Research Article

Estimating reflectance and shape of objects from a single cartoon-shaded image

Hideki Todo[1] (✉), Yasushi Yamaguchi[2]

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 digital illustrations having a cartoon shading style. 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 reflection. To handle this non-photorealistic shading property, we focus on shading analysis of the most fundamental cartoon shading technique. Based on the color map shading representation, we propose a simple method to determine the input shading as that of 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.

Keywords   non-photorealistic rendering; cartoon shading; relighting; quantization

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 used roughly follows a physical illumination model, editing is restricted to 2D drawing operations. We are interested in exploring new interactions which allow relighting of a painted shading style given a single input image.

Reconstructing surface shape and reflectance from a single image is known as the *shape-from-shading* problem [1]. Based on the fundamental problem setting, most relighting approaches assume shading follows a Lambertian model [2–4]. Although these approaches work well for photorealistic images, they often fail to interpret 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 designed shading is typically more quantized than the inherent surface shape and its illumination. This assumption is common in many 3D stylized rendering techniques which use color map representation [5–7] that simply convert smooth 3D illumination to an artistic shading style. As shown in Fig. 1, this simple mechanism 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 such shading.
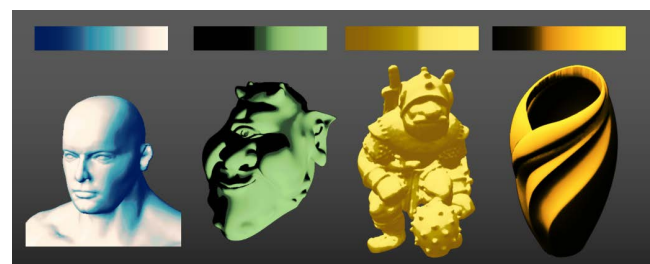
In this paper, we propose a simple shading

1   Tokyo University of Technology, Tokyo, 192-0982, Japan. E-mail: toudouhk@stf.teu.ac.jp (✉).

2   The University of Tokyo, Tokyo, 153-8902, Japan. E-mail: yama@graco.c.u-tokyo.ac.jp.

Fig. 1 Stylized shading styles obtained by color map representation.

analysis method to recover a reasonable shading representation from the input quantized shading. As a first step, we focus on the most fundamental cartoon shading [6]. Our primary assumption is that the main nonlinear factor in the final shading can be encoded by a color map function. With this in mind, we aim to reconstruct a smooth surface field and a nonlinear reflectance property from the input shading. Using these estimated data, our method provides a way to change the illumination of the input image with its quantized shading style. To evaluate our approach, we conducted a simple pilot study using a prepared set of 3D models and color maps with 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

Color mapping is a common approach used to generate stylized appearances in comics or illustrations. In stylized rendering of a 3D scene, the color map representation is used to convert smooth 3D illumination into quantized nonlinear shading effects [5–7]. Similar conversion techniques are used in 2D image abstraction methods for photorealistic images or videos [8–11]. As a starting point, our work follows the basic assumption that stylized shading appearance is based on a smooth surface shape.

Previous shape reconstruction methods for painted illustrations also attempt to recover a smooth surface shape from the limited information provided by feature lines. Lumo [12] generates an approximate normal field by interpolating normals on region boundaries and interior contours. Sýkora et al. [13] extended this approach with a simple set of user annotations to recover full 3D shape for global illumination rendering. CrossShade [14] 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. [15] 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* [16] modeling framework also follows the strategy of modeling normal fields by designing isophote curves. In this work, the interpolation

scheme requires manual editing 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 use a single cartoon-shaded image to provide a shading representation that contains both a shape and a nonlinear color map reflectance.

An entire illumination constraint is considered in the well-known shape-from-shading (SFS) problem [1] for photorealistic images. Since the problem is severely ill-posed, accurate surface reconstruction requires skilled user interaction [3, 4, 17]. The user must specify shape constraints to reduce the solution space of the SFS problem. To reduce user burden, another class of approach suggests rough approximation from luminance gradients [2, 18] that can be tolerated by human perception. However, such approaches assume a photorealistic reflectance model, which often results in large reconstruction errors for the nonlinear shading in digital illustrations.

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   Problem definition

3.1   Shading model assumptions

As proposed in the technique of cartoon shading [6], we assume a color map representation is used to reproduce the artist's nonlinear shading effects. Figure 2 illustrates the basic cartoon shading process. In this model, shading color $c \in \mathbb{R}^3$ is computed as follows:

$$c = M(I) \tag{1}$$

where $I \in \mathbb{R}$ is the luminance value of the
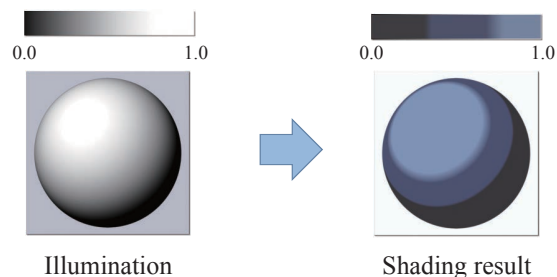

Fig. 2   Cartoon shading process.

illumination, and $M : \mathbb{R} \mapsto \mathbb{R}^3$ is a 1D color map function which converts the luminance value to the final shading color. For a diffuse shading material, we set $I = \boldsymbol{L} \cdot \boldsymbol{N}$, where $\boldsymbol{L}$ is a light vector and $\boldsymbol{N}$ is the surface normal vector. We are interested in manipulating $\boldsymbol{L}$ to $\boldsymbol{L}'$ to produce a new lighting result, i.e., $\boldsymbol{c}' = M(\boldsymbol{L}' \cdot \boldsymbol{N})$.

However, the inverse problem is ill-posed if only shading color $\boldsymbol{c}$ is available. The primary consideration of this paper is that we limit the solution space for other factors while preserving the final shading appearance. Some basic assumptions considered in this paper are as follows.

- **Smooth shape and illumination.** We assume that the surface shape $\boldsymbol{N}$ and the illumination I are smooth and follow a linear relationship. The only nonlinear factor is the color map function M, which is used to produce the stylized shading appearance.

