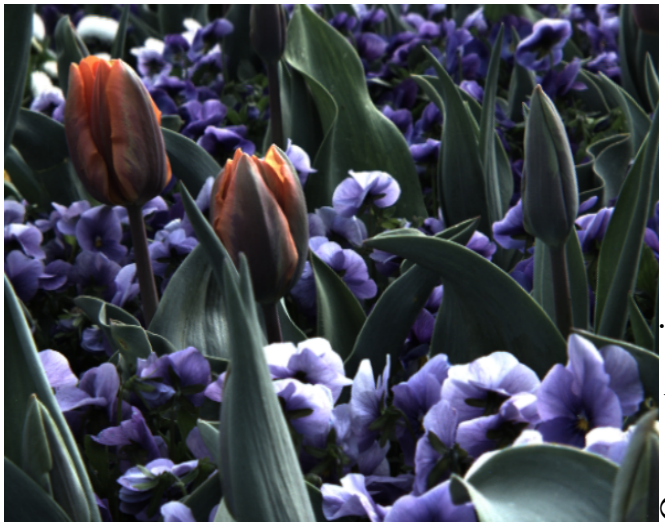
Disparity range: $[-1.43 \;\; 1.42]$

Scene description: This is a real-world example captured in sunlight. This dataset is used in unsupervised training of the autoencoder.

