Interactive HDR Environment Map Capturing on Mobile Devices

Peter Kán

Institute of Software Technology and Interactive Systems, Vienna University of Technology, Vienna, Austria

Abstract

Real world illumination, captured by digitizing devices, is beneficial to solve many problems in computer graphics. Therefore, practical methods for capturing this illumination are of high interest. In this paper, we present a novel method for capturing environmental illumination by a mobile device. Our method is highly practical as it requires only a consumer mobile phone and the result can be instantly used for rendering or material estimation. We capture the real light in high dynamic range (HDR) to preserve its high contrast. Our method utilizes the moving camera of a mobile phone in auto-exposure mode to reconstruct HDR values. The projection of the image to the spherical environment map is based on the orientation of the mobile device. Both HDR reconstruction and projection run on the mobile GPU to enable interactivity. Moreover, an additional image alignment step is performed. Our results show that the presented method faithfully captures the real environment and that the rendering with our reconstructed environment maps achieves high quality, comparable to reality.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Color, shading, shadowing, and texture I.3.3 [Computer Graphics]: Picture/Image Generation—Digitizing and scanning

1. Introduction

Estimation of real illumination plays an important role in many areas of computer graphics, including image-based lighting [Deb98], material reconstruction or augmented reality [GEM07]. In order to capture and store the illumination of the real world, environment maps are commonly used. Therefore, environment map capturing is essential for applications which utilize the real illumination. However, the previous methods for environment map capturing suffer from the requirement of additional hardware, complex setup and necessary post-processing. Additionally, approaches for capturing panoramas on mobile devices were presented [WMLS10, XP10]. However, they do not produce results in high dynamic range (HDR) which is necessary for image-based lighting and material reconstruction. Moreover, many of these methods do not cover the full sphere of light.

This paper presents a novel method for environment map capturing which utilizes a consumer mobile device to capture HDR environment map covering a full sphere. Besides the mobile device, no additional hardware is required. Our method does not need post-processing as all the data are accumulated and converted interactively to radiance. The captured environment map can be directly used for image-based lighting even during the capturing process. In order to properly project the camera image to the environment map, we utilize the inertial measurement unit (IMU) of the mobile device which includes gyroscope, accelerometer and magnetometer. The camera image is projected to the environment map according to the orientation of the device, estimated by the IMU. Geometric and radiometric calibration is performed to accurately project the captured data. The main advantages of our method are its practical use, HDR reconstruction and interactive performance.

High dynamic range of environment maps is of paramount importance because the environmental light contains a broad range of intensities which are not possible to be stored in low dynamic range (LDR) image. In our approach, the HDR data is reconstructed from camera image with automatically varying exposure and known exposure time. First, we estimate the inverse camera response curves [RBS99] of specific camera in the calibration procedure to map the measured sensor response onto HDR radiance in a correct way. The reconstructed HDR radiance values are accumulated into spherical environment map. Therefore, if more images with different exposures are projected onto their overlapping area, the information from all overlapping images is stored. Both projection of a camera image to the environment map and HDR reconstruction run on the mobile GPU to enable interactive performance.

The limited accuracy of built-in mobile sensors causes problems with the alignment of projected camera images. We overcome these problems by using alignment of camera image with environment map based on feature matching. Our approach projects the captured data from the environment map back to the camera frame. Then, the features from this projection are matched to the features from the new camera image. The rotational difference is calculated and the camera orientation is corrected to project the new image with correct alignment. An environment map, captured by our method, can be seen in Figure 1. The main contributions of this paper are:

- A novel method for HDR environment map capturing by a mobile device.
- An accurate radiance estimation and accumulation from moving camera with automatic exposure.
- A method for interactive alignment of new camera data with the previously captured map.

2. Related Work

Environment maps have been widely used in computer graphics for rendering with natural illumination. Therefore, numerous techniques for their capturing have been presented. An environment map can be captured by taking a picture of reflective sphere [Deb98]. Multiple pictures can be combined to obtain full 360 degrees of incoming light. Another approach is to use an additional camera with a fish eye lens [SJW*06, SSI99, GEM07]. Multiple images with different exposures are taken by this camera to build a HDR environment map. The limitations of these techniques are the requirements of additional hardware and post-processing. In contrast to that, our method only needs a consumer mobile phone and it does not require an additional post-processing.

