University of Kentucky Doctoral Dissertations

Graduate School

2011

Improving SLI Performance in Optically Challenging Environments

Eric Dedrick
University of Kentucky, dedrick@vbnm.net

ABSTRACT OF DISSERTATION

Eric Dedrick

The Graduate School
University of Kentucky
2011

Improving SLI Performance in Optically Challenging Environments

---
ABSTRACT OF DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Engineering at the
University of Kentucky

By
Eric Dedrick
Lexington, Kentucky

Director: Dr. Daniel L. Lau,
Professor of Electrical and Computer Engineering
Lexington, Kentucky 2011

ABSTRACT OF DISSERTATION

Improving SLI Performance in Optically Challenging Environments

The construction of 3D models of real-world scenes using non-contact methods is an important problem in computer vision. Some of the more successful methods belong to a class of techniques called structured light illumination (SLI). While SLI methods are generally very successful, there are cases where their performance is poor. Examples include scenes with a high dynamic range in albedo or scenes with strong interreflections. These scenes are referred to as optically challenging environments.

The work in this dissertation is aimed at improving SLI performance in optically challenging environments. A new method of high dynamic range imaging (HDRI) based on pixel-by-pixel Kalman filtering is developed. Using objective metrics, it is show to achieve as much as a 9.4 dB improvement in signal-to-noise ratio and as much as a 29% improvement in radiometric accuracy over a classic method. Quality checks are developed to detect and quantify multipath interference and other quality defects using phase measuring profilometry (PMP). Techniques are established to improve SLI performance in the presence of strong interreflections. Approaches in compressed sensing are applied to SLI, and interreflections in a scene are modeled using SLI. Several different applications of this research are also discussed.

KEYWORDS: SLI, range reconstruction, HDRI, compressed sensing

Author's signature: _____ Eric Dedrick

Date: _____ May 4, 2011

Improving SLI Performance in Optically Challenging Environments

By
Eric Dedrick

Director of Dissertation: Daniel L. Lau

Director of Graduate Studies: Stephen D. Gedney

Date: May 4, 2011

RULES FOR THE USE OF DISSERTATIONS

<u>Name</u>   <u>Date</u>

_____

_____

_____

_____

_____

_____

_____

DISSERTATION

Eric Dedrick

Improving SLI Performance in Optically Challenging Environments

DISSERTATION

A dissertation submitted in partial
fulfillment of the requirements for
the degree of Doctor of Philosophy
in the College of Engineering at the
University of Kentucky

By
Eric Dedrick
Lexington, Kentucky

Director: Dr. Daniel L. Lau,
Professor of Electrical and Computer Engineering
Lexington, Kentucky 2011

Dedicated to my parents and my brother, who have given me much support and encouragement in this endeavor.

ACKNOWLEDGMENTS

I would like to thank all those who have helped me along my career.

First and foremost, I would like to thank Dr. Daniel Lau for his time and efforts throughout my time at the University of Kentucky. He has been an instrumental part of my success.

I would also like to thank Dr. Laurence Hassebrook and Dr. Kevin Donohue for their contributions to this effort. On numerous occasions their expertise was invaluable.

Finally, I would like to thank my recently graduated peers Dr. Kai Liu and Dr. Yongchang Wang. They have been very helpful at every opportunity, and I have been inspired by their enthusiasm and drive for excellence.

CONTENTS

LIST OF TABLES

Chapter 1 Introduction

The research in this dissertation aims to improve the performance of 3D modeling using structured light illumination (SLI) in optically challenging environments. One type of optically challenging environment considered in this dissertation occurs when the dynamic range in scene albedo exceeds what can be captured by an image sensor using a single integration time. Another optically challenging environment examined in this dissertation occurs when multipath interference is present in a scene.

Before discussing details about the research performed in this dissertation, this introduction focuses on providing the reader with broad, non-technical information in order to better understand this topic, the challenges being addressed, and the techniques utilized in developing solutions. This begins with understanding the goals of machine vision.

What is Machine Vision?

Human vision is a highly developed sense. The visual cortex is the largest system in the human brain, and it allows many complex tasks to be performed with seemingly little effort. However, while humans excel at tasks associated with vision, it is machines which excel at tedious, high-throughput tasks associated with automation and inspection. Consequently, there is much interest in allowing computers to "see."

Consider the image shown in Figure 1.1. A human observer has little difficulty identifying this scene as a culture of cells, and can easily differentiate between background clutter and healthy cell colonies. Based on this, one might believe that allowing computers to "see" should be an easy task. However, consider the boxed region of Figure 1.1 presented as it would be "seen" by a computer, shown in Figure 1.2. Although both regions contain the same information, when this information is presented in a manner which makes it inaccessible to the human visual cortex, the identification of healthy cell colonies becomes an extremely challenging task for humans, and the difficulty of giving computers "sight" becomes apparent.

Sometimes distinctions are drawn between several closely related disciplines. The field of image processing aims to alter and enhance the way the information in an image gets presented to the human visual cortex. For example, the contrast between cell colonies and background can be increased, or cell colony edges can be enhanced. Computer vision uses techniques from the field of artificial intelligence to allow a computer to "understand" an image. For example, a set of features describing cell colonies can be created using attributes such as shape and color, and techniques from pattern analysis and machine learning can be used to differentiate healthy cell colonies from background clutter and undesirable colonies. The field of machine vision does not aim to make computers "understand" images, but rather focuses on making high fidelity measurements useful for applications such as remote sensing or industrial inspection. While these fields may appear to be separate disciplines, in practice there are no clear-cut boundaries between these areas, and many applications require the
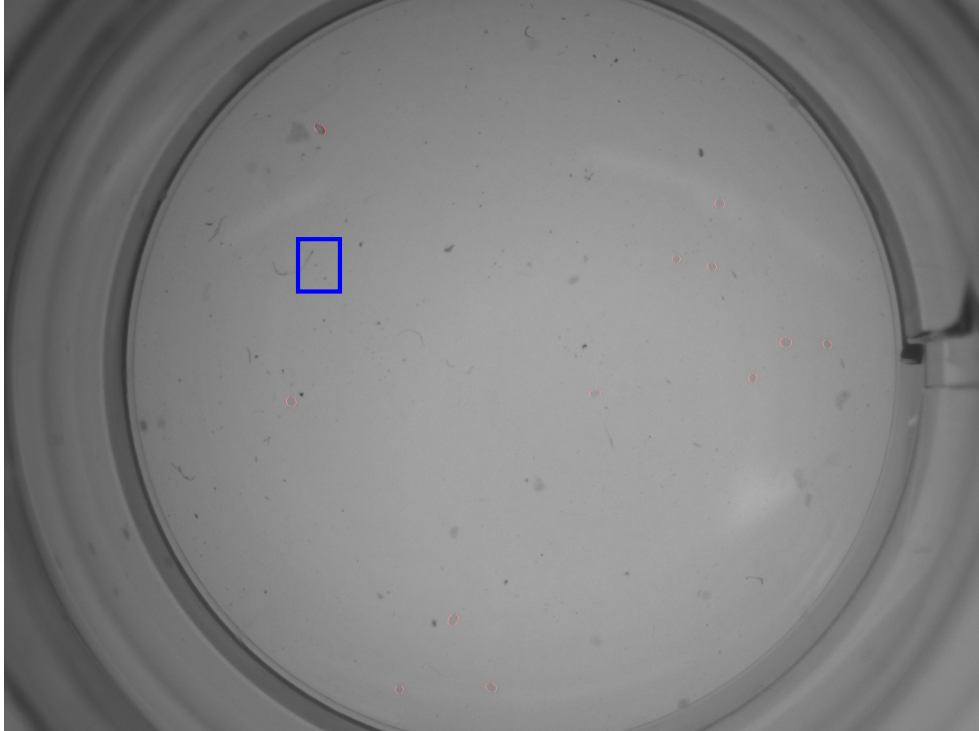
Figure 1.1: An image of a culture of cells presented in a form accessible to the human visual cortex.

use of all three disciplines.

The research in this dissertation focuses on machine vision, specifically techniques for acquiring high fidelity radiance and depth measurements.

3D Capture Technologies

The acquisition of 3D scenes has been an important problem in machine vision for several decades, with many approaches having been explored. Following the organization in [85], these approaches may be classified as

- single camera,

- multi-camera,

- holographic,

- pattern projection, and

- time-of-flight approaches.

Holographic techniques use a coherent light source and capture interference patterns between a reference beam and light scattered by the object. Since they are based on light interference, holographic approaches can offer excellent measurement

```
135 135 137 136 137 140 138 141 141 135 135 138 137 139 139 138 139 140 141 139 141 142 142 139 141 141 140 138 140 139 142 141 138 142
140 140 139 140 143 140 140 141 139 142 141 142 140 143 141 143 138 138 137 139 138 139 138 135 137 138 138 138 138 136 137 138 138 139
139 138 140 139 141 140 142 141 138 140 140 138 138 139 140 139 139 141 140 141 141 140 140 141 141 142 135 138
139 140 141 139 138 140 139 139 138 137 136 141 135 140 139 139 140 136 139 141 141 140 140 140 141 142 140 138 140 140 139 139 140 140
138 138 140 140 140 142 139 139 141 140 142 139 139 139 141 139 141 144 136 137 138 142 139 139 137 137 137 140 136 137 140
140 140 140 139 141 138 142 141 140 142 139 140 139 141 138 141 139 140 137 142 140 139 141 142 143 143 141 144 144 141 136 137 137 139
137 136 138 138 140 136 136 140 138 136 140 138 140 136 139 140 138 140 142 143 140 138 140 142 139 140 141 137 138 140
141 138 142 140 142 141 141 145 141 143 142 144 139 137 139 136 137 139 138 138 140 138 139 139 138 139 137 138 138 140 137 138 138 137
140 139 138 141 139 142 139 140 141 141 141 140 138 139 139 140 139 141 139 139 139 141 137 140 140 140 140 140 137 142 143 144 141 142
141 135 140 140 139 139 137 138 139 134 139 139 139 138 140 138 140 139 141 139 139 139 141 137 140 140 140 140 140 137 142 143 144 141 142
142 142 142 142 144 141 143 142 143 143 138 137 138 138 138 136 139 139 139 137 138 138 139 142 142 140 139 138 139
137 137 142 137 141 140 141 141 141 139 141 138 139 140 140 142 143 142 141 141 140 141 142 144 144 141 138 139 138 138 141 136 135 138
138 136 139 138 138 139 136 140 138 141 139 143 137 141 137 137 140 139 140 140 139 139 142 138 139 139 141 138 139 142 140
140 142 142 140 142 141 142 142 137 138 135 139 138 138 139 137 136 139 138 138 139 139 137 137 138 140 138 140 139 138 137 136 141 142
140 135 137 140 140 141 139 142 139 140 142 140 142 142 141 143 144 139 139 141 140 138 138 139 138 138 138 137 137 137 136 139
138 137 140 140 138 139 137 137 140 140 137 141 139 141 137 141 138 138 140 139 140 142 139 143 138 140 141 142 142 142 140 141 144 141
143 139 143 144 141 139 139 138 139 138 135 135 137 133 136 142 139 138 138 138 138 138 138 140 142 139
144 140 140 140 140 142 140 140 140 139 142 140 141 141 142 140 140 142 141 142 142 143 141 138 141 136 138 137 139 136 137 139 141 142 140
138 139 140 140 141 140 137 136 135 137 135 137 136 140 139 141 141 140 142 138 139 139 138 143 146 141 141 144
141 143 144 140 138 138 138 137 139 138 137 137 140 139 140 138 137 139 137 139 140 137 136 135 135 136 137 142 140 138 140 139 137 138
141 144 140 140 142 143 141 141 138 144 139 146 143 142 139 141 146 142 142 143 137 139 140 139 139 139 138 138 139
137 136 138 137 136 125 127 134 137 136 140 139 139 139 141 137 140 139 139 139 140 141 143 140 141 140 142 140 142 141 142 142 142 143
141 142 140 137 140 137 136 137 139 138 137 139 141 141 138 139 141 137 137 139 135 135 138 127 117 119 127 133 136 140 140 138 141 141 140
142 138 142 139 142 143 141 141 140 141 142 144 141 139 145 146 145 144 140 140 141 137 141 139 136 138 136 139 136 139 135 138 138 140
138 135 127 114 114 124 131 135 138 141 141 142 142 141 141 140 140 141 142 139 141 142 141 140 142 142 143 144 142 142
139 139 136 138 139 138 140 141 140 138 139 139 138 137 138 139 138 136 119 114 121 133 137 139 135 141 140 143 138 140 140 144 141 139
142 142 141 143 142 143 143 143 144 143 144 145 144 145 140 140 138 139 137 140 140 138 134 137 138 133 128
117 120 132 139 139 140 138 137 137 141 141 143 142 142 143 138 141 143 142 142 143 141 141 143 143 143 140 147 144 141 144 144 139 139
139 139 139 140 137 137 139 139 140 137 139 136 132 129 125 127 138 137 138 138 141 141 140 139 139 141 142 138 139 139 141 138 139 142
145 140 143 144 143 143 144 144 144 142 145 144 144 142 137 141 138 140 141 138 140 139 140 142 139 138 137 138 137 136 130 129 132 137
141 139 142 141 139 141 142 141 142 139 140 141 142 143 142 145 146 143 145 144 142 141 142 140 138 138 139
140 141 141 140 141 139 139 138 138 135 138 134 130 131 138 140 141 140 141 141 142 140 141 141 139 140 144 141 142 141 141 140 140 143
144 143 143 141 142 143 142 144 144 144 142 140 136 139 136 139 139 137 138 140 140 136 139 139 134 129 131 136 137 141 142 138
139 140 141 140 140 142 141 137 139 141 141 141 143 143 142 143 142 144 140 144 141 140 141 147 142 143 145 143 142 137 139 140 138 139
140 137 139 139 139 138 135 129 128 131 136 141 140 140 140 141 141 140 138 138 139 137 131 128 132 136 139 138 141 140 142 140
142 141 142 142 144 144 142 144 144 143 138 142 141 139 139 139 141 139 140 138 138 139 137 131 128 132 136 139 138 141 140 142 140
139 143 144 144 141 143 142 142 143 143 141 145 143 141 144 144 145 144 145 141 139 139 139 141 141 142 139
137 139 141 135 132 129 129 133 138 139 138 143 140 143 142 141 141 139 144 142 142 139 143 141 142 142 143 141 141 145 142 141 143 143 141
142 144 141 143 146 143 143 145 141 141 140 141 141 140 139 139 137 132 126 127 131 136 138 140 138 138 143 139 141 141 140 141 143 141
142 141 140 141 140 142 142 144 142 140 142 145 142 141 144 144 142 142 143 140 145 147 142 143 142 141 139 139 139 140 138 142 141 138
138 133 125 131 138 135 136 138 141 141 141 139 142 142 142 143 140 141 141 144 142 142 142 144 143 145 145 142
143 143 141 145 148 144 142 138 140 140 138 140 139 141 143 139 136 126 130 133 136 139 139 142 140 141 140 142 142 142 143 139 142 140
142 143 143 140 141 140 142 140 140 145 142 144 145 143 143 140 140 140 140 144 144 145 143 139 140 140 137 132 130
132 137 137 136 142 140 141 141 142 141 142 141 140 144 138 138 142 144 142 140 142 142 141 141 142 143 144 143 142 143 145 141 143 143
143 144 143 144 136 138 138 141 141 139 142 132 130 135 136 137 140 140 140 141 143 141 138 140 141 137 138 138 139
142 143 141 146 143 142 139 144 143 146 144 143 143 142 144 144 145 144 144 145 139 137 139 141 140 139 141 142 136 133 129 135 138 139
139 140 141 140 140 139 138 142 139 139 137 139 140 143 144 142 140 142 139 140 139 143 141 142 145 142 143 144 144 147 142
145 146 138 141 140 139 139 141 141 140 138 134 134 137 140 141 140 140 138 140 140 141 142 139 142 137 134 132 136 139 143 142 140 140
144 141 142 138 143 144 143 144 145 142 146 143 141 141 141 141 140 138 135 135 136 140 141 141 144 145 144 143 144 144 144 145 146 145 146 143 143 140 141 139
141 141 140 142 138 139 140 131 126 124 131 139 142 142 142 142 138 142 143 144 144 143 145 144 143 144 144 144 145 146 145 146 144 142
140 142 138 140 139 138 136 132 136 135 138 138 139 134 130 129 136 139 142 143 141 142 142 139
143 142 141 144 144 143 144 143 145 142 144 144 146 144 145 146 138 141 139 139 136 138 134 132 134 137 141 139 139 140 141 141 138 141
142 145 141 141 140 137 137 140 141 142 142 143 141 141 145 145 140 149 146 145 144 145 145 145 143 147 141 140
138 140 138 136 131 135 137 139 139 142 139 142 138 140 139 140 139 142 141 141 141 140 142 143 146 142 144 141 141 141 143 141 144 139
146 143 140 144 143 145 144 144 144 145 144 147 139 138 137 134 134 133 130 136 138 138 139 140 140 140 140 140 141 140
142 142 142 144 142 143 142 143 145 143 144 144 141 144 144 144 144 143 145 140 140 143 147 143 145 145 147 145 145 142 146 141 139 136 132
133 135 134 138 140 139 138 139 142 140 140 140 141 143 141 141 143 141 145 146 145 143 145 141 143 142 141 144
143 142 141 145 143 144 144 147 146 143 147 146 139 134 135 131 135 132 136 138 141 142 141 139 139 140 143 142 145 141 140 140 141 143
142 142 145 144 148 145 143 142 141 145 144 145 147 142 144 146 145 147 142 146 145 147 145 133 133 135 136 135
137 138 140 137 142 140 142 141 142 141 141 142 144 143 143 142 143 141 146 148 145 142 142 143 142 143 141 140 144 140 146 143 143 144
143 144 146 147 143 145 143 144 144 145 136 132 136 135 140 139 138 139 143 145 146 140 142 143 142 143 143 144
145 144 143 142 145 142 141 144 141 144 142 141 146 144 145 142 145 145 146 145 146 145 145 145 135 139 139 141 138 140 139 141
142 141 141 141 143 138 141 141 143 143 140 140 144 146 145 143 145 143 144 146 143 145 141 145 146 143 145 146
144 142 144 146 146 145 146 146 140 138 139 141 141 140 138 141 140 141 140 139 141 141 141 140 139 141 144 144 143 141 143 144 142 142
141 143 141 144 139 145 141 142 142 142 143 145 144 144 146 146 140 142 139 139 139 141 139 141 141 139 141 141 141
141 140 141 143 144 143 144 144 141 140 141 143 139 143 147 145 143 143 140 143 141 142 141 143 140 141 142 146 145 144 146 144 144 146
145 143 145 144 145 148 143 141 143 138 138 141 141 142 145 142 142 145 145 144 144 143 142 143
143 142 143 141 145 142 145 142 142 143 144 144 145 145 145 145 145 148 144 146 145 146 141 141 141 143 139 142 140 142 140 140 141 142
140 142 143 143 140 140 143 142 143 145 142 143 144 143 140 143 142 142 142 141 141 142 145 145 146 146 146 143 143
145 145 144 148 139 140 141 140 137 142 142 141 142 140 138 143 139 143 145 142 143 139 143 144 143 145 142 144 144 142 142 143 144 139
143 145 143 146 144 142 141 143 142 145 145 146 143 143 143 141 141 141 141 141 141 141 142 143 143 146
140 142 142 143 143 142 144 140 140 141 143 146 144 141 141 141 144 143 140 142 145 141 140 141 139 139 145 145 146 146 143 143
145 143 143 144 141 141 142 138 141 141 144 143 141 141 145 145 144 143 141 145 145 144 147 146 143 144 142
141 142 140 136 132 134 138 143 142 146 146 144 145 146 145 149 145 145 141 140 143 140 139 140 142 141 142 140 141 142 141 139
142 143 141 145 143 142 143 143 144 139 140 142 147 142 142 136 128 122 121 134 142 147 146 147 146 148 142 143 145 148 145
138 142 145 142 142 141 143 139 142 144 143 143 143 143 143 141 144 142 143 142 142 145 143 141 139 142 142 143 143 141 144 142
133 131 125 122 133 144 146 147 145 146 146 144 146 144 144 146 144 144 140 141 142 144 141 141 141 141 145 143 143 144 144
145 140 142 144 141 143 142 136 134 140 142 144 144 144 140 141 141 140 137 139 146 146 147 147 143 144 144 145 148 146 145 140 140
144 141 143 144 141 142 142 141 140 143 143 144 140 143 142 142 134 138 141 143 143 142 146 144 141 144 142
144 146 140 144 143 138 143 140 144 144 143 142 143 145 145 144 141 147 142 143 145 145 144 145 146 142 142 140 143 142
143 143 140 140 143 140 144 142 141 143 146 143 143 149 143 145 141 142 142 143 142 144 141 142 145 141 145 144 144 148 145 145 146 144
143 145 146 144 144 146 143 145 145 147 147 144 139 143 143 145 141 145 145 145 141 141 141 141 145 143 141 142 143 143 141
142 142 141 146 144 143 144 145 147 146 145 146 145 141 145 145 144 143 147 142 144 143 146 144 145 145 143 141 141 142 143 141 140
143 141 144 142 142 143 140 144 142 143 143 144 144 142 144 144 142 141 141 141 144 142 141 141 144 145 145 144 147 146 143 144 142
143 143 142 144 143 141 144 145 149 144 141 141 142 143 143 142 142 142 141 141 141 144 142 141 141 144 145 143 143 141 144
145 143 145 142 142 144 147 144 147 143 144 143 143 141 146 143 142 142 135 140 144 145 145 146 144 146
```

Figure 1.2: The boxed region of Figure 1.1 presented in a form not accessible to the human visual cortex. Both images contain the same information.

accuracy. However, holography still largely uses analog approaches, and while advances in digital holography have been made, resolution, data storage, and speed are just some of the issues preventing its widespread use [85].

The many single camera techniques available can be broadly classified as shape-from-shading [98], shape-from-texture [43], shape-from-focus [101], and shape-from-motion [62]. Each of these methods has its own advantages and disadvantages. However, 3D recovery from one view is an ill-posed problem, and the performance of these approaches is very much dependent on the validity of their assumptions. While
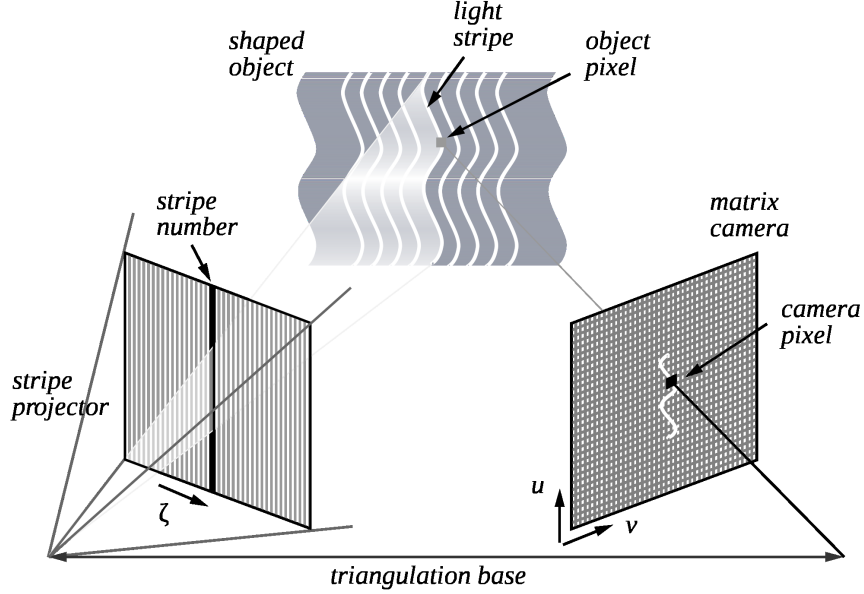
Figure 1.3: An illustration of optical triangulation in a structured light scanner. From [1].

these techniques may be appropriate and useful in some applications, for measurement applications performed under general conditions other techniques are preferable.

3D reconstruction from multi-camera stereovision has received much attention, and one review on the subject is [78]. In these approaches, 3D reconstruction is performed using triangulation between the cameras. After camera calibration, correspondences must be established between pixels in the two cameras. Unfortunately, this can be a very challenging task, especially in low texture environments.

Like the multi-camera approaches, pattern projection methods, also referred to as structured light illumination (SLI) methods, recover 3D coordinates using triangulation between a camera and a projector, shown in Figure 1.3. Correspondences between camera and projector pixels are established by projecting, detecting, and decoding coded patterns, thus overcoming the correspondence problem present in multi-camera stereovision. A number of coded patterns exist, with many summarized in [75, 4]. Some popular approaches are based on projecting binary patterns such as gray codes. Also popular are techniques which project phase-shifted sinusoids such as phase measuring profilometry (PMP). Good surveys on the subject can be found in [4] and [75].

Several time-of-flight systems, also referred to as light detection and ranging (LIDAR) systems, exist. Pulse LIDAR is based upon fast electronics and reconstructs range from

$$T = \frac{2d}{c} \tag{1.1}$$

where T is the measured delay between pulse generation and detection, d is range, and c is the speed of light. One such system, described in [61], uses an avalanche

4

photodiode array capable of single photon counting and performs range reconstruction with sub-millimeter accuracy. When fast electronics are not required, frequency modulated continuous wave (FMCW) or amplitude modulated continuous wave (AMCW) systems can be used. In FMCW systems, the emitted signal increases in frequency at a constant rate and the frequency shift between emitted and detected signal is used to yield time of flight, with range again determined from (1.1). An example of AMCW is the Equinox sensor described in [25], which uses two photodetectors per pixel, one modulated in-phase and the other out-of-phase with the emitted sinusoidal signal, to estimate the phase delay between emitted and received signals. Phase delay is then used to calculate time-of-flight and range.

For most measurement applications, 3D reconstruction is performed using either SLI or LIDAR. LIDAR has received considerable attention in recent years because of its high accuracy, high speed, low power requirements, and mechanical simplicity [2]. However, projector-camera range sensors also have high accuracy, use low-cost commodity hardware, have a fairly simple system design, and can achieve real-time performance [45], so they too are widely used as 3D scanning systems.

1.1 SLI in Optically Challenging Environments

Although SLI techniques are generally very successful, there are some cases where their performance is poor and the intense illumination of the projector can create or exacerbate some problems. Because SLI techniques assume direct path illumination will result in a stronger detected response than interreflections, as is the case with Lambertian surfaces, specular surfaces pose considerable problems. Scenes with strong interreflections create problems because a projected code can be detected at multiple locations in scene, resulting in correspondence ambiguities. Additionally, the decreased SNR can lead to decode errors when binary patterns are used, and cause phase shifts when sinusoidal patterns are used. Additional challenges exist when the dynamic range in albedo exceeds what is captured by the camera. Quantization and saturation can result in a partial or total loss of the projected signal, and due to the blooming effect where electrons in saturated pixels leak to neighboring pixel sites, neighboring pixels may also be affected. For this reason, scenes with non-Lambertian scattering, strong interreflections, and high dynamic range in albedo are referred to collectively as optically challenging environments.

An example where the dynamic range in scene intensity exceeds what is captured by the image sensor is shown in Figure 1.4. One of the captured patterns is shown in Figure 1.4a, with over-exposed and under-exposed regions clearly visible. Several defects in the reconstructed surface are visible in Figure 1.4b–Figure 1.4d. Many of the points in the corrugated aluminum foil are completely lost due to signal saturation. The shape of the polished, conical piece of aluminum is not correctly reconstructed, and many surfaces which should be flat are not. There are numerous other problems such as isolated points at the wrong locations.

Another example of an optically challenging scene is shown in Figure 1.5, where PMP is used to scan a scene containing a mirror and a compact disk placed in front of a uniform white screen. The first defects are due to non-Lambertian scattering by the

(a)

(b)

(c)

(d)
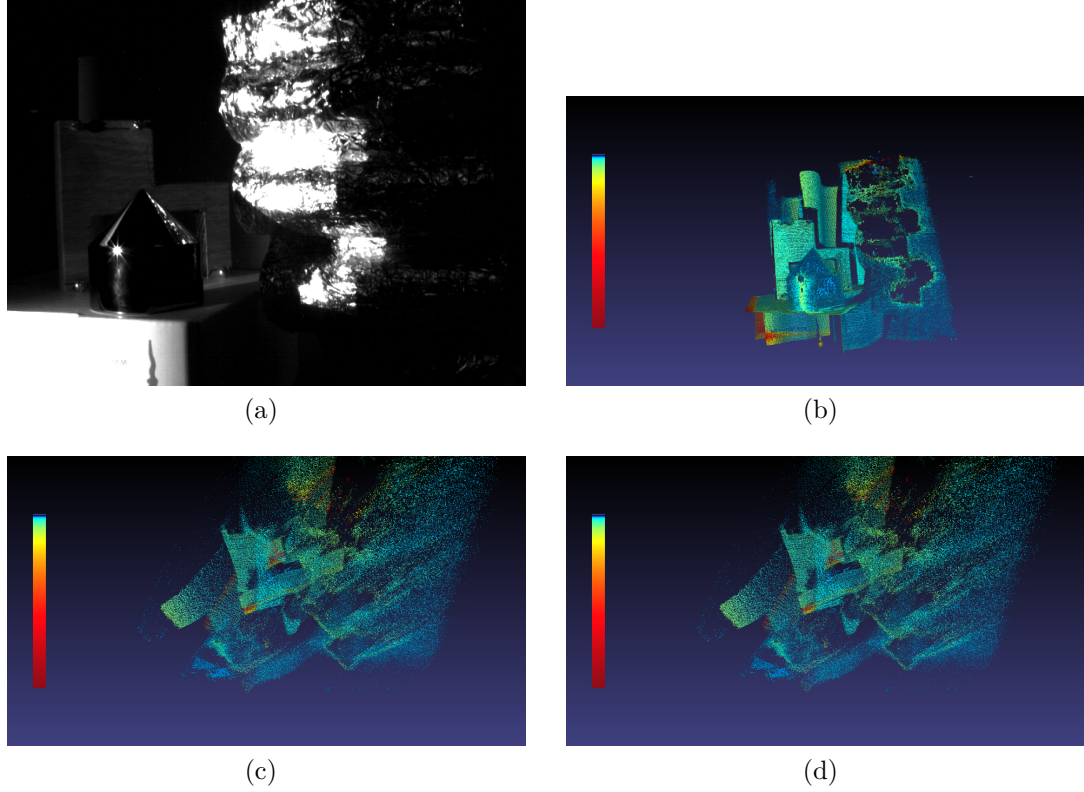
Figure 1.4: Images of (a) an optically challenging scene with high dynamic range in scene intensity, and (b)-(d) several views of the 3D reconstruction.

mirror. Almost all of the direct-path projector illumination incident on mirror (B) is reflected onto the screen at (C) and very little is reflected into the camera. Instead, light from region (A) is reflected by the mirror into the camera. Since the strongest detected codes do not correspond to direct path illumination, the reconstructed depth of the mirror (B) is incorrect. Region (C) has direct-path projector illumination, but also has a strong reflection from the mirror (B). Since the projected patterns are sinusoidal, multipath interference results in a phase shift which is related to the difference in the path lengths. This results in incorrect reconstruction at (C). In (D), reflections along a piece of aluminum are so bright that the dynamic range of the camera is exceeded. The image sensor becomes saturated, and much of the projected sinusoidal pattern is lost, causing poor phase estimates. In (E), codes from other locations in the image are detected due to a reflection on the table surface. Consequently many points which receive no direct illumination are also modeled. Finally, in (F) interreflections are strong in the corner, resulting in a distortion of the shape.

Some efforts have been made to address the problems associated with SLI scanning in optically challenging environments, including approaches dealing with the high dynamic range (HDR) in scene albedo, and approaches which address interreflections. Before discussing interreflections, we examine HDR imaging and range sensing, topics which will first require an understanding of image sensors.
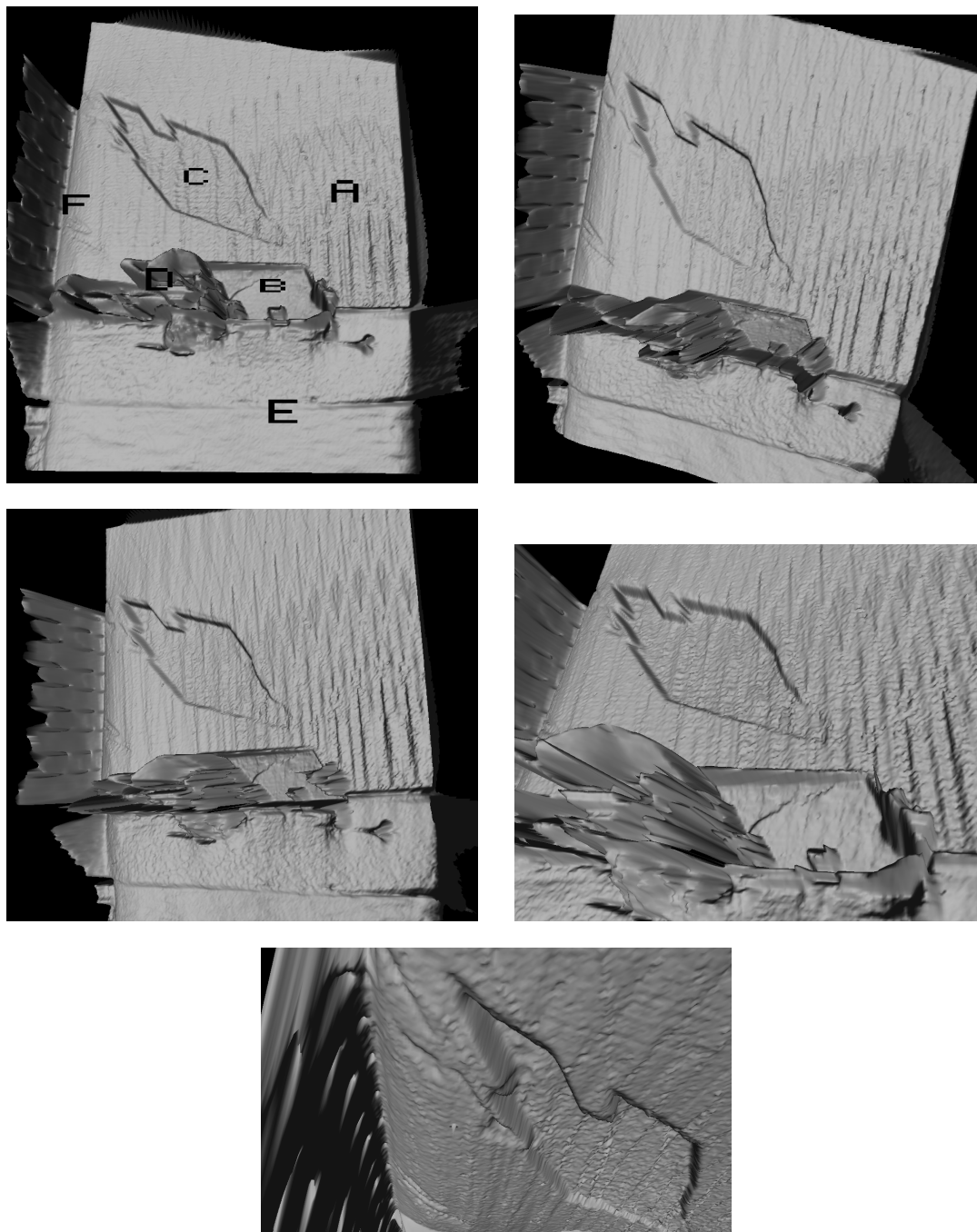
Figure 1.5: Several views of an optically challenging scene with some defects labeled.

Image Sensors

If computers are to "see," then cameras must be their eyes, and if one is focused on measurement applications, knowledge of cameras and image sensors is essential. Many cameras are unsuitable for use in measurement applications for a number of reasons. They may have a non-linear response designed to mimic the response of
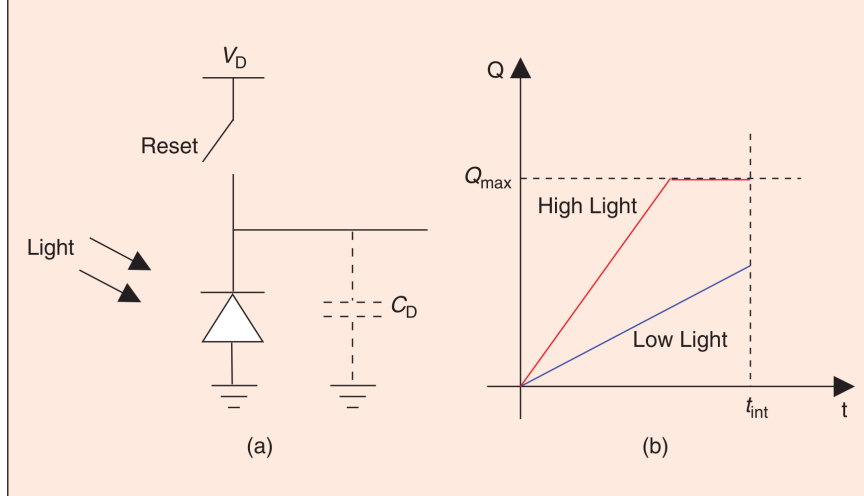
Figure 1.6: Direct integration in an image sensor pixel. From [21].

analog film. Precise control of camera parameters such as exposure time or sensor gain may not be possible. Compression may be applied to the images, and color or image processing operations may be performed for artistic effect resulting in a loss of signal fidelity. While these cameras may be excellent choices for photography applications, machine vision cameras are designed specifically for use in measurement and automation applications. They have a linear response, often are monochromatic, allow precise user control of camera parameters, are high speed, and have features such as networking designed to make them easy to interface with computers.