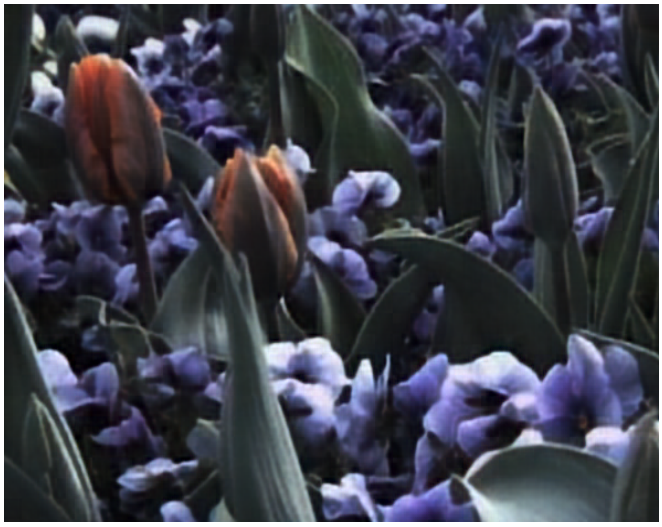
Light field preview



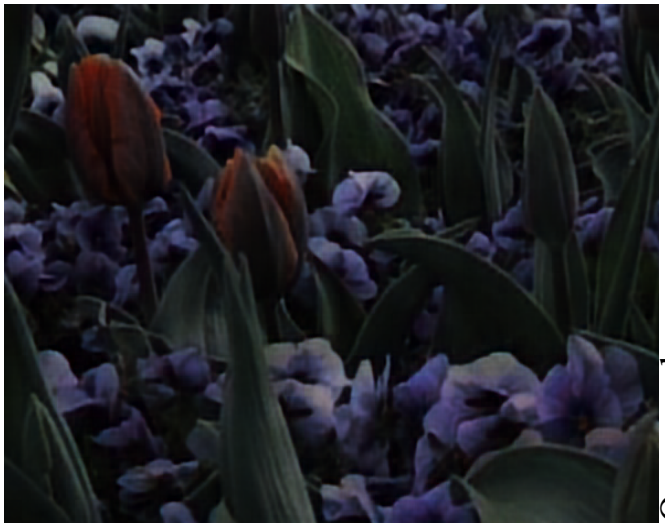
GT center view



Our center view

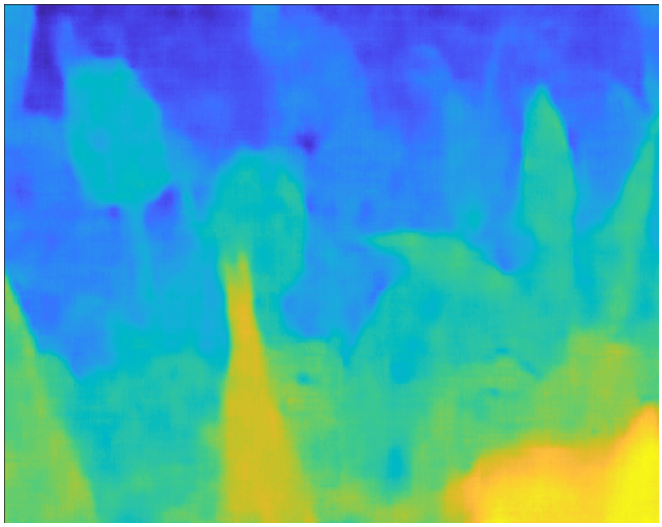


Our diffuse



Our specular

Our disparity



Light field *Koala with saxophone*

Type: Real world light field obtained with Lytro Illum plenoptic camera

Size: 9×9 views

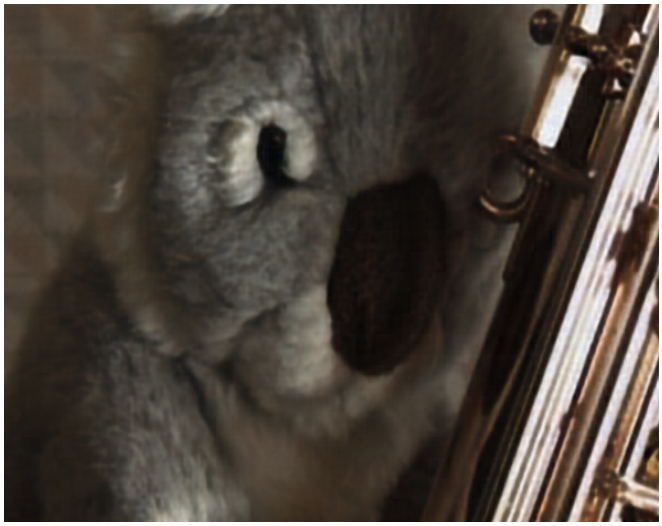
Resolution: 434×625 pixels

Disparity range: $[-1.52 \;\; 0.93]$

Scene description: This is a real-world example captured in approximately white light. It contains an almost Lambertian koala toy and a highly specular saxophone. The data set is also used during unsupervised training.

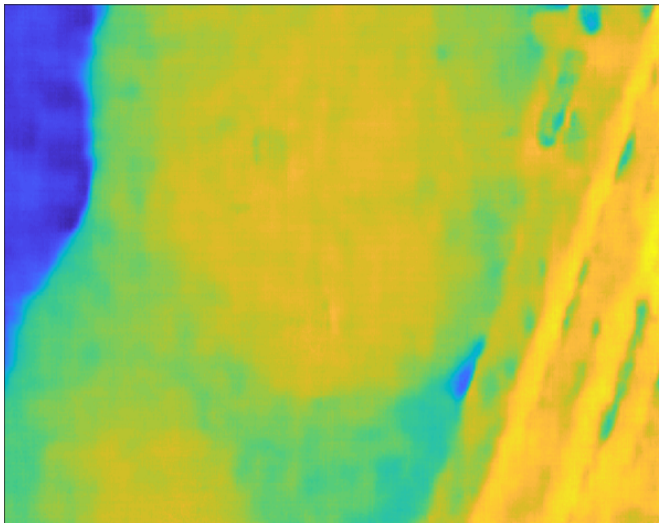
Light field preview


GT center view


Our center view


Our diffuse


Our specular


Our disparity

Light field *Origami*

Type: Real world light field obtained with Lytro Illum plenoptic camera

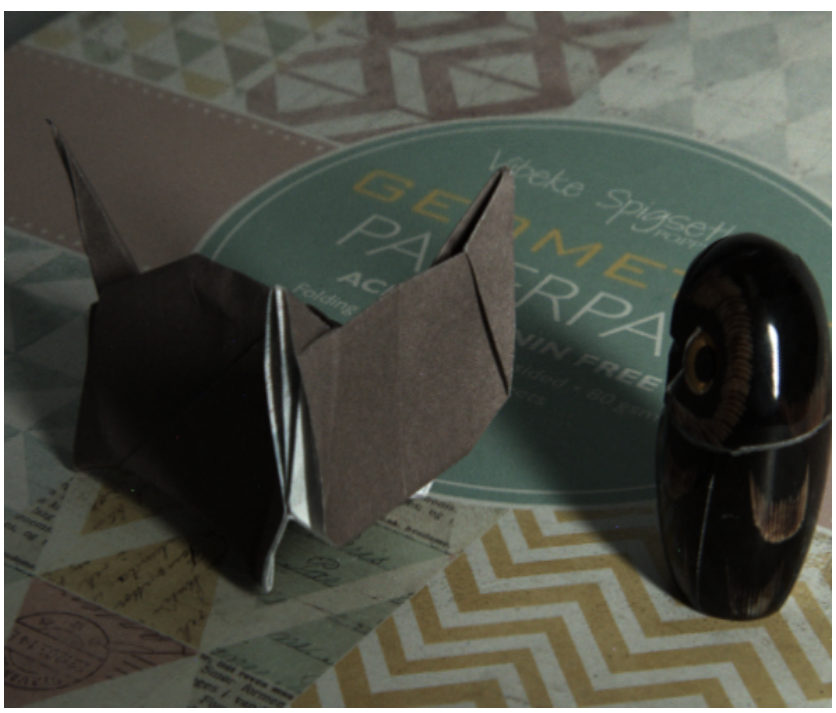
Size: 9×9 views

Resolution: 434×625 pixels

Disparity range: $[-1.76 \;\; 0.67]$

Scene description: This is a real-world example that is also used in the main paper. The light field shows two objects with different material, one is made of paper and another one is from wood with glossy coating.

