論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Reconstruction and Interactive Reformation of the 3D Geometry and Lighting in Single Image
著者(和文)	WuJung-Hsuan
Author(English)	Jung-Hsuan Wu
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第10551号, 授与年月日:2017年3月26日, 学位の種別:課程博士, 審査員:斎藤 豪,亀井 宏行,徳永 健伸,宮崎 純,篠田 浩一
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第10551号, Conferred date:2017/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	要約
Type(English)	Outline

Thesis Outline

This thesis, which is titled as "Reconstruction and Interactive Reformation of the 3D Geometry and Lighting in Single Image," includes five chapters. The following is the brief introduction of each chapter.

Chapter 1 introduces the background and the purpose of this thesis. The diffuse lighting in a real-world scene captured by a digital camera indicates lots of important information about the scene, which includes its 3D structure, its materials, and the light sources which illuminate the scene. However, they are extremely difficult to be recovered from the captured photograph. This thesis presents the methods to semi-automatically estimate and create a pseudo 3D scene from a single photograph. The results of this thesis show that the created pseudo 3D scenes allow the user to interactively edit the image by changing the lighting and by moving, inserting, or removing the objects such as human and car in the image, and visually plausible results are synthesized in real-time during the editing stage. This chapter also briefly introduces the following chapters

In chapter 2, the state-of-the-art research to estimate and reconstruct the 3D geometry and configuration of illumination of the scene in photographs is surveyed and discussed. Different from the previous work presented in this chapter, this thesis only requires a single photograph and a small amount of user inputs. It is the first method that allows editing of existing lighting in an image and does not need any pre-knowledge of the scene, multiple photographs, and special equipments.

In chapter 3, a novel top-down approach which estimates the luminance distribution in a photograph and reconstructs a pseudo 3D scene by adopting a simplified physically based lighting model, which consists of the camera, light source, and environment reflection, is proposed. It only requires a single input photograph and a small amount of user annotations. This method adopts the hill climbing strategy in order to overcome the difficulty of simultaneously estimating the camera parameters and the illumination conditions. It can stably estimate the luminance distribution with light spots on the diffuse surfaces and recover visually plausible environment reflection. The confidence map is also presented to enhance the robustness of the estimating process against occlusion, shadow, the influence of the other light spot, and the highlight-clipping effect. Although this method assumes that the scene consists of Lambertian planes, which are horizontal ground and vertical walls, the experimental results show that even if the estimation error of the results increases, this method still produces visually plausible images for scenes consisting of bumpy and/or non-Lambertian surfaces, which means that this method is robust enough for relighting images. The user study indicates that the relit results synthesized with the pseudo 3D scene created by the proposed method from the images used in the experiment are so realistic that people are hard to distinguish them from real photographs.

In chapter 4, a shadow-based method is presented to extend the method proposed in chapter 3 by estimating and reconstructing more details of the scene from one photograph so that the synthesized images look more realistic. This shadow-based method requires one input photograph, the parameters of camera and light source, and a small amount of user annotations to create pseudo 3D geometries for bumpy surfaces and objects, such as cars and humans, whose cast-shadow is also in the input photograph. The pseudo 3D geometry of a bumpy surface is estimated with its self-shadow, which is the shadow caused by part of the surface itself. The pseudo 3D model of an object is recovered with an iterative optimization method which starts by creating a rough pseudo 3D model depending on the contour of the object in the photograph. The rough pseudo 3D model is then iteratively refined by minimizing the difference between the synthetic shadow and the cast-shadow in the input photograph. The experimental results show that the pseudo 3D models produce visually plausible shadows while it is slightly moved around its original position and that it still looks visually plausible even if the viewing direction is slightly changed.

Chapter 5 summarizes the methods proposed in this thesis and presents the future works and applications of this thesis.

The methods proposed in this thesis have been proved to be able to estimate and reconstruct pseudo 3D scene geometry and the configurations of illumination from a single input photograph with a small amount of user assistance. The reconstructed scenes can be used to synthesize visually plausible images. Potential applications of this thesis include films, video games, digital image editing, image content understanding, automatic scene estimation, and image forgery detection. This thesis is very useful for many fields and industries.