Assuming a machine vision camera is selected for an application, one should understand the properties image sensor being used. Two common classes are the charge coupled device (CCD) and complementary metal-oxide-semiconductor (CMOS) active pixel sensor (APS) image sensor designs. Both CCD and CMOS sensors use reverse biased PN junction photodiodes as photodetectors, converting incident light into a very small amount of photocurrent, usually femtoamperes. As seen in Figure 1.6 (a) after reset to the diode voltage V_D, the current from the photodiode is integrated across the diode capacitance C_D, resulting in (b) charge being accumulated as integration time t_{int} increases. Depending on integration time and light level, the pixel well can attain its maximum capacity Q_{max} and saturate, with additional charge leaking to neighboring pixels. In APS, each pixel has a source follower which performs the charge-to-voltage conversion and isolates the accumulated charge from sensor circuitry.

After integration, sensor readout is performed, shown in Figure 1.7. In a CCD sensor, charge is transferred along both vertical and horizontal charge coupled devices to achieve a raster readout of charge, which is then converted to voltage and digitized. Readout for a CMOS sensor is performed much in the same way as computer random access memory (RAM) using row and column select circuits. Each column of pixels has a separate amplifier and analog-to-digital converter (ADC) .

Each sensor has its own advantages and disadvantages, with many ideas and
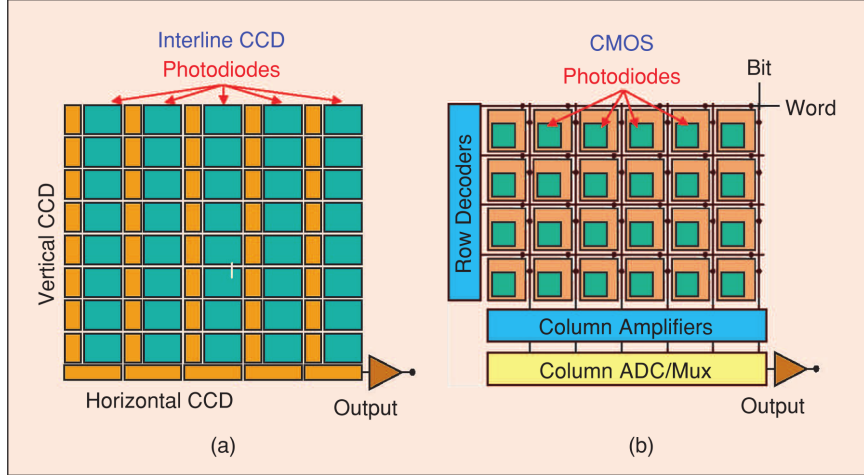
Figure 1.7: Readout design for (a) CCD and (b) CMOS image sensors. From [21].

architectures having been developed [21]. In general, CCD image sensors have smaller pixel sizes and better signal uniformity due to the single amplifier and ADC, but require high power to transfer charge and are slower since an image is read serially. Due to their pixel amplifiers and row and column select transistors, CMOS sensors with APS have larger pixel sizes with smaller photodiode fill factors. Due to their many active circuit elements, column amplifiers, and column ADCs, CMOS sensors also have much greater non-uniformity. However, CMOS sensors are high speed, low power, and are generally lower cost due to their simpler fabrication process.

As just mentioned, CMOS image sensors are generally noisier than CCD image sensors due mainly to the presence of the many column amplifiers and active circuits, and, to a lesser degree, the presence of the read circuitry. Nevertheless, both CCD and CMOS image sensors have temporal noise from a number of sources such as shot noise, reset noise, readout noise, and quantization noise. Fixed pattern noise is also present in these sensors due to variations in dark current, column amplifiers, ADCs, and even non-uniformities in manufacturing. Fixed pattern noise sources are classified as having dark signal non-uniformity (DSNU) if they result in non-uniformity in sensor response offsets, or photo-response non-uniformity (PRNU) if the noise source creates non-uniformities in response gains. The dominant noise source and signal-to-noise ratio (SNR) are dependent on light level, as seen in Figure 1.8.

Another source of fixed pattern noise is due to pixel vignetting [11, 21]. As shown in Figure 1.9, the height of the image sensor results in only perpendicular light striking the photodetector. Due to the nature of many types of camera lens systems, this effect increases for pixels further away from the optical axis, and can be exacerbated by the microlens array placed on the sensor, shown in Figure 1.10. Geometrical optics contributes to this problem and introduces other sources of non-uniformity such as optical vignetting.

The field of solid state image sensors is large, and many details have been omitted from this discussion. An excellent resource on solid state devices is [69], and good starting points on CCD and CMOS image sensors are [21] and [44].

9
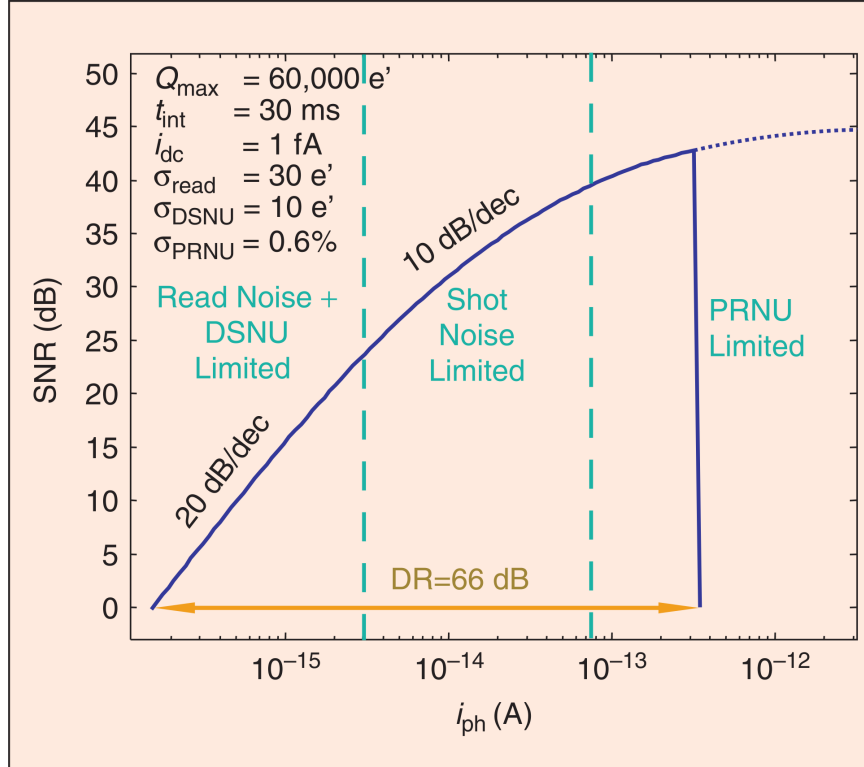
Figure 1.8: SNR as a function of photocurrent for an image sensor showing the dominant noise sources. From [21].
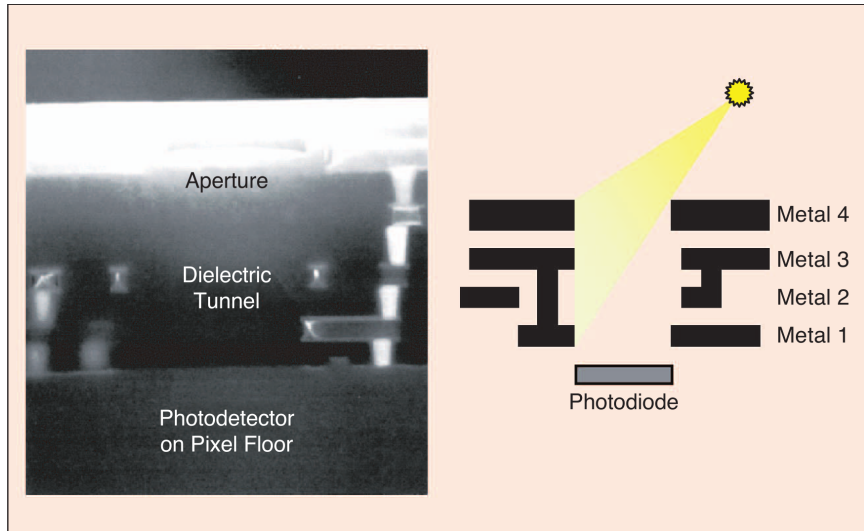
Figure 1.9: An illustration of how pixel vignetting can occur when light is not normal to the sensor. From [21].

Based on this understanding of image sensors, we now examine the challenges which arise when a scene has a dynamic range in intensity beyond that which can be captured by the image sensor. We also examine several strategies which have been
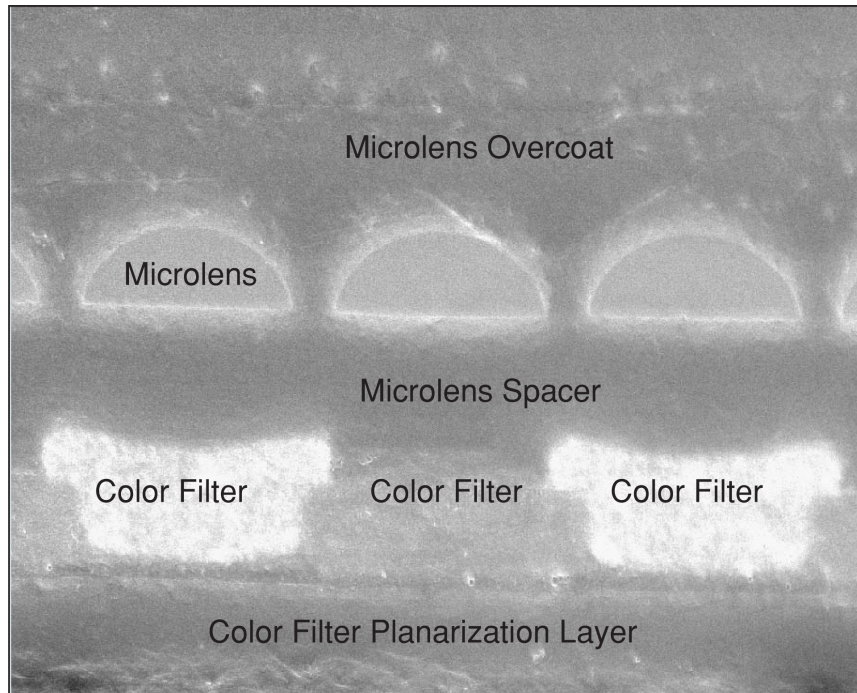
Figure 1.10: A SEM image of the cross section of an image sensor showing the microlens array. From [21].

devised to increase the dynamic range of image sensors.

HDR Imaging

As seen in Figure 1.6, under sufficiently strong illumination, the charge storing capacity of a pixel can be reached, resulting in both saturation, and leakage of charge into neighboring pixel wells, a condition known as "blooming." On the other hand, the signal received by pixels under weak illumination may be lost entirely due to quantization in the ADC. These problems can occur simultaneously at different locations in an image sensor, and when this happens the scene is said to have a high dynamic range of intensities.

Several CMOS sensor designs have been developed to address this, and they are categorized in [84] as

- companding sensors,

- multimode sensors,

- clipping sensors,

- frequency based sensors,

- time-to-saturation sensors,

- global-control-over-integration-time sensors, and

11

- autonomous-control-over-integration-time sensors.

Companding sensors extend the dynamic range of the sensor by using pixels which have a logarithmic pixel response. Multimode sensors have a linear response in low light, but under high illumination have a logarithmic response. Clipping sensors change pixel capacitance during integration. Pixels in frequency based sensors emit pulses when saturated and then reset. Thus, they produce a series of pulses whose frequency is a function of light intensity. Time-to-saturation sensors record the integration time at which a pixel becomes saturated. Global-control-over-integration-time sensors acquire multiple images at different integration times and use the pixel value at the maximum non-saturating integration time to construct a HDR image. Finally, autonomous-control-over-integration-time sensors detect pixels which are about to saturate, reset them, and record the fact that the pixel has "rolled over."

Clearly, many sensor-based approaches to HDR imaging exist, and with so many HDR sensors, most HDR application requirements can probably be met through appropriate sensor selection. For example, a sensor presented in [36] has a dynamic range of 120 dB and operates at up to 1000 frames per second, which should be sufficient for all but the most demanding applications. Despite this, HDR image sensors are not yet common in most cameras, so HDR imaging (HDRI) must be performed using multiple exposure approaches. These techniques construct HDR images by acquiring multiple low dynamic range images at varying integration times. Multi-exposure approaches are examined in more detail in Chapter 2.

Scenes with high dynamic range in intensity are common in photography, but also arise in SLI due to the high intensity light being introduced by the projector. Thus, HDR range sensing has also been studied.

HDR Range Sensing

Some approaches have been developed to improve SLI performance when a scene has a high dynamic range in albedo, and one of the simplest is the work of Zhang and Yao [99]. A PMP scan is performed and the pattern sequence captured by the camera is examined to see if saturation has occurred in any of the camera pixels. If saturation occurred, a new PMP scan is performed with projector intensity reduced globally. This process continues until saturation has been avoided in every camera pixel. Since pixels are processed independently in PMP, phase is determined at each pixel location using only the patterns from the scan in which the pixel has the greatest non-saturating intensity.

Skocaj and Leonardis [82] perform SLI using binary stripe patterns. By projecting stripe patterns with varying projector intensities and performing radiometric calibration of both the camera and the projector, HDR images are constructed for each pattern. These HDR images are then used as inputs to the decoder instead of the usual low dynamic range images, and 3D reconstruction is then performed, resulting in a greatly improved 3D model.

Prior to projecting sinusoidal illumination patterns, Koninckx *et al.* [40] perform a radiometric calibration for both camera and projector, and establish the minimum

and maximum projector intensities for each camera pixel so that under- and overexposure can be avoided. This information is then used along with iterative updates of surface shape to locally adapt the projected patterns to the scene and thus avoid over- or under-exposure, thus allowing patterns to be captured in a single camera image.

While these techniques show that high dynamic range in albedo can be addressed by regulating projector intensity, these approaches have drawbacks. In [99], the problem of under-exposure is not addressed, and addressing it may ultimately require adjustments in camera parameters such as sensitivity and exposure time. Furthermore, no HDR albedo image is formed. Both [82] and [40] require radiometric calibration of both camera and projector, and the locally adaptive patterns in [40] require additional computational cost to generate geometry and pattern updates.

While these approaches reduce the dynamic range of the scene by decreasing illumination from the projector, an alternative would be to increase the dynamic range of the image sensor. This is the approach examined in Chapter 2.

As mentioned earlier, dynamic range is only one issue which can result in an optically challenging environment. Another is the presence of multipath interference in the scene, which often occurs as a result of interreflections.

Scene Interreflections

As seen in Figure 1.5, interreflections can cause errors in SLI scanning. Some approaches have been developed to improve SLI robustness in the presence of interreflections. Xu and Aliaga [95] use an adaptive approach based on masked time-multiplexed gray codes, where patterns and their binary inverses are projected, and the brighter of two determines if a pixel is classified as "on" or "off." The difference between the two intensities determines the quality of the measurement. A large difference in pixel values between the "on" and "off" states indicates a good measurement, while little difference between the two suggests corruption due to an interreflection, and the pixel is then classified as uncertain. Scanning proceeds iteratively with only uncertain pixels being projected. This reduces the global illumination on each pass leading to improvement in SNR for the remaining uncertain pixels.

Other approaches are based on constraint testing. Trucco and Fisher [89] perform SLI using multiple cameras and consistency checks to detect false range data. These consistency checks are based several ideas. The same code should not be detected at multiple locations in the image. A camera can not detect a code from a location on a specular surface whose normal is not oriented towards the camera. If a code is observed by multiple cameras, each should yield a consistent reconstructed point. Finally, a code detected in one camera should be detected in other cameras unless there is an occlusion. These constraints allow a specular surface with concave holes to be successfully scanned.

Park and Kak [65] develop the technique of multipeak range imaging to store multiple candidate range measurements, and then use numerous constraint and consistency tests on scans performed from many viewpoints to differentiate between true range data and spurious reflections. This approach is shown to be successful for many

types of optically challenging objects.

Also noteworthy are approaches based on illumination with polarized light. Clark *et al.* [14] use changes in linear polarization to differentiate between specular reflections and direct illumination. Scanning is performed with a linearly polarized laser stripe and a camera with a polarizing filter, with images captured at multiple angles of polarization. Chen *et al.* [13] combine polarization analysis with another technique [57] to separate direct illumination from subsurface scattering in the scanning of translucent objects.

Although these techniques can allow SLI to be performed in the presence of interreflections, they may require non-standard hardware or have added complexity due to constraint checks or adaptive feedback control. None of these approaches consider illumination with sinusoids. Furthermore, multipath interference may result from more than just interreflections. For example, a single pixel may be illuminated by both foreground and background surfaces due to fractional pixel filling at the edge of an object. Both foreground and background surfaces may also be detected if some motion occurs during scanning.

Finally, a commonly employed solution which addresses both HDR in albedo and scene interreflections is to apply a coating with better optical characteristics to the object being scanned. However, this may introduce surface defects, and in many cases may be impractical or even impossible. Additionally, in computer graphics SLI is used to create light transport models, and in these applications altering the reflectance properties of the target object is unacceptable.

Light Transport and Compressed Sensing

Thus far we have examined SLI as a tool used in machine vision. However, as previously mentioned, it is also used in other fields such as computer graphics. Since accurate models of light transport are required for realistic computer renderings of scenes, these models are developed from real-world scenes imaged under controlled illumination, thus leading to the use of SLI in these applications.

An example of this is the work of Debevec *et al.* [17] in which renderings of the human face are generated from images captured with a light stage. The authors also raise a fundamental problem: light transport at a surface is described by an eight-dimensional quantity called the reflectance field and sampling it poses enormous measurement and storage requirements. Much work is devoted to addressing this challenge. One approach is to simplify the light transport model. In [17], the reflectance field is approximated with a non-local reflectance field of lower dimensionality.

In other works [57, 55] an even simpler light transport model is used where the intensity of a camera pixel is attributed to a single component describing the effects of direct illumination and another term describing global illumination effects.

Some works model light transport as a matrix relating each projector pixel to each camera pixel. Switz *et al.* [79] discuss how this light transport matrix can be determined through impulse sampling and show how it can be used to cancel interreflections. This approach, however, still has considerable sampling and storage requirements, so instead of impulse sampling, others [80, 67] use compressive sampling

approaches to reduce these costs.

These light transport models, while useful in computer graphics, do not aim to reconstruct 3D surfaces. However, they do serve as a basis for our goal of simultaneously obtaining both 3D surface and light transport information. For example, they suggest that compressed sensing approaches can be utilized to reduce the storage and measurement requirements associated with modeling light transport.

Compressed sensing (CS) is a rapidly developing field which aims to reduce the number of measurements required to recover a signal when the signal has a sparse representation in some basis. To better understand CS, we first examine some basic concepts in data compression and then examine the measurement process.

In data compression, a compressible signal $\mathbf{x}$ can be projected into a new basis $\mathbf{\Phi}$ with a new representation $\mathbf{X}$. This may be written as

$$\mathbf{x} = \mathbf{\Phi}\mathbf{X}, \tag{1.2}$$

where each row of $\mathbf{\Phi}$ is a basis vector. With a suitable choice of basis, $\mathbf{X}$ will have a large number of negligible coefficients which can then be truncated to achieve compression. In the appropriate basis, compression can reduce storage and transmission requirements by many orders of magnitude. For example, in image compression $\mathbf{x}$ is a 2D signal obtained from an image sensor, often with several million elements. However, in the wavelet basis only a few thousand terms in $\mathbf{X}$ are required to suitably reconstruct $\mathbf{x}$.

Of course performing dense, evenly spaced sampling only to then discard most of the resulting data is a very inefficient process. It can also be undesirable when there is a cost associated with sampling. For example, attaining high sampling rates can be both difficult and expensive, and in cases such as x-ray computed tomography (CT) imaging, a poor sampling strategy subjects a patient to unnecessary radiation exposure. Reducing the number of required measurements is essential.

The measurement process can be represented as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \tag{1.3}$$

where $\mathbf{y}$ is a set of measurements, $\mathbf{x}$ is a signal or state, $\mathbf{A}$ is a measurement matrix, and $\mathbf{n}$ is additive noise. The fundamental theorem of linear algebra states that for a noise free system, (1.3) can not be solved unless the number of equations is equal to the number of unknowns, and even more equations are required for the solution to be robust to noise. However, under-determined systems are quite common. The field of regularization theory focuses on solving (1.3) when the system is under-determined by introducing *a priori* knowledge, usually gained from physical reasoning about the problem, to supply enough additional constraints to solve the system.

Compressed sensing takes a different approach. As with (1.2), if $\mathbf{x}$ were a compressible signal and a suitable basis were known, only a small number of measurements would be required to recover $\mathbf{x}$. As will be discussed in more detail in Chapter 4, when the matrix $\mathbf{A}$ is chosen such that certain properties are satisfied, $\mathbf{x}$ can be recovered using an optimally minimal number of measurements.

CS has been applied successfully in many fields. However, to date the only application of CS in SLI is [30], where an inhomogeneous participating media is

examined. This application is very different than traditional SLI performed in 3D surface reconstruction. As such, the suitability of CS in traditional SLI is examined in Chapter 4.

1.2 Overview of Dissertation

As we have seen, the use of SLI methods to perform 3D surface reconstruction is widespread and is generally very successful, but performance problems can occur in optically challenging environments. HDR in albedo and multipath interference are two such challenges encountered in SLI scanning. Few solutions have been developed to address these problems, and many of them have significant drawbacks which may make them unsuitable for many applications. Furthermore, these effects can sometimes be subtle and not easily anticipated, raising the concern that defects might go undetected. Additionally, none of these approaches attempt to incorporate any modeling of light transport into the scanning process.

The research in this dissertation aims to address these problems. Unlike other approaches, we postulate that HDR imaging methods used in measurement applications should be based on image sensor models and state estimation approaches. We seek to develop an approach to detect and quantify the presence of multipath interference in PMP. Additionally, we aim to improve SLI performance in the presence of interreflections while also constructing a light transport model. Finally, we examine how recent advances in CS can be applied in SLI.

As previously mentioned, one type of optically challenging scene occurs when the dynamic range in scene albedo results in a range of intensities beyond what can be captured by a image sensor using a single integration time. With sensor-based solutions not yet commonplace, Chapter 2 examines strategies which extend the dynamic range of the image sensor by acquiring multiple images at different integration times, focusing on the suitability of these techniques for use in machine vision applications. Several common and state-of-the-art techniques are presented in detail. We then develop a new approach to HDRI which is based on an understanding of image sensors and pixel-by-pixel Kalman filtering, with model parameters determined through calibration. We then construct objective performance metrics to assess the accuracy and precision of these different HDRI approaches. It is shown that the Kalman filtering based approach achieves as much as a 9.4 dB improvement in SNR, and as much as a 29% improvement in radiometric accuracy over a classic method. Additionally, the effect of sampling strategy on the performance of HDRI techniques is explored.

We have seen that another type of optically challenging scene is one with multipath interference. While other approaches either use a non-standard scanning system or obtain some robustness to interreflections through the use of adaptive binary coding, commonly used sinusoidal patterns are not considered. Consequently, in Chapter 3 we review the process of range reconstruction using PMP. We extend a quality metric from [77] to develop tests which can identify and quantify the presence of multipath interference in PMP scanning. We detect multipath interference arising from a number of sources, specifically interreflections, fractional pixel filling, and motion

during scanning. This allows an operator to detect and quantify the presence of these quality degrading effects and improve the setup of the scene. This information can also be used to suppress or weight noisy measurements in post-processing steps.

As already discussed, SLI is used in areas other than just machine vision. With the goal of achieving realistic appearance of synthetic scenes, the measurement of light transport in real-world scenes is driving the use of SLI in computer graphics. Chapter 4 builds on work in computer graphics to develop an approach which can improve the robustness of 3D reconstruction in the presence of interreflections, while also allowing a model of light transport to be constructed. Both improved robustness and light transport measurement capabilities are demonstrated with a scene containing strong interreflections. We also examine how advances in CS can be utilized in SLI, discuss when these CS techniques may be practical, and show typical results.

Finally, conclusions and extensions of this work are discussed in Chapter 5. An approach to HDR video based on a fading memory filter and weighted pixel voting is used to adaptively determine camera integration times. The Kalman filter based HDRI method of Chapter 2 has been used to perform multispectral imaging of metallic objects of historical significance, so applications related to art conservation are discussed. The Kalman filter based HDRI method has also been used as a core technology in an automated assay selection system. Additionally, the use HDRI in some computer vision algorithms, such as scale invariant feature transform (SIFT) is discussed. Application of SLI approaches developed in Chapter 3 and Chapter 4 to other optically challenging environments, such as those with sub-surface scattering, is also discussed.

To summarize, SLI is a class of approaches used to perform 3D reconstruction of surfaces. In general, it is very successful and widely used. It is not, however, without some limitations. Among these are challenging scenes due to a high dynamic range in albedo and multipath interference resulting from effects such as interreflections. The research in this dissertation develops approaches to robust performance in these optically challenging environments. It also shows how SLI can be used to model light transport due to interreflections, a result useful in computer graphics applications. The use of compressed sensing in SLI is also studied. Some of the research undertaken in this dissertation has already been employed in various machine vision applications. While the research in this dissertation allows SLI to be performed under less restrictive conditions, it is anticipated that this research can be utilized in applications other than those considered in this work, such as the scanning of translucent objects [23], or objects with strong subsurface scattering such as marble sculptures, which is an area of interest to the digital Michelangelo project [38, 39, 41]. This project aims to provide users with a remote and interactive means of exploring sculptures of great cultural significance, while also helping conservationists ensure these works are preserved.

Chapter 2 High Dynamic Range Imaging

It is common in photography for the dynamic range of a real world scene to exceed what can be captured by an image sensor with a single integration time, and this can result in images with under- or over-exposed areas. An example of this from [18] is shown in Figure 2.1.

The brightness of a scene is described by the radiometric quantity radiance, which is the radiant power subtended by a solid angle incident upon a surface in a given direction. Expressed differently, radiance describes the radiant power emitted from a patch in a scene captured by an imaging system with a particular field of view, imaging area, and orientation. Irradiance is the density of radiant power incident upon some surface area. Spectral radiance describes the radiant power at a particular wavelength emitted by a source captured by the imaging system, and spectral irradiance is the radiant power at a particular wavelength incident on a surface area.

The photoresponse of solid state image sensors is wavelength dependent, and is referred to as the quantum efficiency of the sensor. This dissertation follows the convention of [18], where radiance is understood to mean this spectral irradiance weighted by the quantum efficiency of the sensor. Additionally, since camera and image sensor properties such as pixel area, orientation, and field of view are fixed during use, precise radiometric definitions are not required. Since radiance and irradiance are directly proportional, they are used interchangeably in the literature, and either of these quantities may be referred to as intensity. Additionally, absolute radiance is difficult to establish, so radiance is usually understood to be relative to some reference, a convention which this dissertation will also adopt.

Exposure refers to the quantity of light detected by an image sensor pixel, and is related to both radiance and integration time. Similar to the relation in physics between power and work,

$$X = rT, \tag{2.1}$$

where X is exposure, r is radiance, and T is integration time. The reciprocity relation states that if r and T are varied such that X remains constant, the same digital pixel value will result. In common usage, exposure also refers to the image produced by a sensor using a specific integration time. To avoid confusion with X, in this dissertation the term exposure-image will be used.

The dynamic range of a scene describes the ratio of the greatest radiance to lowest radiance, usually expressed in decibels. There is no precise definition of what constitutes high and low dynamic range. In this dissertation, a scene which can be captured by an image sensor with a single integration time without over- or under-exposure is referred to as having low dynamic range (LDR) , and scenes which can not be captured using a single integration time are referred to has having HDR. HDR images are sometimes referred to as radiance maps, but this dissertation refers to them as HDR images.
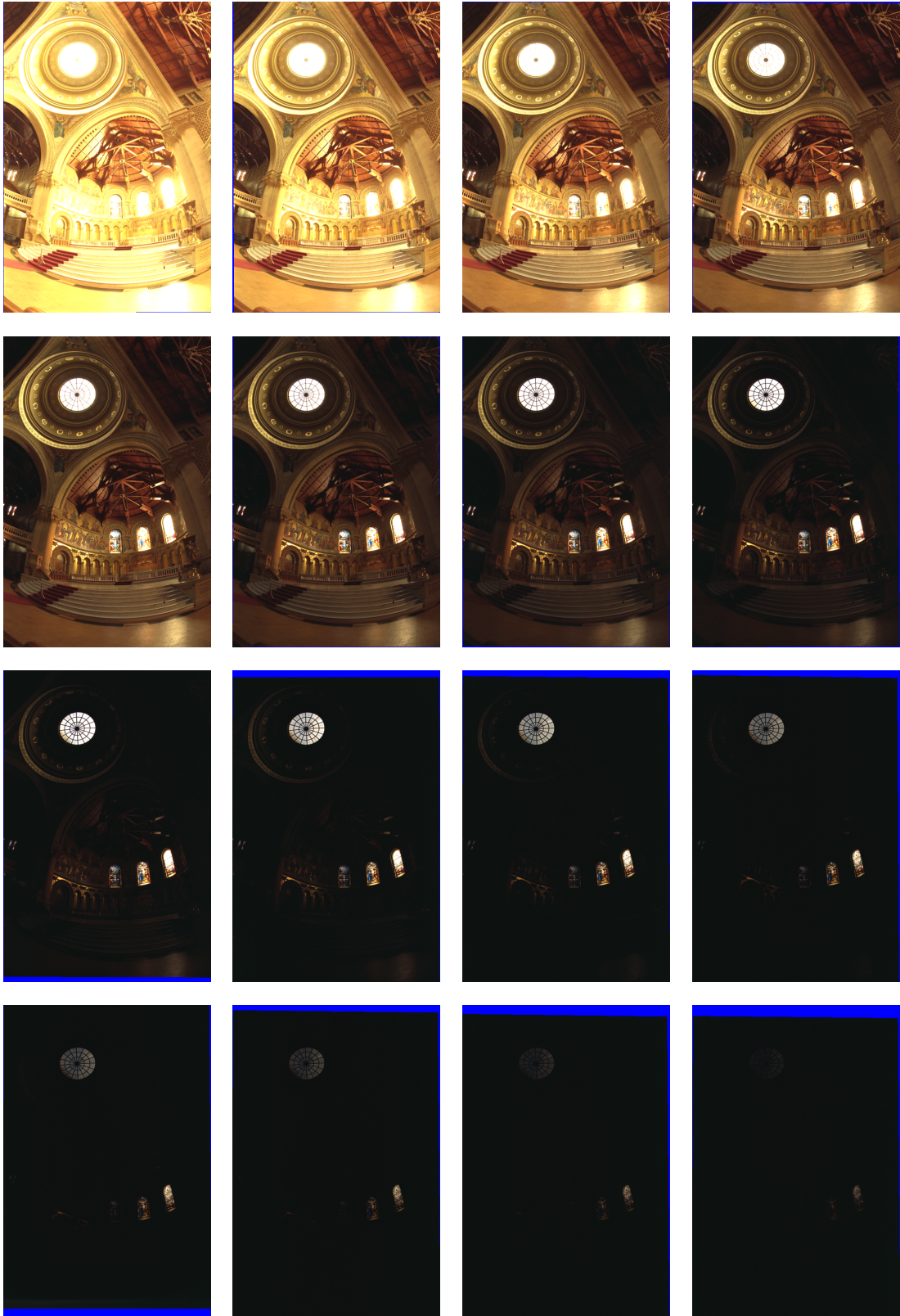
Figure 2.1: Exposure-images from a HDR sequence in [18].

2.1 Summary of HDRI Technologies

Early work aiming to increase the dynamic range of analog cameras was performed by Charles Wyckoff, who created three-layered films with each layer having a different ISO rating [94]. Each layer could be developed independently, and often false color was used to produce HDR images with color coded scene radiance. Famous examples of this were HDR images of nuclear explosions taken during hydrogen bomb testing which appeared on popular magazine covers.

With the shift to digital imaging, many authors [47, 96, 49, 50, 48, 18, 54, 22, 90, 74, 28, 29, 64, 73, 42, 10, 37] replicate Wyckoff's approach by acquiring multiple LDR exposure-images of a static scene using multiple integration times and then fusing them digitally into a single composite HDR image. These methods are referred to collectively as multi-exposure approaches [24].

One of the earliest approaches to HDRI using digital cameras is the work of Mann and Picard [49]. Integration time ratios and digital pixel value mappings between exposure-images are used to obtain a parametric camera response function, and exposure-image fusion is performed using a weighted average. Since confidence in the accuracy of a measurement is related to the derivative of the response function, the derivative of the camera response function is called the certainty function, and these certainty functions are used as weights in exposure-image fusion. Mitsunaga and Nayar [54] also use a parametric response model with model parameters determined through constrained minimization. Some improvements include the use of weighting based on SNR, the requirement that only approximate integration times be known, and the construction of a camera response function using only from pixels not affected by vignetting.

In another early approach, Debevec and Malik [18] use the reciprocity relation, integration times, and smoothness constraints to construct an objective function. Minimization of this objective function is used to construct a non-parametric camera response function. Exposure-image fusion is then performed using a weighted average of the exposure-images. Both response recovery and exposure-image fusion make use of a weighting function which gives maximum weight to middle digital pixel values and linearly decreases to zero at the maximum and minimum.

More recently, Robertson *et al.* [74] use a maximum likelihood (ML) approach for response recovery and exposure-image fusion. A non-parametric camera response function is obtained using constrained minimization. Weights used in exposure-image fusion use certainty functions computed from a smoothed response function, with smoothing performed using cubic splines.

Other works of note include Tsin *et al.* [90], which recovers a camera response function in the presence of camera white balancing. Noise sources receive thorough treatment, and radiance uncertainties are estimated based on the statistics of pooled pixels. Exposure-image fusion is performed by iteratively updating radiance estimates and using a weighting term based on the estimated noise variance. Pal *et al.* [64] develop an approach based on Bayesian estimation. This approach produces iterative estimates of scene radiance and radiance uncertainties and does not assume the camera response function must remain constant between exposure-images. Unlike

other techniques, in the work of Akyuz and Reinhard [3] the goal is not to recover a camera response function, but rather to achieve noise reduction. This is done with an approach based on frame averaging.

Lastly, Richards and Cromwell [73] devise an HDRI technique specifically for temperature measurement with infrared cameras. The camera response function of an InSb image sensor is obtained through calibration using blackbodies of controlled temperatures and emissivities. HDR images are constructed by estimating scene radiance at each pixel solely from the exposure-image with the longest integration time in which the pixel is not saturated. This approach is referred to as superframing and, although designed for use in infrared applications, Madden [47] and Yamada *et al.* [96] show that similar approaches can be successfully applied to visible light sensors.

Some of the more common and state-of-the-art HDR methods are used as benchmarks in performance analysis. These selected approaches are now presented with more detail.

Debevec and Malik

A popular approach in HDRI is the work of Debevec and Malik [18]. The key to this approach is reciprocity and postulating that the digital pixel response Z of an arbitrary pixel is a function of the product of scene radiance r and integration time T,

$$Z = f(rT), \tag{2.2}$$

where f is a camera response function mapping rT into digital pixel values. The camera response function is assumed to be identical for all sensor pixels. Since increasing values of rT are expected to result in increasing values of Z, the camera response function is assumed to be monotonic and have an inverse f^{-1}. Thus,

$$f^{-1}(Z) = rT, \tag{2.3}$$

and

$$\ln f^{-1}(Z) = \ln r + \ln T, \tag{2.4}$$

and with $g(\cdot) = \ln f^{-1}(\cdot)$,

$$g(Z) = \ln r + \ln T. \tag{2.5}$$

Solving (2.5) in the least squares sense means minimizing the sum of the squares of the errors (SSE) . Accounting for all camera pixels i and exposures-images j, the objective function O may be written as

$$O = \sum_i \sum_j \left(g(Z_{ij}) - \ln R_i - \ln T_j \right)^2 . \tag{2.6}$$

A smoothness constraint is imposed on the second derivative of g, with λ being a smoothness parameter,

$$O = \sum_i \sum_j \left(g(Z_{ij}) - \ln R_i - \ln T_j \right)^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} g''(z)^2, \tag{2.7}$$

21

where Z_{min} and Z_{max} are the minimum and maximum digital pixel values and $g''(z)$ is approximated with $g''(z) = g(z-1) - 2g(z) + g(z+1)$. It is assumed that digital pixel values near Z_{min} and Z_{max} have greater measurement errors than middle digital pixel values. Thus a weighting function w is applied, with

$$w(z) = \begin{cases} z - Z_{min}, & z \le \frac{1}{2}(Z_{min} + Z_{max}) \\ Z_{max} - z, & z > \frac{1}{2}(Z_{min} + Z_{max}). \end{cases} \tag{2.8}$$

Thus,

$$O = \sum_i \sum_j \left(w(Z_{ij}) \left[g(Z_{ij}) - \ln R_i - \ln T_j \right] \right)^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} \left[w(z)g''(z) \right]^2. \tag{2.9}$$

Since Z takes on only discrete values, g may be obtained by solving a linear least squares system with the form $\mathbf{Ax} = \mathbf{B}$ with the partitioned structure

$$\left(\begin{array}{ccc|ccc} \mathbf{0} & w(Z_{ij}) & \mathbf{0} & \mathbf{0} & -w(Z_{ij}) & \mathbf{0} \\ & \vdots & & & \vdots & \\ \hline \mathbf{0} & 1 & \mathbf{0} & & \mathbf{0} & \\ \mathbf{0} & w(Z_{ij}) & -2w(Z_{ij}) & w(Z_{ij}) \; \mathbf{0} & \mathbf{0} & \\ & \vdots & & & \vdots & \end{array} \right) \left(\begin{array}{c} g(Z_{min}) \\ \vdots \\ g(Z_{max}) \\ \hline \ln R_i \\ \vdots \end{array} \right)$$
$$= \left(\begin{array}{c} w(Z_{ij}) \ln T_j \\ \vdots \end{array} \right), \tag{2.10}$$

where $\mathbf{0}$ represents a row of zero or more zeros. It is evident how the system (2.10) corresponds to the terms in (2.9), except for the middle partitioned row in $\mathbf{A}$. Since this technique only finds relative radiance, this row sets the gauge such that $R = 0$ for the middle value Z. The over-determined system (2.10) can then be solved using a standard least squares technique such as singular value decomposition (SVD).

