

Femtosecond Transient Imaging

by

Ahmed Kirmani

Integrated Masters in Mathematics and Computing
Indian Institute of Technology Delhi (2008)

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
in partial fulfillment of the requirements for the degree of

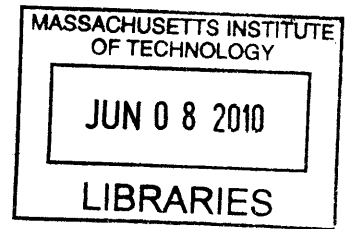
Master of Science in Media Technology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

ARCHIVES



© Massachusetts Institute of Technology 2010. All rights reserved.

Author _____
Program in Media Arts and Sciences
May 7, 2010

Certified by _____
Ramesh Raskar
Associate Professor of Media Arts and Sciences
Program in Media Arts and Sciences
Thesis Supervisor

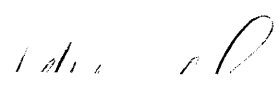
Accepted by _____
Prof. Pattie Maes
Associate Academic Head
Program in Media Arts and Sciences

Femtosecond Transient Imaging

by


Ahmed Kirmani

The following people served as readers for this thesis:



Thesis Reader _____

Vivek Goyal
Associate Professor of Electrical Engineering and Computer Science
Research Laboratory of Electronics at MIT



Thesis Reader _____

Pablo Parrilo
Professor of Electrical Engineering and Computer Science
Laboratory of Information and Decision Systems at MIT

Femtosecond Transient Imaging

by

Ahmed Kirmani

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on May 7, 2010, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Technology

Abstract

This thesis proposes a novel framework called *transient imaging* for image formation and scene understanding through impulse illumination and time images. Using time-of-flight cameras and multi-path analysis of global light transport, we pioneer new algorithms and systems for scene understanding through time images. We demonstrate that our proposed transient imaging framework provides opportunities to accomplish tasks that are well beyond the reach of existing imaging technology. For example, one can infer the geometry of not only the visible but also the hidden parts of a scene, enabling us to *look around corners*.

Traditional cameras estimate intensity per pixel $I(x, y)$. Our transient imaging camera prototype captures a 3D time-image $I(x, y, t)$ for each pixel and uses an ultra-short pulse laser for flash illumination. Emerging technologies are supporting cameras with a temporal-profile per pixel at picosecond resolution, allowing us to capture an ultra-high speed time-image. This time-image contains the time profile of irradiance at a sensor pixel. The speed of light is relevant at these imaging time scales, and the transient properties of light transport come into play. In particular we furnish a novel framework for reconstructing scene geometry of hidden planar scenes. We experimentally corroborated our theory with free space hardware experiments using a femtosecond laser and a picosecond accurate sensing device.

The ability to infer the structure of hidden scene elements, unobservable by both the camera and illumination source, will create a range of new computer vision opportunities.

Thesis Supervisor: Ramesh Raskar

Title: Associate Professor of Media Arts and Sciences, Program in Media Arts and Sciences

Acknowledgements

My deepest gratitude goes to my advisor Ramesh Raskar being an amazing and supportive mentor. I would like to thank him for giving me the opportunity to go to graduate school and nurturing my desire to do research.

I consider myself very fortunate to get the opportunity to work closely with Professors Vivek Goyal, Pablo Parrilo and Mounqi Bawendi. I would like to thank them for their patience and guidance. My interactions with them has made a major and ever lasting impact on my thinking and the style of my research.

I am also very thankful to Prof. Mounqi Bawendi (Chemistry), Prof. Neil Gershenfeld (CBA), Prof. Joseph Paradiso(Media Lab), Dr. Franco Wong (RLE), Dr. Manu Prakash (MIT, now Professor at Stanford) for making available their opto-electronic apparatus that made our experiments possible. I would also like to thank MIT EECS Professors George Verghese, Franz Kaertner, Rajeev Ram for invaluable discussions. MIT undergraduate and graduate students: Tyler Hutchison and Josh Wang who deserve a special mention and Biyeun Buczyk, George Hansel, assisted in carrying out initial experiments.

My co-workers and friends at MIT made the experience really enjoyable and memorable. I would like to single out: Ankit Mohan, Tyler Hutchison, Aditya Bhakta, Andrea Colaco, Jaewon Kim, John Sun, Lav Varshney, Daniel Weller, Parikshit Shah, Vaibhav Rathi, Raghavendra Hosur and Aniruddh Sarkar.

This thesis would not have been possible without the support and sacrifice of my family, in particular my mother and grand mother. I thank them for their love and care. I would like to dedicate this thesis to the memory of my father.

Ahmed Kirmani
MIT

Contents

1	Introduction	15
1.1	Motivation	15
1.2	Problem and Solution	20
1.3	Related Work	21
1.3.1	Theoretical Prior Art	24
1.3.2	Hardware and Experimental Prior Art	26
1.4	Contributions	28
2	Transient Imaging Framework	29
2.1	Formulation and Terminology	30
2.2	Transient Imaging Camera	33
2.2.1	Space Time Impulse Response	33
3	Hidden Geometry Estimation	37
3.1	Inverse Geometry Problem	37
3.1.1	Distances from STIR	38
3.1.2	Structure from Pairwise Distances	39
3.1.3	Scenes with Occluders	41
4	Hardware Prototype and Experiments	45
4.1	Transient Imaging Camera Prototype	45
4.2	Hidden Geometry Reconstruction Experiments	47
4.2.1	Missing direct reflection	47
4.2.2	Looking Around the Corner	47
5	Limitations and Future Work	51
5.1	Theoretical Limitations	51
5.2	Experimental Limitations	52
5.3	Future Work	53
6	Overview, Applications and Conclusion	55
6.1	Summary of Approach	55
6.2	Applications	56
6.3	Conclusion	56

List of Figures

1-1	<i>Can you look around the corner into a room with no imaging device in the line of sight?</i> We have demonstrated that by emitting short pulses (labeled 1-2), and analyzing multi-bounce reflection from the door (4-1), we can infer hidden geometry even if the intermediate bounces (3) are not visible. Our transient imaging camera prototype comprises of (a) Femtosecond laser illumination (b) Picosecond-accurate detectors and (c) an Ultrafast sampling oscilloscope. We measured the Space Time Impulse Response (STIR) of a scene (d) containing a hidden 1-0-1 barcode and reconstructed the hidden barcode positions.	16
1-2	Parameterizing a ray in 3D space by position (x, y, z) and direction (θ, ϕ)	17
1-3	A flowchart describing our approach to scene understanding using time images: flow of information from the transient imaging camera hardware (a) into the transient imaging theoretical framework (b) which uses the time image measurement to estimate scene properties like hidden surface structure and reflectance.	22
1-4	How can you look around corners in this scene?	23
1-5	Popular imaging methods plotted in the <i>Space-Angle-Time axes</i> . With higher dimensional light capture, we expand the horizons of scene understanding. Our paper uses LIDAR-like imaging hardware, but, in contrast, we exploit the multi-path information which is rejected in both LIDAR (LIght Detection And Ranging) and OCT (Optical coherence tomography).	25
2-1	A scene consisting of $M = 5$ patches and the illumination-camera patch p_0 . The patches have different spatial coordinates (z_i^x, z_i^y, z_i^z) , orientations \hat{n}_i and relative visibility between patches v_{ij} . The patches also have different material properties, for instance p_4 is diffused, p_4 is translucent and p_5 is a mirror.	31
2-2	A ray impulse $E_{02}[t]$ directed towards patch p_2 at time t . This ray illuminates p_2 at time instant $t + \delta_{02}$ and generates the directional radiance vector $[L_{20}[t + \delta], L_{21}[t], L_{23}[t]]$. These light rays travel towards the camera p_0 and scene patches p_1 and p_3 resulting in global illumination	32
2-3	Scalar form factors f_{kij} are the proportion of light incident from patch p_k on to p_i that will be directed towards p_j	32
2-4	Transient Imaging Camera captures a 3D time image in response to a ray impulse illumination.	33
2-5	Measuring the STIR of a scene using the transient imaging camera. We successively illuminate a single patch at a time and record a 3D time image. Collection of such time images constitutes the 5D STIR.	35

3-1	Measuring the STIR of a scene with 3 patches using the transient imaging camera. We successively illuminate a single patch and record a 3D time image. Collection of such time images creates a 5D STIR.	38
3-2	(a) Estimating distances in an all-visible scene comprising of 3 rectangles which are discretized as 49 patches. Note that reflectance is not relevant. (b) Original geometry shows the surface normals in green. (c) We used noisy 1 st and 2 nd time onsets (Gaussian noise $\sim \mathcal{N}(\mu, \sigma^2)$, $\mu = \text{device resolution} = 250\text{ps}$ and $\sigma = 0.1$) to estimate the distances using the \mathbf{T}_2 operator (inset shows enlarged view). (d) This is followed by isometric embedding and surface fitting. The reconstruction errors are plotted. Color bar shows %-error in reconstructed coordinate values.	40
3-3	A scene with $M = 4$ patches. Patches p_2 and p_3 are hidden. The blue (first) and green (second) onsets are a result of directly observing visible patches p_1 and p_4 . The pattern of arrival of third onsets depends on the relative distance of the hidden patches p_2 and p_3 from the visible patches. The onsets that correspond to light traversing the same Euclidean distance are readily identified. Once the onsets are labeled, they are used to obtain distances that involve hidden patches.	42
3-4	(a) Estimating distances in scenes with hidden patches. Unknown to the estimation algorithm, the hidden patches are on a plane (shown in black). (b) Original patch geometry. We use 1 st , 2 nd and 3 rd bounce onsets, our labeling algorithm and the \mathbf{T}_3 operator (c) to estimate hidden geometry. (d) The isometric embedding error plot verifies negligible reconstruction error and near co-planarity of patches. Onset noise and color bar schemes are same as Figure 3-2.	44
4-1	Design and Verification of a Transient Imaging Camera. (a) The ray impulses are recorded after being attenuated by a varying neutral density filter. The peak pulse intensity decreases linearly with the attenuation. (b) The intensity of the first bounce from a diffuser obeys the inverse square fall-off pattern. (c) We are able to record pulse intensities that are discernible from the noise floor even after the ray impulse has been reflected by three (2 diffuse and 1 specular) patches. The time shifts are linearly proportional to the multi-path length.	46
4-2	Missing direct reflection (Top): (a) A photo of the setup. (b) Ray diagram describing the light pulse path in 3 raster scans. (c) Plot showing multilateration using the 3 raster scans data: original and reconstructed scene geometries. (d) Oscilloscope data plot showing the TDOA between the 2 nd bounce and the reference signal. Looking around the corner (Bottom): (a) A photo of the setup showing hidden 1-0-1 barcode. The sensors and the laser are completely shielded from the barcode. (b) Ray diagram tracing the paths of 1 st and 3 rd bounces in the 2 raster scans. (c) Plot showing the scene geometry reconstructed using multilateration. (d) Oscilloscope data plot showing the 1 st bounce and the two separately recorded 3 rd bounces, for both raster scans. Note the very small delay ($\leq 200\text{ps}$) between two 3 rd bounce arrivals. Please zoom in the PDF version for details.	49

6-1 Some of the application areas which will benefit from transient imaging and looking around the corner: rescue and planning, vehicle navigation around blind corners, less-invasive bio-medical imaging and a new generation of highly mobile and light weight medical scanners. 57

Chapter 1

Introduction

How can we use time-of-flight cameras for performing multi-path analysis of global light transport? If so, then how can this allow the design of new algorithms and systems for scene understanding through time images? What are the tasks that this new reasoning can accomplish that are well beyond the reach of modern imaging technology? Can we use such a framework for estimating 3D geometry and appearance of not only visible but also hidden parts of a scene? Can we infer object shapes, motion reflection around an opaque occluder? Can you automatically label the objects in the scene based on their material index (wood, mirror, glass etc.) using a single image? We propose a novel theoretical framework for image formation and scene understanding using impulse illumination and time images. We call this framework *Transient imaging*. We also propose the design of an ultrafast imaging device based on this framework which we term as a *Transient imaging camera*. We used our inverse algorithms and measurements collected using the transient imaging camera to estimate the hidden scene geometry from a fixed single viewpoint.

1.1 Motivation

Light is a very rich signal and carries information that allows us to observe and interpret the world around us. Mathematically, light rays may be represented using *Light fields* [1]. The light field or the *Plenoptic function* describes the amount of light traveling in every direction through every

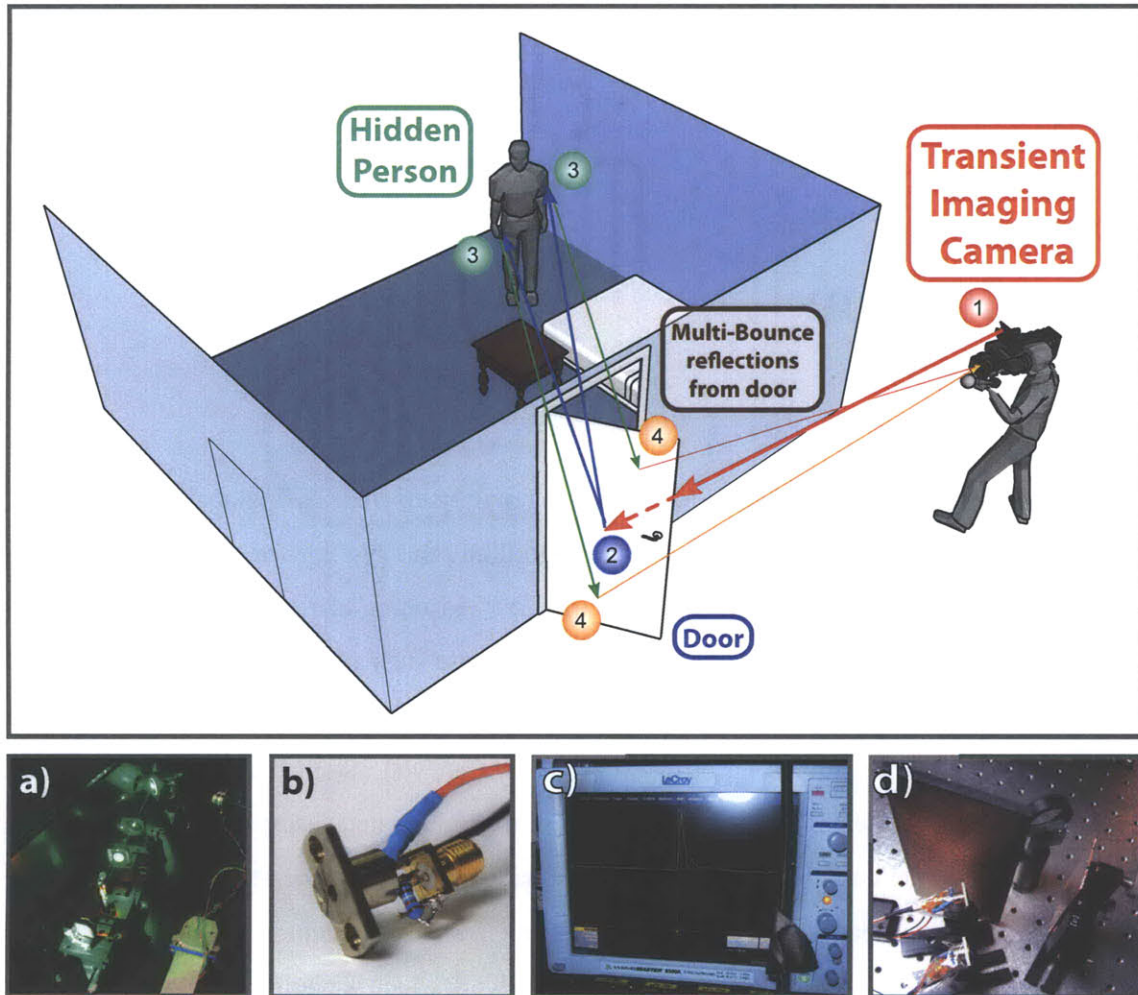


Figure 1-1: *Can you look around the corner into a room with no imaging device in the line of sight?* We have demonstrated that by emitting short pulses (labeled 1-2), and analyzing multi-bounce reflection from the door (4-1), we can infer hidden geometry even if the intermediate bounces (3) are not visible. Our transient imaging camera prototype comprises of (a) Femtosecond laser illumination (b) Picosecond-accurate detectors and (c) an Ultrafast sampling oscilloscope. We measured the Space Time Impulse Response (STIR) of a scene (d) containing a hidden 1-0-1 barcode and reconstructed the hidden barcode positions.

