Hidden Object Doppler: Estimating Motion, Size and Material Properties of Moving Non-Line-of-Sight Objects in Cluttered Environments

by

Rohit Pandharkar

B.Tech. Electronics and Telecommunications, College of Engineering Pune, India
(2009)

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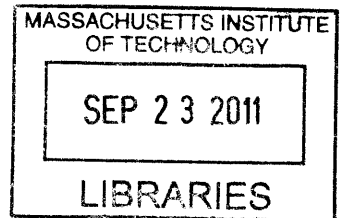in partial fulfillment of the requirements for the degree of

Master of Science in Media Technology

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2011

Author _____
Program in Media Arts and Sciences
May 6, 2011

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

Accepted by _____
Prof. Mitch Resnick
LEGO Papert Professor of Learning Research , Academic head
Program in Media Arts and Sciences

Hidden Object Doppler: Estimating Motion, Size and Material Properties of Moving Non-Line-of-Sight Objects in Cluttered Environments

by

Rohit Pandharkar

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning,
on May 6, 2011, in partial fulfillment of the
requirements for the degree of
Science Masters in Media Technology

Abstract

The thesis presents a framework for Non-Line-of-Sight Computer Vision techniques using wave fronts. Using short-pulse illumination and a high speed time-of-flight camera, we propose algorithms that use multi path light transport analysis to explore the environments beyond line of sight.

What is moving around the corner interests everyone including a driver taking a turn, a surgeon performing laparoscopy and a soldier entering enemy base. State of the art techniques that do range imaging are limited by (i) inability to handle multiple diffused bounces [LIDAR] (ii) Wavelength dependent resolution limits [RADAR] and (iii) inability to map real life objects [Diffused Optical Tomography]. This work presents a framework for (a) Imaging the changing Space-time-impulse-responses of moving objects to pulsed illumination (b) Tracking motion along with absolute positions of these hidden objects and (c) recognizing their default properties like material and size and reflectance.

We capture gated space-time impulse responses of the scene and their time differentials allow us to gauge absolute positions of moving objects with knowledge of only relative times of arrival (as absolute times are hard to synchronize at femto second intervals). Since we record responses at very short time intervals we collect multiple readings from different points of illumination and thus capturing multi-perspective responses allowing us to estimate reflectance properties. Using this, we categorize and give parametric models of the materials around corner. We hope this work inspires further exploration of NLOS computer vision techniques.
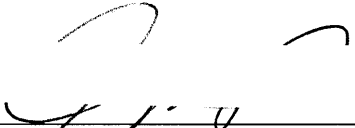
Thesis Supervisor: Prof. Ramesh Raskar
Title: Associate Professor of Media Arts and Sciences, Program in Media Arts and Sciences

Hidden Object Doppler: Estimating Motion, Size and Material Properties of Moving Non-Line-of-Sight Objects in Cluttered Environments

by

Rohit Pandharkar

The following people served as readers for this thesis:

Thesis Reader_____

Prof. Alex 'Sandy' Pentland

Professor of Media Arts and Sciences

Program in Media Arts and Sciences, Massachusetts Institute of Technology

Thesis Reader_____

Prof. Laurent Demanet

Assistant Professor of Mathematics

Dept of Mathematics, Massachusetts Institute of Technology

Preface

As I am about to complete a stepping stone in graduate life by writing this thesis, I would like to reflect on my take on experience at MIT in general and experience while working on projects. I would like to take this opportunity to make a humble attempt at trying to jot down how I felt and what changes I have seen in myself.

Graduate experience at MIT has been wonderful, I could n't have had a better place for graduate studies. Moreover, Media Lab is almost a playground for grown-ups. A 16 screen Samsung Display in the atrium, a play station and a ping pong table and most importantly transparent offices make it a happier place. I have learnt a lot here. I realized how big the world is. I realized what trillion dollar economies are. I was introduced to the silicon valley and Cambridge entrepreneurial ecosystem and also the posh McKinsey offices and hardcore Google development teams. Through conferences I got a chance to see more and more of US, from California to New York and from Seattle to Arizona. My peers from different countries made me aware of cultural lenses that people see this world through. I have seen people forming teams to help in Haiti earthquake relief and with equal enthusiasm discuss the recent Japan earthquake and the Libya issues. I have worked in teams with people thrice my age, with women and with friends who dropped out of MIT to do startups. It was wonderful to have UROPs who taught me ice skating, and friends who gave lessons in salsa. I remember discussing electric vehicles, civil air traffic and a CDMA based light field camera with friends. And of course, dozens of startup ideas that got us excited for a day or two. I had a chance to learn coursework from best of the teachers and researchers like Prof. Gilbert Strang, who have dedicated their lives helping people learn math. This place made me meet venture capitalists, billionaires and business tycoons from all over the world. People from India I probably would not have met in India itself, visited MIT and Media Lab, and I was fortunate enough to meet these luminaries briefly. A few minutes spent with them have reshaped my thinking and made me dream about bigger things. Meeting Ratan Tata, Anand Mahindra, Sunil Bharati Mittal and Desh Deshpande made me see India through the MIT and american lens. I have learnt from professors, books, internet and this place. I would like to take what I have learnt back to my country with me, and try to contribute towards running corporations and institutions that I come in touch with.

From here onwards starts one of the works I did under Prof. Ramesh Raskar, that I am most proud of. When we started, we did not know how to solve this problem, we just had a gut feeling that this could be solved with the tools we had. The journey through this work has been a true research experience, full of uncertainties and excitement. It feels nice to have something that has a solution to the problem I posed while proposing this thesis.

Acknowledgements

My first and foremost thanks to my thesis supervisor, academic advisor and friend, Prof. Ramesh Raskar, for giving me the opportunity to come to graduate school and furthering my budding research interests into a more mature methodical approach. He has not only spent time supervising my research, but has also supervised my growth as a professional and a better human being right from making better presentations to dealing with tougher situations.

I would also like to thank Prof. Alex Sandy Pentland and Prof. Laurent Demanet for guiding me and providing directives that made me explore the topic deeper and wider. I have been lucky to have a thesis committee I had dreamt of, that has my own idols from the areas of computer vision, mathematics and computational imaging. I would also like to thank Prof. Kavita Bala (Cornell university) for her significant guidance in reflectance estimation work.

I would like to thank my parents for making me what I am today. I remember my mother spending hours trying to teach me, when I threw tantrums and did not want to learn from books. My father, literally made every book I wanted available within hours of my realization that I needed the book, especially when even finding stores that had those books was difficult. My parents have invested their time, money and effort in me. I hope to make them feel proud as I graduate. Thank you mom and dad. I would also like to thank Neha Karnad for supporting me as I work on this thesis and simultaneously plan for the life after MIT.

I would also like to thank my lab mates and collaborators (in alphabetical order): Ankit Mohan, Amit Agrawal, Andy Bardagjy, Abhijit Bendale, Sharmeen Borwarek, Amna Cavalic, Kevin Chiu, Tom Cuypers, Otkrist Gupta, Kenichiro Fukushi, Matt Hirsch, Roarke Hosrtmeyer, Tyler Hutchison, Jaewon Kim, Yunhee Kim, Ahmed Kirmani, Douglas Lanman, Everett Lawson, Taya Leary, Dennis Miaw, Prof. Yasuhiro Mukaigawa, Nikhil Naik, Prof. Manuel Oliviera, Alex Olwal, Vitor Pamplona, Erick Passos, Daniel Saakes, Yolanda Spinola, Masaki Suzuki, Andreas Velten, Ashok Veeraraghavan and Jan Zizka. All of them made the Camera Culture group experience memorable. As Ramesh likes to say, all of us are here: 1. To have fun and 2. To change the world.

I would also like to thanks my professors, colleagues and peers from MIT who have contributed to the thought processes as well as un-explanable fun and excitement during my time at MIT. Whether it be climbing up the MIT dome or participating in 100k with crazy ideas or organizing the Media Lab India conference or walks by the Charles river. To mention a few, Prof. Joost Bonsen, Nitin Rao, Hemant Taneja, Prof. Ken Zolot, Prof. Gilbert Strang, Prof. George Barbasthathis, Prof. Sharmila Chatterjee, Arun Paidimarri, Mihir Sarkar, Andrea Colaco and Pranav Mistry.

Rohit Pandharkar, Massachusetts Institute of Technology

Contents

List of Figures

16

18

19

List of Tables

Chapter 1

Introduction

1.1 Research team

Before I begin I would like to introduce briefly the enthusiastic research team which has helped a lot in this thesis work. The research work on hidden object tracking and reflectance estimation using fast camera was done with the help of experimentation team that has contributed towards the data capture as well as conceptualization of the project. Andreas Velten (post-doctoral fellow in Camera Culture Group), Andy Bardagjy and Nikhil Naik (graduate students in Camera Culture Group) and Everett Lawson (graduate student at the Media Lab) have helped with experiments. My advisor Prof. Ramesh Raskar has also contributed strongly towards the ideation, discussing the possible approaches to solve the problem and pointing out the nitty gritties in the algorithms for this research.

1.2 Non Line of Sight Computer Vision

Novel imaging systems often inspire radical advances in computer vision because researchers can then explore vision science using previously unknown morphologies. It takes years for design morphologies to reach a practical level. Soon after that it leads to the development
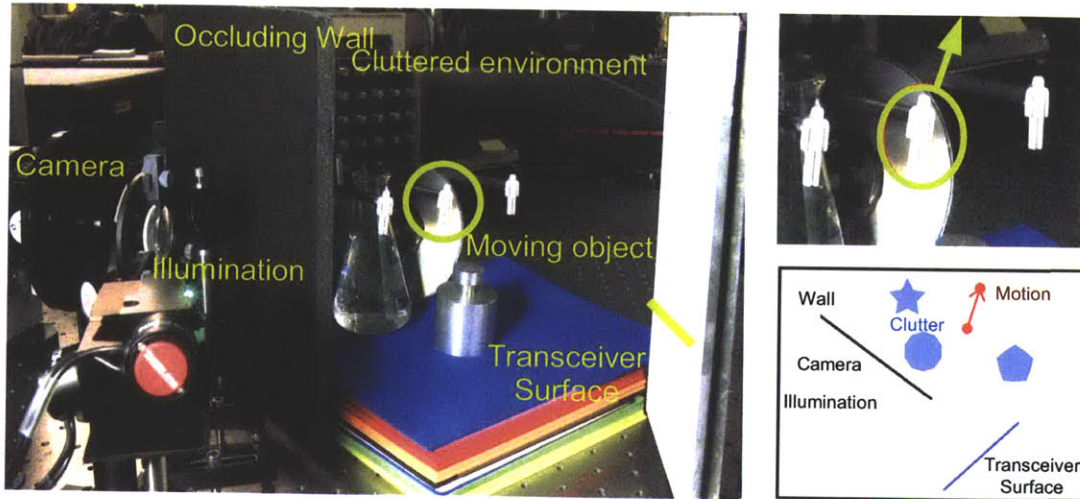
Figure 1-1: Our approach estimates motion and size of moving objects in a cluttered environment containing CDs (diffraction), water (refraction), and metallic objects (specular reflections). The short pulsed laser as well as time-of-flight camera both cannot see the moving object (moving person) beyond the occluding wall. We shine the laser on the transceiver wall that plays the role of a secondary transmitter of light towards moving object as well as primary receiver of light coming back from the object. A one-dimensional time of flight camera records the space-time response over a line on the transceiver wall, for the time window of transient response. Recording two such frames separated by known time T lets us estimate motion as well as size of the moving object in flatland.

the development of novel algorithms and frameworks for visual understanding. For example, infrared cameras allowed researchers explore the thermal nature of light and develop algorithms for bio parameter estimation [8]; similarly, the advent of radar enabled aerial object tracking and has resulted in development of track-before-detect algorithms [44]. With the introduction of time-of flight imaging systems using pico second time-resolved cameras and femto second pulsed lasers, we can explore space time light propagation in a scene. We exploit these novel imagers in this work to perform multi path light transport analysis to track non-line-of-sight (NLOS) moving objects in cluttered environments. In particular, this thesis describes a theoretical model for a NLOS tracking and and size estimation system and the hardware and software implementation of said system.

Estimating motion of NLOS moving objects has a variety of applications such as detecting oncoming vehicles while turning, sensing room occupants during search and rescue, and

monitoring moving machinery in close quarters.

In this work, we estimate motion and size of objects moving in a region that is observable by neither the illumination nor the camera, but are indirectly visible from sparse scene locations (termed as secondary transmitters and primary receivers) allowing us to illuminate them (transmit) as well as record (receive) their space-time impulse responses with a secondary receiver - our time of flight camera.

1.2.1 Thesis statement

In this thesis I plan to infer insights about objects moving out of line of sight using the information from the light that reaches the target and comes back at us through multiple reflections. I achieve this by a short pulsed laser and an ultra fast streak camera that gives time of arrival information and use of sophisticated mathematics that lets me conclude about the properties of hidden moving objects. I hope that this work would enable a platform to explore Non-Line-of-sight computer vision and spawn applications in areas like NLOS medical imaging, NLOS tracking in transportation, defense, surveillance and rescue.

1.2.2 Contributions

1 Tracking hidden moving objects in NLOS cluttered environment and gauging their size
2 Estimating reflectance properties of objects using multi-bounce reflections

Table 1.1: Major contributions of this thesis

The thesis has contribution in two major areas:

(1) NLOS object tracking and size estimation

1. We present an algorithm for estimating motion and absolute locations of NLOS moving objects in cluttered environments through tertiary reflections of pulsed illumination, using only relative time differences of arrival at an array of receivers. Our approach can accommodate global uncalibrated time bias.

25

2. We present a method for estimating size of NLOS moving objects by back projecting extremas of NLOS moving object time responses.

(2) Estimating reflectance properties of hidden objects.

1. We show the ability to infer BRDF from tertiary reflections via de multiplexing.

2. We show combining linear inversion (de multiplexing) and parametric fitting in a single step to speed up the capture process.

3. We analyze the invertibility and show that although it is impractical to recover heterogeneous BRDF using slow cameras, fast ToF cameras can recover the BRDF.

4. We show physical experiments using ordinary as well as ToF cameras, and validate the results.

1.2.3 Scope and Limitations

This work represents an advancement in the field of tracking and size estimation in non-line-of-sight situations. Despite the obvious benefits such technique offers in fields as diverse as medicine, defense and robotics, there are some limitations. The most central limitation is on scene geometry; there must be some diffuse scene features that are both visible to the target and to the time-of-flight camera. Although our system can only deal with tertiary bounces, most of real life scenarios include features that are jointly observable by the target and the camera (the tertiary bounce case).

In addition, our experiments were performed in two dimensional flatland. To extend this technique to three dimensions, one simply needs to sample more points on the jointly observable object to constrain the reconstruction.

For the reflectance capture part, the limitations are that our approach does not necessarily sample all half angles of all materials in the scene. Also, our approach has a limited resolution for small features with varying BRDF or surface normals.

Finally, to achieve centimeter scale reconstructions, the system must have pico second level time resolution and fairly low jitter. Time-of-flight cameras with larger time resolution can be used if the experimenter can accept poorer spatial resolution

1.3 Related Work

Our work is highly influenced by pioneering work in inverse light transport [46, 38, 31, 47, 24]. Our work exploits a 5D light transport to simultaneously deal with cluttered, time of flight and NLOS scenes.

1.3.1 Object Tracking:

Object tracking is a well known problem among the computer vision community. Typically, some features describing object geometry, color, size or texture, can be exploited to gain insight about target pose [51, 57]. Comparing differences between successive frames using optical flow methods is also a popular tracking method [49, 32]. Progress in appearance based tracking has been made through the application of machine learning techniques such as MCMC particle filtering [6] and manifold learning methods [28, 9]. Some researchers have exploited geometric reconstruction techniques such as multi-view stereo [17] or structured light [45] to determine target pose. The proposed method simply tracks one moving target using pose information in a static scene. In the future, our technique might be combined with, say, particle filtering to track multiple targets with overlapping paths.

Video tracking:

Tracking of objects based on video feeds is mostly performed by looking at the differences in the frames. The video tracking proces involves Representation of the target and localizing it. The standard algorithms used for video object tracking are:

Blob tracking: segmenting the interior of the objects as a blob, using optical flows or block-based correlation.

Kernel based tracking: This method is as mean shift tracking. Using iterative localization, this algorithm maximizes the similarity measure. An example of this method is Bhattacharya coefficient method.

Contour tracking: Detection of object boundary

Figure 1-2: Example of a standard motion tracking application used for surveillance at airports

Kalman filter: Kalman filter is an optimal recursive Bayesian filter for linear functions subjected to Gaussian noise.

Particle filter: This method carries out the sampling of underlying state-space distribution of non-linear and non-Gaussian processes.

1.3.2 Active emission methods for range estimation:

Many techniques have been established for range estimation using time-of-flight measurements from active emission in a wide variety of electromagnetic and pressure-wave domains including radio frequency (RADAR), optical (LIDAR), and acoustic (SONAR) [50, 19, 10]. However these techniques are limited by penetrability or lack of robust multi path frameworks. Raskar and Davis proposed the idea of using 5D light transport for non-line-of-sight imaging [42]. Recent work by our group has shown that very simple hidden patterns such as 1-0-1 bar code can be recovered in sanitized conditions [24]. In this work, we track non-line-of-sight targets in *extremely cluttered* environments containing specular, diffuse and reflective objects.

1.3.3 Phased arrays:

It is common practice in the radio-frequency and acoustic domains to form a wavefront sensor by combining measurements from multiple detector elements [34]. Our work leverages these concepts to form a phased array in the environment to recover target location.

1.3.4 Motion estimation of partially occluded objects:

Furthermore, some researchers have attacked the problem of motion estimation through heavy occlusions using techniques such as light field refocusing [22] and background subtraction [21]. In addition, active emission methods have been devised for tracking through occluders [4]. Though these methods have shown promise, none can track objects which are *completely* occluded as we have.
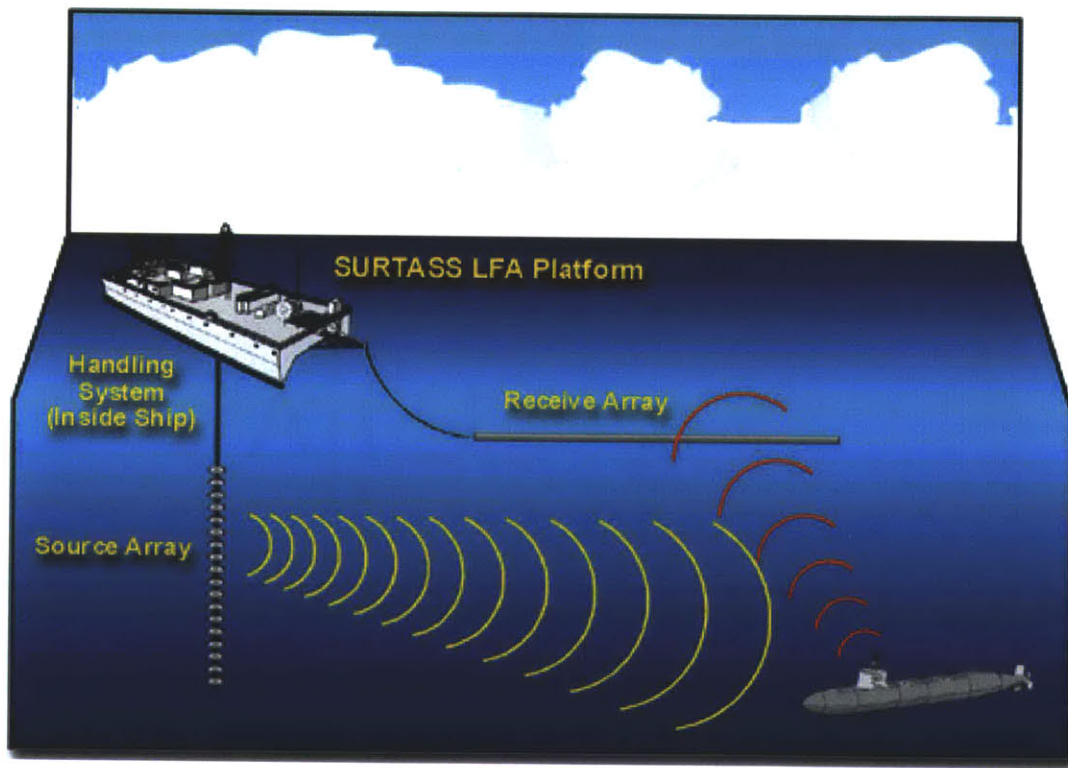
Figure 1-3: Example of a standard sonar tracking application used for navy surveillance

Figure 1-4: Example of radar tracking signals used in military surveillance
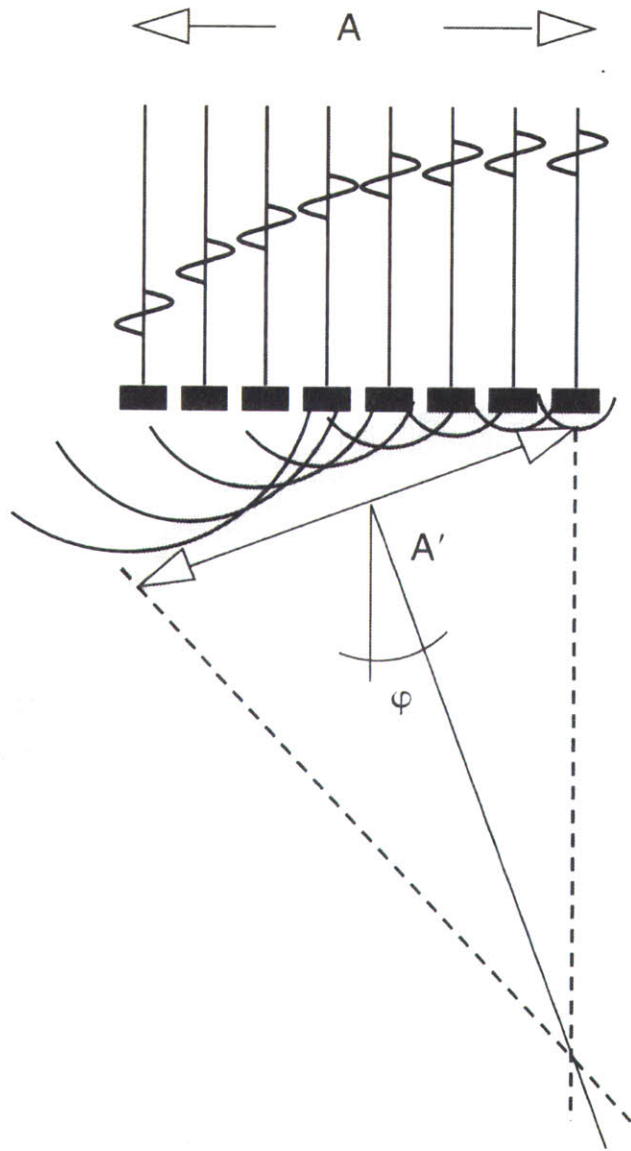
Figure 1-5: Operating principle of phased array. With an array of sensors, once can localize the unknown source based on just the differential times of arrivals compared within the sensors.
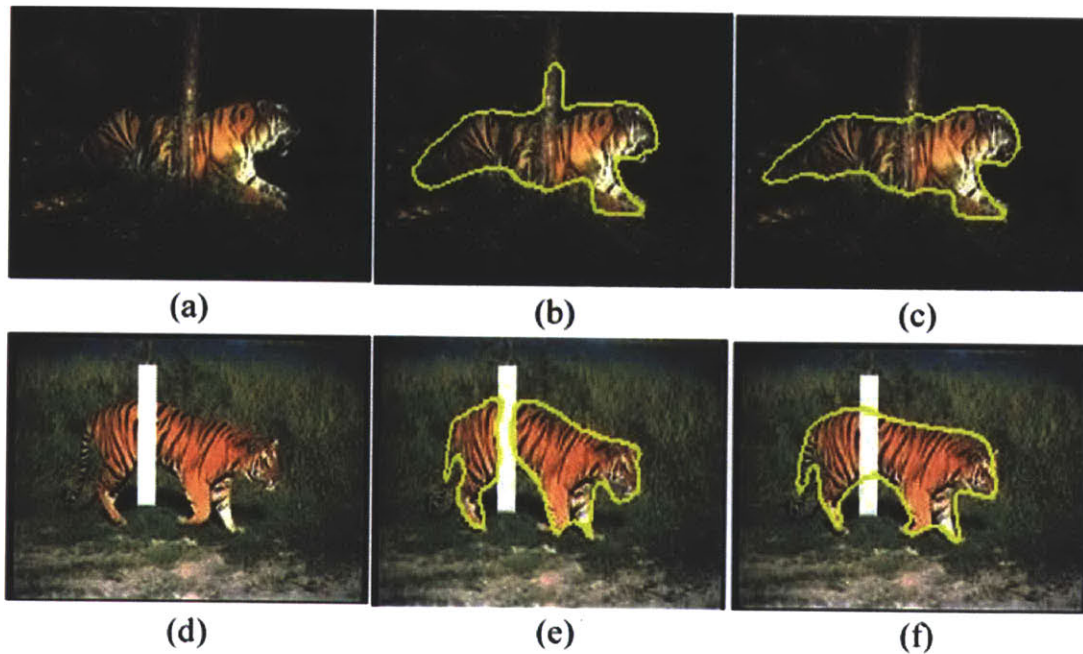
Figure 1-6: An example of estimating partially occluded object using vision principles

1.4 A typical scene setup

For a hypothetical case or real experimentation environment in this thesis work, following essential parts make a typical scene.