Of course solving (2.10) for every pixel would be impractical, so pixels must be sampled from the image such that the system is sufficiently overdetermined. Once g is obtained, $\ln r_i$ can be obtained for each pixel using a weighted average

$$\ln r_i = \frac{\sum_j w(Z_{ij})(g(Z_{ij}) - \ln T_j)}{\sum_j w(Z_{ij})}. \tag{2.11}$$

Mitsunaga and Nayar

Another common HDRI technique is the work of Mitsunaga and Nayar [54]. This technique assumes that a digital pixel value Z is related to exposure X through the camera response function f,

$$Z = f(X), \tag{2.12}$$

with the goal being to find the inverse function f^{-1} so that (2.12) can be inverted. A fundamental assumption of this approach is that f^{-1} may be modeled as a low order polynomial,

$$X = f^{-1}(Z) = \sum_{n=0}^{N} c_n Z^n, \tag{2.13}$$

22

where c_n are the polynomial coefficients and the polynomial order N is an adjustable parameter judiciously chosen using *a priori* knowledge. Different integration times T_q and T_{q+1} have the ratio $R_{q,q+1} = \frac{T_q}{T_{q+1}}$, and so

$$R_{q,q+1} = \frac{E_q}{E_{q+1}} \tag{2.14}$$

$$= \frac{f^{-1}(Z_q)}{f^{-1}(Z_{q+1})} \tag{2.15}$$

$$= \frac{\sum_{n=0}^{N} c_n Z_q^n}{\sum_{n=0}^{N} c_n Z_{q+1}^n}. \tag{2.16}$$

With Q pairs of integration times and P pixels, the sum of the squares of the errors in (2.16) is

$$O = \sum_{q=1}^{Q-1} \sum_{p=1}^{P} \left(\sum_{n=0}^{N} c_n Z_{p,q}^n - R_{q,q+1} \sum_{n=0}^{N} c_n Z_{p,q+1}^n \right)^2. \tag{2.17}$$

Proceeding to minimize O in the usual way,

$$\frac{\partial O}{\partial c_n} = 0, \tag{2.18}$$

and with $f(1) = 1$ as a choice of gauge,

$$c_N = 1 - \sum_{n=0}^{N-1} c_n. \tag{2.19}$$

The polynomial coefficients may be recovered by solving the resulting linear system which has the form $\mathbf{Ax} = \mathbf{B}$.

With

$$z(n, p, q) = Z_{p,q}^n - r_{q,q+1} Z_{p,q+1}^n, \tag{2.20}$$

each element $A_{i,j}$ of $\mathbf{A}$ can be expressed as

$$A_{i,j} = \sum_{n=0}^{N-1} \sum_{q=1}^{Q} \sum_{p}^{P} \left((z(i, p, q) - z(N, p, q)) \left(z(j, p, q) - z(N, p, q) \right) \right), \tag{2.21}$$

each element B_i of $\mathbf{B}$ may be expressed as

$$B_i = \sum_{q=1}^{Q=1} \sum_{p}^{P} z(N, p, q) \left(z(i, p, q) - z(N, p, q) \right), \tag{2.22}$$

and the vector $\mathbf{x}$ containing the polynomial coefficients can be obtained using a linear least squares approach such as SVD.

With the f^{-1} polynomial model recovered, HDRI creation proceeds by obtaining

$$X_{p,q} = f^{-1}(Z_{p,q}). \tag{2.23}$$

Then normalizing by the integration time,

$$r_{p,q} = \frac{X_{p,q}}{T_q}. \tag{2.24}$$

Radiance is then determined by the weighted average of all integration times,

$$r_p = \frac{\sum_{q=1}^{Q} w(Z_{p,q}) r_{p,q}}{\sum_{q=1}^{Q} w(Z_{p,q})}, \tag{2.25}$$

with the weighting function w defined as

$$w(Z) = \frac{f^{-1}(Z)}{f^{-1\prime}(Z)}, \tag{2.26}$$

a weighting function which is based on assigning greater weight where r is not sensitive to small variations in Z.

Robertson *et al.*

The HDRI method developed by Robertson *et al.* [74] is a probabilistic approach based on ML estimation. With Z being a digital pixel value in some exposure-image at an arbitrary pixel location, the camera response function f maps the exposure X into a digital pixel value according to

$$Z = f(X), \tag{2.27}$$

with

$$X = rT + n, \tag{2.28}$$

where r is the radiance at that pixel site, T is the integration time, and n is additive noise. Since in digital imaging there are a number of noise sources, a zero-mean Gaussian model is used for n, with a variance σ^2.

Known Response Function

Temporarily assume that f is known. In this case we use the terminology

$$X_m = f^{-1}(m) \tag{2.29}$$

to mean the exposure when the digital pixel value is m. With noise being considered, X and Z are random variables. In fact the exposure $X_{Z_{ij}}$ and digital pixel values Z_{ij} at each pixel site j in each exposure-image i are considered to be independent identically distributed (IID) random variables. Temporarily defining a weight $w = \frac{1}{\sigma^2}$, the probability P for the entire set of acquired exposure-images is

$$P \propto \exp\left(-\frac{1}{2}\sum_{i,j} w(r_j T_i - X_{z_{ij}})^2\right). \tag{2.30}$$

24

Since maximizing P is same as minimizing $-\ln P$,

$$O = \sum_{i,j} w(r_j T_i - X_{z_{ij}}).$$ (2.31)

Minimizing this objective function O with respect to the radiance r_j leads to the ML estimate $\widehat{r_j}$,

$$\widehat{r_j} = \frac{\sum_i w T_i X_{z_{ij}}}{\sum_i w T_i^2}.$$ (2.32)

Unknown Response Function

When the camera response function is not known, a Gauss-Seidel like approach is used, using (2.31) to first produce an estimate of X_m, and then using this value to estimate r_j.

To avoid the difficulty of also having to estimating the weight function, a fixed functional form is assumed,

$$w(x) = \exp\left(-\frac{4(x - \frac{V}{2})^2}{(\frac{V}{2})^2} \right),$$ (2.33)

where x ranges from the minimum pixel value 0 to the maximum pixel value V.

Based on minimizing (2.31) with respect to X_m, on iteration l the estimate $\widehat{X_m^l}$ is

$$\widehat{X_m^l} = \frac{\sum_{i,j;z_{ij}=m} w(m) T_i \widehat{r_j^{l-1}}}{\sum_{i,j;z_{ij}=m} w(m)},$$ (2.34)

where the summation is performed only over pixels whose digital pixel value is m, and $\widehat{r_j^{l-1}}$ is the estimate of r_j obtained on the previous iteration. Since $w(m)$ is constant for a specific digital pixel value m,

$$\widehat{X_m^l} = \frac{1}{N_m} \sum_{i,j;z_{ij}=m} T_i \widehat{r_j^{l-1}},$$ (2.35)

with N_m being the number of occurrences of digital pixel value m in the exposure-image set. Then minimizing (2.31) with respect to r_j,

$$\widehat{r_j^l} = \frac{\sum_i w(z_{ij}) T_i \widehat{X_{z_{ij}}^l}}{\sum_i w(z_{ij}) T_i^2}.$$ (2.36)

Iterations then continue until convergence, which is based on the rate of decrease of the objective function.

Akyuz and Reinhard

Unlike other approaches, Akyuz and Reinhard [3] do not specify a HDRI technique. Rather, they demonstrate the benefits of performing a frame averaging operation to a window of exposure-images in the exposure-image sequence.

With a known camera response function f, the digital pixel value Z at an arbitrary site in exposure-image i is given by

$$Z_i = f(rT_i), \tag{2.37}$$

where r is radiance and T_i is the integration time of the exposure-image. Thus,

$$r = \frac{f^{-1}(Z_i)}{T_i}. \tag{2.38}$$

If an average value $\bar{r}$ were to be estimating using a weighted average from N subsequent exposure-images with a known weight w,

$$\bar{r} = \frac{\sum_{i=1}^{N} \frac{f^{-1}(Z_i)w(Z_i)}{T_i}}{\sum_{i=1}^{N} w(Z_i)}. \tag{2.39}$$

For each exposure-image in the sequence, the computed r can then be replaced with $\bar{r}$ computed from the subsequent N exposure-images, leading to a new digital pixel value Z_i',

$$Z_i' = f(\bar{r}T_i), \tag{2.40}$$

with the process then repeated for each exposure-image in the sequence.

The weight function employed in this technique is fairly sophisticated, combining a traditional weight function with Hermite polynomial interpolation to decrease the weight of pixels near saturation even further. So (2.39) becomes

$$\bar{r} = \frac{\sum_{i=1}^{N} \frac{f^{-1}(Z_i)g(Z_i)}{T_i}}{\sum_{i=1}^{N} g(Z_i)}. \tag{2.41}$$

For Z ranging from 0 to 255,

$$g(Z_i) = \begin{cases} w(Z_i) & i = 1 \\ w(Z_i)\tau(Z_i) & i \neq 1 \end{cases} \tag{2.42}$$

$$\tau(Z) = \begin{cases} 1 & 0 \leq Z < 200 \\ 1 - 3h(Z)^2 + 2h(Z)^3 & 200 \leq Z < 250 \\ 0 & 250 \leq Z \leq 255 \end{cases} \tag{2.43}$$

$$h(Z) = 1 - \frac{250 - Z}{50} \tag{2.44}$$

$$w(Z) = f^{-1}(Z)f'(Z)\left(1 - \left(\frac{Z}{127.5} - 1\right)^{12}\right). \tag{2.45}$$

Richards and Cromwell

The approach of Richards and Cromwell [73] is designed for use in infrared (IR) imaging systems, and uses the principle of superframing. Unlike other approaches, for a pixel at an arbitrary location j, r_j is not estimated by combining Z_j from multiple exposure-images, but rather is estimated only using the largest non-saturated digital pixel value $Z_j^{\max}$ in the exposure-image sequence and $T_{\max}$ is the integration time of the exposure-image where $Z_j = Z_j^{\max}$. In other words,

$$Z_i^{\max} = f(r_i T_{\max}), \tag{2.46}$$

so

$$r_i = \frac{f^{-1}(Z_i^{\max})}{T_{\max}}. \tag{2.47}$$

Of course this requires knowing f and f^{-1}, which is done through an off-line calibration step. In this IR application the calibration targets are temperature controlled laboratory blackbodies, though if this technique were applied to visible solid state image sensors, a chart such as a white balance card could be used.

HDRI in Temperature Measurement Applications

Central to many HDRI techniques is the recovery of the camera response function or its inverse, where the inverse response function is then used to map pixel values back to scene radiance. Although weighted by the quantum efficiency of the image sensor, if imaging were performed using a single wavelength, the camera response function would map digital pixel values to true spectral radiances. For this reason, HDRI techniques have been used in thermographic applications.

In addition to the superframing approach of [73], HDRI techniques have been used in thermographic applications in Zhao *et al.* [100], where the method in [18] is used to estimate the spectral radiance of a combustion flame. The three-color method of non-contact temperature measurement is then used to combine these spectral radiance estimates with knowledge of blackbody radiation to generate temperature estimates. This allows the temperature profile of the flame to be generated.

While these applications show how HDRI techniques can be useful in measurement applications, most HDRI techniques [47, 96, 49, 50, 48, 18, 54, 22, 90, 74, 28, 29, 64, 42, 10, 37] are not designed for this purpose. For example, several techniques [49, 18, 54, 50, 48, 22, 90, 74, 28, 64] focus on the recovery of unknown and general camera response functions directly from scene exposure-images rather than a calibration using specific targets such as color charts, an approach desirable for computational photography with consumer cameras, but less suitable for measurement applications. Others [29, 42, 37] determine the camera response function based on a previously constructed database of camera response functions. The camera response function is assumed to be constant across the sensor. Although most HDRI techniques try to be robust in the presence of noise, few approaches [90, 64] form a noise model based on calibration. For convenience and generality, weighting is often performed with

certainty functions, which are not based on an understanding of solid state image sensor models as developed in [21] or [32].

To overcome some of these limitations, a new method of HDRI is introduced in Section 2.2, which is postulated to be better suited for use in measurement applications. This performance of this method is compared to techniques discussed previously in this chapter in Section 2.3, where objective metrics are developed to evaluate both the accuracy and precision of these techniques. We show that the Kalman filtering based approach achieves as much as a 9.4 dB improvement in SNR and as much as a 29% improvement in radiometric accuracy over other established methods.

2.2 A Kalman Filtering Approach to High Dynamic Range Imaging for Measurement Applications

Although possibly suitable for photography applications, several of the HDRI methods discussed previously in this chapter are not ideal for use in measurement applications. For example, solid state image sensors suffer from fixed pattern non-uniformity both in response and in noise power. Current approaches do not address this. The weights used in exposure-image fusion should reflect the actual noise power present in the exposure-images, whereas existing approaches are based on certainty functions. For measurement applications, the uncertainty in a measurement also needs to be estimated, a consideration neglected in existing approaches.

Some of these issues are addressed by allowing the camera response function to vary independently across pixel locations on the image sensor, perhaps making "pixel response functions" the more descriptive term. In developing this approach, first pixelwise variations in non-uniformity and measurement noise power are modeled and corrected, thus improving performance of the sensor array. Other shortcomings are addressed through the use of the well established technique of Kalman filtering, which is applied independently at each pixel location to generate radiance estimates.

Camera Calibration

As with many measurement techniques, this approach requires sensor calibration, which is performed using a spectrally flat white balance card. There are several advantages to performing calibration with a balance card. Since it has spatially uniform reflectance, it can be used to correct spatial non-uniformities which may exist due to variations in illumination, optical vignetting, or sensor non-uniformities. Since it is spectrally flat, calibration is insensitive to the type of illuminant used. The balance card defines a reference radiance, and the radiance estimates generated by this Kalman filtering-based HDRI approach are relative to this reference. Thus, reflectance measurements are not sensitive to illuminant intensity. As should be expected, changes in illumination, the optical system, or some camera settings may necessitate re-calibration.

For the HDR applications in this dissertation, we choose a camera known to have a linear camera response. A simple exposure-image taken with this digital camera is
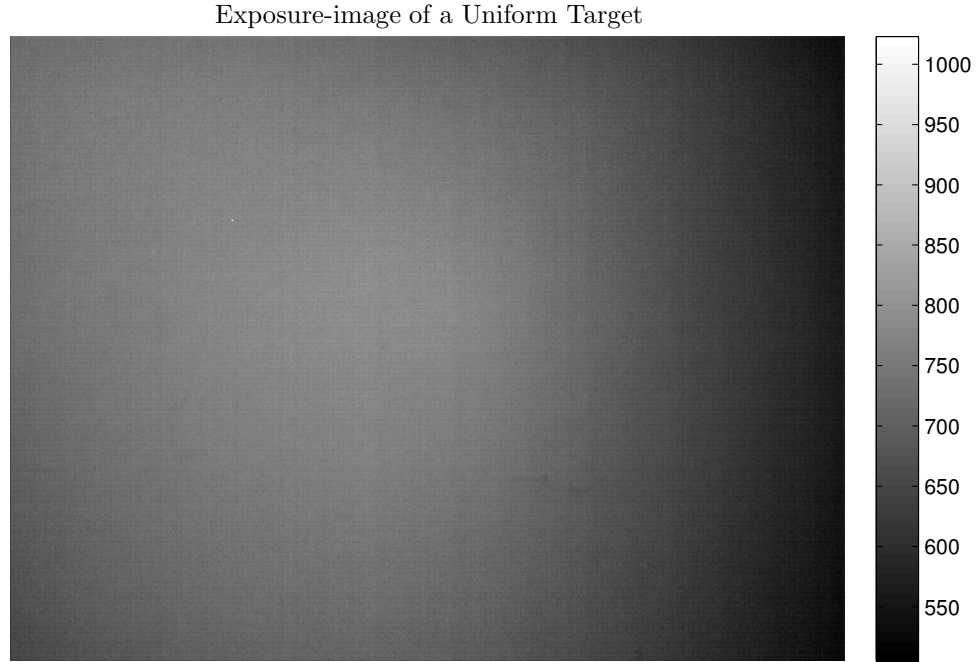
Figure 2.2: An exposure-image of a uniform target taken using a 10 bit camera. Non-uniformities are evident.

shown in Figure 2.2. Spatial non-uniformities due to lighting, vignetting fall-off, and sensor fixed pattern noises are all apparent.

Neglecting noise, the response of each pixel will have the form

$$z = ATr + B, \tag{2.48}$$

where z is the digital pixel value at a particular pixel site, T is the integration time, r is the scene radiance at this pixel location, and A and B are the parameters to be determined for each pixel through calibration. A white balance card defines the reference radiance $r = 1$ at each pixel, and by normalizing measurements to this reference, spatial non-uniformities are folded into the pixel gain and offset coefficients.

By acquiring exposure-images of the calibration card with different integration times, A and B, shown in Figure 2.3a and Figure 2.3b respectively, can be computed at each pixel location using linear regression. Acquiring repeated exposure-images of the calibration card at the same integration time allows the sample variance of z can be calculated at each pixel. Following the approach outlined in [70], knowing the sample variance of z allows A, B, and their uncertainties to be estimated. A goodness-of-fit calculation can also be performed, and in these experiments the validity of the linear pixel model was supported.

Dark current and PRNU are noise sources which can be corrected by folding them into the gain and offset terms of the pixel response functions. Once corrected, only zero-mean noise remains. This noise is due to shot noise and sensor read noise, which includes quantization noise. As seen in Figure 2.4, this noise is not uniform across

Sensor Gain (pixel values/s)

(a)

Sensor Offset (pixel values)
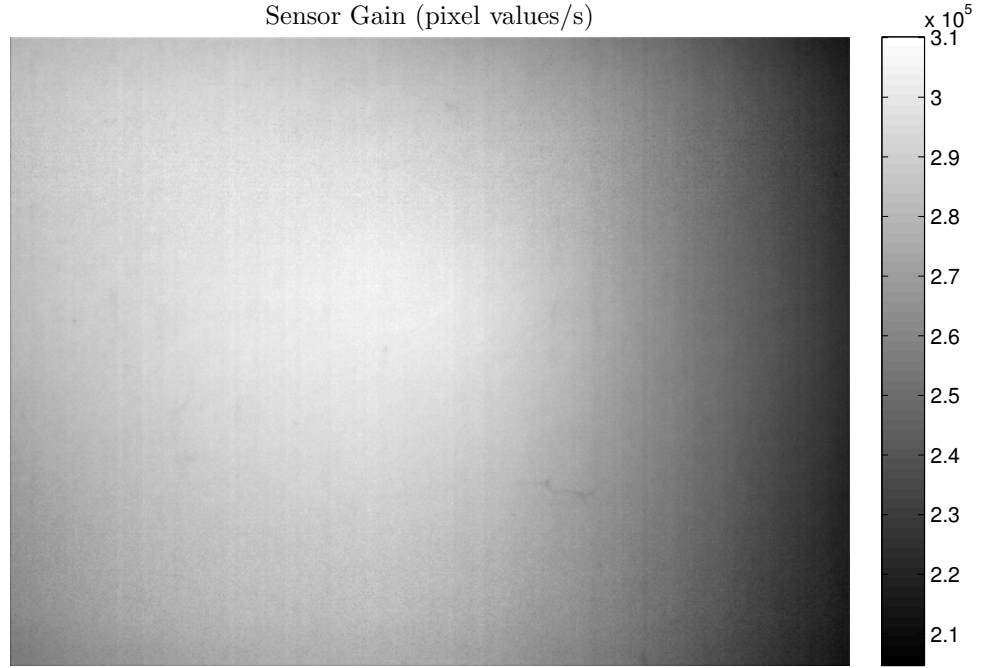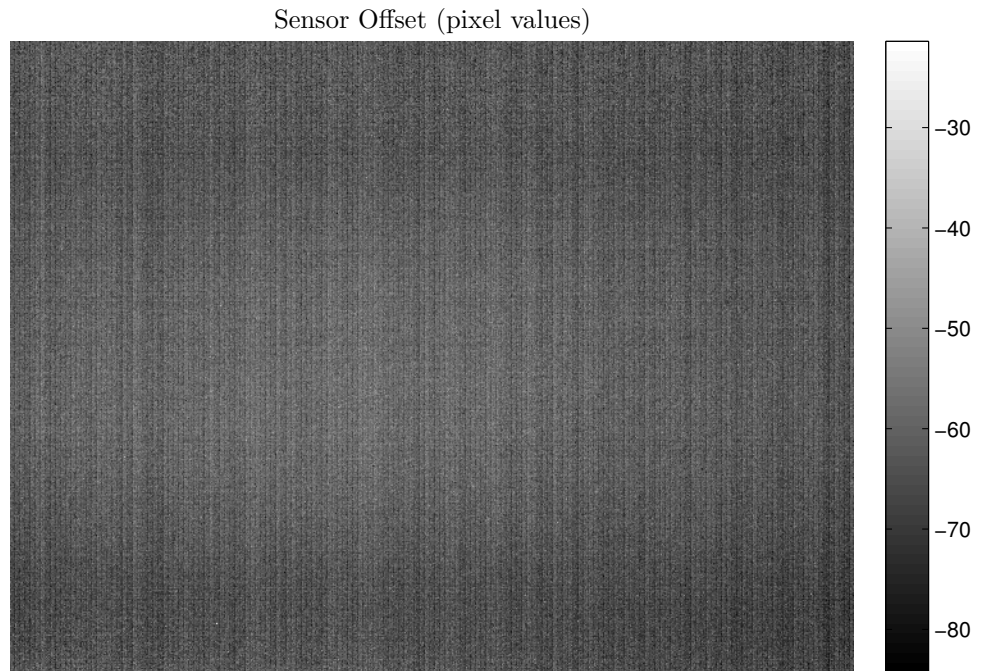
(b)

Figure 2.3: The (a) gain and (b) offset coefficients of the pixel response functions.

the sensor, motivating a pixelwise model of sensor noise power. The power of shot and read noise can be modeled as a linear function of Tr [21, 32], so at a particular

pixel location

$$R = CTr + D, \tag{2.49}$$

with the calibration card again defining $r = 1$. R is the measurement noise power and C and D, shown in Figure 2.5, are parameters to be determined at each pixel location through linear regression. Noise power can be estimated with the experimentally determined sample variance, s_z^2, computed using multiple exposure-images taken with the same integration time. As before, performing the fit according to [70] requires knowing the variance of s_z^2. This can be determined by assuming pixel outputs are corrupted by Gaussian noise of variance σ_z^2, approximating σ_z^2 with s_z^2, and applying the variance of sample variance formula given in [92]. Again, based on the uncertainties on the fitted parameters and a goodness-of-fit test, we determine the pixel noise models to be acceptable for our system.

Adding both process noise w and measurement noise n to the pixel response model,

$$z = AT(r + w) + B + n, \tag{2.50}$$

and so

$$\sigma_z^2 = A^2T^2Q + R, \tag{2.51}$$

where Q is the process noise power at a particular pixel location. If σ_z^2 is again approximated with s_z^2 and R is determined from the pixel noise model (2.49), then any residual error between s_z^2 and R can be attributed to A^2T^2Q. So process noise power can be estimated from this residual error. Although Q can be modeled as a function of signal level, a simpler approach is used in which it is assumed to be constant. It is estimated as the maximum residual error between R and s_z^2. In other words, it is assumed that there is always sufficient process noise to account for all of the discrepancies between R as determined by the model and the experimentally determined s_z^2. The magnitude of the resulting process noise power is shown in Figure 2.6. In this simple approach, any outlier in the calibration data can result in a large estimate of process noise power, which did occasionally occur in some isolated pixels. If desired, more advanced models for estimating process noise can be used and further discussion of this can be found in [76], [97], and [93], among others.

A shot noise process is described by a Poisson distribution, and, as shown in Figure 2.7, as the expected number of occurrences increases, the Poisson distribution quickly begins to resemble a Gaussian distribution. If the lowest pixel values were omitted and only the linear region of the pixel response functions were used, the system would have a linear measurement model corrupted by what is approximately additive white Gaussian noise. The Kalman filter is known to be an optimal state estimator under these conditions, so this approach performs radiance estimation based on pixel-by-pixel Kalman filtering, and directs us in our of choice of cameras to select one with a linear camera response function.

The camera selected was a Prosilica GC640 [72] with a Micron MT9V203 CMOS image sensor [53]. Since this is a monochrome camera which can operate in a fully manual mode, it eliminates the need to compensate for features such as automatic exposure control, automatic gain control, and automatic white balance. The 10-bit

raw image output helps reduce quantization noise and does not introduce compression artifacts. Integration times are controlled with microsecond resolution, and since CMOS image sensors have anti-blooming capabilities [44], blooming is not addressed in the image processing software. The GC640 also has a camera response function that is linear over nearly all of its operating range. Measurements falling outside of this linear operating region simply are not used. Although we use a monochrome camera, color camaras can be accommodated by simply applying the calibration and radiance estimation procedures to each color channel independently.

Radiance Estimation

Having selected a camera with a linear response function and using an operating region where a Gaussian noise model may be applied, we base this HDRI approach on Kalman filtering since it is known to be an optimal state estimator under these assumptions. In state estimation, the state $\mathbf{x}$ of a linear system, expressed in state space form, evolves from discrete time $k-1$ to k according to the process model

$$\mathbf{x}_k = \mathbf{\Phi}_{k-1}\mathbf{x}_{k-1} + \mathbf{\Gamma}_{k-1}\mathbf{u}_{k-1} + \mathbf{w}_{k-1}, \tag{2.52}$$

where $\mathbf{\Phi}$ governs the time evolution of the system, $\mathbf{\Gamma}$ is a weight applied to the control $\mathbf{u}$, and $\mathbf{w}$ is additive white Gaussian process noise. A measurement of the state performed at time k is given by the measurement model

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \tag{2.53}$$

where $\mathbf{z}$ is an observation vector, $\mathbf{H}$ relates state to observation, and $\mathbf{v}$ is additive white Gaussian measurement noise. When $E\left\{\mathbf{w}_k\mathbf{w}_k^T\right\} = \mathbf{Q}_k$, $E\left\{\mathbf{v}_k\mathbf{v}_k^T\right\} = \mathbf{R}_k$, and $E\left\{\mathbf{v}_k\mathbf{w}_k^T\right\} = 0$, where $E\{\cdot\}$ is the expected value operator, optimal estimates of the system state $\hat{\mathbf{x}}$ may be generated recursively with a Kalman filter, commonly expressed as

$$\begin{aligned}
\hat{\mathbf{x}}_k(-) &= \mathbf{\Phi}_{k-1}\hat{\mathbf{x}}_{k-1}(+) + \mathbf{\Gamma}_{k-1}\mathbf{u}_{k-1} & (2.54)\\
\mathbf{P}_k(-) &= \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1}(+)\mathbf{\Phi}_{k-1}^T + \mathbf{Q}_{k-1} & (2.55)\\
\mathbf{K}_k &= \mathbf{P}_k(-)\mathbf{H}_k^T\left(\mathbf{H}_k\mathbf{P}_k(-)\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} & (2.56)\\
\hat{\mathbf{x}}_k(+) &= \hat{\mathbf{x}}_k(-) + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_k(-)\right) & (2.57)\\
\mathbf{P}_k(+) &= \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_k(-)\left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)^T \\
&\quad + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T & (2.58)
\end{aligned}$$

where the $(-)$ and $(+)$ designations refer to *a priori* and *a posteriori* estimates respectively, $\mathbf{P}$ is an estimate of the covariance $E\{(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T\}$, $\mathbf{I}$ is the identity matrix, and $\mathbf{K}$ is the Kalman gain. With (2.58) expressed in Joseph form, $\mathbf{P}_k(+)$ will remain correct even if non-optimal Kalman gains are used.

The reciprocity relation states that the response of a pixel is a function of the product of radiance r and integration time T [18]. With the assumptions of a static scene, independent pixel responses, static offset and gain parameters, and lack of

control signals, the general process and measurement models given in (2.52) and (2.53) take on simpler scalar forms,

$$r_k = r_{k-1} + w_{k-1}, \tag{2.59}$$

and

$$z_k = AT_k r + B + n_k, \tag{2.60}$$

where A and B are gain and offset parameters of a particular pixel in the sensor determined through calibration. These simplifications in process and measurement models lead to simplifications in the Kalman filter which can be used to estimate the radiance at this pixel location,

$$\begin{align}
\hat{r}_k(-) &= \hat{r}_{k-1}(+) \tag{2.61} \\
P_k(-) &= P_{k-1}(+) + Q_{k-1} \tag{2.62} \\
K_k &= AT_k P_k(-) \left(A^2 T_k^2 P_k(-) + R_k\right)^{-1} \tag{2.63} \\
\hat{r}_k(+) &= \hat{r}_k(-) + K_k \left(z_k - AT_k \hat{r}_k(-) - B\right) \tag{2.64} \\
P_k(+) &= \left(1 - K_k AT_k\right)^2 P_k(-) + K_k^2 R_k. \tag{2.65}
\end{align}$$

When many exposure-images are available, initial estimates of $\hat{r}$ and P are less important, but when working with a small exposure-image sequence, starting with reasonable initial estimates will greatly increase the rate of convergence.

In applying this approach at a given pixel location, the first valid measurement is used to generate an initial state estimate by directly computing radiance using (2.48), and taking $P = \frac{R}{A^2 T^2}$, as computed from (2.63), and (2.65) with $P_k(-) = \infty$. New pixel measurements obtained from subsequent exposures-images are then fed into the scalar Kalman filter given by (2.61)-(2.65). This procedure is performed independently for each pixel, each with its own Kalman filter. The filter given by (2.61)-(2.65) is a computationally efficient way to generate recursive estimates of scene radiance and its uncertainty at a particular pixel location. Like other HDRI techniques, after certain one-time-cost initializations have been performed, constructing the HDR images requires only a simple calculation at each pixel for each exposure-image in the sequence. Thus, computational complexity is $O(NM)$, where N is the number of pixels in an exposure-image and M is the number of exposure-images in the sequence.

2.3   Performance Analysis

Since the techniques of [18] and [73] have been used in measurement applications, we seek to compare the performance of the proposed method developed in Section 2.2 to these and other methods developed in Section 2.1. In evaluating performance, an exposure-image sequence of a Gretag-Macbeth color chart, illuminated by a 60 watt incandescent source, is used. This exposure-image sequence is shown in Figure 2.8. Performance is assessed using objective metrics, with SNR used as a measure of precision and a radiance ratio test used as a measure of accuracy.

In applying [18], high frequency noise is observed in the response function when the tunable smoothness parameter λ is not sufficiently large. To avoid introducing this high frequency noise into the resulting HDR image, $\lambda = 100$ is used in all subsequent analysis. Because of prohibitive computational costs, data reduction was performed in response recovery. The response curve was constructed using 500 randomly selected pixel locations having at least 250 unique pixel values. The linear system had a rank of 1,401.

In [74], the unsmoothed, non-parametric response was found to be essentially linear. Rather than smoothing the non-parametric response with a cubic spline approximation, the response was smoothed using linear regression. This smoothing operation was required to avoid introduction of high frequency noise present in the recovered response into both certainty functions and the resulting HDR image.

In applying [54], only a first order polynomial was required to model the response function. Care was taken to select only pixels free of vignetting for response recovery. However, we note that the use of a single camera response function applied at all pixel locations will not remove any non-uniformities which may be present.

In [3], the response function obtained from [18] was used. A cluster size of 4 was used, and the weighting functions used in Hermite interpolation were scaled so they would be appropriate for our 10-bit camera.

These camera response functions are shown in Figure 2.9. These responses are similar to those presented in [100], which also appear to be linear with very little high frequency noise, though their responses appear to deviate from a linear model at the lowest pixel values. Since [100] makes no mention of non-uniformity correction having been performed, we apply [18, 74, 54, 3] directly to the input exposure-images without correction and, since these approaches do not have a measurement noise power model, that information is unused.

In applying [73], we note that infrared cameras commonly employ non-uniformity correction techniques such as calibration-based approaches [68] or scene-based methods [31, 87]. For this reason, we apply non-uniformity correction and compute radiance according to (2.48). Since [73] also does not have a measurement noise power model, this information is again unused.

In all techniques, radiance is estimated using only pixel values which fall within the same region of the camera response (pixel values between 10 and 1000 inclusively) and pixels values outside this range are not used. Consequently, not all exposure-images have usable pixel values at all pixel sites. The number of usable samples at each pixel location in the exposure-image sequence in Figure 2.8 is shown in Figure 2.10. The resulting HDR images are shown in Figure 2.11. The uncertainty estimates also generated by the Kalman approach are shown in Figure 2.12.

Differences between the HDR images in Figure 2.11 are subtle, and all methods would probably be acceptable to a human observer. However, since these methods are being considered for use in measurement applications, objective metrics should be applied. Uniformity is first assessed using SNR. Specifically, we take the mean of the radiance, $\overline{r}$, of the six fully visible patches in the first row of the input images as signal amplitude, and the standard deviation of the radiance, σ_r, is taken to be the

Table 2.1: SNR of the exposures shown in Figure 2.8.

Exposure time (ms)	Patch					
	1	2	3	4	5	6
0.5	–	–	–	–	5.4401	18.2712
1.0	–	–	–	14.2343	21.5028	26.7798
2.5	–	–	21.6881	28.8705	31.5550	34.5962
5.0	–	21.2310	29.5692	34.7955	35.6194	–
6.5	10.2332	25.6217	32.5179	37.2303	36.9629	–
8.0	15.3800	27.5444	33.8581	39.0126	–	–
15.5	24.6397	33.5083	38.2832	–	–	–
23.0	27.7790	35.6693	–	–	–	–
35.5	29.3652	38.8439	–	–	–	–
65.5	30.9778	–	–	–	–	–

Table 2.2: SNR of HDR images generated from the sequence in Figure 2.8.

	Patch					
	1	2	3	4	5	6
Debevec	30.3454	34.6026	35.4377	37.6564	34.4921	33.2228
Akyuz	29.1482	34.7759	37.7086	39.7170	37.6245	34.2089
Mitsunaga	30.7343	36.8106	36.9837	38.1079	36.3369	33.6362
Robertson	30.5381	36.6753	37.2196	39.3475	36.2155	33.8717
Richards	31.4998	40.4434	41.2946	42.6610	42.4634	38.9171
Kalman	31.5019	40.5026	41.5586	43.4604	43.9762	41.5586
Kalman (no process noise)	29.4480	40.1853	42.0776	34.9087	39.2212	39.1335

noise amplitude. The SNR is then computed as

$$\text{SNR} = 20 \log_{10} \left(\frac{\overline{r}}{\sigma_r} \right). \tag{2.66}$$

Table 2.1 lists the SNRs of the original exposure-image sequence while Table 2.2 shows the SNRs of the various HDRI techniques.

Examining Table 2.1, the SNR improves as exposure time increases until samples fall outside the usable linear region of the camera response function. This makes the weighting function in [18] seem inappropriate and could explain why the SNR of this technique is actually less than the SNR of the last available single exposure-image. With [3, 54, 74] using either the same or very similar response functions as [18], we suspect their improved SNR performance over [18] is due to both improved weighting and exposure-image fusion models. On the other hand, since [73] generates its radiance estimates using only the samples with maximum integration times prior to saturation, the improvement in SNR over the best individual exposure-

Table 2.3: The Weighted SNR of each patch for the input exposures shown in Figure 2.8.

	Patch					
	1	2	3	4	5	6
Kalman	59.1533	67.9350	69.8581	71.5991	73.2186	68.3230
Kalman (no process noise)	45.1755	64.5994	70.3566	56.0852	63.5300	67.5510

images can only be attributed to non-uniformity correction. In the Kalman filtering approach, some improvement can be achieved over [73] if the process noise considered in Figure 2.6a is applied. Kalman filtering provides as much as a 9.4 dB improvement in SNR over the classic method of [18].

It is worth noting that not all of the information provided by the Kalman filter has been utilized. In addition to radiance estimates, the Kalman filter provides estimates of the uncertainty in pixel radiance. For reasons such as non-uniformities in measurement noise power or an outlier in the exposure-image sequence, equal confidence in pixel radiance should not be assumed. Rather than giving all pixels equal weight, some applications may benefit from weighting pixels by the inverse of their Kalman filter variance estimates. This weighted average and weighted variance can be used in a weighted SNR metric, given in Table 2.3.

Looking now at the accuracy of the radiance estimates, since relative radiance is easier to obtain than absolute radiance, we have devised a radiance ratio test comparing our measured luminance values to the known CIELAB coordinates of the Gretag-Macbeth color chart. These L* values were converted to relative luminance according to [26] and averaged over each patch. Dividing these relative luminances by the lightest reference checker yielded a ratio independent of the white point normalization factor. Since radiance is proportional to luminance at every pixel, the experimentally determined radiance ratios were compared to the known luminance ratios, which served as ground truth.

Examining Table 2.4, Kalman filtering with and without process noise appears to perform roughly equivalently, as does [73]. However, [18, 3, 54, 74] all yield radiance ratios which are consistently larger than ground truth, and we see the Kalman based approach offers as much as a 29% improvement in accuracy over [18]. We also see that, for all but the darkest patch, the accuracy of both the Kalman and superframing approaches appear to be similar to that of an I1 spectrophotometer, a measurement tool commonly used in color science applications.

We note that, from examining Figure 2.12 and Figure 2.11f, it seems the uncertainty in relative radiance is closely related to the relative radiance itself. To explore this, we let $P_k(-) = P_0$, and by substituting (2.63) into (2.65) and iterating k times,

$$P_k(+) = \frac{P_0 \prod_{i=1}^{k} R_i}{P_0 A^2 \left(\prod_{i=1}^{k} R_i \right) \left(\sum_{i=1}^{k} \frac{T_i^2}{R_i} \right) + \prod_{i=1}^{j} R_i}. \tag{2.67}$$

Table 2.4: Relative reflectances of HDR images generated from the sequence in Figure 2.8.

