Pacific Graphics (2016)
E. Grinspun, B. Bickel, and Y. Dobashi (Editors)

Short Paper

Reflectance and Shape Estimation for Cartoon Shaded Objects

H. Todo[†1,2] and Y. Yamaguchi[1,2]

[1]The University of Tokyo, Japan
[2]JST CREST, Japan

Abstract

Although many photorealistic relighting methods provide a way to change the illumination of objects in a digital photograph, it is currently difficult to relight a cartoon shading style in digital illustrations. The main difference between photorealistic and cartoon shading styles is that cartoon shading is characterized by soft color quantization and nonlinear color variations that cause noticeable reconstruction errors under a physical reflectance assumption such as Lambertian. To handle this non-photorealistic shading property, we focus on the shading analysis of the most fundamental cartoon shading technique. Based on its color map shading representation, we propose a simple method to decompose the input shading to a smooth shape with a nonlinear reflectance property. We have conducted simple ground-truth evaluations to compare our results to those obtained by other approaches.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

Despite recent progress in 3D computer graphics techniques, traditional cartoon shading styles remain popular for 2D digital art. Artists can use a variety of commercial software (e.g. Adobe® Photoshop, Corel® Painter) to design their own expressive shading styles. Although the design principle roughly follows a physical illumination model, its editing is restricted to 2D drawing operations. We are interested in exploring new interactions for relighting a painted shading style from a single input image.

Reconstructing a surface shape and reflectance from a single image is known as the Shape-From-Shading problem [HB89]. Based on the fundamental problem setting, most relighting approaches assume a Lambertian shading model for shading representation [KRFB06, OZM*06, WSTS08]. Although these approaches work well for photorealistic images, they often fail to decompose cartoon shading styles in digital illustrations.

The main difference between photorealistic and cartoon shading styles is that cartoon shading is characterized by nonlinear color variation with soft quantization. The final shading is typically more abstracted from the original surface and its illumination. This assumption is common in many 3D stylized rendering techniques using color map representation [LMHB00, SMGG01, BTM06, MFE07] that simply convert smooth 3D illumination to an artistic shading style. As shown in Figure 1, this simple mechanism

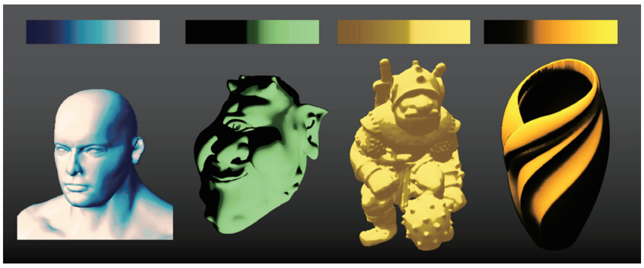

Figure 1: *Stylized shading styles obtained by color map representation.*

can produce a variety of shading styles with different quantization effects. However, such stylization processes make it more difficult for shading analysis to reconstruct a surface shape and reflectance from the designed shading.

In this paper, we propose a simple shading analysis method to decompose nonlinear shading to a reasonable shading representation. As a stating point, we focus on the most fundamental cartoon shading [LMHB00]. Our primary assumption is that the main nonlinear factor of the final shading can be encoded by a color map function. Therefore, we aim to reconstruct a smooth surface field and a nonlinear reflectance property from the input shading. Based on these estimated data, our method provides a way to relight the input shading. To evaluate our approach, we conducted a simple pilot study using a prepared set of 3D models and color maps with

† Present affiliation: Tokyo University of Technology, Japan

DOI: 10.2312/pg.20161331

a variety of stylization inputs. The proposed method was quantitatively compared to related approaches, which provided several key insights regarding relighting stylized shading.

2. Related Work

Previous shape reconstruction methods for painted illustrations also attempt to recover a smooth surface shape from the limited information of feature lines. Lumo [Joh02] generates an approximated normal field by interpolating normals on region boundaries and interior contours. Sýkora et al. [SKv*14] extended this approach with a simple set of user annotations to recover a full 3D shape for global illumination renderings. CrossShade [SBSS12] enables the user to design cross-section curves for better control of the constructed normal field. The CrossShade technique was extended by Iarussi et al. [IBB15] to construct generalized bend fields from rough sketches in a bitmap form. However, these approaches only focus on shape modeling from the boundary constraints. The recently proposed Inverse Toon Shading [XGS15] modeling tool also follows the strategy of modeling normal fields by designing isophote curves. In this work, the interpolation scheme assumes manual editing tasks to design two sets of isophotes with different illumination conditions for robust interpolation. In addition, reliable isophote values are also assumed. In contrast, our objective is to decompose a single cartoon shaded image to a shading representation that requires both a shape and a nonlinear color map reflectance.