1. Hidden, Non-line-of-sight moving object

2. Secondary 'Trans-receiver wall' or a diffuse surface that can look at the target object

3. Laser and Streak camera, that cannot directly look at target object, but can look at a secondary diffuse surface that in turn can look at the target objects

The short pulse laser is aimed at the secondary trans-receiver wall, is reflected from this wall to the object and back, towards the streak camera, thus making it a multi path analysis problem. The 1-7 shows such a real life scenario.
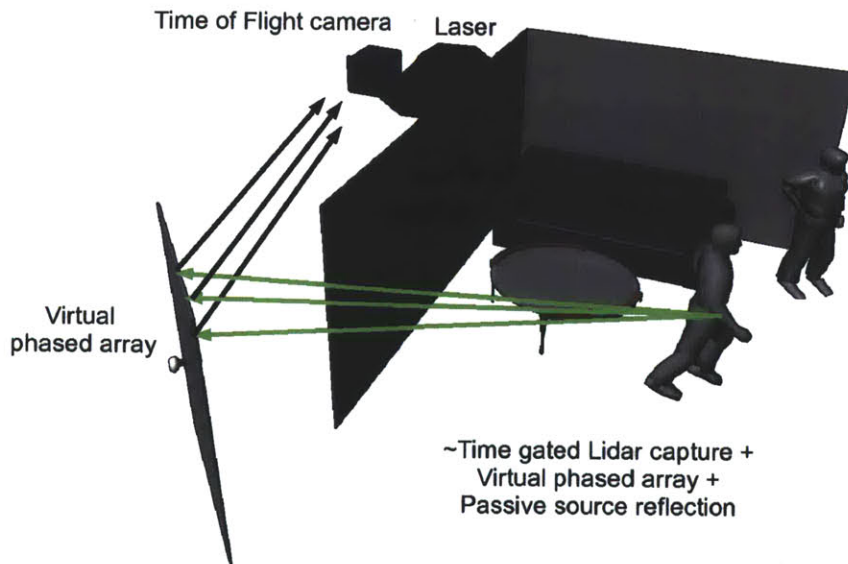
Figure 1-7: Imagine being able to track a person hiding in a room, without entering the room, without having to confront a psychopathic assassin or a baby caught up in a house on fire. A short pulsed laser flash and a trillion frames per second camera allows storing space-time impulse responses of the scene out of line of sight. Using these streak images, we infer location, size and velocity of the hidden moving objects. Moreover, we can also categorize material of the object based on its reflectance properties.
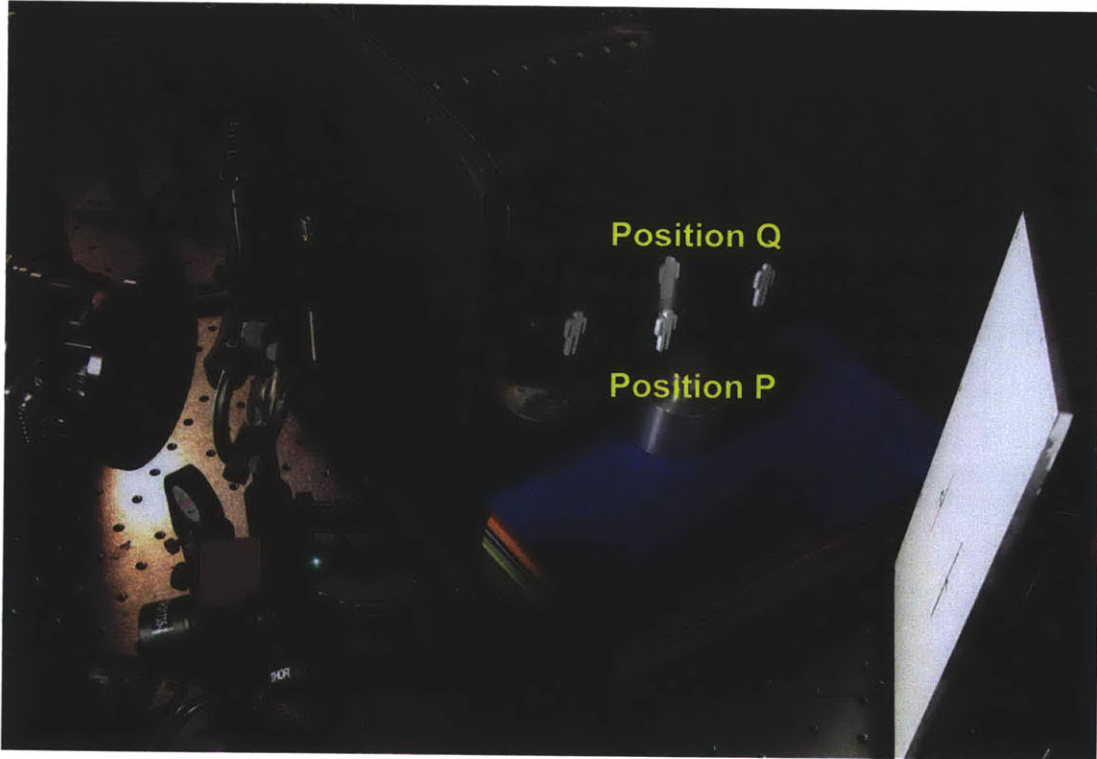
Figure 1-8: A typical experimental scene: Real scene depicting motion of a human figure in the center moving from position P to Q as a linear motion in the flatland comprising of object O, R_i line and points L and C_i. The cluttered environment consists of CD (diffractive reflection), metallic cylinders (specular reflection) and transparent water container (refraction).

Chapter 2

Algorithms and Methodology

2.1 Time of flight signal capture

Using time of arrival information at different wavelengths is a concept relatively well-known. Researchers have been using ToF signal capture for Gated Lidars, ToF tomography machines and similar ranging techniques. For the sake of completeness, we still elaborate a few basic principles of ToF signal capture.

2.1.1 Concept of slicing out scenes

Time of flight capture involves sending a short pulse or burst of signal and capturing the response at finer time resolution. In a non ToF system, the time response is typically integrated over the whole capture interval, however for a ToF system the time interval levels are pretty low, to the tune of a few micro or nano seconds.

It is very intuitive to imagine that the time of flight system thus results in capturing the response of a slice of the scene that arrives during a particular time interval. The given illustration shows the concept.
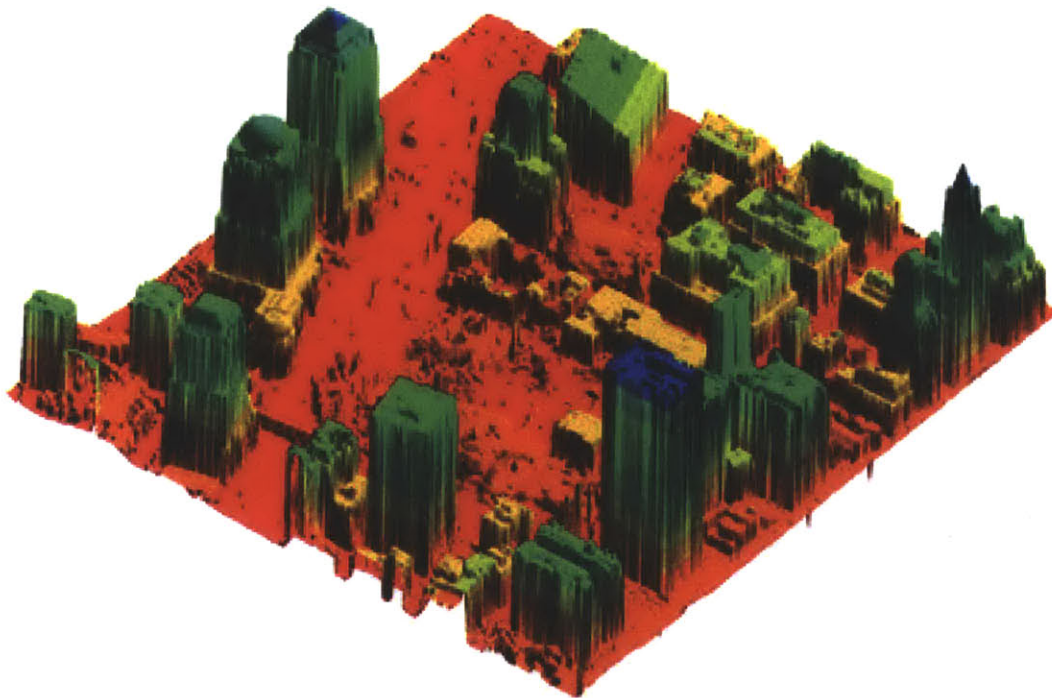
Figure 2-1: With the time of arrival information, one can capture the depth/range information in the form of sliced out scenes. Source: http://en.wikipedia.org/wiki/ TOF camera
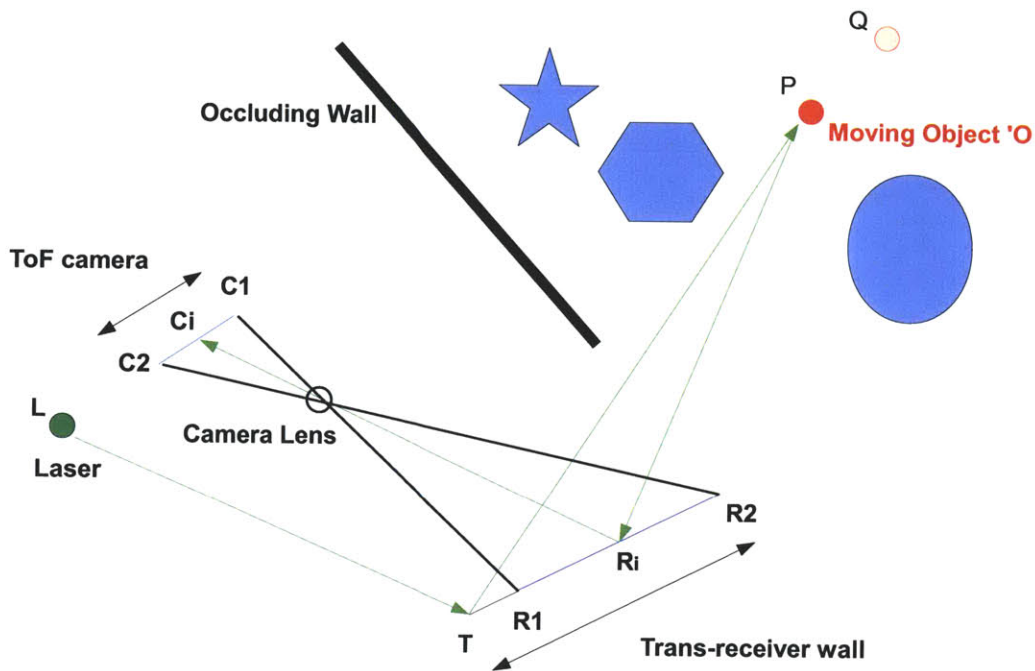
Figure 2-2: Schematics and terminology of our framework that includes time-of-fight camera, illumination, and a NLOS moving object in a cluttered environment.

2.2 Motion estimation in cluttered NLOS environments

2.2.1 Understanding Time Differentials

In this section we explain the theoretical forward model for multi path light transport from illumination to capture of NLOS objects under considered situations. Figure 2-2 shows positions of laser illumination source L and 1D time-of-flight camera C. The occluder wall prohibits the laser as well as camera to look at the moving object O beyond the wall directly. However, it is possible to look at object O through an intermediate bounce at the transceiver wall. This wall is named so because it plays the role of secondary transmitter of light coming from laser to illuminate the object as well as the role of primary receiver of light coming back from the object letting camera work as a secondary receiver. Our

39

approach allows us to time-gate the space-time capture of the moving objects, masking responses from the background clutter.

Since we use a time-of-flight camera with very short integration time, the speed of object is extremely slow compared to image capture time. Thus, we can assume the real life moving object to be stationary during this capture time.

We use only intra frame differential times of arrival relative to the arrival time of first photon for estimating the object locations. Hence, we will now focus on multi path analysis for a stationary object scenario to see how differential times of arrival can be exploited for locating the object.

Consider a single frame capture by time-of-flight camera. The space-time response of object O is made up of photon arrival times for an array of secondary receivers along line C_1 to C_2. Each of these arrival times for a particular receiver C_i is determined by the path a ray took from laser to C_i.

Let us consider the two different paths from laser to C_0 and C_i through a given secondary transmitter Tx that also go through primary receivers R_0 and R_i respectively. Both R_0 and R_i lie on the primary receiver array on the trans-receiver wall.

Let us also assume that $Path_0$ is the shortest distance path among all possible discretized paths, thus it corresponds to time of arrival of first photon from object at streak camera. $Path_i$ represents any other possible non-shortest distance path. Multi path representations of both these paths can be described as follows:

$$Path_0 = LT + TO + OR_0 + R_0C_0 \quad Path_i = LT + TO + OR_i + R_iC_i \qquad (2.1)$$

If we were to represent relative times of arrival for any non-shortest distance path with respect to shortest time of arrival, it can be written as

$$\Delta Path_i = Path_i - Path_0 \quad \Delta Path_i = [OR_i + R_iC_i] - [OR_0 + R_0Q_0] \qquad (2.2)$$

This shows that the differential times of arrival are dependent on two path disparities:

1. Difference between paths from object O to primary receivers R_0 and R_i

2. Difference between paths from R_0 and R_i to C_0 and C_i respectively.

Out of these, latter part of differences due to RC pairs are within line of sight, so they can be calculated earlier using the time of flight camera. Hence we subtract these time differences, resulting in differential time of arrival:

$$\delta Path_i = OR_i - OR_0 \tag{2.3}$$

From the third bounce capture time-of-flight camera output, we calculate $\Delta Path_i$ and process them to obtain $\delta Path_i$. In Section 2.3.1, we explain how we can get absolute locations of object O at the image capture time using only differential time of arrival: $\delta Path_i$.

2.3   Motion Estimation Algorithm

2.3.1   Theoretical model for point objects

For the sake of simplicity, we first present the model for point object estimation. This approach allows us to look at the problem in a simpler manner without getting into the complexities of signal overlap and superposition. In this subsection, we will treat objects in the scene that can be reduced to a point allowing us to associate the signals coming from object as coming from a point source.

Thus this now becomes a source localization problem, albeit to be solved in a multi path environment.

Lets assume that the object moved along a line in x-y plane from position P to Q between this period. We have two space-time images $[I(x,t)]$, $Frame_i$ and $Frame_{i+1}$, one when

41

object O was at location P and the other one when the object was at location Q. Our approach is to locate the object at each of these instances (position P and position Q), letting us also estimate the motion vector PQ.

Algorithm:

1. Subtract $Frame_{i+1}$ from $Frame_i$.

$$S = Frame_{i+1} - Frame_i \qquad (2.4)$$

This gives us the space-time impulse responses of moving object O exclusively, as the static cluttered environment responses get canceled out. With this subtracted image, we have the response of the object in position P with positive values and response of the object in Q with negative values, provided that the two responses do not overlap. We segment out individual responses by treating each signed value response set as an individual response from each location. This process is shown in Figure 2-3. We only consider pixel values which are smaller than ϵ in the other frame.

$$response_P = S.*(S > 0).*(Frame_{i+1} \leq \epsilon) \quad response_Q = -S.*(S < 0).*(Frame_i \leq \epsilon)$$
$$(2.5)$$

Note: We would like to point out here that the trivial approach of positive and negative extracts after subtractions to get separated streaks works only when the displacements are larger than the actual size of the object. This approach fails to give results directly when the streaks being subtracted overlap. However this problem can be solved by observing multiple frames and subtracting non overlapping frames and then interpolating the track for intermediate positions.

2. Having extracted individual space-time response profiles corresponding to first and second location of moving object, we now find each of these locations by making use of the extracted responses.

For a given space-time response:
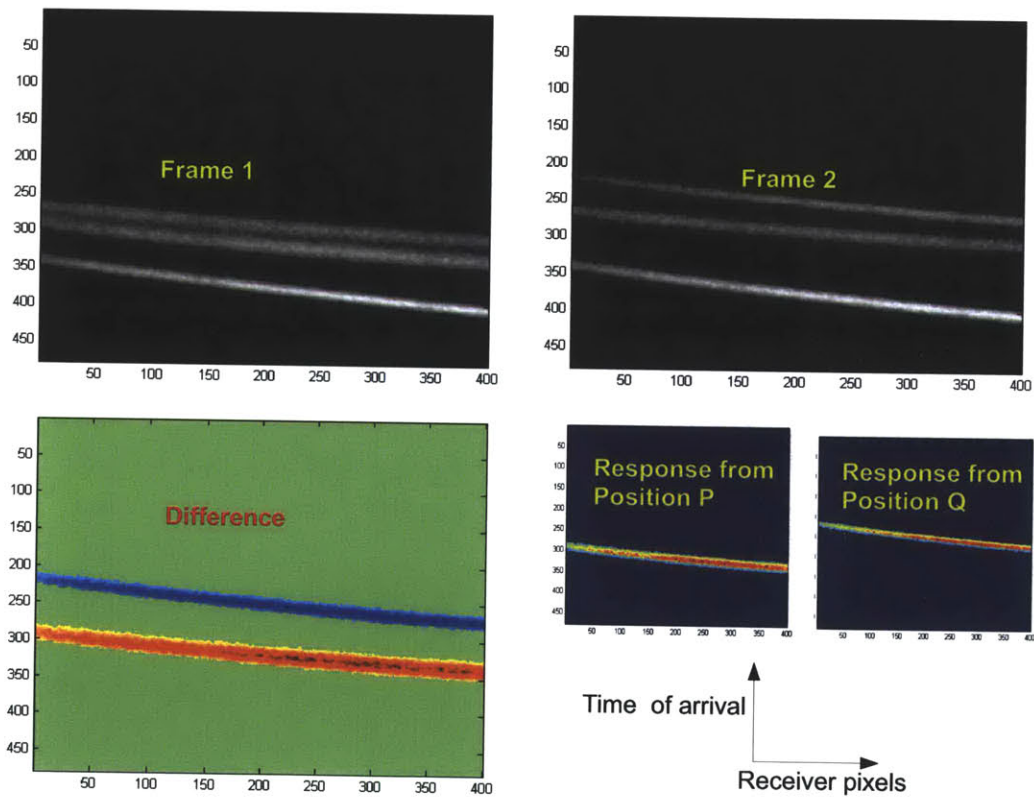
Figure 2-3: Response from location P and response from location Q are extracted from the two frames composed of superposition of responses from clutter and the target by taking a difference of two frames.

(a) Since the time values have an unknown bias due to unsynchronized laser and camera operation, we set our time origin at time-of-arrival of first photon from moving object in the response profile.

Hence, now every other time of arrival is with respect to time-of-arrival of first photon returning from moving object.

Moreover, from this changed time origin response profile, we pick only the first onsets arriving at each receiver, leading to a thin lined response curve. This is equivalent to treating the object as single point object for motion estimation purposes. We will revisit and change this when we consider size estimation.

(b) We back project the wave fronts using these differential time-of-arrival by drawing locus circles around each primary receiver R_i as shown in Figure 2-4. Each of these i^{th} circle essentially gives locus of the possible points where the photon that later arrived at R_i could have been when first photon reached R_0.

(c) Clearly, as all photons on these newly back projected locus circles originated from object O, they have to be equidistant from a particular point, and that equal distance in theory equals OR_0 (which we do not know in practice).

