Ground-truth dataset and evaluation for intrinsic images

Roger Grosse, Micah Johnson, Edward Adelson, and William Freeman

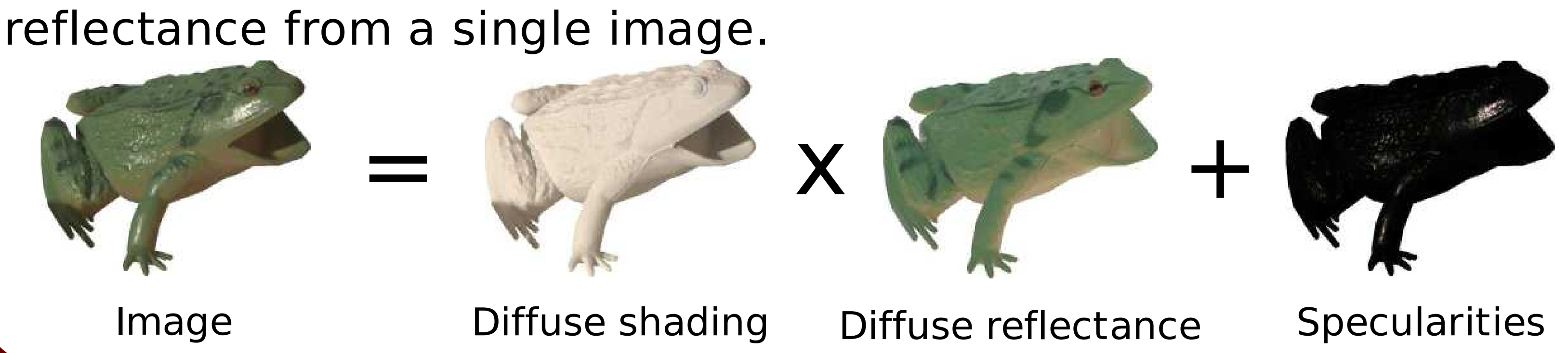
Massachusetts Institute of Technology

Contact: rgrosse@mit.edu

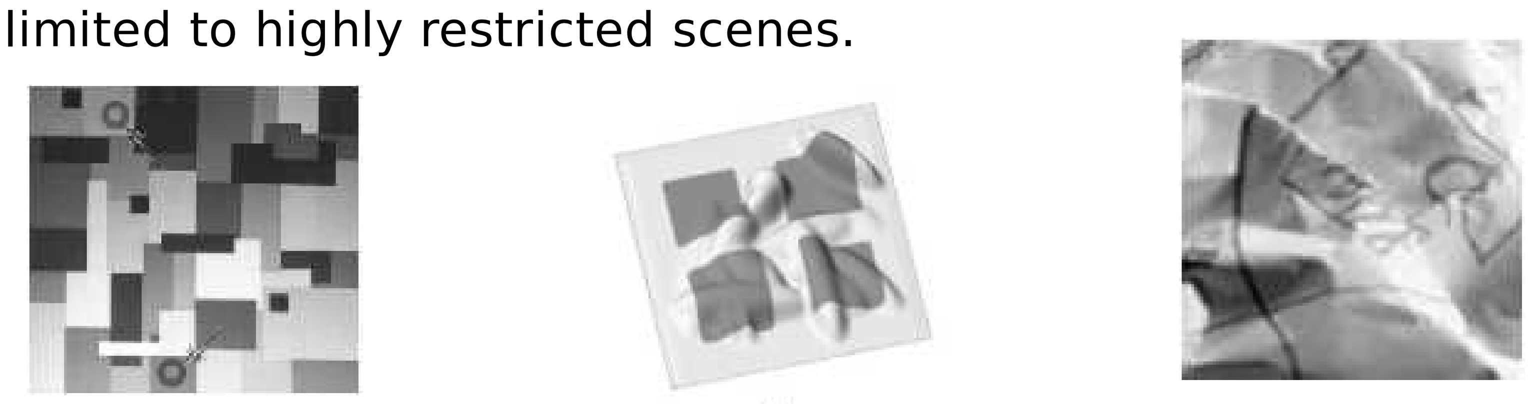
Motivation

Every pixel depends on illumination, geometry, and reflectance. Many vision applications can benefit from separating these factors.

Goal of intrinsic image decomposition: recover shading and reflectance from a single image.