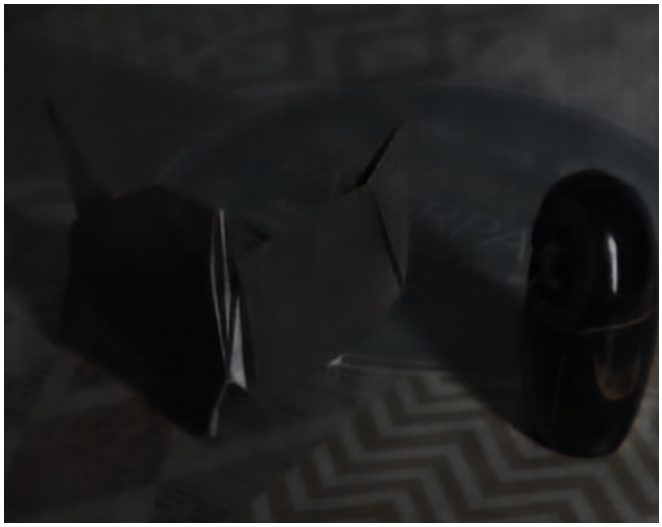
Light field preview



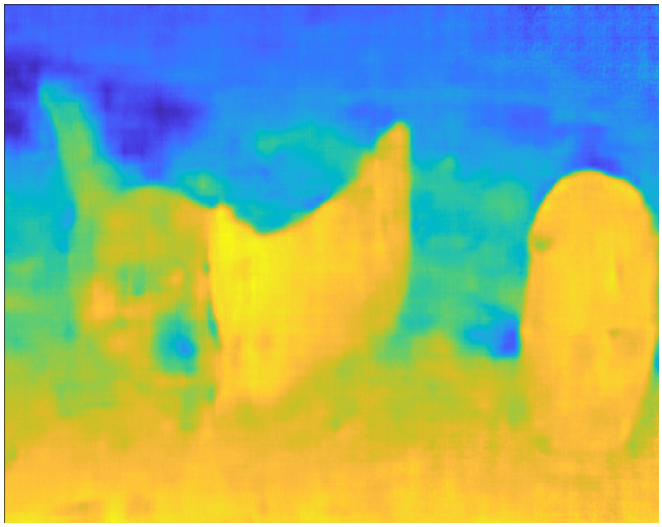
GT center view

Our center view

Our diffuse



Our specular



Our disparity

Light field *Hedgehog*

Type: Real world light field obtained with Lytro Illum plenoptic camera

Size: 9×9 views

Resolution: 434×625 pixels

Disparity range: $[-1.82 \;\; 0.84]$

Scene description: Another example with a highly specular object that is used for unsupervised training of the network.

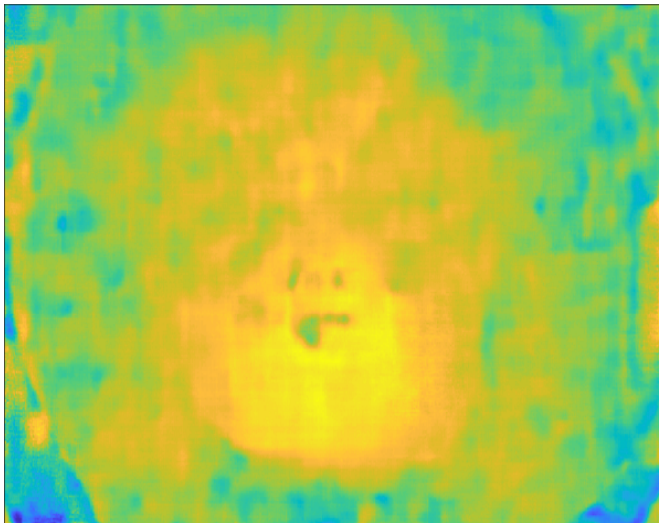
Light field preview

GT center view

Our center view

Our diffuse

Our specular



Our disparity

Light field *Saxophone and Koala*

Type: Real world light field obtained with Lytro Illum plenoptic camera

Size: 9×9 views

Resolution: 434×625 pixels

Disparity range: $[-2.25 \;\; 0.63]$

Scene description: This data set is used in the main paper, here we show larger images compared to the main manuscript. This data set was never presented to the network and is used for testing.

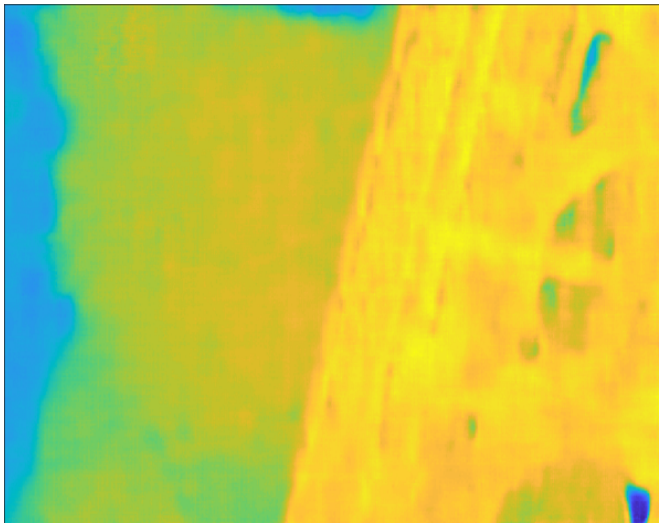
Light field preview

GT center view

Our center view

Our diffuse

Our specular



Our disparity

Light field *Maria*

Type: Real world light field from HCI benchmark [3]