3. Thus, taking the front side convex hull of all these back projected locus circles and solving a constrained least squares problem for a point that gives least distance variation from all these points results in solution of the absolute location of object.

$$min \frac{1}{2}||Mx - D||_2^2, \quad \text{such that} \quad lb < x < ub \tag{2.6}$$

where lb and ub define lower and upper bounds on locations within the occluded region and

$$x = [location_x; location_y] \quad \text{and} \quad M(i,:) = [m_i, 1] \tag{2.7}$$

and m_i is the slope of normal to the convex hull curve at i^{th} point on the curve.

$$D(i) = y_i + m_i x_i \tag{2.8}$$

44
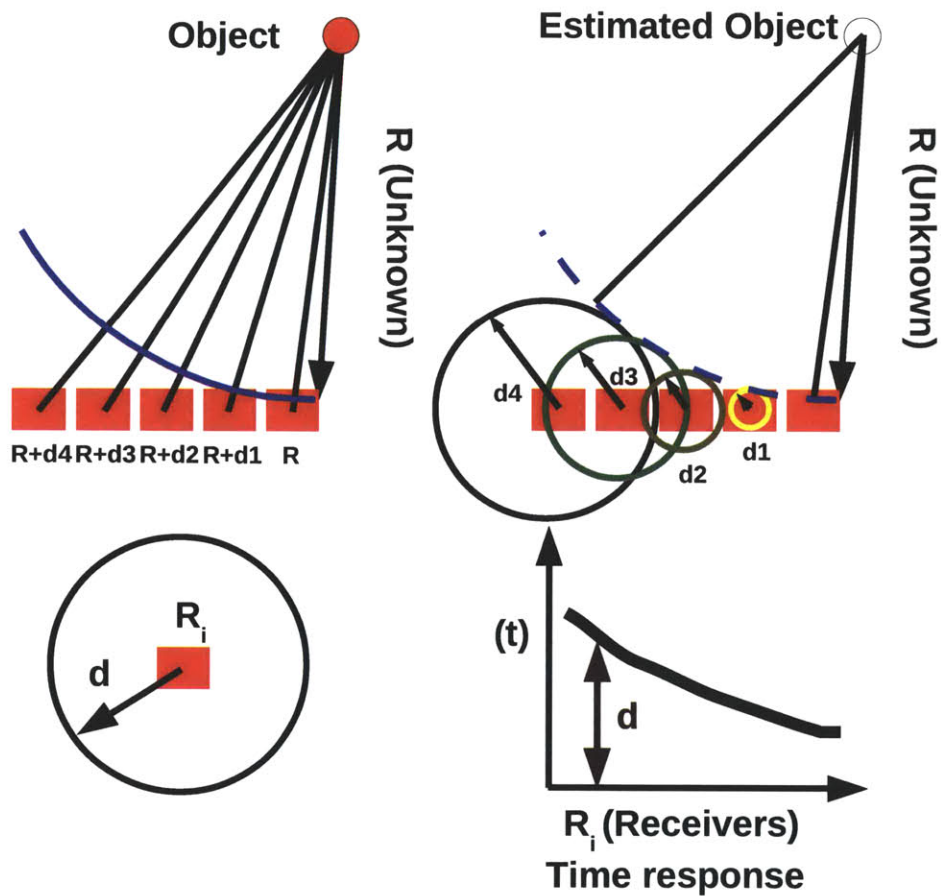
Figure 2-4: Our algorithm makes use of differential times-of-arrival by treating the transceiver wall as a phased array. (top) Forward model for phased array based capture. (middle) Locus of differential time-of-arrival wavefront. (bottom) Estimated object localization using constrained least squares.

where y_i and x_i are coordinates of ith point on the curve.

$$[location_x, location_y] = \text{estimated target location} \qquad (2.9)$$

4. Having done this for $Frame_i$ and $Frame_{i+1}$, we have estimated locations P and Q of the moving object. Hence, we can also compute motion vector PQ. We can also estimate the velocity from the relation

$$v = \frac{|PQ|}{T} * \frac{PQ}{|PQ|} \qquad (2.10)$$

where $\frac{PQ}{|PQ|}$ is the unit vector along motion direction.

Note that the constrained least square model in spatial domain explained here is the simplest tool that serves the purpose. However, more sophisticated techniques can be also applied from the area of source localization in communication theory and acoustics namely, Generalized Cross-Correlation PHAse Transform (GCPHAT)[25] and Adaptive Eigenvalue Decomposition (AED)[3].

2.3.2 Discussions on point object model

Note that the model explained above is a point object model. Almost every real life object is a non-point object. Thus this approach will give the tentative centers of locations of the objects but will not be able to track object as a point cloud. Why does this happen? This is the result of the superposition or mixing of the streak images or the space time impulse responses of the points on the object. However this point object model is till valid and useful to take us towards a continuous object model made up of a point cloud. We will now briefly see how the approach can be re-engineered to solve for a point cloud.

2.3.3 Theoretical model for real-life objects: Superposition principle

The solution for a point cloud of objects can be solved by setting the distance locus error below an adaptive threshold.

Thus, if $f(g_i)$ (g_i being parametrized version of closed solution) is the function that defines the closed point cloud of solution for the object existence, where f(g_i) is the frontal contour boundary in the flatland of the experimental scene; and if b is the streak image data vectorized,

the closed form solution becomes:

$$maximize(g1, g2, g3...gN) \text{ subject to } ||A * f(g_i) - b||_2 < \epsilon \qquad (2.11)$$

where ϵ is the error threshold adaptively set to the error level that covers a point cloud. For example, this error can be the error for the solution of the point-object problem.

If the parametric function is assumed to be an ellipse with the ellipse origin, tilt angle and the major and minor axes, (x_0, y_0, θ, a, b), we get elliptical contour solutions with the extreme curves.

Extreme curves define the left right and front back extreme constraints for the boundary contour of the object existence.

The extreme curves in such a case can be selected as:

1. Lowermost curve: front extrema

2. Uppermost curve: back extrema

3. Left top to right bottom curve: right extrema

4. Left bottom to right top curve: left extrema
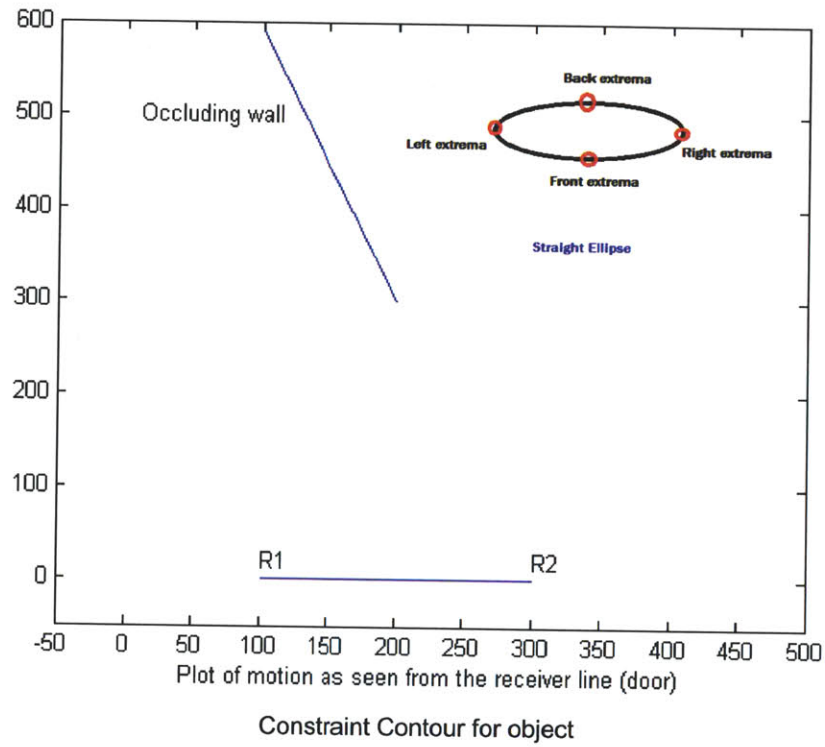
Figure 2-5: If the streaks arriving from left right front and back extremas are normal, we can expect a normal thin ellipse as shown in this figure.
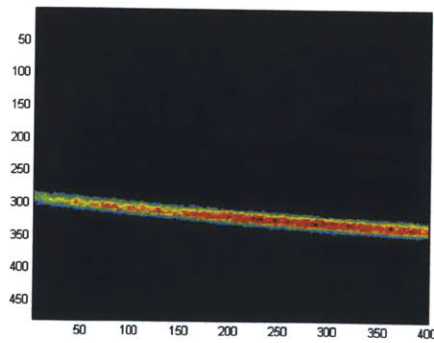
Figure 2-6: If the streaks arriving from left right front and back extremas are normal, we can expect a normal thin ellipse.[Corresponding streak image]

48

Constraint Contour for object
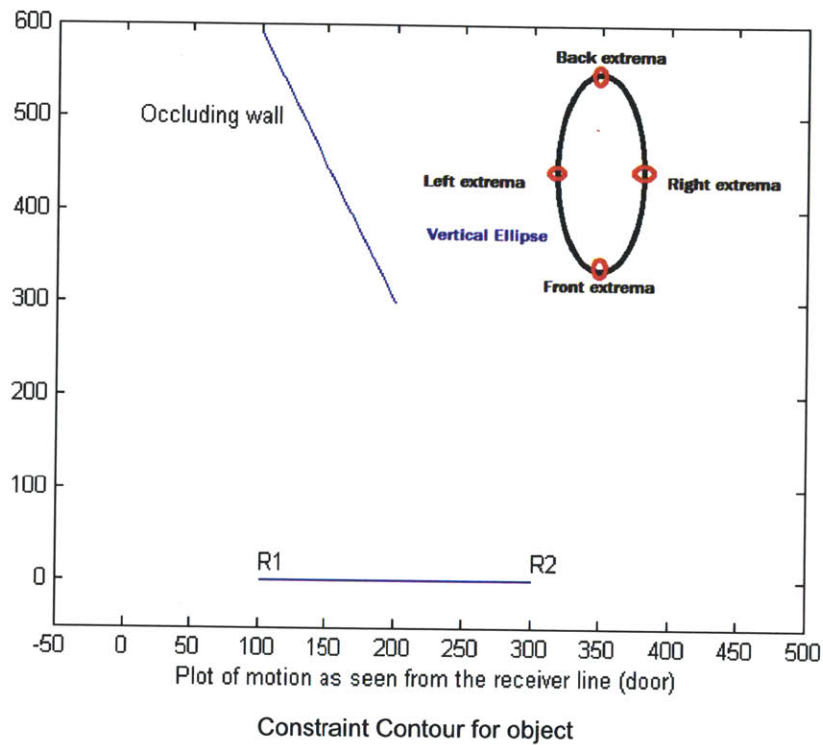
Figure 2-7: If the streaks arriving from left right extremas are normal but have a strong difference in the front and back layer extremas (a consistently thick banana), we can expect a vertically elongated ellipse as shown in this figure.
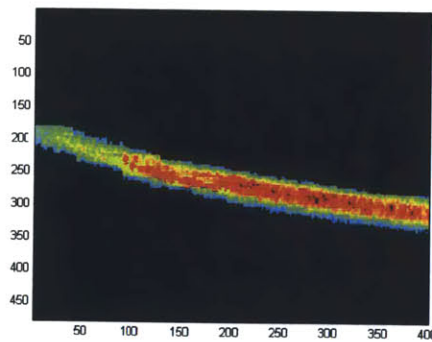
Figure 2-8: If the streaks arriving from left right extremas are normal but have a strong difference in the front and back layer extremas (a consistently thick banana), we can expect a vertically elongated ellipse. [Corresponding streak image]

49
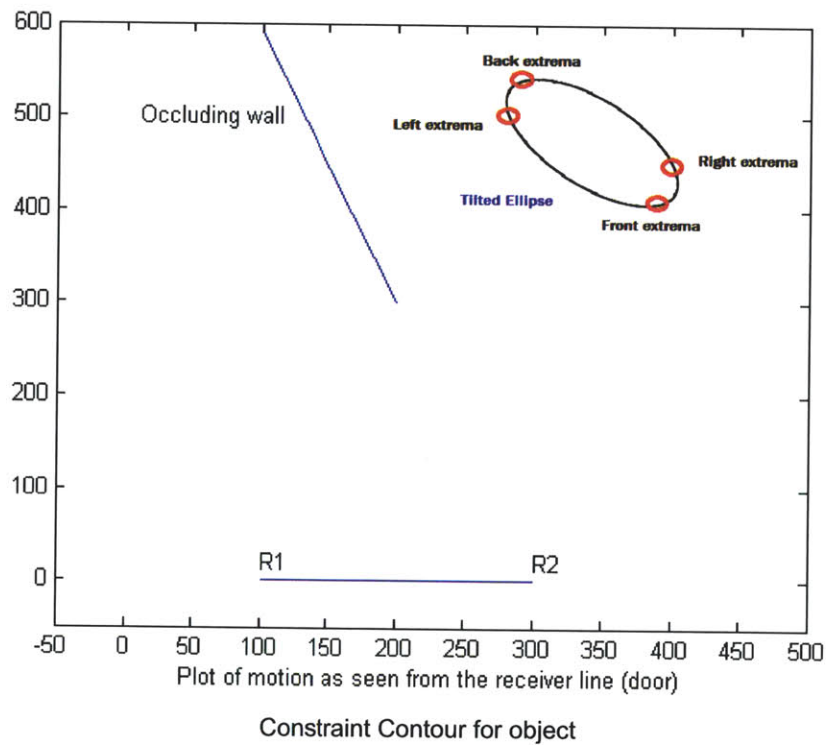
Figure 2-9: If the streaks arriving from left right extremas are asymmetric such that the left side of streaks much thicker than the right side with a strong difference in the front and back layer extremas (a thick banana), we can expect an ellipse tilted towards in a clockwise manner as shown in this figure.
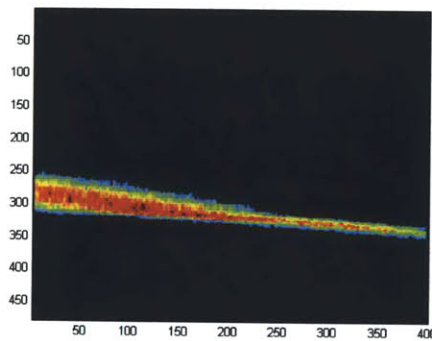
Figure 2-10: If the streaks arriving from left right extremas are asymmetric such that the left side of streaks much thicker than the right side with a strong difference in the front and back layer extremas (a thick banana), we can expect an ellipse tilted towards in a clockwise manner. [Corresponding streak image]

50

2.4 Method for estimating moving object size

2.4.1 Theoretical model for front layer size estimation

We highlight the fact that, so far, we have used the first onset at the primary receivers, while in reality as shown in Figure 2-3 actual extracted space-time responses have a finite width for each receiver resulting in a thick banana shaped profile instead of a thin lined profile curve.

We now consider these complete thick profiles. The thicker profile appearance results from the fact that these photons don't just arrive from the closest point on object (first onset) but also from all other points that the object is made up of. Thus we now extract two extreme profile curves that lead to back projection of two ends of the object as seen from the transceiver wall as shown in Figure 2-11. This is tantamount to finding the diameter of circle of confusion created if we were to solve the least squares problem for all possible choices of profile curves from the banana shaped profile.

Like in the multi point object tracking, if the parametric function of object contour is assumed to be an ellipse with the ellipse origin, tilt angle and the major and minor axes, (x_0, y_0, θ, a, b), we get elliptical contour solutions with the extreme curves.

Note that doing multi point object tracking and object sizing are both effectively achieving the same thing, which is getting the constraint on the object existence contour.

Extreme curves define the left right and front back extreme constraints for the boundary contour of the object existence.

The extreme curves in such a case can be selected as:

1. Lowermost curve: front extrema

2. Uppermost curve: back extrema

3. Left top to right bottom curve: right extrema

Size Contour Constraint for the object based on all extremum arcs

Solving for all possible arcs inside time banana

Back extremum arc

Right extremum arc

Left extremum arc

Front extremum arc

Recorded banana shaped profile
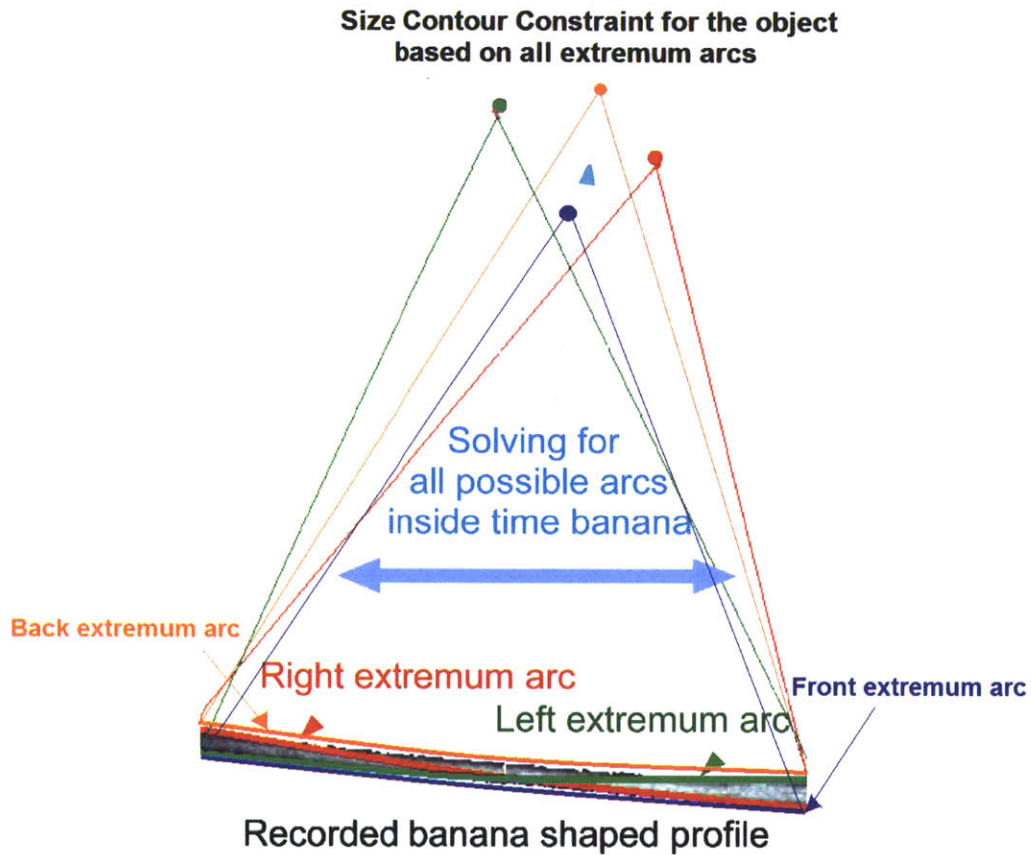
Figure 2-11: We also estimate size of moving object as seen from the primary receivers, by back projecting the extremas of time responses received. Locus of back projections of all possible arcs gives point cloud of the estimated size.

4. Left bottom to right top curve: left extrema

It should be noted that with this approach we estimate the object existence constraint contour as seen from the trans-receiver wall or the secondary surface.

2.4.2 Scope and limitations

(a) Back layer recovery limits

We note that with this technique only the front, back and left, right extremas of the object can be traced with high confidence level. Rest of the points in the point cloud obtained

are with lower confidence level as the arc information within the banana shape cannot be disambiguated. Thus the internal points except the extremities are 'possible candidates' but not the definite points in existence. Concavities even on the front cannot be traced back. This happens because all the concavities produce streaks that get mixed/superimposed with the streaks received because of adjacent points. Thus all one has is a banana shaped profile. The most legible information from the banana are the extremity curves as shown in the diagram. Using this we get the extrema information of points that are pat of the object positions.

(b) Benefits of front layer width estimation

Although the technique can estimate only front layer width, this information in itself is very useful. In most of the cases over a period of time one can take front side width measurements and gauge the real size range of the object. This approach fails only in cases when the target object is moving in only one orientation all the time, and has strongly disproportionate length in the back layer. For this to happen: (a) The object has to keep moving with a constant orientation with respect to the secondary surface, which is not possible as if its moving with a normal pose, its orientation w.r.t the secondary surface keeps on changing and hence it is possible to see further parts of the object as it moves. (b) It should have a disproportionate length in the blind spot areas, that is in the back layer. Both of these possibilities are rare to find, thus in most of the real-life cases this approach serves the purpose.

Chapter 3

Hardware Setup

3.1 Femto second laser and pico second streak camera

3.1.1 Femto second laser

A femto second laser is a laser that can emit short light pulses of durations well below 1 ps, (1 femto second $= 10^{-15}$s). Such a short pulse is achieved by a phenomenon called passive mode locking. This laser belongs to the category of ultra short pulse lasers.

Some of the vital characteristics of these types of lasers are:

1. Pulse duration: This is the duration for which light energy is emitted out of the laser in a pulse.

2. Pulse repetition rate. Since these lasers are very high frequency lasers in the sense that they emit many short pulses at a high frequency (to the tune of MHz), the pulse repetition rate define the total energy emitted out averaged out over a second.

3. Time bandwidth product (TBP): It shows whether the spectral width is greater than required for the given pulse duration for a laser. The pulse quality involves other

aspects such as details of the temporal and spectral pulse shape, such as the presence of temporal or spectral side lobes.

For this thesis work, we used a 80 fs pulse duration, 790 nm wavelength, 800 mW Average output power pulses of a femto second laser.

