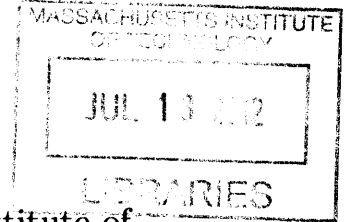


**Algorithms for Reconstruction of hidden 3D
shapes using diffused reflections**

ARCHIVES



by

Otkrist Gupta

B.Tech. in Computer Science and Engineering, Indian Institute of
Technology Delhi, 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 Arts and Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

© Massachusetts Institute of Technology 2012. All rights reserved.

Author
Program in Media Arts and Sciences,
School of Architecture and Planning
May 07, 2012

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

Accepted by
Mitchel Resnick
LEGO Papert Professor of Learning Research,
Academic Head,
Program in Media Arts and Sciences

Algorithms for Reconstruction of hidden 3D shapes using diffused reflections

by

Otkrist Gupta

Submitted to the Program in Media Arts and Sciences,
School of Architecture and Planning
on May 07, 2012, in partial fulfillment of the
requirements for the degree of
Master of Science in Media Arts and Sciences

Abstract

This thesis aims at discovering algorithms to recover the geometry of hidden objects from tertiary diffuse scattering, given time of flight information. We focus on using ultra high speed capture of photons to accurately determine information about distance light travelled and using it to infer hidden geometry. We aim at investigating issues such as the feasibility, uniqueness (in solution domain) and invertibility of this problem. We also aim at formulating the forward and inverse theory of secondary and tertiary diffuse scattering using ideas from tomography. We aim at developing tomography based approaches and sparsity based methods to recover 3D shapes of objects "around the corner". We analyze multi-bounce propagation of light in an unknown hidden volume and demonstrate that the reflected light contains sufficient information to recover the 3D structure of the hidden scene. We formulate the forward and inverse theory of secondary and tertiary scattering reflection using ideas from energy front propagation and tomography. We show that using careful choice of approximations, such as Fresnel approximation, greatly simplifies this problem and the inversion can be achieved via a backpropagation process. We provide a theoretical analysis of the invertibility, uniqueness and choices of space-time-angle dimensions using synthetic examples. We show that a 2D streak camera can be used to discover and reconstruct hidden geometry. Using a 1D high speed time of flight camera, we show that our method can be used to recover 3D shapes of objects "around the corner".

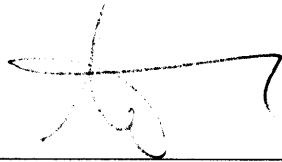
Thesis Supervisor: Ramesh Raskar
Title: Associate Professor of Media Arts and Sciences,
Program in Media Arts and Sciences

**Algorithms for Reconstruction of hidden 3D shapes using
diffused reflections**

by

Otkrist Gupta

The following person served as a reader for this thesis:



Reader: _____

Assistant Prof. Laurent Demanet
Assistant professor of Applied Mathematics
MIT Department of Mathematics

**Algorithms for Reconstruction of hidden 3D shapes using
diffused reflections**

by
Otkrist Gupta

The following person served as a reader for this thesis:

Reader: W. Michael Bove, Jr.

V. Michael Bove, Jr.
Principal Research Scientist
MIT Media Lab

Acknowledgements

I would like to thank my advisor, Prof. Ramesh Raskar, for his continued guidance, support, valuable insights and discussions crucial to solving this problem over the past year.

I would like to thank Dr. Thomas Willwacher for working with me on the mathematical formulations. Furthermore, I thank Dr. Andreas Velten, Prof. Mounji Bawendi and Prof. Ashok Veeraraghvan for their help with hardware and formulation of problem. I would also like to thank Dr. Christopher Barsi, Roarke Horstemeyer, Rohit Pandharkar, Andy Bardagjy and Nikhil Naik for their helpful comments. This work was funded by the Media Lab Consortium Members, DARPA through the DARPA YFA grant, and the Institute for Soldier Nanotechnologies and U.S. Army Research Office under contract W911NF-07-D-0004. A. Veeraraghavan was supported by the NSF IIS Award No. 1116718. We would like to thank Amy Fritz for her help with acquiring and testing data and Chinmaya Joshi for his help with the preparation of the figures.