- **Monotonic function for color map.** For the color map function M, we assume a monotonic relation between image luminance I_c (obtained from $\boldsymbol{c}$) and surface illumination I. This assumption is important to simplify our problem definition as a variation of a photorealistic relighting problem.

- **Diffuse lighting for illumination.** We analyze all shading effects as due to diffuse lighting. We do not explicitly model specular reflections and shadows in our shading analysis experiments.

4 Methods

Figure 3 illustrates the main process of the proposed shading analysis and relighting approach. Here we provide the primary objective and summarize each step.

- **Initial normal estimation.** First, an initial normal field $\boldsymbol{N}_0$ is required as input for the reflectance estimation and normal refinement steps. Since the reflectance property is not available, we simply approximate a smooth rounded normal field from the silhouette.

- **Reflectance estimation.** Given the initial normal field $\boldsymbol{N}_0$, we estimate a key light direction $\boldsymbol{L}$ and a color map function M which best fit $\boldsymbol{c} = M(\boldsymbol{L} \cdot \boldsymbol{N}_0)$. This decomposition result roughly matches the original shading $\boldsymbol{c}$ for the given $\boldsymbol{N}_0$.

- **Normal refinement.** Since the estimated decomposition does not satisfy $\boldsymbol{c} = M(\boldsymbol{L} \cdot \boldsymbol{N}_0)$, we refine the surface normal $\boldsymbol{N}_0$ to $\boldsymbol{N}$ to reproduce the original shading $\boldsymbol{c}$.
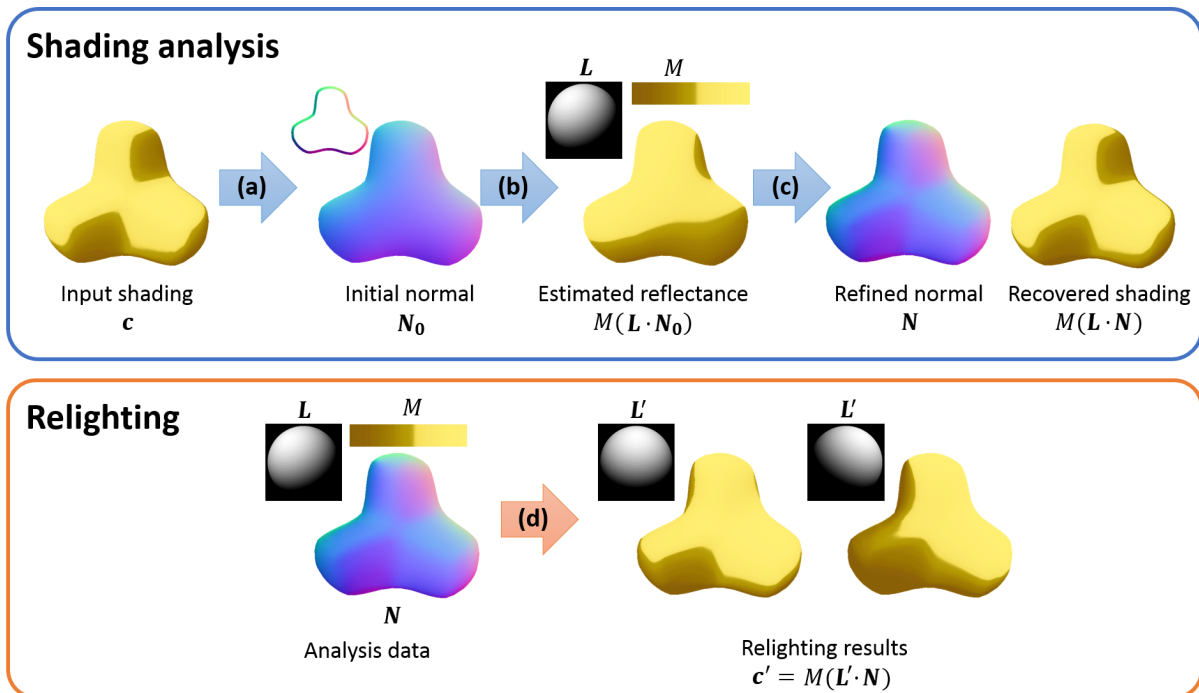

Fig. 3 Method overview. (a) Initial normal estimation to approximate a smooth rounded normal field. (b) Reflectance estimation to obtain a light and a color map. (c) Normal refinement to modify the initial normal by fitting the shading appearance. (d) Relighting to provide lighting interactions based on the shading analysis data.

- **Relighting.** Based on the above analysis results, the proposed method can relight the given input illustration. We change the light vector $\boldsymbol{L}$ to $\boldsymbol{L}'$ to obtain the final shading color $\boldsymbol{c}' = M(\boldsymbol{L}' \cdot \boldsymbol{N})$.

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

4.1 Initial normal estimation

For the target region Ω, we can obtain a rounded normal field $\boldsymbol{N}_0$ from the silhouette inflation constraints [12, 13]:

$$\begin{cases} \boldsymbol{N}_0(p) = \boldsymbol{N}_{\partial\Omega}(p), & p \in \partial\Omega \\ \Delta\boldsymbol{N}_0(p) = 0, & p \in \Omega \end{cases} \quad (2)$$

where $\boldsymbol{N}_{\partial\Omega} = (N_{\partial\Omega x}, N_{\partial\Omega y}, 0)$ is the normal constraint from the silhouette $\partial\Omega$. These normals are propagated to the interior of Ω using a diffusion method [19]. As shown in Fig. 4, we can obtain a smooth initial normal field $\boldsymbol{N}_0$ as a rounded shape.

4.2 Reflectance estimation

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

The reflectance estimation process takes the original color $\boldsymbol{c}$ and the initial normal $\boldsymbol{N}_0$ as inputs to estimate the light direction $\boldsymbol{L}$ and the color map function M. We assume that the scene is illuminated by a single key light direction (i.e., $\boldsymbol{L}$ 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 step was significantly affected by the input material style and shape. Our simple experiment is summarized in the Appendix. Since $\boldsymbol{L}$ 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 $\boldsymbol{L}_t$ to observe errors caused by the other estimation steps.

