2012 IEEE 24th International Conference on Tools with Artificial Intelligence

Texture and Shadow Insensitive Metric for Image-based Reconstruction

Rui Wu, Xu Zhao, Zhong Zhou*, Wei Wu
State Key Laboratory of Virtual Reality Technology and Systems
School of Computer Science and Engineering, Beihang University, Beijing, China
Email: {wurui, zhaoxu, zz, wuwei}@vrlab.buaa.edu.cn

***Abstract*—This paper proposes an accurate metric for image-based 3d reconstruction without ground truth. Specially, our metric is insensitive to texture changing and shadows, which are commonly occurred in real world scenes. Based on the inter-reflected rendering model, we improve the accuracy of previous irradiance-based metric. Additionally, we estimate the reflectance of each vertex on the surface to support the case with varying reflectance. We also consider the difference between estimated and observed irradiance in our metric to further eliminate the boundary effect of texture changing or self-shadow. Experiments on both indoor and outdoor datasets illustrate the effectiveness of our metric. Our evaluation results are not only more accurate than the results of previous metrics, but also insensitive to the texture and shadow.**

I. Introduction

Automatic and accurate reconstruction of 3D objects and environments from multiple images plays a significant role in many applications. Multi-view stereo (MVS) is one of the most attractive image-based reconstruction methods. Over the last few years, a number of high-quality MVS algorithms have been developed, and the state of the art is improving rapidly, especially after the benchmark datasets were proposed [1], [2]. 3D reconstructions can be evaluated by geometric comparison with ground truth models acquired from the 3D laser scanner. However, in practice, it is impractical for many indoor/outdoor scene modeling applications to access such ground truth and measure their recovered models.

On the other hand, for accurate 3D reconstruction, MVS methods usually need a moderate number of images as input. It is more and more convenient to obtain a sufficient number of images, like from webcam, personal mobile device, video, Internet and so on. However, it is hard to determine how many images are needed and where they are located. As noted by Hornung et al. [3], more input images do not definitely lead to better recovered accuracy, sometimes even may reduce the reconstruction quality. All above issues require a new metric for image-based 3D reconstruction, which can not only evaluate the reconstruction results without any ground truth datasets, but also guide the reconstruction process to add proper new images to improve the accuracy.

Our work is based on the previous work proposed by Zhao et al. [4]. They presented an irradiance-based metric to measure the image-based 3D reconstruction results. Their metric can automatically indicate the regions which are hard to recover and it is independent of any reconstruction algorithm as well. Their basic idea is based on the assumption that the irradiance on the object's surface change continuously. Therefore they can use the irradiance gradient to recognize the regions where irradiance changes greatly and treat them as reconstruction errors. Nevertheless, their assumption only works well for convex objects with constant reflectance. Except for reconstruction errors, other factors like texture changing or self-shadow caused by self-occlusion could also result in the discontinuous irradiance. These phenomenons are commonly occurred in real world scenes but neglected by Zhao's method, which will badly affect their metric accuracy.

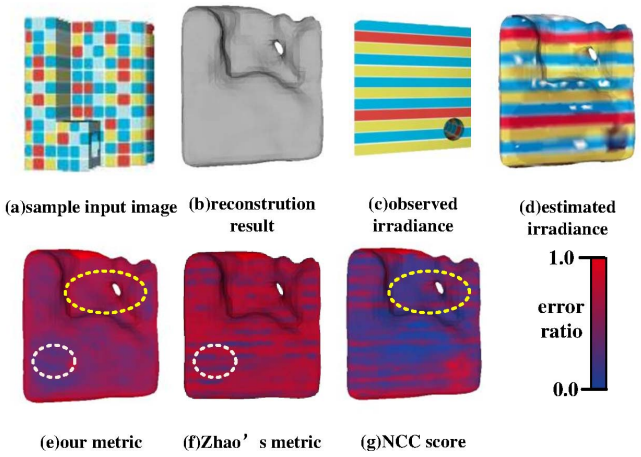

Fig. 1. 3D reconstruction evaluation by different metrics on synthetic data. 12 images (a) are inputted to reconstruct 3D proxy model (b). Combined with shading, we estimate irradiance around the surface (d) to compute our texture and shadow insensitive metric (e). Compared with the traditional photo-consistency metric (g), our metric can indicate the wrongly reconstructed regions. Compared with previous irradiance-based metric (f), our metric avoid influence of texture changing. For illustration, red means high error ratio.