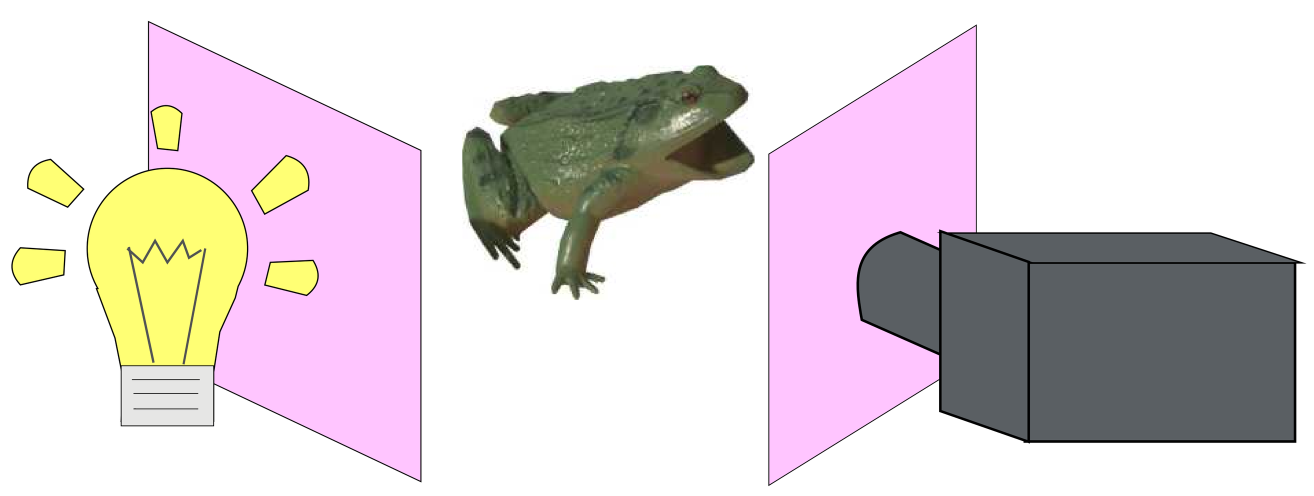


Many techniques have been proposed, but quantitative evaluation has been limited to highly restricted scenes.



We have created a ground-truth dataset with a variety of complex objects, and we used this to compare curent state-of-the art algorithms.