Color map estimation. Given the smooth

Fig. 4 Initial normal field obtained by silhouette inflation.

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

As shown in Fig. 5, isophote pixels of I_0 do not provide the same color as $\boldsymbol{c}$. Therefore, a straight forward minimization of $\sum_{\Omega} \|\boldsymbol{c} - M(I_0)\|^2$ produces a blurred color map M.

To avoid this invalid correspondence between I_0 and $\boldsymbol{c}$, we force monotonicity by sorting the target pixels in dark-to-bright order as shown in Fig. 6. From the sorted pixels, we can obtain a valid correspondence between luminance range $[I_i, I_{i+1}]$ and each shading color $\boldsymbol{c}_i$ in the same luminance order. As a result, a color map function M is recovered as a lookup table for obtaining $\boldsymbol{c}_i$ from $[I_i, I_{i+1}]$. We also construct the corresponding inverse map M^{-1}, which is an additional lookup table to retrieve the luminance range $[I_i, I_{i+1}]$ from a shading color $\boldsymbol{c}_i$.

4.3 Normal refinement

As shown in the right image of Fig. 6, the shading result of $M(\boldsymbol{L} \cdot \boldsymbol{N}_0)$ does not match $\boldsymbol{c}$ perfectly. Here we consider refining normal $\boldsymbol{N}_0$ to reproduce the
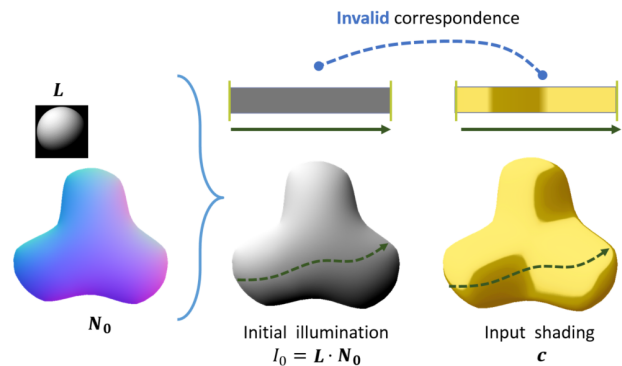
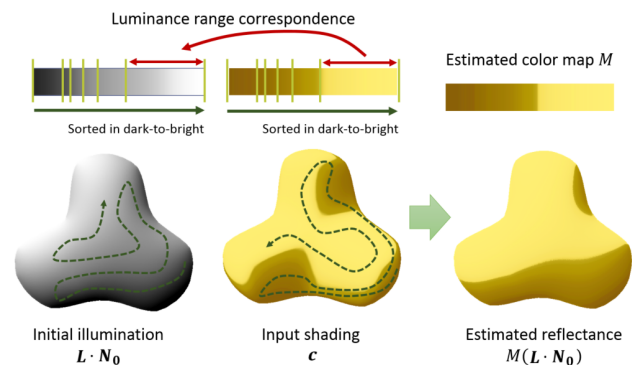**Fig. 5** Invalid correspondence between the initial illumination I_0 and the input shading $\boldsymbol{c}$.

Fig. 6 Color map estimation. Given the set of illumination $\boldsymbol{L} \cdot \boldsymbol{N}_0$ and original color $\boldsymbol{c}$, a color map function M is estimated by matching the range of luminance orders.

original color c by minimizing the following objective function:

$$E_N(\boldsymbol{N}) = \sum_{\Omega} \| \boldsymbol{c} - M(\boldsymbol{L} \cdot \boldsymbol{N}) \|^2 + \lambda \sum_{\Omega} \| \Delta \boldsymbol{N} \|^2 \quad (3)$$

where $\sum_{\Omega} \| \boldsymbol{c} - M(\boldsymbol{L} \cdot \boldsymbol{N}) \|^2$ forces the shading function to match the input shading, $\sum_{\Omega} \| \Delta \boldsymbol{N} \|^2$ is a smoothness constraint, and λ is a regularization factor for the smoothness constraint. Estimating $\boldsymbol{N}$ from Eq. (3) is not straightforward due to the non-linear function of M.

To address this issue, we provide the following complementary objective function to Eq. (3):

$$E_N'(\boldsymbol{N}) = \sum_{\Omega} \| M^{-1}(\boldsymbol{c}) - \boldsymbol{L} \cdot \boldsymbol{N} \|^2 + \lambda \sum_{\Omega} \| \Delta \boldsymbol{N} \|^2 \quad (4)$$

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 the Gauss–Seidel method with successive over relaxation until convergence to a local minimum. When the inverse function M^{-1} is simply defined from the image luminance I_c, Eq. (4) is the same as the photorealistic formulation suggested in a previous study [4]:

$$E_N'(\boldsymbol{N}) = \sum_{\Omega} \| k_{\mathrm{d}}' I_c - \boldsymbol{L} \cdot \boldsymbol{N} \|^2 + \lambda \sum_{\Omega} \| \Delta \boldsymbol{N} \|^2 \quad (5)$$

where $k_{\mathrm{d}}' = 1/k_{\mathrm{d}}$ is the reciprocal of the diffuse reflectance constant. In our case, we can define the inverse function M^{-1} from the estimated color map function M.