Entire illumination constraint is considered in a well-known Shape-from-Shading (SFS) problem [HB89] for photorealistic images. Since the problem is severely ill-posed, accurate surface reconstruction requires skilled user interactions [OZM*06, WTBS07, WSTS08]. The user must specify the shape constraints to reduce the solution space of the SFS problem. To reduce user burden, another class of approach proposed a rough approximation from luminance gradients [KRFB06, LMJH*10] that can be tolerated by human perception. However, such approaches assume a photorealistic reflectance model, which often fails to analyze nonlinear shading in digital illustrations with large reconstruction errors.

Motivated by these considerations, we attempt to leverage limited cartoon shading information to model a smooth surface shape and nonlinear reflectance to reproduce the original shading appearance.

3. Methods

Figure 2 illustrates the main process of our shading analysis and relighting approach. The process includes 4 basic steps: initial normal estimation, reflectance estimation, normal refinement, and relighting. Assuming cartoon shading [LMHB00] for input shading $\mathbf{c}$, we will decompose the shading into a key light direction $\mathbf{L}$, a surface normal field $\mathbf{N}$, and a color map function M to fit $\mathbf{c} = M(\mathbf{L}\cdot\mathbf{N})$.

In the following sections, each step of the proposed shading analysis and relighting approaches is described in detail.

3.1. Initial Normal Estimation

For the target region Ω, we can obtain a rounded normal field $\mathbf{N}_0$ from the silhouette inflation constraints [Joh02, SKv*14]:

$$\begin{array}{rcll} \mathbf{N}_0(p) & = & \mathbf{N}_{\partial\Omega}(p) & p \in \partial\Omega \\ \Delta\mathbf{N}_0(p) & = & 0 & p \in \Omega \end{array}, \tag{1}$$

where $\mathbf{N}_{\partial\Omega} = (N_{\partial\Omega x}, N_{\partial\Omega y}, 0)$ is the normal constraint on the silhouette $\partial\Omega$. Note that these normals are propagated in the interior of Ω using a diffusion method [OBW*08]. As shown in Figure 3, we can obtain a smooth initial normal field $\mathbf{N}_0$ as a rounded shape.

3.2. Reflectance Estimation

Once the initial normal field $\mathbf{N}_0$ is obtained, our system estimates reflectance factors based on the cartoon shading representation $\mathbf{c} = M(\mathbf{L}\cdot\mathbf{N})$.

The reflectance estimation process takes the original color $\mathbf{c}$ and the initial normal $\mathbf{N}_0$ as inputs to estimate the light direction $\mathbf{L}$ and the color map function M. We assume that the scene is illuminated by a single key light direction (i.e., $\mathbf{L}$ is the same for the entire image). The color map function M is estimated for each target object.

In the early stage of our experiments, we observed that the key light estimation was significantly affected by the input material style and shape. Our simple experiment is summarized in the supplemental material. Since the $\mathbf{L}$ is a key factor in the following estimation steps, we assume that a reliable light direction is provided by the user. In our evaluation, we used a predefined ground-truth light direction $\mathbf{L}_t$ to observe errors caused by the other estimation steps.

Color Map Estimation

Given the smooth illumination result $I_0 = \mathbf{L}\cdot\mathbf{N}_0$, we estimate a color map function M to fit $\mathbf{c} = M(I_0)$.

Figure 4 shows an overview of our color map estimation process. Since I_0 and $\mathbf{c}$ are not in a strict monotonic relation, a straight forward minimization of $\sum_\Omega \|\mathbf{c} - M(I_0)\|^2$ produces a blurred color map M. To avoid an overlapping relation of I_0 and $\mathbf{c}$, we force monotonicity by sorting the target pixels in dark-to-bright order. From the sorted pixels, we can simply recover a color map function M by referencing the same luminance order of I_0 and $\mathbf{c}$.