As illustrated in Fig. 1, both our and Zhao's metric can correctly indicate the poorly reconstructed region marked by the yellow circle where NCC score cannot. However, Zhao's metric is seriously affected by texture changing. The corresponding evaluation result (Fig. 1(f)) shows the unreasonable stripe pattern at the boundary of different textures, for example, boundary areas marked by the white circle.

To address the above problem, we come up with a more accurate metric for image-based 3D reconstruction. The metric

*Corresponding author. Email:zz@vrlab.buaa.edu.cn

1082-3409/12 $26.00 © 2012 IEEE
DOI 10.1109/ICTAI.2012.140

computation is similar to the previous irradiance gradient, but we use the inter-reflected rendering model instead and encapsulate the radiance self-transfer function to improve the accuracy of irradiance estimation. It also enables us to indicate the 3D reconstruction geometric errors from multiple images under arbitrary unknown illumination. Considering the scene with varying reflectance, we compute the reflectance of each vertex to better estimate the surface irradiance in the Spherical Harmonics domain. In addition, we consider the difference between estimated and observed irradiance in our metric to avoid the boundary effect of texture changing and self-shadow.

The contributions of our paper include:

(1) We take the self-casting shadows and inter-reflection into consideration, thus our approximation of the lighting and irradiance is more close to the reality and further leads to a more accurate metric than previous ones.

(2) We estimate the reflectance of each vertex to gain an irradiance field compatible with input images. By this way our metric is not constrained to Lambertian objects only, but also works well on diffuse objects with varying reflectance.

(3) We consider the difference between estimated and observed irradiance during the metric calculation. Therefore, our anisotropic irradiance gradient metric can provide robust performance to the boundary of texture and shadow.

II. Related Work

MVS intends to reconstruct a complete 3D object model from a collection of images taken from known camera viewpoints. On one hand, an accurate and refined reconstruction model needs sufficient input images; on the other hand, overmuch images which contains redundant details would take too long time to handle; even worse, sometimes more input images doesn't lead to more accurate results. As a result, an evalution of the recovered model is of significant importance to balance the accuracy against efficiency.

Laser scan can output realistic 3D models with scarce errors, thus its output can be regarded as the ground truth. To compare MVS reconstruction result with the ground truth is a intuitive idea to measure it. Seitz et al. [1] propose a evaluation MVS methodology. They capture and calibrate multi-view image datasets with high-accuray gound truth. While Seitz et al. present indoor datasets acquired under controlled lighting condition, Strecha et al. [2] propose a collection of high-resolution images and LIDAR data of outdoor expansive scene. They calculate the relative error occurance for the reference view and the error distribution based on the ground truth. Researchers can submit their reconstruction results online [5], [6] in order to evaluate their recovered model.

Researches study factors that influence the recovered model and present several empirical metrics. The effect properties on image-based reconstruction process can be divided into three categories. (1) *Image quality*. Hornung et al. [3] point out that the number of input images increases to perform a robust reconstruction in case of the projection error, illumination variation and the blurred or noisy images. (2) *Relationship between images*. Goesele et al. [7] compute a depth map and merge the result to form the model. They choose neighbor views between which angle is small enough to provide abundant shared feature points, as well as enough parallax to keep up the reconstruction accuracy. (3) *Relationship between image and model*. Dunn et al. [8] pointed out that input images with higher sampling rate will bring about more accurate recovered model. They seek to find next best view that has small view angle of the focused area.

In the field of view planning, researchers try to select the view or image sets that have the greatest contribution to the accuracy of the 3D model. Photo-consistency, the coherence of pixel intensities in different views, is often used [3]. Besides, in their key frame extraction algorithm, Ahmed et al. [9] compute the root mean reprojection error for each frame then compute the min, max, mean, and standard deviation statistics over the entire sequence. And Dunn et al. [8] estimate the structure estimation uncertainty and select views that best reduce it. The key idea of metrics used in view planning applications is to attempt a optimal solution to achieve the optimization target, such as maximize photo-consistency or minimizing reprojection error. Unfortunately, the target is set on the basis of reconstruction process; thus these metrics are coupled with some certain MVS methods and can't be applied to a wide range.