Figure 7 illustrates the illumination constraints for the normal refinement process. From the color map estimation process described in Section 4.2, the luminance range $[I_i, I_{i+1}]$ is known for each shading color $\boldsymbol{c}_i$. Therefore, the illumination is restricted by the following conditions:

$$\boldsymbol{L} \cdot \boldsymbol{N}(p) \in [I_i, I_{i+1}], \quad p \in \boldsymbol{C}_i \quad (6)$$

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

We solve the problem by minimizing the following energy:

$$E_N(\boldsymbol{N}) = E_I(\boldsymbol{N}) + \lambda \sum_{\Omega} \| \Delta \boldsymbol{N} \|^2 \quad (7)$$

where $E_I(\boldsymbol{N}) = \sum_i \sum_{\boldsymbol{C}_i} P_i(\boldsymbol{L} \cdot \boldsymbol{N})$ is the luminance range constraint with penalty functions P_i. We define P_i for each $\boldsymbol{C}_i$ as follows:
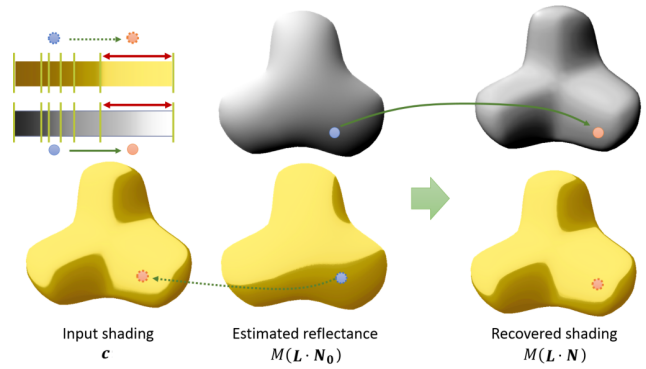

Fig. 7 Illumination constraints for normal refinement. The initial illumination result is modified by luminance range constraints derived from M^{-1}.

$$P_i(I) = \begin{cases} 0, & I_i \leqslant I \leqslant I_{i+1} \\ \| I - I_i \|^2, & I < I_i \\ \| I - I_{i+1} \|^2, & I > I_{i+1} \end{cases} \quad (8)$$

The normal $\boldsymbol{N}$ is updated iteratively from the estimated initial normal $\boldsymbol{N}_0$ in Gauss–Seidel iterations. Here we chose $\lambda = 1.5$ to obtain the refinement result. Compared to the initial normal $\boldsymbol{N}_0$, the refined normal $\boldsymbol{N}$ better fits the original color c.

4.4 Relighting

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

$$\boldsymbol{c}' = M(\boldsymbol{L}' \cdot \boldsymbol{N}) \quad (9)$$

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

5 Evaluation of shading analysis

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

5.1 Experimental design

To generate a variety of stylized appearance, we first prepared shape and color map datasets (see Fig. 8).

Shape dataset. We prepared 20 ground-truth 3D models having varying 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
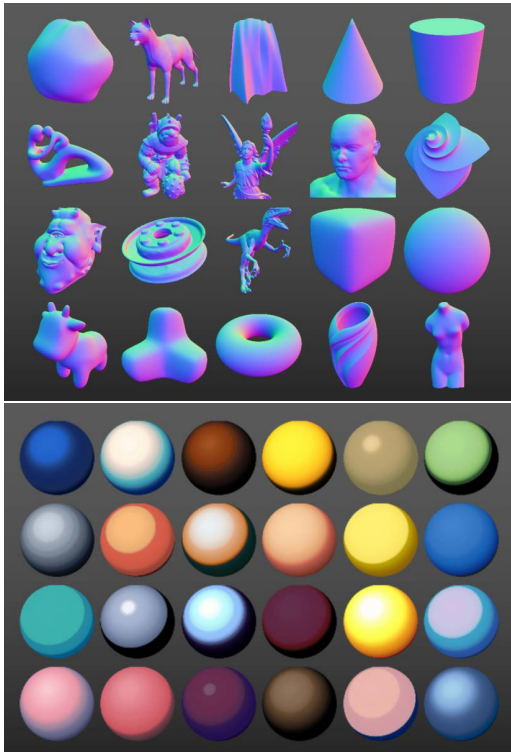
from a specific view point to generate a 512×512 normal field.

Color map dataset. To better understand real situations, we extracted 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 extract more than 100 material areas from different digital illustrations sources. From the extracted color maps, we selected 24 distinctive color maps with different quantization effects.

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

5.2 Comparison of reflectance models

We first compared the visual difference between our target cartoon shading model and a common photorealistic Lambertian model as shown in Fig. 9. To obtain an ambient color $\boldsymbol{k}_\mathrm{a}$ and a diffuse reflectance color $\boldsymbol{k}_\mathrm{d}$ for the Lambertian shading representation $\boldsymbol{c} = \boldsymbol{k}_\mathrm{a} + \boldsymbol{k}_\mathrm{d}I$, we minimized $\|M(I) - (\boldsymbol{k}_\mathrm{a} + \boldsymbol{k}_\mathrm{d}I)\|^2$ with the input color map function M. The color difference suggests that cartoon
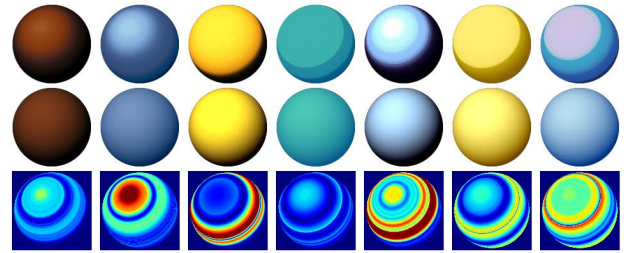

Fig. 9 Comparison of reflectance models. Top: color map materials selected from our dataset. Middle: Lambertian material fitted to the corresponding color map. Bottom: color difference between the color map materials and Lambertian materials. The materials are listed according to the color difference.

shading includes some nonlinear parts, which cannot be described by a simple Lambertian model. We will discuss how this nonlinear reflectance property affects the estimation results.

5.3 Shading analysis

Figure 10 summarizes a comparison of our estimation results with ones from Lumo [12] and the Lambertian assumption [4]. To simulate Lumo we used the silhouette inflation constraints of the initial normal estimation in Eq. (2). For the Lambertian assumption, we used the illumination constraint in Eq. (5) 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 (Section 4.2) to reproduce the original shading appearance.

As shown in Fig. 10, Lumo cannot produce the details of illumination due to the lack of inner shading constraints. The Lambertian assumption recovers the original shading appearance well; however, the estimated normal field is overfitted 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 both similar to the ground-truth.

Figure 11 summarizes the shading analysis results for different material settings. Although our method cannot recover the same shape from different quantization styles, the estimated normal field is smoother than the input shading.