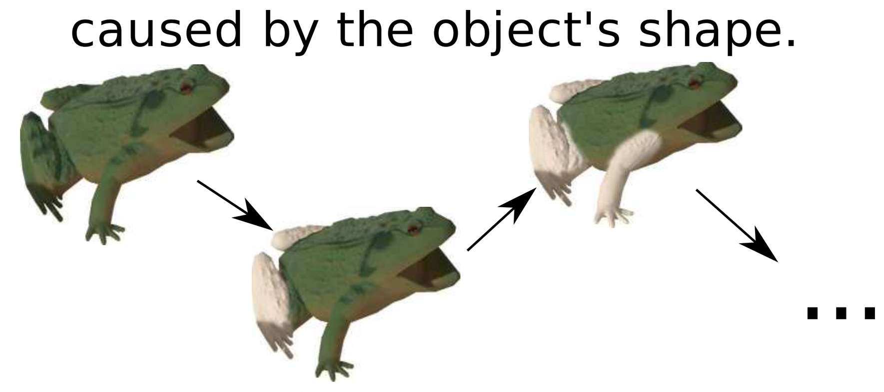
Procedure



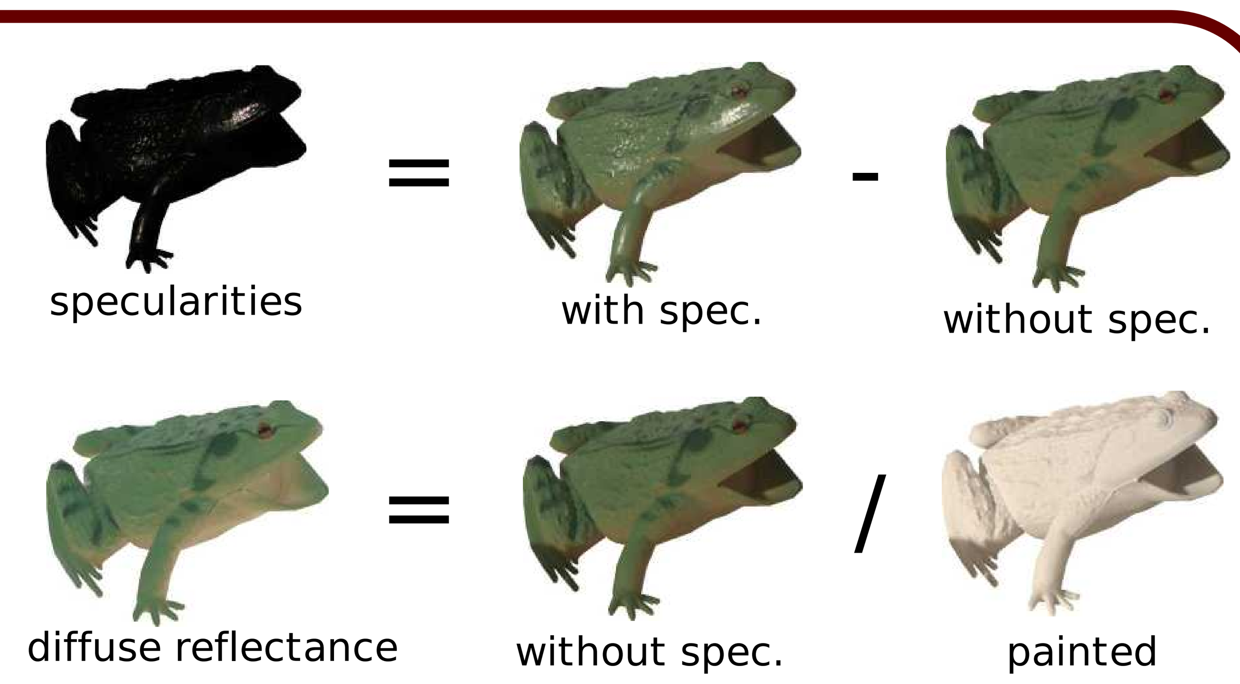
Specular reflections preserve light's polarity, while diffuse reflections do not.

We applied polarizing filters over the light source and camera. Rotating the filters toggles the specularities.

To retrieve the shading image, we painted the object a uniform shade of gray so that all edges were caused by the object's shape.