Zhao et al. [4] propose a metric to measure the image-based 3D reconstruction results without ground truth. Considering the uncertainty of topology, texture and soft shadow, they formulate a anisotropic irradiance gradient field to indicate the regions where reconstruction error occurs. What's more, their method is decoupled of the MVS method. Experiments on both synthetic data and real world images have validate the effectness of their metric.

Our metric is an extension of irradiance gradient metric, while our technique differs from it in two aspects: combination of shape from shading (SfS) and estimation of irradiance.

SfS methods use shading cues to compute per-pixel surface orientation instead of sparse depth to estimate shape from a single image. SfS has the complementary nature compared with MVS, and much work has been done on combining these two techniques. Samaras et al. [10] iteratively estimate both shape and illumination given multiple views taken under fixed illumination. Yu et al. [11] propose two algorithms for modeling non-Lambertian objects illuminated by distant light sources. Both explicitly model the reflectance using either a Phong or Torrence-Sparrow model, and then optimize the estimated shape and reflectance. Different from the above, the irradiance-based metric intends to measure the reconstruction result rather than estimate the shape. In the work of Zhao et al., they use the gradient of shading values, the estimated irradiance, to point out the changes of topology of object. However, we consider the fact shading tends to be more accurate for higher frequency shape components, and use gradient of the difference between estimated and observed irradiance instead.

Basri et al. [12] and Ramamoorthi et al. [13] introduce a theoretical result that reflection from a curved surface, lit by

an environment map, could be seen as a spherical convolution of the incident illumination and the reflective properties or BRDF of the surface. The convolution theory also implies a product formula in the frequency domain, in terms of spherical harmonic basis coefficients. Sloan et al. [14] introduce the term precomputed radiance transfer. In spherical harmonics domain, they precompute the shadow and inter-reflection components of the object to compute irradiance on vertices and reduce the convolution to a dot product. Except for expressing lighting in the spherical harmonics domain, wavelet [15] is also employed to gain a representation in higher frequency environment. Precomputation-based rendering method focused on a real-time irradiance computation under given illumination; while the irradiance-based metric method attempts to estimate the irradiance under the guidance of images of the scene. Different from rendering model Zhao et al. applied, we consider inter-reflection and estimate corresponding outgoing irradiance.

III. The Irradiance-based Metric

Image-based modeling can be seen as the inverse process of rendering. As for a 3D Lambertian model, if given fixed incident lighting, the reflected light intensity, the irradiance r, is determined by normal of the surface and supposed to be continuous. Under the assumption irradiance r changes continuously on the object's real surface, if error occurs, gradient of irradiance becomes significant, as in Fig. 2.

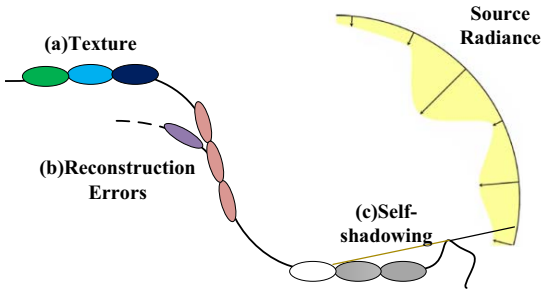

Fig. 2. The basic idea of our metric for measuring 3D reconstruction. Assume the irradiance on the surface changes continuously, the reconstruction errors (b) occur when irradiance changes greatly. However, continuity of irradiance may also be broken by texture changing (a) and self-shadowing (c).

Nevertheless, the assumption above proposed by Zhao et al. only works well for convex objects with constant reflectance. Except for reconstruction errors, other factors could also result in the discontinuous irradiance. Greger et al. [16] point out that due to occlusions, irradiance on surfaces become discontinuous in transition zones between transition zones between umbra and penumbra, and between penumbra and unshadowed regions. And on the boundary of changed texture, discontinuity occurs too. What's more, in the analytic methods, Arvo [17] carefully consider inter-reflection because number of bounces may be large due to occlusion and local geometric complexity. But as Zhao et al. presented, inter-reflections and self-shadows are neglected.