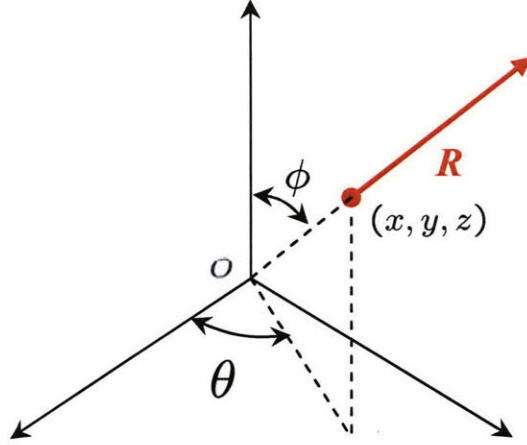


Figure 1-2: Parameterizing a ray in 3D space by position (x, y, z) and direction (θ, ϕ) .

point in space at any wavelength at any point in time. Michael Faraday was the first to propose that light should be interpreted as a field, much like the magnetic fields. The phrase light field was coined by Alexander Gershun in a classic paper on the radiometric properties of light in three-dimensional space (1936). The phrase has been redefined by researchers in computer graphics as the plenoptic function $L(x, y, z, \theta, \phi, \lambda, t)$ which is a function of three spatial dimensions, two angular dimensions, wavelength and time (see Figure 1-2). The time parameter is usually ignored since traditionally all measurements are made in steady state, where the ray intensities do not change over time. The plenoptic function encodes information about the scene and has a lot of structure which governed by the scene properties. The simplest example is the encoding of the reflectance of scene elements in the light field. A ordinary digital camera captures this information into a 2D color image $I(x, y, c)$ by marginalizing $L(x, y, z, \theta, \phi, \lambda, t)$ and sampling the wavelength as:

$$I(x, y, c) = \int_{t=0}^{T_{exp}} \int_z \int_{\theta} \int_{\phi} L(x, y, z, \theta, \phi, \lambda, t) \delta(\lambda - c) \quad (1.1)$$

In a traditional camera, the light incident at a pixel is integrated along angular, temporal and wavelength dimensions during the exposure time to record a single intensity value. Traditional cameras sense a limited 2D projection of the complete light field of the scene as defined by Equation 1.1.

Many distinct scenes result in identical projections, and thus identical pixel values. Hence it is very difficult to use traditional imaging to estimate properties such as the depth of a mirror, the position of a sheet of glass, or the overall scale of a scene because these properties are not directly observable in the RGB values reported by a traditional camera. Another seemingly impossible problem is to estimate the structure and geometry of scene elements that are occluded from the camera and the light source. Since no light signal from the hidden scene parts will directly reach the sensor, the reflectance information of hidden scene elements gets blurred with the visible scene information in a non recoverable manner. If we can resolve this information, it allow us to look around corners without being in the direct line of sight (see Figure 1-1). In a room-sized environment, a microsecond long exposure (integration) time is long enough for a light impulse to fully traverse all the possible multi-paths introduced due to inter-reflections between scene elements and reach steady state. Traditional video cameras sample light very slowly compared to the time scale at which the transient properties of light come into play.

It is also possible to sample other geometric dimensions of the plenoptic function. Light field cameras [2, 3] sample both the incident angle of light and the spatial location. They have paved the way for powerful algorithms to perform many unconventional tasks such as extended depth of field, refocusing after a photo has been taken and depth estimation from fixed viewpoint. We propose that time sampling the incident light will provide even more powerful tools. The research community has extensive experience building complex models of light transport, visibility, and imaging, but none of these models take into account the transient nature of light and attempt to solve practical vision problems using it. Our work seeks to investigate the theoretical models necessary to explain imaging under these conditions, build hardware to validate these theories, and then develop the hardware into a generally useful tool.

This problem is inspired by the cumulative technological advances in ultra-fast imaging and photonics in the past decades. We exploit the fact that it is now possible to capture light at extremely short time scales. Light travels about 1 foot/nanosecond¹, and we already have commercial ultra-fast imaging hardware that is capable of sampling light at even sub-picosecond scales. Ultra-fast illumination sources such as femtosecond lasers have been available for decades. The combination

¹ 1 ns = 10⁻⁹ seconds. Precisely light travels 0.3mm in 1 picosecond = 10⁻¹² seconds.

of the ultra-fast sensing and illumination is the backbone of modern imaging systems that are based on time-of-flight, such as LIght Detection and Ranging (LIDAR) [4–8] for depth and 3-dimensional geometry sensing, two-photon microscopy and Optical Coherence Tomography ([9,10]). However, all these techniques so far have only addressed the problem of structural reconstruction (for example 3-dimensional geometry of scenes and shape of biological samples). Note that it is well known that light scatters in a scene and takes a complex multipath from illumination source \rightarrow scene elements \rightarrow sensor [11–13]. This process is termed as *light transport*. But because light travels so fast, we need ultra-fast light sampling to resolve these multipaths.

Ordinary cameras do not operate at short enough exposures times; in an ordinary real world scene, given 1 millisecond (1000 foot in light distance travel), light starting from the source would completely traverse all the possible distinct multipath in the scene and reach a steady state. Hence all the light multipath is lost during ordinary exposure time scales (integration of light incident on the sensor over time). Nevertheless, camera based 2-dimensional intensity images $I(x, y)$ have long been used to observe and interpret scenes. New sensors and algorithms for scene understanding will clearly benefit many application areas such as robotics, industrial applications, user interfaces and surveillance. We propose a new camera design that captures a 3D time-image $I(x, y, t)$. This 3-dimensional time image is created by continuously time sampling the incoming irradiance at each pixel at location (x, y) . Note that we had already mentioned that the plenoptic function as well the illumination are all functions of time and the steady state light transport is a special case in which all of these functions are constant over time.

The goal of an imaging device is to capture the scene it is exposed to, and to reconstruct the most faithful computer processable representation of the actual scene. We need theoretical models so that we may be able to make inferences and estimations from that recorded data. From the imaging perspective, there are extremely good models for image formation in steady state light transport using the ordinary 2-dimensional cameras, and algorithms for the inverse problem of scene understanding. These have been well whetted and well understood (for example, the pinhole model and homography [14]). But no such models exist for time image formation ($I(x, y, t)$) and for studying time delayed multipath light propagation. Since we are now dealing with signal waveforms $I(x, y, t)$ which are a function of time rather than just intensities, it is natural to ask signal processing ques-

tions, and formulate physically realistic models based on ideas from dynamical systems, filtering and sampling theory. Moreover, in order to develop a practical imaging implementation based on our theory we need a good characterization of noise in time image formation, and the notions of uncertainty and sensitivity. We are proposing a completely new domain where all the traditional notions related to system parameters, design and analysis need to be extended for time resolved computational imaging. We propose to study, model and exploit the benefits of temporal multi-path diversity in light transport for the development of novel signal processing theory and practical imaging devices based on our framework.

1.2 Problem and Solution

Our methods exploit the dynamics of light transport. Critical to our work is the distinction between steady-state and transient light transport. Traditional cameras capture 2D images which are projections of light field generated by steady-state light transport. This corresponds to the familiar case of global light transport equilibrium, in which the speed of light is infinity and it is conveniently assumed that light takes no time to traverse all the different multi-paths in the scene to reach the state of steady state light transport. In transient imaging, we eschew this simplification. An impulse of light evolves into a complex pattern in time. The dynamics of transient light transport can be extremely complex, even for a simple scene. Recent advances in ultra-high speed imaging have made it possible to sample light as it travels 0.3 millimeter in 1 picosecond (1 picosecond = 10^{-12} seconds). At this scale, it is possible to reason about the individual paths light takes within a scene. This allows direct observation of properties such as distance, and the difference between primary and secondary reflections.

In our formulation we considered a scene composed of small discrete planar facets with unknown 3D positions and distances, also containing hidden elements. Our camera model comprised of a generalized sensor and a pulsed illumination source. Each sensor pixel observes a unique patch in the scene. It also continuously time samples the incoming irradiance, creating a 3D *time image*, $I(x_i, y_i, t)$. The pulsed illumination source generates arbitrarily short duration and directional impulse rays. Unlike a traditional 2D camera pixel, which measures the total number of photons, a

transient imaging camera sensor measures photon arrival rate as a function of time at each pixel. The *Ray impulse response* captured using a transient imaging camera is a 5D function. By modeling the propagation and interaction of the light rays with scene elements at femtosecond time scales (1 femtosecond = 10^{-15} seconds), we estimate scene properties like depth and 3D geometry.

We propose a conceptual framework for scene understanding through the modeling and analysis of global light transport using time images. We explore new opportunities in multi-path analysis of light transport using time-of-flight sensors. Our approach to scene understanding is four-fold (also see Figure 1-3):

- 1 Measure the scene's transient photometric response function using the directional time-of-flight camera and active impulse illumination.
- 2 Estimate the structure and geometry of the scene using the observed STIR.
- 3 Use the estimated structure and geometry along with *a priori* models of surface light scattering properties to infer the scene reflectance.
- 4 Higher order inference engines can be constructed that use the estimated scene properties for higher level scene abstraction and understanding.

We corroborate our theoretical framework with experiments conducted using a prototype transient imaging camera. Our experiments demonstrate feasibility but not a full-fledged imaging apparatus. In particular, we intend this prototype (Figure 1-1) to show that it is feasible to reason about multi-bounce global transport using the 5D STIR. Our solution to the problem of single viewpoint looking around corners using transient imaging can be summarized in the pseudo-equation: 2-dimensional single viewpoint projection + 1-dimensional time sampling = 4-dimensional multiple viewpoint illumination and sensing.

1.3 Related Work

Transient imaging is a new imaging domain. To our best knowledge, neither the transient light transport theory nor the presented imaging experiments have ever been conceived in literature. The

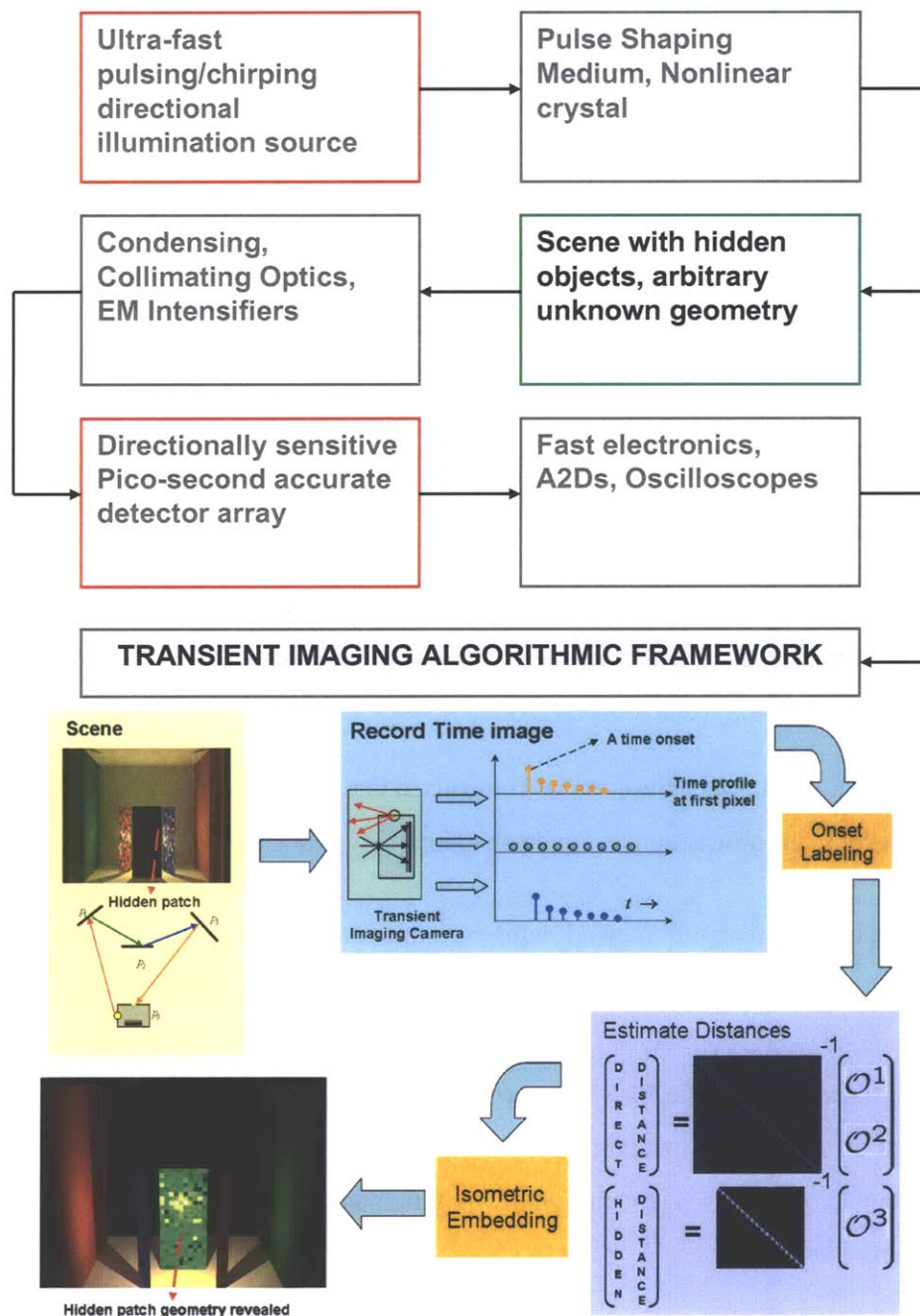


Figure 1-3: A flowchart describing our approach to scene understanding using time images: flow of information from the transient imaging camera hardware (a) into the transient imaging theoretical framework (b) which uses the time image measurement to estimate scene properties like hidden surface structure and reflectance.