HDR reconstruction is necessary for environment maps to capture the high contrast of real light. Several methods were presented to reconstruct HDR radiance from multiple images [DM97, RBS99]. These methods calculate the HDR radiance as a weighted sum of input images captured by the static camera with varying exposure times. Additionally, they reconstruct the camera response function to model the nonlinearities of the sensor response. The response function is taken into account in HDR reconstruction.

An important related research has been done in the field of panorama capturing. Different panorama representations have been presented including cylindrical [WMLS10], spherical [DWWW08] and others [SS97]. An advantage of panorama capturing is that it can be easily implemented on mobile devices. Similar to our method in terms of orientation tracking is the work of DiVerdi et al. [DWH09]. The previous methods for panorama capturing suffer from the limitation of low dynamic range data and therefore are not suitable for light capturing. Our method reconstructs the environment maps in HDR to accurately capture the real light.

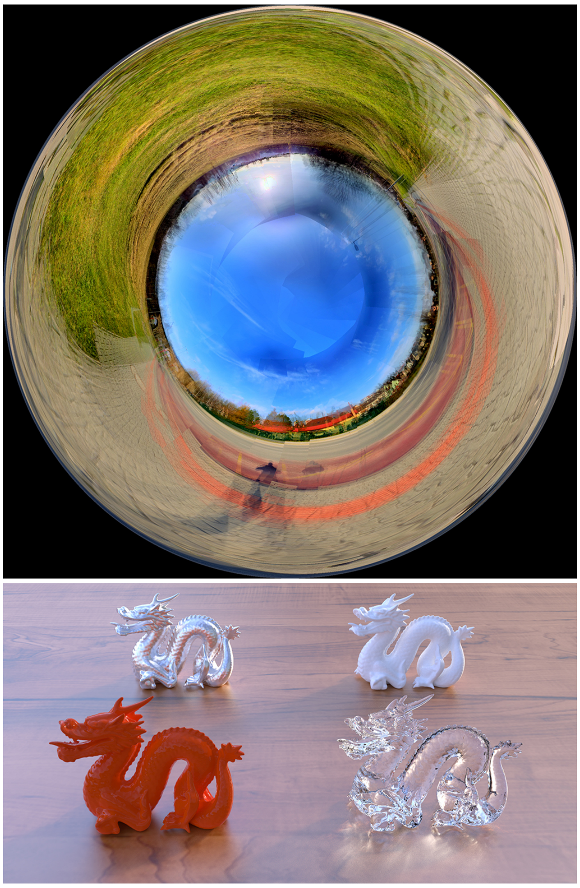

Figure 1: *(Top) An environment map captured by our method. (Bottom) A rendering with the captured map by path tracing. Importance sampling was used to efficiently sample the environment map.*

3. Environment Map Capturing

Our method for environment map capturing projects each camera image onto the spherical environment map and reconstructs HDR radiance from accumulated image data. The projection of the image is based on the current orientation of the mobile device. This orientation is calculated from the data of built-in IMU which includes gyroscope, accelerometer and magnetometer. The camera image is converted to the radiances and it is stored as float values. The weighted radiance and weights from multiple images are accumulated into a spherical environment map. Finally, the accumulated data are divided by the accumulated weights to get the radiance.

3.1. Camera Image Projection

In order to reconstruct the environment map, covering the full sphere of directions, we project multiple images with different device orientations onto one 2D texture. This texture represents the spherical environment map. We use the mapping to the sphere similar to Debevec [Deb98]. Distinctly, our map is oriented to have positive z axis (up vector) projected to the image center because we prefer to have a consistent image deformation along horizontal direction. We assume purely rotational motion of the mobile device. The error, caused by small translational motion can be compensated by image alignment (Section 4), especially for large environments. The projection of each image onto the environment map texture depends on the orientation of the mobile device and intrinsic camera parameters. We assume a camera with a fixed focal length. Therefore, the intrinsic parameters of the camera are reconstructed once in the calibration and can be reused for each projection which uses the same camera.

The spherical projection of the camera image to the environment map is calculated in the following way. For each pixel of the resulting environment map, we calculate the light direction $\vec{d}$ which corresponds to this pixel. The direction $\vec{d}$ is calculated by rotating the vector $(0,0,-1)$ ϕ degrees around x axis and θ degrees around z axis. The angles ϕ and θ are calculated by the following equations:

$$\theta = tan^{-1}(u,v), \quad \phi = \pi\sqrt{u^2+v^2} \tag{1}$$

The *uv* coordinates represent the position in the environment map texture and they are in range (-1,1). In the next step, the intersection of a ray with direction $\vec{d}$ and the image plane is calculated. The camera image data from this intersection is accessed, projected onto HDR radiance, and stored at the pixel of the environment map with image coordinates (u,v). If the environment map already contains values from previous frames, the new projected image data are accumulated. In order to avoid sharp discontinuities around image edges, we apply an image blending with linear falloff from the image center.