3.3. Normal Refinement

As shown in the right image of Figure 4, the shading result of $M(\mathbf{L}\cdot\mathbf{N}_0)$ does not match $\mathbf{c}$ perfectly. Here we consider refining the normal $\mathbf{N}_0$ to reproduce the original color $\mathbf{c}$ by minimizing the following objective function:

$$E_N(\mathbf{N}) = \sum_\Omega \|\mathbf{c} - M(\mathbf{L}\cdot\mathbf{N})\|^2 + \lambda \sum_\Omega \|\Delta\mathbf{N}\|^2, \tag{2}$$

where $\sum_\Omega \|\mathbf{c} - M(\mathbf{L}\cdot\mathbf{N})\|^2$ forces the shading function to match the input shading, $\sum_\Omega \|\Delta\mathbf{N}\|^2$ is a smoothness constraint, and λ is a regularization factor for the smoothness constraint. Estimating $\mathbf{N}$ from Equation 2 is not straightforward due to the non-linear function of M.

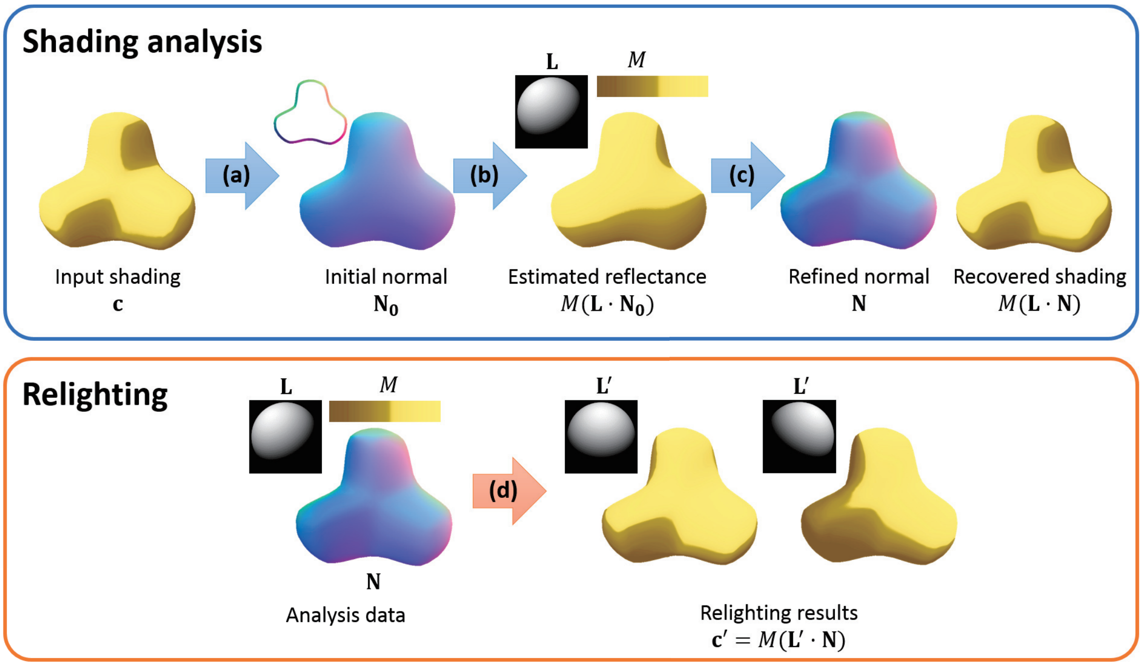

Figure 2: *Method overview. (a) Initial normal estimation to approximate a smooth rounded normal field. (b) Reflectance estimation to obtain a light and a color map. (c) Normal refinement to modify the initial normal by fitting the shading appearance. (d) Relighting to provide lighting interactions based on the shading analysis data.*