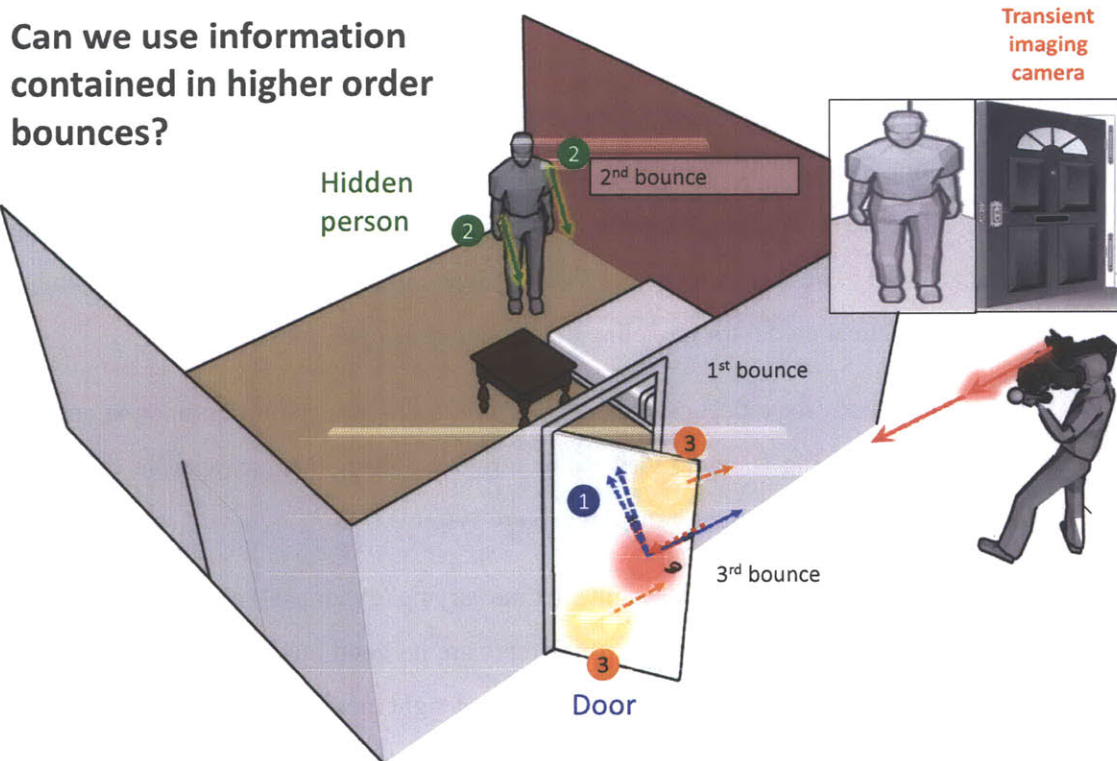


Figure 1-4: **How can you look around corners in this scene?**

This figure describes how light travels in a scene containing some hidden elements. The transient imaging camera has a flash capable of sending out extremely short and highly directional pulses of light such as shown here in red. A laser is an example of such a device. This short pulse will bounce off the wooden door say and it will spread out and scatter in all different directions. We call this first reflection or first scattering of light from the door, the first bounce. Some of the light from this 1st bounce travels directly back towards the camera and could form an image of the door. The rest of the light travels inside the room and some of it hits the hidden person and creates a second scattering which we call the second bounce. Clearly because there is no camera that has the hidden person in the line of sight, the light from the second bounce cannot be used to create a photo of the person or estimate his location in the room. However, some of the light from the second bounce travels back towards the diffuse door to create a third bounce. Now because the door is visible to the camera, the light from third bounce can be captured at the camera and perhaps if we are smart enough, it may be used to create a photo of the hidden person. Note that only a very small fraction of the light from the hidden person reaches back to the camera having undergone severe attenuation during the 3 bounces. Most of the light reaching the camera is directly from the door and hence the signal that contains the information regarding the hidden person is buried in the signal from the door. The final image captured is hence of the door and NOT of the hidden person. If we replace the wooden diffused door with a mirror, we will be able to, very clearly see what's inside the room. This is because mirror does not scatter the information which means that mirror is like an all-pass filter whereas a diffuser is scatters the light information in all directions and behaves as a very lossy low pass filter. We are going to demonstrate a way using time resolved computational imaging to turn that wooden diffused door into a mirror. We do this by exploiting the fact that light travels at a finite speed. Combining a novel theory that we have proposed for analyzing time delayed of propagation of light, with ultrafast sensing and active femtosecond illumination we can use information contained in higher order light bounces to reconstruct hidden scene geometry from a fixed viewpoint.

basic differences between transient imaging and related imaging methods are summarized as:

- Theoretical models and experimental methods to analyze global light transport do not perform temporal multipath analysis. They are all based on steady state image capture using traditional cameras. There are no models for time variant light transport primarily because it was thought it would not lead to development of practical hardware implementations. Transient imaging is a novel theoretical model for modeling the dynamics of light transport.
- Data driven methods for reflectance capture require multi-view point illumination and cameras. Transient imaging is single view point, single exposure and can achieve competitive results compared to the 8D light transport capture systems.
- Techniques such as LIDAR which exploit the time varying (dynamical) nature of the plenoptic function do not consider the multipath since there are no good models for time variant light multipath. In fact all ultra-fast techniques are line of sight methods for maximizing the SNR through multipath rejection. LIDAR is also not robust to detecting highly specular objects.
- Optical Coherence Tomography (OCT) works in millimeter sized biological samples. Transient imaging achieves high quality 3D reconstruction for real world multi-scale scenes using free space femtosecond laser illumination.

For the sake of completeness we also present a brief survey of related work comparing in detail the key differences between existing systems and our approach. Also see Figure 1-5 for a quick comparison of transient imaging with other popular imaging methods.

1.3.1 Theoretical Prior Art

Models for Global Light Transport: Light often follows a complex path between the emitter and sensor. A description of steady-state light transport in a scene is referred to as the *rendering equation* [15]. Extensions have been described to include time in light transport [16]. In [17], Raskar and Davis proposed inverse analysis using a 5D time-light transport matrix to recover geometric and photometric scene parameters. In addition, Smith et. al. [18] proposed a modification of the

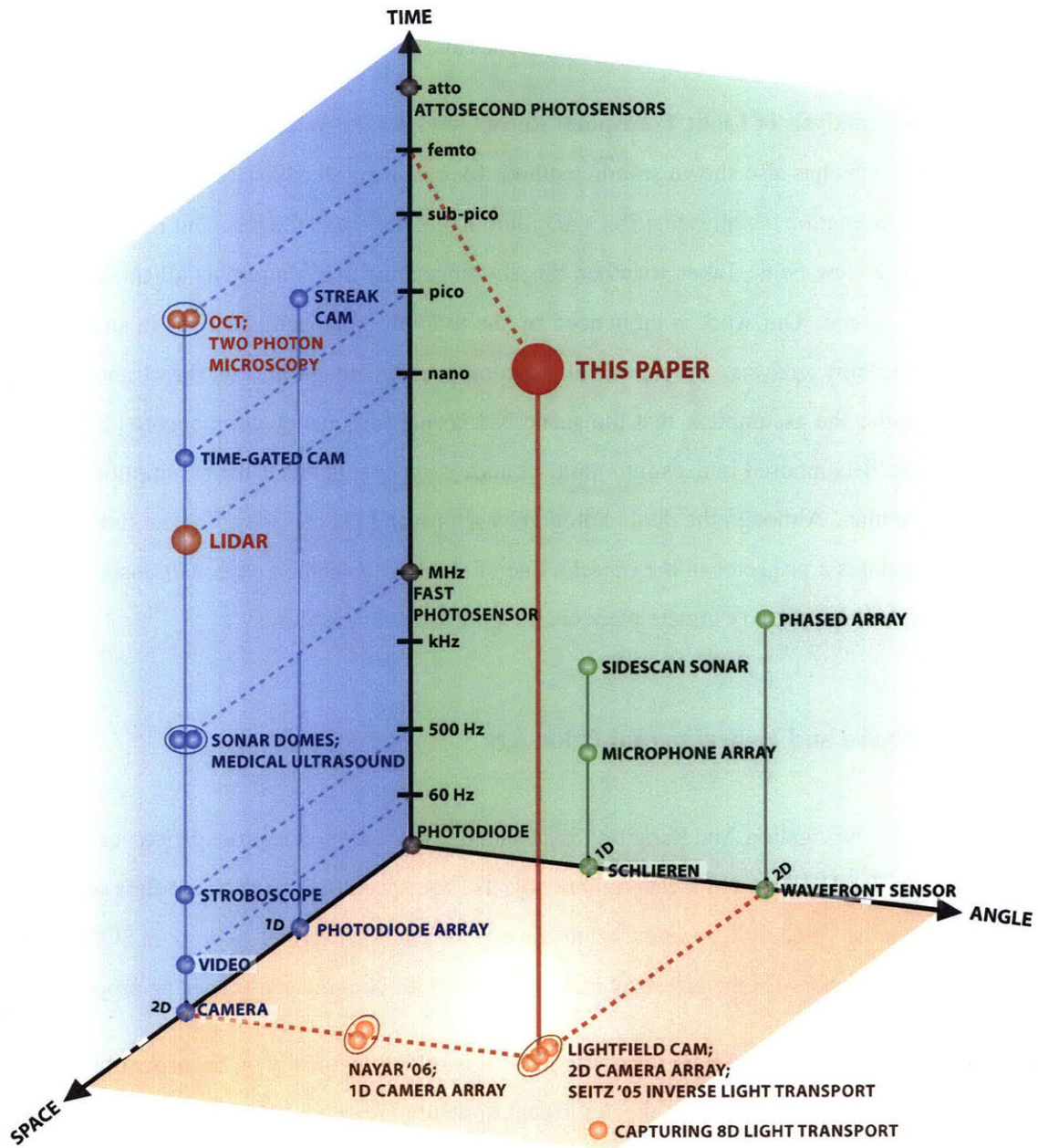


Figure 1-5: Popular imaging methods plotted in the *Space-Angle-Time* axes. With higher dimensional light capture, we expand the horizons of scene understanding. Our paper uses LIDAR-like imaging hardware, but, in contrast, we exploit the multi-path information which is rejected in both LIDAR (Light Detection And Ranging) and OCT (Optical coherence tomography).

rendering equation via a transient rendering framework. Accurate measurement of physical scene properties is called inverse-rendering [19]. Complex models have been developed for reconstructing specular [20], transparent [21], Lambertian [22] scenes and joint lighting and reflectance [23].

Capturing and Analysis of Light Transport: Recent work in image-based modeling and computational photography has also shown several methods for capturing steady-state light transport [13]. The incident illumination is represented as a 4D illumination field and the resultant radiance is represented as a 4D view field. Taken together, the 8D reflectance field represents all time-invariant interaction of a scene. Our work is influenced by the following pioneering efforts in steady-state global light transport analysis. Nayar et. al. decomposed an image into its direct and indirect components under the assumption that the scene has no high-frequency components [11]. Seitz et. al. [12] have decomposed images into multi-bounce components under the assumption that the scene is Lambertian. Although the dual photography approach [13] can see an object hidden from a camera, it requires a projector in the object’s line of sight. Our method exploits transient, rather than steady-state transport, to estimate more challenging scene properties.

1.3.2 Hardware and Experimental Prior Art

SONAR: (SOund Navigation And Ranging), is a technique that uses sound propagation in a medium such as air or water to detect and locate remote objects. The speed of sound is six orders of magnitude slower than the speed of light, and therefore easier to detect. Nevertheless work in SONAR has produced intricate models of the effects of many surfaces with complicated scattering properties.

LIDAR: (LIght Detection And Ranging) systems modulate light, typically on the order of nanoseconds, and measure the phase of the reflected signal to determine depth [4]. Flash LIDAR systems use a 2D imager to provide fast measurement of full depth maps [5, 6]. Importantly, a number of companies (Canesta, MESA, 3DV, PMD) are pushing this technology towards consumer price points. The quality of phase estimation can be improved by simulating the expected shape of the reflected signal or estimating the effect of ambient light [7]. Separately detecting multiple peaks in the sensor response can allow two surfaces, such as a forest canopy and a ground plane, to be detected, and waveform analysis can detect surface discontinuities [8].

LIDAR systems have been very successful in some circumstances, but they work well only for certain types of surface reflectance, and do little to help estimate other global properties such as relationship between scene patches. In addition they are used in restrictive configurations by carefully placing emitters near the receivers. What is needed is a generalized sensor which fundamentally records a greater portion of the light transport in a scene. This sensor could then be used to design new algorithms and specialized sensing methods.

Time-gated imaging captures a gated image, $I(x, y, t_\delta)$, by integrating the reflected pulse of light over extremely short windows. Multiple captures at incremental time windows, t_δ , allow the time image $I(x, y, t)$ to be captured at up to 100 picosecond accuracy. Nanosecond windows are used for imaging tanks at the range of kilometers and picosecond gating allows imaging in turbid water. It is possible to construct the transient photometric response function using gated imagers, e.g. Busck et. al show a TPRF measured to 100 picosecond accuracy.

Streak cameras: Streak cameras are ultrafast photonic recorders which deposit photons across a spatial dimension, rather than integrating them in a single pixel. Using a 2D array, $I(x, y_\delta, t)$ can be measured. Sweeping the fixed direction, y_δ , allows $I(x, y, t)$ to be captured. Picosecond streak cameras have been available for decades [24]. Modern research systems can function in the attosecond range [25].

Femtosecond Imaging: Optical coherence tomography (OCT) [9], an interferometric technique, and two-photon microscopy [10], using fluorescence, allow high-quality, micrometer-resolution 3D imaging of biological tissue. Both these methods are based on pulsed femtosecond illumination. Femtosecond windows also allow ballistic photons to be separated from scattered photons while imaging in biological tissue.

Advanced cameras and sensors observe and interpret scenes by relying on computational geometry and signal processing tools. Common methods use a suitable EM spectrum to see through a specific type of an occluder. We instead propose an approach that exploits light bouncing and reflecting from objects around the corner. This makes our approach independent of the type of occluder. Experiments in this proposal are the first attempt at free-space use of femto-laser illumination in contrast to their established use in optical fibers or millimeter-scale biological samples. All the

existing methods based on time sampling of light make no use of global light transport reasoning to infer scene characteristics. They instead image in a single direction time-gated window to improve SNR and reject multi-path scattering. This paper shows that complex global reasoning about scene content is possible given a measured multi-path time profile.