3.1.2 Pico second accurate streak camera

[Source: Hamamatsu corporation: C5680 series data sheet]
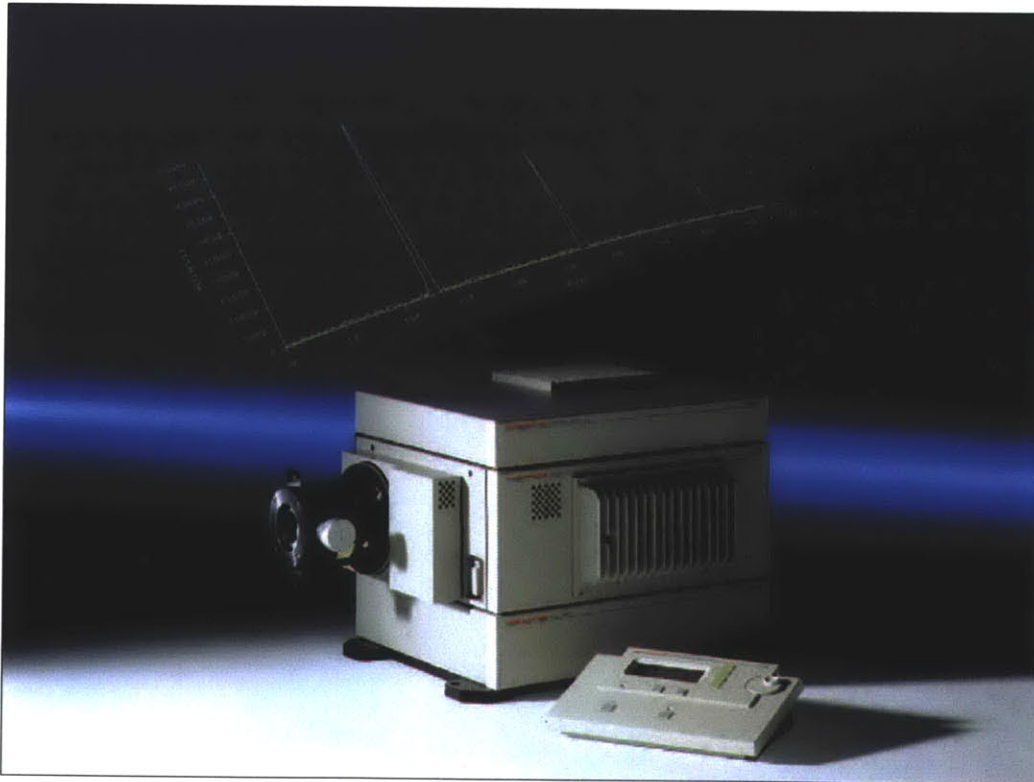
Figure 3-1: Figure shows Hamamatsu C5680 series streak camera that allows space time information capture by photon time of arrival and line of arrival. Source: Hamamatsu C5680 streak camera data sheet

(a) Operating principle The light pulse as a response of the scene that enters the camera, gets focused onto the photo cathode tube of the camera with the help of a lens. The streak

tube converts the information inherent in the arriving photons into electrons and these electrons are then accelerated and conducted through the tube which registers the time offsets as well as variation in energy (intensity). The electrons are deflected in different angles at different speeds finally touching the Micro Channel plate (MCP). When electrons arrive at the Micro Channel Plate, they are multiplied by hundreds of times and are then projected towards the phosphor screen, which again converts them back to light. The fluorescence image corresponding to the time of arrival of these photons is then created at the screen. This the vertical axis of the camera serves as a temporal axis.
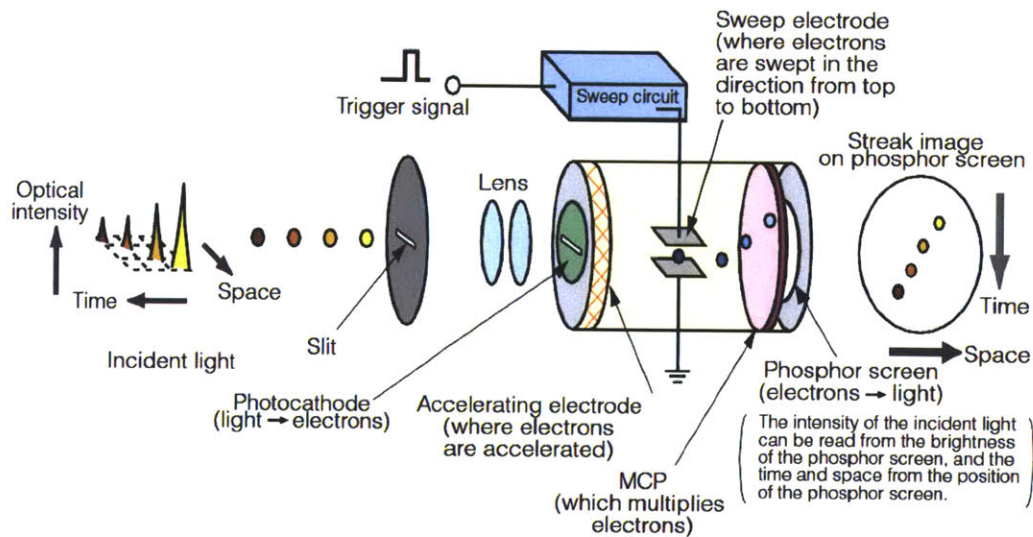
Figure 3-2: Figure shows a diagram explaining the components vital in the streak camera operating principle as described in the text. Source: Hamamatsu C5680 streak camera data sheet

(b) Camera hardware configuration The camera is operated in a streak mode with the selection of: C5680 main unit, selection of the input optics, selection of streak tube and the output format. The display ends up showing streak images with one axis being a one dimensional space and other dimension being time in picoseconds. The MCP gain for the camera was found to be non-linear and thus was characterized and normalized by the modeling performed.

① **C5680 main unit (with power supply and camera controller)**

③ **Function expansion unit**

② **Sweep unit**

Selection of C5680 main unit
Selection of input optics
Selection of streak tube
Selection of output format

Selection of sweep unit

Selection of function expansion unit

[Suffix (Model No.)]
One of the following suffixes is appended to the model number of the C5680, depending on the type of streak tube and output format used.

C5680-2x

X=1 Accommodates 200 nm to 850 nm, 1 MCP
2 Accommodates 300 nm to 1600 nm, 1 MCP
3 Accommodates 115 nm to 850 nm, 1 MCP
4 Accommodates 200 nm to 900 nm, 1 MCP
5 Accommodates 200 nm to 850 nm, 2 MCP

Figure 3-3: Figure shows a diagram that shows the functional configuration in which the streak camera was used. Source: Hamamatsu C5680 streak camera datasheet

3.2 Experimental scene setup

3.2.1 Experimental scene

Figure 1-1 shows our experimental setup that has optical-bench-scale replica of a real life scene of estimating motion and size of objects moving inside a region with no direct access through a camera or flash. Both camera and laser source are blocked by the black wall, with an indirect light access possible through a bounce on the diffused wall surface shown as transceiver wall.

For our experiments, we construct a time-of-flight camera by combining a femto second laser and a streak camera. In particular, we use a Ti:Sapphire laser to illuminate a spot on the transceiver wall with 80 fs, 790 nm wavelength, 800 mW Average output power pulses and a Hamamatsu C5680 universal streak camera that can resolve up to 10 ps level time differences of arrival. We capture the time resolved photon arrival profile over a 15 cm long one dimensional space giving a 2D space vs time image for a nanosecond wide period as one capture frame. This one dimensional space observed is a line of array of primary receivers that can see moving objects. The 2D images captured are termed as space-time impulse responses ($I(x,t)$) of the scene.

We create a cluttered environment of miscellaneous objects in the region behind the wall. In the experiment we move objects coated with diffuse lambertian reflecting paint along a line in x-y plane and we take one nanosecond-long space-time images of scene at time instants separated by $T = 1$ sec. Each of these space-time images has a resolution of 10 ps on time axis and 0.2 mm on spatial (receiver array) axis. We refer to each of these space-time images as $Frame_i$ and $Frame_{i+1}$.

3.2.2 Mitigating uncalibrated time bias:

Since the exposure time of time-of-flight camera is much shorter than actual travel time for path from laser to camera, there has to be a time shift between laser and camera trigger. Any absolute time shifts are impossible to calibrate because small mechanical shift in devices can create multiple pico second of errors. So similar to uncalibrated pose (i.e. unknown spatial relationship) in many computer vision algorithms, we choose uncalibrated global time shift (unknown temporal relationship) based recovery. We assume that this time shift is constant across a set of measurements if they are taken within a few seconds of each other. However we do not know value of this global time shift. Thus absolute times of arrival are unavailable. Nonetheless, our approach uses only differential times of arrival relative to arrival time of the first photon from the moving target.

3.2.3 Choice of scale

At present we have worked on an optical bench scale environment that is 1m by 1m wide. This choice of scale is for several reasons. Most important reason being the laser intensity fall off at longer distances. For a longer distance the lasers required will be of very high power (we did not have access to such high powered lasers). Moreover, at present stage such high powered lasers are not practical to be used. However over a period of time we expect higher powered lasers to reduce in size as well as cost.

The ultra fast imaging devices today are quite bulky. The laser sources and high speed cameras fit on a small optical bench and need to be carefully calibrated for triggering. However,

there is a parallel research in femto second solid state lasers and they will greatly simplify the illumination source. Pico second accurate single pixel detectors are now available for under $100. Building an array of such pixels is non-trivial but comparable to thermal-IR cameras. Nevertheless, in the short run, we are building applications where portability is not as critical. For endoscopes, the imaging and illumination can be achieved via coupled fibers.

3.2.4 Clutter and other factors that affect multi path capture

Following factors affect the camera signal capture:

1. Occlusions: If there are objects present within the shortest path between target moving object and receiving secondary surface, light coming from the target object gets blocked out leading to zero signal reception from the occluded areas.

2. Dust particles: Participating media in the environment of operation also behave similar to occlusions and harm the signal capture. Presence of strong haze in the environment can severely affect signal capture. At present we have no mathematical modeling around the problems caused by participating media however concepts from communication theory can be borrowed to deal with multi path fading and propagation.

3. Distances traversed by light: The laser intensity falls off with the squared power of distance traveled. In a four path traversal system as used in this work, each of these distances affect the intensity falloff.

4. Angles of incidence: There is a fall off proportional to the cosine of angle of incidence when light falls on a surface. In our case there are three bounces off a surface contributing to cosine fall offs.

5. Ambient light: While operating in bright conditions the ambient light strongly overpowers the laser intensities, especially the returning laser signal which is to the tune of a millionth fraction of what was beamed by laser.

Chapter 4

Results

4.1 Temporal and spatial resolution performance:

Using a camera with 10 ps temporal resolution does not directly translate to 3 mm accuracy because of the error bounds within the least squares problem solution. Thus we recover up to cm level accurate results.

4.2 Result datasets

We tested our framework on following real-life scenarios.

4.2.1 Moving human figure

Figure 4-2 shows results for a scenario where a human figure moves from position P to Q in a room that has other people present in it. The cluttered environment consists of CD (diffractive reflection), metallic cylinders (specular reflection) and transparent water container (refraction). Motion from P and Q is a linear motion in flatland comprising of the points shown in Figure 2-2. Our results correctly estimate motion and size of the moving human figure and are accurate up to cm level.
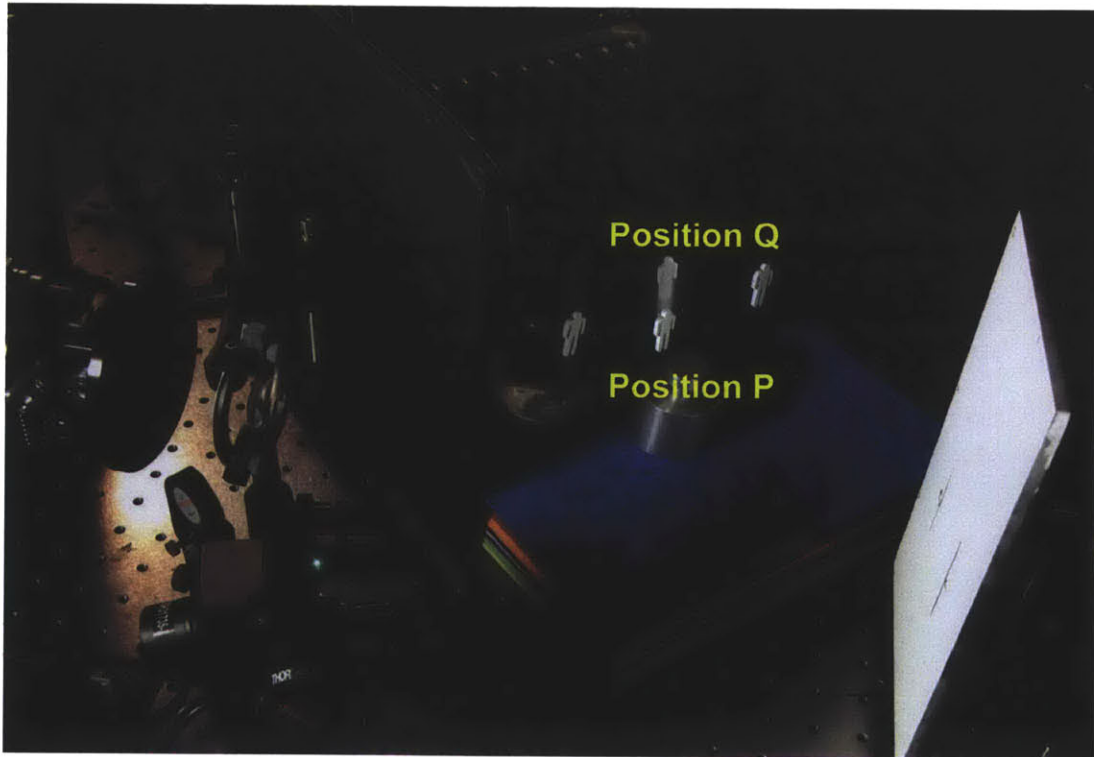
Figure 4-1: Real scene depicting motion of a human figure in the center moving from position P to Q as a linear motion in the flatland comprising of object O, R_i line and points L and C_i. The cluttered environment consists of CD (diffractive reflection), metallic cylinders (specular reflection) and transparent water container (refraction).
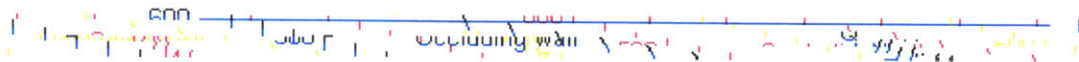
Figure 4-2: Motion estimation results (P,Q) for moving human figure in cluttered surroundings, compared with the ground truth (M,N). Scale: 100 units = 7cm.
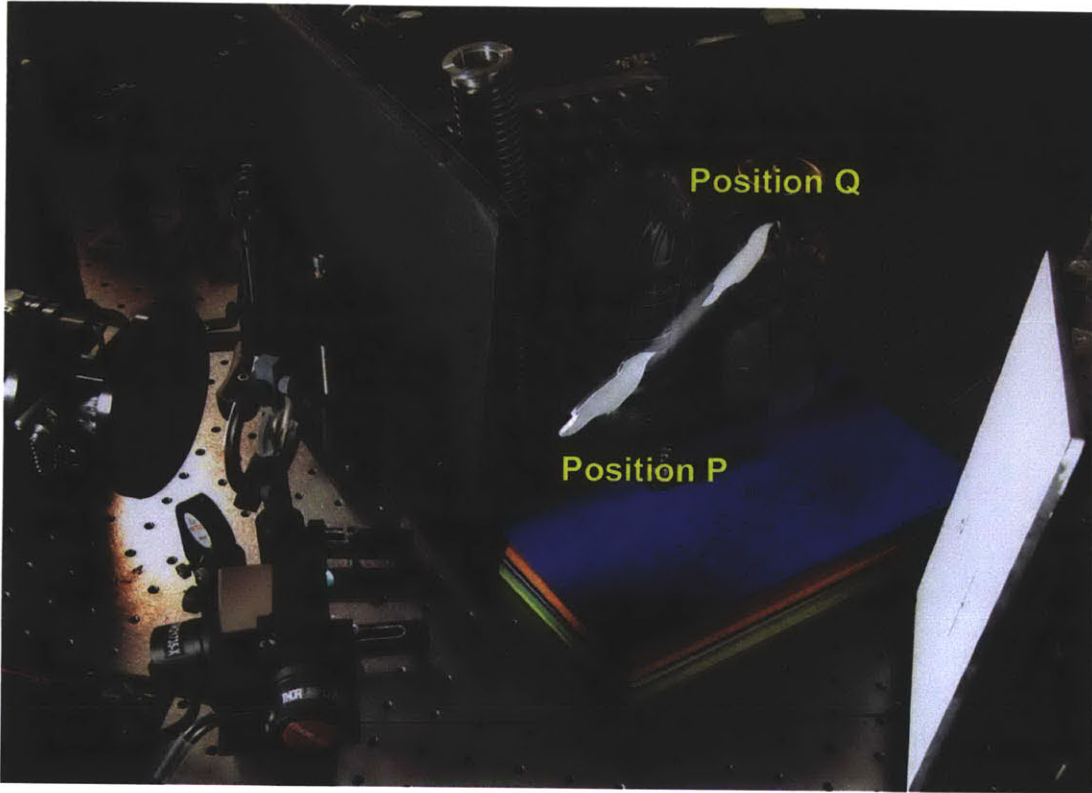
Figure 4-3: Real scene depicting motion of a bird moving from position P to Q as a linear motion in the flatland comprising of object O, R_i line and points L and C_i. The cluttered environment consists of lens (refraction), metal flowerpot (specular reflection) and water bottle (refraction). Scale: 100 units = 7cm.

4.2.2 Moving bird in cluttered surroundings

Figure 4-4 shows results for a scenario where a bird moves from position P to Q in a NLOS region filled with clutter like CD, lens and water bottle. Motion from P and Q is a linear motion in flatland comprising of the points shown in Figure 2-2. Our results correctly estimate motion and size of the moving bird and are accurate up to cm level.
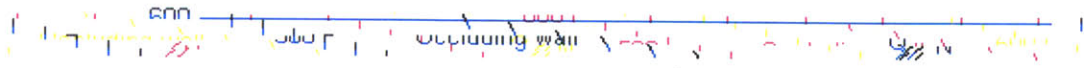
Figure 4-4: Motion estimation results (P,Q) for moving bird in cluttered surroundings, compared with the ground truth (M,N)

4.2.3 Dual frame and Multi frame motion tracking

We note here that successive dual frame tracking is not the most optimal method. In cases when the target moving object displacements are smaller than the object size, the estimates do not exactly correspond to the free space object points as the streaks from location P and location Q cannot be disambiguated in such cases.

However, with a multi frame longer duration of observation this problem gets easily solved as the interpolation can lead to intermediate positions. Here we exploit the fast nature of signal capture in streak cameras as the motion can still be treated as linear over the longer period of observation.

We show here the results of dual frame tracking and multi frame tracking and how multi frame tracking gives ideal solution.

4.2.4 Varying size objects: Size estimation

We varied object sizes from 1 cm to 5 cm in increments of 1cm and analyzed the performance metrics for these size changes. We show our conclusions and results in following descriptions.

Size and Displacement dependence of performance: A very small object creates an overall small Signal energy imprint in the captured signal and hence is difficult to deal with, similarly, low displacement creates overlapping time responses which are difficult to separate when the two frames show intensity variations as elaborated in Figure 4-16.