We also compute the mean squared error (MSE) to compare estimated results quantitatively (see Figs. 12–15). In each comparison, we used the same shape and changed materials for computing the shape estimation errors.

Note that our method tends to produce smaller errors for simple rounded shapes but the errors
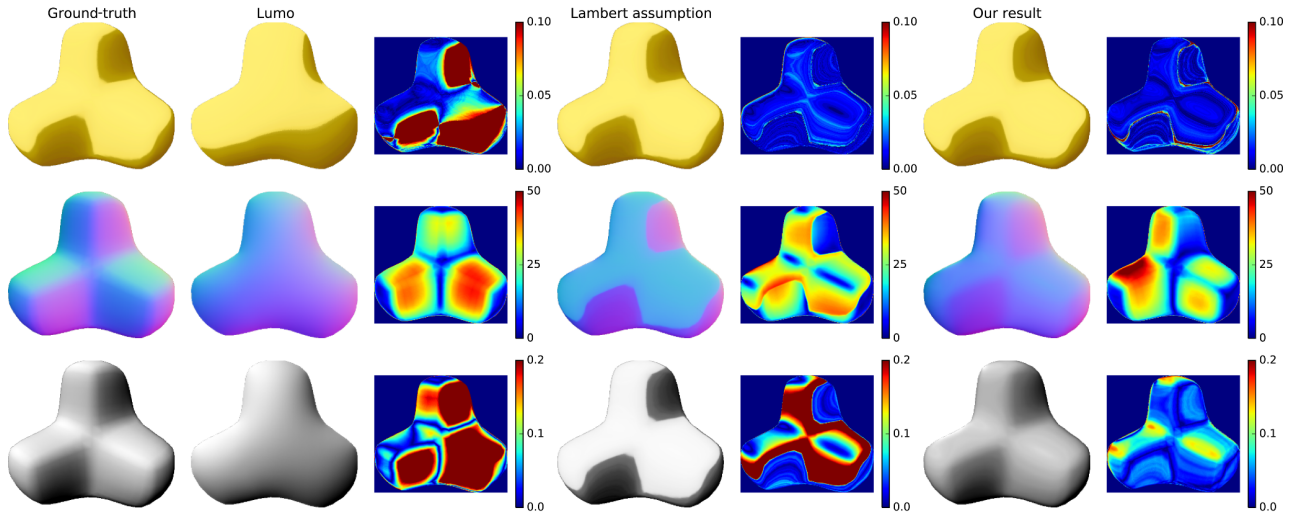
Fig. 10 Comparison of shading analysis results with Lumo [12] and Lambertian assumption [4]. The proposed method reproduces the original shading appearance similar to the Lambertian assumption with a smooth normal field as in Lumo.
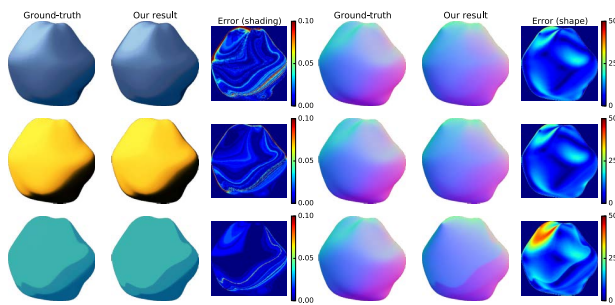

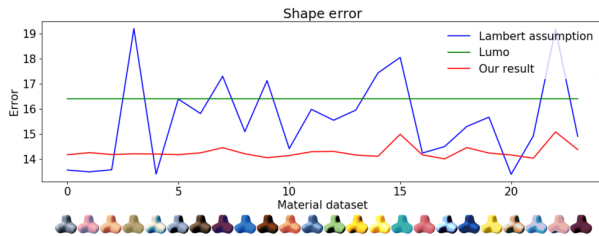**Fig. 11** Shading analysis results for different color map materials.

Fig. 12 Errors of estimated shape depending on input material (simple shape *Three Box*).
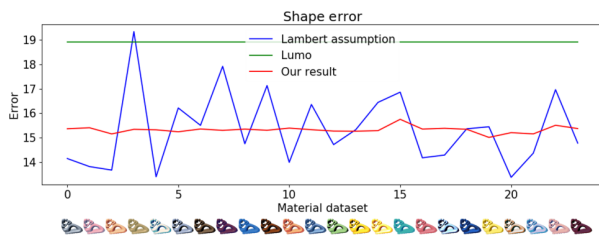

Fig. 13 Errors of estimated shape depending on input material (medium complexity shape *Fertility*).

become larger than the Lambertian assumption for more complex shapes. For a complex shape like the *Pulley* shown in Fig. 15, even the Lambertian assumption results in large errors. Since initial
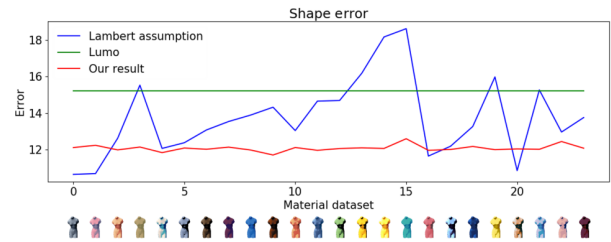

Fig. 14 Errors of estimated shape depending on input material (medium complexity shape *Venus*).
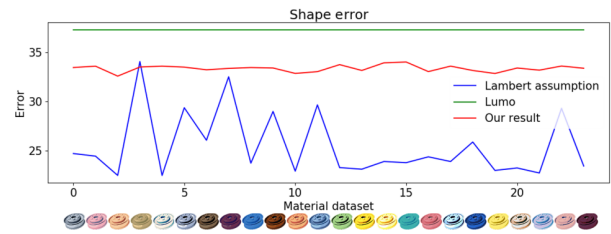

Fig. 15 Errors of estimated shape depending on input material (complex shape *Pulley*).

normal estimation errors become large in such cases, our method fails to recover a valid shape when only minimizing the appearance error. We provide further discussions on initial normal estimation errors in Section 7.

Though the estimated shape may not be accurate, our method successfully reduces the influence of the material difference in all comparisons. Thanks to the proposed shading analysis based on the cartoon shading model assumption, our method regulates estimated reflectance properties for various quantization settings.