1.4 Contributions

We made the following specific contributions through this thesis:

- Developed a theoretical framework called *Inverse transient light transport* for estimating the geometry of scenes that may contain elements occluded from both the camera and illumination.
- We proposed a transient imaging camera model which time samples incident light continuously and uses spatio-temporal impulse illumination as the light source.
- We built a proof-of-concept hardware prototype comprising a femtosecond laser and a directionally sensitive, picosecond accurate photo sensor array, intensified CCD cameras and picosecond streak cameras to demonstrate the practical use of transient imaging theory. Using our prototype, we demonstrated all the key functionalities required in an imaging device: geometry, photometry, multi-bounce (multi-path) light separation and free-space functioning.
- We also demonstrated how our proposed imaging model enables novel scene understanding which allows us to look around the corner without any device in the line of sight. In particular, we reconstruct the structure and geometry of hidden planar lambertian scenes using our transient imaging theory and camera prototype. We have verified this through both simulation and experiments.

Chapter 2

Transient Imaging Framework

In this chapter we describe the theoretical foundations of transient imaging. Critical to our work is the distinction between steady-state and transient light transport. Steady-state light transport assumes equilibrium in global illumination. It corresponds to the familiar case in computer vision or graphics, in which the speed of light is conventionally assumed to be infinite, taking no time to cross any distance. The irradiant flux (rate of incident photons) is conveniently assumed to be constant and not a function of time. Videos may be interpreted as a sequence of images of different but static worlds because the exposure time of each frame is sufficiently long. They sample light very slowly compared to the time scale at which the transient properties of light come into play. In a room-sized environment, a microsecond exposure (integration) time is long enough for a light impulse to fully traverse all the possible multi-paths introduced due to inter-reflections between scene elements and reach steady state.

In our transient light transport framework, we assume that the speed of light is some finite value and light takes a finite amount of time to travel from one scene point to the other. As light scatters around a scene, it takes different paths, and longer paths take a longer time to traverse. Even a single pulse of light can evolve into a complicated pattern in time. The dynamics of transient light transport can be extremely complex, even for a simple scene. The theory of light transport describes the interaction of light rays with a scene. Incident illumination provides first set of light rays that travel towards other elements in the scene and the camera. The direct bounce is followed by a complex pattern of

inter-reflections whose dynamics is governed by the scene geometry and material properties of the scene elements. This process continues until an equilibrium light flow is attained.

2.1 Formulation and Terminology

We consider a scene \mathcal{S} (figure 2-1) composed of M small planar facets (patches with unit area) p_1, \dots, p_M with geometry $G = \{Z, D, N, V\}$ comprising of the patch positions $Z = [z_1, \dots, z_M]$ where each $z_i \in \mathbb{R}^3$; the distance matrix $D = [d_{ij}]$ where $d_{ij} = d_{ji}$ is the Euclidean distance between patches p_i and p_j ; the relative orientation matrix $N = [\hat{n}_1, \dots, \hat{n}_M]$ consists of unit surface normal vectors $n_i \in \mathbb{R}^3$ at patch p_i with respect to a fixed coordinate system; and the visibility matrix $V = [v_{ij}]$ where $v_{ij} = v_{ji} = 0$ or 1 depending on whether or not patch p_i is occluded from p_j . For analytical convenience, we consider the camera (observer) and illumination (source) as a single patch denoted by p_0 . All the analysis that follows can be generalized to include multiple sources and the observer at an arbitrary position in the scene.

We introduce a variation of the rendering equation [15], [26] in which we represent the time light takes to traverse distances within the scene by a finite delay. Let t denote a discrete time instant and $\{L_{ij} : i, j = 0, \dots, M\}$ be the set of radiances for rays that travel between scene patches. *Transient light transport* is governed by the following dynamical equation which we term as *transient light transport equation*:

$$L_{ij}[t] = E_{ij}[t] + \sum_{k=0}^M f_{kij} L_{ki}[t - \delta_{ki}] \quad (2.1)$$

Equation 2.1 states that the scalar ray radiance $L_{ij}[t]$ leaving patch p_i towards p_j at time t is the sum of emissive radiance $E_{ij}[t]$ and the form factor weighted delay sum of the radiances from other patches. For simplicity, let the speed of light $c = 1$. Then the propagation delay δ_{ij} is equal to the distance d_{ij} (see figure 2-2). We assume that all delays δ_{ij} are integer multiples of a unit delay. The scalar weights f_{kij} or *form factors* denote the proportion of light incident from patch p_k on to

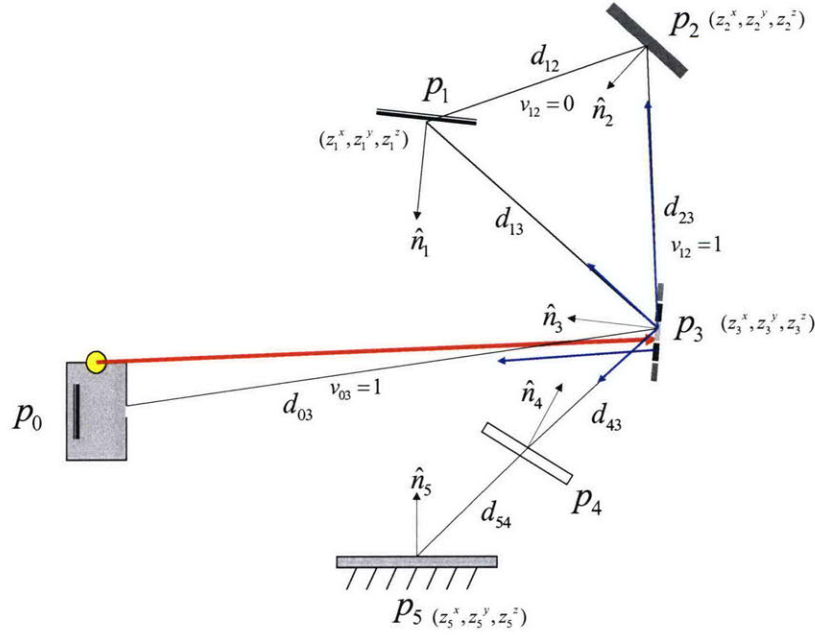


Figure 2-1: A scene consisting of $M = 5$ patches and the illumination-camera patch p_0 . The patches have different spatial coordinates (z_i^x, z_i^y, z_i^z) , orientations \hat{n}_i and relative visibility between patches v_{ij} . The patches also have different material properties, for instance p_4 is diffused, p_4 is translucent and p_5 is a mirror.

p_i that will be directed towards p_j :

$$f_{kij} = \rho_{kij} \left(\frac{\cos(\theta_{in})\cos(\theta_{out})}{\|z_i - z_j\|^2} v_{ki}v_{ij} \right)$$

where ρ_{kij} is the directional reflectance which depends on the material property and obeys Helmholtz reciprocity ($\rho_{kij} = \rho_{jik}$), θ_{in} is the incident angle and θ_{out} is the viewing angle (see figure 2-3). Additionally, if the patch does not interact with itself and then $f_{kij} = 0$ for $k = i$ or $i = j$. We assume that the scene is static and material properties are constant over the imaging interval. The source and observer patch p_0 does not participate in inter-reflections: $f_{i0j} = 0$; $i, j = 0, \dots, M$.

We model illumination using the emitter patch p_0 . All other patches in the scene are non-emissive, $E_{ij}[t] = 0$: $i = 1, \dots, M; j = 0, \dots, M; t = 0, \dots, \infty$. Illumination is the set of radiances $\{E_{0j}[t] : \forall j = 1, \dots, M; t = 0, \dots, \infty\}$ representing the light emitted towards all scene patches at different time instants. The outgoing light at patch p_i is the vector of directional radiances $\mathbf{L}[i, t] = [L_{i0}[t], \dots, L_{iM}[t]]$ and for the entire scene we have the transient light transport vector $\mathbf{L}[t] =$

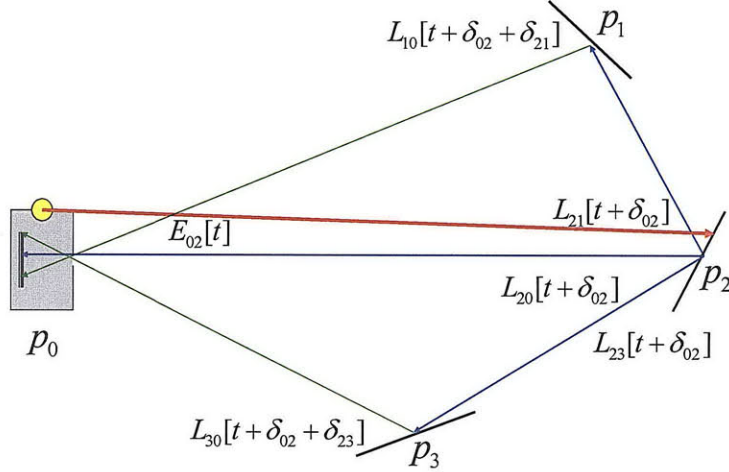


Figure 2-2: A ray impulse $E_{02}[t]$ directed towards patch p_2 at time t . This ray illuminates p_2 at time instant $t + \delta_{02}$ and generates the directional radiance vector $[L_{20}[t + \delta_{02}], L_{21}[t + \delta_{02}], L_{23}[t + \delta_{02}]$. These light rays travel towards the camera p_0 and scene patches p_1 and p_3 resulting in global illumination

$[L[1, t], \dots, L[M, t]]$ which contains $(M(M - 1) + M)$ scalar irradiances. We can only observe the projection of $L[t]$ that is directed towards the camera p_0 . At each time t we record a vector of M intensity values $L^c[t] = [L_{10}[t - \delta_{10}], \dots, L_{M0}[t - \delta_{M0}]]^T$.

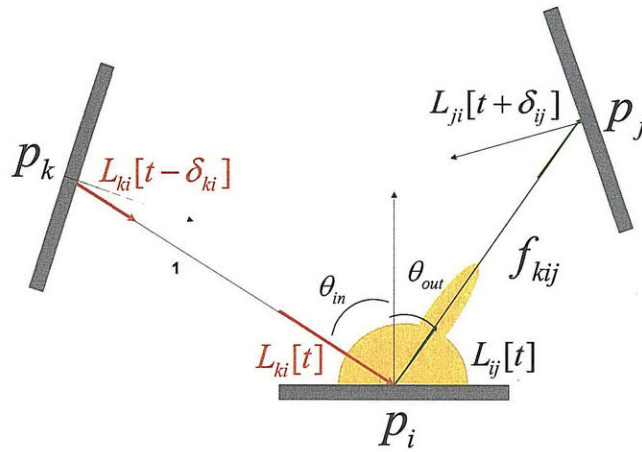


Figure 2-3: Scalar form factors f_{kij} are the proportion of light incident from patch p_k on to p_i that will be directed towards p_j

2.2 Transient Imaging Camera

The transient imaging camera model comprises a generalized sensor and a pulsed illumination source. Each sensor pixel observes a unique patch in the scene. It also continuously time samples the incoming irradiance, creating a 3D *time image*, $I(x_i, y_i, t)$ (see Figure 2-4). The pixel at sensor position (x_i, y_i) observes the patch p_i over time. The pulsed illumination source generates arbitrarily short duration and directional impulse rays. The direction of an impulse ray aimed at patch p_i is specified by (θ_i, ϕ_i) . The sensor and illumination are synchronized for precise measurement of Time Difference Of Arrival (TDOA).

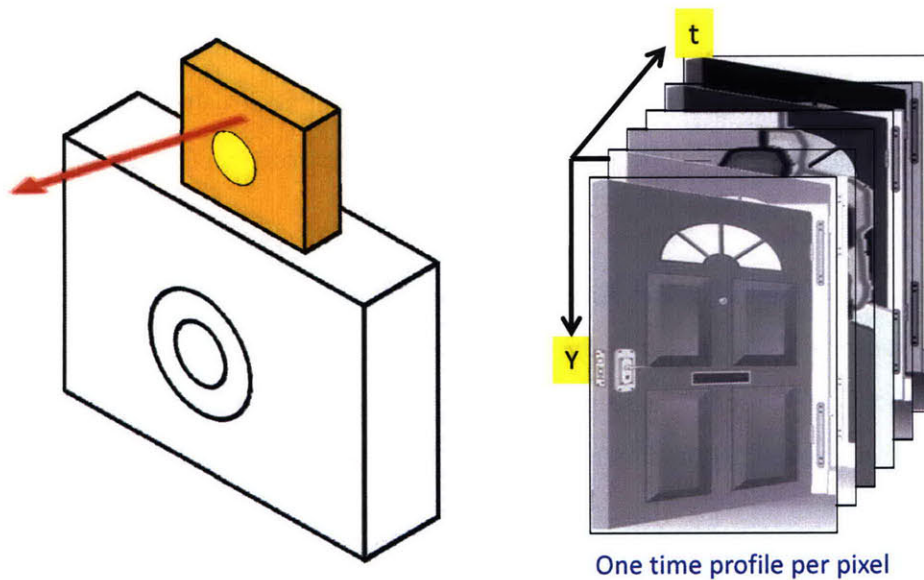


Figure 2-4: Transient Imaging Camera captures a 3D time image in response to a ray impulse illumination.

2.2.1 Space Time Impulse Response

The Space Time Impulse Response or the *STIR* of a scene S denoted by $\mathbf{STIR}(S)$ is a collection of time images, each captured with an impulse ray aimed the direction (θ_j, ϕ_j) , illuminating a single scene patch p_j . This is a 5D function: $\mathbf{STIR}(x_i, y_i, \theta_j, \phi_j, t)$. We measure the STIR as follows (see Figure 3-1):

1 For each patch $p_j : j = 1, \dots, M$.

1a Illuminate p_j with an impulse ray (θ_j, ϕ_j) .

1b Record time image $\{I(x_i, y_i, t) : i = 1 \dots M; t = 0 \dots T\} = \mathbf{STIR}(x_i, y_i, \theta_j, \phi_j, t)$.

In still scenes with static material properties, transient light transport is a Linear and Time Invariant (LTI) process. Hence, the STIR of the scene characterizes its appearance under any general illumination from a given camera view. Note that this is similar to the characterization of a LTI system using its time impulse response except that the notion of an impulse in our case extends to both space and time. We note that the transient light transport equation (Equation 2.1) describes a MIMO linear time invariant (LTI) system. One of the important properties of an LTI system is that scaled and time shifted inputs will result in a corresponding scale and time shift in the outputs. Hence the STIR of the scene can be used to completely characterize its behavior under any general illumination.

We propose a conceptual framework for scene understanding through the modeling and analysis of global light transport. The transient imaging camera maps a scene to its space time impulse response (STIR), but how can we infer the scene structure given its space time impulse response. In the Chapter 3 we propose an algorithmic framework for looking around corners, i.e., for reconstructing hidden scene structure for planar layered lambertian scenes.

