

# A NEW GROUND-TRUTH DATASET FOR INTRINSIC IMAGE RESEARCH

Nestmeyer T., Kiefel M., Gehler P.

{tnestmeyer, mkiefel, pgehler}@tue.mpg.de

Max Planck Institute for Intelligent Systems, Tübingen, Germany



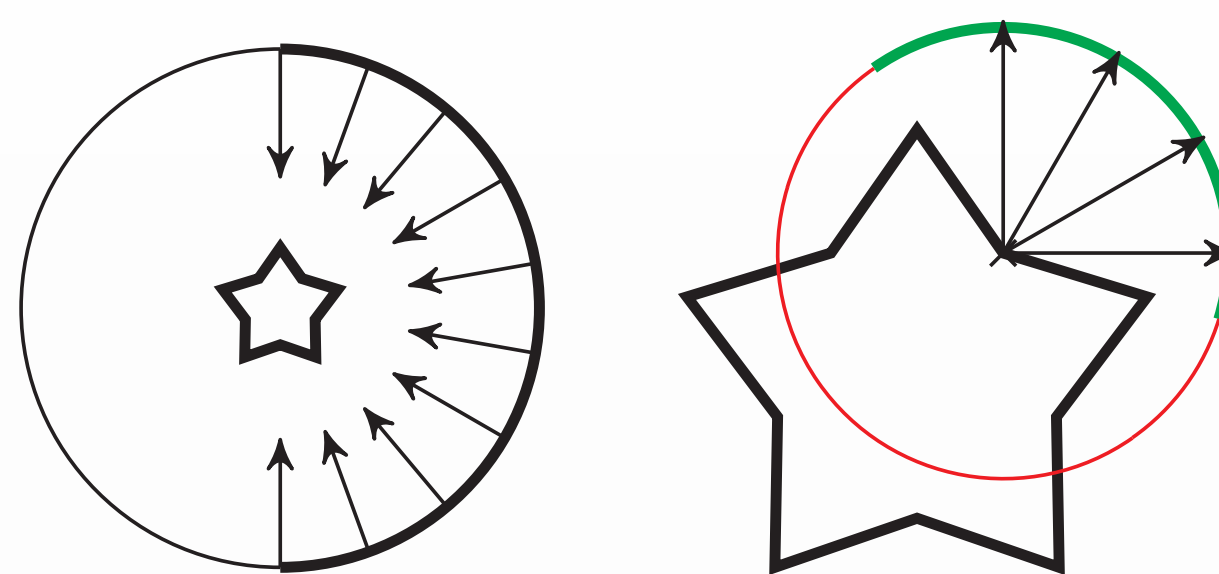
MAX-PLANCK-GESELLSCHAFT

## Abstract

When taking a picture, the image is formed by an interplay of light with the reflectance properties of the objects in a scene. Recovering these physical properties from a single image is underconstrained. Nevertheless, we believe that this intermediate representation simplifies other tasks, e.g. detection or scene understanding. Since current state-of-the-art ground-truth datasets are rather limited, we would like to create a new dataset to further enable intrinsic image research.

## Formation of an image

What we actually see is a combination of the reflectance properties of an object and the intensity and color of the light source.



Assumptions:

- We know the color of the light.
- Light source is infinitely far away.
- There are no interreflectances.

Light can be described by a function defined on the surface of a sphere. Spherical harmonics are suited for this purpose.