Our metric eliminates the influence of factors other than the geometrical errors and provide a texture and shadow insensitive metric following the inter-reflected rendering model of concave objects. We estimate the reflectance of the surface to gain a compatible estimation of irradiance. Also based on the shading technique, we calculate the difference between the estimated and observed irradiance and compute its gradient instead of gradient of estimated irradiance.

IV. The Metric Calculation

Our metric is intended to indicate the regions of the reconstruction results which challenge the establishment of MVS correspondence; what is more, the metric should be insensitive to the texture and shadow of the object, as to correctly point out the areas and provide robust performance. The illumination is assumed to be fixed and distant, but is otherwise general and known. We represent the illumination using Spherical Harmonics (SH). To keep the problem tractable, we henceforth assume the objects to be diffuse. We take self-shadow and inter-reflections into account so as to get an accurate metric.

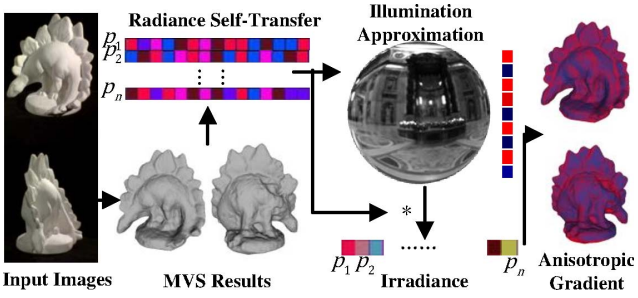

Fig. 3. The outline of our metric calculation pipeline. Multiple images are inputted to generate the initial 3D triangle mesh model by any multi-view stereo reconstruction algorithms. Then we encapsulate the shadow and inter-reflection by radiance self-transfer function, which enable us to accurately estimate the incident illumination in the Spherical Harmonics domain. We further calculate the irradiance of each vertex considering the cases of varying texture and shadow and finally get our anisotropic irradiance gradient metric.

The workflow of our method is shown in Fig. 3. The following subsections explain the computation of the gradient field in detail. As we will describe, we handle the changes in texture by attributing them into the variations of the albedos, and estimating the albedos with the input images. And we get rid of the shadows effect as we compare the gradient difference in the observed irradiance and the estimated irradiance.

A. The Rendering Equation

Assuming all objects in the scene are non-emitters, the image rendering equation [18], [14] can be defined as equation below

$$B(x,\omega_o) = B_{DS}(x,\omega_o) + B_{DI}(x,\omega_o), \tag{1}$$

where B_{DS} is the outgoing irradiance due to the illumination from the environment and B_{DI} is the outgoing irradiance because of inter-reflection of the object, namely

$$\begin{aligned} B_{DS} &= \int_{\Omega} \rho(x,\omega_i,\omega_o) L(x,\omega_i) V(x,\omega_i) H(x,\omega_i) d\omega_i, \\ B_{DI} &= \int_{\Omega} \rho(x,\omega_i,\omega_o)\, \bar{L}(x,\omega_i)(1 - V(x,\omega_i)) H(x,\omega_i) d\omega_i \end{aligned} \tag{2}$$

where $B(x, \omega_o)$ is the irradiance at vertex x to the direction w_o and the symbol w_i is the incident angle. The symbol Ω represents the domain of all possible directions. ρ is the bidirectional reflectance distribution function(BRDF) of the surface. $L(x, \omega_i)$is the incident radiance at vertex x, and $\bar{L}(x, \omega_i)$ is the radiance from the object itself in the direction w_i. $V(x, \omega_i)$is a visibility function. $H(x, \omega_i)$ is the clamped cosine kernel, which $H(x, \omega_i) = \max(\omega_i \cdot \mathbf{n}(x), 0)$, and $\mathbf{n}$ is the surface normal.