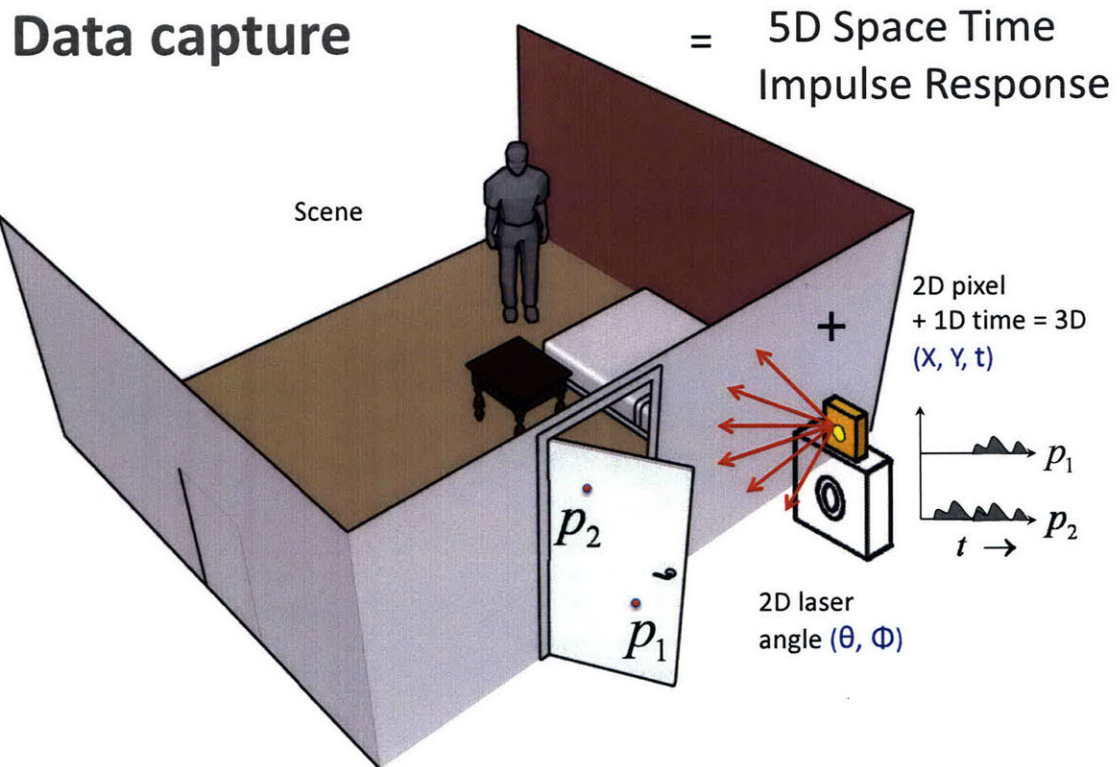


Figure 2-5: Measuring the STIR of a scene using the transient imaging camera. We successively illuminate a single patch at a time and record a 3D time image. Collection of such time images constitutes the 5D STIR.

Chapter 3

Hidden Geometry Estimation

Unlike traditional time-of-flight imaging, our goal is to compute the direct distances, d_{0i} , using the first bounce, and the pairwise distances, d_{ij} . Instead of using intrinsic camera calibration, we exploit second and higher order bounces to estimate scene geometry. First, we use the onset information contained in the STIR to estimate pairwise distances. Then we compute a robust isometric embedding to determine patch coordinates.

3.1 Inverse Geometry Problem

We develop our formulation for a scene with the following strict assumptions:

- 1 Each patch is visible from all the other patches ($v_{ij} = 1, \forall i, j$). If not, then we image locally with a set of patches for which this is true.
- 2 The reflectance of each patch p_i has a non-zero diffuse component. This assumption ensures that we are able to estimate direct distances d_{0i} .

In Section 3.1.3, we discuss the extension of our transient imaging framework to scenes consisting of patches hidden from the camera and illumination.

3.1.1 Distances from STIR

Define $\mathcal{O}^1 = \{O_i^1 | i = 1, \dots, M\}$ as the set of *first onsets*: the collection of all time instants, O_i^1 , when the pixel observing patch p_i receives the first non-zero response while the source illuminates the same patch p_i (Figure 3-1). O_i^1 is the time taken by the impulse ray originating at p_0 directed towards p_i to arrive back at p_0 after the first bounce; this corresponds to the direct path $p_0 \rightarrow p_i \rightarrow p_0$. Similarly, we define $\mathcal{O}^2 = \{O_{ij}^2 | i, j = 1, \dots, M; j \neq i\}$ as the set of *second onsets*: the collection of times when the transient imaging camera receives the first non-zero response from a patch p_i while illuminating a different patch p_j (Figure 3-1). This corresponds to the multi-path $p_0 \rightarrow p_j \rightarrow p_i \rightarrow p_0$. $O_{ij}^2 = O_{ji}^2$. It is straightforward to label the onsets in \mathcal{O}^1 and \mathcal{O}^2 because they correspond to the first non-zero responses in STIR time images.

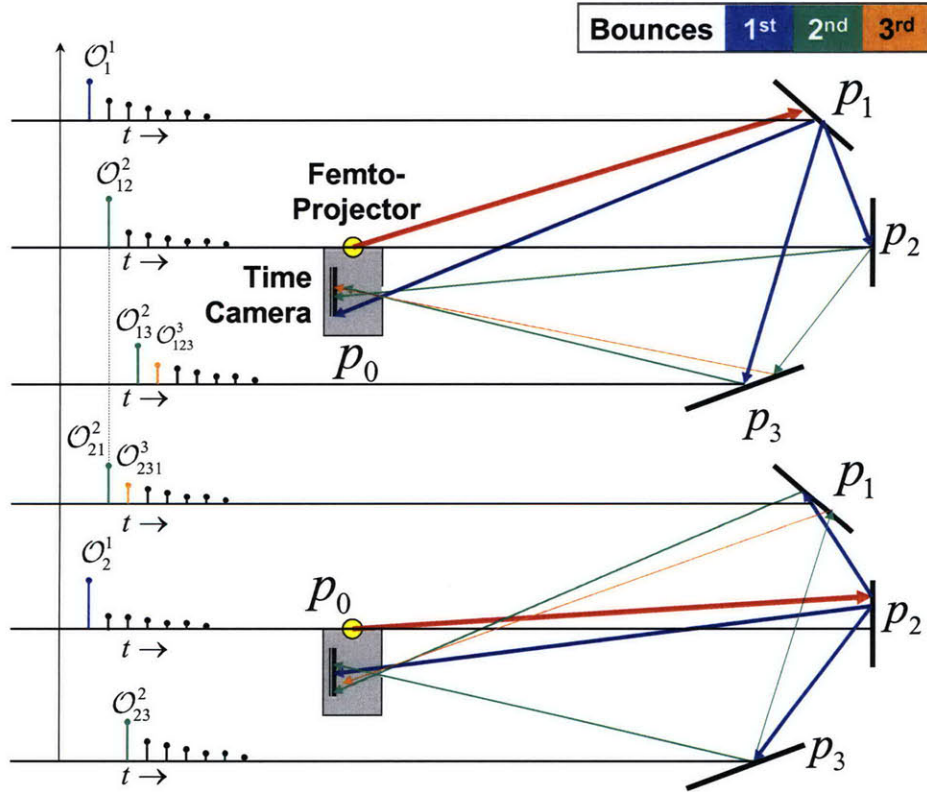


Figure 3-1: Measuring the STIR of a scene with 3 patches using the transient imaging camera. We successively illuminate a single patch and record a 3D time image. Collection of such time images creates a 5D STIR.

In order to compute \mathbf{D} using \mathcal{O}^1 and \mathcal{O}^2 , we construct the forward distance transform, \mathbf{T}_2 , of size $(M(M+1)/2 \times M(M+1)/2)$ which models the sum of appropriate combinations of path lengths contained in the distance vector $\mathbf{d} = \text{vec}(\mathbf{D})$ and relates it to the vector of observed onsets \mathbf{O} . Then we solve the linear system $\mathbf{T}_2 \mathbf{d} = \mathbf{O}$ to obtain distance estimates $\hat{\mathbf{d}}$. As an example, consider a scene with 3 patches ($M = 3$) as shown in Figure 3-1. The linear system for this scene is constructed as:

$$\begin{bmatrix} 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} d_{01} \\ d_{12} \\ d_{13} \\ \hline d_{02} \\ d_{23} \\ \hline d_{03} \end{bmatrix} = c \begin{bmatrix} O_1^1 \\ O_{12}^2 \\ O_{13}^2 \\ \hline O_1^2 \\ O_{23}^2 \\ \hline O_1^3 \end{bmatrix}$$

For any M , matrix \mathbf{T}_2 is full rank and well-conditioned. Due to synchronization errors, device delays and response times the observed onsets have measurement uncertainties which introduce errors in distance estimates. We use the redundancy in second onset values ($O_{ij}^2 = O_{ji}^2$) to obtain multiple estimates, $\hat{\mathbf{d}}$, and reduce error by averaging them.

3.1.2 Structure from Pairwise Distances

The problem of estimating scene structure, \mathbf{Z} , from pairwise distance estimates, \mathbf{D} , is equivalent to finding an *isometric embedding* $\hat{\mathbf{Z}} \subset \mathbb{R}^{M \times 3} \rightarrow \mathbb{R}^3$ (Algorithm 1, [27]). For computational convenience we take p_0 to be the origin ($z_0 = (0, 0, 0)$). A computer simulation that recovers the scene structure from noisy distance estimates using the isometric embedding algorithm is shown in Figure 3-2. We used the estimated coordinates, $\hat{\mathbf{Z}}$, iteratively to recompute robust distance estimates. The use of convex optimization to compute optimal embeddings in the presence of distance uncertainties is explained in [27].

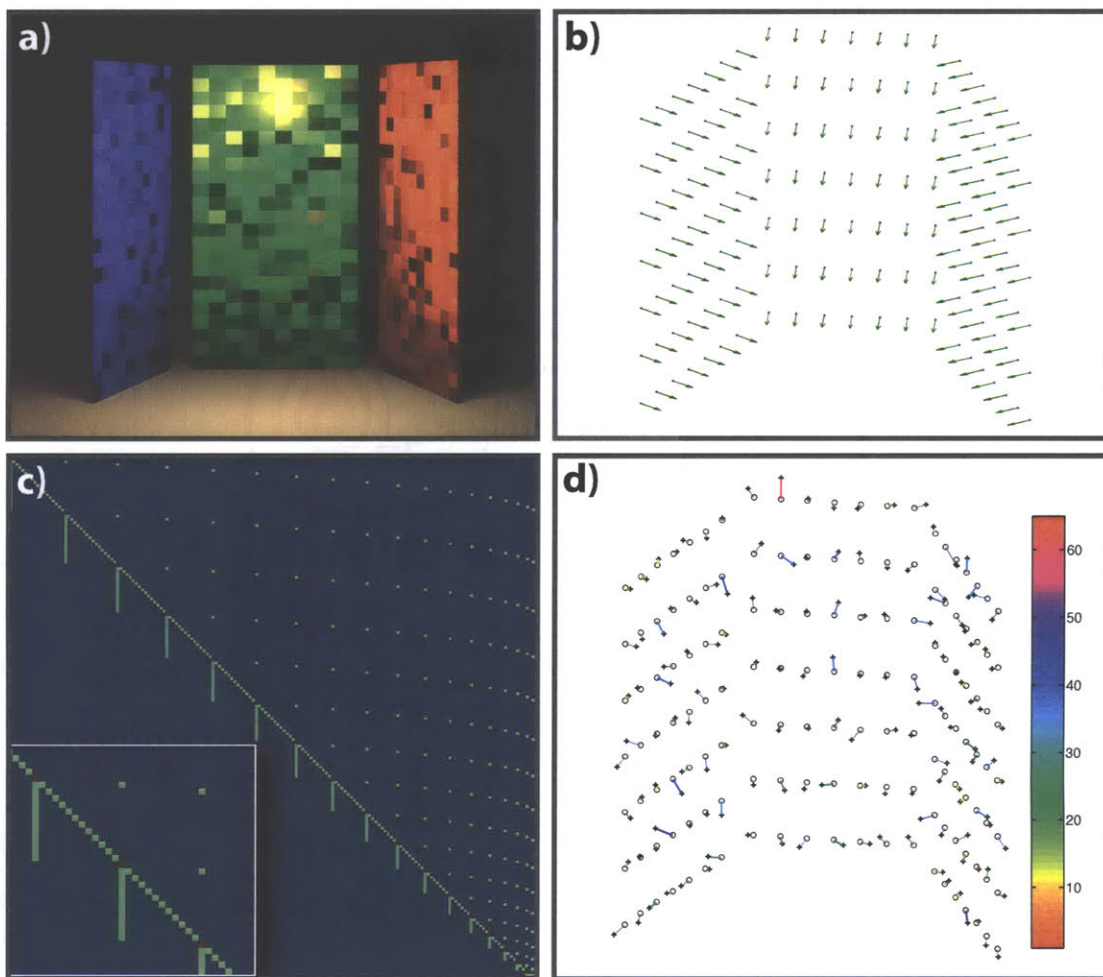


Figure 3-2: (a) Estimating distances in an all-visible scene comprising of 3 rectangles which are discretized as 49 patches. Note that reflectance is not relevant. (b) Original geometry shows the surface normals in green. (c) We used noisy 1st and 2nd time onsets (Gaussian noise $\sim \mathcal{N}(\mu, \sigma^2)$, $\mu = \text{device resolution} = 250\text{ps}$ and $\sigma = 0.1$) to estimate the distances using the \mathbf{T}_2 operator (inset shows enlarged view). (d) This is followed by isometric embedding and surface fitting. The reconstruction errors are plotted. Color bar shows %-error in reconstructed coordinate values.

Algorithm 1 ISOEMBED [$\hat{\mathbf{D}}$]

1. Compute $h_{ij} = \frac{1}{2} (d_{0i}^2 + d_{0j}^2 - d_{ij}^2)$. Construct Gram matrix $\mathbf{H}_{M \times M} = [h_{ij}]$
 2. Compute the SVD of $\mathbf{H} = U\Sigma V^T$
 3. Pick 3 largest eigenvalue-vectors $\Sigma_3^{3 \times 3}, U_3^{M \times 3}, V_3^{3 \times M}$
 4. Compute embedding $Z_e = (\Sigma_3)^{1/2} V_3$
 5. Rotate and translate to align $\hat{\mathbf{Z}} = \mathbf{R}Z_e + \mathbf{T}$
-

3.1.3 Scenes with Occluders

We now consider a scene that contains a set of patches (say H) hidden from both the camera and the source. Hidden surface estimation is viewed as two sub-problems:

- 1 Labeling third onsets.
- 2 Inferring distances to hidden patches from integrated path lengths.

To estimate the structures of the hidden patches, we make the following three strong assumptions:

- 1 The number of hidden patches is known or assumed.
- 2 All third bounces arrive before fourth and higher order bounces.
- 3 No two or more distinct third bounces arrive at the same time in the same time profile $\text{STIR}(x_i, y_i, \theta_j, \phi_j, t = 0 \dots T)$.