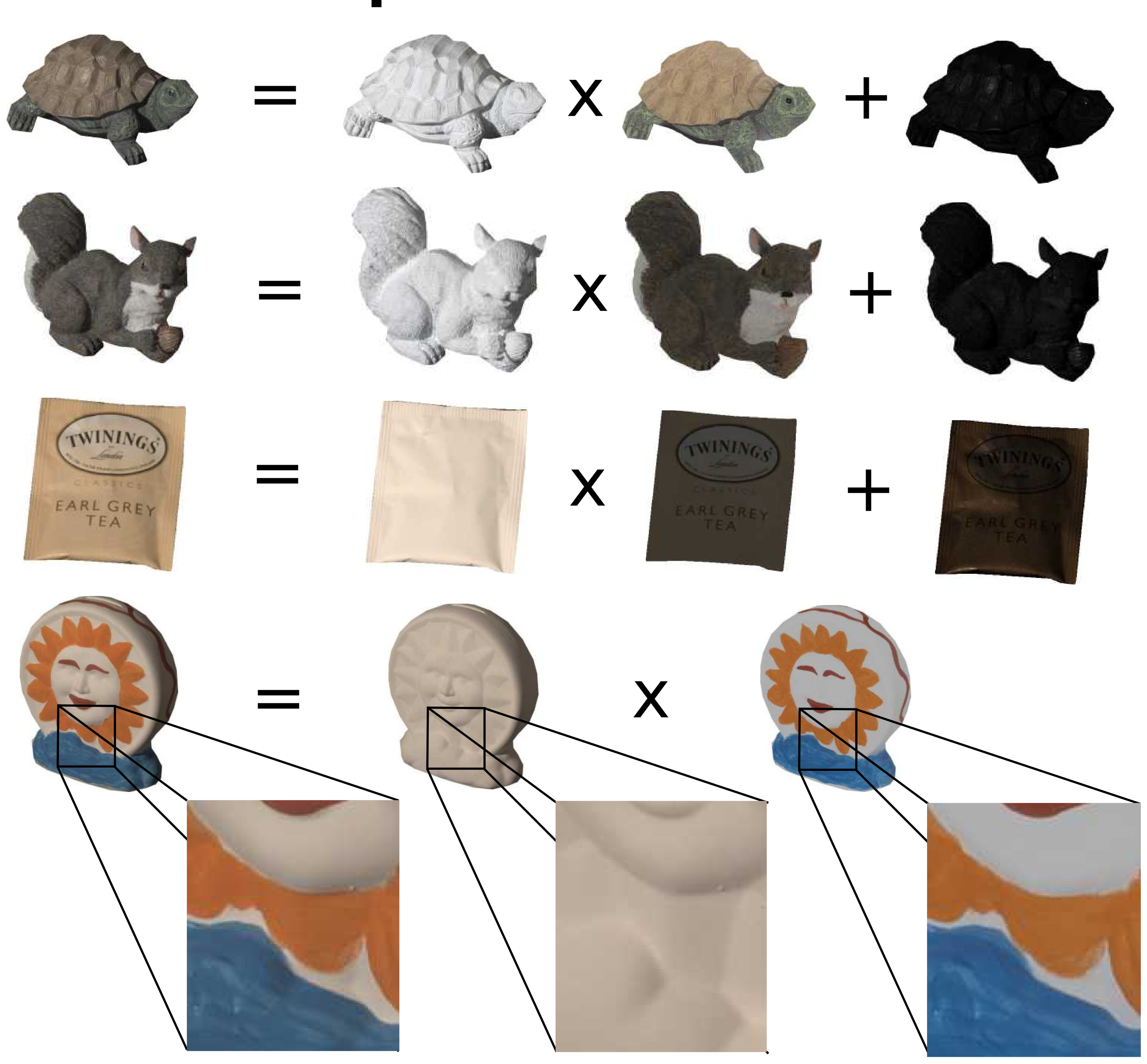


We quantified the interreflections by incrementally painting convex sections and found them to be extremely small.



Arithmetic on the photographs gives us the specularity and reflectance images.

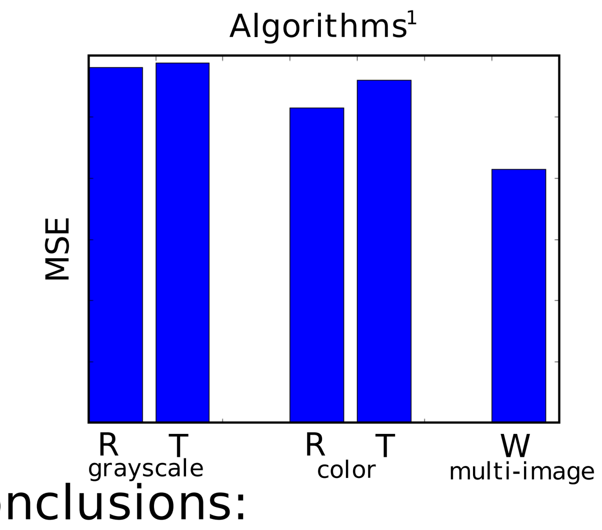
Examples

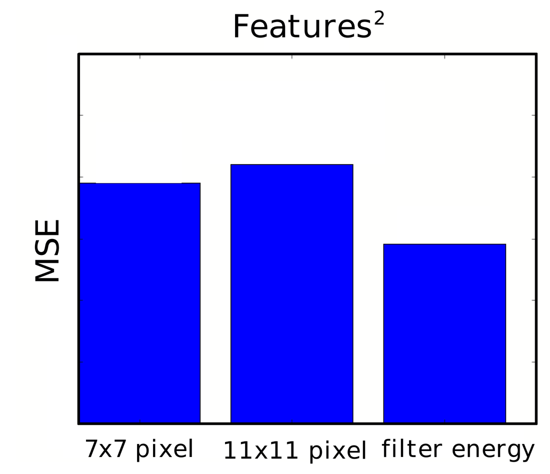


Experiments

We quantitatively evaluated several existing intrinsic image algorithms on our dataset.

R: Retinex, a simple heuristic which thresholds image gradients
T: The approach of Tappen et al. (2005), which trains a gradient classifier from computer generated images and shares information globally using belief propagation
W: Weiss (2001), using multiple images of the same object





Basic conclusions:

- Surprisingly, the simple Retinex algorithm does as well as or better than Tappen et al. (2005)'s more complex approach. This may be due to their specialized training set.
- The task is easier when algorithms may use color.
- Texture features based on local filter energy generalize better to new objects than features based on raw pixels.

References

M. F. Tappen, W. T. Freeman, and E. H. Adelson. Recovering intrinsic images from a single image. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(9): 1459-1472, 2005.

M. F. Tappen, E. H. Adelson, and W. T. Freeman. Estimating intrinsic component images using non-linear regression. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, volume 2, pages 1992-1999, 2006.

Y. Weiss. Deriving intrinsic images from image sequences. In Proceedings of the International Conference on Computer Vision, volume 2, pages 68-75, 2001.

[1] MSE of the steerable pyramid representation of the algorithms' outputs, log-pixel intensities, sigma = 2 pixels. Used 3-fold cross-validation such that images most similar to a given test image were not included in the training data. Input to Weiss's algorithm included specularities because we did not apply polarizing filters in the moving-lights condition.

[2] Comparison used an exemplar-based learning algorithm similar to Tappen et al. (2006).