We assume that lighting is distant, so incident radiance is the same over the object surface and we have $L(\omega_i)$ for short. And the object is diffuse, so the same amount of light is reflected into every outgoing direction, and the BRDF can be described as the albedo $\rho(x)$. Then for each outgoing direction we have

$$\begin{aligned} B(x) &= (\rho(x)/\pi) \int_{\Omega} L(\omega_i) V(x, \omega_i) H(x, \omega_i) d\omega_i \\ &+ (\rho(x)/\pi) \int_{\Omega} \bar{L}(x, \omega_i)(1 - V(x, \omega_i)) H(x, \omega_i) d\omega_i, \end{aligned} \tag{3}$$

Different from the rendering equation Zhao et al. applied, we add the term B_{DI} to represent the outgoing irradiance due to inter-reflection. Besides, the irradiance gradient method treats albedo $\rho(x)$ as a constant, and simply scales the incident lighting $L(\omega_i)$ by the albedo to form a scaled illumination. Further, in their calculation process, the scaled illumination is considered as the incident lighting.

B. Radiance Self-Transfer

An initial mesh model of the object is reconstructed with the help of MVS. Based on this model, our method calculates the radiance self-transfer. The object's response to its environment can be viewed as transfer function, mapping incoming to outgoing radiance. By representing both incident radiance and transfer functions in the SH domain, the light transport can be reduced to a simple dot product between their coefficients

$$B(x) = \sum_{k=1}^{n^2} M_k(x) L_k, \tag{4}$$

where $n-1$ is the order of the SH, $M_k(x)$ is the transfer vector, of which each component represents the linear influence that a lighting basis function L_k has on shading at vertex x. In this case, the inter-reflected diffuse transfer process should be divided into two pass. The first pass is to calculate the shadow diffuse transfer, which $M_1(x) = V(x, \omega_i) H(x, \omega_i)$. The second pass is to calculate the inter-reflection transfer based on the first pass, which $M_2(x) = (1 - V(x, \omega_i)) H(x, \omega_i)$.

C. Irradiance Estimation

Given the radiance transfer coefficients, to estimate the incident illumination coefficients, we first need to approximate the albedos, $\rho(x)$, of the surface. We ascribe changes in texture to albedos, and estimate them from images. By taking albedos into account, the estimated irradiance matches with the observed, thus our metric wipes off the effect of texture.

Assume $Q(x_i)$ is the set of cameras that can see the i-th vertex x_i, P_c is the projection matrix for camera c, and

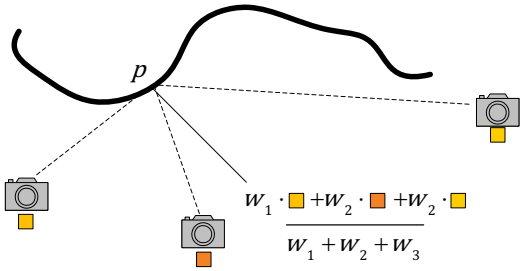

Fig. 4. The computation of reflectance of each vertex on the surface. Different viewpoints are used to estimate the reflectance. The viewpoint which directly faces the vertex will contribute more through a larger weight.

$I_C(P_C(x_i))$ represents the image intensity corresponding to x_i and captured by camera c. Then as in Fig. 4. the estimated albedo of xi is respectively

$$\rho(x_i) = \sum_{c \in Q(x_i)} w_c I_c(P_c(x_i)) / \sum_{c \in Q(x_i)} w_c \tag{5}$$

in which

$$w_c = \frac{vec(x_i, cam_c) \cdot \mathbf{n}(x_i)}{\|vec(x_i, cam_c)\| \cdot \|\mathbf{n}(x_i)\|}, \tag{6}$$

where cam_c is the spatial position of the c-th camera, $vec(x_i, cam_c)$ is the vector from the vertex to the camera, and $\|\|$ means the l_2-norm of a vector. The weight term makes sure that the image taken orthogonally with the surface normal contributes little to the estimation of the albedo.

Based on the transfer coefficients and albedo of each vertex, we estimate the lighting coefficients by minimizing

$$\hat{l} = \arg\min_{l} \sum_{i} \sum_{c \in Q(i)} \left| \sum_{k=1}^{n^2} M_k(x_i) l_k - I(P_c(x_i)) \right| \tag{7}$$

After we've got the current estimated geometry and illumination, the next step is to relight the object, that is, to calculate the outgoing irradiance according to Eq. 3.

Prior to the original metric, after the accurate calculation of transfer function and estimation of albedos, our irradiance field matches with the ground truth. This way, we can ascribe the cause of significant difference of observed and estimated irradiance can be think as the geometrical errors, but not the insufficient precision of our estimation process.