The second assumption is true for scenes that have no inter-reflection amongst hidden patches. The third assumption is generally valid because we measure the STIR one patch at a time. If a patch, p_i , is hidden from p_0 , then the first and second onsets involving p_i cannot be observed, i.e the vector of distances $\mathbf{d}_H = [d_{ij}] : p_i \in H, j = 0, \dots, M$ cannot be estimated using just \mathcal{O}^1 and \mathcal{O}^2 . Hence, we need to consider the set of *third onsets*, $\mathcal{O}^3 = \{O_{ijk}^3 : i, j, k = 1, \dots, M; i \neq j; j \neq k\}$, that corresponds to third bounces. Note that there are $O(M)$ first onsets, $O(M^2)$ second onsets and $O(M^3)$ third onsets. Also, Euclidean geometry imposes that $O_{ijk}^3 = O_{kji}^3$. Labeling the onsets

contained in \mathcal{O}^3 is non-trivial. An important generalization of the hidden patches scenario is to estimate distances in the case of multiple interactions between hidden patches. If a hidden patch has at most N inter-reflections with the other hidden patches, then we need to utilize onsets that correspond to up to $(N + 3)$ bounces i.e. the sets $\mathcal{O}^1, \mathcal{O}^2, \dots, \mathcal{O}^{N+3}$.

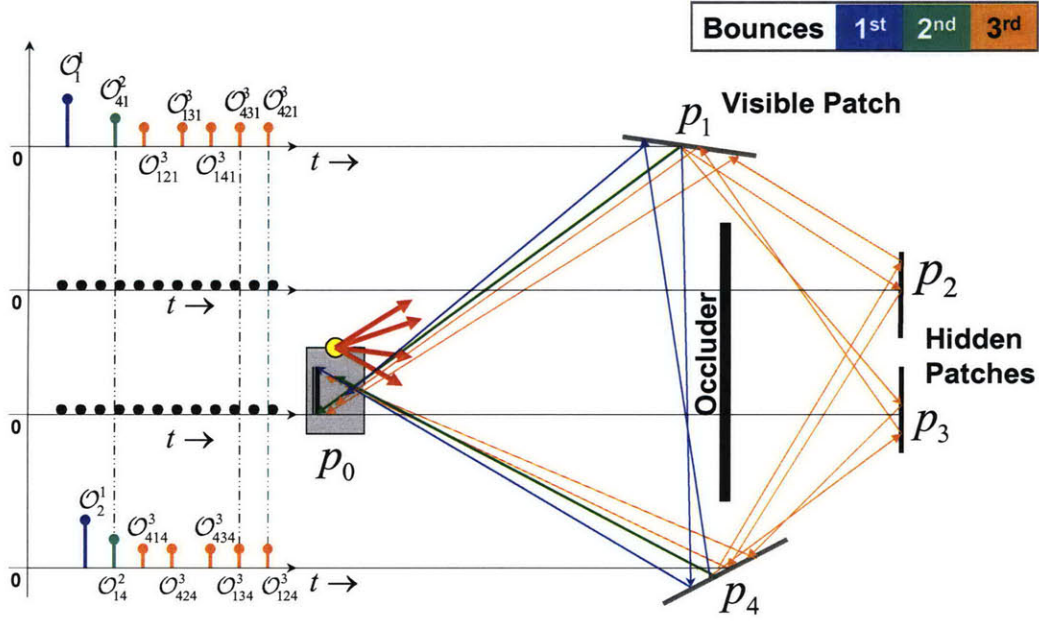


Figure 3-3: A scene with $M = 4$ patches. Patches p_2 and p_3 are hidden. The blue (first) and green (second) onsets are a result of directly observing visible patches p_1 and p_4 . The pattern of arrival of third onsets depends on the relative distance of the hidden patches p_2 and p_3 from the visible patches. The onsets that correspond to light traversing the same Euclidean distance are readily identified. Once the onsets are labeled, they are used to obtain distances that involve hidden patches.

As a simple example, consider the scene in Figure 3-3. Assume that the patches p_2 and p_3 are hidden. We first compute the distances involving visible patches, d_{01}, d_{04}, d_{14} as explained in Section 3.1.2. The distances (d_{21}, d_{24}) and (d_{31}, d_{34}) are not directly observable. Once these distances are estimated, d_{02}, d_{03} and d_{23} can be computed using multilateration. Now, we apply our labeling algorithm to identify third onsets. The onsets, O_{141}^3 and O_{414}^3 , are readily labeled using TDOA, since we know the distances to patch p_1 and p_4 . The onsets $O_{121}^3, O_{131}^3, O_{424}^3, O_{434}^3, O_{124}^3, O_{134}^3, O_{421}^3, O_{431}^3$ are disambiguated using the facts that $O_{421}^3 = O_{124}^3, O_{431}^3 = O_{134}^3$ and the onsets arrive in different time profiles of the $\text{STIR}(\mathcal{S})$. We sort the remaining onsets based on their arrival times and label

them based on the *a priori* assumption of the proximity of hidden patches to visible patches. In this example, w.l.o.g we assume that p_2 is closer to p_1 than p_3 . Hence, the onset O_{121}^3 arrives earlier than O_{131}^3 (see onset arrival profile in Figure 3-3). This labeling procedure can be generalized for multiple hidden patches:

- 1 Estimate the distances to all the visible scene patches (Section 3.1.2) and use the arrival times to label all third bounce onsets corresponding to visible geometry.
- 2 Fix an arbitrary ordering of hidden patches based on their proximity to some visible patch.
- 3 Use arrival times to identify the third onset pairs corresponding to same path length ($O_{ijk}^3 = O_{kji}^3$). Label them with the ordering of step 2.
- 4 Sort the remaining onsets according to their arrival times and use step 2 ordering to label them.

We construct the distance operator, \mathbf{T}_3 , that relates third bounces arrival times involving hidden patches, \mathbf{O}_H , and the distances to the hidden patches, \mathbf{d}_H . We solve the resulting linear system $\mathbf{T}_3\mathbf{d}_H = \mathbf{O}_H$ and obtain the complete distance set, \mathbf{D} . We then estimate the structure, \mathbf{Z} , as discussed in Section 3.1.2. An example of reconstructing hidden 3D geometry is shown in Figure 3-4.

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} d_{21} \\ d_{24} \\ d_{31} \\ d_{34} \end{bmatrix} = c \begin{bmatrix} O_{121}^3 - O_1^1 \\ O_{124}^2 - (O_1^1 + O_4^1)/2 \\ O_{131}^3 - O_3^1 \\ O_{134}^2 - (O_1^1 + O_4^1)/2 \end{bmatrix}$$

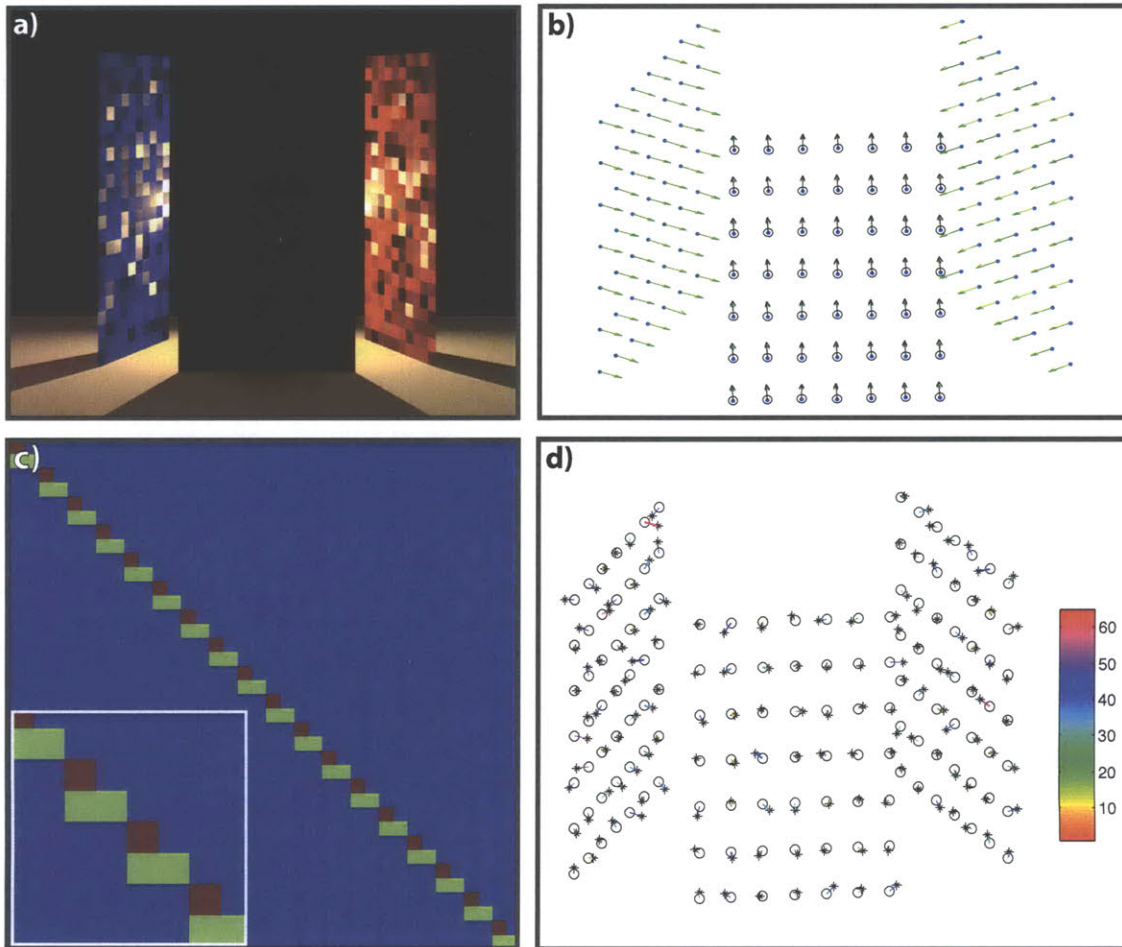


Figure 3-4: (a) Estimating distances in scenes with hidden patches. Unknown to the estimation algorithm, the hidden patches are on a plane (shown in black). (b) Original patch geometry. We use 1st, 2nd and 3rd bounce.onsets, our labeling algorithm and the T_3 operator (c) to estimate hidden geometry. (d) The isometric embedding error plot verifies negligible reconstruction error and near co-planarity of patches. Onset noise and color bar schemes are same as Figure 3-2.

Chapter 4

Hardware Prototype and Experiments

4.1 Transient Imaging Camera Prototype

The framework developed in Section 3.1 was corroborated with experiments conducted using a prototype transient imaging camera. We demonstrated the feasibility with proof of concept experiments but not a full-fledged imaging apparatus. In particular, this prototype (Figure 1-1) was developed with the intent to show that it is feasible to reason about multi-bounce global transport using the STIR.

A commercially-available reverse-biased silicon photo sensor (Thorlabs FDS02, \$72) was used as the ultrafast sensor. This sensor has an active area of 250 microns in diameter and a condensing lens to gather more light. A 5 GHz ultrafast oscilloscope digitized the photo-currents. The least count was 50 ps (1.5cm light travel). The ray impulse source was a modelocked Ti-Sapphire laser with a center wavelength of 810 nm, that emitted 50 femtosecond long pulses at a repetition rate of 93.68 MHz. The spatial bandwidth of these pulses greatly exceeds the response bandwidth of the sensor. The average laser power was 420 milliwatts, corresponding to a peak power of greater than 85 kW.

We were required to sample the incident light with picosecond resolution and be highly sensitive to a low photon arrival rate. Our depth resolution is limited by the response time of the detector and digitizer (250 ps, 7.5cm light travel). The high peak power of our laser was critical for reg-

istering SNR above the dark current of our photo sensor. Also, our STIR acquisition times are in nanoseconds, which allows us to take a large number of *exposures* and time average them to reduce Gaussian noise. In absence of a 2D photo sensor array, we emulated directionality by raster scanning the scene with a steerable laser and sensor.

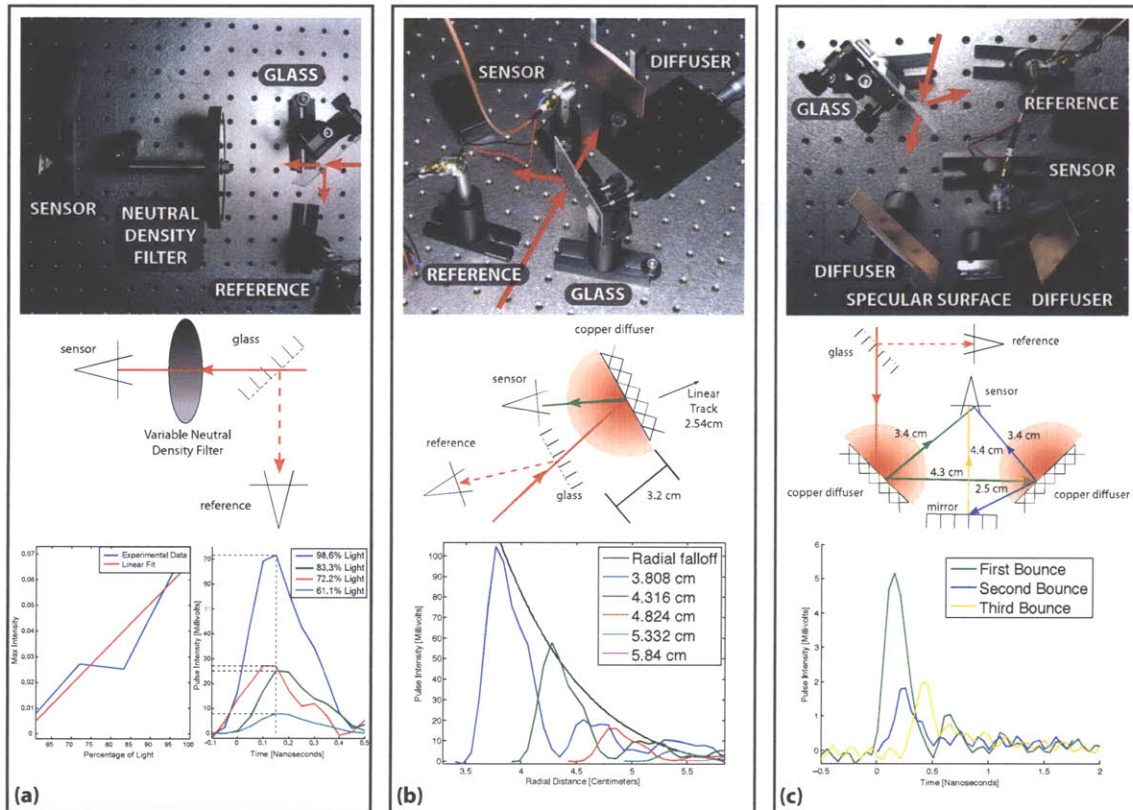


Figure 4-1: Design and Verification of a Transient Imaging Camera. (a) The ray impulses are recorded after being attenuated by a varying neutral density filter. The peak pulse intensity decreases linearly with the attenuation. (b) The intensity of the first bounce from a diffuser obeys the inverse square fall-off pattern. (c) We are able to record pulse intensities that are discernible from the noise floor even after the ray impulse has been reflected by three (2 diffuse and 1 specular) patches. The time shifts are linearly proportional to the multi-path length.