Contents

1	Introduction	23
1.1	Motivation	23
1.2	Problem Description	24
1.3	Related Work	26
1.4	Contribution	28
2	Theoretical Foundations	33
2.1	Modelling Propagation of a Light Pulse for Multiple Bounces	33
2.1.1	Space-Time Warping for Bounce Reduction	35
2.1.2	Scattering of a pulse	36
2.1.3	Hyperbolic Contribution	38
2.2	Forward model: Elliptical Tomographic Projection	38
2.2.1	Elliptical tomography problem description	39
2.2.2	Challenges and missing cones	39
3	Inversion Analysis	43
3.1	Inverse Algorithm: Filtered Back Projection	43
3.1.1	Overview of the algorithm	43
3.1.2	Phase 2: Data Preprocessing	44
3.1.3	Phase 3: 3D Reconstruction	45
3.2	Inversion using parameter estimation	47
3.2.1	A remark about the filtering step	48

4	Sparsity Based Reconstruction	49
4.1	Linearizing The system	49
4.2	Methods	51
4.2.1	Sparse Formulation	52
4.2.2	Matching Pursuit Algorithms	52
4.2.3	Fixed Point Methods	53
4.2.4	Projective Gradient with Lasso	54
5	Experiments and Results	63
5.1	Hardware assembly	63
5.1.1	Results	67
5.1.2	Performance Evaluation	68
6	Future Directions	81
6.1	Conclusion	83

List of Figures

1-1	Plenoptic function parametrizes world pixels by location, wavelength and orientation. (Source: Wikipedia)	25
1-2	Setup (Left). Forward model (Center) The laser illuminates the surface S and each point $s \in S$ generates an energy front. The spherical energy front contributes to a hyperbola in the space-time streak photo, I_R . (Right) Spherical energy fronts propagating from a point create a hyperbolic space-time curve in streak photo.	26
1-3	Dual Photography: <i>Helmholtz Reciprocity</i> . (Please refer to [35]) . . .	28
1-4	TADAR : <i>millimeter wavelength imaging</i> . Source [2]	29
1-5	LIDAR <i>Time of flight of photons</i> .(Source [3])	30
2-1	Forward Model. (Left) The laser illuminates the surface S and each point $s \in S$ generates a energy front. The spherical energy front contributes to a hyperbola in the space-time streak photo, I_R . (Right) Spherical energy fronts propagating from a point create a hyperbolic space-time curve in streak photo.	33
2-2	A space time transform on a raw streak photo allows us to convert 4 segment problem into a sequence of 2 segment problems. The toy scene is a small $1\text{cm} \times 1\text{cm}$ patch creating a prominent (blurred) hyperbola in warped photo. Backpropagation creates low frequency residual but simple thresholding recovers the patch geometry.	35

2-3	<p>Reconstruction Algorithm An illustrative example of geometric reconstruction using streak images. (a) <i>Data capture</i>. The hidden object consists of a 2 cm × 2 cm square white patch. The captured streak images are displayed in the top row. (b) <i>Contributing voxels in Cartesian space</i>. For recovery of hidden position the possible locations in Cartesian space that could have contributed to the streak image pixels p, q, r are ellipsoids in 3D. If there is a single world point contributing intensity to all 3 pixels, the corresponding ellipses intersect. The white bar corresponds to 2 centimetres. (c) <i>Backprojection and heatmap</i>. We use a back-projection algorithm that finds overlaid ellipses corresponding to all pixels. (d) Backprojection using all pixels in a set of 59 streak images. (e) <i>Filtering</i>. After filtering with a second derivative, the patch location and 2 centimeter lateral size are recovered. Joint work with Andreas Velten, Thomas Willwacher, Ashok Veeraraghavan, Mounsi G. Bawendi and Ramesh Raskar [41].</p>	41
3-1	<p>The top left figure shows streak images being generated by near field sources. On bottom left we see effect when this sources travel farther away, The rightmost figure depicts how we can analytically predict single sources using multiple sensor laser combinations. Notice how the accuracy is affected if lasers shift.</p>	44
4-1	<p>How each world voxel contributes to a streak inside data. The entire streak image is composed by linear sum of all the streaks. If there is no occlusion we can turn the forward model into a linear system and can generate a corresponding matrix representation. Vectorized world voxels (denoted by c_i) on multiplication with linear transform matrix A yield streak image. The right hand is sum of streak images A_j corresponding to each voxel.</p>	50

4-2	Our approach can handle steep angles, sharp corners, concave as well as convex features.Reconstruction of (a) Wedge (b) S shape and a (c) Sphere using 8 streak images and COSaMP matching pursuit algorithm. Referenced from joint work with Ashok Veeraraghvan in [41]. .	51
4-3	Reconstruction of a spade using plain linear inversion techniques. The matrix A was generated using methods described in section 4.1. We multiplied both sides with A^T to make problem computationally feasible. We used pseudo inverse implementation in MATLAB to find the solution.	56
4-4	This figure shows how the reconstruction quality deteriorates by down sampling world voxels. (Left) Zero down-sampling (Middle) one third down-sampling (Right) One fifth down-sampling.	56
4-5	Above figure shows how spade evolves with every few iteration of SPGL. Notice that the interior points get picked first. The last image is result from 30 iterations.	57
4-6	Methods like SPGL requiring only single parameter have an optimal parameter value which results in correct reconstruction. The above figure shows the variation in error with variation in tau. This is in fact the pareto curve described in [6] between one norm of solution and two norm of residual.	57
4-7	This figure shows how spade reconstruction changes with increasing tau. For lesser values of tau only interior points get selected and rest of shape remains unknown. As tau is increased the solution converges to a spade.	58
4-8	. Anaysis of reconstruction if random noise is added to parameters. The reconstruction is pretty robust to random salt and pepper noise in laser locations.	58
4-9	Random Salt and pepper noise in homography and its effect on reconstruction.	59
4-10	Effect on reconstruction with systemic noise.	59