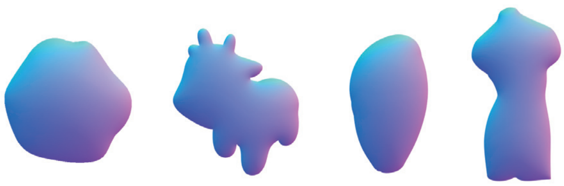

Figure 3: *Initial normal field obtained by silhouette inflation.*

To address this issue, we provide the complementary objective function of Equation 2:

$$E'_N(\mathbf{N}) = \sum_{\Omega} \|M^{-1}(\mathbf{c}) - \mathbf{L}\cdot\mathbf{N}\|^2 + \lambda \sum_{\Omega} \|\Delta\mathbf{N}\|^2, \quad (3)$$

where $M^{-1} : \mathbb{R}^3 \mapsto \mathbb{R}$ is the inverse function of M to change the appearance constraint into the illumination constraint. Since the constraint becomes a simple quadratic function, it can be minimized using Gauss-Seidel method with successive over relaxation until convergence to a local minimum.

Figure 5 illustrates the illumination constraints for the normal refinement process. Through the color map estimation process in Section 3.2, luminance range $[I_i, I_{i+1}]$ is known for each shading color $\mathbf{c}_i$. Therefore, the illumination is restricted by the following conditions:

$$\mathbf{L}\cdot\mathbf{N}(p) \in [I_i, I_{i+1}] \quad p \in \mathbf{C}_i, \quad (4)$$

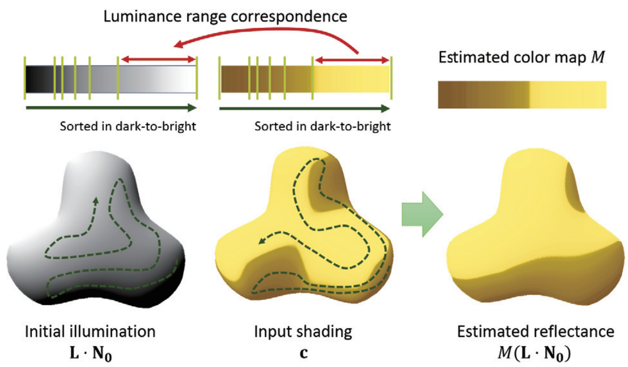

Figure 4: *Color map estimation. Given the set of illumination* $\mathbf{L}\cdot\mathbf{N}_0$ *and original color* $\mathbf{c}$*, a color map function* M *is estimated by matching the range of luminance orders.*

where $\mathbf{C}_i := \{p \in \Omega | \mathbf{c}(p) = \mathbf{c}_i\}$ is the quantized color area and illumination $\mathbf{L}\cdot\mathbf{N}(p)$ is constrained to $[I_i, I_{i+1}]$.

We solve the problem by minimizing the following energy:

$$E_N(\mathbf{N}) = \rho E_I(\mathbf{N}) + \lambda \sum_{\Omega} \|\Delta\mathbf{N}\|^2, \quad (5)$$

where $E_I(\mathbf{N}) = \sum_i \sum_{\mathbf{C}_i} P_i(\mathbf{L}\cdot\mathbf{N})$ is the luminance range constraint with penalty functions P_i, and ρ is the weight for $E_I(\mathbf{N})$. We define

P_i for each $\mathbf{C}_i$ as follows:

$$P_i(I) = \begin{cases} 0 & I_i \leq I \leq I_{i+1} \\ \|I - I_i\|^2 & I < I_i \\ \|I - I_{i+1}\|^2 & I > I_{i+1} \end{cases} \quad (6)$$

The parameter ρ can be adjusted to balance fitting against smoothness. The normal $\mathbf{N}$ is updated iteratively from the estimated initial normal $\mathbf{N}_0$ in the Gauss-Seidel iterations. Here we chose $\rho = 20.0$ and $\lambda = 30.0$ to obtain the refinement result. Compared to the initial normal $\mathbf{N}_0$, the refined normal $\mathbf{N}$ better fit the original color $\mathbf{c}$.

