LaTeX Author Guidelines for CVPR Proceedings

Anonymous CVPR submission

Paper ID ****

Abstract

We present a novel method for inferring a hidden moving scene from faint shadows cast on a diffuse surface by an unknown occluder. We model the system as a simple 2D convolution between the scene and the occluder. Thus, we formulate the task of jointly recovering the unknown scene and unknown occluder as a blind deconvolution problem.

Blind deconvolution is a notoriously ill-posed problem. To help solve it, we assume the hidden scene contains motion, which helps us solve the problem for two reasons. First, each frame of the observation gives us the occluder convolved with a different kernel, providing many different "views" of the same object. Second, motion in realistic scenes is typically sparse, which we make use of in our reconstruction.

We demonstrate our method to be effective with simulations and experiments under a variety of settings. We also explore the dual problem, in which there is a fixed scene and moving occluder.

1. Introduction

Imaging scenes that are not directly visible, also called non-line-of-sight (NLoS) imaging, is a difficult and often ill-posed problem. Recently, it has become an area of active study [citations]; methods that rely on visible light to image hidden scenes usually presume that there is something directly visible to both the observer and the hidden scene (see Fig. **??** for an illustration of such a scenario). In this work, we refer to this visible area as the *observation plane.*

Past methods that rely on human-visible light to image hidden scenes can be divided into one of two categories: *active* methods, which introduce light into the scene and make use of known or measured properties of the introduced light, such as time of return, to image the hidden scene [citations]. *Passive* methods, on the other hand, rely exclusively on ambient light from the scene, such as secondary reflections on the observation plane, to infer the contents of the hidden scene [citations]. In general, active methods have are more powerful and have a wider variety of tools available with which to do imaging, but passive methods are more widely applicable, since they can be deployed even without the help of lasers or time-of-flight cameras. However, passive methods do usually rely on the hidden scene being ambiently lit and cannot be used to image dark scenes. The methods presented in this paper are passive methods, and as such assume the hidden scene to be lit.

Passive methods suffer from the fact that in real-world settings, only a two-dimensional array of observations can be observed (see e.g. Fig. **??**), but the scene producing those observations is three-dimensional. Hence, the problem in such cases is inherently ill-posed. Past methods have resolved this issue by either assuming the scene lies on a lower-dimensional manifold, thereby only reconstructing only a lower-dimensional projection of the scene [citations], or making use of a strong spatial prior over realistic scenes to reconstruct [citations]. Our method falls into this former category, as we assume that both the scene and occluder lie on parallel, flat planes. This allows us to model the shadows cast on the observation planes as a simple convolution of these two planes (see Fig **??** for an illustration).

Although there has by now been plenty of previous work demonstrating that it is possible to use the presence of an occluder to infer the structure of a hidden scene [citations], this work, to our knowledge, is the first to do so in a *blind* manner, meaning that we know nothing *a priori* about the structure of the occluder. Past work that exploits occlusion either uses scene calibration to get a precise picture of the occluder before system can work [citations] or is limited to situations in which the occluder has some basic, common shape, like a pinhole, pinspeck, or edge [citations]. The blind nature of this problem compounds the already daunting challenge of non-line-of-sight imaging. However, we hope that this will make our method widely applicable in a variety of situations in which occluders are complex but pre-calibration is not an option, such as traffic or search and rescue [citations].

2. Background

This work draws inspiration from past work from two broad categories: the first is non-line-of-sight imaging, particularly occlusion-aided non-light-of-sight imaging, and the second is past work in blind deconvolution. To our knowledge, this work is the first to synthesize these two well-studied areas of research into an algorithm that does something novel: get a two-dimensional view of a hidden scene, with only minimal assumptions about the hidden scene and unknown occluder.

2.1. Non-line-of-sight imaging

2.2. Blind deconvolution

3. Scenario

3.1. Setup

Our model of the scenario consists of three elements: a hidden moving scene, an occluder, and the observation plane. We model each of these elements as 2D planes parallel to each other. See Fig. **??** for an illustration.

The hidden scene is presumed to be a collection of diffuse reflectors, shining light uniformly in all directions towards the occluder and observation plane. The hidden scene is also presumed to contain some motion. The unknown occluder is presumed to be a set of perfectly black, planar objects. We assume the hidden scene, unknown occluder, and observation planes to each be a substantial distance apart, relative to their sizes.

The observation plane is presumed to be perfectly Lambertian. In simulations, we also presume the observation plane to be white and uniform, and that all of the light reaching the observation plane comes from the scene; in experiment, we use mean-subtraction to account for non-white, non-uniform observations with ambient “nuisance” light sources, a method also employed in other work (e.g. [citations]). This allows us to apply our method to most realistic scenarios with minimal adaptations to the core algorithm. We explore the effect of other deviations from the idealized sceenario we present here in Section **??**.

3.2. Light Propagation

We model the propagation of light from the scene through the occluder as a simple 2D convolution of the scene with the occluder. This follows from the assumptions laid out in the previous sections, in particular the ones that posit that the three scenario elements are distant parallel planes. Ours is not the first work to model Lambertian occluder-based light propagation as a convolution; indeed, an identical assumption is made in [citations], where it is explained in depth why the assumptions we make imply such a model for light propagation. In this paper, we limit our explanation of this phenomenon to the illustration in Fig. **??** and accompanying caption, and describe how robust this assumption is in the real world in Section **??**.

In simulations, we assume that we see the full convolution of the scene and the occluder on the wall. If the scene is a plane of size $x_s \times y_s$ and the occluder a plane of size $x_o \times y_o$, this corresponds to an observation of $(x_s + 2x_o) \times (y_s + 2y_o)$. See Fig. **??** for an illustration of why this is.

4. Occluder Estimation

5. Scene Reconstruction

6. Deviations

7. Results

8. Conclusion