4.3 Performance Analysis

4.3.1 Size and displacement dependence of performance

Figure 4-16 gives an elaborate explanation of the size and displacement performance of the approach. The best case for operation is when the inter frame displacements are larger than
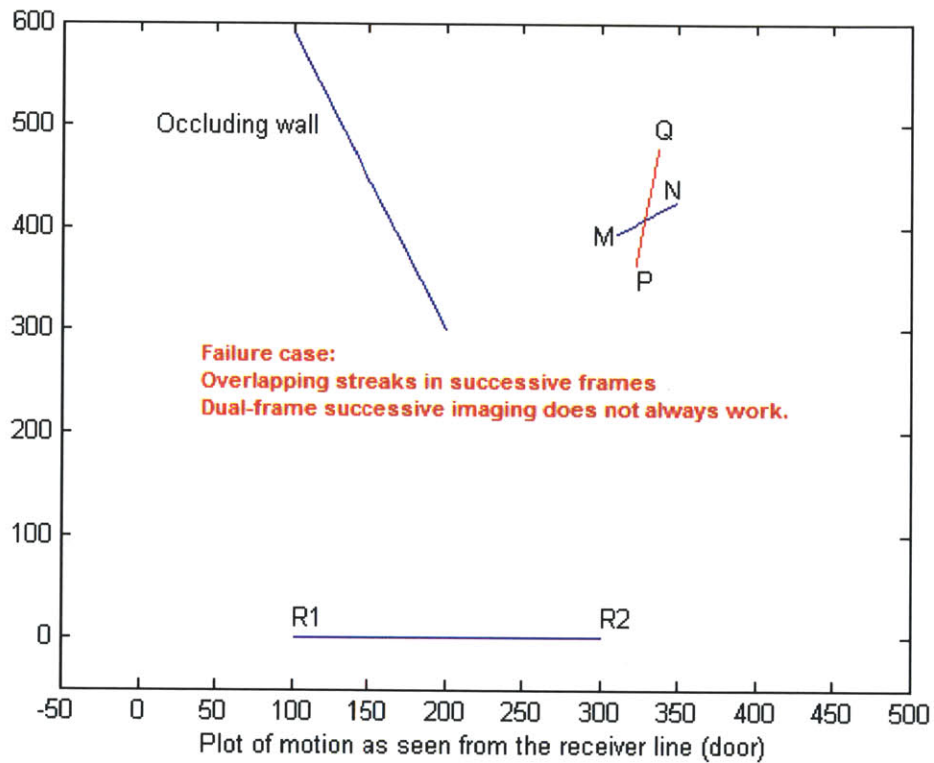
Figure 4-5: Motion estimation results (P,Q) for moving bird in cluttered surroundings, compared with the ground truth (M,N). This result is a failure case of dual-frame successive tracking where the displacement of object between successive frames is smaller than the object size.
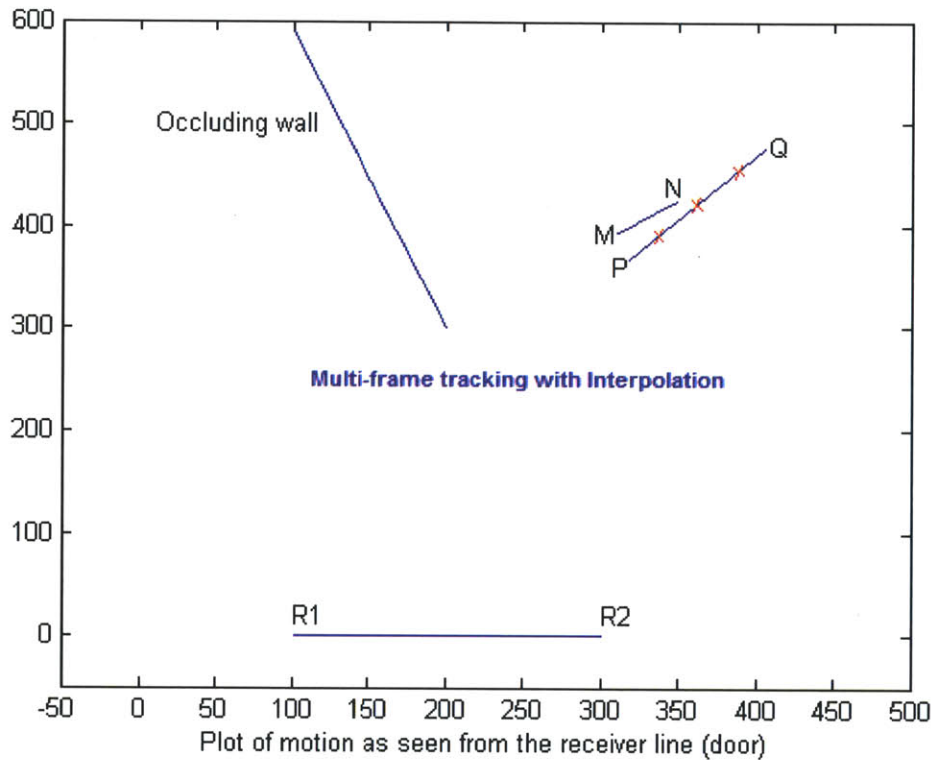
Figure 4-6: Motion estimation results (P,Q) for moving bird in cluttered surroundings, compared with the ground truth (M,N). This result shows that multi frame tracking with interpolation works in cases of dual-frame successive tracking failure where the displacement of object between successive frames is smaller than the object size. Observing the object for a longer time enables larger displacements to be captured and interpolating within them gives intermediate positions under the assumption that motion is linear within the few nanoseconds of observation. For example in this figure, using ith and (i+4)th frame can give the three intermediate positions. The advantage being that the (i+4)th frame and ith frame do not have an overlap in object positions even if the the object size is slightly larger.
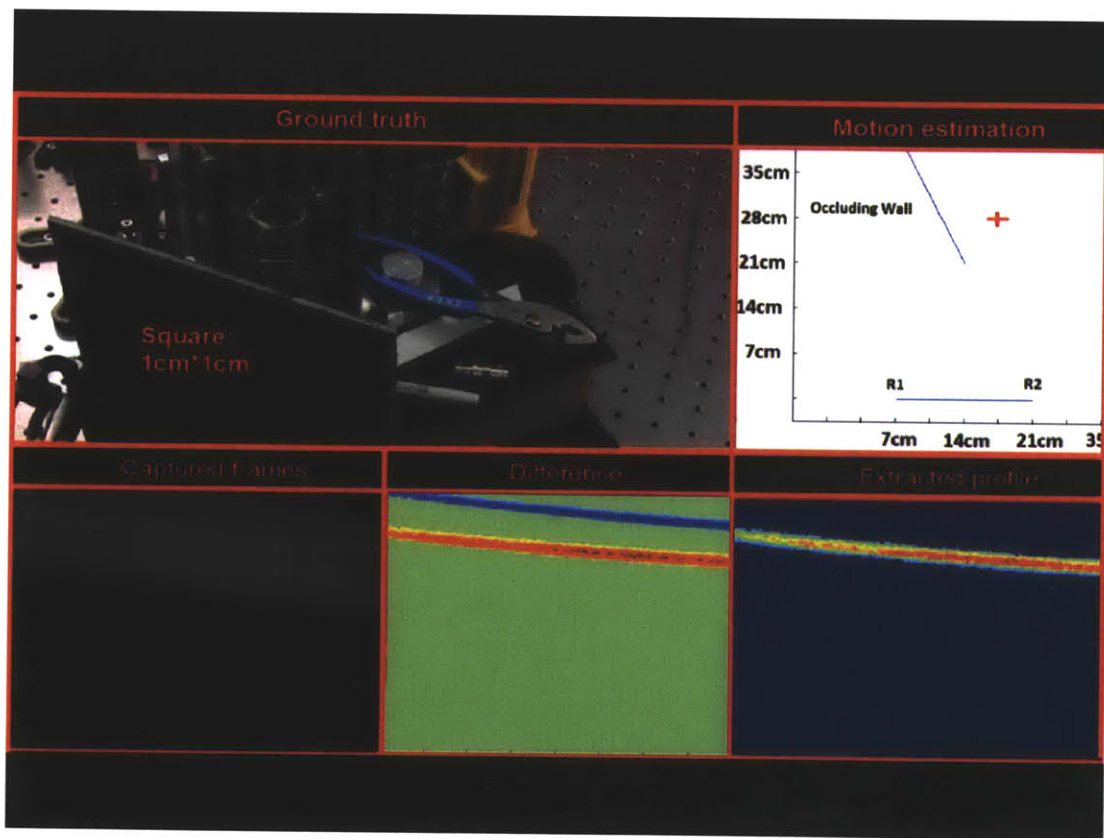
Figure 4-7: Motion estimation results (P, Q) for moving object in cluttered surroundings, compared with the ground truth (M, N). Scale: 100 units = 7cm. result from Frame 1-2.
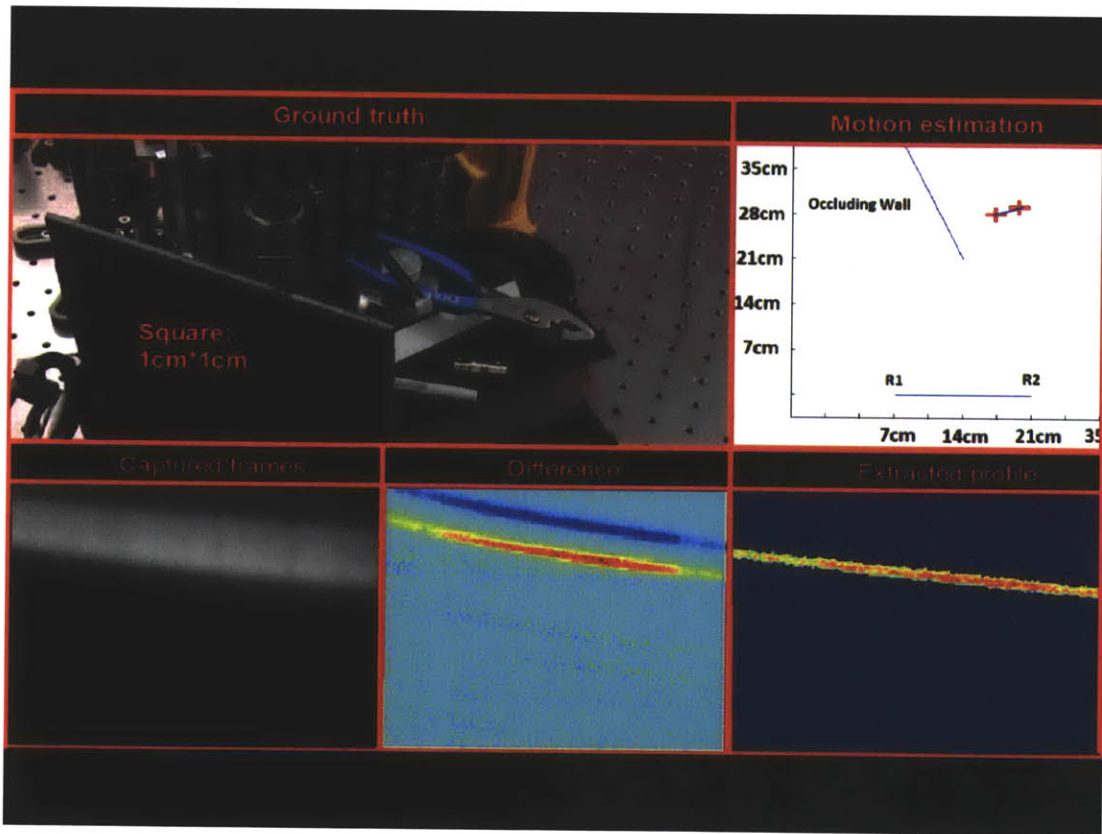
Figure 4-8: Motion estimation results (P,Q) for moving object in cluttered surroundings, compared with the ground truth (M,N). Scale: 100 units = 7cm. result from Frame 2-3.
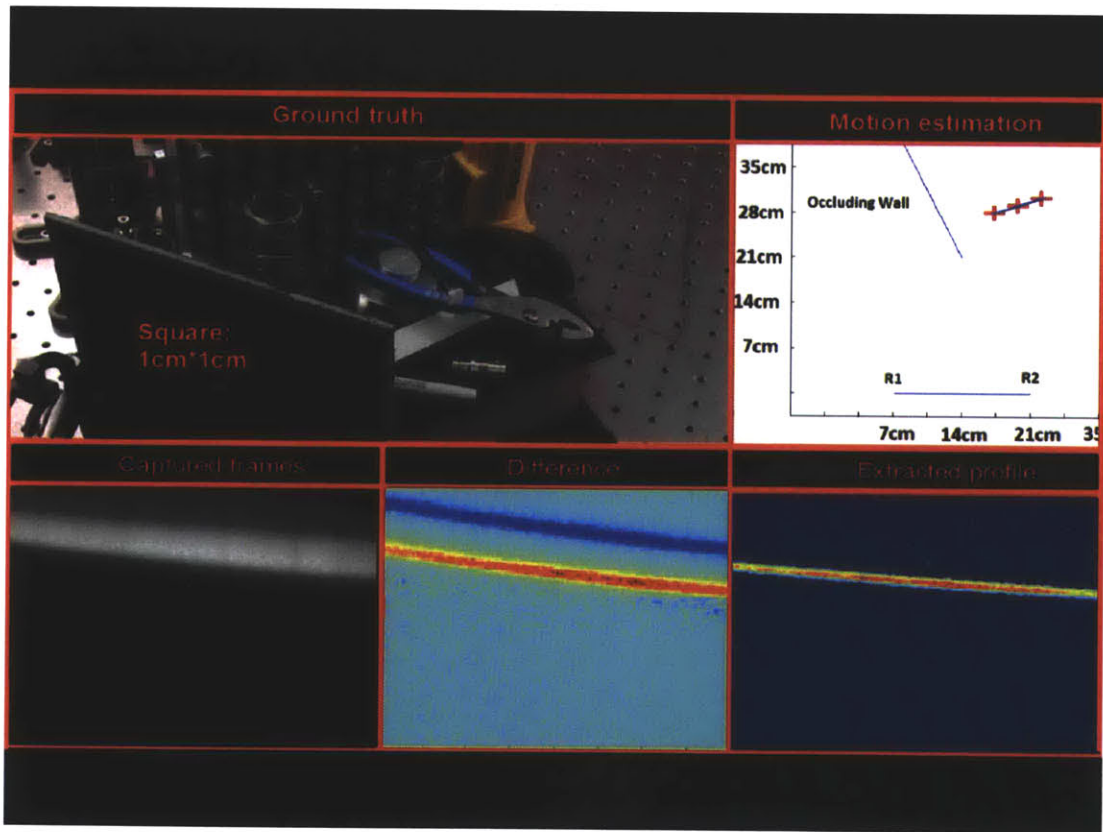
Figure 4-9: Motion estimation results (P,Q) for moving object in cluttered surroundings, compared with the ground truth (M,N). Scale: 100 units = 7cm. result from Frame 3-4.
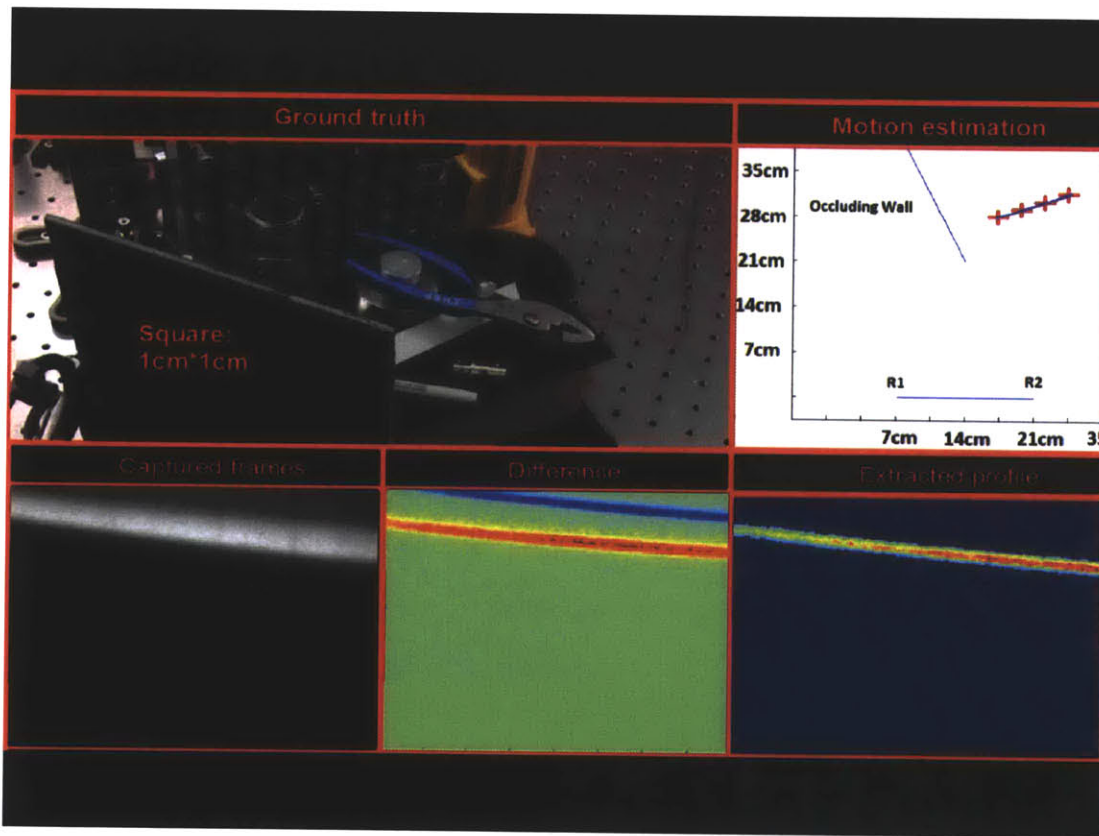
Figure 4-10: Motion estimation results (P,Q) for moving object in cluttered surroundings, compared with the ground truth (M,N). Scale: 100 units = 7cm. result from Frame 4-5.
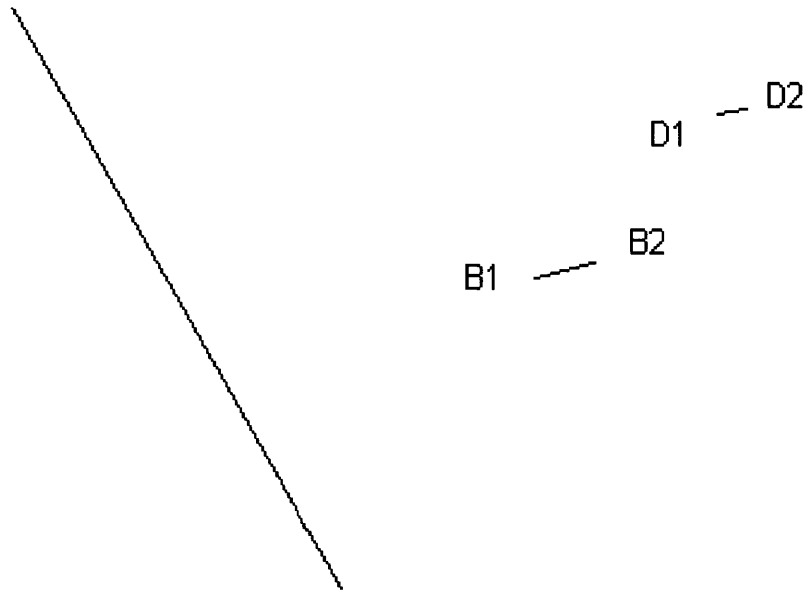
Figure 4-11: Size estimation results B_1B_2, D_1D_2 for moving bird and moving human figure. Sizes indicate the upper bounds on the size as seen from the transceiver wall.

the object size. To extend the abilities to the areas where the inter frame displacements are smaller than object size, one can use multiframe tracking with interpolation.

4.3.2 Resolution

The resolution of objects tracked is limited by the experimental as well as mathematical inversion performance. The camera used has a time slice limit of 2 picoseconds. Thus the theoretical spatial limit is 0.6 mm. However the mathematical inversion errors as well as noise in signal capture reduce the resolution accuracy. At present our results are accurate up to centimeter level.