Four proof-of-concept experiments (Figure 4-1) were conducted in flatland (2D) to demonstrate the following key properties of a transient imaging camera: free space functioning, linearity, multi-path light collection, inverse square intensity falloff and time invariance. Synchronization was achieved by triggering our pulses based on a reference photo sensor. A small part of the laser pulse was

deflected into a reference photo sensor using a semi-reflective glass patch and all pulse arrivals (onsets) were measured as TDOA with respect to the reference pulse.

4.2 Hidden Geometry Reconstruction Experiments

The transient imaging prototype was used along with the algorithmic framework developed in Section 3.1 to estimate geometry for objects that do not reflect any light to camera due to specularity or occlusion.

4.2.1 Missing direct reflection

Consider the example shown in Figure 4-2(top) comprising a mirror and a diffuser. In traditional cameras it is difficult to estimate the distance to a specular surface because there is no direct reflection received at the camera. Using transient imaging analysis, estimated the distances to specular surfaces by observing indirect bounces. If we aim the laser, L , towards a mirror (in a known direction) it will strike an unknown point on M . The reflected light will then illuminate points on the diffuser. Separately, the position and the depth of the diffuser, x , was estimated via stereo triangulation (using the known angle of the laser beam) or ToF (Section 3.1.2). When the laser illuminates M , the total path length sensed at a pixel observing D is $(z + y + x)$. Since x is known, the point M is obtained using conic multilateration. Note that, in dual photography [13], we create the dual image, i.e. the projector view, but that does not allow 3D estimation. We conducted 3 raster scans and assumed $z_1 = z_2 = z_3 = z$. The path lengths $z_i + x_i + y_i, i = 1, 2, 3$ were estimated using TDOA. In this experiment, we incurred a position error of 1.1662 cm and a maximum distance error of 7.14% in coordinate reconstruction by multilateration.

4.2.2 Looking Around the Corner

We demonstrated an example of multi-path analysis in a scene that contains patches which were not visible to either the camera or the illumination source. Consider the ray diagram shown in

Figure 4-2(bottom). Only light rays that have *first* bounced off the diffuser reach the hidden patches P_1, P_2, P_3 . Light that is reflected from the hidden patches (second bounce) can only reach the camera once it is reflected off the diffuser again (third bounce). The position and depth of the points on the diffuser were estimated using first bounce onsets. We then raster scanned across the diffuser and measured the time difference of arrival (TDOA) between the first and third bounce onsets. We imaged a hidden 1 – 0 – 1 barcode using the first and third bounces off of a single diffuser. We used sensors, S_1 and S_2 , and a femtosecond laser source, L , neither of which had the barcode in their line of sight. The patches P_1 and P_3 were ground mirrors and P_2 was free space. The mirrors were aligned to maximize the SNR required for registering a third bounce. The maximum separation between P_1 and P_3 was limited to 5 cm because of SNR considerations. The first bounce, LD_1S_1 , was recorded by S_1 , and the two third bounces from the hidden patches, $LD_1P_1D_4S_2$ and $LD_1P_3D_3S_2$, arrived at S_2 within 200 ps of each other. Our current sensor was not fast enough and could only record the sum of the two third bounces. The two bounces can be recorded more accurately with a faster picosecond sensor or separated using deconvolution using S_2 's impulse response. As a proof of concept, we computed a high quality estimate by blocking P_1 and P_3 , one at a time. The reconstruction results are shown in Figure 4-2(bottom). We incurred a maximum error of 0.9574 cm in coordinate reconstruction.

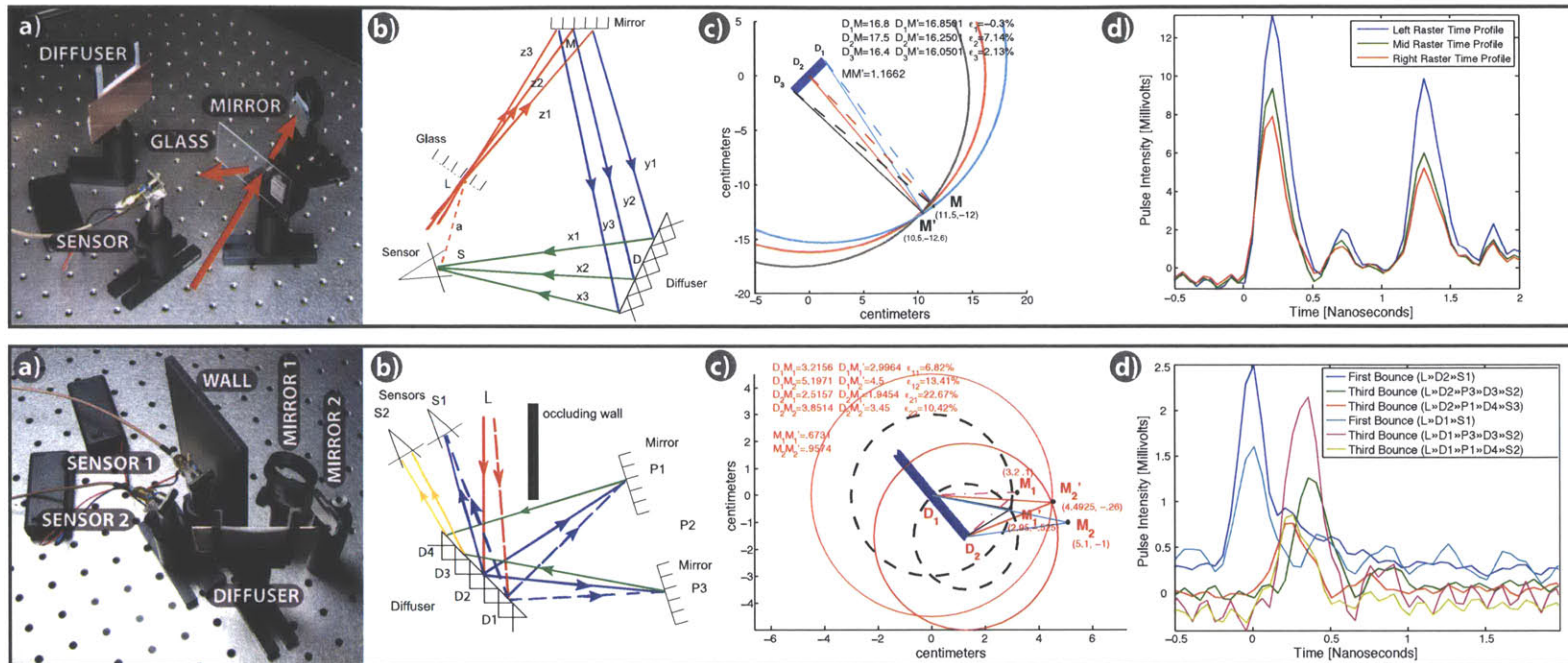


Figure 4-2: **Missing direct reflection (Top):** (a) A photo of the setup. (b) Ray diagram describing the light pulse path in 3 raster scans. (c) Plot showing multilateration using the 3 raster scans data: original and reconstructed scene geometries. (d) Oscilloscope data plot showing the TDOA between the 2nd bounce and the reference signal. **Looking around the corner (Bottom):** (a) A photo of the setup showing hidden 1-0-1 barcode. The sensors and the laser are completely shielded from the barcode. (b) Ray diagram tracing the paths of 1st and 3rd bounces in the 2 raster scans. (c) Plot showing the scene geometry reconstructed using multilateration. (d) Oscilloscope data plot showing the 1st bounce and the two separately recorded 3rd bounces, for both raster scans. Note the very small delay (≤ 200 ps) between two 3rd bounce arrivals. Please zoom in the PDF version for details.

Chapter 5

Limitations and Future Work

Transient imaging is in a preliminary stage of theoretical modeling and experimentation. Our work, as it stands, has several limitations which make it challenging to generalize the transient imaging method to more complex, general scenes. However, with sophisticated modeling and advanced instrumentation, it is possible to alleviate these restrictions.

5.1 Theoretical Limitations

Fundamentally, transient imaging is an inverse problem. As with all inverse problems, there are inherent issues related to robustness and success rate. The inverse light transport problems posed in Section 3.1 have degenerate cases in which multiple solutions (scenes) exist for the same observable STIR. Some of these issues can be alleviated with the use of suitable prior models for scene geometry. The resulting parameter estimation problems may be solved using optimization based regularization schemes and robustness constraints. But it may still not be possible to guarantee 100% accuracy while using transient methods. Importantly, although our additional time-domain data is noisy, it still restricts the class of solutions to a greater extent than using the more limited data of traditional cameras.

Novel noise models for the transient imaging camera are required to account for uncertainties due to light-matter interplay. If two or more scene patches are occluded from each other ($v_{ij} = 0$, $i, j \neq$

0), our theoretical model fails. This problem is circumvented by using our transient imaging framework locally, with a subset of scene patches that satisfy our assumptions. The number of STIR measurements grow polynomially with number of patches, but the onset labeling complexity is exponential in the number of bounce orders used for inversion. Our framework will benefit from optimization-based onset labeling to account for time arrival uncertainties. We made a set of strong *a priori* assumptions for hidden surface estimation. Statistical regularization schemes, along with scene geometry priors, will allow us to extend transient reasoning to complex scenes where hidden surfaces may involve local scattering.

5.2 Experimental Limitations

Our approach shares practical limitations with existing active illumination systems in terms of power and use scenarios: the sources must overcome ambient illumination, and only scenes within finite distance from the camera can be imaged. In addition, we require precise high-frequency pulsed opto-electronics at a speed and quality that are not currently available at consumer prices. An important challenge we face is to collect strong multi-path signals. Some of the other immediate technological challenges we face in implementing transient imaging for real scenarios include focusing light at extremely high sampling speeds, single photon sensitive detectors etc. Hardware solutions to some of our challenges may not be available immediately but recent advances trends in photonics and ultrafast sensing have indicated progress in that direction. Transient imaging may pose as an impetus to rapid advancements in device physics and photonics.

Direct reflections from scene elements are significantly stronger than light which has traveled a more complex path, possibly reflecting from diffuse surfaces. Every bounce off a diffuse surface creates considerable light loss and, thus, impacts the SNR. Thus a major challenge is to collect strong multi-path signals (requiring single photon sensitivity) with ultra-fast time sampling. Commercial solutions, such as Intensified CCD cameras, allow image acquisition at very low light levels and at relatively high gating speeds (200ps or lower). The illumination source must be powerful enough to overcome ambient light. We expect that solid state lasers will continue to increase in power and frequency, doing much to alleviate these concerns but in the meanwhile we may make heavy

use of existing LIDAR and OCT hardware in order to demonstrate preliminary applications of this technique. Also, our current method will not work for scenes which have arbitrarily placed highly specular objects, though reasoning may be improved with the use of appropriate priors.

Isolating onsets in practice is noisy, as onsets do not arrive at discrete instances; rather, they arrive as a continuous time profile. Though we have assumed a discrete patch model, future research should include continuous surface models and utilize tools in differential geometry to model the transport in general scenes. Additionally, the ray impulses are low pass filtered by the sensor response. All these reasons cause a broad temporal blur, rather than a sharp distinct onset. Defocus blur causes some scene patches to correspond to the same camera pixel. We alleviate this by working within the camera's depth-of-field.

5.3 Future Work

Emerging trends in femtosecond accurate emitters, detectors and nonlinear optics may support single-shot time-image cameras. Upcoming low-cost solid state lasers will also support ultra-short operation. The key contribution here is exploration of a new area of algorithms for solving hard problems in computer vision based on time-image analysis. New research will adapt current work in structure from motion, segmentation, recognition and tracking to a novel time-image analysis that resolves shapes in challenging conditions.

We propose to build a hardware prototype that functions in general scenes with better SNR and high spatio-temporal resolution. We need to develop a robust algorithmic inference framework to process the captured data-sets and estimate scene geometry. Through collaboration with applied optics and photonics researchers, we hope to use emerging solid state lasers and optics to build a single-shot time-image camera and port algorithms to modern processors. In the longer term, multiple problems need to be addressed to scale the solution beyond a room-sized environment. Ultimately, a hybrid detector made up of a transient imager and a long wavelength imager will support better scene understanding in estimating hidden objects.

Chapter 6

Overview, Applications and Conclusion

6.1 Summary of Approach

The goal of this thesis was to explore the opportunities in multi-path analysis of light transport. We developed the theoretical basis for analysis and demonstrated potential methods for recovering scene properties for a range of simple scenarios. Emerging trends in femtosecond accurate emitters, detectors and nonlinear optics may support single-shot time-image cameras. Upcoming low-cost solid state lasers will also support ultra-short operation. The key contribution was the exploration of a new area of algorithms for solving hard problems in computer vision based on time-image analysis.

If the STIR of the scene is available then it can be directly used to obtain the geometry G of the scene by solving the inverse geometry problem discussed in Section 3.1. This procedure is summarized as:

Algorithm 2 INVERSELIGHTTRANSPORT [STIR(\mathcal{S})]

1. Solve inverse geometry using **STIR**(\mathcal{S})
 - Estimate the distance matrix D using time onsets
 - Compute the coordinate set Z using isometric embedding
 - Compute the surface normals N using smoothness assumption
-

The goal of transient imaging is to explore the opportunities in multi-path analysis of light transport. We developed the theoretical basis for analysis and demonstrated potential methods for recovering scene properties for a range of practical, real world scenarios including guidance and control, and, rescue and planning.

6.2 Applications

Transient imaging will be useful in a range of applications where the inverse imaging problem is intractable with today's sensors such as back-scatter reconstruction in medical imaging through scattering medium (including in-contact medical imaging), tracking beyond line of sight in surveillance and robot path planning, segmentation by estimating material index of each surface in a single photo, micro machining and optical lithography.

With extended observable structure we will enable better layout understanding for fire and rescue personnel, and car collision avoidance at blind corners, (also see Figure 6-1). Other applications of non line-of-sight imaging include real time shape estimation of potential threats around a corner, computing motion parameters of hidden objects, rescue and planning in high risk environments, blind assistance, inspection of industrial objects (with hidden surfaces).

Some of the theoretical methods and results we have developed may be used to augment existing LIDAR imaging for improved 3D reconstruction and better scene understanding. The theoretical and experimental tools and techniques developed for transient imaging could be extended to other domains such as acoustics, ultrasound and underwater imaging. It may be hard to identify other such application areas immediately but as transient imaging becomes a widely used tool, researchers and engineers will find new uses for its methods.