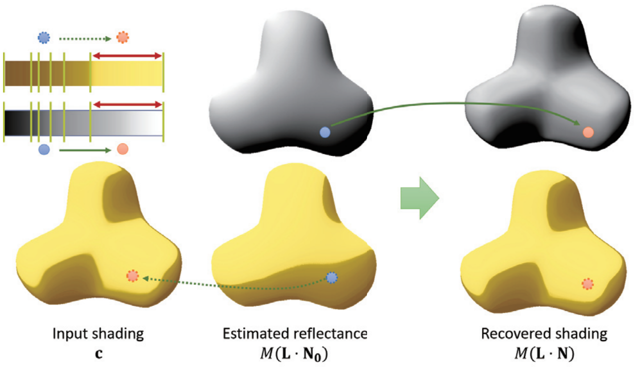

Figure 5: *Illumination constraints for normal refinement. The initial illumination result is modified by luminance range constraints derived from* M^{-1}.

3.4. Relighting

Based on the cartoon shading representation $\mathbf{c} = M(\mathbf{L} \cdot \mathbf{N})$, our system enables lighting interactions for the input illustration. We can obtain a relighting result $\mathbf{c}'$ by changing the light vector $\mathbf{L}$ to $\mathbf{L}'$ as follows:

$$\mathbf{c}' = M(\mathbf{L}' \cdot \mathbf{N}), \quad (7)$$

where the estimated factors M and $\mathbf{N}$ are preserved in the relighting process.

4. Evaluation of Shading Analysis

To evaluate our shading analysis approach, we conducted a simple pilot study via a ground-truth comparison. We compared our estimated results with several existing approaches and ground-truth inputs.

4.1. Experimental Design

To generate a variety of stylized appearance, we first prepared shape and color map datasets (see Figure 6).

Shape dataset. We prepared 20 ground-truth 3D models with different shape complexity and recognizability. This dataset includes 7 simple primitive shapes and 13 other shapes from 3D shape repositories. Each ground-truth model is rendered from a specific view point to generate a 512×512 normal field.

Color map dataset. To better understand real situations, we extracted the color maps from existing digital illustrations. We selected a small portion of a material area with a stroke. Then the selected pixels were simply sorted in luminance order to obtain a color map. We tried to select more than 100 material areas from different digital illustration sources. From the extracted color maps, we selected 24 distinctive color maps with different quantization effects.

Given the set of ground-truth normal field $\mathbf{N}_t$ and color map M_t, a final input image was obtained by $\mathbf{c}_t = M_t(\mathbf{L}_t \cdot \mathbf{N}_t)$. Note that we also provide a ground-truth light direction $\mathbf{L}_t$ in our evaluation process.

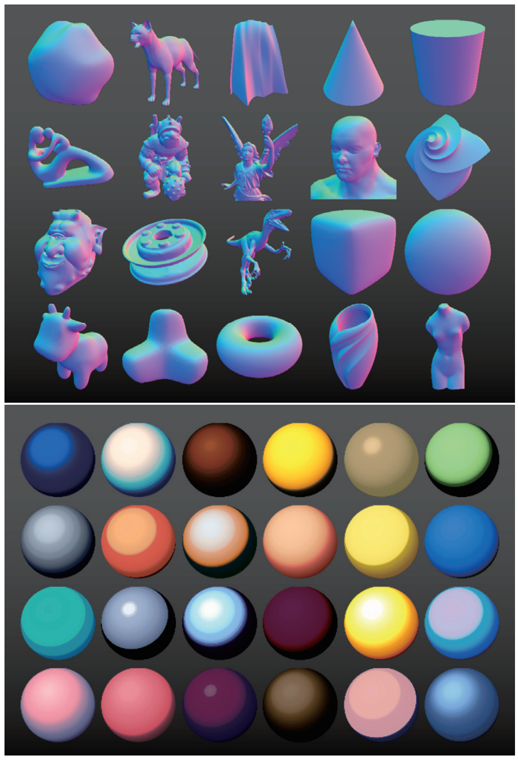

Figure 6: 20 *ground-truth 3D shapes and* 24 *color maps in our datasets.*

4.2. Shading Analysis

Figure 7 summarizes the comparison of our estimation results with Lumo [Joh02] and the Lambert assumption of [WSTS08]. To simulate Lumo we used the silhouette inflation constraints of the initial normal estimation in Equation 1. For the Lambert assumption, we used the illumination constraint in Equation 3 with a small value $\lambda = 1.0$ to fit the input image luminance I_c. In all examples we used our color map estimation method in Section 3.2 to reproduce the original shading appearance.