3.2. HDR Reconstruction

High dynamic range of reconstructed radiance data is essential to accurately represent the high contrast of real light. Therefore, low dynamic range data, captured by the image sensor, have to be converted to the HDR radiance. We use the approach presented by Robertson et al. [RBS99]. The inverse camera response function is reconstructed for each color channel in the calibration process. Then, we use these functions for the HDR reconstruction in the following procedure. A mobile camera is set to the auto-exposure mode to correctly adapt exposure time to the local portion of the scene where the camera is pointing. If two images overlap, they usually have slightly different exposure time due to the adapting auto-exposure. For i^{th} image, captured by the mobile camera, we know its exposure time t_i. Therefore, the radiance L_j can be calculated from the pixel values y_{ij} as:

$$L_j \approx \frac{\sum_i w_{ij} t_i f^{-1}(y_{ij})}{\sum_i w_{ij} t_i^2} \tag{2}$$

y_{ij} is the pixel value from the i^{th} image which is projected to the j^{th} pixel of the environment map. The radiance L_j corresponds to the j^{th} pixel of the environment map. $f^{-1}()$ is the inverse camera response function. The weight w_{ij} is calculated according to [RBS99]. The accumulation of the sum in the numerator of the fraction in Equation 2 is done for each color channel when a new image is captured and projected to the environment map. The denominator of this fraction is accumulated to the alpha value. Finally, each color channel is divided by the alpha value to obtain the reconstructed HDR radiance L_j. The HDR reconstruction runs on GPU and the environment map is accumulated to the floating point framebuffer to achieve interactive speed.

4. Image Alignment

The estimation of device orientation by built-in IMU sensors introduces drift which causes errors in the alignment between the projected camera images. Therefore, an additional alignment correction is beneficial. Similarly to [WMLS10], we employ a feature matching to align a new camera image with the captured environment map. In our approach, we render the camera view with the current orientation from the environment map by projecting the environment map to the camera frame. Then, we detect and match the image features from this rendered view with the captured camera image. In the next step, the rotational correction is calculated from the directions corresponding to the matched points. Finally, the orientation of the device is corrected by the calculated rotation. This approach improves the alignment of the overlapping images in the environment map. In some cases, an insufficient number of features are matched causing an imperfect image alignment. Nevertheless, the reconstructed environment map is beneficial for image-based lighting or material reconstruction, because the imperfect alignment is not visible after the light reflection on non-specular surface.

5. Results

In order to assess the capability of our method to capture the real illumination, we compare a rendered model of the dragon to the real dragon, printed on a 3D printer. The captured environment map is used to illuminate the rendered dragon by the image-based lighting. Additionally, differential rendering [Deb98] is used to composite the model into the real image. The reconstructed environment map and the comparison of rendered to real dragon can be seen in Figure 2. Light reflections on the rendered dragon are visually similar to the ones on the real dragon. The brightness of reflections is correct due to HDR environment map. Moreover, the direction and shape of the virtual shadows are similar to the real ones. Current disadvantage of our method is the imperfect alignment of images in some situations. Nevertheless, the result is visually acceptable, particularly if the reconstructed map is used for image-based lighting.

