POINTAR: Efficient Lighting Estimation for Mobile Augmented Reality

Yiqin Zhao
Worcester Polytechnic Institute
yzhao11@wpi.edu

Tian Guo
Worcester Polytechnic Institute
tian@wpi.edu

1. Lighting Estimation and Its Challenges

In this poster, we describe the problem of lighting estimation in the context of mobile augmented reality (AR) applications and our proposed solution. Lighting estimation refers to recovering scene lighting with limited scene observation and is critical for realistic 3D rendering. As a long-standing challenge in the fields of both computer vision and computer graphics, the difficulty of lighting estimation is exacerbated for mobile AR scenarios. When interacting with mobile AR applications, users would trigger the placement of virtual 3D objects into any position or orientation in their surrounding environments. In order to present a more realistic effect, such objects need to be rendered with appropriate lighting information. However, lighting, especially in the indoor scenes, can vary both spatially and temporally.

At a high level, to obtain the lighting information at a given position in the physical environment, one would use a 360-degree panoramic camera that can capture incoming lighting from every direction. However, this leads three challenges in the mobile augmented reality scenarios. **First**, the AR application needs to estimate the lighting at the target location, i.e., where the 3D object will be rendered, based on the photo captured by mobile camera at the device location. **Second**, as the mobile camera often only has a limited field of view (FoV), i.e., less than 360 degree, the AR application needs to derive the lighting information outside the FoV. **Lastly**, as lighting information is used for rendering, the estimation should be fast enough and ideally to match the frame rate of 3D object rendering.

2. Our Proposed Solution: POINTAR

State-of-the-art lighting estimation approaches [1–3] are purely neural networks based and did not consider the unique challenges of executing on mobile devices. Gardner et al. proposed a simple transformation, without using the depth information, that can lead to image distortion. Garon et al. improved the spatial lighting estimations with a two-branches neural network that was reported to perform well on a laptop GPU but not on mobile devices. Song et al. further improved the estimation accuracy by decomposing the pipeline into differentiable sub-tasks. However, the overall network was large in size and computational complexity which makes it ill-suited for running on mobile phones.

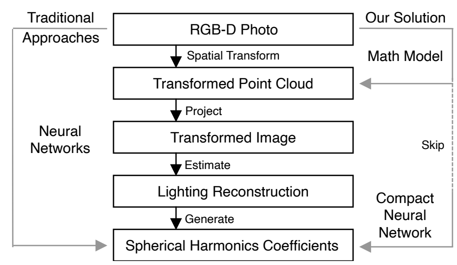

Figure 1. Traditional approaches vs. our proposed solution.

We propose a *hybrid lighting estimation* for mobile AR with the promise of realistic rendering effects and fast estimation speed. We rethink and redefine the lighting estimation pipeline by leveraging an efficient mathematical model and a compact deep learning model. Combined, POINTAR tackles the *view position spatial transformation* and the *limited view lighting reconstruction* problems. Specifically, we first transform the captured RGB-D image to a point cloud at the target location by simulating camera movement with a math model; we then train a compact neural network that predict lighting information (in the form of spherical harmonics coefficients) from the transformed point cloud, shown in Figure 1. In summary, our proposed POINTAR uses the fast math model instead of neural networks and skips a portion of the traditional pipeline therefore saving inference execution time, while maintaining good accuracy.

Results and Demo We plan to include experiment results of POINTAR in estimating lighting information for an open-source dataset Matterport3D during our poster presentation. We also plan to include a demo that showcases the effect our lighting estimation on the rendered 3D objects.

References

[1] Marc-André Gardner et al. Learning to predict indoor illumination from a single image. *ACM Transactions on Graphics*, 2017.
[2] Mathieu Garon et al. Fast spatially-varying indoor lighting estimation. *CVPR*, 2019.
[3] Shuran Song et al. Neural illumination: Lighting prediction for indoor environments. *CVPR*, 2019.