The Hand as a Shading Probe

Yasuhiro Yao* Harumi Kawamura† Akira Kojima‡
Nippon Telegraph and Telephone Corporation

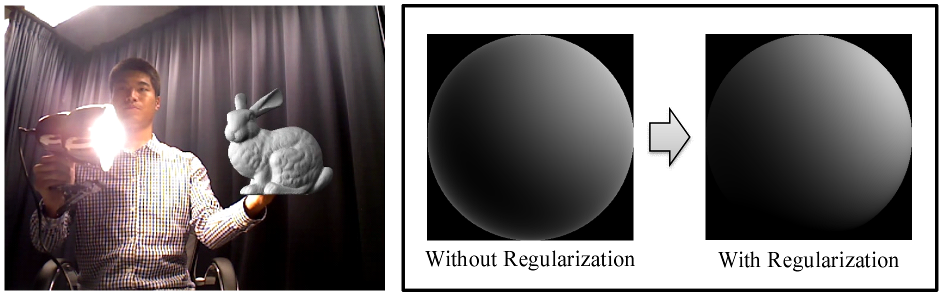

Figure 1: *The system calculates shading matched to the scene from a hand automatically in real time (left). The ringing artifact seen at the bottom left boundary, which typically appears when there are not enough sample points, disappeared by imposing regularization (right).*

1 Introduction

To render computer graphics (CG) objects realistically in the real world, it is important to match their shadings to the scene. Existing methods to achieve such realism are not applicable to realtime consumer augmented reality (AR). Debevec's method requires a mirrored sphere, which ordinary consumers do not have [Debevec 1998]. Karsh's method requires user annotation, which makes it difficult to use in a realtime application [Karsch et al. 2011].

Here, we propose a shading derivation method utilizing the human hand as a light probe. We chose to use the hand since hands are readily available if people are present. The system we implemented the method into automatically detects the hand, calculates shading in realtime, and is easy to use. The method is suitable for consumer applications since it can be implemented into a system merely by using a commercial RGB-D sensor.

2 Our Method

The method is based on regularized spherical harmonic (SH) regression that is modified from [Yao et al. 2012] to enable it to work in realtime. For every point of a hand, we take the luminance value L_i (i is an index) and surface normal (θ_i, ϕ_i), and calculate SH weights w_l^m, $(0 \leq l \leq 2)$ as follows.

$$\underset{w_l^m,\ 0 \leq l \leq 2}{\arg\min} \{e(w_l^m) + \lambda r(w_l^m)\}$$

$$e(w_l^m) = \sum_i \left(L_i - \sum_{l=0}^{2} \sum_{m=-l}^{l} w_l^m Y_l^m (\theta_i, \phi_i) \right)^2 \qquad (1)$$

$$r(w_l^m) = (w_0^0)^2 + \frac{9}{4} \sum_{m=-1}^{1} (w_1^m)^2 + 16 \sum_{m=-2}^{2} (w_2^m)^2$$

,where $e(w_l^m)$ is an error term and $\lambda r(w_l^m)$ is a regularization term. The term $r(w_l^m)$ is designed so that w_0^0, w_1^m, and w_2^m are equally penalized when their magnitudes take the ratio of $\hat{A}_l$ as rel. (2), where $\hat{A}_l$ is described in [Ramamoorthi and Hanrahan 2001].

$$|w_0^0| : |w_1^m| : |w_2^m| = \hat{A}_0 : \hat{A}_1 : \hat{A}_2 = 1 : \frac{2}{3} : \frac{1}{4} \qquad (2)$$

*e-mail: yao.yasuhiro@lab.ntt.co.jp
†e-mail: kawamura.harumi@lab.ntt.co.jp
‡e-mail: kojima.akira@lab.ntt.co.jp

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07

According to [Ramamoorthi and Hanrahan 2001], the ratio of SH weight magnitudes should be around the ratio of $\hat{A}_l$ when they model irradiance or diffuse shading. The regularization is imposed so that the weights will satisfy this condition. The regularization weakens ringing artifacts that typically appear when there are not enough sample points (see Fig.1 (right)).

Since the minimization operation shown as (1) above can calculate SH weights in a single step, our method is suitable for realtime applications. This is different from the method reported in [Yao et al. 2012] which requires iterative calculation.

3 Implementation

We developed a system that combined our method with hand tracking by using an RGB-D sensor, ASUS Xtion Pro Live. The system tracks a hand and derives its area by using hue and depth differences. Then, from the luminance values and surface normals of the points in the area, it uses our method to calculate shading in realtime and then uses the shading to render CG objects (see Fig.1 (left)). Since this system also utilizes the hand as a geometrical probe, it can place CG objects at the position where the hand was with shading matched to the position. To use this system, users only need to place a hand in front of an RGB-D sensor. This ease of use creates possibilities for using our method in consumer AR applications involving persons such as virtual fitting.

References

DEBEVEC, P. 1998. Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH '98*, ACM, New York, NY, USA, 189–198.

KARSCH, K., HEDAU, V., FORSYTH, D., AND HOIEM, D. 2011. Rendering synthetic objects into legacy photographs. *ACM Trans. Graph. 30*, 6 (Dec), 157:1–157:12.

RAMAMOORTHI, R., AND HANRAHAN, P. 2001. An efficient representation for irradiance environment maps. In *SIGGRAPH '01*, ACM, New York, NY, USA, 497–500.

YAO, Y., KAWAMURA, H., AND KOJIMA, A. 2012. Shading derivation from an unspecified object for augmented reality. In *ICPR '12*, IAPR, 57–60.