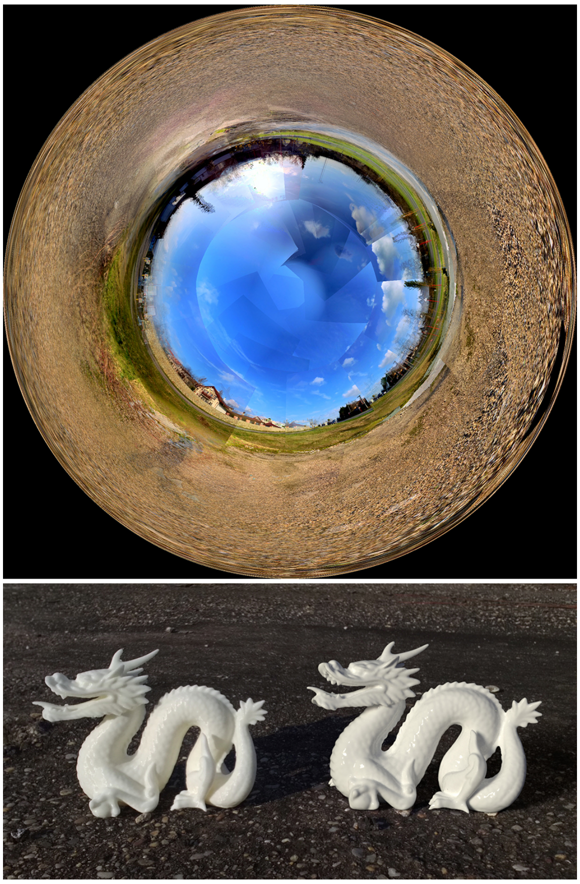

Figure 2: *(Top) The environment map captured by our method and tonemapped by exposure and gamma correction. (Bottom) The comparison of real (left) and rendered (right) dragons. The rendered dragon is illuminated by the captured environment map. The material was set manually.*

The average speed of image capturing, alignment, and projection to HDR environment map is 2 fps. The bottleneck of our approach is the slow camera framerate for taking pictures. A solution would be to use a live video. However, current mobile phones do not provide exposure time for preview video and the exposure time is essential for HDR calculation. We expect our method to run in real time on future devices that will provide exposure time for camera preview because the average time of image projection and HDR reconstruction is 117 ms with alignment and 1,3 ms without it. We used NVIDIA Shield tablet in our experiments.

6. Conclusion

This paper presents a novel method for HDR environment map capturing by a consumer mobile phone. Our method does not require any additional hardware and therefore is very practical for capturing of any environment. We project the camera images onto the spherical map based on the orientation of the device. Additionally, we reconstruct the HDR radiance from the captured and accumulated image data. Our approach is extended with image alignment method based on feature matching which corrects the orientation given by device sensors. The results show that the proposed method is beneficial for rendering with natural light. Our method can be valuable for many fields, particularly image-based lighting, material reconstruction and augmented reality.

7. Acknowledgments

We thank Georg Gerstweiler for his help with 3D printing and Annette Mossel for proofreading. The dragon model is the courtesy of Stanford Computer Graphics Laboratory. This research was funded by Austrian project FFG 843484.

References

[Deb98] DEBEVEC P.: Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH '98* (NY, USA, 1998), ACM, pp. 189–198. 1, 2, 3

[DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH '97* (NY, USA, 1997), ACM Press, pp. 369–378. 2

[DWH09] DIVERDI S., WITHER J., HÃŰLLERER T.: All around the map: Online spherical panorama construction. *Computers & Graphics 33*, 1 (2009), 73 – 84. 2

[DWWW08] DENG X., WU F., WU Y., WAN C.: Automatic spherical panorama generation with two fisheye images. In *WCICA 2008* (June 2008), pp. 5955–5959. 2

[GEM07] GROSCH T., EBLE T., MUELLER S.: Consistent interactive augmentation of live camera images with correct near-field illumination. In *VRST* (NY, 2007), ACM, pp. 125–132. 1, 2

[RBS99] ROBERTSON M., BORMAN S., STEVENSON R.: Dynamic range improvement through multiple exposures. In *ICIP* (1999), vol. 3, pp. 159–163. 1, 2, 3

[SJW*06] STUMPFEL J., JONES A., WENGER A., TCHOU C., HAWKINS T., DEBEVEC P.: Direct HDR capture of the sun and sky. In *SIGGRAPH 2006 Courses* (NY, USA, 2006), ACM. 2

[SS97] SZELISKI R., SHUM H.-Y.: Creating full view panoramic image mosaics and environment maps. In *Annual Conference on Computer Graphics* (1997), pp. 251–258. 2

[SSI99] SATO I., SATO Y., IKEUCHI K.: Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *IEEE Transactions on Visualization and Computer Graphics 5*, 1 (1999), 1–12. 2

[WMLS10] WAGNER D., MULLONI A., LANGLOTZ T., SCHMALSTIEG D.: Real-time panoramic mapping and tracking on mobile phones. In *IEEE VR* (2010), pp. 211–218. 1, 2, 3

[XP10] XIONG Y., PULLI K.: Fast panorama stitching for high-quality panoramic images on mobile phones. *Consumer Electronics, IEEE Transactions on 56*, 2 (2010), 298–306. 1