4.3.3 SNR

Understandably, the signal to noise ratio uis very poor in these experiments. As one can imagine, very little light comes back after three bounces off diffuse surfaces. Moreover the
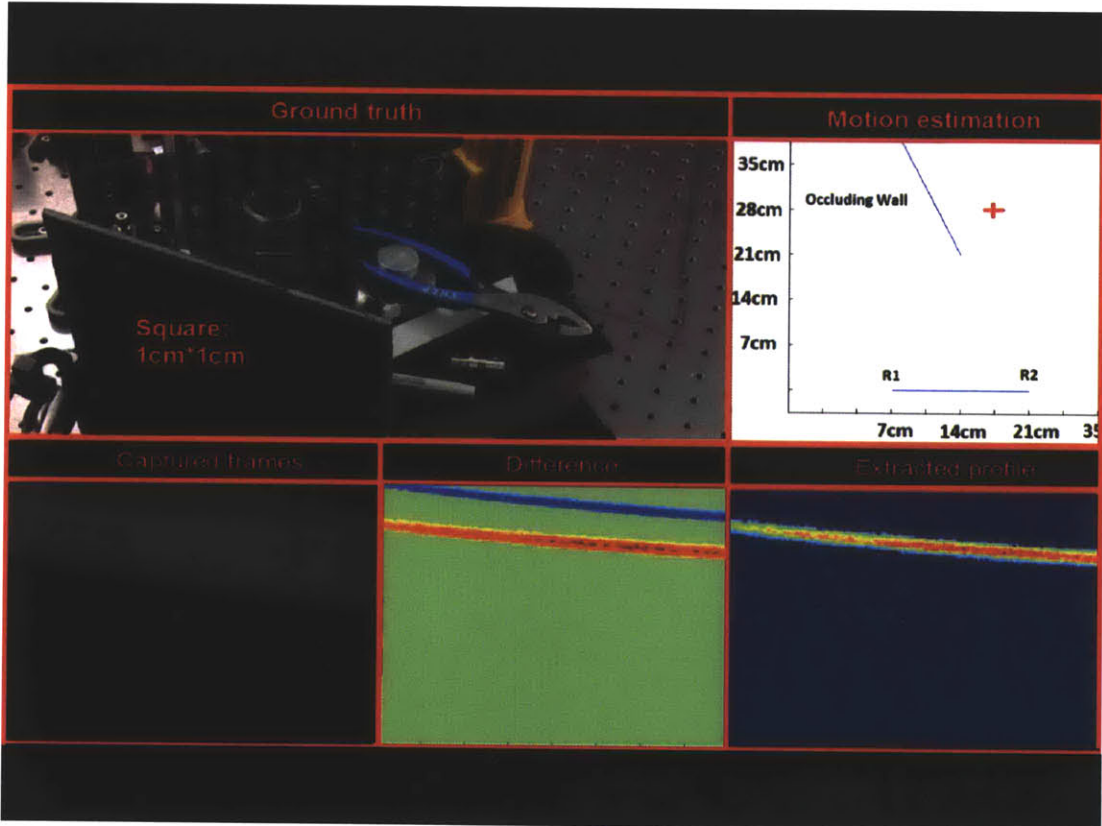
Figure 4-12: Diffused 1cm*1cm square shaped object moving at a constant velocity was imaged using our ToF camera and the location of the object was spotted. Figure above shows the ground truth image, the time responses at camera, the frame difference image, the extracted profile and the location of the spot computed using the tracking algorithm described.
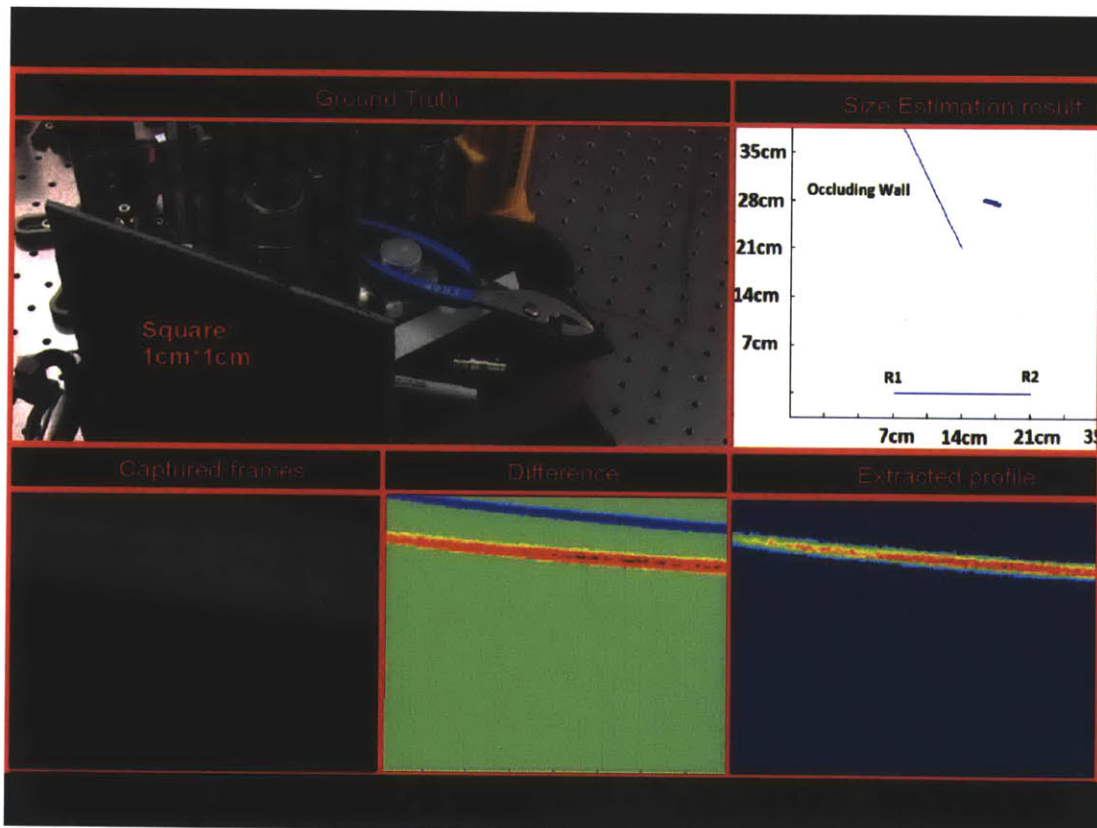
Figure 4-13: Diffused square objects with centimeter level increments on size were used for size estimation experiments. Our algorithm estimates sizes of objects up to centimeter level accuracy. 1cm by 1 cm object.
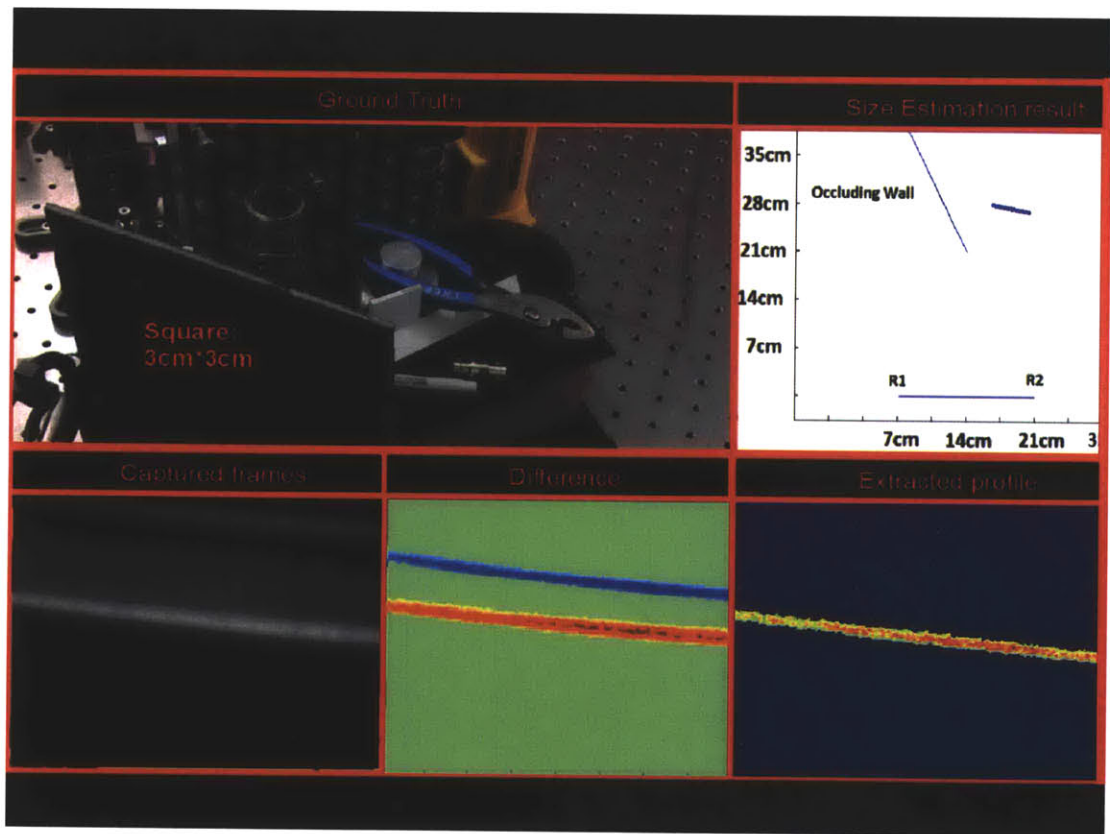
Figure 4-14: Diffused square objects with centimeter level increments on size were used for size estimation experiments. Our algorithm estimates sizes of objects up to centimeter level accuracy. 3 cm by 3 cm object.
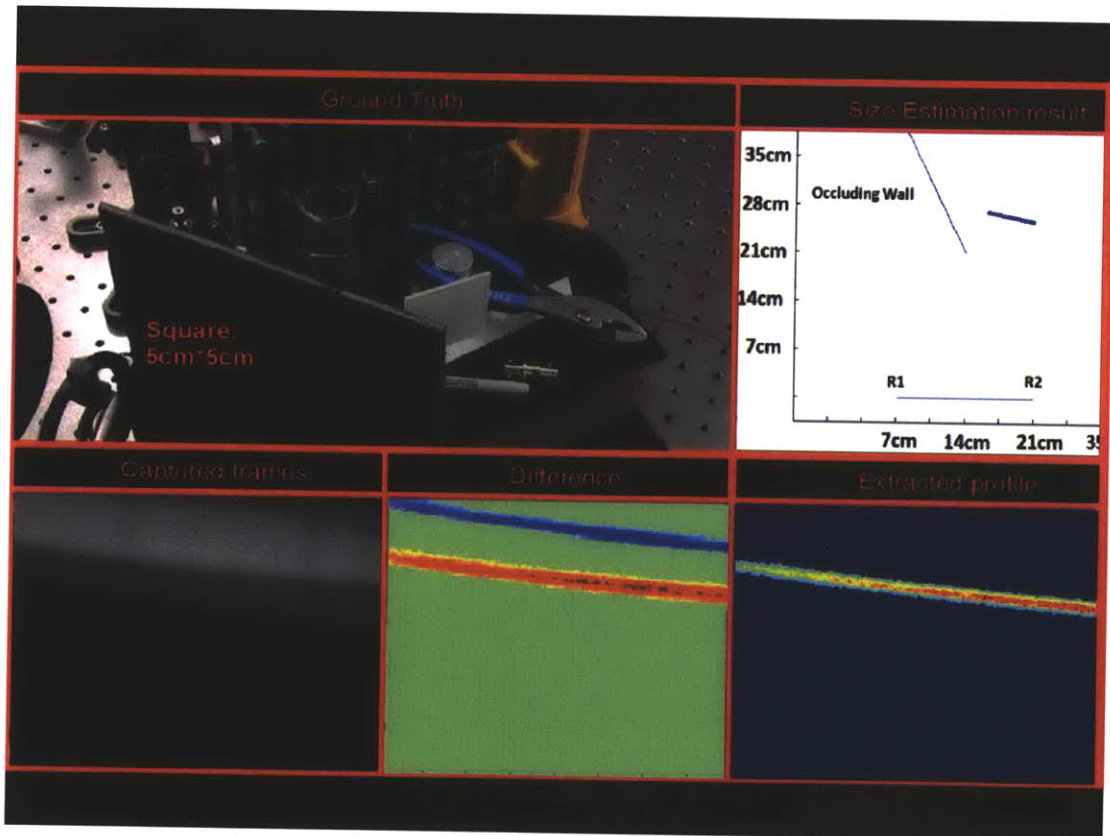
Figure 4-15: Diffused square objects with centimeter level increments on size were used for size estimation experiments. Our algorithm estimates sizes of objects up to centimeter level accuracy. 5cm by 5 cm object.
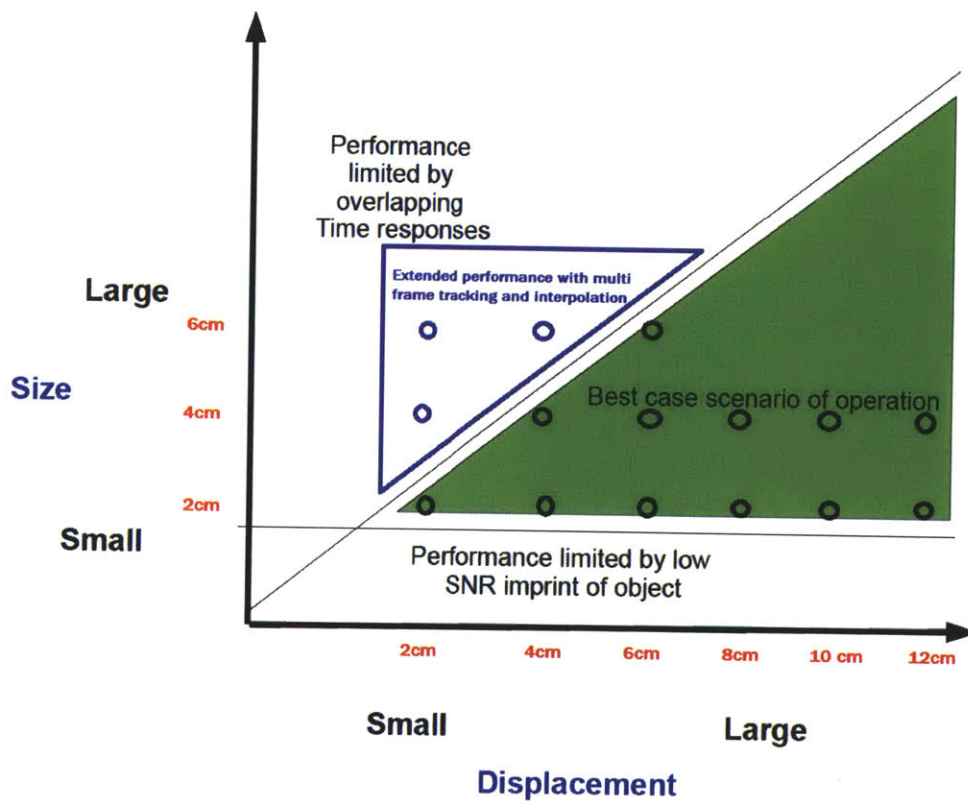
Figure 4-16: Performance matrix showing key factors that boost or deter performance of our approach for size and displacement of target. The black data points are real points for ideal scenario of operation when the displacements are larger than 2cm and are larger than the size of the object. The blue points show the extended region of operation achieved by using multi frame capture and interpolation for linear motion (over a few nanoseconds).

incident intensity itself is limited by the pulsed laser intensity as well as pulse repetition rate. Quantitatively, we get only a millionth of the energy back.

4.3.4 Objects contributing to same pixels in streak images

Our approach infers information from streak images that can be associated with points in free space. However ambiguities arise when there are multiple objects at the same distance from the receiving sensor pixel and thus end up mixing their space-time impulse responses. These space-time impulse responses are essential for locating the point in free space. if they are mixed, with the current system there is no clear way to disambiguate them. Thus, such a case will be a failure on the approach. However, such a scenario rarely arises as the difference of frames reduce regions of interests to moving objects only, and the same path lengths within vicinity are acceptable as they contribute to the point cloud as explained in the real-life non point object solution approach earlier in this chapter.

Ways to disambiguate overlapping streaks are coding in time and space. Here we would like to refer to Coded exposure and coded aperture approaches in computational photography. An equivalent of these techniques used with pulsed laser and streak capture can enable clearer disambiguation of these streaks.

4.3.5 Overlap within inter-frame displacements

If an object is moving very slowly, its position can overlap within the successive frames of capture. In such a case, $frame_{i+1}$ - $frame_i$ will not result in clear disambiguation of the object positions. In such a case, its best to use interpolation in multiple frames and using frames that are further apart in displacement of the object.

4.3.6 Participating media

Presence of dust, haze or other participating media Beverley affects the streak signal capture. In our experiments we used a cleaner environment. However while dealing with

participating media, one can borrow concepts from communication theory to deal with multi path fading and propagation.

One way to deal with the participating media is to send pilot signal bits as used in multi path fading approaches to figure out models for the fading effect models as well as multi path echoes.

Chapter 5

Towards reflectance estimation

This work was done in collaboration with Andreas Velten (post doctoral fellow in Camera Culture group) and Nikhil Naik (graduate student at MIT Media Lab), Shuang Zhao (graduate student from Cornell University) and Prof. Kavita Bala (Professor in Computer Science from Cornell University). I would like to thank my collaborators for the experimental data collections and insights into the signal capture and the methodology. I would also like to acknowledge my advisor Prof. Ramesh Raskar for the numerous discussions on experimental design setup and the methodology. My contributions in the project are towards the problem definition, algorithmic mathematical model development and developing the code to invert data captured to achieve reflectance model values.

5.1   Overview

We demonstrate a new technique that allows a camera to rapidly acquire reflectance properties of objects in a scene from a single viewpoint, over relatively long distances and without encircling equipment. Our system exploit interreflections among scene points to indirectly observe a BRDF at a scene point from multiple incoming and outgoing directions. To scale this to a general scene, we introduce the concept of time-of-flight reflectance estimation. We demonstrate a laser and time-of-flight camera based combined hardware framework, that

measures material properties by indirectly illuminating an object, and observing reflected light from a laser source indirectly. Our time-of-flight camera captures time responses to a short pulse illumination, observed over another tertiary surface. This lets us collectively acquire dense angular but low spatial sampling of the BRDF, within a limited solid angle range -all from a single viewpoint.

We experimentally show that this method can recover parametric BRDF models for homogeneous and heterogeneous patches. We analyze the constraints and performance of the acquisition system, and discuss several plausible application scenarios. As compared to lengthy or highly calibrated BRDF acquisition techniques, our goal is not to mimic their performance but provide a unique tool for photographers to capture meaningful scene information using a fast and portable device.

5.2 Introduction to reflectance capture using multi path analysis

The goal of computational photography is to capture sufficient information about the scene so as to provide ultimate post-capture control. Examples include the ability to manipulate exposure, focus or lighting. In this thesis, we show how to capture the reflectance in a scene from a single viewpoint without instrumenting the room. Knowing how a material reflects light in multiple incoming and outgoing directions has long been a problem of interest in computer graphics. The simplest measurement of a surfaces' material would capture its albedo. More generally the Bidirectional Reflectance Distribution Function (BRDF) characterizes light reflection from surfaces. BRDFs, once known, are useful in rendering, relighting, as well as for matching and material identification. In this paper, our goal is to capture scene reflectances by densely sampling a finite cone of incoming and outgoing directions. We assume that the scene has a few different materials (on the order of 10s) that we can acquire quickly, and conveniently using a portable device that can operate at a distance from the surface.

For this we exploit two phenomena (a) interreflections and (b) time of flight imaging. There

82

is a long and rich history of acquisition of spatially varying BRDFs, bidirectional texture functions, and parametrized BRDFs [5, 27, 36, 30, 7]. However, typical acquisition techniques assume carefully calibrated setups with long acquisition times. In this thesis we exploit fast (femto second) time-of-flight imaging and indirect observation to achieve very fast acquisition of materials "in the wild". We propose a compact device that can be used directly in the field without instrumenting the space with encircling equipment or coming in close proximity of the surface. This fits in the general spirit of computational photography to allow one to capture meaningful properties and then allow powerful post-capture operations, in this case to relight or edit materials.

5.2.1 New abilities in reflectance capture

We present a new technique for BRDF acquisition using multi path light transport. Our approach uses indirect viewing (instead of directly viewing the material being measured) with 3-bounce scattering, coupled with time-of-flight imaging to capture reflectances. To the best of of our knowledge this is the first approach to use indirect measurement in this manner. The inherent challenge we address is to decode material measurements in the presence of mixing over angular and space dimensions. Using time-of-flight (ToF) principles, ours is probably the first solution to rapidly and remotely recover the BRDF.

Another benefit of this approach is that by indirect viewing (by observing measurements indirectly on a large receiver surface) we can acquire a large angular sampling of the material in a single step. This is in contrast with traditional approaches require a dense angular sampling to ensure that specular peaks are not missed.

- We show the ability to infer BRDF from tertiary reflections via de multiplexing.

- We show combining linear inversion (de multiplexing) and parametric fitting in a single step to speed up the capture process.

- We analyze the invertibility and show that although it is impractical to recover heterogeneous BRDF using slow cameras, fast ToF cameras can recover the BRDF.

- We show physical experiments using ordinary as well as ToF cameras, and validate the results.

5.2.2 Advantages and limitations of multi bounce reflectance capture

Benefits:

1. Our approach enables extremely fast BRDF acquisition in seconds (in seconds with our prototype, in principle in nanoseconds).

2. Our approach does not need instrumentation of the target material

3. Our approach detects specular peaks in a single measurement.

4. Since we use indirect reflections, our approach allows for a remote BRDF capture without having to be very close to the target material.

5. We also translate our approach into a device design that can allows non-laboratory BRDF capture in real life

   environments.

6. We show results on material editing in real life scenes, using our fast BRDF acquisition approach.

Limitations:

1. Our approach does not necessarily sample all half angles of all materials in the scene.

2. Our approach has a limited resolution for small features with varying BRDF or surface normals.

Capture Dimensionality The fact that the captured data is only one dimensional reduces the diversity of our collected data but is no serious limitation. The direction missing from the camera can be probed by the laser that can be scanned in 2 dimensions across the scene.

Signal to Noise Ratio and Scalability   The signal to noise ratio (SNR) of the combined streak camera system is about 1000:1. It can be increased dramatically by using the camera in photon counting mode.

Another way to improve the SNR is to bundle the same laser power into fewer pulses at a lower repetition rate but with the same pulse length.

Intensity Limiting   Higher output power generally increases the intensity of the light striking the scene and can cause safety issues and in extreme cases damage the scene material. The intensity can be lowered by increasing the laser beam diameter.

Time Resolution   A more challenging problem is timing jitter. Electronic circuits do not operate on timescales shorter than several picoseconds. While the concept of the streak camera circumvents this problem almost entirely, it still requires a periodic electronic trigger signal at the repetition rate of the laser. Timing instabilities in this signal limit our time resolution to a few picoseconds.

Limited Angles   Another problem is that by using whatever scene geometry is available to us we may not be able to sample every incoming and outgoing angle of a material and the reconstructed BRDFs will consequently be incomplete. This problem can be minimized by using every laser and camera angle that is available to us. Nevertheless we may only reconstruct a subset of the set of all BRDF values for all angles and all materials in the scene. It is important to not that the BRDF portions required to render the scene from the perspective of the camera and light source are contained in that collected subset.

Dynamic Range   When collecting specular materials the brightness of the specular peak can hinder a detection of the remaining BRDF due to limited dynamic range. The diffuse light reflected off a mirror for example will be overpowered by the specular peak of the mirror unless a laser and camera position can be found where the specular peak is outside the captured image.

Color Since the laser operates at a single wavelength our images are monochrome and taken in the near infrared. Colored images could be taken with a white light super continuum source or a set of 3 lasers at different colors.

5.3 Prior art in multibounce reflectance capture

Nicodemus et al. [41] introduced the Bidirectional Reflectance Distribution Function (BRDF), which characterizes light reflection from surfaces. BRDF acquisition has received much attention with efforts to acquire isotropic and anisotropic BRDFs, spatially varying BRDFs and BTFs [54, 35, 5, 29, 30, 36, 40, 27]. These acquisition processes are often lengthy requiring extensive laboratory settings with calibration, and hours to days of acquisition. [56] presents a recent survey on the state of the art.

Capturing a full 6D BRDF (for spatially varying), 4D BRDF, or 3D BRDF (for isotropic materials) requires measuring a large number of samples of the material. To decrease acquisition costs, many techniques focus on matching to simple parametric models. (e.g., [58, 37, 29, 30, 15]). Recently, various techniques decrease acquisition costs by exploiting various properties of BRDFs like reciprocity, separability, spatial smoothness, and compressibility [53, 7, 59, 48, 14, 12]. Various lighting configurations and variations have been considered including polarization [33, 13], structured illumination [55, 48].

Most BRDF techniques directly view and image the material sample and image it, with some notable exceptions [16, 26, 18]. We indirectly view the sample through a diffuse reflector. Hawkins [18] developed the dual light stage and image a diffuse environment to increase angular measurements.

The goal of this work is BRDF recovery from the measured light transport. Nayar et al [39] proposed an iterative photometric stereo algorithm, which is the first work that uses interreflections as a useful cue for shape and reflectance estimation. Recently there is a large interest in recovering scene properties such as geometry and albedo from multiple bounces of light [38, 46, 47, 42, 31], and also on recovering shape and material recovery
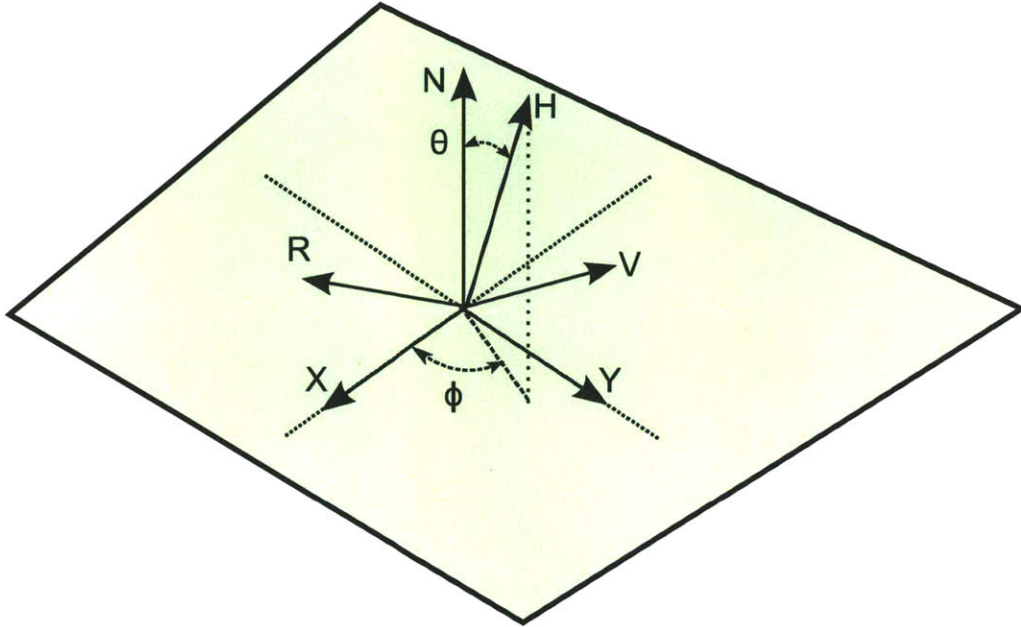
Figure 5-1: Geometry of reflection: Half vector H has spherical coordinates (θ,ϕ) and lies in the plane defined by incident vector R and outgoing vector V. Ref. Ashikhmin dBRDF.

simultaneously [20].

Several approaches use parametric BRDF models. Gardner[11], the surface is scanned with a linear light source and captured from a fixed view, fit to an isotropic BRDF model. Lensch et al.[29, 30] reconstruct SVBRDFs using clustering and a known parametric model (Lafortune) with linear combinations of the BRDFs at each point. Goldman et al [15] use image-based techniques to acquire low-dimensional isotropic Ward model.

Our work is inspired from the previous works in multi bounce and inversion methods. [38, 46, 47].

	Slow Camera	1D time-camera	2D time-camera
Single Point	Direct lookup, 4D BRDF	3D BRDF	4D BRDF
Homogeneous Patch	Goode cone coverage but blurred response	Coverage and sharp recovery	
Segmented Reflectance	Ill-conditioned		
Status	Using simple camera	High Speed ToF camera	Simulation Only
Challenges	Long exposure	Low SNR	Slow inversion

Figure 5-2: Exploiting interreflections and time domain. While a traditional still camera can recover single point BRDF, 1D or 2D time camera can recover large cones with sufficient high frequency details.

5.4 Multi-bounce reflectance capture

5.4.1 Geometry of Acquisition

We propose a simple geometric setup for the acquisition of BRDFs. Figure 5-4 shows the configuration in our hardware device. It includes a laser that illuminates points on the source S. The source S and receiver R are both assumed to be known Lambertian materials. In our equations we will use the index i for surface S using i, j for the receiver R, and k for the patch P. A camera indirectly views the surface R.

Our goal is to estimate the reflectance of patches on surface P. We shine light on each of the points on surface S and capture the photon arrivals (a) using a still camera (Section 5.4.2), or (b) using a fast (time-of-flight camera) (Section 5.5), and use this information to estimate the reflectances of patches P.