	Patch					
	1	2	3	4	5	6
Debevec	0.0587	0.1271	0.2495	0.4500	0.6718	1.0000
Akyuz	0.0569	0.1257	0.2485	0.4487	0.6613	1.0000
Mitsunaga	0.0584	0.1267	0.2503	0.4504	0.6648	1.0000
Robertson	0.0576	0.1257	0.2466	0.4446	0.6604	1.0000
Richards	0.0481	0.1030	0.2050	0.3798	0.5946	1.0000
Kalman	0.0488	0.1045	0.2082	0.3861	0.6028	1.0000
Kalman (no process noise)	0.0494	0.1060	0.2106	0.3834	0.6054	1.0000
I1	0.0270	0.0960	0.2030	0.3780	0.6110	1.0000
Published Ratios	0.0340	0.0967	0.2098	0.3978	0.6454	1.0000

Dividing by P_0 and then letting $P_0 = \infty$ results in

$$P_k(+) = \frac{\prod_{i=1}^{k} R_i}{A^2 \left(\prod_{i=1}^{k} R_i \right) \left(\sum_{i=1}^{k} \frac{T_i^2}{R_i} \right)}, \quad (2.68)$$

which further simplifies to

$$P_k(+) = \frac{1}{A^2 \sum_{i=1}^{k} \frac{T_i^2}{R_i}}. \quad (2.69)$$

Referring back to (2.49), we experimentally verified that D is typically less than 1% of CTr such that

$$P_k(+) = \frac{C}{A^2 \sum_{i=1}^{k} T_i}. \quad (2.70)$$

So it is now evident that the uncertainties of the radiance estimates are influenced by the integration times, and the maximum integration time that can be used is determined by the scene radiance itself, explaining the similarity between Figure 2.12 and Figure 2.11f.

One common way of sampling a HDR scene is to double the integration time between each exposure-image. In this case, (2.70) simplifies to

$$P_k(+) = \frac{C}{A^2 \left(2^{k+1} - 1 \right) T_1}. \quad (2.71)$$

Here we see that samples obtained from the longest integration time alone are responsible for almost half of the reduction in variance. When employing this sampling strategy, it would not be surprising for the performance of the superframing approach to be similar to that of the Kalman filtering based approach.

Since it appears the suitability of the superframing approach is closely related to the sampling strategy employed, we consider a different sampling strategy. Figure 2.13 shows a sequence of exposure-images with a much lower dynamic range,

Table 2.5: SNR of exposures shown in Figure 2.13.

Exposure Time (ms)	Patch			
	3	4	5	6
2.0	18.0096	26.7564	30.9232	33.0210
2.0	18.3829	26.8251	30.7348	33.3043
2.0	17.5777	26.2512	30.7608	32.6487
2.5	21.6881	28.8705	31.5550	34.5962
2.5	21.6647	29.3805	31.6248	35.3026
2.5	20.5622	29.0543	31.1387	34.2279

Table 2.6: The SNR of HDR images generated from the input sequence in Figure2.13.

	Patch			
	3	4	5	6
Debevec	28.7037	33.1889	34.2265	35.2383
Akyuz	26.5613	31.4348	33.8282	35.6908
Mitsunaga	26.7198	32.1197	34.0297	36.0123
Robertson	28.7348	32.7417	34.1816	35.2534
Richards	28.2253	34.0987	36.7021	39.7394
Kalman	29.6610	37.0182	39.8893	44.2158
Kalman	35.1463	41.1453	42.6616	45.2699
(no process noise)				

and integration times are repeated. The SNRs of these input exposure-images are shown in Table 2.5. Table 2.6 shows the SNRs of the different HDRI approaches, and Table 2.7 shows the results of the radiance ratio test.

Once again, [18, 3, 54, 74] generally did not achieve the same performance as [73] or Kalman approaches in either SNR or radiance ratio. While both [73] and Kalman filtering appear to perform similarly in radiance ratio, improvements in SNR can be seen in the Kalman filtering approach which are, once again, very much dependent on the process noise power assumed to be present. We attribute this improvement in SNR to the data smoothing properties of the Kalman filter. Figure 2.14 and Figure 2.15 show the relative radiance estimates generated for an individual pixel across the exposure-image sequence. Data smoothing can be clearly seen, especially in the absence of process noise. Without process noise, the reduction in radiance uncertainty due to increased sampling is also apparent.

In summary, while many HDRI applications may be suitable for use in computational photography, they are not appropriate for use in measurement and machine vision applications. As such, a Kalman filtering based approach was developed in Section 2.2, and compared to several other HDRI approaches in Section 2.3, with the Kalman filter based approach generally offering improved performance. The importance of sensor calibration has been shown, and the influence of sampling

Table 2.7: Relative reflectances of HDR images generated from the sequence in Figure 2.13.

			Patch			
	1	2	3	4	5	6
Debevec	–	–	0.2509	0.4548	0.6754	1.0000
Akyuz	–	–	0.2310	0.4438	0.6724	1.0000
Mitsunaga	–	–	0.2133	0.4308	0.6641	1.0000
Robertson	–	–	0.2498	0.4547	0.6754	1.0000
Richards	–	–	0.2053	0.3913	0.6177	1.0000
Kalman	–	–	0.2027	0.3880	0.6150	1.0000
Kalman (no process noise)	–	–	0.2056	0.3912	0.6184	1.0000
I1	0.0270	0.0960	0.2030	0.3780	0.6110	1.0000
Published Ratios	0.0340	0.0967	0.2098	0.3978	0.6454	1.0000

strategy on HDRI performance has been explored. It is likely no single HDRI technique will prove ideal for all applications. However, the maturity and flexibility of state estimation techniques is likely to help facilitate the use of HDRI in a variety of measurement applications.
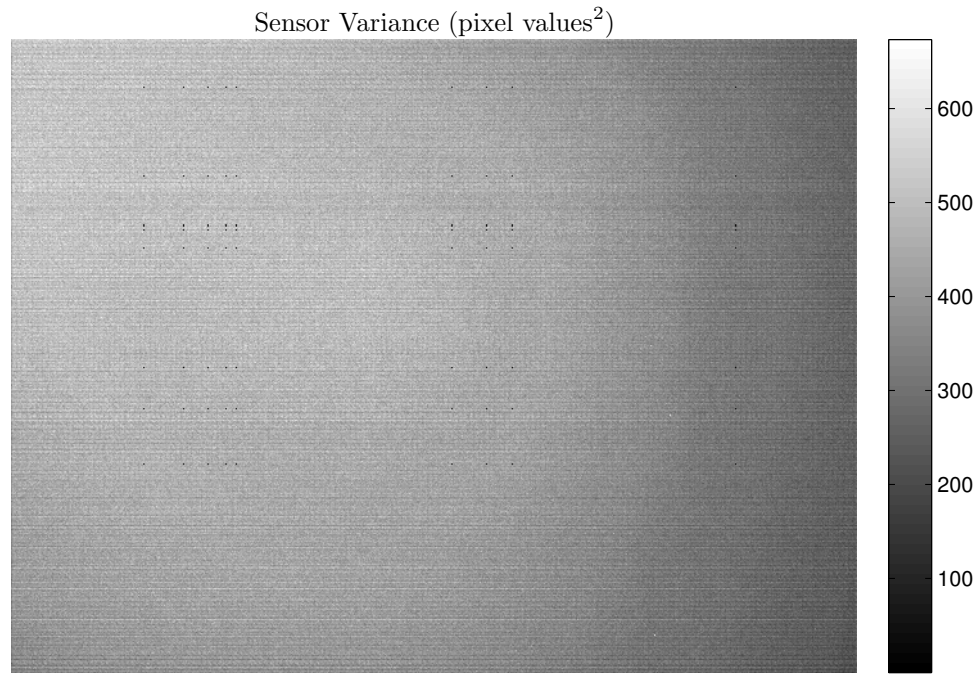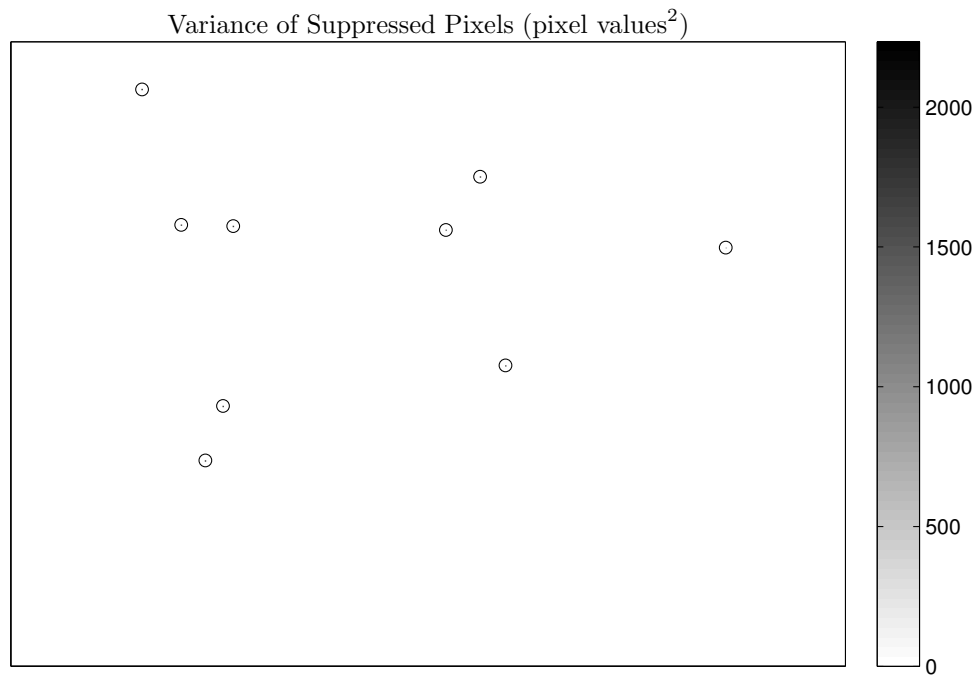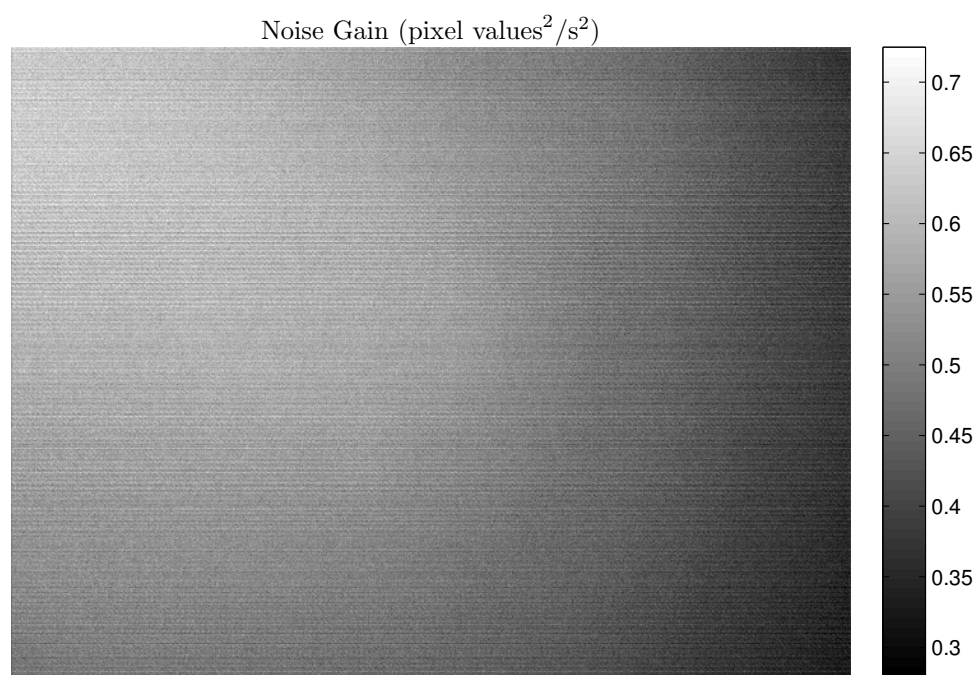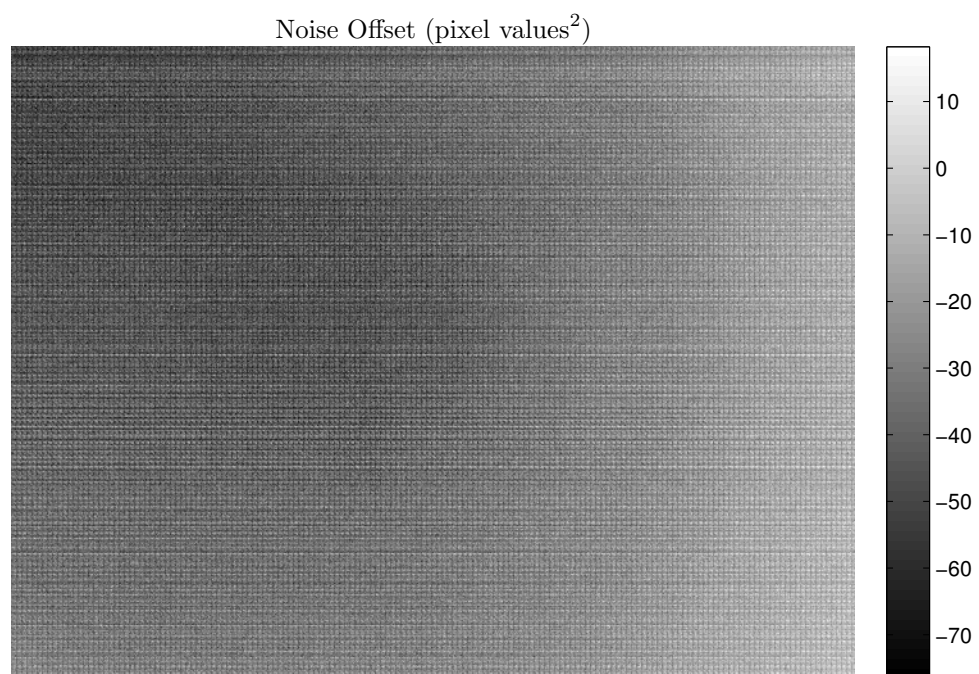
Figure 2.4: The (a) variance of a sequence of 49 exposure-images taken with identical camera parameters. This non-uniformity motivates a pixelwise measurement noise power model. Shown in (b) are a few isolated pixels with large variances which were suppressed in (a).
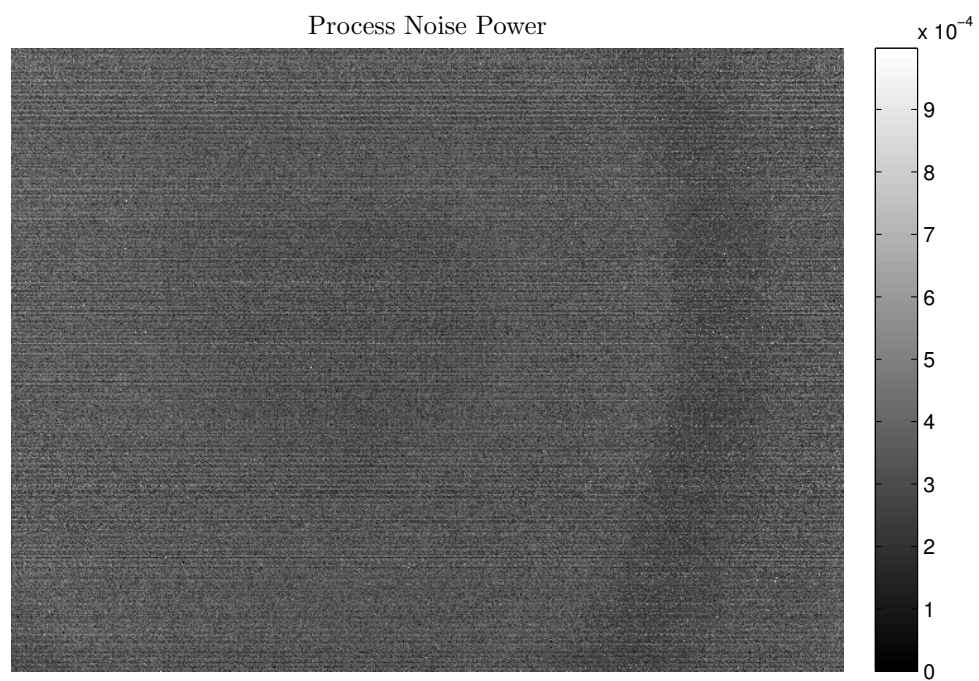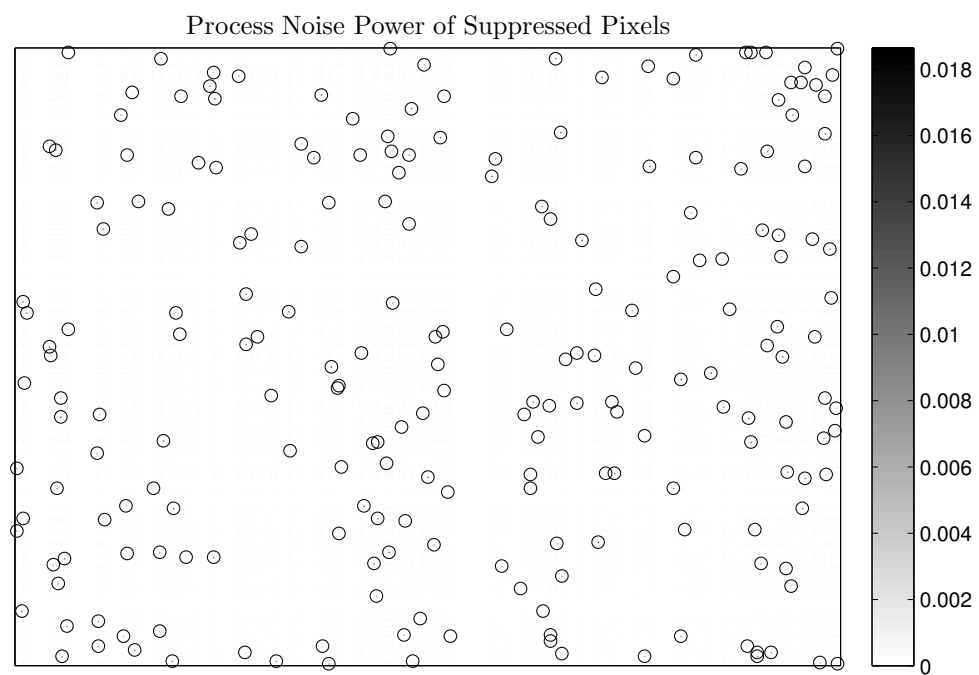
Noise Gain (pixel values^2/s^2)

(a)

Noise Offset (pixel values^2)

(b)

Figure 2.5: The (a) gain and (b) offset coefficients of the measurement noise power model.

Process Noise Power

(a)

Process Noise Power of Suppressed Pixels

(b)

Figure 2.6: The (a) estimated process noise power Q. Like r, Q is also normalized by the reference radiance. Shown in (b) are isolated pixels whose process noise power was suppressed in (a).
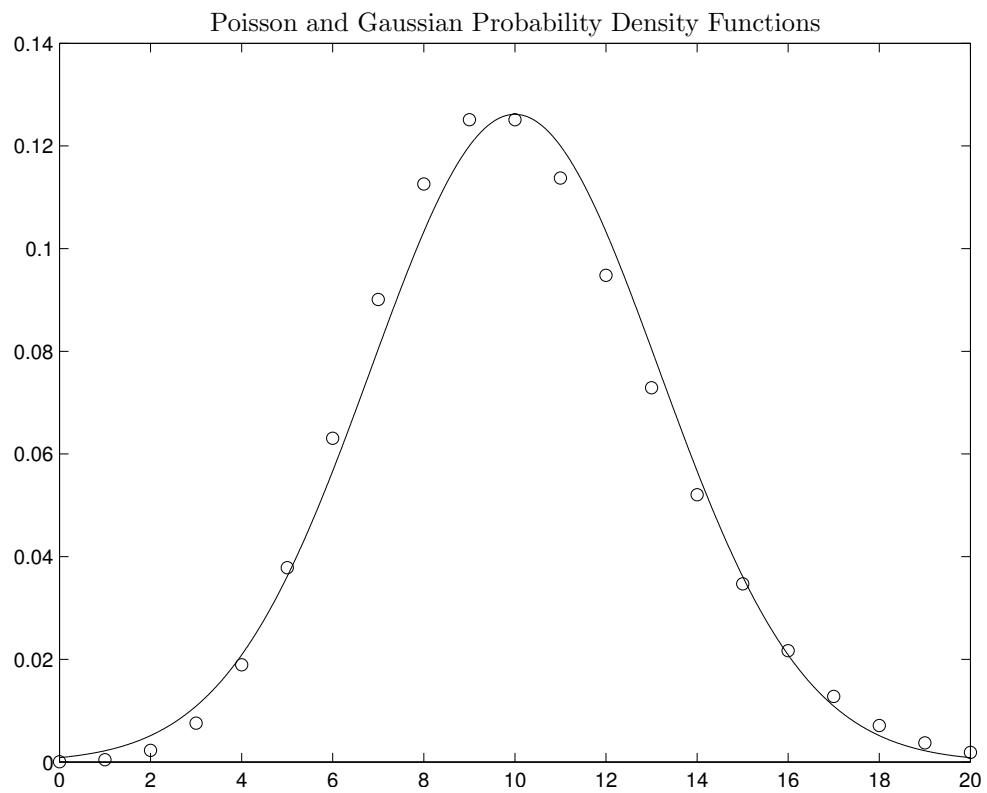
Figure 2.7: The probability density function of a Poisson distribution (open circles) overlaid with the probability density function of a Gaussian distribution. Both distributions have a mean and variance of 10.
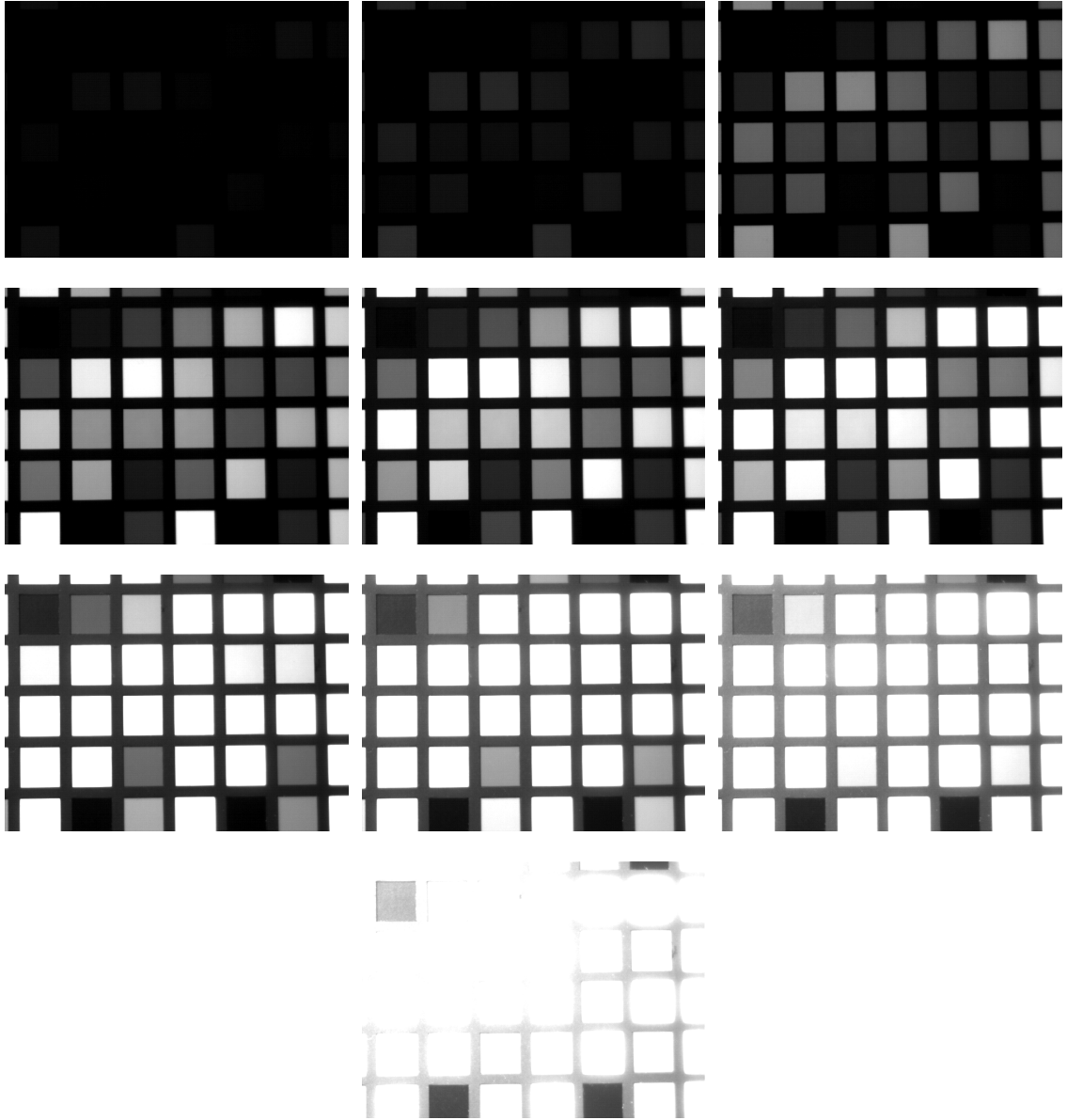
Figure 2.8: The exposure-images used as inputs to HDRI algorithms. Integration times were 0.5, 1.0, 2.5, 5.0, 6.5, 8.0, 15.5, 23.0, 35.5, and 65.5 ms.
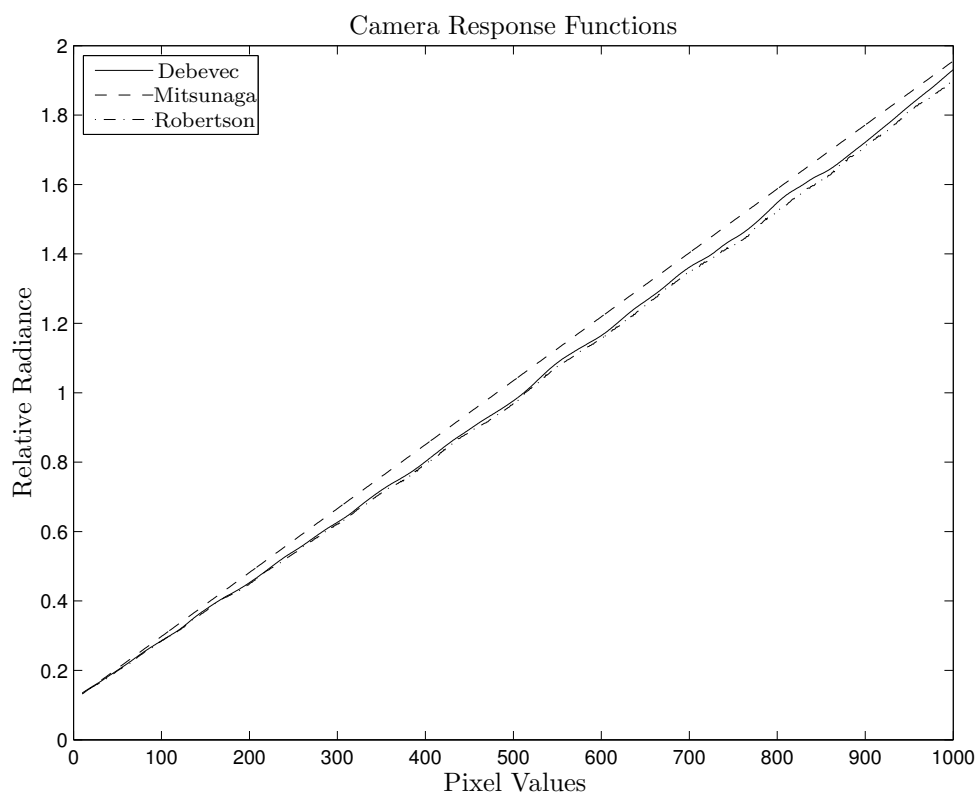
Figure 2.9: The camera response functions computed by the Debevec, Mitsunaga, and Robertson methods.
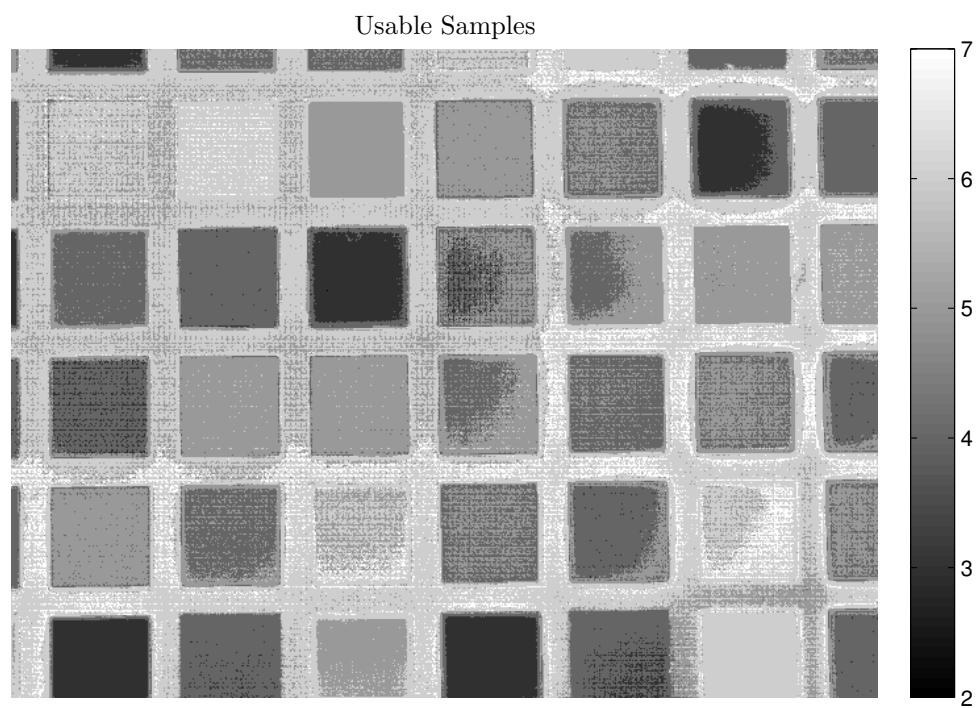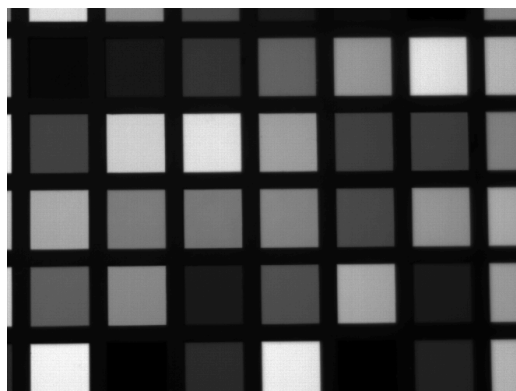
Figure 2.10: The number of usable samples in the exposure-image sequence at each pixel site.
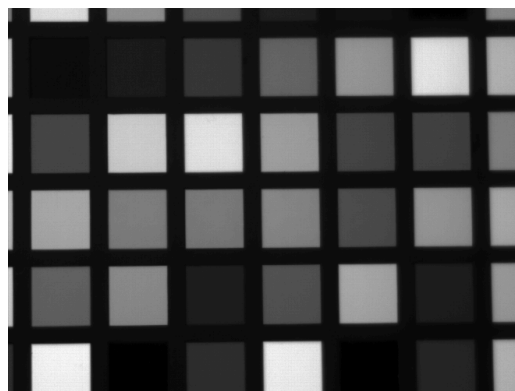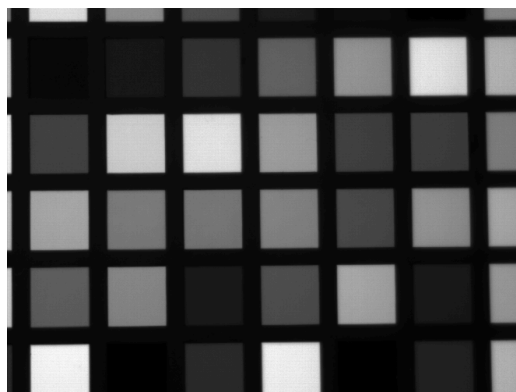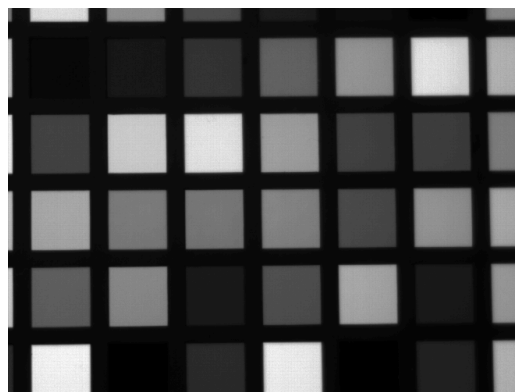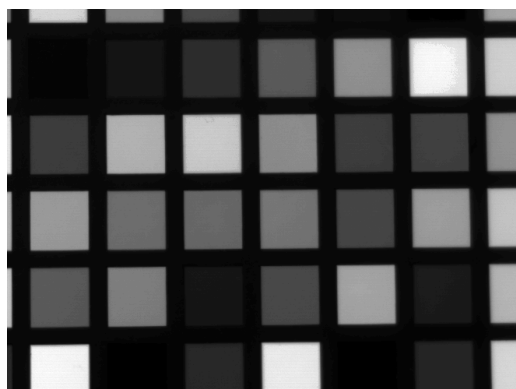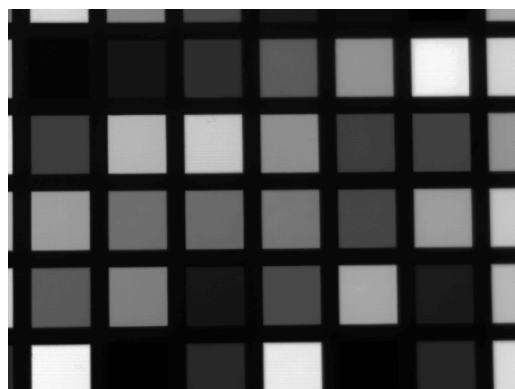
Figure 2.11: Radiance estimates generated using (a) Debevec, (b) Akyuz, (c) Robertson, (d) Mitsunaga, (e) Richards, and (f) Kalman filtering.
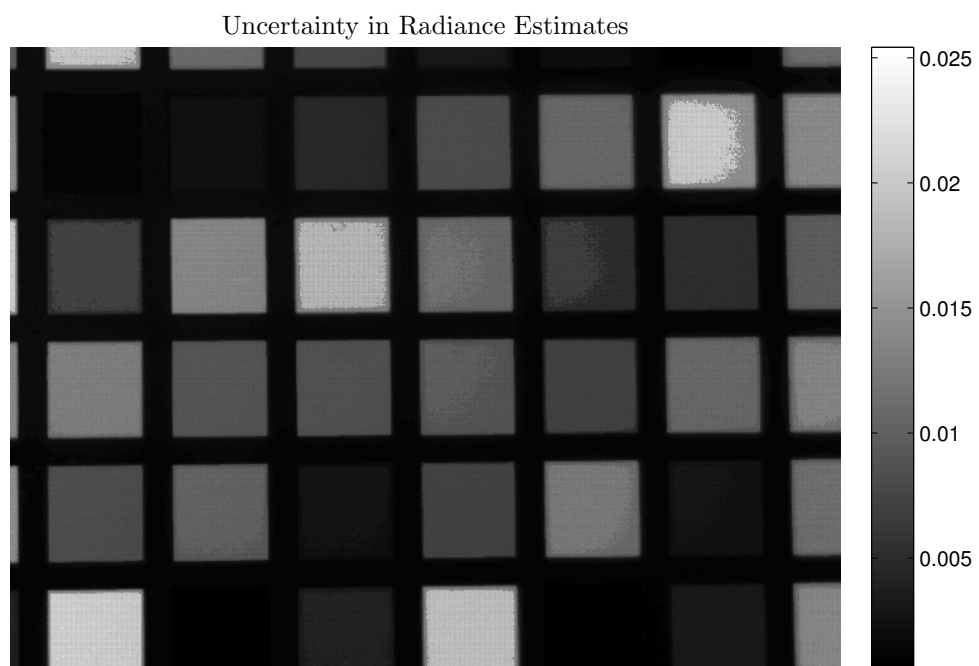
Figure 2.12: Estimates of the uncertainty in relative radiance generated by the Kalman filtering approach.
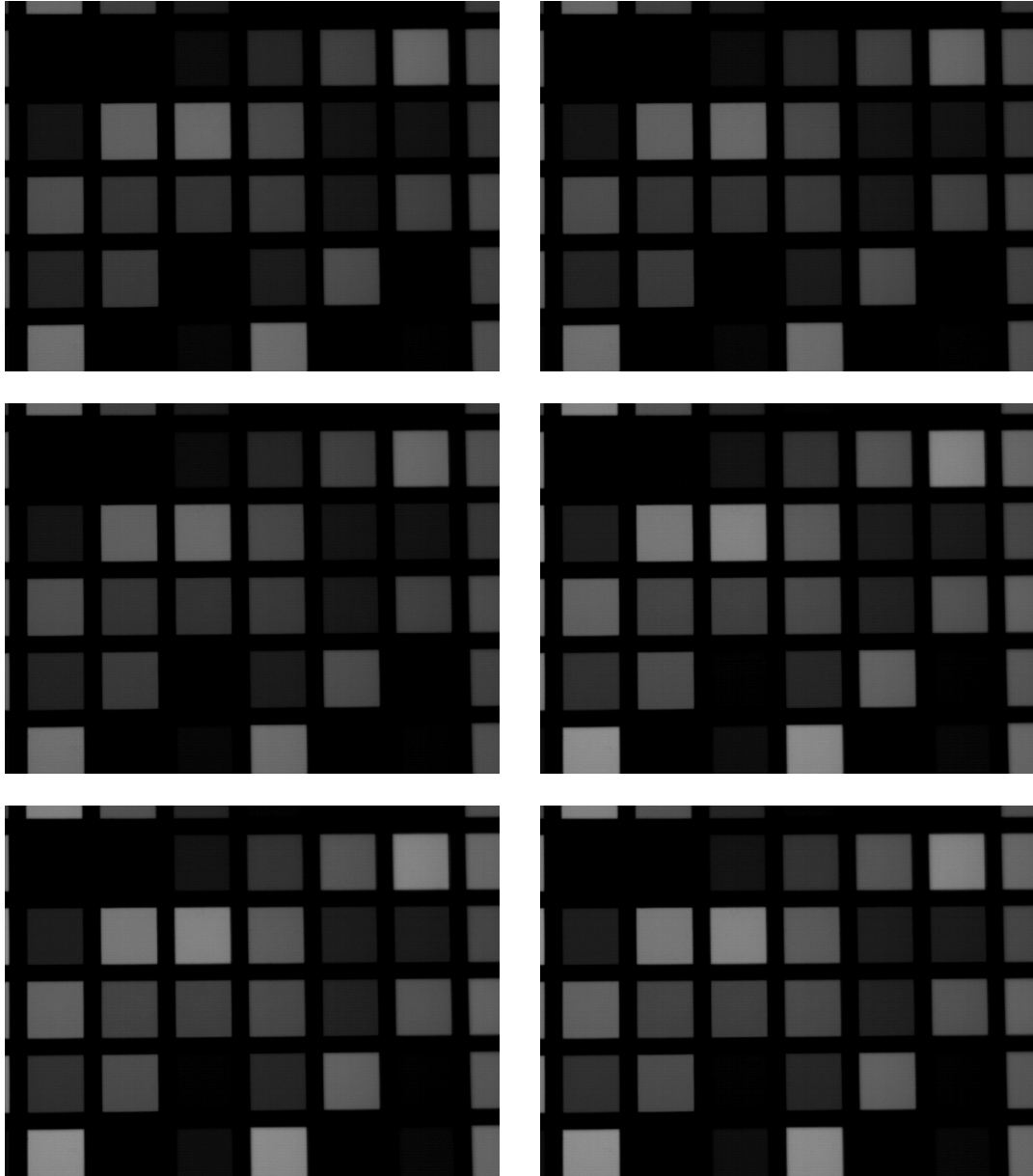
Figure 2.13: A second set of exposure-images used as inputs to the different HDRI techniques. Integration times are 2, 2, 2, 2.5, 2.5, and 2.5 ms.
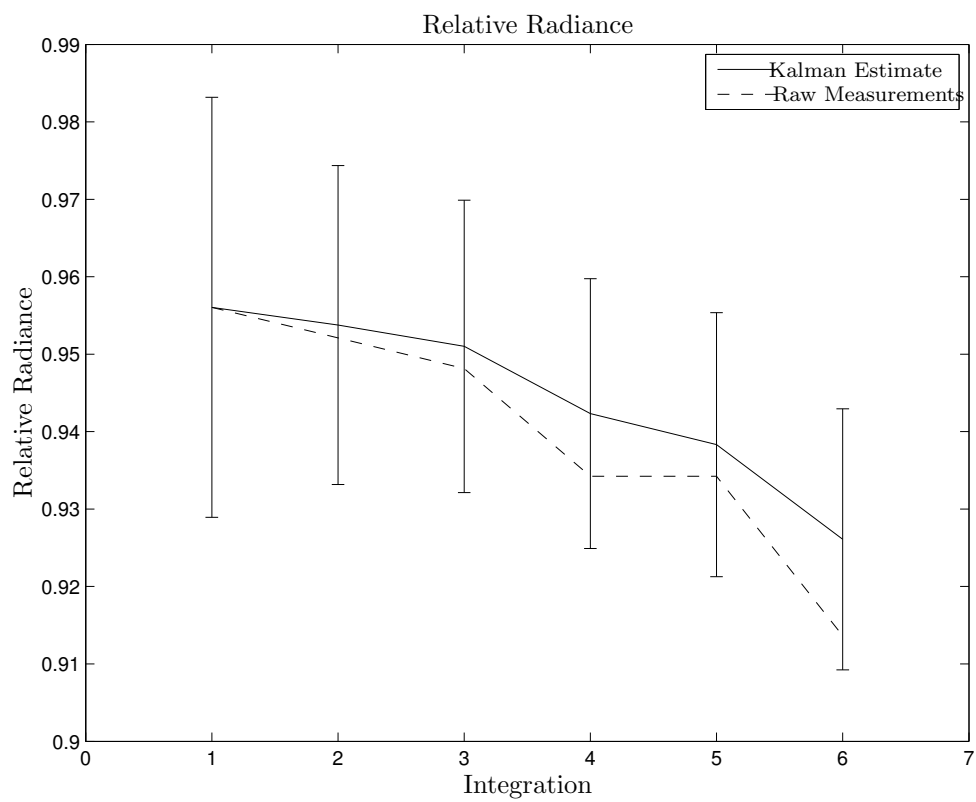
Figure 2.14: Relative radiance estimates of the Kalman filtering approach.
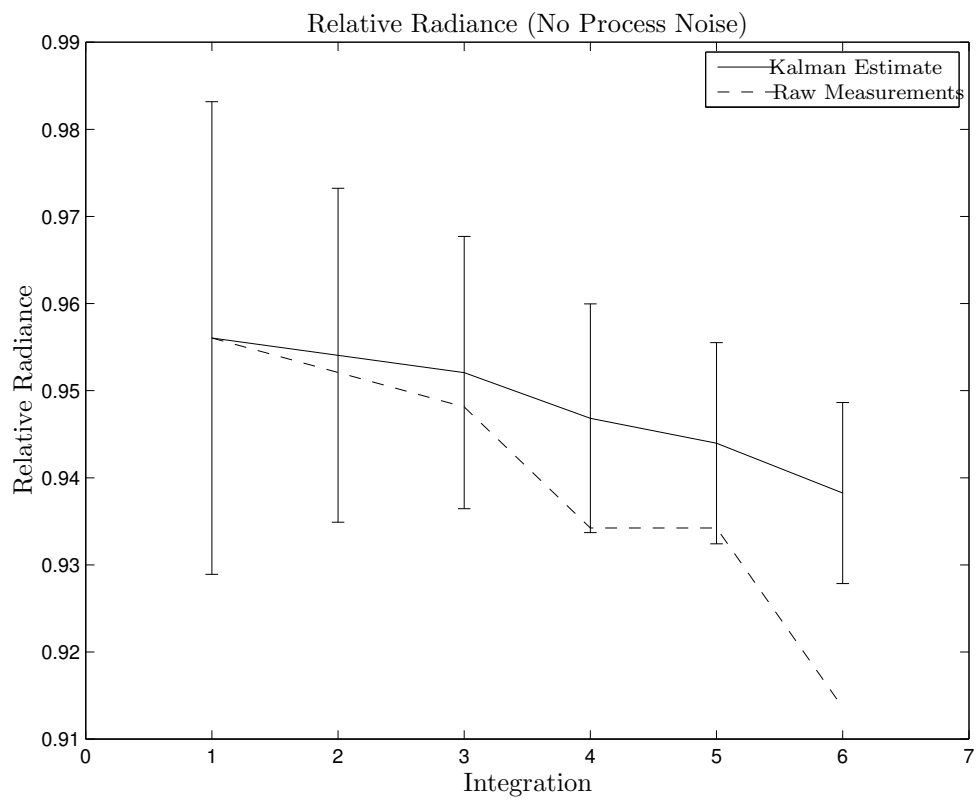
Figure 2.15: Relative radiance estimates of the Kalman filtering approach with no process noise.