As shown in Figure 7, Lumo cannot produce the details of illumination due to the lack of inner shading constraints. The Lambert assumption recovers the original shading appearance well; however, the estimated normal field is over-fitted to the quantized illumination. Although our method distributes certain shading errors near the boundaries of the color areas, it produces a relatively smooth normal field and illumination that are similar to the ground-truth.

Further evaluations with different material and shape settings are summarized in the supplemental materials.

4.3. Relighting

Figure 8 and supplemental videos summarize the comparison of our relighting results with Lumo [Joh02] and the Lambert assumption of [WSTS08]. In all examples, we first estimate shading representations in the shading analysis step and then use the analysis data to produce relighting results.

Similar to the discussion in the previous evaluation for the shading analysis, our method and the Lambert assumption can preserve the original shading appearance for the input shading. However, the Lambert assumption tends to be strongly affected by the input initial illumination so that dynamic illumination changes from the input light directions are less noticeable in the relighting results. On the other hand, our method and Lumo can produce the dynamic illumination changes as in the ground-truth relighting results. Even our method cannot fully recover the details of the ground-truth shape, our shading decomposition result can provide both the dynamic illumination change and the details of the target shape.

5. Discussion and Future Work

In this paper, we have demonstrated a new shading analysis framework for cartoon shaded objects. The shading transition of the relighting results is improved by the proposed shading analysis. We incorporate the color map shading representation in our shading analysis approach, which enables the shading decomposition to a smooth normal field and a nonlinear color map reflectance. We have introduced a new way to provide lighting interaction with digital illustrations; however, there are several things left to accomplish.

First, the light estimation result is significantly affected by the input shading. More robust cartoon-shading-friendly estimation approaches are preferred. A perceptually motivated approach [LMGH*13] might be suitable. Second, currently, we use only silhouette constraints for the initial normal field approximation. As suggested by Lumo [Joh02], interior contours can improve the initial normal field. Even though we require a robust edge detection process to define suitable normal constraints for various illustration styles, this direction is a promising future work that may yield a more pleasing initial normal field. Another limitation is that the current formulation only minimizes the appearance error. Since the estimated shape is an important factor for relighting animation, we plan to integrate user constraints [OZM*06, WSTS08, SBSS12] for initial normal estimation.

We are currently investigating how our method may be applicable to real digital illustrations. Our initial attempt with simple character illustrations are demonstrated in the supplemental videos and material. While our initial experiments produced possible shading transitions via the diffuse shading assumption, our method cannot fully encode additional specular and shadow effects. Therefore, incorporating such specular and shadow models is an important future work for more practical situations. Such shading effects are often designed with more non-photorealistic principles; however, we hope that our approach will provide a promising direction for new 2.5D image representations of digital illustrations.

Acknowledgements

We would like to thank Tatsuya Yatagawa, Hiromu Ozaki, Tomohiro Tachi, and Takashi Kanai for their valuable discussions and suggestions. Additional thanks go to the AIM@SHAPE Shape Repository, Keenan's 3D Model Repository for the input 3D models used in this work. This work was supported in part by the JST CREST project and JSPS KAKENHI Grant Numbers JP15H05924.

References

[BTM06] Barla P., Thollot J., Markosian L.: X-toon: An extended toon shader. In *Proceedings of NPAR 2006* (New York, NY, USA, 2006), ACM, pp. 127–132. 1

[HB89] Horn B. K. P., Brooks M. J. (Eds.): *Shape from Shading*. MIT Press, Cambridge, MA, USA, 1989. 1, 2

[IBB15] Iarussi E., Bommes D., Bousseau A.: Bendfields: Regularized curvature fields from rough concept sketches. *ACM Trans. Graphics (SIGGRAPH 2015) 34*, 3 (May 2015), 24:1–24:16. 2

[Joh02] Johnston S. F.: Lumo: Illumination for cel animation. In *Proceedings of NPAR 2002* (New York, NY, USA, 2002), ACM, pp. 45–52. 2, 4, 5, 6

[KRFB06] Khan E. A., Reinhard E., Fleming R. W., Bülthoff H. H.: Image-based materival editing. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2006) 25*, 3 (July 2006), 654–663. 1, 2