Size: 9×9 views

Resolution: 926×926 pixels

Disparity range: $[-1.04 \;\; 0.37]$

Scene description: A wooden statue with textured background. For more details about this data set we refer to the benchmark [3]

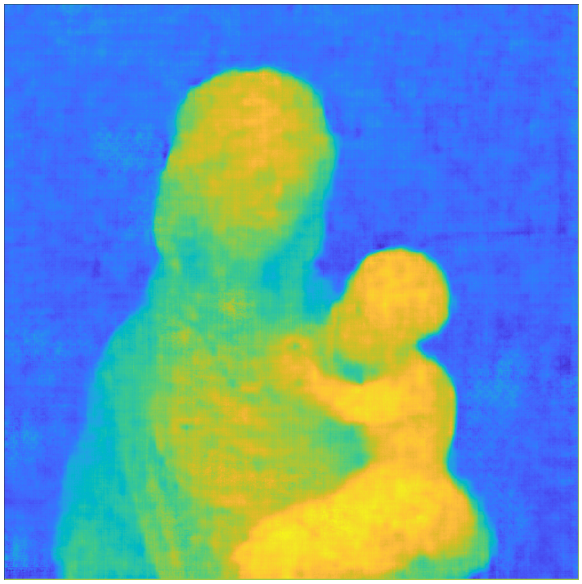
Light field preview

GT center view

Our center view

Our diffuse

Our specular



Our disparity

Light field *Cube*

Type: A real-world light field from the HCI benchmark [3]

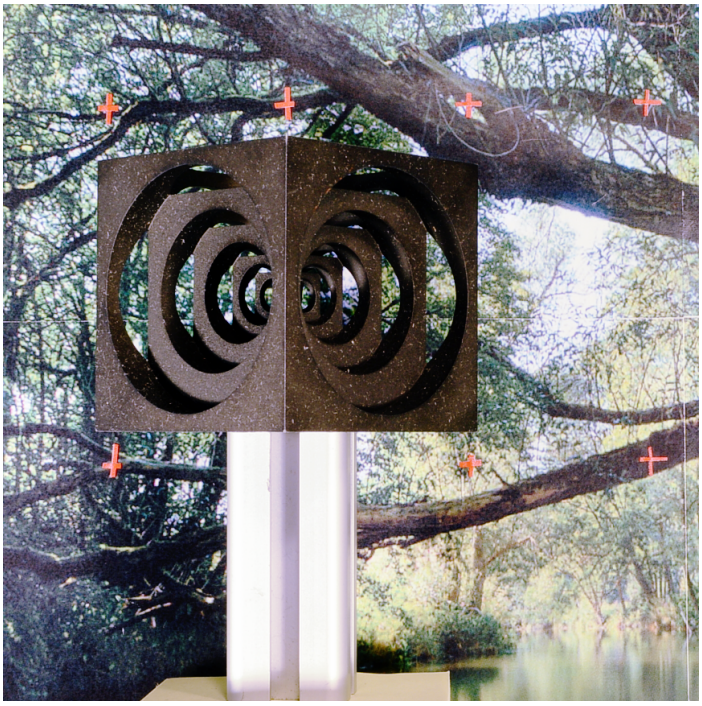
Size: 9×9 views

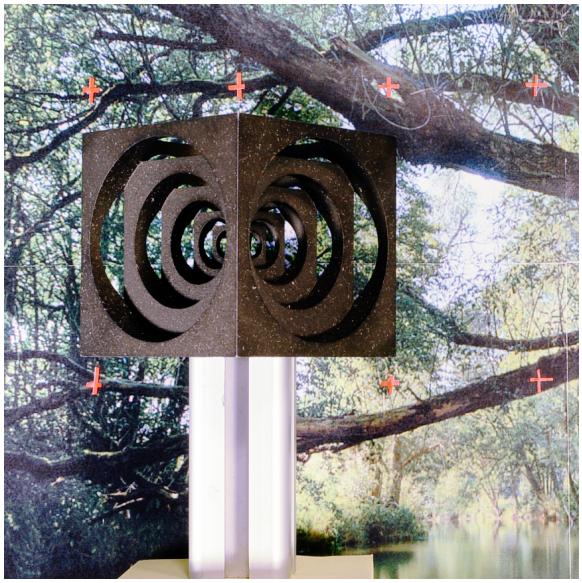
Resolution: 898×898 pixels

Disparity range: $[-1.59 \;\; 0.93]$

Scene description: Data set from HCI benchmark [3]that is used in the main paper. Here we present larger images and reconstructed center view of the light field.