4-11	Variation in A with systemic noise. As we can see even small amount of homography error can change A significantly.	60
4-12	Effect on reconstruction with systemic noise in A	60
4-13	(Above) Hidden MIT around the corner (Below) Reconstruction using SPGL	61
5-1	The experimental setup. (B). The laser pulses strike the wall at a point l and some of the scattered light strikes the hidden object (e. g. at s), returns to the wall (w) and is collected by the camera (c). Both the galvo scanner and the camera are controlled by a computer. Done in collaboration with Andreas Velten, Thomas Willwacher, Ashok Veer- araghavan, Mounji G. Bawendi and Ramesh Raskar [41].	65
5-2	Resolution in depth. Left: Distance estimation. Time here is measured in mm of traveled distance at the speed of light $1\text{ mm}\approx 0.3\text{ ps}$. Middle: Error is less than 1 mm. Right: Plot of intensity as a small patch is moved perpendicular to the first surface.	66
5-3	Resolution in lateral dimension measured with a chirp pattern. (Left) Setup with chirp pattern (occluder removed in this photo) (Middle) Raw streak photo from streak camera (Right) The blue curve shows reconstruction of the geometry and indicates that we can recover fea- tures with 0.5 cm in lateral dimensions in the given scenario. The curve shows the confidence in the backprojected values.	66
5-4	Performance degradation due to temporal jitter.	67

5-5 Reconstruction of a scene consisting of a big disk, a triangle and a square at different depth. (Left) Ground truth. (Middle) Reconstruction, front view. (Right) Reconstruction, side view. Note that the disk is only partially reconstructed, and the square is rounded of, while the triangle is recovered very well. This illustrates the diminishing resolution in directions parallel to the receiver plane towards the borders of the field of view. The blue planes indicate the ground truth. The gray ground planes and shadows have been added to help visualization. Referenced from joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41]. 68

5-6 Challenges in imaging around the corner with a conventional, low temporal resolution laser and camera. (a) A setup with hidden mannequin but using a red continuous laser and a Canon 5D camera. (b) An image of the wall recorded with the Canon 5D camera with the room lights turned off and no hidden object present. (The recorded light is due to the reflections from walls behind the laser and camera.) (c) An image recorded with the hidden mannequin present. The increased light level on the wall is marginal, is low spatial frequency and shows no noticeable high frequency structure. (d) An image of the wall with the hidden mannequin moved away from the wall by 10 cm. The reduction in light level on the wall has no visible structure. (e) The difference between image in (b) and (c) using a false color map. (f) The difference between (b) and (d). (g) The difference between (c) and (d). (h) The plot of intensities along the centered horizontal scanline of each of the images (b=red, c=black, d=blue). 69

5-7	<p>Limitations of heatmap after the backprojection algorithm. (Left) Propagation of a single point and its backpropagation for a flatland case. Reconstruction using backpropagation shows that one can recover a sharp peak, but it is surrounded by a low frequency residual. Y and Z sections are shown. The resolution in Y is lower than the resolution in Z. (Right) Cross Correlation of the streak images corresponding to nearby voxels (3D case with 1D streak image). Notice that the Y-resolution is worse because the sensor is 1D along the x-axis. Results from joint work with Andreas Velten in [41].</p>	70
5-8	<p>Reconstruction of a planar object in an unknown plane in 3D. (Left) The object. (Middle Left) 2D Projection of the filtered heatmap. (Middle Right) A 3D visualization of the filtered heatmap. (Right) Reconstruction using sparsity based methods. The gray ground plane has been added to aid visualization. Referenced from joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41].</p>	70
5-9	<p>Reconstruction of a wooden man, painted white. Center - reconstruction using simple back projection based methods. Right - reconstruction using sparse reconstruction methods. Results from joint work with Andreas Velten in [41].</p>	70
5-10	<p>Depiction of our reconstruction algorithm for a scene consisting of two birds in different planes. (a) Photographs of the input models. (b) 9 out of 33 <i>streak images</i> used for reconstruction. (c) The raw (unfiltered) <i>backprojection</i>. (d) The <i>filtered</i> backprojection, after taking a second derivative.(e,f) 3D renderings in Chimera.</p>	71