## Goals

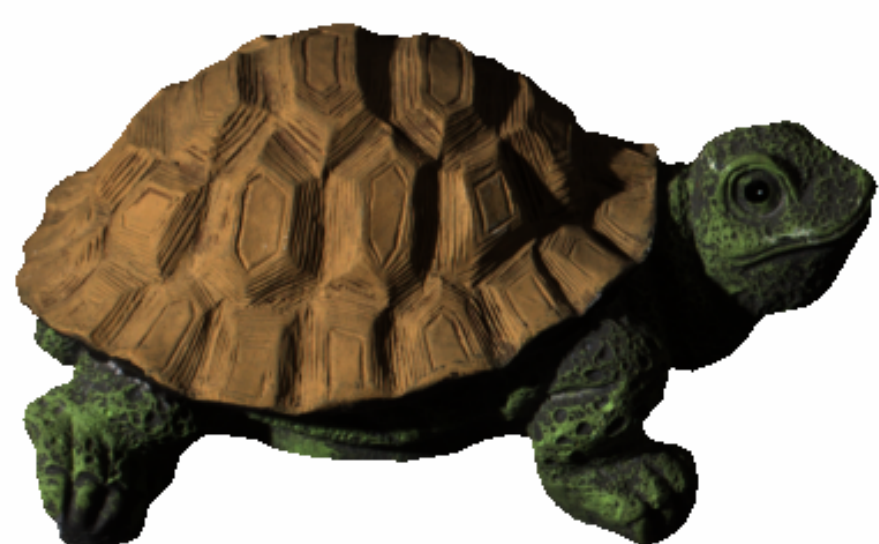
We want to create a new ground-truth dataset that has

- a big enough amount of images to practically do learning
- good resolution (current technical limitation is 4 MP)
- also provide video with ground-truth reflectance and shading (during 3D scanning of objects)
- objects with color gradients (that does not come from shading with colored light)

We are certain that being able to separate the reflectance of objects in images or videos can lead to performance increases in a broad range of other computer vision tasks, when using this information as input.

## Intrinsic Images

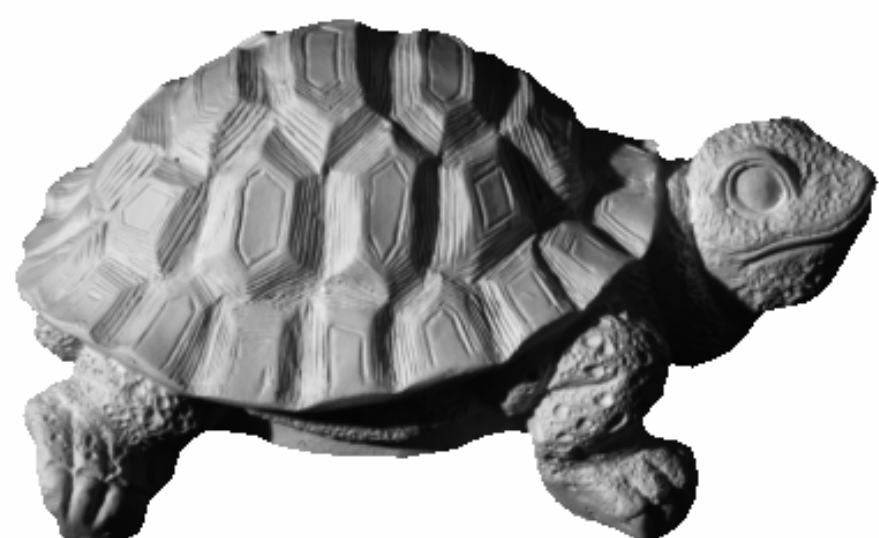
Decompose a single image into its reflectance and shading components:



input image



reflectance



shading

## MIT intrinsic

The current state-of-the-art ground-truth dataset is the MIT intrinsic dataset [1]. It has a lot of limitations:

- only 20 images
- small resolution (about 0.2 MP)
- only static images (no video)
- objects are not everyday things
- only one perspective each (although different light settings)
- no geometry information for shape ground-truth

## Our Method

We want to use a lightstage that enables us to get 200 images per object. Additionally, we would like to record and reconstruct the geometry of a scene with an Artec 3D scanner<sup>a</sup>.



Artec Eva



Example scan

In combination we hope to be able to extract a ground-truth dataset that gives researchers the opportunity to test their methods on a broader scale than currently possible.

<sup>a</sup><http://www.artec3d.com/>

## References

- [1] Roger Grosse, Micah Johnson, Edward H. Adelson and William T. Freeman. "Ground-truth dataset and baseline evaluations for intrinsic image algorithms" in ICCV 2009.