5.4 Relighting

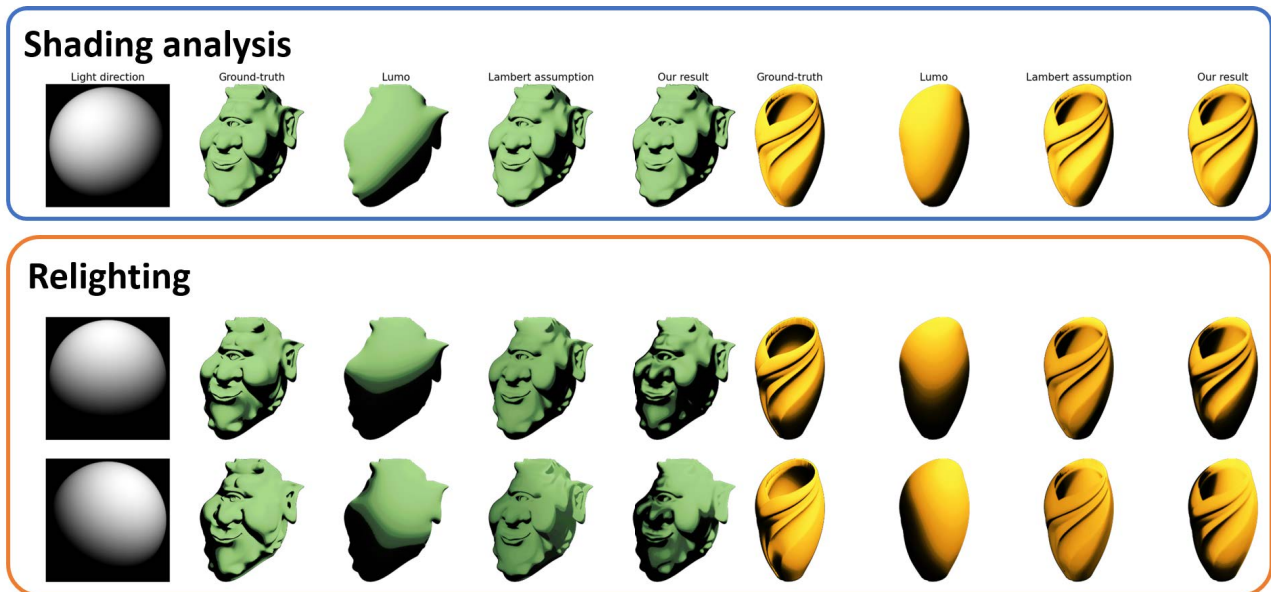Figure 16 and the supplemental videos in

Fig. 16 Comparison of our relighting results with those from Lumo [12] and using the Lambertian assumption in Ref. [4]. The shading analysis shows the estimated shading results from the input ground-truth light direction and shading. The analysis data are used to produce the following relighting results. Our method can produce dynamic illumination changes from the input light directions as in Lumo, which are less noticeable in the Lambertian assumption. The details of the shapes are also preserved in our method.

the Electronic Supplementary Material (ESM) summarize a comparison of our relighting results with those from Lumo [12] and using the Lambertian assumption in Ref. [4]. In all examples, we first estimate the shading representations in the shading analysis step. Then we use the analysis data to produce relighting results.

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

6   Real illustration examples

We have tested our shading analysis approach on different shading styles using three real illustrations.

Figure 17 shows relighting results for the one of them, the others are included in the supplemental videos in the ESM. The material regions are relatively simple, but each material region is painted with different quantization effects.

To apply our shading analysis and relighting methods, we first manually segmented material regions for the target illustration. We also provide a key light direction L for the target illustration, which is needed for our reflectance estimation step.

Fig. 17 Relighting sequence using the proposed method. Non-diffuse parts are limited to static transitions with simple residual representation.

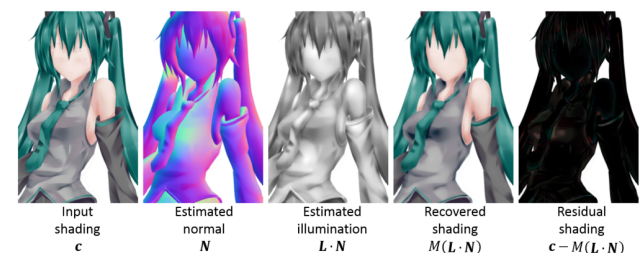| Input shading c | Estimated normal N | Estimated illumination $L \cdot N$ | Recovered shading $M(L \cdot N)$ | Residual shading $c - M(L \cdot N)$ |

Fig. 18 Reflectance and shape estimation results for a real illustration. Non-diffuse parts are encoded as residual shading.

Figure 18 illustrates the elements of reflectance and shape estimation results for the illustration. Compared to the ideal cartoon shading in our evaluations, a material region in the real examples may include non-diffuse parts. As suggested by a photorealistic illumination estimation method [20], we encode such specular and shadow effects as residual differences $\Delta \boldsymbol{c} = \boldsymbol{c} - M(\boldsymbol{L} \cdot \boldsymbol{N})$ from our assumed shading representation $\boldsymbol{c} = M(\boldsymbol{L} \cdot \boldsymbol{N})$. Finally, we obtain relighting results as $\boldsymbol{c} = M(\boldsymbol{L}' \cdot \boldsymbol{N}) + \Delta \boldsymbol{c}$ by changing the light direction $\boldsymbol{L}'$.

As shown in Fig. 17 and the supplemental videos in the ESM, the residual representation can recover the appearance of the original shading. We also note that our initial experiment produced possible shading transitions for diffuse lighting, while specular and shadow effects are relatively static.

7 Discussion and future work

In this paper, we have demonstrated a new shading analysis framework for cartoon-shaded objects. The visual appearance of the relighting results is improved by the proposed shading analysis. We incorporate color map shading representation in our shading analysis approach, which enables shading decomposition into 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.