Chapter 3 Detecting Multipath Scattering in PMP

PMP is a highly successful and commonly employed technique used to perform 3D surface reconstruction. However, defects can occur when multipath interference is present in a scene. Multipath interference may result from interreflections, or from other sources. For example, when there is an abrupt discontinuity in surface height, such as at a step edge or occlusion, due to fractional pixel filling a pixel may be partially illuminated by both foreground and background surfaces. Both foreground and background surfaces may also illuminate a pixel if some motion occurs during scanning. Consequently, multipath interference can be difficult to anticipate and control. The resulting distortions can be subtle, raising the possibility that these effects may be undetected during scanning.

After introducing range reconstruction and PMP, we build on a quality metric used in [77], and demonstrate that multipath scattering can be both identified and quantified in PMP scanning.

Range Reconstruction

In SLI, a range sensor is formed using a calibrated camera-projector system. When a pinhole model holds, the relationship between camera coordinates and world coordinates is

$$\mathbf{X_c} = \mathbf{M^{wc}}\mathbf{X_w}, \tag{3.1}$$

with

$$\mathbf{X_c} = \begin{pmatrix} x_c \\ y_c \\ w_c \end{pmatrix}, \tag{3.2}$$

and

$$\mathbf{X_w} = \begin{pmatrix} x_w \\ y_w \\ z_w \\ 1 \end{pmatrix}. \tag{3.3}$$

The 3D world coordiates are (x_w, y_w, z_w), the camera coordinates are (x_c, y_c), w_c is a scale factor, and $\mathbf{M^{wc}}$ is a 3x4 matrix determined through calibration. Similarly, projector coordinates and world coordinates are related according to

$$\mathbf{X_p} = \mathbf{M^{wp}}\mathbf{X_w}, \tag{3.4}$$

with

$$\mathbf{X_c} = \begin{pmatrix} x_p \\ y_p \\ w_p \end{pmatrix}, \tag{3.5}$$

where (x_p, y_p) are projector coordinates, w_p is a scale factor, and $\mathbf{M^{wp}}$ is another 3x4 matrix determined through calibration.

From (3.1) and (3.4),

$$x_c = \frac{m_{11}^{wc}x_w + m_{12}^{wc}y_w + m_{13}^{wc}z_w + m_{14}^{wc}}{m_{31}^{wc}x_w + m_{32}^{wc}y_w + m_{33}^{wc}z_w + m_{34}^{wc}} \tag{3.6}$$

$$y_c = \frac{m_{21}^{wc}x_w + m_{22}^{wc}y_w + m_{23}^{wc}z_w + m_{24}^{wc}}{m_{31}^{wc}x_w + m_{32}^{wc}y_w + m_{33}^{wc}z_w + m_{34}^{wc}} \tag{3.7}$$

$$x_p = \frac{m_{11}^{wp}x_w + m_{12}^{wp}y_w + m_{13}^{wp}z_w + m_{14}^{wp}}{m_{31}^{wp}x_w + m_{32}^{wp}y_w + m_{33}^{wp}z_w + m_{34}^{wp}} \tag{3.8}$$

$$y_p = \frac{m_{21}^{wp}x_w + m_{22}^{wp}y_w + m_{23}^{wp}z_w + m_{24}^{wp}}{m_{31}^{wp}x_w + m_{32}^{wp}y_w + m_{33}^{wp}z_w + m_{34}^{wp}}. \tag{3.9}$$

With

$$\mathbf{C} = \begin{pmatrix} m_{11}^{wc} - m_{31}^{wc}x_c & m_{12}^{wc} - m_{32}^{wc}x_c & m_{13}^{wc} - m_{33}^{wc}x_c \\ m_{21}^{wc} - m_{31}^{wc}y_c & m_{22}^{wc} - m_{32}^{wc}y_c & m_{23}^{wc} - m_{33}^{wc}y_c \\ m_{11}^{wp} - m_{31}^{wp}x_p & m_{12}^{wp} - m_{32}^{wp}x_p & m_{13}^{wp} - m_{33}^{wp}x_p \\ m_{21}^{wp} - m_{31}^{wp}y_p & m_{22}^{wp} - m_{32}^{wp}y_p & m_{23}^{wp} - m_{33}^{wp}y_p \end{pmatrix} \tag{3.10}$$

and

$$\mathbf{D} = \begin{pmatrix} m_{34}^{wc}x_c - m_{14}^{wc} \\ m_{34}^{wc}y_c - m_{24}^{wc} \\ m_{34}^{wp}x_p - m_{14}^{wp} \\ m_{34}^{wp}y_p - m_{24}^{wp} \end{pmatrix}, \tag{3.11}$$

we see that world coordinates can be recovered as the solution to

$$\mathbf{C} \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} = \mathbf{D}. \tag{3.12}$$

Since (3.12) is overdetermined, it can be solved using techniques such as pseudoinverse, but the system can also be solved with only one of x_p or y_p, so scanning need only be performed in one dimension, thus reducing the number of required patterns.

PMP

In PMP, a scene is illuminated by a projected sequence of 1D sinusoidal patterns P_n. Assuming scanning is performed in the x direction,

$$P_n = A \sin\left(2\pi f x_p - 2\pi \frac{n}{N}\right) + A, \tag{3.13}$$

where A represents both the DC offset and the amplitude of the projected sinusoid, f represents the spatial frequency of the projected pattern, x_p is the projector column, n is the pattern number, and N is the total number of patterns in the sequence. In this paper we use the convention that f is an integer, with it understood to be divided by the projector width. In the literature, sometimes the quantity $2\pi\frac{n}{N}$ is termed the phase of the sinusoid, with PMP being referred to as a phase shifting technique. It is also common to refer to $2\pi f x_p$ as the phase of the sinusoid, and this is the convention which is used in this dissertation.

For each projected pattern P_n, a camera pixel measurement y_n must be obtained for each camera pixel location. The radiance at an arbitrary location in the scene is

$$BA \sin \left(2\pi f x_p - 2\pi \frac{n}{N}\right) + BA + F, \tag{3.14}$$

where B is the albedo and F is the ambient background radiance at that scene location. With a linear camera response and background intensity subtraction, the camera pixel measurement at an arbitrary location is

$$y_n = C \sin \left(2\pi f x_p - 2\pi \frac{n}{N}\right) + C, \tag{3.15}$$

with C being magnitude in camera pixel values. The camera pixel measurement y_n may correspond to raw data from the camera, or it may be obtained using HDR imaging techniques, which may be advisable since these techniques can perform noise reduction, eliminate saturation and quantization effects, and compensate for a non-linear camera response.

In order to perform surface reconstruction, x_p must be determined for each camera pixel location. Expanding (3.15),

$$
\begin{aligned}
y_n &= C \sin \left(2\pi f x_P\right) \cos \left(-2\pi \frac{n}{N}\right) + \\
& \quad C \cos \left(2\pi f x_p\right) \sin \left(-2\pi \frac{n}{N}\right) + C.
\end{aligned}
\tag{3.16}
$$

The measurement sequence at this camera pixel location can be expressed as

$$\mathbf{Y} = \mathbf{AX}, \tag{3.17}$$

where

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}, \tag{3.18}$$

$$\mathbf{A} = \begin{pmatrix} \cos\left(-2\pi\frac{1}{N}\right) & \sin\left(-2\pi\frac{1}{N}\right) & 1 \\ \cos\left(-2\pi\frac{2}{N}\right) & \sin\left(-2\pi\frac{2}{N}\right) & 1 \\ \vdots & \vdots & \vdots \\ \cos\left(-2\pi\frac{N}{N}\right) & \sin\left(-2\pi\frac{N}{N}\right) & 1 \end{pmatrix}, \tag{3.19}$$

and

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = \begin{pmatrix} C \sin\left(2\pi f x_p\right) \\ C \cos\left(2\pi f x_p\right) \\ C \end{pmatrix}. \tag{3.20}$$

Assuming a sufficient number of patterns are used, (3.17) can be solved using standard techniques such as pseudoinverse, and x_p can then be found according to

$$x_p = \frac{1}{2\pi f} \arctan \left(\frac{X_1}{X_2}\right), \tag{3.21}$$

noting that phase unwrapping must also be used if $f > 1$. Many approaches to phase unwrapping have been developed, and some resources on the subject can be found in [34, 6, 52]. Having determined x_p, (3.12) may be solved and 3D world coordinates obtained for each camera pixel.

3.1 Identifying Multipath Effects in PMP

Absent multipath scattering, the state estimates obtained using PMP can be arranged to form an invariant quantity. Examining (3.20),

$$
\begin{aligned}
\frac{X_1^2 + X_2^2}{X_3^2} &= \frac{C^2 \sin^2 (2\pi f x_p) + C^2 \cos^2 (2\pi f x_p)}{C^2} \\
&= 1.
\end{aligned}
\tag{3.22}
$$

As phase is varied (3.22) will map to the unit circle, so we refer to this quality measure as the unit circle metric. In [77], deviations from unity are used to assess the quality of the sinusoidal fit.

Suppose instead of being illuminated by a single projector column x_p, multipath scattering caused a camera pixel to be illuminated by two projector columns. With x_1 and x_2 the projector columns and with C and D the response magnitudes, following the PMP process would result in the estimated state

$$
\mathbf{X} = \begin{pmatrix} C \sin (2\pi f x_1) + D \sin (2\pi f x_2) \\ C \cos (2\pi f x_1) + D \cos (2\pi f x_2) \\ C + D \end{pmatrix}.
\tag{3.23}
$$

The numerator of the unit circle metric would be

$$
\begin{aligned}
X_1^2 + X_2^2 &= (C \sin (2\pi f x_1) + D \sin (2\pi f x_2))^2 + \\
&\quad (C \cos (2\pi f x_1) + D \cos (2\pi f x_2))^2 \\
&= C^2 + D^2 + 2CD \cos (2\pi f (x_1 - x_2)).
\end{aligned}
\tag{3.24}
$$

With a third projector column x_3 of camera pixel intensity E detected, the PMP process would estimate the state

$$
\mathbf{X} = \begin{pmatrix} C \sin (2\pi f x_1) + D \sin (2\pi f x_2) + E \sin (2\pi f x_3) \\ C \cos (2\pi f x_1) + D \cos (2\pi f x_2) + E \cos (2\pi f x_3) \\ C + D + E \end{pmatrix},
\tag{3.25}
$$

and the numerator of the unit circle metric would be

$$
\begin{aligned}
X_1^2 + X_2^2 &= C^2 + D^2 + E^2 + \\
&\quad 2CD \cos (2\pi f (x_1 - x_2)) + \\
&\quad 2CE \cos (2\pi f (x_1 - x_3)) + \\
&\quad 2DE \cos (2\pi f (x_2 - x_3)).
\end{aligned}
\tag{3.26}
$$

A similar structure exists for any number of multipaths. In addition to the unit circle metric, $X_1^2 + X_2^2$ is expected to remain constant as frequency is varied unless multipath interference is present. We refer to this quantity as the magnitude-constancy metric, and it allows the presence of multipath interference to be detected by performing PMP using multiple spatial frequencies.

In summary, from (3.24) and (3.26) we see that the presence of signal corruption can be identified in a PMP scan with the use of the unit circle metric. Multipath interference can be specifically identified by examining the magnitude-constancy metric obtained from PMP scans at multiple spatial frequencies. Furthermore, since (3.24) resembles the classic beats formula $A\cos 2\pi(f_2 - f_1)t$, by varying spatial frequency in fixed intervals, it would be possible to estimate the difference between the projector coordinates of the two paths using the fast Fourier transform (FFT).

3.2   Experiments and Results

Having introduced PMP quality metrics, we now discuss experiments designed to detect and quantify multipath interference. In these experiments, PMP is performed using a calibrated camera-projector system [33, 72] using the Kalman filtering based HDRI method developed in Chapter 2. As discussed previously, since only one projector coordinate needs to be determined to perform surface reconstruction, scanning is performed only in the x_p direction.

Interreflections

The scene being imaged consists of a mirror placed in front of a flat surface with an interreflection clearly visible. Magnitude, phase, and unit circle metric from a scan using $f = 1$ are shown in Figure 3.1.

PMP is also performed while f is varied from 4 to 64 in steps of 4. Some of the resulting magnitude images are shown in Figure 3.2, some of the resulting phase images are shown in Figure 3.3, and some of the unit circle metric images are shown in Figure 3.4.

With PMP performed at 16 evenly spaced frequencies, the FFT is performed in order to examine the magnitude-constancy metric. The magnitudes of the first few FFT bins are shown in Figure 3.5, and the greatest non-DC component is shown in Figure 3.6.

Sub-Pixel Multipath Effects

This experiment aims to detect and quantify multipath scattering due to sub-pixel imaging, where a foreground object and a background object both fill only a fraction of the same pixel. The foreground object in the scene is a fine thread, less than a single pixel thick along most of its length, and the background object is a solid white foam board. The distance between foreground and background objects increases from bottom to top. Magnitude, phase, and unit circle metric from a scan using $f = 1$ are shown in Figure 3.7, with an area of thread and shadow enlarged in Figure 3.8.

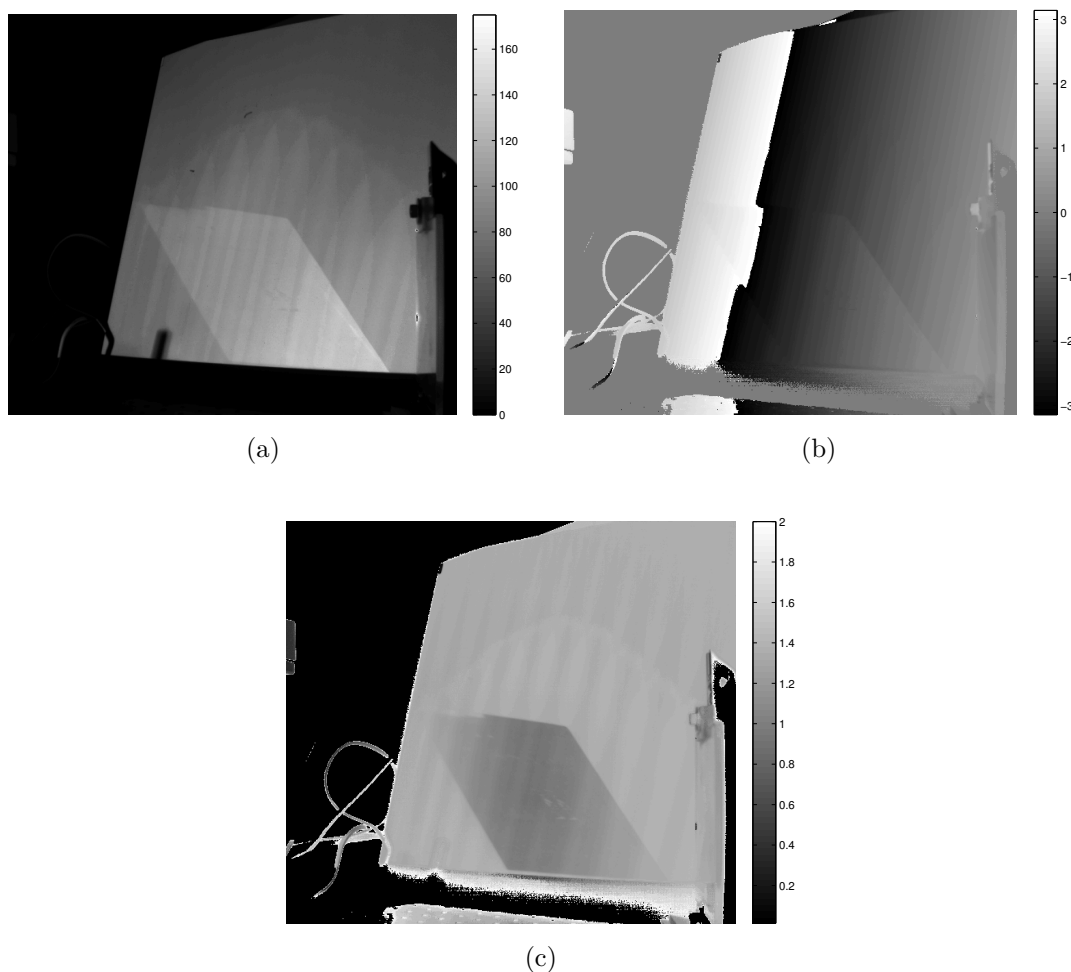(a)   (b)

(c)

Figure 3.1: The (a) magnitude, (b) phase, and (c) unit circle metric of a scene with an interreflection sampled using $f = 1$.

As before, PMP scanning is also performed varying f from 4 to 64 in steps of 4. The unit circle metric for $f = 64$ is shown in Figure 3.9, and Figure 3.10 shows the greatest non-DC FFT bins of the magnitude-constancy metric.

Motion Multipath Effects

This experiment seeks to detect multipath effects resulting from motion during scanning. The scene being imaged uses a toy truck as the foreground object and a white foam board as the background object. The truck is displaced from left to right during scanning. Magnitude, phase, and unit circle metric from a scan with $f = 1$ are shown in Figure 3.11.

Discussion

The phase corruption which can occur in PMP scanning due to multipath scattering is evident in Figure 3.1b, with the distortion especially clear at the $-\pi, \pi$ discontinuity. These corrupted pixels are also evident in Figure 3.1c.

In Figure 3.2, magnitude varies little during the frequency sweep. In Figure 3.3, distortion due to the interreflection is evident in all frequencies, and this is also seen in Figure 3.4. Additionally, even though an interreflection is present, when the path length difference is an integer multiple of 2π, the two paths are in phase and the interreflection can not be detected by simply applying the unit circle metric at a single frequency. However, phase distortion does not occur where the interreflection has the same phase as the direct path.

In Figure 3.5, we see that even small path length differences can be seen in Figure 3.5a. The background, which is prominent in Figure 3.5a, becomes much less apparent as the path length difference increases, being barely discernible in Figure 3.5h. This demonstrates that some diffuse scattering is present in foam board.

In Figure 3.6, the magnitude of the path length difference in projector coordinates between the interreflection and the direct path has been determined, although well understood aliasing effects are evident. Each FFT bin represents a path length difference of 16 projector columns, with a maximum detectable path length difference of 128 projector columns.

As seen in Figure 3.7 and Figure 3.8, sub-pixel multipath effects can also cause artifacts in PMP scanning. In Figure 3.8c and Figure 3.9, the unit circle metric shows PMP artifacts due to sub-pixel multipath scattering and the shadow. In Figure 3.10, performing PMP using multiple frequencies reveals an increasing path length difference between foreground and background moving upward. As before, each FFT bin represents a path length difference of 16 projector columns, with a maximum detectable path length difference of 128 projector columns.

Finally, Figure 3.11 shows how inter-frame motion can induce multipath scattering artifacts in PMP scanning. These artifacts can be detected with the unit circle metric, with some pixels in Figure 3.11c showing a significant deviation from unity.

In summary, in this chapter we have seen that PMP performed in the presence of multipath interference is optically challenging. We have also seen that multipath interference can occur in numerous subtle ways, making it difficult to anticipate and control. To help detect and quantify the presence of quality problems in PMP, we have developed constraint checks and have demonstrated them to be effective in several optically challenging scenes.
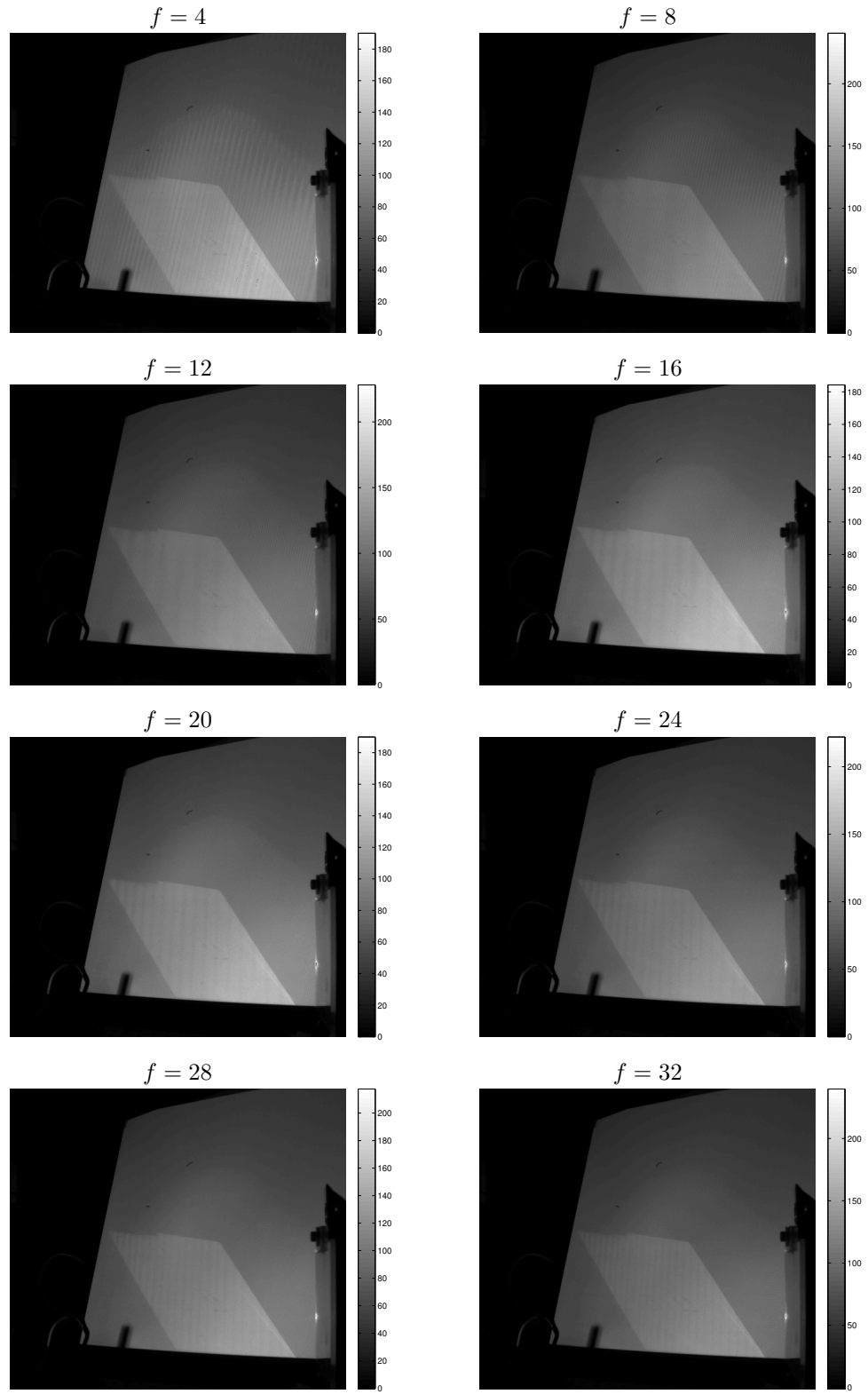
Figure 3.2: Magnitudes for the first few frequencies as f was varied from 4 to 64 in steps of 4.
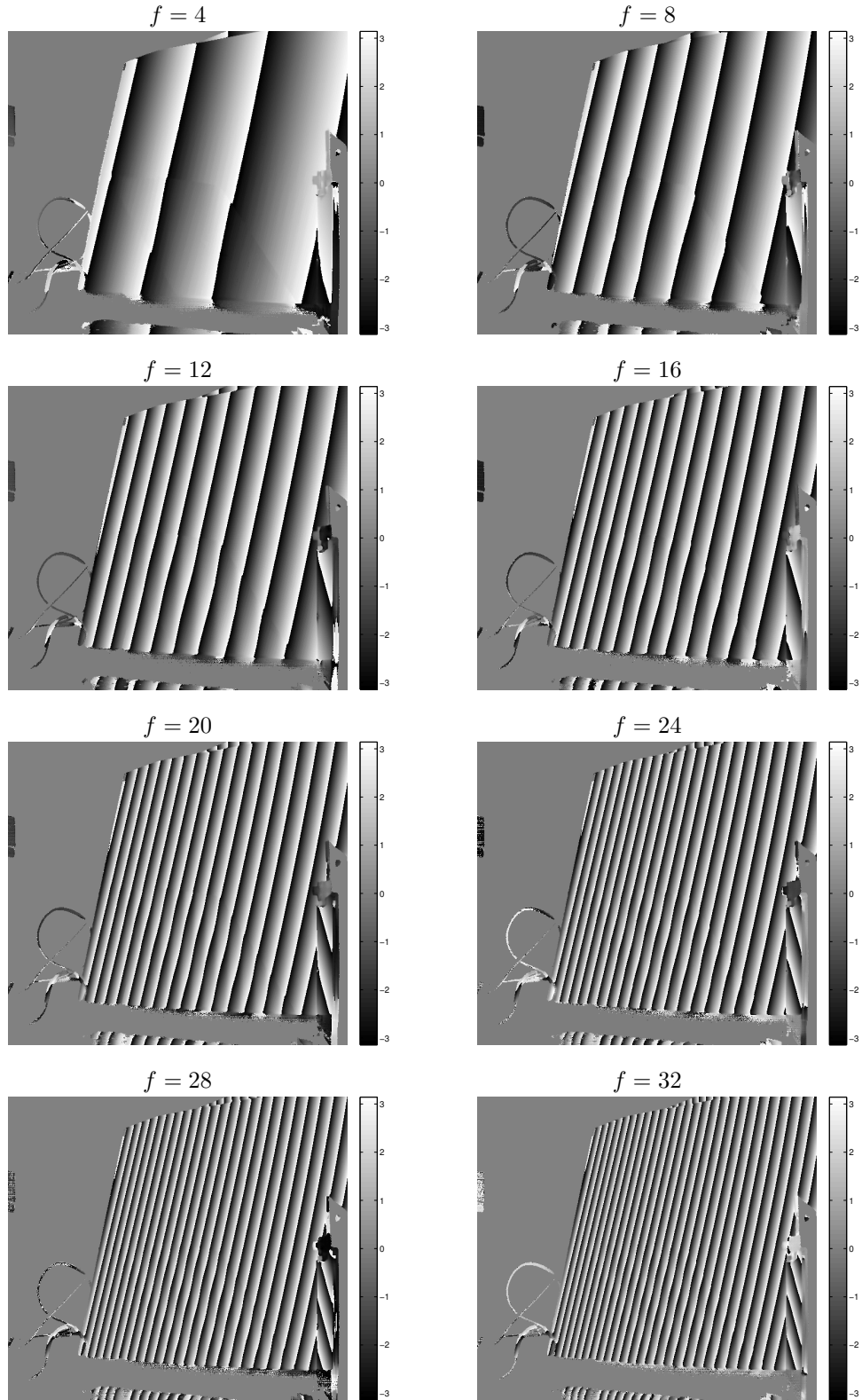
Figure 3.3: Phase of the first few frequencies as f was varied from 4 to 64 in steps of 4.
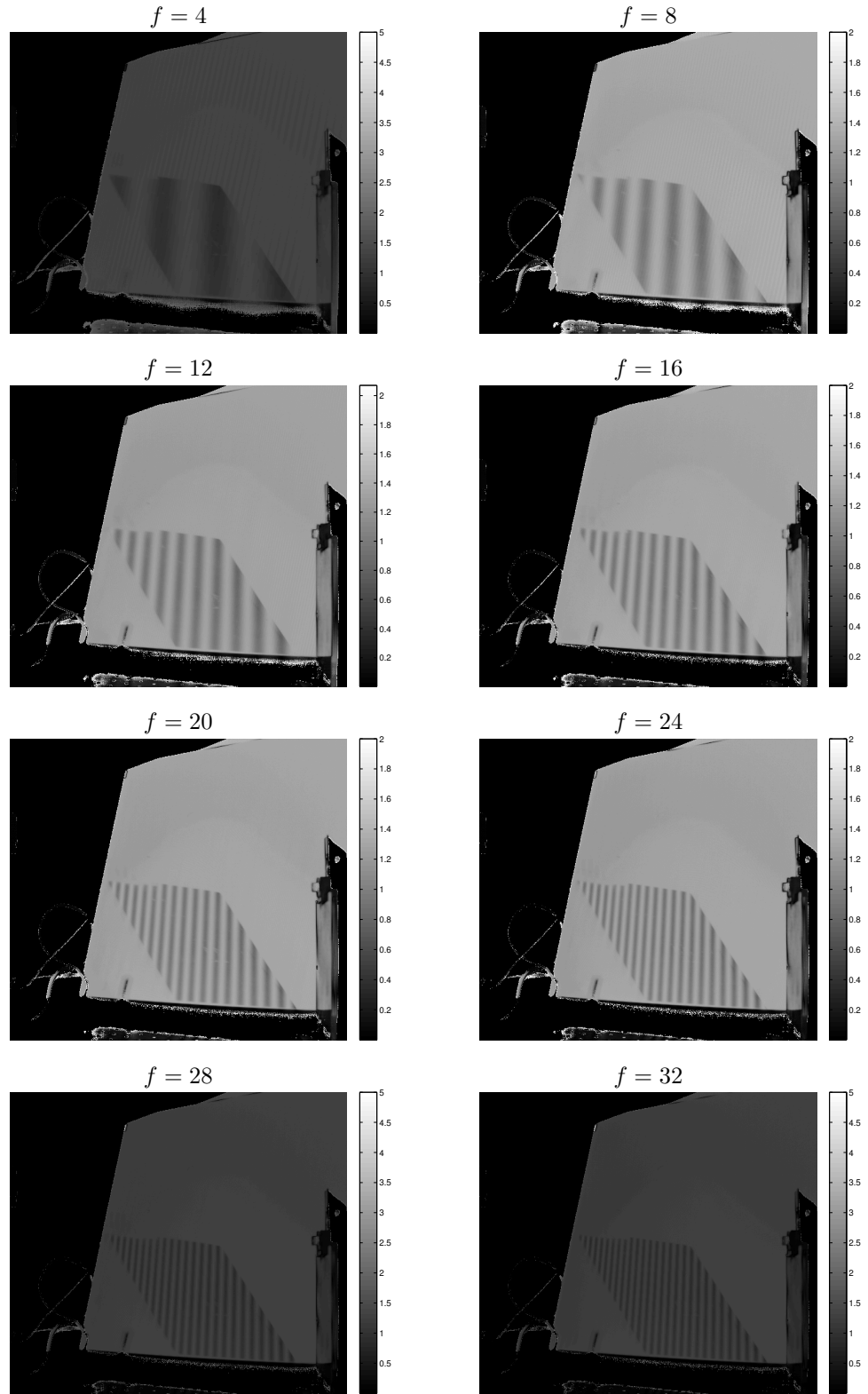
Figure 3.4: The unit circle metric of the first few frequencies as f was varied from 4 to 64 in steps of 4.
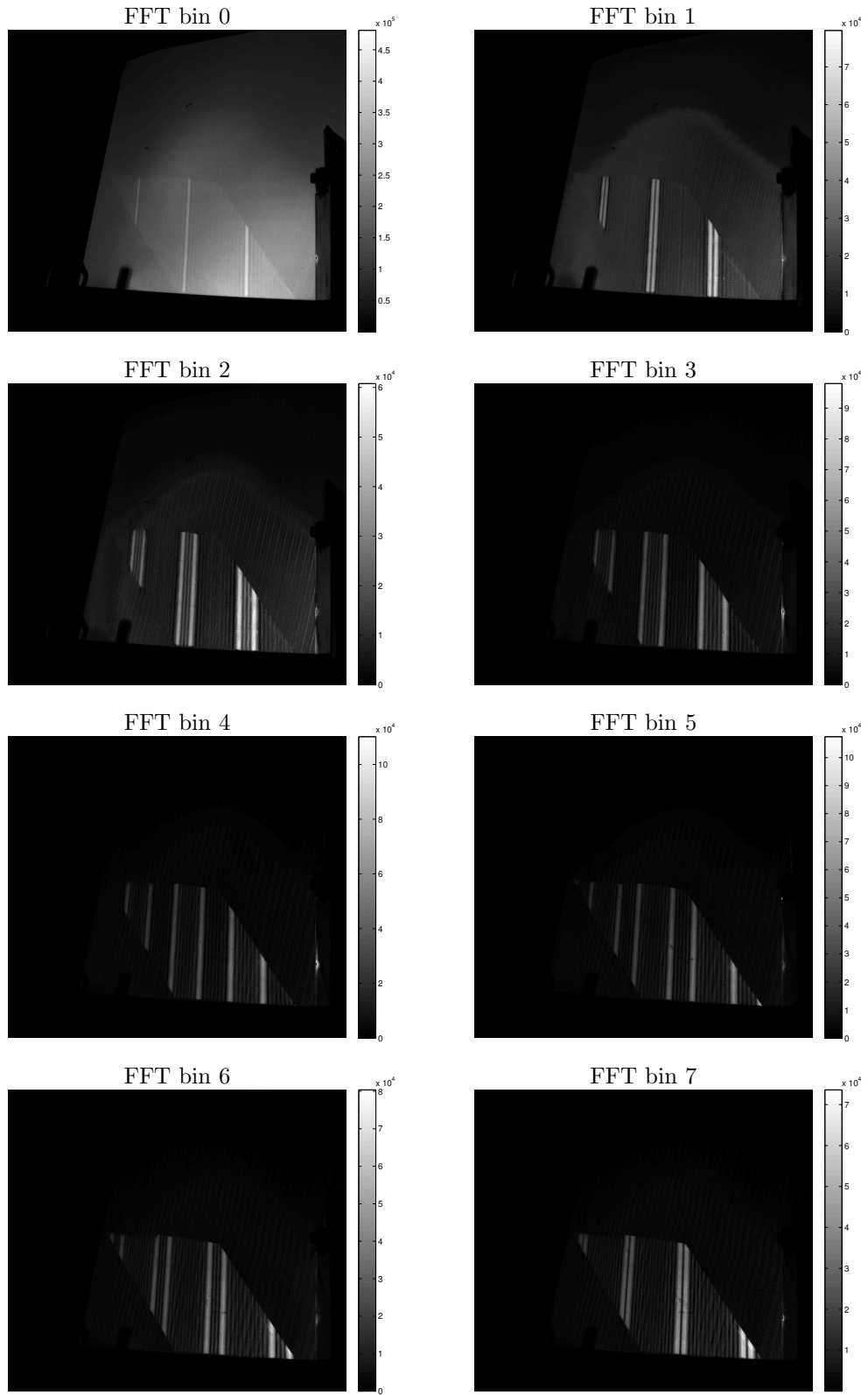
Figure 3.5: The magnitudes of the first few FFT bins.