[LMGH*13] Lopez-Moreno J., Garces E., Hadap S., Reinhard E., Gutierrez D.: Multiple light source estimation in a single image. *Computer Graphics Forum 32*, 8 (2013), 170–182. 5

[LMHB00] Lake A., Marshall C., Harris M., Blackstein M.: Stylized rendering techniques for scalable real-time 3d animation. In *Proceedings of NPAR 2000* (New York, NY, USA, 2000), ACM, pp. 13–20. 1, 2

[LMJH*10] Lopez Moreno J., Jimenez J., Hadap S., Reinhard E., Anjyo K., Gutierrez D.: Stylized depiction of images based on depth perception. In *Proceedings of NPAR 2010* (New York, NY, USA, 2010), ACM, pp. 109–118. 2

[MFE07] Mitchell J., Francke M., Eng D.: Illustrative rendering in Team Fortress 2. In *Proceedings of NPAR 2007* (New York, NY, USA, 2007), ACM, pp. 71–76. 1

[OBW*08] Orzan A., Bousseau A., Winnemöller H., Barla P., Thollot J., Salesin D.: Diffusion curves: A vector representation for smooth-shaded images. *ACM Trans. Graphics (SIGGRAPH 2008) 27*, 3 (Aug. 2008), 92:1–92:8. 2

[OZM*06] Okabe M., Zeng G., Matsushita Y., Igarashi T., Quan L., Yeung Shum H.: Single-view relighting with normal map painting. In *Proceedings of Pacific Graphics 2006* (2006), pp. 27–34. 1, 2, 5

[SBSS12] Shao C., Bousseau A., Sheffer A., Singh K.: Crossshade: Shading concept sketches using cross-section curves. *ACM Trans. Graphics (SIGGRAPH 2012) 31*, 4 (July 2012), 45:1–45:11. 2, 5

[SKv*14] Sýkora D., Kavan L., Čadík M., Jamriška O., Jacobson A., Whited B., Simmons M., Sorkine-Hornung O.: Ink-and-ray: Bas-relief meshes for adding global illumination effects to hand-drawn characters. *ACM Transactions on Graphics 33*, 2 (Apr. 2014), 16:1–16:15. 2

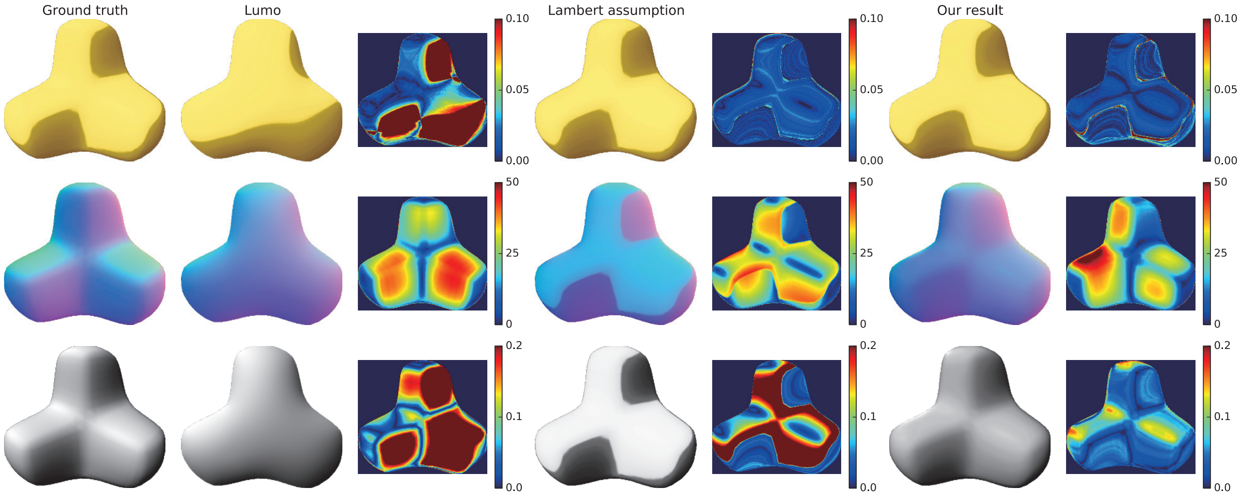

Figure 7: *Comparison of our shading analysis results with Lumo [Joh02] and the Lambert assumption of [WSTS08]. Our method reproduces the original shading appearance similar to the Lambert assumption with a smooth normal field as in Lumo.*