Firstly, our method requires a reliable light direction which is provided by the user. Since the light estimation method in the Appendix is significantly affected by the input shading, more friendly and robust cartoon shading estimation approaches are needed. We consider that a perceptually motivated approach [21] might be suitable.

Secondly, the method minimizes the appearance error, because a shading image is the only input. This results in an under-constrained problem to estimate both shape and reflectance. Actually, our method achieves almost the same appearance as the input. As shown in Fig. 19, the proposed method cannot recover the input shape even if the material has Lambertian reflectance with full illumination constraints. Although the recovered shape satisfies appearance similarity with the color map that is
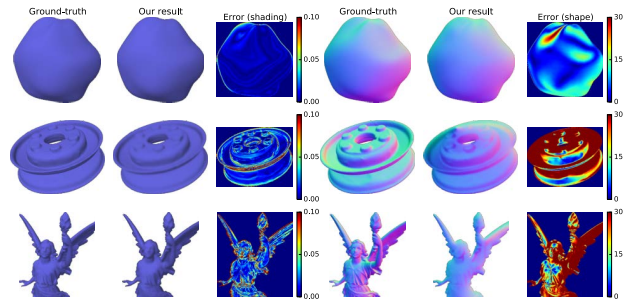

Fig. 19 Shape analysis results for Lambertian reflectance. *Blob* (top): small errors in shape and shading. *Pulley* (middle): large errors in shape. *Lucy* (bottom): large errors in shading.

estimated in advance, we need a better solution space to obtain a plausible shape. Since a desirable shape is typically different for different users, we plan to integrate user constraints [3, 4, 14] for normal refinement. More robust iterated refinement cycles of shape and reflectance estimations are to be desired.

Another limitation is that our initial normal field approximation assumes the shape to be convex. This causes errors noticeable in complex shapes such as the *Pulley*, as shown in Fig. 19. Currently, we also plan to incorporate interior contours for concave constraints as suggested by Lumo [12]. Even though we require a robust edge detection process to define suitable normal constraints for various illustration styles, this is a promising direction for future work that may yield a more pleasing initial normal field.

Although large collections of 2D digital illustrations are available online, we cannot directly apply our method since we require manual segmentation. A crucial area of future research is to automate albedo estimation, as suggested by intrinsic images [22, 23]. While our initial experiments with manual segmentation 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 using non-photorealistic principles; however, we hope that our approach will provide a promising direction for new 2.5D image representations of digital illustrations.

Appendix   Light estimation

In the early stage of our experiments, we tried to

estimate the key light direction $\boldsymbol{L}$ from the input shading $\boldsymbol{c}$ and the estimated initial normal $\boldsymbol{N}_0$.

As suggested by Ref. [4], we approximate the problem using Lambertian reflectance $I_c = k_{\mathrm{d}}\boldsymbol{L}\cdot\boldsymbol{N}_0$, where the diffuse term $\boldsymbol{L}\cdot\boldsymbol{N}_0$ is simply scaled by the diffuse constant k_{d}. For the input illumination I_c, we compute the luminance value from the original color $\boldsymbol{c}$ as the L component in *Lab* color space. We estimate the light vector $\boldsymbol{L}$ by minimizing the following energy:

$$E_L(\boldsymbol{L}') = \sum_{\Omega} \|I_c - \boldsymbol{L}'\cdot\boldsymbol{N}_0\|^2 \qquad (10)$$

where $\boldsymbol{L}'$ is given by $\boldsymbol{L}' = k_{\mathrm{d}}\boldsymbol{L}$. We finally obtain the unit light vector $\boldsymbol{L} = \boldsymbol{L}'/\|\boldsymbol{L}'\|$ by normalizing $\boldsymbol{L}'$. The diffuse reflectance constant k_{d} is optionally computed from $k_{\mathrm{d}} = \|\boldsymbol{L}'\|$.

Figure 20 summarizes our experiment for light estimation. In this experiment, we give a single ground-truth light direction $\boldsymbol{L}_{\mathrm{t}}$ (top left) to generate the input cartoon-shaded image $\boldsymbol{c}_{\mathrm{t}}$ and then estimate a key light direction $\boldsymbol{L}$ by solving Eq. (10).

It can be observed that the estimated results look consistent with near-Lambertian materials (the left 3 maps) but inconsistent with more stylized materials (the right 3 maps). Another important factor is the shape complexity. The estimated light direction is
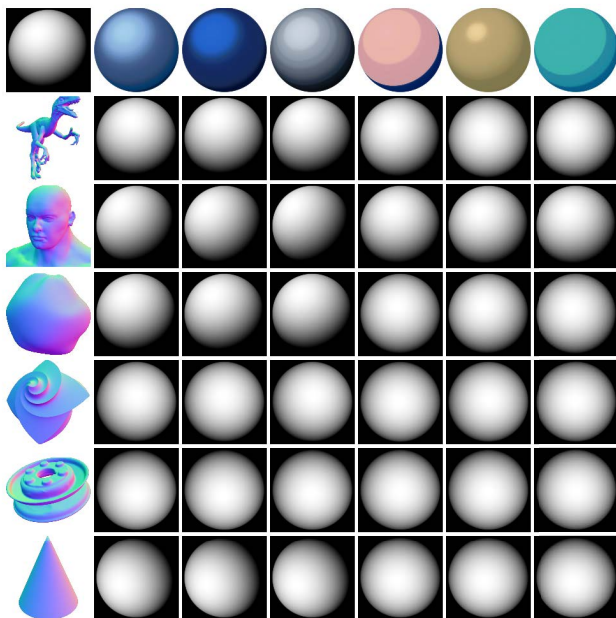

Fig. 20 Light estimation error. Top left: input ground-truth light direction $\boldsymbol{L}_{\mathrm{t}}$. Top row: input color map materials shaded from the $\boldsymbol{L}_{\mathrm{t}}$. The left 3 maps have small average errors; the right 3 maps have large average errors. Left column: input 3D models. The top 3 models have small average errors; the bottom 3 models have large average errors.

relatively consistent with rounded smooth shapes. However, the light estimation error becomes quite large when the input model contains many crease edges, especially around the silhouette.