Light field preview



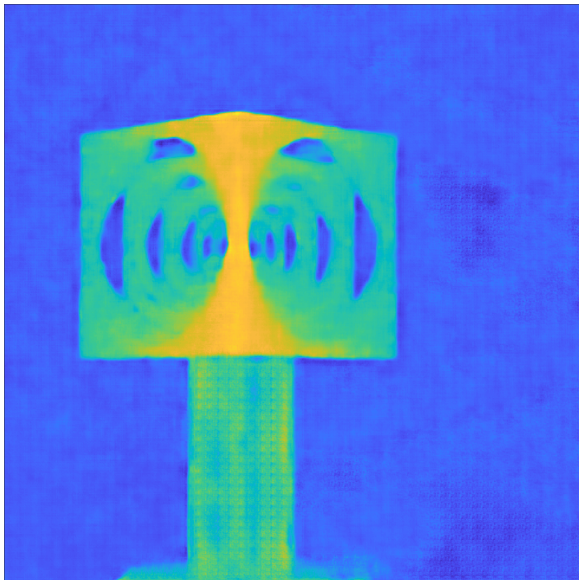


GT center view

Our center view

Our diffuse

Our specular



Our disparity

Light field *Bicycle benchmark*

Type: Synthetic test scene from light field benchmark [1]

Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-1.7 \;\; 1.7]$

Scene description: A scene with complicated geometries and shapes which is difficult for both reflection separation and disparity estimation.

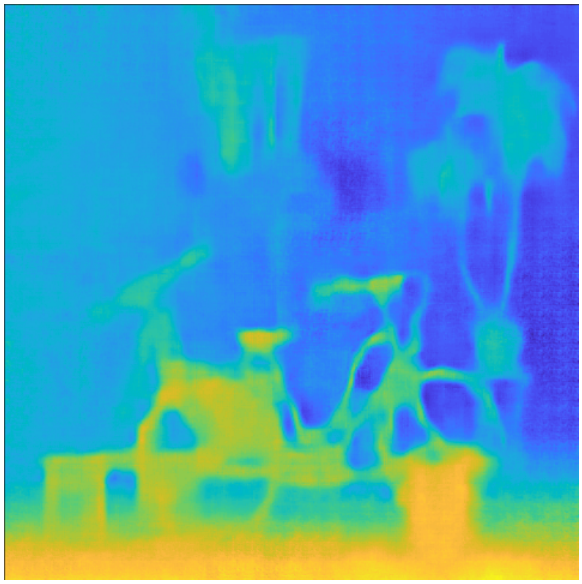
Light field preview

GT center view

Our center view

Our diffuse

Our specular



Our disparity

Light field *Cotton benchmark*

Type: Synthetic test scene from light field benchmark [1]

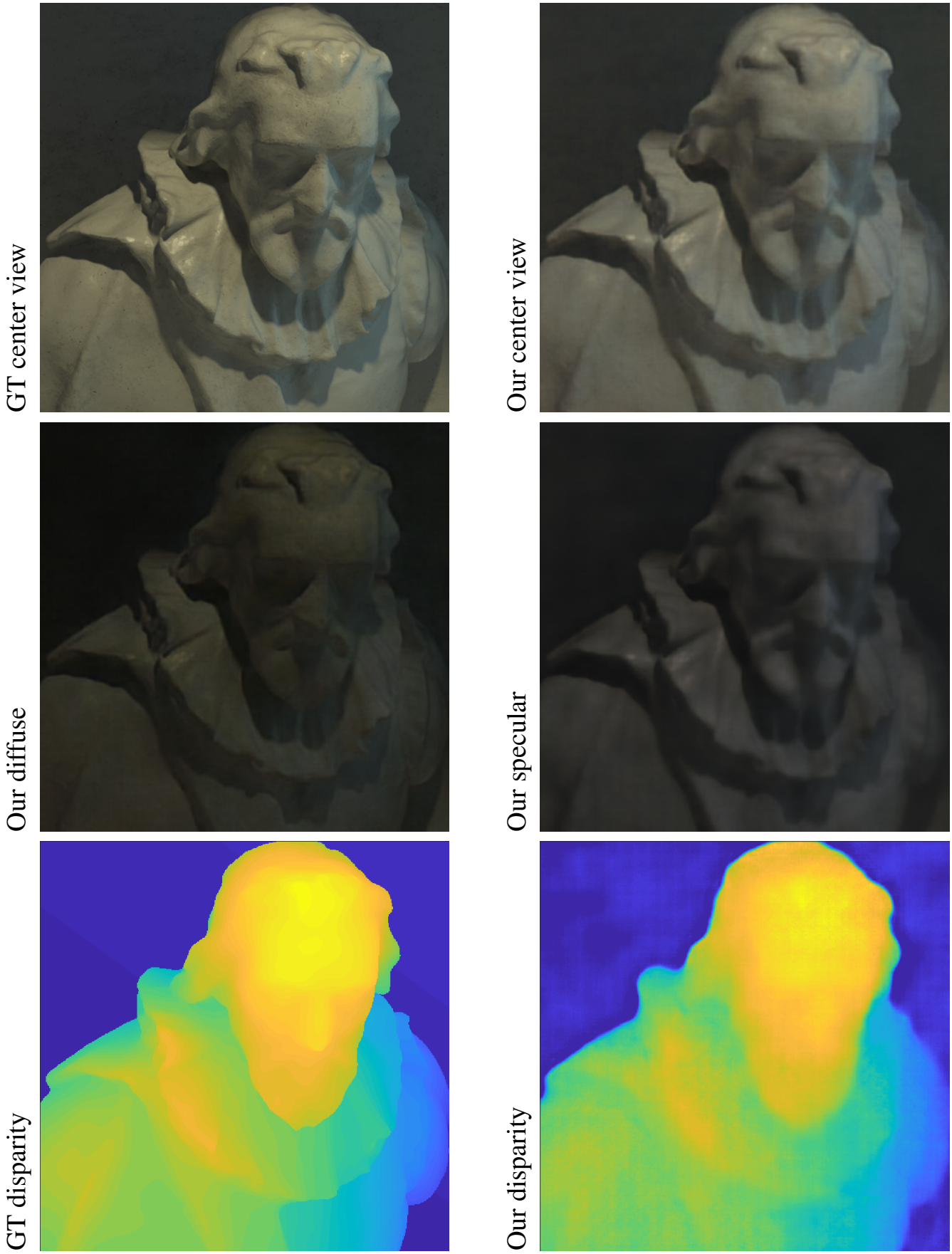
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-1.46 \;\; 1.32]$