D. Anisotropic Gradient Field Calculation

Shape from shading methods compute per-pixel surface orientation instead of sparse depth, thus they can recover fine details from shading variations. So the gradients of the shading values, the estimated irradiance, accurately point out the changes of topology of object. And because shading tends to be more accurate for higher frequency shape components, we compare the gradients of the observed and estimated irradiance instead of directly comparing irradiance values. What's more, in this way, our metric is insensitive to self-cast shadows.

For each vertex, the gradient is defined as

$$G(x_i) = \sum_{x_j \in N(x_i)} \sum_{c \in Q(x_i, x_j)} w(x_i, x_j) \cdot (d(x_i) - d(x_j))^2 \quad (8)$$

where i and j are the indices of the vertex, $N(x_i)$ is the set of the neighbor of the vertex x_i. The weight function is negative correlation to the distance between x_i and x_j and defined as $w(x_i, x_j) = radius/dis(x_i, x_j)$. $d(x_i)$ is the difference between the observed and estimated irradiance, namely

$$d(x_i) = B(x_i) - I(P_c(x_i)) \quad (9)$$

V. Experiments

We validate our algorithm using a synthetic model, shown in Fig. 1, and two real world datasets: a plaster dinosaur (Fig. 5) and a hall(Fig. 6). The synthetic model is a machine component with gridded color. For the synthetic model, we use 4 simulated point light-sources. The images of plaster dinosaur, one of the Middlebury datasets, are taken under controlled studio environment. The plaster dinosaur has a slightly varying albedo and casts shadows onto itself. The other real world image sets, the hall, is captured outdoor, under general yet uncontrolled illumination. We use Furukawa's method [19] to generate the initial patch models. Then to produce the mesh model, we apply the Poisson surface [20] for the synthetic object and the dinosaur, and touch expand method [21]for the hall. We estimate the lighting and irradiance in SH domain and set the SH order between 3 to 6 as a compromise of efficiency against accuracy. We use Direct X to calculate the transfer function of each vertex. For lighting estimation, we use Lebvenberg-Marquardt method to solve the minimization problem in Eq. 7.

A great number of MVS measures try to evaluate the visual compatibility of a reconstruction with the input images. While most of the measures work based on the pixel intensity or the color of the image and establish correspondences between different views. This kind of measures are often called photo-consistency metric. In our experiment, we also compare the performance of our metric with a basic photo-consistency measure method, the the Normalized Cross-Correlation (NCC).

Besides, we conduct experiments to prove the advantages of our metric over the metric proposed by Zhao et al.. In the experiment, we run comparison on 3D model with texture and self-shadow and evaluate the performance of both metrics.

A. Effectiveness Validation

We evaluate the effectiveness of our metric both on synthetic and indoor/outdoor real world scenes.

The synthetic dataset was generated by rendering 12 images at pixel resolution 390×390. Fig. 1(a) displays an image as an example. We set the order of SH to 3. By comparison, we can see the hole area is poorly recovered. As in Fig. 1(d), the irradiance model matches with the synthetic one, which reveals that our lighting and irradiance estimation method gives convincing and accurate result. Fig. 1(e) shows our anisotropic gradient field correctly indicates that there is reconstruction errors in the hole region.

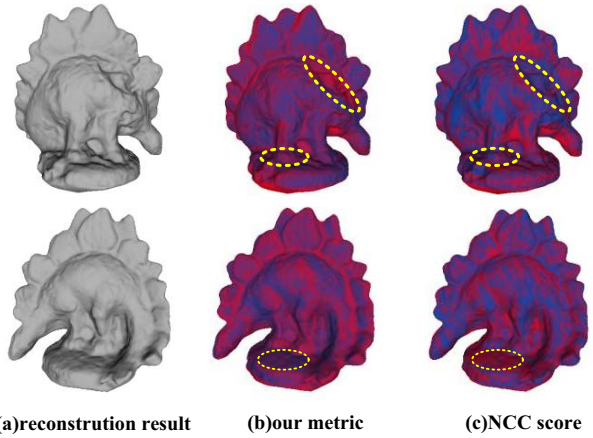

Fig. 5. Evaluation by our metric on real-world indoor dino dataset. The hollow area on the back and the toe area with lots of details are accurately indicated by our metric (red means high error ratio).