Path Length Differences (FFT bins)
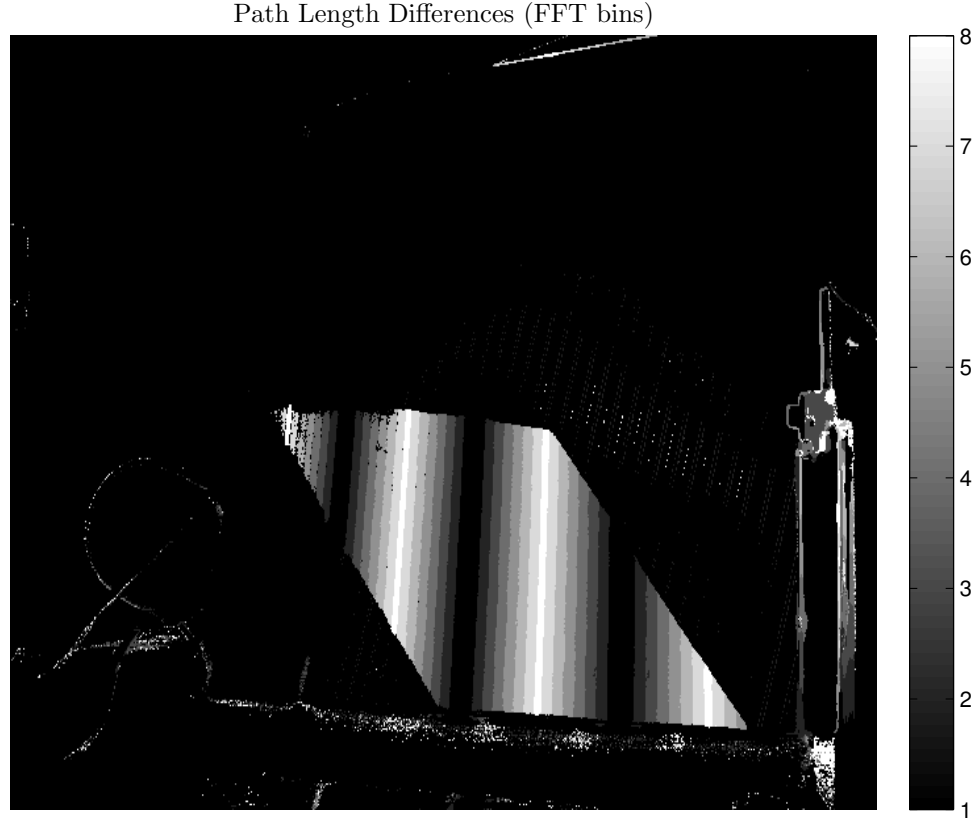
Figure 3.6: The path length differences of projector coordinates in FFT bins. Each bin corresponds to a path length difference of 16 projector columns, and the maximum detectable path length difference is 128 projector columns.

Magnitude   Phase   Unit Circle Metric

(a)   (b)   (c)

Figure 3.7: The (a) magnitude, (b) phase, and (c) unit circle metric of a scene with a sub-pixel foreground object acquired using $f = 1$.

Magnitude	Phase	Unit Circle Metric
(a)	(b)	(c)
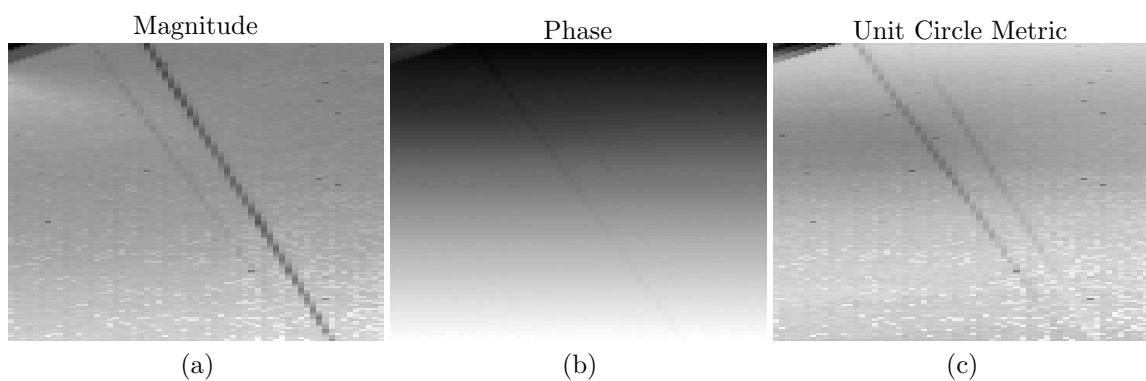
Figure 3.8: An enlarged view of Figure 3.7 in a region containing both the foreground object (left) and its shadow (right).
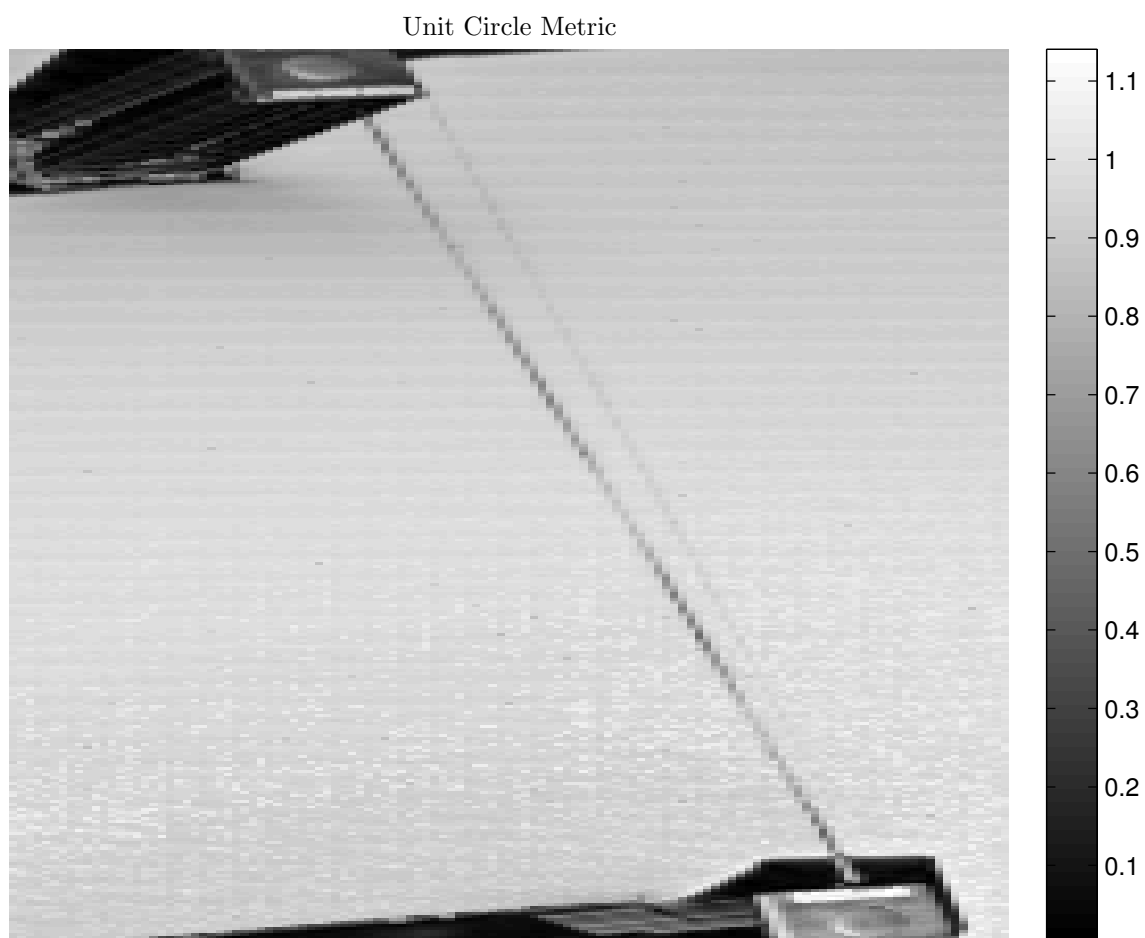
Unit Circle Metric

Figure 3.9: Unit circle metric for $f = 64$.
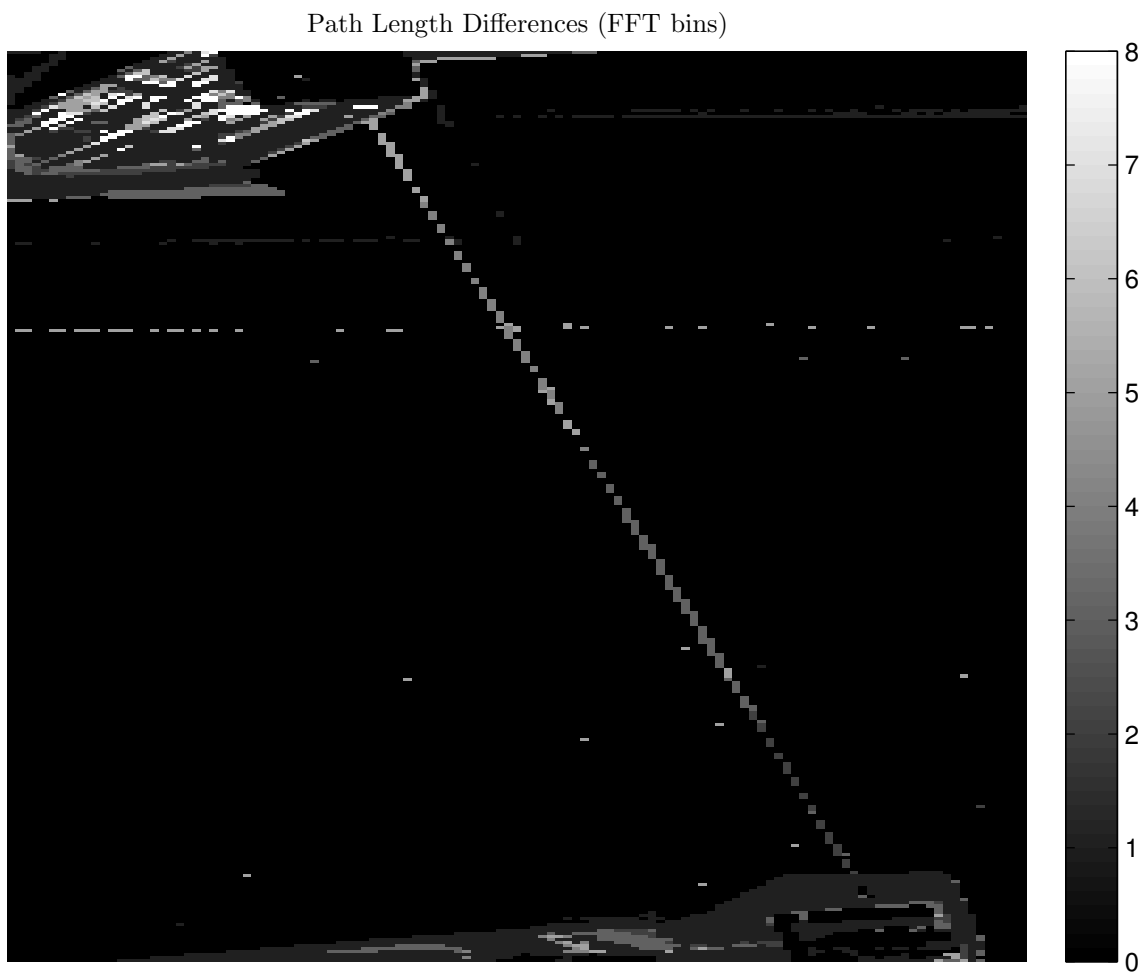
Figure 3.10: The path length differences between foreground and background objects in FFT bins. Each bin corresponds to a path length difference of 16 projector columns, and the maximum detectable path length difference is 128 projector columns.
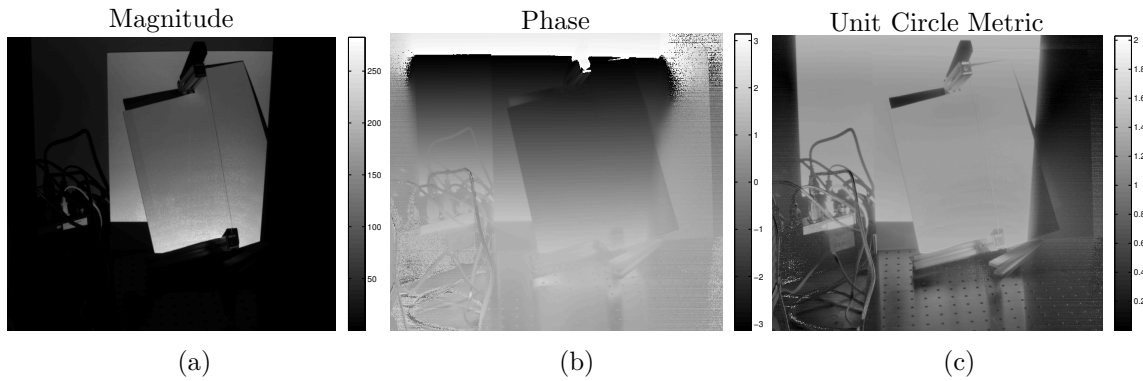

| (a) | (b) | (c) |
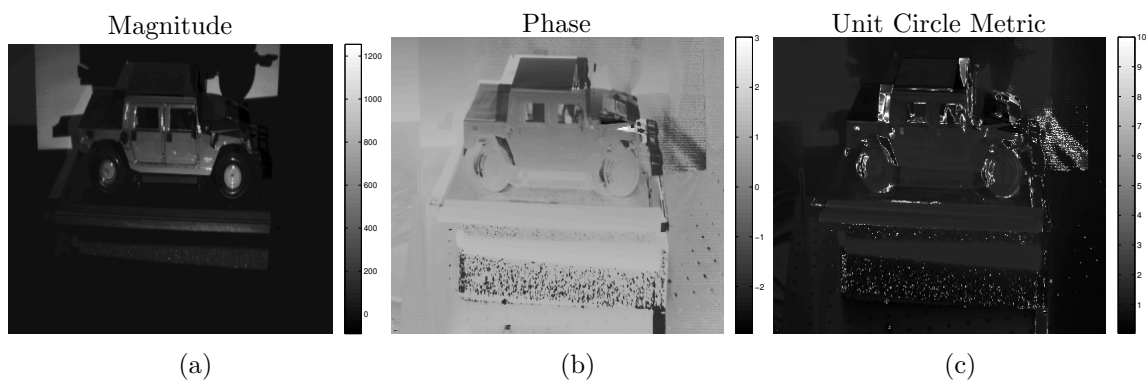
Figure 3.11: The (a) magnitude, (b) phase, and (c) unit circle metric of a scene with motion occurring between frames of a scan acquired using $f = 1$.

Chapter 4 Modeling Multipath Scattering in Structured Light Scanning

SLI methods are commonly utilized techniques in machine vision. However, in general they do not perform well when interreflections are present in a scene, and the few SLI approaches which can be successful in these environments have significant drawbacks. Building on SLI techniques used in computer graphics, we develop methods which can improve the robustness of 3D reconstruction in the presence of interreflections and demonstrate how recent advances in compressed sensing can be applied to SLI. We also demonstrate how interreflections in a scene can be modeled, which brings us to a discussion of light transport.

Light Transport

With ambient light subtracted and an appropriate HDR technique employed, a linear relation exists between scene radiance and camera pixel response. If any existing non-linearities in the projector are corrected then this linear relation can be extended to include the projector and the relation between the intensity of projector pixels and camera pixels can be expressed as

$$\mathbf{C} = \mathbf{TP}, \tag{4.1}$$

where $\mathbf{C}$ is a vector of camera pixel responses, $\mathbf{P}$ is a vector of projector pixel intensities, and $\mathbf{T}$ is known as the transport matrix.

Since usually there are millions of pixels in cameras and projectors, $\mathbf{T}$ requires trillions of bytes of storage. However, as discussed in Chapter 3, range reconstruction only requires identifying the correspondence between camera pixels and projector rows or columns. Thus, only one dimensional signals need to be projected either along projector columns, $\mathbf{P_x}$, or projector rows, $\mathbf{P_y}$. Using these one dimensional signals, the light transport model can be simplified to

$$\mathbf{C_x} = \mathbf{T_x P_x} \tag{4.2}$$

and

$$\mathbf{C_y} = \mathbf{T_y P_y}, \tag{4.3}$$

where $\mathbf{C_x}$ and $\mathbf{C_y}$ are vectors containing the reponse of each camera pixel to $\mathbf{P_x}$ and $\mathbf{P_y}$, and $\mathbf{T_x}$ and $\mathbf{T_y}$ are transfer matricies relating camera pixel responses to projector columns and rows. Since projectors commonly have a width and height of about a thousand pixels, the storage requirements for $\mathbf{T_x}$ and $\mathbf{T_y}$ are practical to achieve.

In SLI applications, the projected signals $\mathbf{P_x}$ and $\mathbf{P_y}$ are known, $\mathbf{C_x}$ and $\mathbf{C_y}$ are measured, and $\mathbf{T_x}$ and $\mathbf{T_y}$ are to be determined. $\mathbf{T_x}$ and $\mathbf{T_y}$ may be found one row at a time by solving

$$\mathbf{R}_i = \mathbf{Q T}_i \tag{4.4}$$

for each camera pixel i, where $\mathbf{R}_i$ is a vector containing the camera pixel measurements, each row of $\mathbf{Q}$ is the pattern projected during a measurement, and $\mathbf{T}_i$ is

the unknown row of $\mathbf{T_x}$ or $\mathbf{T_y}$. Of course finding $\mathbf{T_x}$ or $\mathbf{T_y}$ from (4.4) using a general approach for most projectors would require sampling with approximately one thousand patterns, but recognizing that $\mathbf{T_x}$ and $\mathbf{T_y}$ are expected to be mostly sparse allows the number of patterns to be reduced through compressed sensing techniques.

Compressed Sensing

Compressed sensing (CS) aims to reduce the number of equations required to solve a linear system by applying sparsity assumptions. Compressed sensing also gives guidance on how measurement acquisition should be performed.

Solving the matrix equation

$$\mathbf{y} = \mathbf{Ax} + \mathbf{n} \tag{4.5}$$

where $\mathbf{y}$, $\mathbf{x}$, and $\mathbf{n}$ are column vectors and $\mathbf{A}$ is a matrix is fundamental to many fields. Commonly, $\mathbf{y}$ represents measurements, $\mathbf{x}$ represents states to be estimated, $\mathbf{A}$ is a measurement matrix, and $\mathbf{n}$ represents additive noise. The fundamental theorem of linear algebra states that for a noise free system, (4.5) can not be solved unless the number of equations is equal to the number of unknowns, and even more equations are required for the solution to be robust to noise. However, under-determined systems are quite common.

The field of regularization theory seeks to solve these ill-posed problems by introducing *a priori* knowledge, usually gained from physical reasoning about the problem. For a least-squares solution, the residual error is to be minimized,

$$\min_{\mathbf{x}} ||\mathbf{Ax} - \mathbf{y}||_2^2. \tag{4.6}$$

where the subscript of 2 indicates the Euclidean norm, since the general p-norm is defined as

$$||\mathbf{x}||_p = \left(\sum_{j=1}^{N} |x_j|^p \right)^{\frac{1}{p}} \tag{4.7}$$

where N is the length of $\mathbf{x}$. When the system is under-determined, additional equations can be introduced by applying constraints,

$$||\mathbf{\Gamma x}||_2^2. \tag{4.8}$$

$\mathbf{\Gamma x}$ is often called the stabilizing functional. The matrix $\mathbf{\Gamma}$ imposes *a priori* assumptions, commonly made about the smoothness, stability, or likelihood of a solution [71]. Both (4.6) and (4.8) can be combined,

$$\min_{\mathbf{x}} ||\mathbf{Ax} - \mathbf{y}||_2^2 + \lambda ||\mathbf{\Gamma x}||_2^2. \tag{4.9}$$

This allows a trade-off between residual error and a property such as smoothness to be made in proportion according to the regularization parameter λ. This has already been seen exactly in this context in Chapter 2 in the fitting of camera response

functions. A very common approach to this problem is Tikhonov regularization [60], where the function to be minimized is

$$||\mathbf{Ax} - \mathbf{y}||_2^2 + ||\mathbf{\Gamma x}||_2^2. \tag{4.10}$$

Note that if $\mathbf{\Gamma}$ is the identity matrix, we are seeking a minimum Euclidean norm solution.

In several applications, the goal is to find the sparsest solution,

$$\min_{\mathbf{x}} ||\mathbf{x}||_0 \text{ s.t. } \mathbf{y} = \mathbf{Ax}. \tag{4.11}$$

Unfortunately solving (4.11) has been proven to be NP-hard [56]. However, it has been shown [19] that for most cases (4.11) is equivalent to minimizing the L_1 norm,

$$\min_{\mathbf{x}} ||\mathbf{x}||_1 \text{ s.t. } \mathbf{y} = \mathbf{Ax}. \tag{4.12}$$

This is called the least absolute shrinkage and selection operator (LASSO) [86] or basis pursuit [12], and

$$\min_{\mathbf{x}} ||\mathbf{y} - \mathbf{Ax}||_2^2 + \lambda ||\mathbf{x}||_1 \tag{4.13}$$

is referred to as basis pursuit denoising [12]. These problems can be solved by techniques in linear or convex programming, such as the well known simplex method [12, 7]. This is shown graphically in Figure 4.1. The linear system is solved when solutions with progressively larger norms are considered and the norm intersects the constraint surface. Note how for L_1 and concave norms like $L_{\frac{1}{2}}$ this leads to a sparser solution, which is not the case for higher order norms.

Another useful approximation to (4.11) is

$$\min_{\mathbf{x}} ||\mathbf{y} - \mathbf{Ax}||_2 \text{ s.t. } ||\mathbf{x}||_0 \leq k \tag{4.14}$$

for a given number of non-zero terms k. Among the approaches which solve (4.14) is orthogonal matching pursuit (OMP) [88] . There are many variants of OMP designed to have improved theoretical properties [59, 20, 58]. These methods are collectively referred to as greedy algorithms, and are the most computationally efficient approaches available.

While these works show that sparse signals can be recovered using L_1 minimization or greedy approaches, several questions remain. For example, how many equations are required to find a solution with k non-zero terms? How sensitive is the solution to noise? Under what conditions can exact recovery be accomplished? While some specifics depend on the problem instance, the work in [9, 8, 15, 27] has shown that optimal reconstruction will occur when $\mathbf{A}$ satisfies the restricted isometry property (RIP),

$$(1 - \delta_k) ||\mathbf{x}||_2^2 \leq ||\mathbf{Ax}||_2^2 \leq (1 + \delta_k) ||\mathbf{x}||_2^2, \tag{4.15}$$

for all possible solutions $\mathbf{x}$ with k non-zero terms, called k-sparse solutions, for some constant $\delta_k \in (0, 1)$. Basically, RIP is satisfied when the columns of $\mathbf{A}$ are approximately orthogonal, the length of $\mathbf{x}$ is preserved by $\mathbf{A}$, and δ_k is not too close
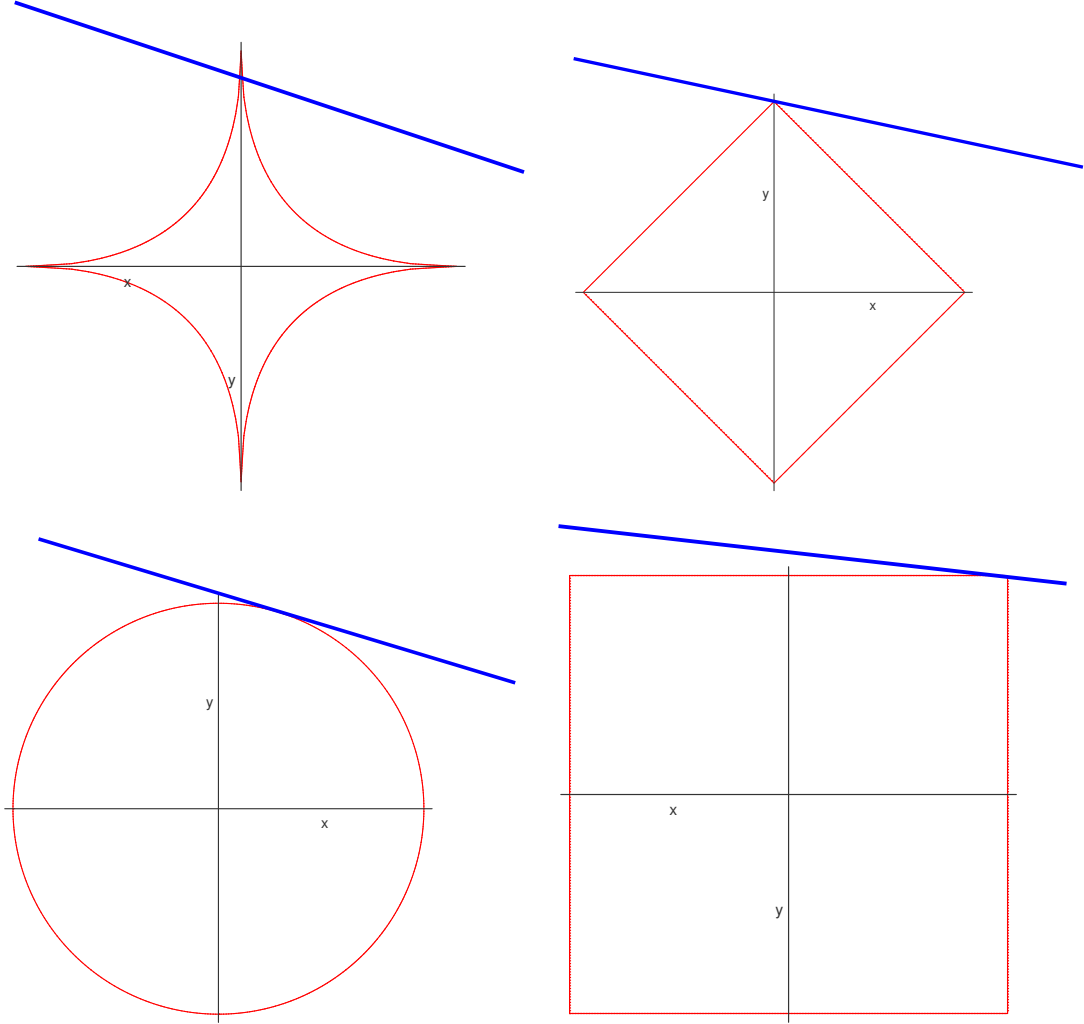
Figure 4.1: Unit balls for the (a) $L_{\frac{1}{2}}$, (b) L_1, (c) L_2, and (d) L_∞ norms. The blue line represents a constraint surface.

to 1. Equivalently, the nullspace property states that optimal recovery can occur when there are no sparse vectors in the nullspace of $\mathbf{A}$. At the time of writing, the only matrices known to satisfy RIP and the nullspace property are random matrices, such as Gaussian, Bernoulli, and random partial Fourier matrices. Whether any other matrices exist which satisfy the RIP is an open question.

The significance of this work to measurement applications is not to be understated. Consider an application such as photography. In general, photographs are compressible in the wavelet domain and generally a few thousand wavelet coefficients are sufficient to produce a satisfactory image. However, many millions of measurements are performed in a rectangular array, transformed into the wavelet basis, and then the vast majority of the data is lost due to truncation. This extremely inefficient process can be very undesirable when there is a cost associated with these measurements, such as the sampling rates of detectors, or radiation exposure to patients in the

case of x-ray computed tomography (CT) imaging. Compressed sensing directs us how to acquire measurements in the compressed domain such that the number of measurements can be minimized while still allowing suitable reconstruction to occur when the solution is expected to be sparse.

OMP

OMP is a particularly efficient approach to CS. Using OMP, a k-sparse vector $\mathbf{x}$ with a total vector length of d can be recovered with $O(k \ln d)$ total measurements. It's speed and simple implementation make it an attractive CS technique, and it is a method selected for use in this research.

Referring again to (4.5), in applying OMP, $\mathbf{A}$ is known and also satisfies certain properties [88], $\mathbf{y}$ is obtained from measurements, and the sparsity level k of $\mathbf{x}$ is known from *a priori* knowledge. OMP proceeds in a number of steps.

1. A residual error vector $\mathbf{r}$ is initialized as $\mathbf{r}_0 = \mathbf{y}$. A k-dimensional vector $\mathbf{\Lambda}$, which will be used to store the index of non-zero entries in $\mathbf{x}$, is initialized to empty. A $N \times k$ matrix $\mathbf{\Phi}$, which is used to store rows of $\mathbf{A}$, is initialized to zeros. If the columns of $\mathbf{A}$, denoted as $\mathbf{A}_j$, are not normalized, then normalization must also be performed. The iteration counter t is initialized to one.

2. The columns of $\mathbf{A}$ are correlated with $\mathbf{r}_{t-1}$ to find the index of the column with the greatest response, λ_t,

$$\lambda_t = \operatorname*{argmax}_{j} |\langle \mathbf{r}_{t-1}, \mathbf{A}_j \rangle| . \tag{4.16}$$

3. With λ_t identified, it is saved at location t in $\mathbf{\Lambda}$, and the corresponding column $\mathbf{A}_{\lambda_t}$ is saved in column t of $\mathbf{\Phi}$.

4. The current estimate of the solution $\mathbf{x}_t$ is obtained by seeking a solution with minimum residual error,

$$\mathbf{s}_t = \operatorname*{argmin}_{s} \|\mathbf{y} - \mathbf{\Phi s}\|_2 . \tag{4.17}$$

5. Update the residual error,
$$\mathbf{r}_t = \mathbf{y} - \mathbf{\Phi s}_t. \tag{4.18}$$

6. Increment t and return to Step 2 while the $t < k$.

When this algorithm has finished, the estimate $\widehat{\mathbf{x}}$ has non-zero values at the indices in $\mathbf{\Lambda}$. These non-zero values are the corresponding entries in $\mathbf{s}_t$.

Since we are seeking a sparse solution, the efficiency of solving the least squares problem in Step 4 can be improved using an algorithm such as LSQR [63], though the performance limiting factor is the matched filter of Step 3, which is $O(Nd)$ since $\mathbf{A}$ is a dense matrix.

70

To date, the only application of CS in SLI is [30], where an inhomogeneous participating media is examined. This application is very different than traditional SLI performed in 3D surface reconstruction. As such, the suitability of CS in traditional SLI is examined. We now describe and present results from an experiment designed to perform robust surface reconstruction using SLI in the presence of strong interreflections. A compressed sensing approach is used. We discuss its performance, and compare results to those obtained from a least squares approach.

4.1 Compressive SLI

A range sensor is formed in the usual way from a calibrated camera-projector system [33, 72]. As discussed in Chapter 2, due to the dynamic range of scene radiance, patterns are captured using the Kalman filtering HDRI method developed in Chapter 2. As mentioned in Chapter 3, only x_p need be determined to perform surface reconstruction, so patterns are based on illuminating projector columns. Since interreflections are present, light transport is modeled as (4.2), thus accounting for the fact that camera pixels may be illuminated by multiple projector columns. With *a priori* knowledge that $\mathbf{T_x}$ is expected to be sparse, OMP, with a LSQR solver [63], is used to solve for the light transport model from (4.4). In the projected patters, column intensities are fully on or fully off according to a Bernoulli distribution, thus ensuring RIP holds while also eliminating the need to perform gamma correction in the projector.

The scene being imaged consists of a mirror placed in front of a flat white foam board, with interreflections clearly visible. A HDR background image of the scene acquired with all projector pixels having an intensity of 0. This ambient background, shown in Figure 4.2 is subtracted from all subsequent measurements so (4.2) holds.

Again, the solution vector $\mathbf{T_x}$ represents the response of a camera pixel to each projector column. While the strongest response is usually due to projector direct path illumination, due to diffuse scattering projector columns neighboring those of the direct path may also induce a strong response, and this causes peak broadening in $\mathbf{T_x}$. In determining the magnitude and location of the strongest response, a peak finding strategy is employed where maximum response is taken as the peak location, and magnitude is found by integrating over the peak. It is assumed that maximum intensity corresponds to direct path illumination, and that remaining peaks are due to interreflections.

In solving for $\mathbf{T_x}$, two systems are considered. In the first, (4.2) is a system with 100 patterns solved for 10 non-zero entries, and in the second the system has 400 patterns used to solve for 20 non-zero entries. Additionally, a performance benchmark is obtained by solving (4.2) using SVD to obtain a least squares solution to a full rank system obtained using 1024 patterns.

The response $\mathbf{T_x}$ obtained for the 100 pattern, 400 pattern, and 1024 pattern least squares approaches is shown for a single representative pixel location in Figure 4.3.

After $\mathbf{T_x}$ has been determined for each camera pixel in the image, peak analysis is performed to find the magnitude and projector column location of the three strongest peaks, shown in Figure 4.4, Figure 4.5, and Figure 4.6.

Figure 4.2: A scene with strong interreflection used in the compressive SLI experiment. This image represents the ambient background intensity of the scene, with camera pixel intensities shown in logarithmic scale.

In examining these results, it is clear that OMP with 100 patterns is not sufficient to yield acceptable results. In Figure 4.3a, the noise floor is very high and the location of the strongest peak does not correspond to the projector column of direct path illumination. In the decoded image Figure 4.4d, we expect to see a smooth progression in projector columns from left to right. This can be observed, but there is a high level of noise, again indicating that identifying the code corresponding to direct path illumination was difficult. In Figure 4.4b and Figure 4.4c there is only slight evidence that an interreflection is be present in the scene. In Figure 4.4e and Figure 4.4f there appears to be no success in identifying the code of the interreflection.