Scene description: A synthetic scene with a single object made of concrete.

Light field preview


GT center view
Our center view
Our diffuse
Our specular
GT disparity
Our disparity

MSE = 1.38

Light field *Antinous benchmark*

Type: Synthetic test scene from light field benchmark [1]

Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-2.83 \;\; 2.58]$

Scene description: An additional scene from the light field benchmark, similar to the previous data set, but with different illumination.

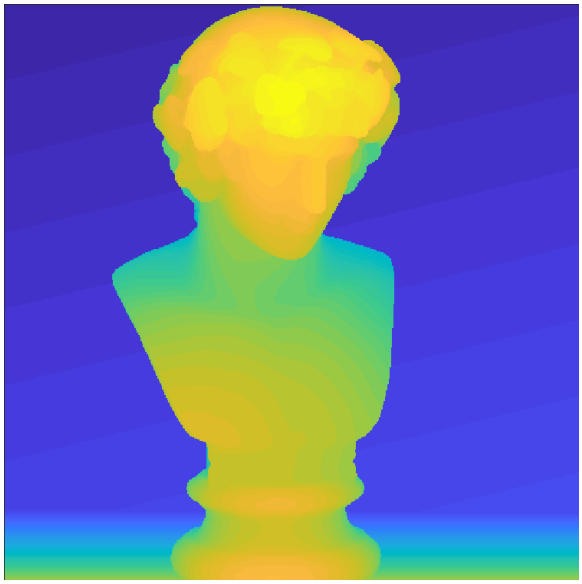
Light field preview

GT center view

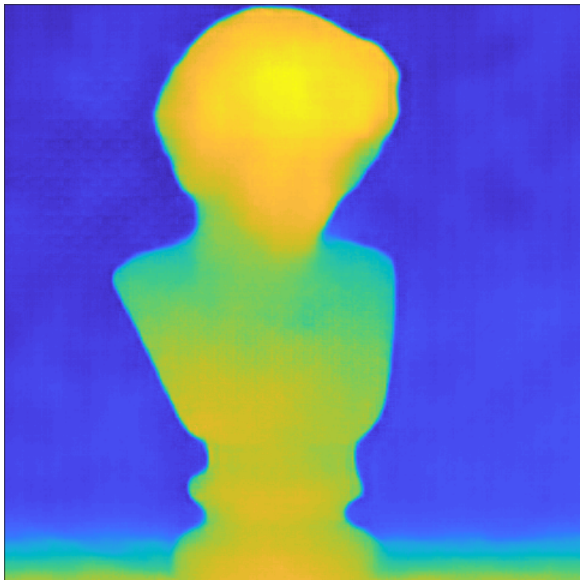
Our center view

Our diffuse

Our specular



GT disparity



Our disparity

MSE = 9.25

Light field *Vinyl benchmark*

Type: Synthetic test scene from light field benchmark [1]