The Middlebury "DinoSparseRing" dataset consists of 16 views sampled on a ring at pixel resolution 640×480. In this experiment, we set the SH order to 4. Our metric sheds light on reconstruction errors, hollow area on the back, and topologically challenging areas, toe area with occlusion around, as in Fig. 5(b).

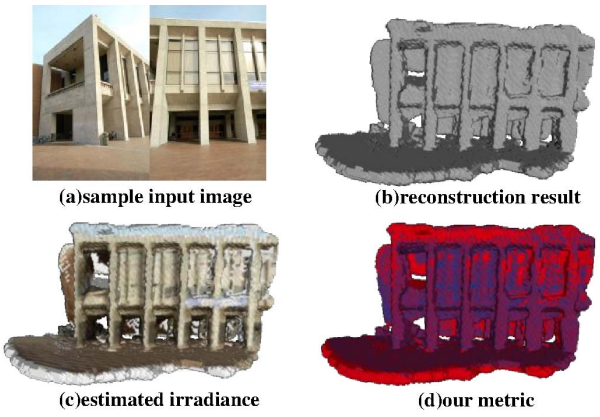

Fig. 6. Evaluation by our metric on real-world outdoor hall dataset. Our metric can successfully be applied into the outdoor image-based reconstruction applications. The poorly reconstructed areas like floor and right window are automatically pointed out by our metric.

The hall dataset contains 61 uncalibrated images at the pixel resolution of 3008×2000 captured under uncontrolled outdoor illumination, displayed in Fig. 6. As for the images captured under uncontrolled illumination, we successfully perform irradiance estimation (Fig. 6(c)) and further anisotropic gradient calculation (Fig. 6(d)). As shown in the reconstructed model, the windows of the upper right is poorly recovered; our metric point our the area automatically.

B. Comparison with NCC Score

Our metric can achieve more accurate and robust performance compared to the NCC as our metric combines the photo-consistency with the shading method, and provides details about the geometry directly. For example, the base of the plaster dinosaur is purely a plane. But the NCC score of this area raises, due to lacking of intensity variation and some shadows casted on it. But our method achieved convincing consequence as it point out that this area is of low gradient. Besides, the back of the model is of repeated texture, so the multi-view method outputs inaccurate matches and treats them as consistent in color. As shown in the top left of Fig. 1(f), NCC score of hole area is high, but with mistakes.

C. Comparison with Zhao's Metric

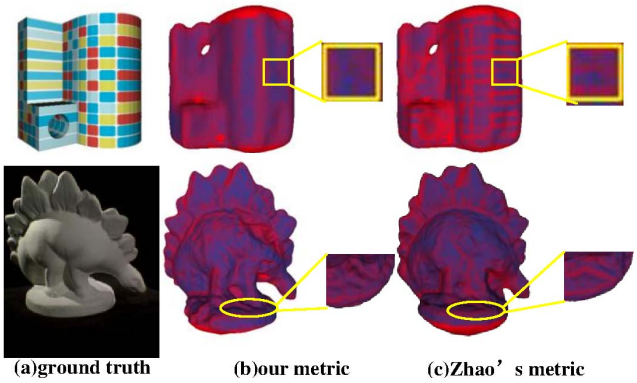

Fig. 7. Comparison our metric with Zhao's metric on the synthetic dataset to show the texture and shadow insensitive characteristic.

Experiments were conducted to test and verify the improvement gained by our metric compared with the previous irradiance gradient metric. The advantages of our metric lie in the robustness against texture changing and self-shadow.

As shown in Fig. 7, boundaries of different colors on a smooth surface was not indicated as a geometrically challenge area with high gradient. And on the base of the plaster dinosaur, a boarder of the umbra and penumbra is displayed in Fig. 7(b) and our method shows insusceptible performance. While, as shown in Fig. 7(c) the metric provided by Zhao et al. [4] fails to give the right result.

VI. Conclusion

This paper presents a texture and shadow insensitive metric for image-based 3D reconstruction. The metric is based on the anisotropic irradiance gradient which can automatically evaluate 3D reconstructions without ground truth. The experimental results demonstrate that we further improve the metric accuracy by encapsulating the inter-reflection rendering model into the radiance transfer function. We also explicitly estimate the reflectance of each vertex on the surface and consider the difference between the estimated and observed irradiance to make our metric insensitive to the texture and shadow.