The results for OMP with 400 patterns are much better. In Figure 4.3a, the location of the strongest peak correctly corresponds to the projector column of direct path illumination, and this peak is well above the noise floor. In Figure 4.5a, we see that the interreflections have been successfully separated from direct path illumination. This is also evident in Figure 4.5d, where we see that the codes corresponding to direct path illumination have been successfully recovered. Furthermore, beyond the ability to remove the corrupting effects of multipath scattering, in Figure 4.5b and Figure 4.5e we see some success in identifying both the magnitude and code of the interreflection, though Figure 4.5c and Figure 4.5f show that the interreflections
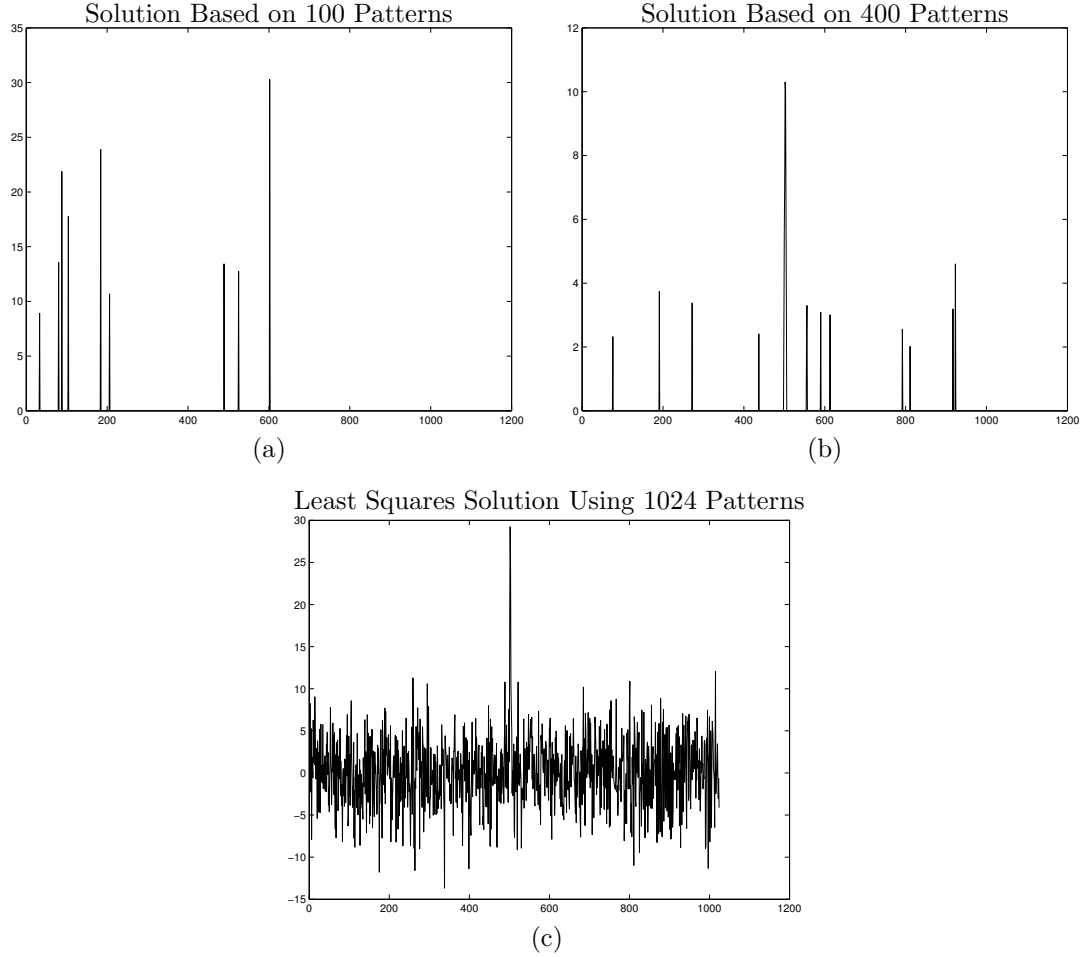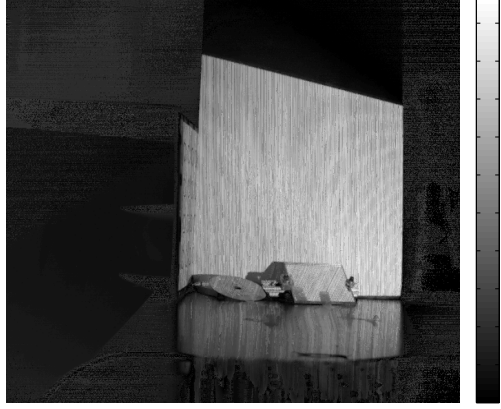
Figure 4.3: Sparse solution vectors for pixel (250, 400) based on (a) 100 patterns with 10 non-zero entries, (b) 400 patterns with 20 non-zero entries, and (c) 1024 patterns using a full rank least squares approach.

do not only appear as the second peak, but also in the third peak as well.

In Figure 4.3c, we see that the least squares approach can also recover the location of the peak corresponding to direct illumination, although the noise floor is high. Unlike OMP, negative solutions are common. From Figure 4.6a and Figure 4.6d we see that interreflections have been removed from the direct path illumination and the codes corresponding to direct illumination have been recovered, though some noise is present. In Figure 4.6b, Figure 4.6c, Figure 4.6e, and Figure 4.6f it is clear that virtually no information about the presence of the interreflection has been recovered.

Clearly, given sufficient measurements, the results from OMP can be impressive. Unfortunately its computational costs are considerable, with the required computations generally taking many hours, if not days, to complete on a single processor computer depending on sparsity and the number of patterns selected. *A priori* knowledge about the expected sparsity of the solution is a very powerful constraint which can help OMP deliver better results with fewer measurements than a least

Magnitude of Peak 1, 100 Patterns (dB)   Magnitude of Peak 2, 100 Patterns (dB)

(a)   (b)
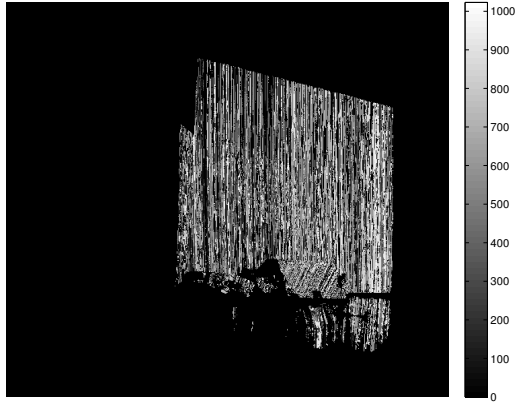
Magnitude of Peak 3, 100 Patterns (dB)   Projector Column of Peak 1, 100 Patterns

(c)   (d)
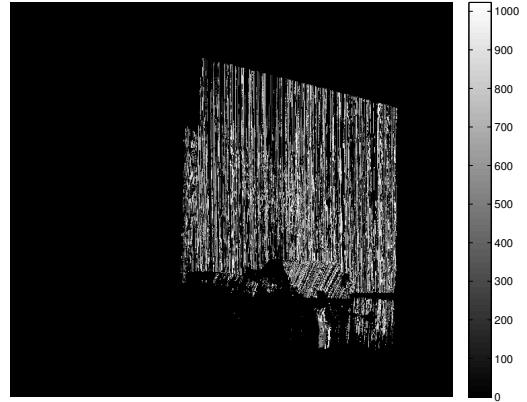
Projector Column of Peak 2, 100 Patterns   Projector Column of Peak 3, 100 Patterns

(e)   (f)

Figure 4.4: The magnitude and projector columns of the three greatest peaks obtained with 100 patterns and 10 non-zero entries.
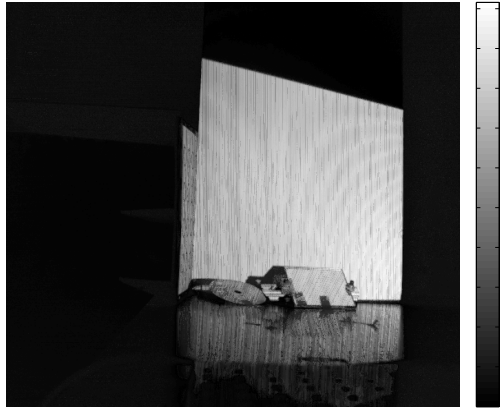
Figure 4.5: The magnitude and projector columns of the three greatest peaks obtained with 400 patterns and 20 non-zero entries.

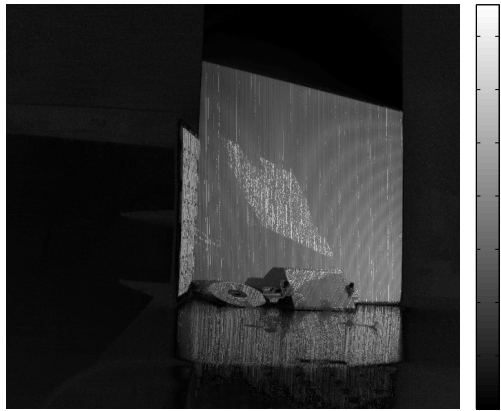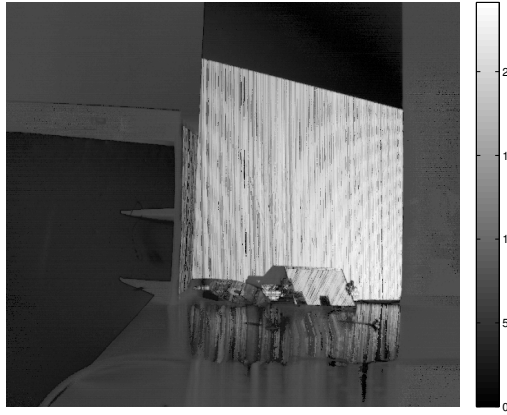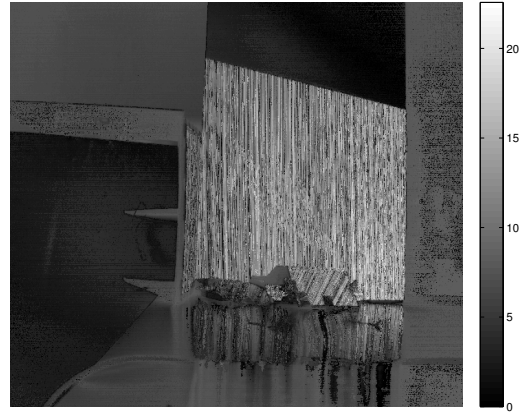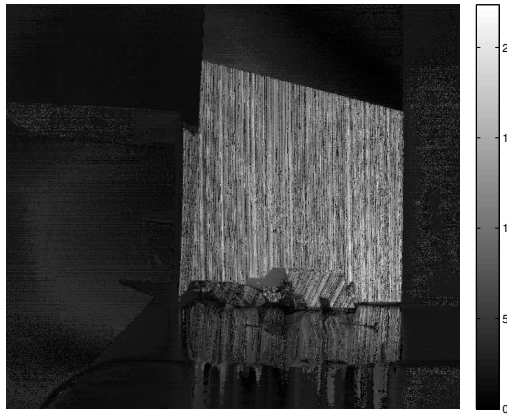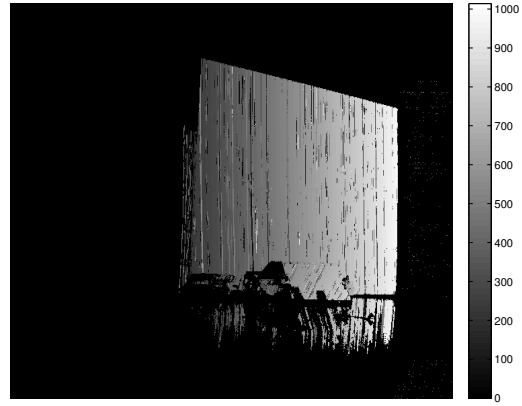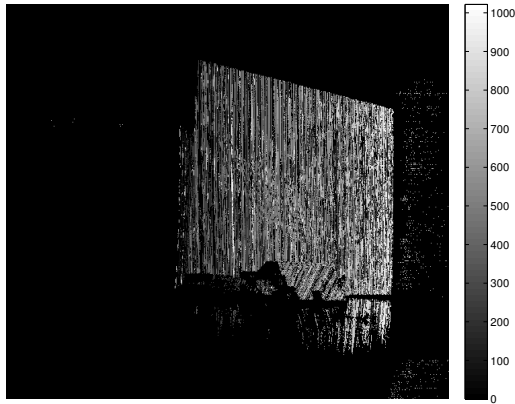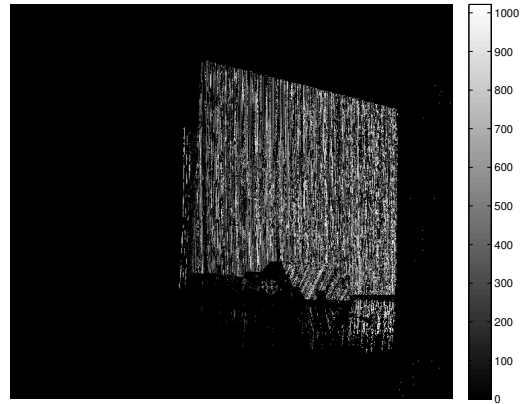Figure 4.6: The magnitude and projector columns of the three greatest peaks obtained with 1024 patterns using a least squares approach.

squares approach in this application. Unfortunately, due in part to the peak broadening due to diffuse scattering, there is a limit on the sparsity of the solution. This and computational requirements make compressed sensing approaches somewhat impractical for this application. Building on the success of OMP in recovering information about the reflection, we present a light stripe based approach designed not just to improve surface reconstruction, but also to model the interreflections present in a scene.

4.2 Modelling Interreflections using SLI

As in Section 4.1, a range sensor is formed using the same camera-projector system using the same HDR imaging methods. However, unlike before, modeling interreflections requires establishing correspondences between camera pixels and projector pixels, not just projector columns. To avoid having a code for each projector pixel, we assume projector pixel coordinates can be obtained by pairing response peaks from a horizontal and a vertical scan, thus greatly reducing the number of required patterns. Of course this requires the responses due direct and indirect illumination have the same magnitude ordering in both the horizontal and vertical scan, an assumption which is reasonable. Also unlike before, where the responses $\mathbf{T_x}$ and $\mathbf{T_y}$ would be obtained by solving (4.2) and (4.3), scans are performed by turning on only one projector row or column at a time. In this way light stripe scanning allows $\mathbf{T_x}$ and $\mathbf{T_y}$ to be directly measured.

With $\mathbf{T_x}$ and $\mathbf{T_y}$ having been obtained for each camera pixel, the same peak finding approach discussed in Section 4.1 is applied. The magnitudes of the strongest peaks in $\mathbf{T_x}$ and $\mathbf{T_y}$ are averaged and their locations paired to form the magnitude of the response corresponding to direct illumination from a specific projector pixel. The same is done for remaining peaks to find the magnitude and projector pixel coordinates corresponding to indirect illumination.

After establishing the magnitudes and projector coordinates corresponding to both direct and indirect illumination for each camera pixel location, correspondences are established between pairs of camera pixels. This is done by searching for the best match between the projector coordinate of direct illumination at one camera pixel location and the projector coordinate of indirect illumination at another pixel location. If the Euclidean distance between these projector coordinates is less than a threshold, an interreflection between the two camera pixels is recorded. When world coordinates are obtained for this camera pixel pair, a 3D vector of light transport at this interreflection is obtained.

The scene being imaged is formed by an aluminum box placed in a corner. As in Section 4.1, a HDR image of the background ambient illumination is created and shown in Figure 4.7. Interreflections are clearly visible between the box and both sides of the corner. Additionally, reflections are evident on the table top surface, though these pixels do not receive any direct illumination from the projector.

The responses $\mathbf{T_x}$ and $\mathbf{T_y}$ obtained from the horizontal and vertical light stripe scans is shown for a single representative pixel in Figure 4.8.

77

Background Intensity (dB)

Figure 4.7: A HDR image of the ambient intensity of a scene with strong interreflections used in the impulse sampling experiment.
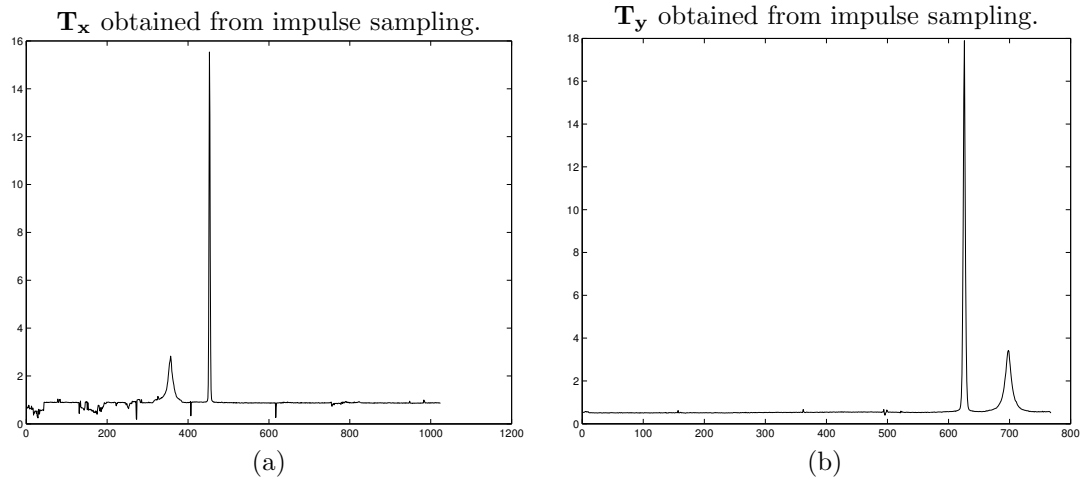

Figure 4.8: Response of pixel (300, 400) based on (a) horizontal light stripe scanning, and (b) vertical light stripe scanning.

With $\mathbf{T_x}$ and $\mathbf{T_y}$ determined for each camera pixel in the image, peak analysis is performed to find the magnitude and projector column location of the three strongest

peaks, shown in Figure 4.9 and Figure 4.10.

Some of the interreflections found in the scene are shown in Figure 4.11.

In Figure 4.8, we see a very clear primary peak corresponding to direct illumination and a smaller, broader peak due to an interreflection in both $\mathbf{T_x}$ and $\mathbf{T_y}$. As expected, peak broadening is especially pronounced in the peak corresponding to the interreflection. These directly measured responses show much improved SNR over those responses obtained as solutions to linear systems.

In both Figure 4.9a and Figure 4.10a we see that interreflections have been completely separated from direct illumination. Reflections are clearly visible in the table top since this area is not directly illuminated by the projector, making the reflections from the screen the strongest peaks. We see in Figure 4.9d and Figure 4.10d that the detected codes have not been corrupted by the presence of interreflections. Interreflections are clearly evident in Figure 4.9b and Figure 4.10b, especially between the aluminum box and screen and in the corner region. The codes of these interreflections can be seen in Figure 4.9e and Figure 4.10e. In Figure 4.9c and Figure 4.10c, the intensity of the third peak is a few decibels below the second peak with noise clearly present. Thus, in constructing interreflections only the first two response peaks are used.

The results of the correspondence search used to model interreflections is shown in Figure 4.11, with only a few of the interreflections being shown. The reflections of the screen on the table top is very evident, as are the interreflections between the screen and the aluminum box, and some interreflections can be seen in the corner region. Interestingly, at the top of the box a couple of false interreflections are also shown. The top of the interreflection on right screen should correspond with the very top edge of the metal box. However, the area of the left screen just above the shadow has very similar projector coordinates to the top edge of the box, and it has a greater intensity under direct illumination than the box, so it has been selected as the source of the interreflection.

In summary, optically challenging scenes may have strong interreflections. Traditional SLI approaches seek to perform 3D surface reconstruction and generally do not perform well when interreflections are present. SLI approaches used in computer graphics model light transport with a transfer matrix and do not seek to perform 3D surface reconstruction. Building on these approaches, we have demonstrated how both 3D surfaces and interreflections can be obtained using SLI, and we have seen how CS can be used to reduce storage requirements required in this approach.

Magnitude of Peak 1 in $\mathbf{T_x}$ (dB)

Magnitude of Peak 2 in $\mathbf{T_x}$ (dB)

(a)

(b)

Magnitude of Peak 3 in $\mathbf{T_x}$ (dB)

Projector Column of Peak 1 in $\mathbf{T_x}$

(c)

(d)

Projector Column of Peak 2 in $\mathbf{T_x}$

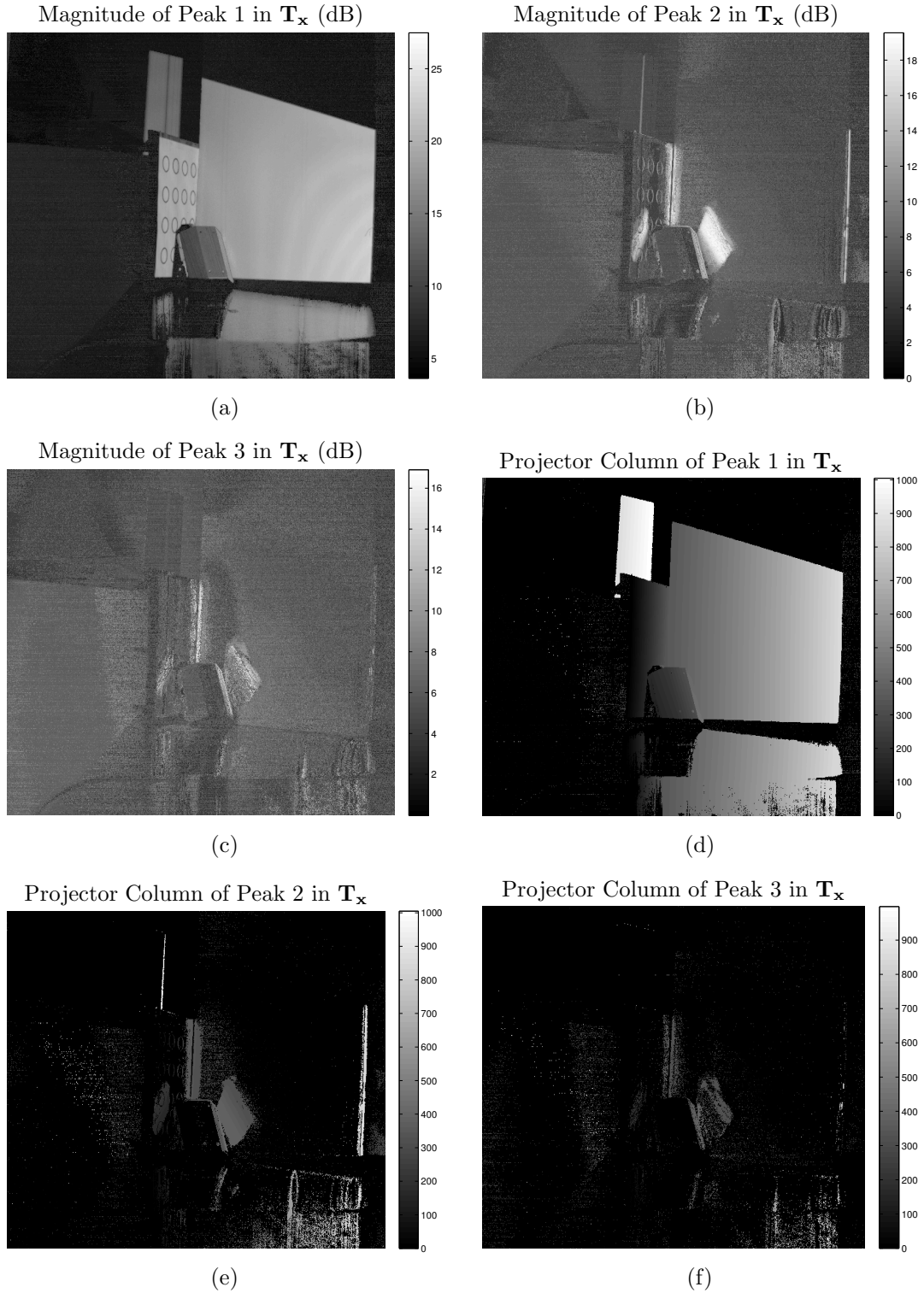Projector Column of Peak 3 in $\mathbf{T_x}$

(e)

(f)

Figure 4.9: The magnitude and projector columns of the three greatest response peaks obtained from horizontal scanning.
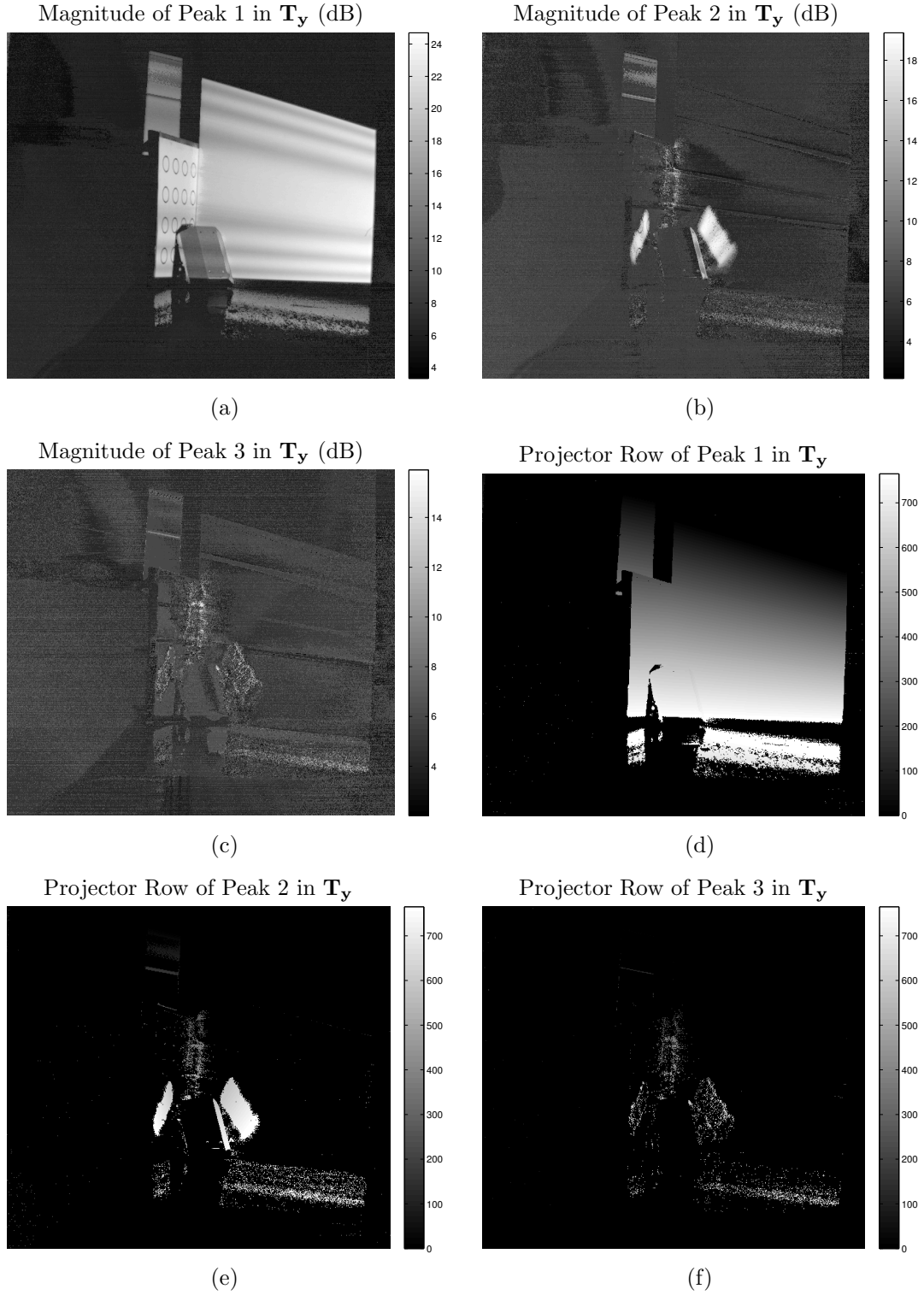
Figure 4.10: The magnitude and projector rows of the three greatest response peaks obtained from vertical scanning.
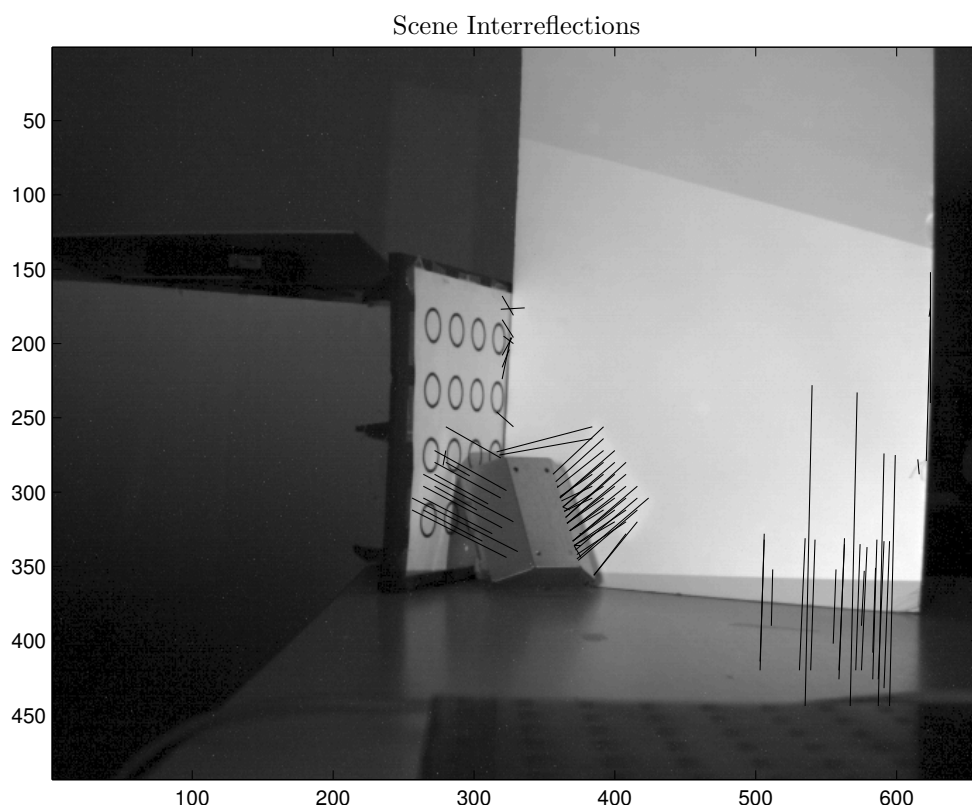
Figure 4.11: A few of the interreflections present in the scene.

Chapter 5 Conclusions and Future Work

This chapter concludes the research performed in this dissertation. In this chapter several current applications of the technologies developed in this research are discussed. These include HDR video, multispectral HDRI for artifact conservation, HDRI and non-uniformity correction in an assay systems, and multispectral HDRI and computer graphics operators. Additionally, we discussion future applications such as the 3D capture of scenes with significant sub-surface scattering.

5.1 Fading Memory Filter and HDR Video

HDR video using multi-exposure techniques is particularly challenging since it requires balancing competing demands. Since the scene has high dynamic range, not all pixels can be sampled with each exposure-image. Furthermore, dark pixels require long integration times, which reduces the frame rate of the camera. Even worse is the fact that, since motion is occurring, establishing pixel correspondences between exposure-images requires motion estimation, and motion estimators generally require integration time to be constant, a requirement which can not be satisfied with traditional multi-exposure HDR methods. Few approaches have been developed to achieve multi-exposure HDR video [35, 91].

In [35], an exposure control algorithm is used. Based on histogram analysis, the exposure control algorithm attempts to capture the entire dynamic range of a scene using two exposure-images of differing integration times, referred to as the long and short exposure-images. Video sampling is performed by alternating between long and short integration times. Motion estimation and image warping are performed prior to exposure-image fusion. Finally, tone mapping is performed so that the image can be viewed on a display. Only exposure-image acquisition and the histogram based exposure control algorithm are performed in real time.

In [91], histogram based techniques are used to perform calibration, response function estimation, HDR reconstruction, registration, and ghost removal. A key idea behind these techniques is that, even though pixel values may differ between exposure-images with different integration times, the monotonic ordering of pixel values in exposure-image histograms is preserved.

Unlike these approaches, we propose a technique based on the Kalman-like fading memory filter and use the differences between predicted and actual measurements, the innovations, as a cue of motion. Using a controller based on weighted pixel voting, integration times are selected such that dense sampling is performed for moving objects, while sparser sampling occurs for objects deemed to be stationary.

Fading Memory Filter

As discussed in [81], a fading memory filter weights recent measurements more heavily than prior measurements. With the same process and measurement models as (2.52)

and (2.53), the fading memory filter is

$$
\begin{align}
\hat{\mathbf{x}}_k(-) &= \boldsymbol{\Phi}_{k-1}\hat{\mathbf{x}}_{k-1}(+) + \boldsymbol{\Gamma}_{k-1}\mathbf{u}_{k-1} \tag{5.1} \\
\mathbf{P}_k(-) &= \alpha^2 \boldsymbol{\Phi}_{k-1}\mathbf{P}_{k-1}(+)\boldsymbol{\Phi}_{k-1}^T + \mathbf{Q}_{k-1} \tag{5.2} \\
\mathbf{K}_k &= \mathbf{P}_k(-)\mathbf{H}_k^T \left(\mathbf{H}_k\mathbf{P}_k(-)\mathbf{H}_k^T + \mathbf{R}_k\right)^{-1} \tag{5.3} \\
\hat{\mathbf{x}}_k(+) &= \hat{\mathbf{x}}_k(-) + \mathbf{K}_k \left(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_k(-)\right) \tag{5.4} \\
\mathbf{P}_k(+) &= \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_k(-)\left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)^T \\
&\quad + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T. \tag{5.5}
\end{align}
$$

In other words, the fading memory filter in (5.1)-(5.5) is the same as the Kalman filter in (2.54)-(2.58) except that in (5.2), *a priori* covariance estimates are scaled by α^2. With this scaling, $\mathbf{P}$ is no longer the estimate of the uncertainty of $\mathbf{x}$, but when $\mathbf{P}$ is increased by scaling, *a priori* state estimates are assigned greater uncertainty. This places greater weight on new measurements.

As with (2.61)-(2.65), the fading memory filter used in this application takes on a simpler form,

$$
\begin{align}
\hat{r}_k(-) &= \hat{r}_{k-1}(+) \tag{5.6} \\
P_k(-) &= \alpha^2 P_{k-1}(+) + Q_{k-1} \tag{5.7} \\
K_k &= AT_k P_k(-)\left(A^2 T_k^2 P_k(-) + R_k\right)^{-1} \tag{5.8} \\
\hat{r}_k(+) &= \hat{r}_k(-) + K_k \left(z_k - AT_k\hat{r}_k(-) - B\right) \tag{5.9} \\
P_k(+) &= \left(1 - K_k AT_k\right)^2 P_k(-) + K_k^2 R_k. \tag{5.10}
\end{align}
$$

The weight α was chosen as

$$
\alpha = 15 + \left\| \frac{z_k - AT_k\hat{r}_k(-) - B}{AT_k\hat{r}_k(-) + B} \right\|_2, \tag{5.11}
$$

a value strongly related to the measurement innovations.

Pixel States

As just discussed, in HDR video one can not expect a pixel to receive usable measurement updates in each exposure-image. Consequently, a pixel may be in one of three states: actively tracking, tracking partially lost, and tracking completely lost. A pixel in the actively tracking state receives measurement updates with every exposure-image, and this pixel has its state information updated according to (5.6)-(5.10). A pixel with tracking completely lost is assumed to have no usable state information since a measurement update has not occurred for some time. In this implementation a pixel is classified as having tracking completely lost if it has not received a usable measurement update from the last 5 measurements. The state is

completely re-initialized on the next usable measurement update according to

$$r = \frac{z - B}{AT} \qquad (5.12)$$

$$R = CTr + D \qquad (5.13)$$

$$P = \frac{R}{A^2T^2} \qquad (5.14)$$

$$\alpha = 1.01 \qquad (5.15)$$

and is then re-classified as tracking. A pixel with tracking partially lost is one which has not received a usable measurement update, but does not meet the criteria for being considered completely lost. It's state estimates are assumed to be valid.

Pixel Voting

After all pixel states are updated, a controller must decide what integration time to select for the next exposure-image. This is done by voting in a discrete parameter space, with 64 choices of integration time. Votes are weighted by innovations and increase exponentially the longer a pixel goes without a measurement update. Pixels which are classified as having tracking completely lost initially vote for the shortest integration time, and progressively vote for longer integration times until they receive a useful measurement update. This procedure aims to balance the competing demands inherent in HDR video.

HDR video is generated for a scene with a moving matchbox and a scene of a flickering candle having dynamic ranges of 70 dB and 95 dB respectively. Frame rates were generally 10-20 fps, and video frames were tonemapped according to [66]. Figure 5.1 shows representative exposure-images from the two videos and Figure 5.2 shows the integration times selected for the video frames. Note that integration times are generally chosen to expose pixels corresponding to the moving object, with other pixels relying on *a priori* information for their radiance estimates. The dominant artifacts in the HDR videos are discontinuous motion and ghosting. These artifacts are to be expected since pixels affected by motion do not always receive immediate measurement updates. Unlike other approaches, a motion estimator is not used, though it could be used to predict the time evolution of the scene. This would likely to improve performance of the technique.

5.2   Multispectral HDRI of Metallic Artifacts

Another application for the HDRI methods developed in this dissertation is in the area of art and artifact conservation. As discussed in [5] and [16], in art conservation applications there is much interest in performing multispectral image analysis. Multispectral information can not only be used to provide highly accurate color information, but can also be useful in analysis, authentication, and restoration applications. While multispectral imaging may be desirable, this is difficult to accomplish using traditional methods when a scene has a high dynamic range of intensities. One such

(a)

(b)

Figure 5.1: Frames from a from a HDR video sequence of (a) a moving matchbox and (b) a candle.

scene is a copper scroll of historical significance, and another is a gold foil with embossed text, shown in Figure 5.3 and Figure 5.4.