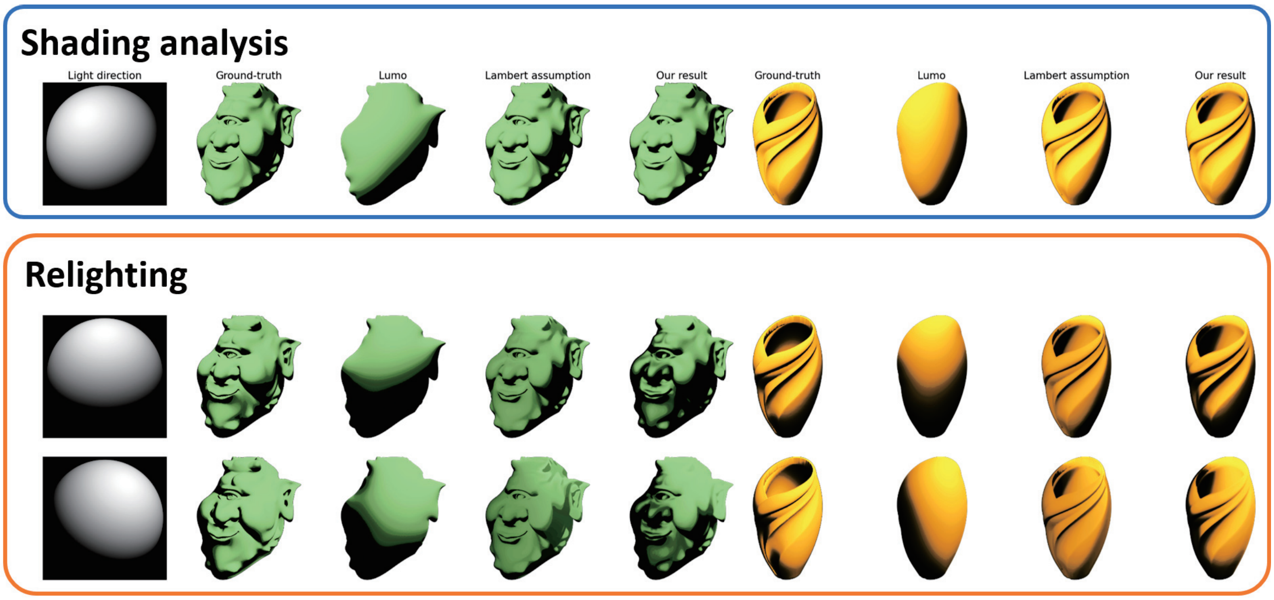

Figure 8: *Comparison of our relighting results with Lumo [Joh02] and the Lambert assumption of [WSTS08]. The shading analysis results show the recovered shading from the input ground-truth light direction and shading. The analyzed data are used to produce the next relighting results. Our method can produce dynamic illumination changes from the input light directions as in Lumo, which are less noticeable in the Lambert assumption. The details of the shapes are also preserved in our method.*

[SMGG01] Sloan P.-P. J., Martin W., Gooch A., Gooch B.: The lit sphere: A model for capturing npr shading from art. In *Proceedings of Graphics interface 2001* (Toronto, Ont., Canada, Canada, 2001), Canadian Information Processing Society, pp. 143–150. 1

[WSTS08] Wu T.-P., Sun J., Tang C.-K., Shum H.-Y.: Interactive normal reconstruction from a single image. *ACM Transactions on Graphics (SIGGRAPH Asia 2008) 27*, 5 (Dec. 2008), 119:1–119:9. 1, 2, 4, 5, 6

[WTBS07] Wu T.-P., Tang C.-K., Brown M. S., Shum H.-Y.: Shapepalettes: Interactive normal transfer via sketching. *ACM Transactions on Graphics (SIGGRAPH 2007) 26*, 3 (July 2007). 2

[XGS15] Xu Q., Gingold Y., Singh K.: Inverse toon shading: Interactive normal field modeling with isophotes. In *Proceedings of SBIM 2015* (Aire-la-Ville, Switzerland, Switzerland, 2015), Eurographics Association, pp. 15–25. 2