6.3 Conclusion

Transient imaging opens up a completely new problem domain and it requires its own set of novel solutions. Although we have demonstrated initial feasibility by proof-of-concept experiments, our

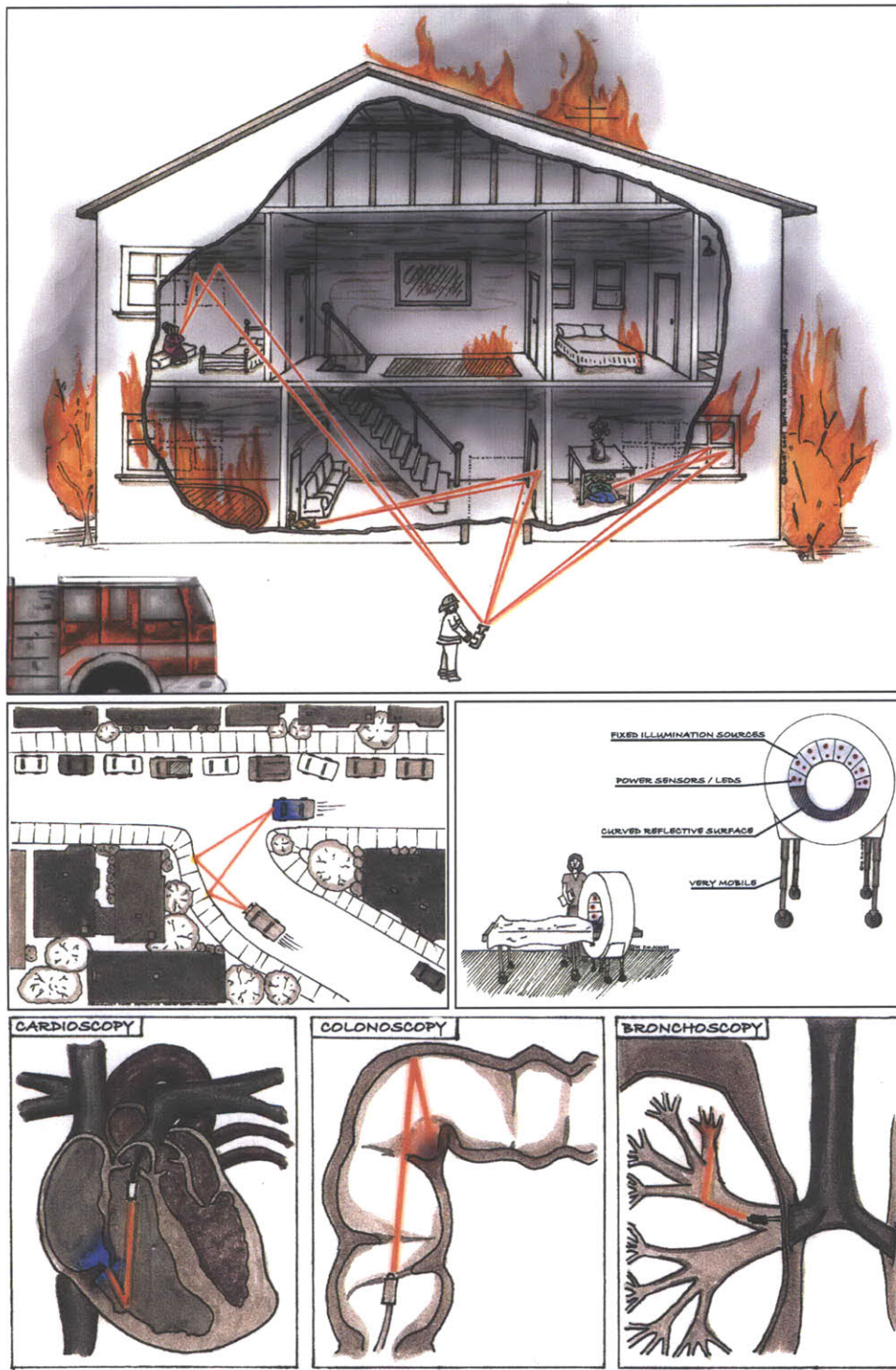


Figure 6-1: Some of the application areas which will benefit from transient imaging and looking around the corner: rescue and planning, vehicle navigation around blind corners, less-invasive biomedical imaging and a new generation of highly mobile and light weight medical scanners.

research poses more questions than it answers. There are several unexplored interesting possibilities and future directions. Before transient imaging becomes a field ready tool, we require significant theoretical and experimental research solving the real world inverse geometry and reflectance problems. In order to tap the full potential of transient imaging, we need to build dedicated hardware systems including, long range, semiconductor femtosecond lasers, sub-picosecond accurate, single photon sensitive detector arrays and powerful on board analog/digital signal processing. This process involves bringing together advanced algebraic, differential geometrical models of real world scene formation, cutting edge and experimental research in applied physics and semiconductors together with state of art on board computational signal processing.

Bibliography

- [1] Wikipedia Reference on Light fields, ", http://en.wikipedia.org/wiki/Light_field.
- [2] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocussing", in *ACM SIGGRAPH*, 2007.
- [3] R. Ng, L. Marc, B. Mathieu, D. Gene, H. Mark, and H. Pat, "Light field photography with a hand-held plenoptic camera", in *Stanford University Computer Science Tech Report CSTR*, 2005.
- [4] GW Kamerman, "Laser Radar [M]. Chapter 1 of Active Electro-Optical System, Vol. 6, The Infrared and Electro-Optical System Handbook", 1993.
- [5] GJ Iddan and G. Yahav, "3D imaging in the studio (and elsewhere...)", in *Proc. SPIE*, 2001, vol. 4298, pp. 48–55.
- [6] R. Lange and P. Seitz, "Solid-state time-of-flight range camera", *Quantum Electronics, IEEE Journal of*, vol. 37, no. 3, pp. 390–397, 2001.
- [7] H. Gonzalez-Banos and J. Davis, "Computing depth under ambient illumination using multi-shuttered light", in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, 2004, vol. 2.

- [8] N. Vandapel, O. Amidi, and JR Miller, “Toward laser pulse waveform analysis for scene interpretation”, in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, 2004, vol. 1.
- [9] J M Schmitt, “Optical coherence tomography (oct): A review”, *IEEE Quan. Elec.* '99.
- [10] Denk W, Strickler J, and Webb W, “Two-photon laser scanning fluorescence microscopy”, *Science* '90.
- [11] 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 *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, New York, NY, USA, 2006, pp. 935–944, ACM.
- [12] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos, “A theory of inverse light transport”, in *Proc. Tenth IEEE International Conference on Computer Vision ICCV 2005*, 17–21 Oct. 2005, vol. 2, pp. 1440–1447.
- [13] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch, “Dual photography”, in *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, New York, NY, USA, 2005, pp. 745–755, ACM.
- [14] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [15] James T. Kajiya, “The rendering equation”, in *SIGGRAPH '86*, New York, NY, USA, 1986, pp. 143–150, ACM.
- [16] James Arvo, “Transfer equations in global illumination”, in *Global Illumination, SIGGRAPH 93 Course Notes*, 1993.
- [17] R Raskar and J Davis, “5d time-light transport matrix: What can we reason about scene properties”, *Int. Memo* '07.
- [18] A Smith, J Skorupski, and J Davis, “Transient rendering”, *UC Santa Cruz TR UCSC-SOE-08-26*, Feb '08.

- [19] G. Patow and X. Pueyo, “A Survey of Inverse Rendering Problems”, *Computer Graphics Forum*, vol. 22, no. 4, pp. 663–687, 2003.
- [20] Kiriakos N. Kutulakos and Eron Steger, “A theory of refractive and specular 3d shape by light-path triangulation”, *International Journal of Computer Vision*, vol. 76, pp. 13–29, 2007.
- [21] Nigel J. W. Morris and Kiriakos N. Kutulakos, “Reconstructing the surface of inhomogeneous transparent scenes by scatter trace photography”, in *Proceedings of the 11th International Conference on Computer Vision*, 2007.
- [22] S. K. Nayar, K. Ikeuchi, and T. Kanade, “Shape from interreflections”, in *Third International Conference on Computer Vision*, 1990.
- [23] R. Ramamoorthi and P. Hanrahan, “A signal-processing framework for inverse rendering”, in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. ACM New York, NY, USA, 2001, pp. 117–128.
- [24] A. Campillo and S. Shapiro, “Picosecond streak camera fluorometry—A review”, *Quantum Electronics, IEEE Journal of*, vol. 19, no. 4, pp. 585–603, 1983.
- [25] J. Itatani, F. Quéré, GL Yudin, M.Y. Ivanov, F. Krausz, and PB Corkum, “Attosecond Streak Camera”, *Physical Review Letters*, vol. 88, no. 17, pp. 173903, 2002.
- [26] David S. Immel, Michael F. Cohen, and Donald P. Greenberg, “A radiosity method for non-diffuse environments”, pp. 133–142, 1986.
- [27] Jon Dattorro, *Convex Optimization & Euclidean Distance Geometry*, Lulu.com, July 2006.
- [28] A Kirmani, T Hutchison, J Davis, and R Raskar, “Looking around the corner using transient imaging”, in *ICCV '09*.
- [29] Gregory J. Ward, “Measuring and modeling anisotropic reflection”, in *SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1992, pp. 265–272, ACM.
- [30] M. Levoy, “Light fields and computational imaging”, in *IEEE Computer*, 2006.

- [31] Paul Debevec, Tim Hawkins, Chris Tchou, Haarm-Pieter Duiker, Westley Sarokin, and Mark Sagar, “Acquiring the reflectance field of a human face”, in *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 2000, pp. 145–156, ACM Press/Addison-Wesley Publishing Co.
- [32] Gaurav Garg Garg, Eino-Ville Talvala, Marc Levoy, and Hendrik P. A. Lensch, “Symmetric photography: Exploiting data-sparseness in reflectance fields”, in *Rendering Techniques 2006: Eurographics Symposium on Rendering*, Thomas Anenine-Möller and Wolfgang Heidrich, Eds., Nicosia, Cyprus, June 2006, pp. 251–262, Eurographics Association.
- [33] S.R. Marschner, *Inverse rendering for computer graphics*, PhD Dissertation Cornell University Ithaca, NY, USA, 1998.
- [34] Vincent Masselus, Pieter Peers, Philip Dutré, and Yves D. Willems, “Relighting with 4d incident light fields”, in *SIGGRAPH '03: ACM SIGGRAPH 2003 Papers*, New York, NY, USA, 2003, pp. 613–620, ACM.
- [35] Srinivasa G. Narasimhan, Mohit Gupta, Craig Donner, Ravi Ramamoorthi, Shree K. Nayar, and Henrik Wann Jensen, “Acquiring scattering properties of participating media by dilution”, in *SIGGRAPH '06: ACM SIGGRAPH 2006 Papers*, New York, NY, USA, 2006, pp. 1003–1012, ACM.
- [36] Gordon Wetzstein and Oliver Bimber, “Radiometric compensation through inverse light transport”, in *Proc. 15th Pacific Conference on Computer Graphics and Applications PG '07*, Oct. 29 2007–Nov. 2 2007, pp. 391–399.
- [37] Yizhou Yu, Paul Debevec, Jitendra Malik, and Tim Hawkins, “Inverse global illumination: recovering reflectance models of real scenes from photographs”, in *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1999, pp. 215–224, ACM Press/Addison-Wesley Publishing Co.
- [38] Canesta, ”, <http://canesta.com/>.
- [39] MESA Imaging, ”, <http://www.mesa-imaging.ch/>.
- [40] 3DV Systems, ”, <http://www.3dvsystems.com/>.

- [41] PMD Technologies, ”, <http://www.pmdtec.com/>.
- [42] Hamamatsu, ”, www.hamamatsu.com/.
- [43] G.T. Russell, J.M. Bell, P.O. Holt, and S.J. Clarke, “Sonar image interpretation and modelling”, *Autonomous Underwater Vehicle Technology, 1996. AUV '96., Proceedings of the 1996 Symposium on*, pp. 317–324, Jun 1996.
- [44] Sina Farsiu, James Christofferson, Brian Eriksson, Peyman Milanfar, Benjamin Friedlander, Ali Shakouri, and Robert Nowak, “Statistical detection and imaging of objects hidden in turbid media using ballistic photons”, *Appl. Opt.*, vol. 46, no. 23, pp. 5805–5822, 2007.
- [45] BB Das, KM Yoo, and RR Alfano, “Ultrafast time-gated imaging in thick tissues: a step toward optical mammography”, *Optics letters*, vol. 18, no. 13, pp. 1092–1094, 1993.
- [46] MA Hofton, JB Minster, and JB Blair, “Decomposition of laser altimeter waveforms”, *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 38, no. 4 Part 2, pp. 1989–1996, 2000.
- [47] J.B. Blair, D.L. Rabine, and M.A. Hofton, “The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and topography”, *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 54, no. 2-3, pp. 115–122, 1999.
- [48] M. Kawakita, K. Iizuka, T. Aida, H. Kikuchi, H. Fujikake, J. Yonai, and K. Takizawa, “Axis-Vision Camera (real-time distance-mapping camera)”, *Appl. Opt.*, vol. 39, no. 22, pp. 3931–3939, 2000.
- [49] W. Schroeder, E. Forgher, and S. Estable, “Scannerless laser range camera”, *Sensor Review*, vol. 19, no. 4, pp. 28–29, 1999.
- [50] R. Gvili, A. Kaplan, E. Ofek, and G. Yahav, “Depth keying”, *SPIE Elec. Imaging*, vol. 5006, pp. 564–574, 2003.
- [51] R. Miyagawa and T. Kanade, “CCD-based range-finding sensor”, *Electron Devices, IEEE Transactions on*, vol. 44, no. 10, pp. 1648–1652, 1997.

- [52] J. Busck and H. Heiselberg, “Gated Viewing and High-Accuracy Three-dimensional Laser Radar”, *Applied Optics*, vol. 43, no. 24, pp. 4705–4710, 2004.
- [53] P. Andersson, “Long-range three-dimensional imaging using range-gated laser radar images”, *Optical Engineering*, vol. 45, pp. 034301, 2006.
- [54] EA McLean, HR Burris Jr, and MP Strand, “Short-pulse range-gated optical imaging in turbid water”, *Applied Optics LP*, vol. 34, no. 21, 1995.
- [55] Lennart Ljung, *System Identification: Theory for the User*, Prentice Hall Information and System Sciences Series, 1987.
- [56] Per Christian Hansen, *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998.
- [57] Tohru Katayama, *Subspace Methods for System Identification*, Springer Communications and Control Engineering, 2005.
- [58] Michel Verhaegen and Vincent Verdult, *Filtering and System Identification: A Least Squares Approach*, Cambridge University Press, 2007.
- [59] Thomas Kailath, *Linear Systems*, Prentice Hall Information and System Sciences Series, 1979.
- [60] Yimin Kang, Han-Din Liu, Mike Morse, Mario J. Paniccia, Moshe Zadka, Stas Litski, Gadi Sarid, Alexandre Pauchard, Ying-Hao Kuo, Hui-Wen Chen, Wissem Sfar Zaoui, John E. Bowers, Andreas Beling, Dion C. McIntosh, Xiaoguang Zheng, and Joe C. Campbell, “Monolithic germanium/silicon avalanche photodiodes with 340 ghz gainbandwidth product”, *Nature Photonics*, vol. 3, pp. 59–63, 2008.