Note that in comparison with other BRDF approaches, we indirectly observe BRDF mea-
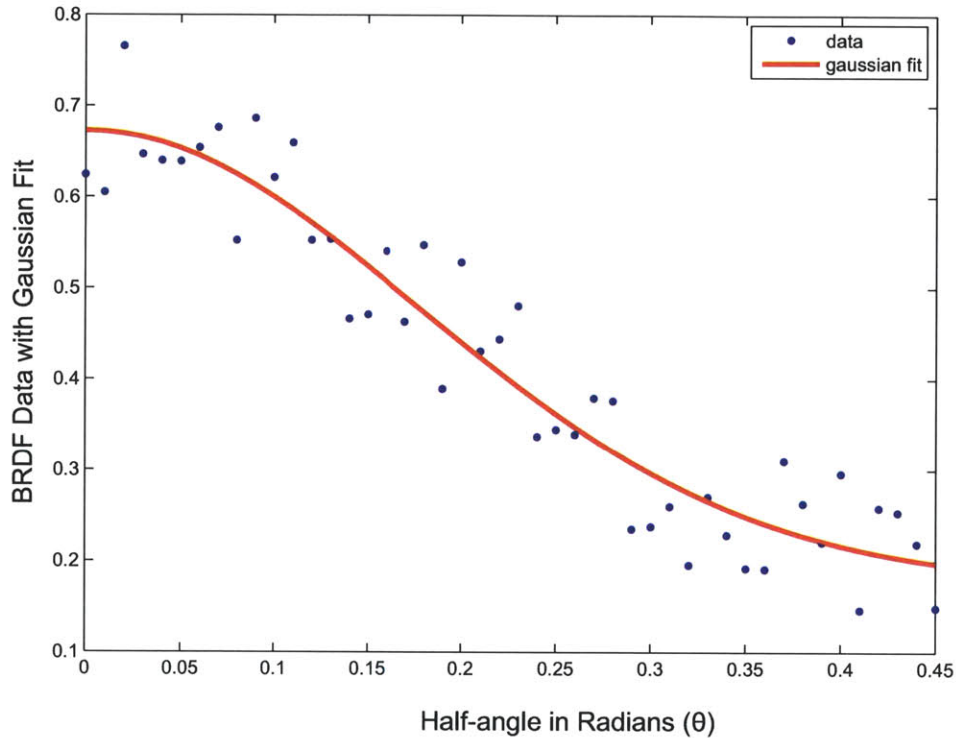
Figure 5-3: We solve for multiplexed coefficients and the fit for the parametric model in a single step. Unlike current methods that first estimate subsampled values of coefficients and then fit a parametric model, our single step approach exploits the reflected intensities. The figure shows a Gaussian fit using a three parametric model applied to observed data for an 1-D segmented BRDF.

surements by imaging a Lambertian material R. This enables taking many BRDF measurements simultaneously thus greatly accelerating acquisition. However, this approach is only practicable with the time of flight camera which is able to disambiguate between measurements across time, thus enabling accurate reconstruction.

5.4.2 Reflectance estimation using a still camera

In this section, we explain how reflectance estimation works in our hardware setup using a still camera and projector. We use the phrase still camera to refer to a 'regular camera'

89

operating at 30 frames per second, in contrast to our fast 'time-of-flight' camera that can run at half-a-trillion frames per second, and can capture nanosecond images.

Consider the setup in Figure 5-4. Assume that there are I points on surface S, J points on surface R, and K points on surface P.

Patch-based BRDF: Consider the light from point i reflecting off k on P, and arriving at the receiver point j on R. Imaging the receiver R gives us measurements of:

$$\sum_k W f_r(i,k,j) \frac{cos(\omega_{ik})cos(\omega_{kj})}{(d_{ik}d_{kj}*)^2} \tag{5.1}$$

where, $f_r(i,k,j)$ is the BRDF at point k with incoming angle defined by the vector ik (ω_{ik}) and outgoing angle (ω_{kj}) (see Figure 5-4), d_{ik}, and d_{kj} are the distances between i and k and j and k respectively, and W is a weighting factor that includes the intensity from the laser modulated by the known reflectances of S and R and other known geometry terms related to the location of the laser and eye wrt S and R respectively.

As the laser sweeps over a 2D range of i values over S, we measure a range of values over a 2D range of j values on R. Note that given the geometry, there are restrictions on the range of incoming and outgoing directions that are measurable. One issue that arises is that the light at j could also arrive directly without bouncing off K directly from i; this is eliminated in our experimental setup by including an occluder that blocks light from traveling directly to R from S.

When the patch area P is finite, reflectances are blurred by a convolution kernel based on the patch area as shown in Figure 5-8. But we can still compute f_r values in the following way. Assume for now a 1D still camera reading a line on the surface R. For each i, this line is represented as $C(i, j_1...j_J)$. The measurements above can be formulated as a linear system of equations as below:

$$\mathbf{B} = \mathbf{GF} + \mathbf{n} \tag{5.2}$$

where, B is the vectorized representation of the 1D images over all i and j, F is a vectorized representation of the unknown BRDF values $f_r(i,k,j)$, and G is the matrix representing
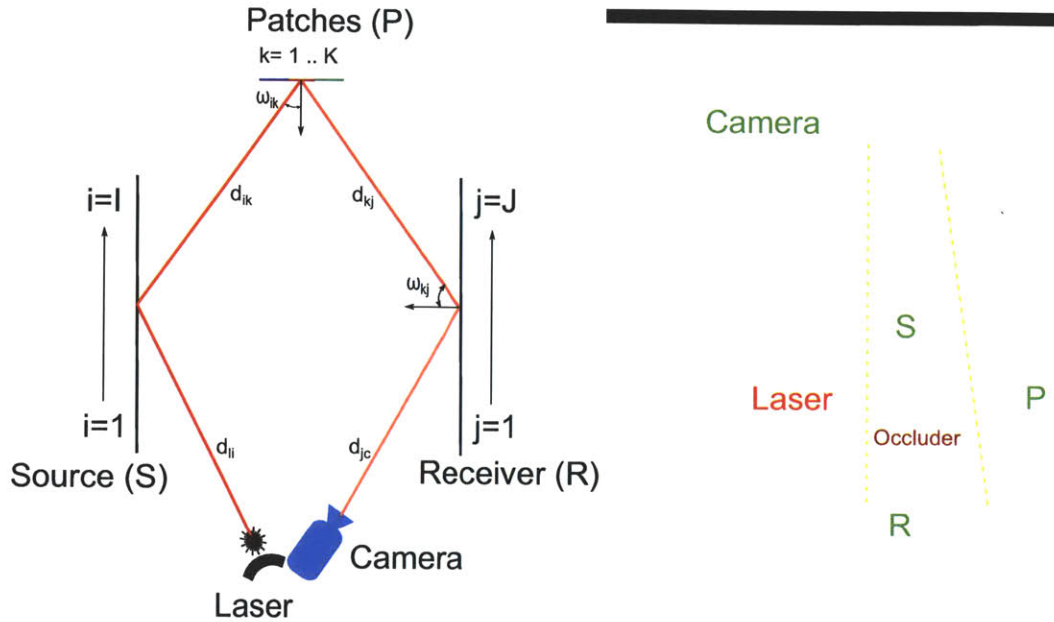
Figure 5-4: We show that scattering can be employed to recover limited cone BRDF of object patches. We record the intensities captured by a receiver R at positions ranging from $j=1$ to J for different positions of the coherent light source on the source S at positions from $i=1$ to $i=I$. The source, object patches and receiver are arranged in an U-shaped configuration. Both the source and receiver have completely diffuse BRDFs. Limited cones of reflectance of the object patches,denoted by k=1 to k=K, are obtained from the observed intensities. (left) The still camera configuration (right) Photo of an actual setup

the physical geometry terms including the cosine and distance terms from Equation 5.1. n represents the noise from camera capture. The matrix G is sparse. Figure 5-5 shows the visual representation of this linear system.

Each i, j combination provides slightly different blurred values of the BRDF, and it should be possible to solve for F. But SNR limits the recovery of the BRDF. Instead we assume a low dimensional parametric model of the BRDF and recover the parameters of this BRDF. We use the half-angle parametrization proposed by Rusinkiewicz [43], and use the BRDF proposed in [1], and used in [14] to measure distributions of the BRDF. Ashikhmin et al. show that using such a fitting process for limited cone data can be effective. We compute
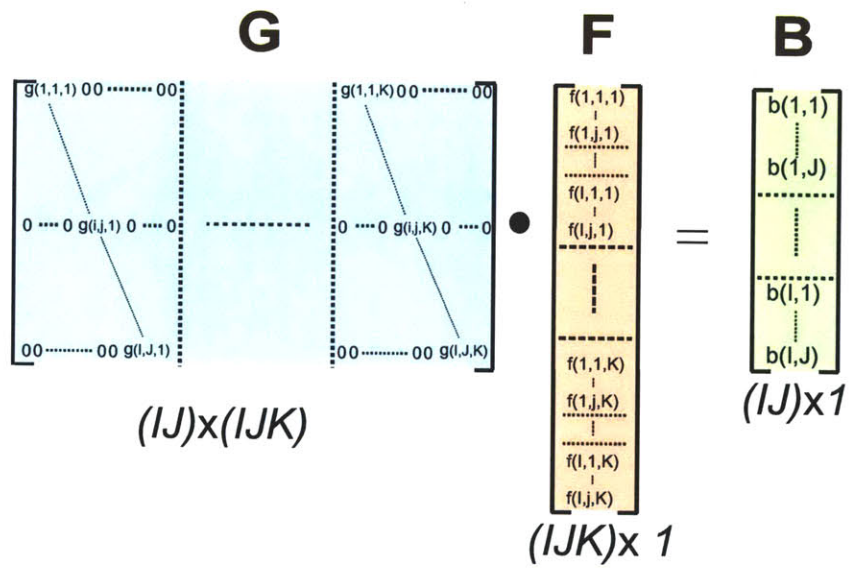
Figure 5-5: The limited cone reflectance values of object patches are obtained from a linear system representation made-up of the BRDF being measured and form factors. The figure shows the linear system for a still camera setup - G includes the form factors, F is the BRDF vector arranged by half-angle values. The resultant vector B is made up of observed radiance values.
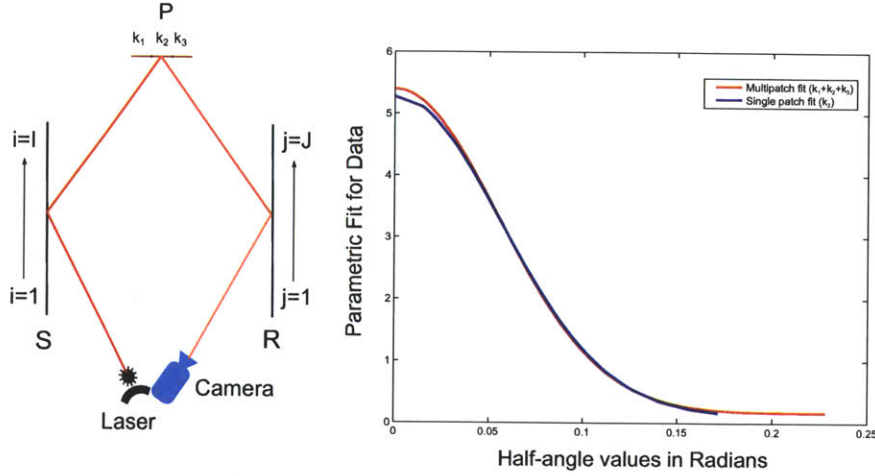
Figure 5-6: For a homogeneous material, the parametric fit obtained from a 'multi patch' made of K homogeneous patches is a combination of parametric fits from identical individual curves in slightly shifted angular ranges. The figure shows a parametric fit inverted from a linear system made up from a combination of three planar patches and the parametric fit inverted from a linear system created when only the central patch is present. The two fits may not match exactly due to the noise added to B in the capture process.

the half angle $h(i, j, k)$ for each measurement and parametrize the BRDF as:

$$f(i, j, k) = \frac{k_d}{\pi} + k_s p(h(i, j, k)) \tag{5.3}$$

where the unknowns k_d, k_s are the diffuse and specular reflectance respectively, and $p(h)$ is the distribution parametrized by the half angle vector. Various distributions $p(h)$ have been published in graphics literature [40, 36, 2]. Since our measurements have relatively limited cones of angles, we assume isotropic BRDFs and fit a single Gaussian lobe to the distribution as done in [52, 7]. In the future we could use their technique of NDF synthesis to fill holes if needed. Thus our BRDF estimation problem reduces to estimating $3K$ unknowns. In cases where P is homogeneous we can recover these parameters as shown in Figure 5-5.

5.4.3 Heterogeneous Multi patch BRDF

In the case where each of the K patches on P has a different BRDF, the number of parametric unknowns is $3K$, while the number of measurements available is still the same as

Copper	Acrylic	Paper
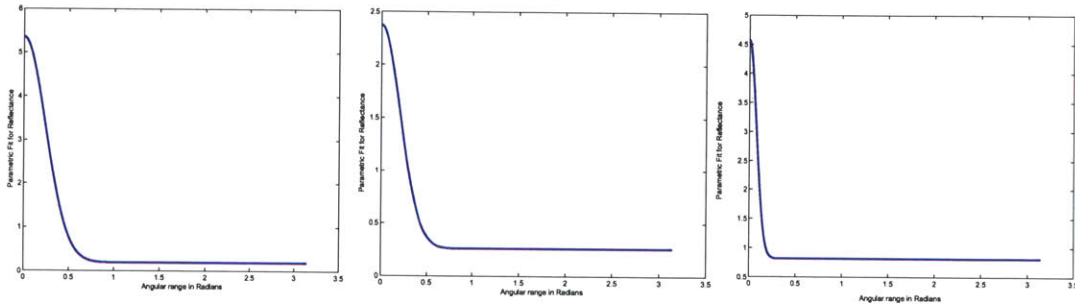
Figure 5-7: We obtain parametric fits for three different materials using the still camera setup and render the results to compare them with each other. The sub figures show the rendered results and 1-D plots of the parametric model against half-angle. (a) Rendered result for Copper and its parametric fit. (b)Rendered result for Red Acrylic and its parametric fit. (c) Rendered result for Paper and its parametric fit.

before: $I \times J$. Each observed value $b(i,j)$ is a linear combination of K values given by:

$$b(i,j) = g(i,j,1) * f(i,j,1) + ... + g(i,j,K) * f(i,j,K) \qquad (5.4)$$

It is impossible to solve for $f(i,j,k)$ for each patch individually as this is the only observation available for this half-angle combination for the K patches. Even in the parametrized space, we now have to solve for $3K$ variables (as compared to 3 in the case of a homogeneous patch). The system becomes under determined in this case as the values in F are a linear combination of K unknown BRDFs in different angular cones, while the number of observations remain the same. This can be seen in Figure 5-8.

Thus, there is a need to disambiguate the mixed combinations of P arriving at each point of R. This motivates the major contribution of our paper. Our time of flight camera allows

94

us to separate these integrated BRDF values by acquiring measurements in nanosecond time frames. This extra dimension of measurement (time) lets us separate out BRDF measurements for fast acquisition.
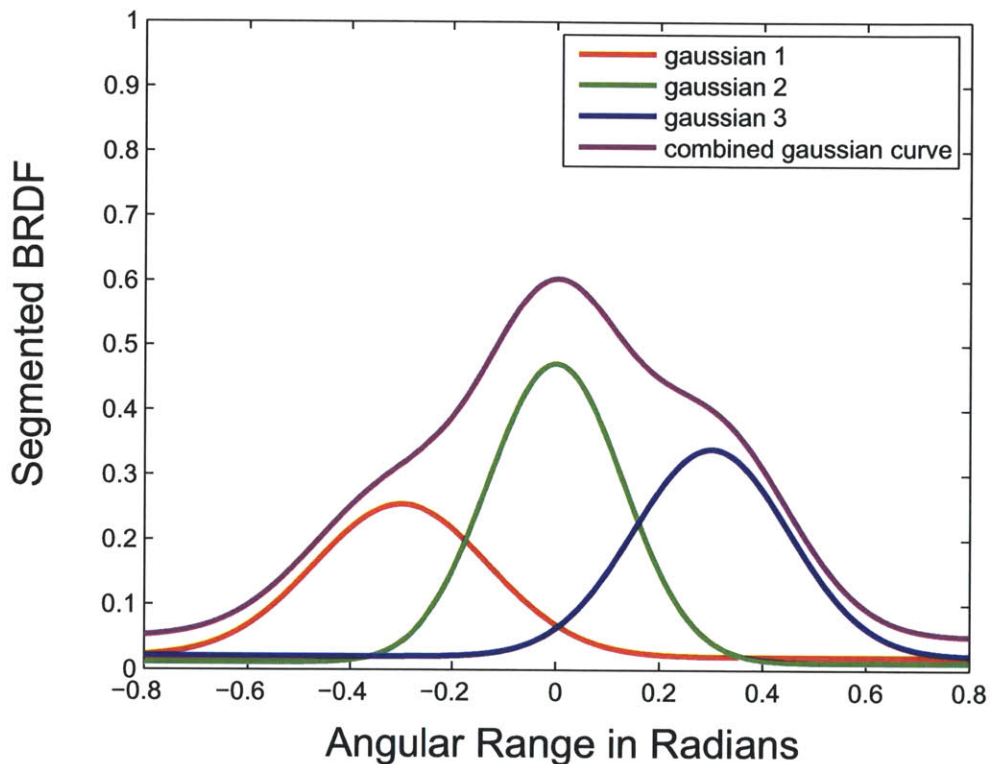
Figure 5-8: For a still camera, we obtain a linear combination of reflectance curves of heterogeneous patches, from which the individual curves cannot be recovered as number of observations remain the same as in case of multiple homogeneous patches.

5.5 Time-of-flight Multi-Bounce BRDF

In this section we describe how the time-of-flight camera addresses the disambiguation problem and enables rapid acquisition of materials.
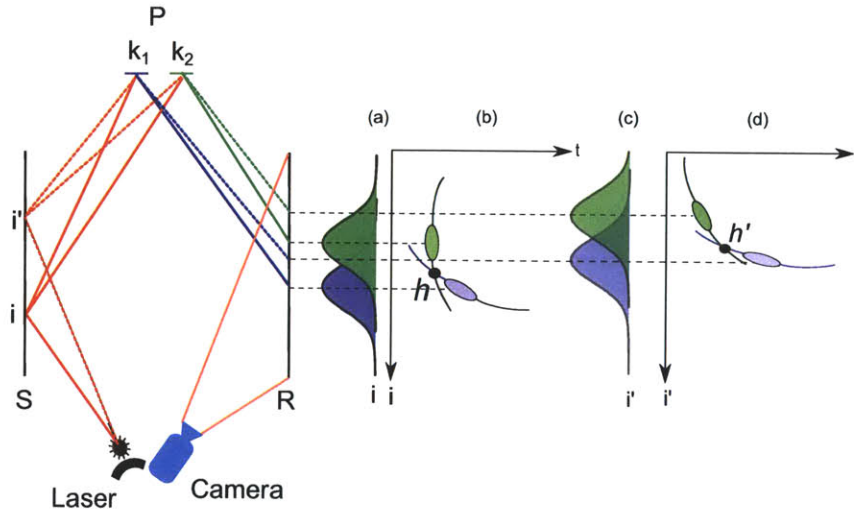
Figure 5-9: The time-of-flight camera disambiguates between multi-reflections from closely placed patches with heterogeneous BRDFs. (a) In the still camera, the two BRDFs, depicted by blue and green curves respectively, get entangled with each other for source position i. This entanglement does not change by a significant amount as the source moves to some other position i' as seen in (c). Hence disentanglement of heterogeneous BRDFs is ill-posed. (b) In contrast, in case of the ToF camera, the two BRDFs are completely disambiguated corresponding to source position i, except for a few points such as h, which have the same path length in the given geometry. However, a different position i' shown in (d), helps to disambiguate the information at point h as it is not entangled any more. (It may still have entanglement at some other point h'.) The whole limited cone BRDF is recovered using a few source positions.

5.5.1 Geometry of Acquisition

The geometry of the setup used for the time-of-flight (ToF) camera is the same U-shape setup employed by the still camera. We use identical steps for image capture using multiple incident laser positions, where the laser scans over a 2D surface S. The fast laser is able to switch rapidly enabling fast scanning of S. The key difference is the use of the time-of-flight camera as the imaging device. This camera can measure light at femto second time intervals; thus light traversing different path lengths arrive at different instants of time that are measureable at the imaging device. The speed of light is $3 \times 10^8 m/second$, therefore the distance traveled in 1 femto second is 0.3 micrometers.

Consider a pixel j on the receiver R: as light from the laser takes different paths reflecting
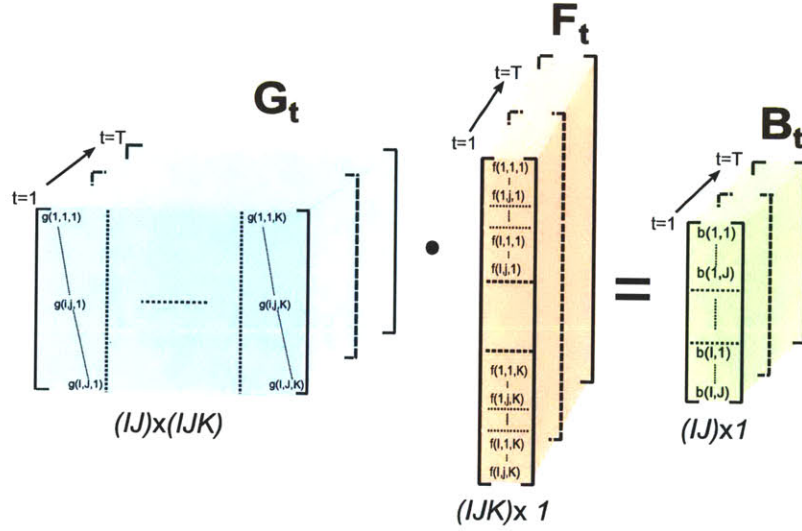
Figure 5-10: The linear system for a time-of-flight camera gets extended along the time dimension as observed intensities at different time instances are separated. Each linear system consisting of $g(i, j, k)$, the fall-off matrix, the BRDF vector $x(i, j, k)$ and observed intensity vector $b(i, j)$ is created at each time instance and solved for the reflectance vector of the patch contributing at that time instance.

off different points k on P, it will arrive at different times at the receiver. For now our camera is able to capture images with 1 spatial dimension (and the temporal dimension). We call these images "streak images". See Figure 5-11 for an example streak image for a small patch P. In comparing with the still camera image we note that if we add all the data across time for a single pixel j and a single source location i we obtain the still camera image value for that (i, j) combination. Also note that with the time-of-flight camera it is possible to observe the specular peak directly from the streak image.