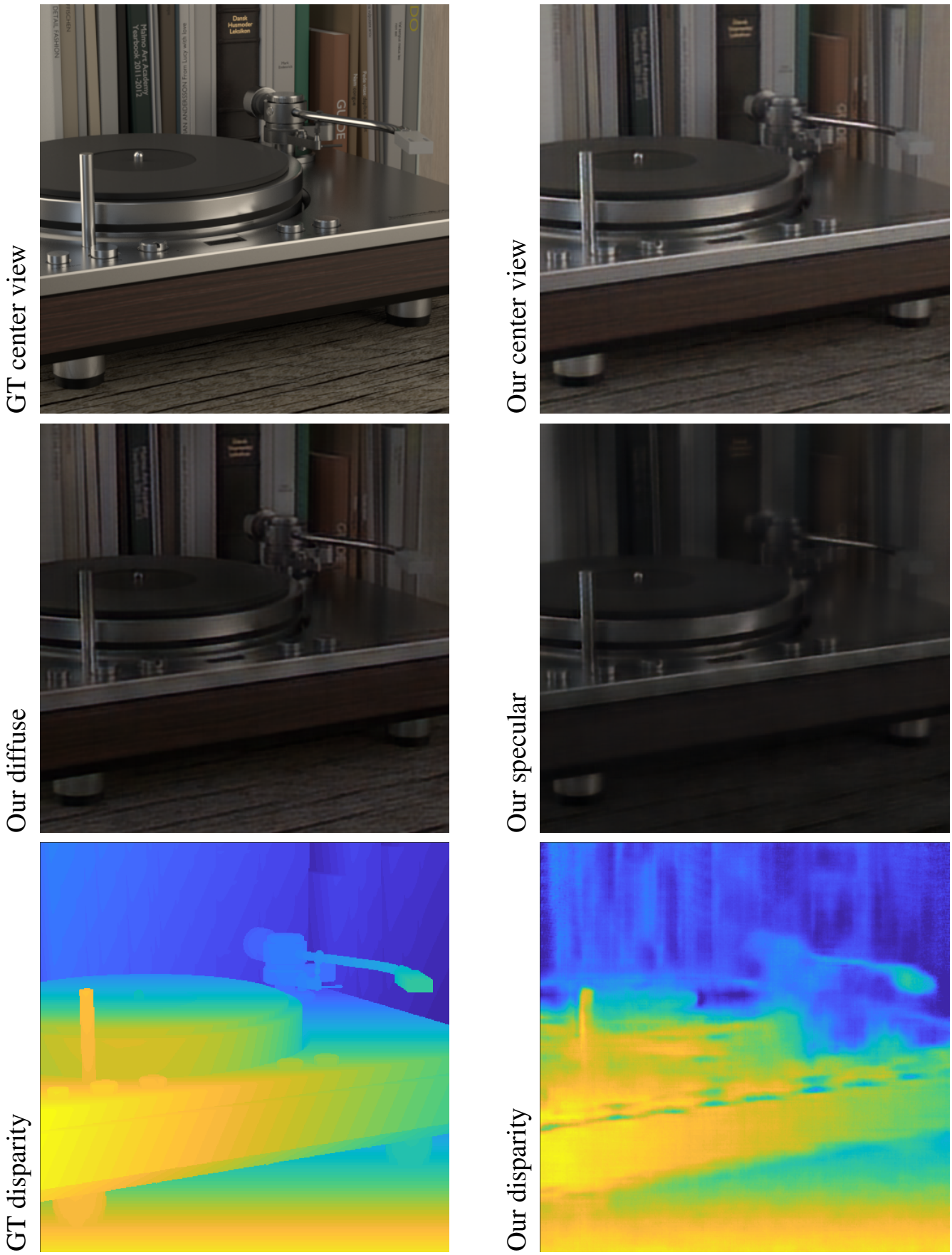
Size: 9×9 views

Resolution: 512×512 pixels

Disparity range: $[-1.49 \;\; 0.94]$

Scene description: The foreground object has metal parts that are specular. This is a synthetic data set from the additional scenes presented in the benchmark [1].

Light field preview


GT center view
Our center view
Our diffuse
Our specular
GT disparity
Our disparity

MSE = 4.4

Light field *Amethyst*

Type: Real world Stanford data set [2]