In this application scene illumination can be controlled and calibration is permissible, which makes the Kalman filtering based HDRI method of Chapter 2 an excellent
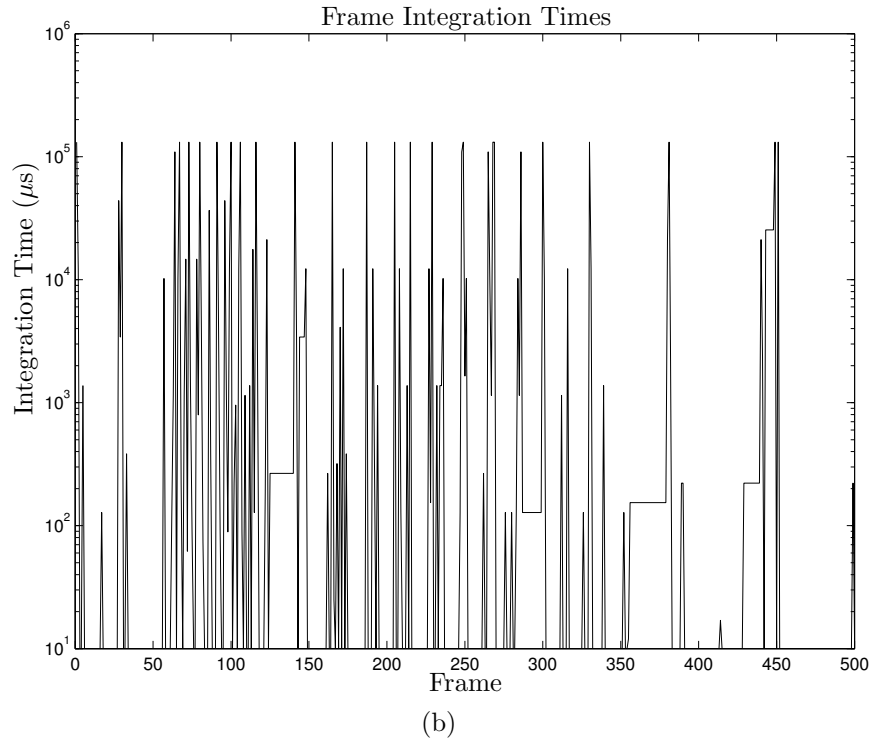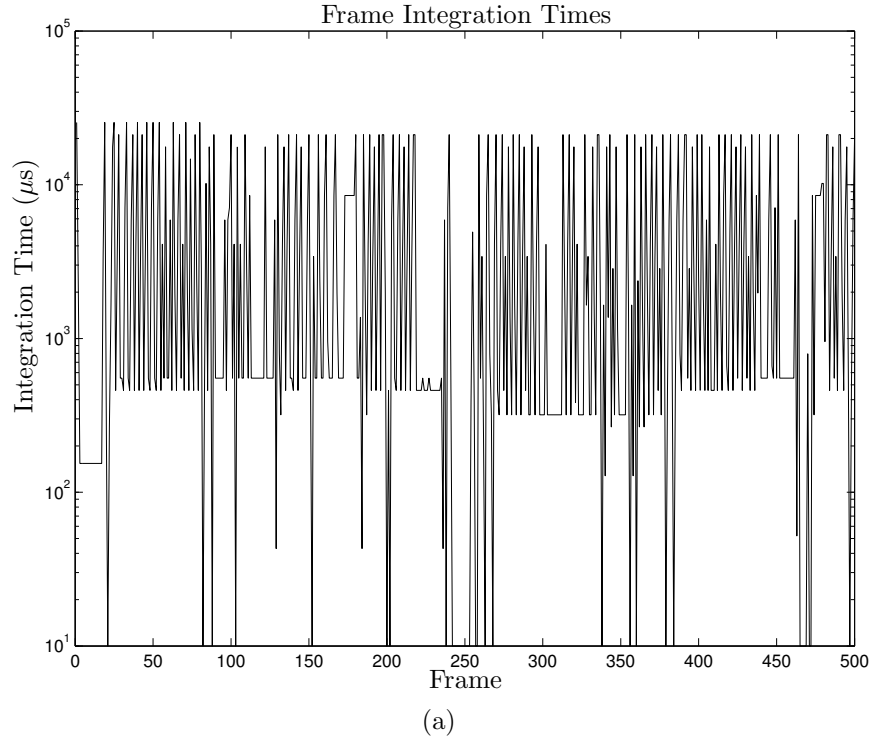
Figure 5.2: Integration times chosen for the (a) matchbox and (b) candle videos sequences.

approach to achieving multispectral HDRI. In this application, HDRI is performed as described in Chapter 2, except that multiple interference filters are used. These interference filters each have a 20 nm spectral bandwidth, with the set ranging from 400 nm to 700 nm. Since a uniform target is used to correct for spatial non-uniformity in illumination and since the calibration is known to be spectrally flat under different illuminants, multispectral HDRI can be achieved in a manner which is independent of illumination conditions. Based on the calibration target, different spectral channels can be appropriately scaled, thus compensating for both the quantum efficiency of the image sensor and the spectral power of the illuminant. Furthermore, non-uniformity correction in the Kalman filtering based HDRI approach allows for robustness in the presence of non-uniformities, which may be caused by optical vignetting or be due to other sources such as dirt on lenses and filters.

5.3   HDRI in an Assay System

The Kalman filtering based HDRI approach has also been used as a component technology in a assay selection system. This application is designed to detect healthy cell colonies while rejecting both poor cell colonies and background clutter. Much of this project is rooted in pattern analysis and machine learning. However, the non-uniformity correction in the Kalman filtering based HDRI approach is an important pre-processing step. Improved uniformity improves the performance of many other image processing operations such as edge detection, resulting in fewer spurious edges. Since image processing operations are also performed using radiance estimates with floating point precision, image processing algorithms also have improved performance due to the reduction of quantization noise.

5.4   Multispectral HDRI and SIFT

Another application of the Kalman filter based HDRI approach is [83], where multi-spectral HDRI is performed on a pig heart, shown in Figure 5.6. This work postulates that multispectral HDRI can lead to an increased number of features detected with the local scale-invariant feature transform (SIFT) [46]. This feature detection is important in biomedical applications, such as the modeling heart dynamics. Results demonstrate that multispectral HDRI can yield an increased number of detected features, and an optimal set of wavelengths is determined.

In this dissertation defects which can occur when performing SLI in optically challenging environments are shown. One such environment is a scene with a high dynamic range in albedo. In addressing this issue it is demonstrated that many existing HDRI approaches, which are suitable for use in computational photography, are not ideal for measurement or machine vision applications. For this reason a HDRI technique is developed based on image sensor models and Kalman filtering. The precision and accuracy of the Kalman filtering approach is assessed and compared to five reference techniques and to ground truths, including an I1 spectrophotometer. This dissertation shows the importance of sensor calibration and non-uniformity

correction, and examines how the performance of a technique can be influenced by the sampling strategy employed.

One of the strengths of Kalman filtering is the number of extensions and variants which have been devised. While a basic Kalman filter may not be appropriate for all applications, it can usually be extended to meet the needs of these new challenges. The problem of HDR video is an example of this. This dissertation also shows the Kalman filtering based HDRI approach applied to several problems, including multispectral HDRI in a artifact conservation application. It is particularly well suited for this application.

While different SLI approaches can be used when the environment is known to be optically challenging, as shown earlier, some problems, such as multipath interference, can occur in subtle ways, making them difficult to anticipate and control. To help detect the presence of quality problems in PMP, tests are developed which allow quality problems such as multipath interference to be detected and quantified.

As has also been shown, optically challenging scenes may have strong inter-reflections, and some applications require these interreflections to be modeled. This dissertation develops a SLI approach based on light stripe scanning which is robust in the presence of interreflections and demonstrates how this approach can be used to model light transport. Also examined are newly developed approaches in compressed sensing which are utilized in SLI. For these reasons, it is likely that this approach can be useful in modeling light transport in other optically challenging scenes such as translucent objects, such as in [23].

All of the research in this dissertation was performed with the goal of advancing the state-of-the art in SLI scanning so that SLI approaches, with their long history of successes, can also be applied to optically challenging environments. This goal is important for future computer graphics, machine vision, and measurement applications.
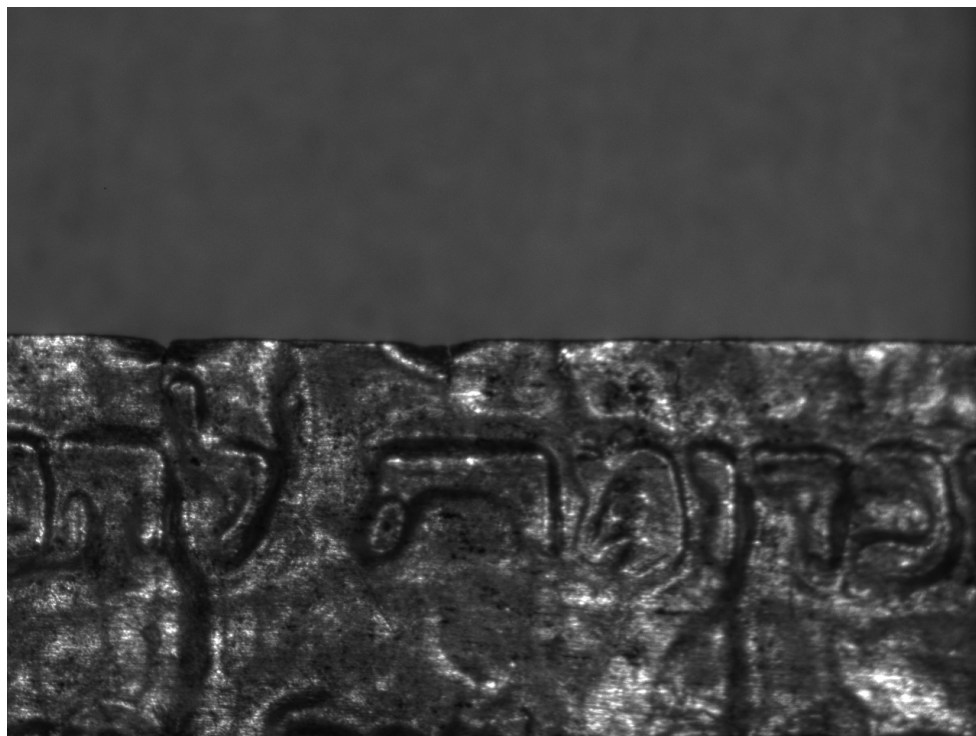
(a)
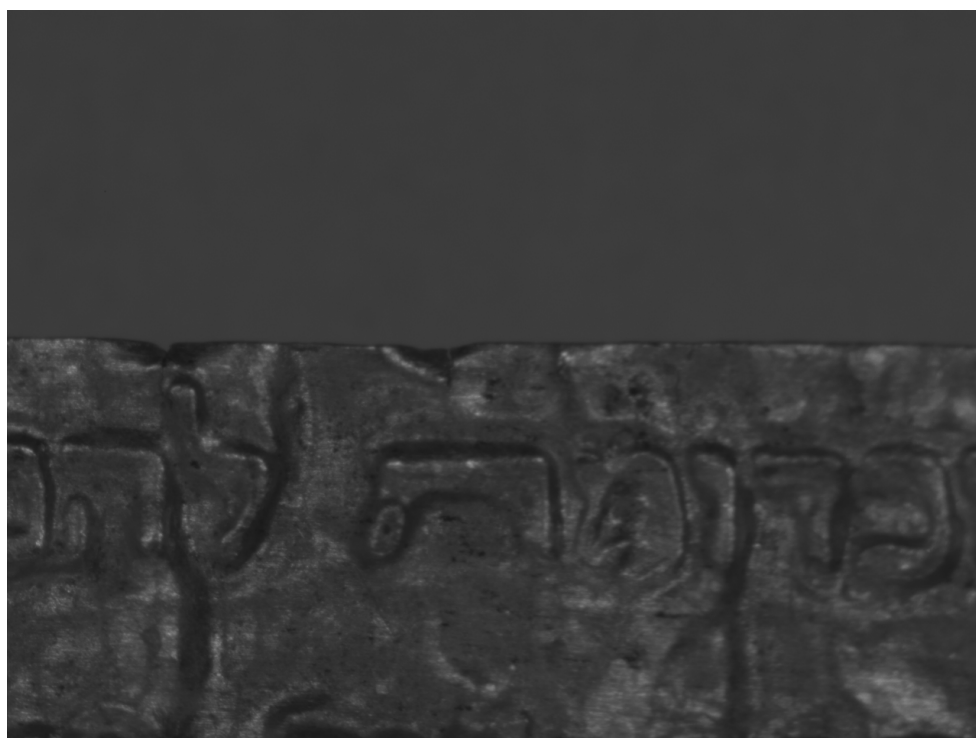
(b)

Figure 5.3: A HDR image of a piece of a gold artifact tonemapped (a) according to [51], and (b) according to [66].

(a)

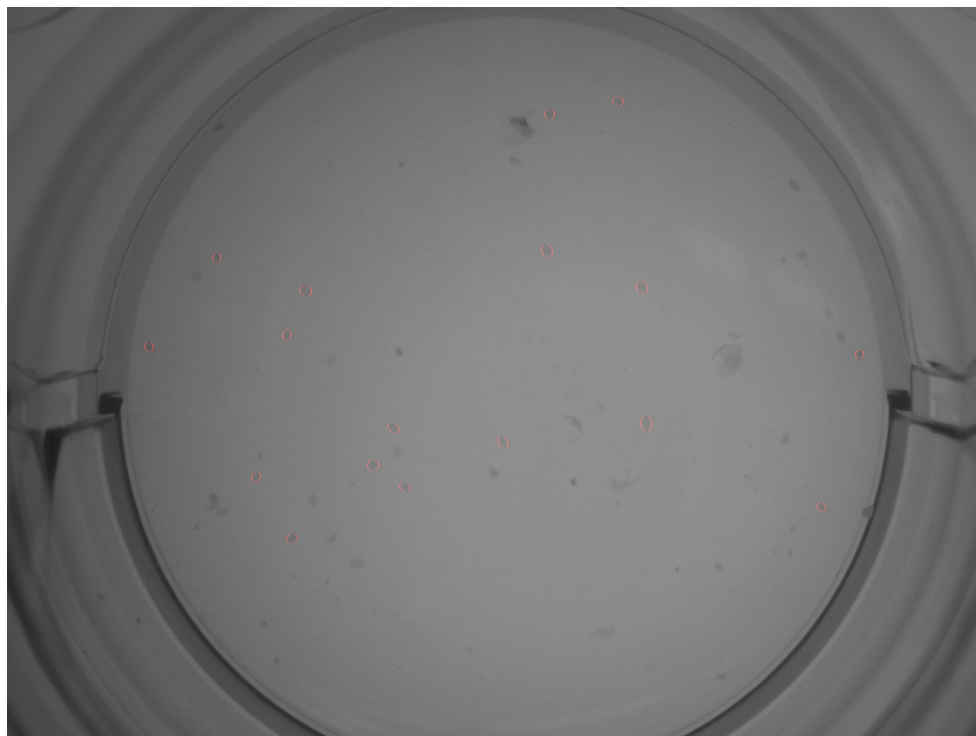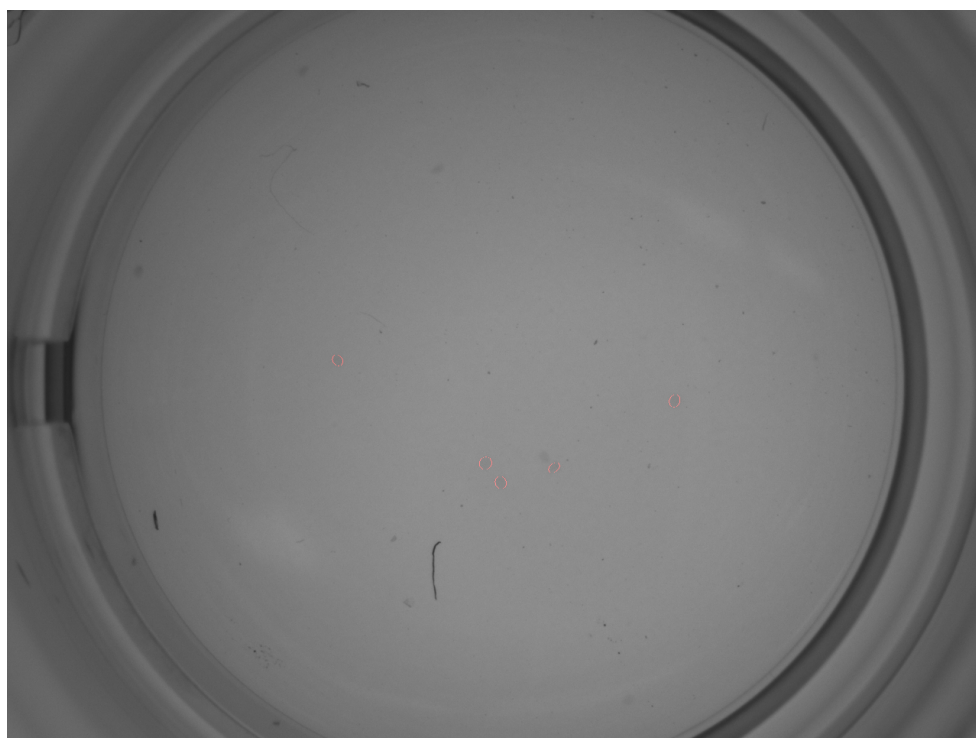
(b)

Figure 5.4: An HDR image of a different piece of the gold artifact in Figure 5.3, tonemapped (a) according to [51], and (b) according to [66].
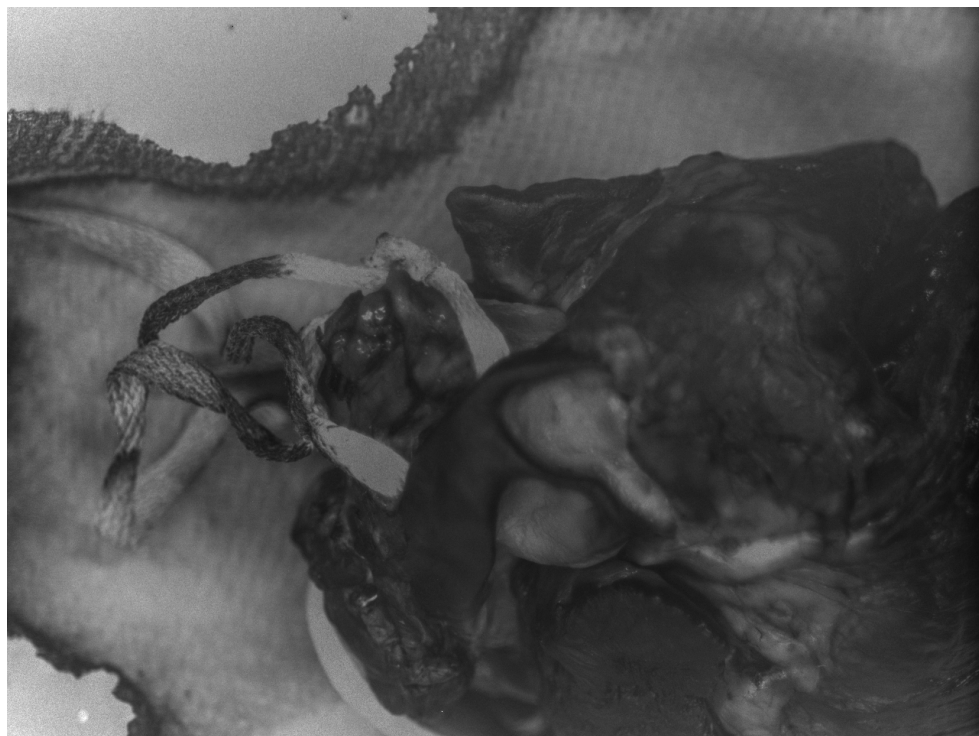
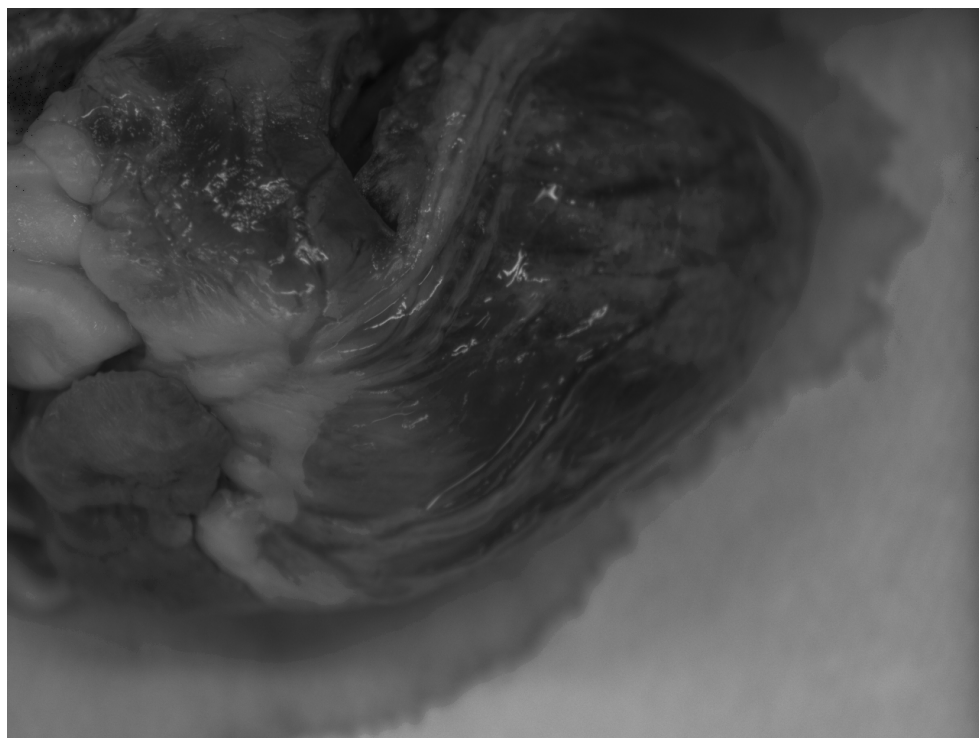(a)

(b)

Figure 5.5: Two images of cell colonies from the assay system. Red outlines indicate colonies classified as "good."

(a)

(b)

Figure 5.6: Two multispectral HDR images of a different view of a pig heart, tonemapped according to [51].

Bibliography

[1]   File:1-stripesx7.svg.   http://en.wikipedia.org/wiki/File:1-stripesx7.svg.   [On-line; accessed April 2, 2011].

[2]   M.D. Adams. Coaxial range measurement - current trends for mobile robotic applications. *Sensors Journal, IEEE*, 2(1):2 –13, February 2002.

[3]   Ahmet Oguz Akyuz and Erik Reinhard. Noise reduction in high dynamic range imaging. *Journal of Visual Communication & Image Representation*, 18(5):366–376, 2007.

[4]   J. Batlle, E. Mouaddib, and J. Salvi. Recent progress in coded structured light as a technique to solve the correspondence problem: a survey, 1998.

[5]   Roy S. Berns. The science of digitizing paintings for color-accurate image archives: A review. *Journal of Imaging Science and Technology*, 45, 2001.

[6]   J.M. Bioucas-Dias and G. Valadao. Phase unwrapping via graph cuts. *Image Processing, IEEE Transactions on*, 16(3):698 –709, march 2007.

[7]   E.J. Candes and T. Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203 – 4215, 2005.

[8]   E.J. Candes and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406 –5425, dec. 2006.

[9]   Emmanuel J. Candès, Justin K. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 59(8):1207–1223, August 2006.

[10]   Frank M. Candocia and Daniel A. Mandarino. A semiparametric model for accurate camera response function modeling and exposure estimation from comparametric data. *IEEE Transactions on Image Processing*, 14(8):1138–1150, 2005.

[11]   Peter B. Catrysse, Xinqiao Liu, and Abbas El Gamal. Qe reduction due to pixel vignetting in cmos image sensors. *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications*, 3965:420–430, 2000.

[12]   Scott S Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, (20):33–61, 1998.

[13]   Tongbo Chen, H.P.A. Lensch, C. Fuchs, and H.-P. Seidel. Polarization and phase-shifting for 3d scanning of translucent objects. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1 –8, June 2007.

[14] James Clark, Emanuele Trucco, and Lawrence B. Wolff. Using light polarization in laser scanning. *Image Vision Comput.*, 15:107–117, February 1997.

[15] Albert Cohen, Wolfgang Dahmen, and Ronald Devore. Compressed sensing and best k-term approximation. *J. Amer. Math. Soc*, pages 211–231, 2009.

[16] Bruno Cornelis, Ann Dooms, Frederik Leen, Adrian Munteanu, and Peter Schelkens. Multispectral imaging for digital painting analysis: a gauguin case study. volume 7798, page 77980I. SPIE, 2010.

[17] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar. Acquiring the reflectance field of a human face. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 145–156, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[18] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH'97 Conference Proceedings*, August 1997.

[19] David L. Donoho. For most large underdetermined systems of equations, the minimal l1-norm near-solution approximates the sparsest near-solution. *Communications on pure and applied mathematics*, 59(7):907–934, 2006.

[20] David L. Donoho, Yaakov Tsaig, Iddo Drori, and Jean luc Starck. Sparse solution of underdetermined linear equations by stagewise orthogonal matching pursuit. Technical report, 2006.

[21] A. El Gamal and H. Eltoukhy. Cmos image sensors. *IEEE Circuits & Devices Magazine*, 21, 2005.

[22] H. Farid. Blind inverse gamma correction. *Image Processing, IEEE Transactions on*, 10(10):1428–1433, October 2001.

[23] Michael Goesele, Hendrik P. A. Lensch, Jochen Lang, Christian Fuchs, and Hans-Peter Seidel. Disco: acquisition of translucent objects. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH 2004, pages 835–844, New York, NY, USA, 2004. ACM.

[24] Michael Goeselle and Karol Myszkowski. *HDR Applications in Computer Graphics*. Springer Berlin Heidelberg, 1997.

[25] S.B. Gokturk, H. Yalcin, and C. Bamji. A time-of-flight depth sensor - system description, issues and solutions. In *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW '04. Conference on*, page 35, 2004.

[26] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing, 2nd ed.* Prentice Hall, 2002.

[27] R. Gribonval and M. Nielsen. Sparse representations in unions of bases. *Information Theory, IEEE Transactions on*, 49(12):3320 – 3325, 2003.

[28] Michael D. Grossberg and Shree K. Nayar. Determining the camera response from images: What is knowable? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1455–1467, 2003.

[29] Michael D. Grossberg and Shree K. Nayar. Modeling the space of camera response functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:1272–1282, 2004.

[30] Jinwei Gu, Shree Nayar, Eitan Grinspun, Peter Belhumeur, and Ravi Ramamoorthi. Compressive structured light for recovering inhomogeneous participating media. In David Forsyth, Philip Torr, and Andrew Zisserman, editors, *Computer Vision ECCV 2008*, volume 5305 of *Lecture Notes in Computer Science*, pages 845–858. Springer Berlin / Heidelberg, 2008.

[31] Russel C. Hardie, Majeed M. Hayat, Earnest Armstrong, and Brian Yasuda. Scene-based nonuniformity correction with video sequences and registration. *Appl. Opt.*, 39:1241–1250, 2000.

[32] Glenn E. Healey and Raghava Kondepudy. Radiometric CCD camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(3):267–276, March 1994.

[33] InFocus Corporation, 27700B SW Parkway Ave. Wilsonville OR, 97070. *InFocus LP70+ User's Guide*, January 2004.

[34] T. R. Judge and P. J. Bryanston-Cross. A review of phase unwrapping techniques in fringe analysis. *Optics and Lasers in Engineering*, 21(4):199 – 239, 1994.

[35] Sing Bing Kang, Matthew Uyttendaele, Simon Winder, and Richard Szeliski. High dynamic range video. *ACM Trans. Graph.*, 22(3):319–325, 2003.

[36] Sam Kavusi and Abbas El Gamal. Folded multiple-capture: an architecture for high dynamic range disturbance-tolerant focal plane array. volume 5406, pages 351–360. SPIE, 2004.

[37] Dae-Woong Kim and Ki-Sang Hong. High dynamic range global mosaic. In *Computer Vision –ACCV 2006*, pages 744–753. Springer Berlin / Heidelberg, 2006.

[38] David Koller, Michael Turitzin, Marc Levoy, Marco Tarini, Giuseppe Croccia, Paolo Cignoni, and Roberto Scopigno. Protected interactive 3d graphics via remote rendering. *ACM Trans. Graph.*, 23:695–703, August 2004.

[39] David Koller, Michael Turitzin, Marc Levoy, Marco Tarini, Giuseppe Croccia, Paolo Cignoni, and Roberto Scopigno. Protected interactive 3d graphics via remote rendering. In *ACM SIGGRAPH 2004 Papers*, SIGGRAPH '04, pages 695–703, New York, NY, USA, 2004. ACM.

[40] T.P. Koninckx, P. Peers, P. Dutre, and L. Van Gool. Scene-adapted structured light. volume 2, pages 611 – 618 vol. 2, jun. 2005.

[41] M. Levoy. The digital michelangelo project. In *3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on*, pages 2 – 11, 1999.

[42] Stephen Lin and Lei Zhang. Determining the radiometric response function from a single grayscale image. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.*, pages 66–73, June 2005.

[43] T. Lindeberg and J. Garding. Shape from texture from a multi-scale perspective. In *Computer Vision, 1993. Proceedings., Fourth International Conference on*, pages 683 –691, May 1993.

[44] Dave Litwiller. Ccd vs. cmos: Facts and fiction. *Photonics Spectra*, pages 154–158, 2001.

[45] Kai Liu, Yongchang Wang Wang, Daniel L. Lau, Qi Hao, and Laurence G. Hassebrook. Dual-frequency pattern scheme for high-speed 3-d shape measurement. *Opt. Express*, 18:5229–5244, 2010.

[46] D.G. Lowe. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, volume 2, pages 1150 –1157 vol.2, 1999.

[47] Brian C. Madden. Extended intensity range imaging. Technical report, GRASP Laboratory, University of Pennsylvania, 1993.

[48] S. Mann and R. Mann. Quantigraphic imaging: Estimating the camera response and exposures from differently exposed images. *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 1:I–842–I–849, 2001.

[49] S. Mann and R. W Picard. Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. Technical Report 323, M.I.T. Media Lab Perceptual Computing Section, Boston, Massachusets, 1994. Also appears, IS&T's 48th annual conference, Cambridge, Massachusets, May 1995.

[50] Steve Mann. Comparametric equations with practical applications in quantigraphic image processing. *IEEE Transactions on Image Processing*, 9:1389–1406, 2000.

[51] Rafal Mantiuk, Karol Myszkowski, and Hans-Peter Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.*, 3:286–308, July 2006.

[52] O. Marklund. An anisotropic evolution formulation applied in 2-d unwrapping of discontinuous phase surfaces. *Image Processing, IEEE Transactions on*, 10(11):1700 –1711, nov 2001.

[53] Micron Technologies, Inc. *1/2-inch VGA (with Freeze-Frame) CMOS Image Sensor*.

[54] T. Mitsunaga and S.K. Nayar. Radiometric self calibration. *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, 1:374–380, 1999.

[55] O. Nasu, S. Hiura, and K. Sato. Analysis of light transport based on the separation of direct and indirect components. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1 –2, June 2007.

[56] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM Journal on Computing*, 24(2):227–234, 1995.

[57] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pages 935–944, New York, NY, USA, 2006. ACM.

[58] D. Needell and J.A. Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301 – 321, 2009.

[59] Deanna Needell and Roman Vershynin. Uniform uncertainty principle and signal recovery viaregularized orthogonal matching pursuit. *Foundations of Computational Mathematics*, 9:317–334, 2009. 10.1007/s10208-008-9031-3.

[60] Arnold Neumaier. Solving ill-conditioned and singular linear systems: A tutorial on regularization. *SIAM Rev.*, 40:636–666, September 1998.

[61] C.L. Niclass, A. Rochas, P.A. Besse, and E. Charbon. A cmos single photon avalanche diode array for 3d imaging. In *Solid-State Circuits Conference, 2004. Digest of Technical Papers. ISSCC. 2004 IEEE International*, pages 120 – 517 Vol.1, 2004.

[62] John Oliensis. A critique of structure-from-motion algorithms. *Computer Vision and Image Understanding*, 80(2):172 – 214, 2000.

[63] Christopher C. Paige and Michael A. Saunders. Lsqr: An algorithm for sparse linear equations and sparse least squares. *ACM Trans. Math. Softw.*, 8:43–71, March 1982.

[64] C. Pal, R. Szeliski, M. Uyttendaele, and N. Jojic. Probability models for high dynamic range imaging. *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2:II–173–II–180, 2004.

[65] J. Park and A.C. Kak. 3d modeling of optically challenging objects. *Visualization and Computer Graphics, IEEE Transactions on*, 14(2):246 –262, March-April 2008.

[66] Sumanta N. Pattanaik, Jack Tumblin, Hector Yee, and Donald P. Greenberg. Time-dependent visual adaptation for fast realistic image display. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, SIGGRAPH '00, pages 47–54, New York, NY, USA, 2000. ACM Press/Addison-Wesley Publishing Co.

[67] Pieter Peers, Dhruv K. Mahajan, Bruce Lamond, Abhijeet Ghosh, Wojciech Matusik, Ravi Ramamoorthi, and Paul Debevec. Compressive light transport sensing. *ACM Trans. Graph.*, 28:3:1–3:18, February 2009.

[68] David L. Perry and Eustace L. Dereniak. Linear theory of nonuniformity correction in infrared staring sensors. *Optical Engineering*, 32(8):1854–1859, 1993.

[69] Robert F. Pierret. *Semiconductor Device Fundamentals*. Addison Wesley Longman, 1996.

[70] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.

[71] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C*. Cambridge University Press, 1992.

[72] Prosilica Inc. *GC640 User Manual*.

[73] Austin A. Richards and Brian K. Cromwell. Superframing: Scene dynamic range extension of infrared cameras. volume 5612, 2004.

[74] M. A. Robertson, S. Borman, and R. L. Stevenson. Estimation-theoretic approach to dynamic range enhancement using multiple exposures. *Journal of Electronic Imaging*, 12:219–228, 2002.

[75] Joaquim Salvi, Jordi Pags, and Joan Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37(4):827 – 849, 2004. Agent Based Computer Vision.

[76] Shuichi Sasa. Robustness of a kalman filter against uncertainties of noise covariances. volume 4, pages 2344–2348 vol.4, Jun 1998.

[77] D Scharstein and R Szeliski. High-accuracy stereo depth maps using structured light. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 195–202, 2003.

[78] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 519 – 528, 2006.

[79] Steven M. Seitz, Yasuyuki Matsushita, and Kiriakos N. Kutulakos. A theory of inverse light transport. In *Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2*, ICCV '05, pages 1440–1447, Washington, DC, USA, 2005. IEEE Computer Society.

[80] Pradeep Sen and Soheil Darabi. Compressive Dual Photography. *Computer Graphics Forum*, 28(2):609 – 618, 2009.

[81] D. Simon. *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. John Wiley & Sons, Inc., 2006.

[82] D. Skocaj and A. Leonardis. Range image acquisition of objects with non-uniform albedo using structured light range sensor. volume 1, pages 778 –781 vol.1, 2000.

[83] Pratik Sogani. The Paradox of Overfitting. Master's thesis, Center for Biomedical Engineering, University of Kentucky, 2010.

[84] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht. Wide-dynamic-range cmos image sensors, comparative performance analysis. *Electron Devices, IEEE Transactions on*, 56(11):2446 –2461, 2009.

[85] E. Stoykova, A.A. Alatan, P. Benzie, N. Grammalidis, S. Malassiotis, J. Ostermann, S. Piekh, V. Sainov, C. Theobalt, T. Thevar, and X. Zabulis. 3-d time-varying scene capture technologies, a survey. *Circuits and Systems for Video Technology, IEEE Transactions on*, 17(11):1568 –1586, 2007.

[86] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58:267–288, 1996.

[87] Sergio N. Torres and Majeed M. Hayat. Kalman filtering for adaptive nonuniformity correction in infrared focal-plane arrays. *J. Opt. Soc. Am. A*, 20:470–480, 2003.

[88] J.A. Tropp and A.C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *Information Theory, IEEE Transactions on*, 53(12):4655 –4666, Dec. 2007.

[89] E. Trucco and R.B. Fisher. Acquisition of consistent range data using local calibration. In *Robotics and Automation, 1994. Proceedings., 1994 IEEE International Conference on*, pages 3410 –3415 vol.4, May 1994.

[90] Yanghai Tsin, Visvanathan Ramesh, and Takeo Kanade. Statistical calibration of CCD imaging process. In *IEEE International Conference on Computer Vision (ICCV'01)*, pages 480–487, July 2001.

[91] Luiz Velho. Histogram-based hdr video. In *SIGGRAPH '07: ACM SIGGRAPH 2007 posters*, page 62, New York, NY, USA, 2007. ACM.

[92] Eric W. Weisstein. Sample variance distribution. http://mathworld.wolfram.com/SampleVarianceDistribution.html.

[93] G Welch and G Bishop. An introduction to the kalman filter. Technical report, 2001.

[94] C. W Wyckoff. An experimental extended response film. *S.P.I.E. NEWSLETTER*, JUNE-JULY 1962.

[95] Yi Xu and D.G. Aliaga. An adaptive correspondence algorithm for modeling scenes with strong interreflections. *Visualization and Computer Graphics, IEEE Transactions on*, 15(3):465 –480, May. 2009.

[96] Keiichi Yamada, Tomoaki Nakano, Shin Yamamoto, Eisaku Akutsu, and Keiji Aoki. Wide dynamic range vision sensor for vehicles. In *IEEE International Conference on Vehicle Navigation and Information Systems*, pages 405–408, 1994.

[97] Paul Zarchan and Howard Musoff. *Fundamentals of Kalman Filtering: A Practical Approach, 2nd ed.* American Institute of Aeronautics and Astronautics, Inc., 2005.

[98] Ruo Zhang, Ping-Sing Tsai, J.E. Cryer, and M. Shah. Shape-from-shading: a survey. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(8):690 –706, August 1999.

[99] Song Zhang and Shing-Tung Yau. High dynamic range scanning technique. *Optical Engineering*, 48(3):033604, 2009.

[100] Hui Zhao, Huajun Feng, Zhihai Xu, and Qi Li. Research on temperature distribution of combustion flames based on high dynamic range imaging. *Optics & Laser Technology*, 39(7):1351–1359, October 2007.

[101] D. Ziou. Passive depth from defocus using a spatial domain approach. In *Computer Vision, 1998. Sixth International Conference on*, pages 799 –804, January 1998.

Vita

Eric Dedrick was born on August 20, 1975, in Lexington, KY. He graduated from Hanover College with a B.A. in Physics in 1998. He then graduated from Purdue University with a M.S. in Physics in 2001. He was a software engineer with Lexmark International for four years.