Note that a pixel j on the receiver R at a time t measures the following:

$$\sum_{k \in K'} f_r(i, k, j) \frac{cos(\omega_{ik})cos(\omega_{kj})}{(d_{ik}d_{kj}*)^2} \tag{5.5}$$

where, $K' \subset K$ consists of all points k such that the path length d for the laser light to arrive at j is the same. That is, at a given instant t light arrives at j from all points k
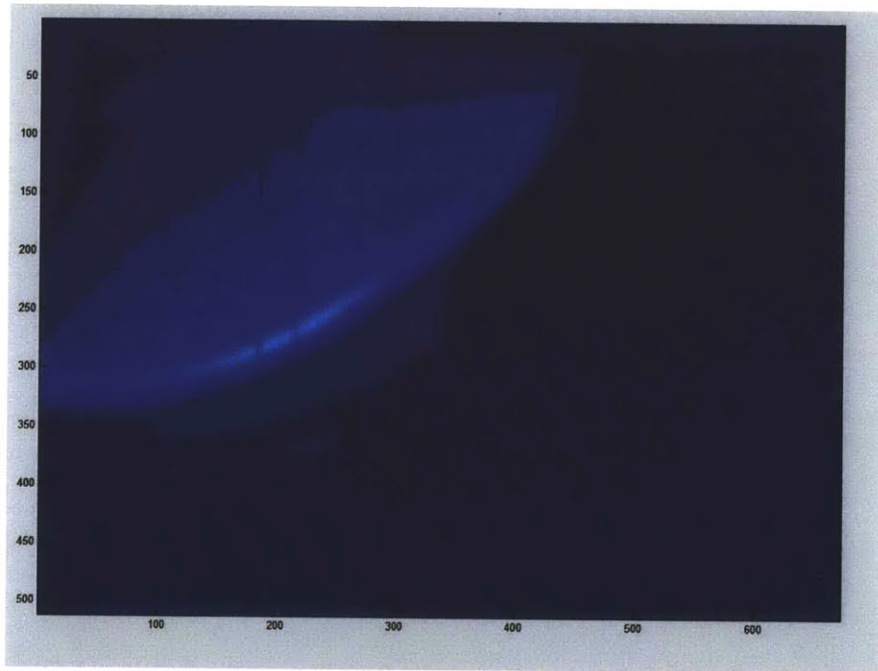
Figure 5-11: A sample streak image from the streak camera. The x axis represents a line in space, and y axis represents the time of arrival in pico seconds. Note: The camera by default has time axis starting from bottom up, so the time images should be read from bottom up. These streak images are received after three bounces of a femto second laser pulse in the shown experimental setup.

which have an equal path length along $i \to k \to j$:

$$d = d_{ik} + d_{kj} = Ct \tag{5.6}$$

where C is some constant.

Figure 5-12: We obtain parametric fits for three different materials using the fast camera setup and render the results to compare them with each other. The figures show the rendered results and 1-D plots of the parametric model against half-angle. (a) Rendered result for Copper and its parametric fit. (b)Rendered result for Red Acrylic and its parametric fit. (c) Rendered result for Paper and its parametric fit.

5.5.2   Inverse System for Time of Flight Images

Consider a case where the surface P is a 'multi patch' made up of K individual patches. Each point on the receiver R receives light from different patches at at different time instances. Thus we can disambiguate between light paths by making use of the time information in the streak images.

Since we have a time-of-flight camera, it is possible for us to obtain the entire geometric configuration of the scene. Given the full geometry of the scene, we can compute the exact time slots at which light will arrive from each individual patch at pixel j. Using these time slots, we extract the streaks corresponding to each patch from the streak images for a given laser position. Using these individual streaks from each laser position, we populate the B vector and the geometry information is again used to calculate the appropriate fall-off

99

factors in the G matrix. Thus our linear system for the still camera gets stretched along the time dimension. This can be seen in Figure 5-10.

As before, using a parametric model for the reflectance of the patches, we can solve the linear system of equations; though in this case each time slot can be treated separately. This is essentially equivalent to placing each of these patches individually at their positions and capturing the data using the still camera. This process requires capturing I images for each patch. So a total of $I \times K$ images need to be captured for K patches. In contrast, for a ToF camera only I 2-D streak images are used for extracting the reflectance curves of K patches. This reduces the whole acquisition process time to a few nanoseconds.

5.5.3 Disambiguation using time

The streak images obtained for different patches are not always completely separated. In the case of where the two light paths arriving at the same point have the same path length, there will be a linear mixing observed at that point in the streak image. However if the dimensions of the 'multi patch' are not comparable with those of the source and receiver walls, there will be only a few such entanglements. This will mean that there will be entangled observations for a few half-angle values for each individual patch. However as we have large number of observations and only three unknowns per patch, we are able to obtain fits for the parametric model of each patch using a few source positions I.

In fact, more generally, for many half angle values where entanglement does not occur, it is possible to use the streak images to directly measure out the BRDF directly without having to resort to fitting to a parametric representation.

Figure 5-13: The setup involves a high speed ToF camera and a pico second accurate camera.

5.6 Experiments

5.6.1 Ultra-Fast Lasers and Detectors

Since their initial conception in 1960 lasers have become rapidly cheaper, smaller and more robust. While the first lasers tended to be large room filling devices or could only operate at small powers and for fractions of a second to avoid overheating, today's diode lasers deliver superior performance with microscopic sizes at the cost of less than a dollar and driven by batteries. Wall plug efficiencies of continuously operating laser diodes can be above 60%. Blue lasers, for example where first demonstrated as dye lasers in the early 70s. The first practical blue diode laser was built in 1996 and only seven years later the first blue laser was available in a commercial product

There is no fundamental reason for this low efficiency and new technologies will continuously improve it.

The ultra-fast detection of light - while more challenging than the generation of ultra-fast pulses - adheres to the same basic reasoning. The fundamental problem here is that ultimately the information is processed in the form of electronic signals which are too slow to represent information on pico second or femto second timescales. Optical gathering, processing, amplification and storage of the information becomes necessary. Developing

methods to do this is one of the primary goals of today's research in computer processors, telecommunication and integrated optics. Significant and rapid progress is being made in these areas and will benefit efforts in transient imaging.

5.6.2 Experimental Setup

Our light source is a mode-locked Ti:Sapphire laser. It delivers pulses of about 50 femto seconds length at a repetition rate of 75 MHz. The laser wavelength is centered at 795 nm. A small portion of the laser beam is split off with a glass plate and is used to synchronize the laser and streak camera. Another portion is split off the main beam and after attenuation is directed to the scene to serve as a timing reference. The main beam is sent to the scene via a system of two galvanometer steered mirrors. It can be directed to scan most of the visible scene.

The camera is a Hamamatsu C5680 streak camera that captures a one dimensional image, .i e. a line in the scene, with a time resolution of about 10 picoseconds and a quantum efficiency of about 10%. The position and viewing direction of the camera is fixed. The streak camera images incoming light in a screen where photons excite electrons that are accelerated and deviated in a vacuum tube similar to a CRT TV. The time dependent deviation of the one dimensional image introduces the time resolution. On the far end the electrons hit a phosphor and are converted back to light and recorded by a slow CCD readout camera.

5.7 Multi bounce reflectance estimation: Conclusion

In this work, we have demonstrated a device setup for High speed photography for segmented scene reflectance capture using indirect reflections. We have also elaborated the underlying theory behind showing how the time of flight capture converts an under constrained problem into an invertible system.

We believe that this approach of BRDF acquisition has potential to enable fast, remote BRDF capture devices. Ability to capture BRDF in a few nanoseconds and without having the instrument the environment like in traditional BRDF capturing techniques can spur more applications like real-time material editing[23], where a person can take photographs of a wooden material and see how his sofa will look like if he used that material.

While fast and smaller solid state lasers are coming, merging them with fast imaging devices is a clear logical step.

This approach sets stage for reducing the time between BRDF acquisition by bringing in speed and portability. Our approach can also be used when capturing complete BRDF is not the ultimate goal but sampling a part of it for material detection with sonar wavelengths can suffice.

Chapter 6

Overview, Applications and Conclusions: NLOS Computer Vision

6.1 Applications

6.1.1 NLOS Motion tracking

The ability to track moving objects that cannot be seen can enable applications into following broad areas:

1. Intelligent Transport: A car taking a blind turn can get alerts for vehicles speeding around the corner. An augmented reality dashboard can enable drivers have a better idea of things that cannot be seen. An advanced futuristic version of multi path analysis based light transport can enable looking through haze also.

2. Medical- internal body imaging with scopes: Medical imaging surgeries that involve probing internal body can use such a multi path light transport based approach to
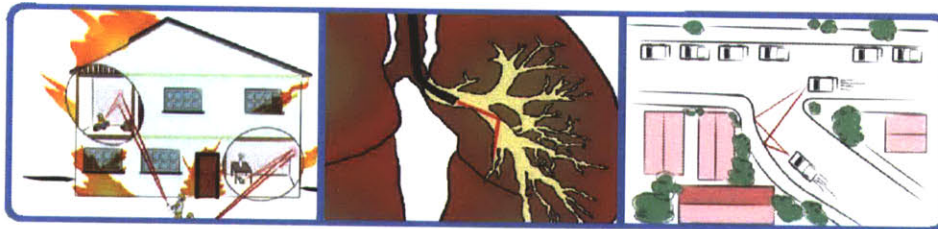
Estimating motion of NLOS moving objects in an unknown clutter

Location   **Size**   **Velocity**

Applications

Search & rescue   **Medicine**   **Transport**

Figure 6-1: Enabled to track location size and velocity of NLOS objects, one can deploy applications like remote search and rescue, medical imaging and intelligent transport

extend the reach of these probes. Surgeries like bronchioscopies, colonoscopies, laparoscopies etc. can use the NLOS tracking technique that can let surgeons look further and deeper. Consider the constraints on diagnostic endoscopy. Great progress in imaging hardware has allowed a gradual shift from rigid to flexible to digital endoscopes. Digital scopes put image sensors directly at the tip of the scopes. However, there is a natural limit to their reach due to constraints in the dimensions of human body that leave very little room for guiding the imager assemblies. Making imagers smaller is challenging due to the diffraction limits posed on the optics as well as due to sensor-noise limits on the sensor pixel size. In many scenarios, we want to avoid the maze of cavities and serial traversal for examination. We want the precise location and size of a lesion when deciding for or against application of limited or extended surgical procedures. Ideally we should be to explore multitude of paths in a simultaneous and parallel fashion. We use transient imaging to mathematically invert the data available in light reflected in complex optical reflections. We can convert elegant optical and mathematical insights into unique medical tools

3. Espionage: Probing into enemy bunkers from longer distance in an intrusive manner is possible when this technology is used in real life cases like espionage. One can track a criminal hiding in an enemy bunker with a small duration laser scan. Practically a few seconds of suck a laser pulsed scan is sufficient to track objects inside the area.

4. Reconnaissance and rescue: A house on fir or a nuclear plant under radiation alert can be probed using this technology. Moreover, looking through thick forests, waterfalls, caves is also possible when used for long range scenarios.

6.1.2 Fast, Touch-free, into-the-wild reflectance estimation

1. Reconnaissance: It is possible to judge what type of material exists inside a hidden environment by making use of the NLOS reflectance capture technology. Imagine being able to detect circuits/ bombs in hidden environment by tracing particular materials or being able to trace a human existence in a cave by probing of the cave by a speleologist.

2. Point of sale information to customer (fabrics, furniture, metals): Imagine a scenario where someone buying furniture from Ikea store can be given a real time make over of the look of his house or car upholstery by using fast capture of material reflectance and replicating it using fast rendering. Fast real time and into the wild touch free reflectance field capture can have strong influence on the retail industry.

3. Changes in material mapping and new environment visualization: Architects, builders and military personnel can map newer environments by a fast mapping of the environment in existence as well as visualization of the new scenario in presence of adding a construction of a sample material. For example, an architect can visualize how will a particular metallic finish look go in the building facade given the other materials and lighting in the area.

6.2 Overview and conclusions

Ours is the first solution to compute parameters of a non-line of sight object in a highly cluttered environment. In eye-safe near-visible wavelength spectrum, there is currently no other optical approach (in computer vision or otherwise) to address arbitrarily shaped NLOS objects among cluttered scenes. Even beyond visible spectrum, currently there is no practical mechanism to see through dense occluders (e.g. a concrete wall or a metal plate) using an image forming directional wavelength. Earlier work in exploiting multi path analysis for NLOS scenes was limited to probing a few points to recover geometry of a simplistic planar object in uncluttered environments. Our novel virtual phased array based approach makes the estimation robust by exploiting an array of points. In this paper, we proposed a new method and also demonstrated it by building a working physical prototype.

Novel imaging systems often inspire radical advances in to the development of novel algorithms and frameworks for visual understanding. We believe that time-of-flight cameras that can measure higher dimensional light transport provide a unique opportunity for computer vision to solve seemingly impossible problems such as tracking a non-line of sight

object. We hope this work inspires future work in other areas of computer vision such as NLOS recognition, NLOS segmentation and data-driven machine learning approaches.

6.3 Funding acknowledgements

Bibliography

[1] Michael Ashikhmin. Distribution-based brdfs. *Technical Report Stony Brook*, 2005.

[2] Michael Ashikhmin, Simon Premože, and Peter Shirley. A microfacet-based BRDF generator. *SIGGRAPH*, Jul 2000.

[3] J. Benesty. Adaptive eigenvalue decomposition algorithm for passive acoustic source localization. *The Journal of the Acoustical Society of America*, 107:384, 2000.

[4] J Busck and H Heiselberg. Gated viewing and high-accuracy three-dimensional laser radar. *Applied optics*, 43(24):4705–4710, 2004.

[5] Kristin Dana, Bram van Ginneken, Shree Nayar, and Jan Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics*, 18(1):1–34, 1999.

[6] F Dellaert. Monte carlo em for data-association and its applications in computer vision. 2001.

[7] Yue Dong, Jiaping Wang, Xin Tong, John Snyder, Yanxiang Lan, Moshe Ben-Ezra, and Baining Guo. Manifold bootstrapping for svbrdf capture. In *ACM SIGGRAPH 2010 papers*, SIGGRAPH '10, pages 98:1–98:10, New York, NY, USA, 2010. ACM.

[8] YKN Eddie. Is thermal scanner losing its bite in mass screening of fever due to sars? *Med. Phys.*, 32:93, 2005.

[9] B Ferris, D Fox, and N Lawrence. Wifi-slam using gaussian process latent variable models. *Proceedings of IJCAI 2007*, Jan 2007.

[10] Dr Takashi Fujii and Tetsuo Fukuchi. Laser remote sensing. page 888, Jan 2005.

[11] Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. Linear light source reflectometry. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):749–758, 2003.

[12] Gaurav Garg, Eino-Ville Talvala, Marc Levoy, and Hendrik P. A. Lensch. Symmetric photography: Exploiting data-sparseness in reflectance fields. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering*, pages 251–262, June 2006.

[13] Abhijeet Ghosh, Tongbo Chen, Pieter Peers, Cyrus A. Wilson, and Paul Debevec. Circularly polarized spherical illumination reflectometry. In *ACM SIGGRAPH Asia 2010 papers*, SIGGRAPH ASIA '10, pages 162:1–162:12, New York, NY, USA, 2010. ACM.

[14] Abhijeet Ghosh, Wolfgang Heidrich, Shruthi Achutha, and Matthew O'Toole. A basis illumination approach to brdf measurement. *Int. J. Comput. Vision*, 90:183–197, November 2010.

[15] Dan B Goldman, Brian Curless, Aaron Hertzmann, and Steven M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. In *IEEE International Conference on Computer Vision*, 2005.

[16] Jefferson Y. Han and Ken Perlin. Measuring bidirectional texture reflectance with a kaleidoscope. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):741–748, 2003.

[17] Richard Hartley and Andrew Zisserman. Multiple view geometry in computer vision. page 655, Jan 2003.

[18] Tim Hawkins, Per Einarsson, and Paul E. Debevec. A dual light stage. In Kavita Bala and Phil Dutre, editors, *Rendering Techniques*, pages 91–98, 2005.

[19] Richard P. Hodges. Underwater acoustics: Analysis, design and performance of sonar. page 366, Jan 2010.

112

[20] Michael Holroyd, Jason Lawrence, Todd Zickler, and Greg Humphreys. A photometric approach for estimating normals and tangents. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia)*, 27(5), 2008.

[21] Y Huang and I Essa. Tracking multiple objects through occlusions. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2:1051–1058, 2005.

[22] N Joshi, S Avidan, W Matusik, and D Kriegman. Synthetic aperture tracking: tracking through occlusions. *computer.org*, Jan 2007.

[23] E.A. Khan, E. Reinhard, R.W. Fleming, and H.H. B
"ulthoff. Image-based material editing. *ACM Transactions on Graphics (TOG)*, 25(3):654–663, 2006.

[24] A Kirmani, T Hutchison, J Davis, and R Raskar. Looking around the corner using transient imaging. *ICCV*, 2009.

[25] C. Knapp and G. Carter. The generalized correlation method for estimation of time delay. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 24(4):320–327, 1976.

[26] Sujit Kuthirummal and Shree K. Nayar. Multiview radial catadioptric imaging for scene capture. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pages 916–923, New York, NY, USA, 2006. ACM.

[27] Jason Lawrence, Aner Ben-Artzi, Christopher DeCoro, Wojciech Matusik, Hanspeter Pfister, Ravi Ramamoorthi, and Szymon Rusinkiewicz. Inverse shade trees for non-parametric material representation and editing. *ACM Transactions on Graphics (SIGGRAPH 2006)*, 25(3), 2006.

[28] ND Lawrence and J Quiñonero-Candela. Local distance preservation in the gp-lvm through back constraints. *Proceedings of the 23rd international conference on Machine learning*, pages 513–520, 2006.

113

[29] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatially varying materials. In *Eurographics Workshop on Rendering*, pages 63–70, 2001.

[30] Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. Image-based reconstruction of spatial appearance and geometric detail. *ACM Transactions on Graphics*, 22(2):234–257, 2003.

[31] S Liu, TT Ng, and Y Matsushita. Shape from second-bounce of light transport. *Computer Vision–ECCV 2010*, pages 280–293, 2010.

[32] BD Lucas and T Kanade. An iterative image registration technique with an application to stereo vision. *International joint conference on artificial intelligence*, 3:674–679, 1981.

[33] Wan-Chun Ma, Tim Hawkins, Charles-Felix Chabert, Mark Bolas, Pieter Peers, and Paul Debevec. A system for high-resolution face scanning based on polarized spherical illumination. In *ACM SIGGRAPH 2007 sketches*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.

[34] R Mailloux. Phased array antenna handbook. *Boston*, Jan 1994.

[35] S. Marschner, S. Westin, E. Lafortune, K. Torrance, and D. Greenberg. Image-Based BRDF measurement including human skin. In *Eurographics Rendering Workshop 99*, pages 139–152, 1999.

[36] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):759–769, 2003.

[37] David McAllister. *A Generalized Surface Appearance Representation for Computer Graphics*. PhD thesis, UNC Chapel Hill, 2002.

[38] S Nayar, G Krishnan, M Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM SIGGRAPH 2006*, Jan 2006.

114

[39] S.K. Nayar, K. Ikeuchi, and T. Kanade. Shape from interreflections. *International Journal of Computer Vision*, 6(3):173–195, 1991.

[40] A Ngan, F Durand, and W Matusik. Experimental analysis of BRDF models. *EGSR*, Jan 2005.

[41] F. E. Nicodemus, J. C. Richmond, and J. J. Hsia. Geometrical considerations and reflectance. *National Bureau of Standards*, October 1977.

[42] R Raskar and J Davis. 5d time-light transport matrix: What can we reason about scene properties. *Internal memo*, 2007.

[43] Szymon Rusinkiewicz. A new change of variables for efficient BRDF representation. In *Eurographics Workshop on Rendering*, pages 11–22, 1998.

[44] D Salmond and B Birch. A particle filter for track-before-detect. *American Control Conference*, Jan 2002.

[45] D Scharstein and R Szeliski. High-accuracy stereo depth maps using structured light. *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, 1, 2003.

[46] S Seitz, Y Matsushita, and K Kutulakos. A theory of inverse light transport. *Computer Vision*, Jan 2005.

[47] P Sen, B Chen, G Garg, and S Marschner. Dual photography. *ACM Transactions on Graphics*, Jan 2005.

[48] Pradeep Sen, Billy Chen, Gaurav Garg, Stephen R. Marschner, Mark Horowitz, Marc Levoy, and Hendrik P. A. Lensch. Dual photography. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):745–755, 2005.

[49] J Shi and C Tomasi. Good features to track. *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, pages 593–600, 2002.

[50] Merrill I. Skolnik. Introduction to radar systems. page 772, Jan 2002.

[51] Richard Szeliski. Computer vision: Algorithms and applications. page 812, Jan 2010.

[52] Jiaping Wang, Peiran Ren, Minmin Gong, John Snyder, and Baining Guo. All-frequency rendering of dynamic, spatially-varying reflectance. In *ACM SIGGRAPH Asia 2009 papers*, SIGGRAPH Asia '09, pages 133:1–133:10, New York, NY, USA, 2009. ACM.

[53] Jiaping Wang, Shuang Zhao, Xin Tong, John Snyder, and Baining Guo. Modeling anisotropic surface reflectance with example-based microfacet synthesis. In *ACM SIGGRAPH 2008 papers*, SIGGRAPH '08, pages 41:1–41:9, New York, NY, USA, 2008. ACM.

[54] Gregory J. Ward. Measuring and modeling anisotropic reflection. In *SIGGRAPH*, pages 265–272, 1992.

[55] Andreas Wenger, Andrew Gardner, Chris Tchou, Jonas Unger, Tim Hawkins, and Paul Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. *ACM Transactions on Graphics (SIGGRAPH 2005)*, 24(3):756–764, 2005.

[56] Tim Weyrich, Jason Lawrence, Hendrik P. A. Lensch, Szymon Rusinkiewicz, and Todd Zickler. Principles of appearance acquisition and representation. *Foundations and Trends in Computer Graphics and Vision*, 4(2):75–191, 2009.

[57] R White and DA Forsyth. Combining cues: Shape from shading and texture. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2:1809–1816, 2006.

[58] Yizhou Yu, Paul Debevec, Jitendra Malik, and Tim Hawkins. Inverse global illumination: recovering reflectance models of real scenes from photographs. In *Computer Graphics (SIGGRAPH 1999)*, pages 215–224, 1999.

[59] Todd Zickler, Sebastian Enrique, Ravi Ramamoorthi, and Peter Belhumeur. Reflectance sharing: Image-based rendering from a sparse set of images. In *Proceedings of*

the Eurographics Symposium on Rendering, pages 253–264. Eurographics Association, 2005.