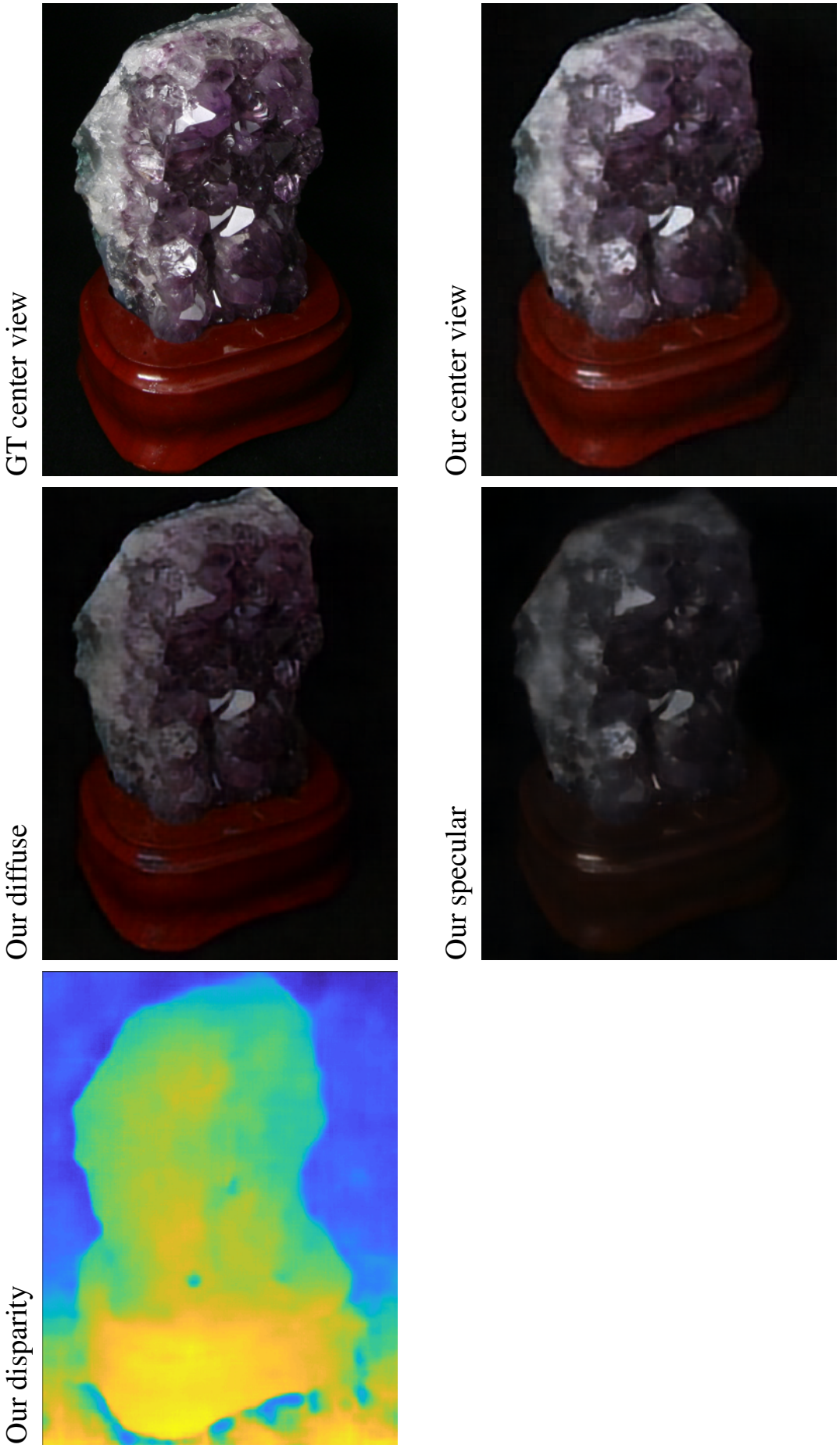
Size: 9×9 views

Resolution: 512×384 pixels

Disparity range: $[-2.46 \;\; 1.4]$

Scene description: A real-world scene with lots of specularity and complex geometry.

Light field preview


GT center view
Our center view
Our diffuse
Our specular
Our disparity

Light field *LegoTruck*

Type: Real world Stanford data set [2]

Size: 9×9 views

Resolution: 480×640 pixels

Disparity range: $[-2.03 \;\; 1.18]$

Scene description: A real-world scene with object made from Lego.

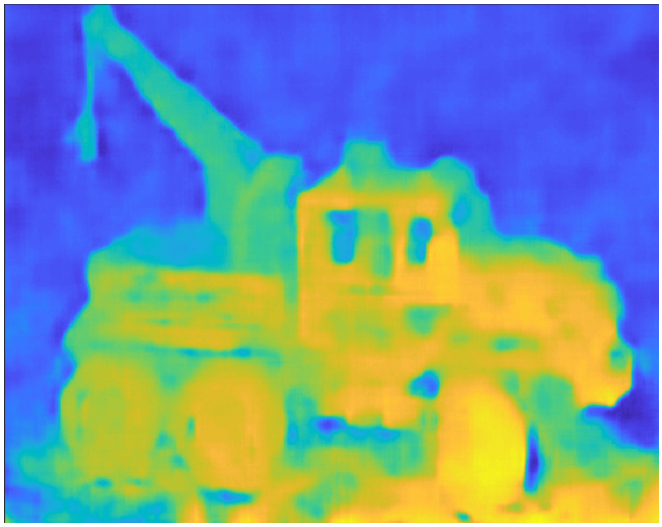
Light field preview

GT center view

Our center view

Our diffuse

Our specular



Our disparity

Video from data set: *Amethyst*

- Input light filed
- **Our results**
- Diffuse
- Specular
- Reconstructed

Video from data set: *Flowers*

- Input light filed
- **Our results**
- Diffuse
- Specular
- Reconstructed

Video from data set: *First test example*

- Input light filed
- **Our results**
- Diffuse
- Specular
- Reconstructed

Video from data set: *Third test example*

- Input light filed
- **Our results**
- Diffuse
- Specular
- Reconstructed

Video from data set: *Hedgehog*

- Input light filed

- **Our results**
- Diffuse
- Specular
- Reconstructed