The result suggests that we require additional constraints to improve light estimation. In this paper, we simply provide a ground-truth light direction for evaluation, or a user-given reliable light direction for relighting real illustration examples.
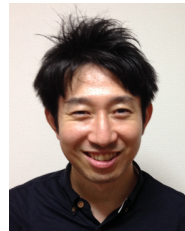
Acknowledgements

References

[1] Horn, B. K. P.; Brooks, M. J. *Shape from Shading*. Cambridge, MA, USA: MIT Press, 1989.

[2] Khan, E. A.; Reinhard, E.; Fleming, R. W.; Bülthoff, H. H. Image-based material editing. *ACM Transactions on Graphics* Vol. 25, No. 3, 654–663, 2006.

[3] Okabe, M.; Zeng, G.; Matsushita, Y.; Igarashi, T.; Quan, L.; Shum, H.-Y. Single-view relighting with normal map painting. In: Proceedings of Pacific Graphics, 27–34, 2006.

[4] Wu, T.-P.; Sun, J.; Tang, C.-K.; Shum, H.-Y. Interactive normal reconstruction from a single image. *ACM Transactions on Graphics* Vol. 27, No. 5, Article No. 119, 2008.

[5] Barla, P.; Thollot, J.; Markosian, L. X-toon: An extended toon shader. In: Proceedings of the 4th International Symposium on Non-Photorealistic Animation and Rendering, 127–132, 2006.

[6] Lake, A.; Marshall, C.; Harris, M.; Blackstein, M. Stylized rendering techniques for scalable real-time 3D animation. In: Proceedings of the 1st International Symposium on Non-Photorealistic Animation and Rendering, 13–20, 2000.

[7] Mitchell, J.; Francke, M.; Eng, D. Illustrative rendering in *Team Fortress 2*. In: Proceedings of the 5th International Symposium on Non-Photorealistic Animation and Rendering, 71–76, 2007.

[8] DeCarlo, D.; Santella, A. Stylization and abstraction of photographs. *ACM Transactions on Graphics* Vol. 21, No. 3, 769–776, 2002.

[9] Kang, H.; Lee, S.; Chui, C. K. Flow-based image abstraction. *IEEE Transactions on Visualization and Computer Graphics* Vol. 15, No. 1, 62–76, 2009.

[10] Kyprianidis, J. E.; Döllner, J. Image abstraction by structure adaptive filtering. In: Proceedings of EG UK Theory and Practice of Computer Graphics, 51–58, 2008.

[11] Winnemöller, H.; Olsen, S. C.; Gooch, B. Real-time video abstraction. *ACM Transactions on Graphics* Vol. 25, No. 3, 1221–1226, 2006.

[12] Johnston, S. F. Lumo: Illumination for cel animation. In: Proceedings of the 2nd International Symposium on Non-Photorealistic Animation and Rendering, 45–52, 2002.

[13] 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* Vol. 33, No. 2, Article No. 16, 2014.

[14] Shao, C.; Bousseau, A.; Sheffer, A.; Singh, K. CrossShade: Shading concept sketches using cross-section curves. *ACM Transactions on Graphics* Vol. 31, No. 4, Article No. 45, 2012.

[15] Iarussi, E.; Bommes, D.; Bousseau, A. Bendfields: Regularized curvature fields from rough concept sketches. *ACM Transactions on Graphics* Vol. 34, No. 3, Article No. 24, 2015.

[16] Xu, Q.; Gingold, Y.; Singh, K. Inverse toon shading: Interactive normal field modeling with isophotes. In: Proceedings of the Workshop on Sketch-Based Interfaces and Modeling, 15–25, 2015.

[17] Wu, T.-P.; Tang, C.-K.; Brown, M. S.; Shum, H.-Y. ShapePalettes: Interactive normal transfer via sketching. *ACM Transactions on Graphics* Vol. 26, No. 3, Article No. 44, 2007.

[18] Lopez-Moreno, J.; Jimenez, J.; Hadap, S.; Reinhard, E.; Anjyo, K.; Gutierrez, D. Stylized depiction of images based on depth perception. In: Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering, 109–118, 2010.

[19] Orzan, A.; Bousseau, A.; Barla, P.; Winnemöller, H.; Thollot, J.; Salesin, D. Diffusion curves: A vector representation for smooth-shaded images. *Communications of the ACM* Vol. 56, No. 7, 101–108, 2013.

[20] Kholgade, N.; Simon, T.; Efros, A.; Sheikh, Y. 3D object manipulation in a single photograph using stock 3D models. *ACM Transactions on Graphics* Vol. 33, No. 4, Article No. 127, 2014.

[21] Lopez-Moreno, J.; Garces, E.; Hadap, S.; Reinhard, E.; Gutierrez, D. Multiple light source estimation in a single image. *Computer Graphics Forum* Vol. 32, No. 8, 170–182, 2013.

[22] Grosse, R.; Johnson, M. K.; Adelson, E. H.; Freeman, W. T. Ground truth dataset and baseline evaluations for intrinsic image algorithms. In: Proceedings of IEEE 12th International Conference on Computer Vision, 2335–2342, 2009.

[23] Rother, C.; Kiefel, M.; Zhang, L.; Schölkopf, B.; Gehler, P. V. Recovering intrinsic images with a global sparsity prior on reflectance. In: Proceedings of Advances in Neural Information Processing Systems 24, 765–773, 2011.

Hideki Todo is an assistant professor in the School of Media Science at Tokyo University of Technology. He received his Ph.D. degree in information science and technology from the University of Tokyo in 2013. His research interests lie in the field of computer graphics in general, particularly non-photorealistic rendering.

Yasushi Yamaguchi, Dr. Eng., is a professor in the Graduate School of Arts and Sciences at the University of Tokyo. His research interests lie in image processing, computer graphics, and visual illusion, including visual cryptography, computer aided geometric design, volume visualization, and painterly rendering. He has been serving as a president of Japan Society for Graphic Science and as a vice president of International Society for Geometry and Graphics.

Other papers from this open access journal are available free of charge from http://www.springer.com/journal/41095. To submit a manuscript, please go to https://www.editorialmanager.com/cvmj.