Although our method is devoted to extend the application range of previous irradiance-based metric, one limitation still comes from the assumption of the diffuse material. In future, we would like to further extend our method to be applied to more general materials. Besides, the Spherical Harmonics act as a low-pass filter on the incident illumination. We will try to employ wavelet as a representation of lighting instead.

VII. Acknowledgment

This work is supported by the National 863 Program of China under Grant No.2012AA011801, the Natural Science Foundation of China under Grant No.61170188, and the National 973 Program of China under Grant No. 2009CB320805.

References

[1] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *CVPR*, Washington, DC, USA, 2006, pp. 519–528.
[2] C. Strecha, W. von Hansen, L. Van Gool, P. Fua, and U. Thoennessen, "On benchmarking camera calibration and multi-view stereo for high resolution imagery," in *CVPR*, Anchorage, AK, 2008, pp. 1–8.
[3] A. Hornung, B. Zeng, and L. Kobbelt, "Image selection for improved multi-view stereo," in *CVPR*, ser. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008, pp. 1–8.
[4] X. Zhao, R. Wu, Z. Zhou, and W. Wu, "A new metric for measuring image-based 3d reconstruction," in *ICPR(to appear)*, Tsukuba Science City, JAPAN, 2012.
[5] S. M. Seit and B. C. et al. (2006) Middlebury multi-view stereo evaluation. [Online]. Available: http://vision.middlebury.edu/mview/
[6] C. Strecha and W. von Hansen et al. (2008) Epfl passive stereo evaluation. [Online]. Available: http://cvlab.epfl.ch/ strecha/multiview/denseMVS.html
[7] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, "Multi-view stereo for community photo collections," in *ICCV*, Brazil, 2007.
[8] E. Dunn and J. M. Frahm, "Next best view planning for active model improvement," in *BMVC*, London, UK, 2009, pp. 1–11.
[9] M. T. Ahmed, M. N. Dailey, J. L. Landabaso, and N. Herrero, "Robust key frame extraction for 3d reconstruction from video streams," in *VISAPP*, Angers, France, 2010, pp. 231–236.
[10] D. Samaras, D. Metaxas, P. Ascalfua, and Y. G. Leclerc, "Variable albedo surface reconstruction from stereo and shape from shading," in *CVPR*, Hilton Head, SC, USA, 2000, pp. 480–487.
[11] T. Yu, N. Xu, and N. Ahuja, "Object-centered surface reconstruction: Combining multi-image stereo and shading," *IJCV*, vol. 73, no. 2.
[12] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *PAMI*, vol. 25, no. 2, pp. 218–233, 2003.
[13] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance determining the illumination from images of a convex lambertian object," *Journal of The Optical Society of America*, vol. 18, no. 10, pp. 2448–2459, 2001.
[14] P. Sloan, J. Kautz, and J. Snyder, "Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments," *ACM SIGGRAPH 2002*, vol. 21, no. 3, pp. 527–536, 2002.
[15] R. Ng, R. Ramamoorthi, and P. Hanrahan, "All-frequency shadows using non-linear wavelet lighting approximation," *ACM SIGGRAPH 2003*, vol. 22, no. 3, pp. 376–381, 2003.
[16] G. Greger, P. Shirley, P. M. Hubbard, and D. P. Greenberg, "The irradiance volume," *IEEE Computer Graphics and Applications*, vol. 18, no. 2, pp. 32–43, 1998.
[17] J. R. Arvo, "Analytic methods for simulated light transport," Ph.D. dissertation, Yale University, 1995.
[18] J. T. Kajiya, "The rendering equation," in *Computer Graphics*, New York, USA, 1986, pp. 143–150.
[19] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multi-view stereopsis," *PAMI*, vol. 32, no. 8, pp. 1362–1376, 2010.
[20] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Symposium on Geometry Processing*, Switzerland, pp. 61–70.
[21] V. Lempitsky and Y. Boykov, "Global optimization for shape fitting," in *CVPR*, Minnesota, USA, 2007.