5-11 The laser beam (red) is split to provide a synchronization signal for the camera (dotted red) and an attenuated reference pulse (orange) to compensate for synchronization drifts and laser intensity fluctuations. The main laser beam is directed to a wall with a steering mirror and the returned third bounce is captured by the streak camera. An Occluder inserted at the indicated position does not significantly change the collected image. Referenced from joint work with Andreas Velten in [41]. 72

5-12 **Smooth Object Reconstruction** Reconstruction of two sinusoids using SPGL1. (a) Photo of the object. The sinusoid is approximately 1.5 cm tall with total length of 5 cm. (b) Nine of the 60 raw streak images. (c) *Side View*. Visualization of one cross sections from reconstruction. (d) *3D cross-sections*. Chosen cross sections along depth (z) projected on the x-y plane reveals the hidden shape contour. (e) *Top View*. (f) *Confidence map*. A rendered point cloud of reconstruction values after soft threshold. 73

5-13 **Complex Object Reconstruction** in multiple poses. (a) Photo of the object. The mannequin is approximately 20 cm tall and is placed about 25 cm from the diffuser wall. (b) Nine of the 60 raw streak images. (c) *Heatmap*. Visualization of the heatmap after backprojection. The hidden shape is barely discernible. (d) *Filtering*. The second derivative of the heatmap along depth (z) projected on the x-y plane reveals the hidden shape contour. (e) *Depth map*. Color encoded depth shows the left leg and right arm closer in depth compared to the torso and other leg and arm. (f) *Confidence map*. A rendered point cloud of confidence values after soft threshold. (i) The stop-motion animation frames from multiple poses to demonstrate reproducibility. Shadows and the ground plane in images (f-i) have been added to aid visualization. Joint work with Andreas Velten, Thomas Willwacher, Ashok Veeraraghavan, Mounji G. Bawendi and Ramesh Raskar [41]. 74

5-14 Reconstruction of a hidden wood mannequin using sparse reconstruction techniques	75
5-15 Demonstration of the depth and lateral resolution. (a) The hidden object to be recovered are three letters, I, T, I at varying depths. The "I" is 1.5 cm in wide and all letters are 8.2 cm high. (b) 9 of 60 images collected by the streak camera. (c) Projection of the heatmap on the x-y plane created by the back projection algorithm. (d) Filtering after computing second derivative along depth (z). The color in these images represents the confidence of finding an object at the pixel position. (e) A rendering of the reconstructed 3D shape. Depth is color coded and semi-transparent planes are inserted to indicate the ground truth. The depth axis is scaled to aid visualization of the depth resolution. Referenced from joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41].	76
5-16 Reconstruction of hidden ITI(MIT) using sparse reconstruction techniques such as SPGL. Note that this result is from a technique entirely different from 5-15. (a) The hidden MIT. (b) Selected streak images. (c) Reconstructed volume with all sections merged into one image. (d) Volumetric slices of reconstructed hidden object. (e) The final result rendered in 3D (side view). Notice how accurately depth gets recovered. (f) Rendered MIT from front.	77

5-17	Results of a multi-pose stop motion animation dataset after filtered backprojection and soft-thresholding. A hidden model of a <i>man with a ball</i> is captured in various poses. The rendering shows the sequence of reconstructions created by our filtered backprojection algorithm and demonstrates the ability to remove low-frequency artifacts of backprojection. The mislabeled voxels remain consistent across different poses indicating stability of our capture and inversion process. Shadows are introduced to aid visualization. Results from joint work with Andreas Velten in [41].	78
5-18	Results from reconstruction of two cards using simple backprojection techniques for real world data. (Left) The actual object. (Right) Reconstructed object.	78
5-19	Results from reconstruction of a wedge using simple backprojection techniques, and carving in real world data. (Left) The actual object. (Right) Reconstructed object.	79
5-20	Results from reconstruction of a man using CoSAMP in simulation. (Left) The actual object. (Right) Reconstructed object.	79
5-21	Results from reconstruction of a man using CoSAMP in simulation. (Left) The actual object. (Right) Reconstructed object.	79

Chapter 1

Introduction

Current computer graphics and vision involve study of light coming from various points in *visible in-line-of-sight* world. Traditional imaging from cameras captures a fraction of light from scene, focuses it and integrates it on a sensor. A very high dimensional representation of scene involves parametrizing light intensity for each wavelength, in every direction coming from every *visible* point in the world. Concepts such as plenoptic function [4] can be used to describe response of an environment to current lighting conditions.

When we capture an image, light from non line of sight objects can also be imaged after reflections from visible objects. The effect of such light on our image is ignored because of its considerably lower intensity. The contribution of these photons is rendered insignificant by integration over much larger periods of time than time light took to move around in a scene. Of course this strategy is ignorant and wasteful of data. In this thesis we address the challenging problem of using these photons to resolve geometry of hidden objects.

1.1 Motivation

Very recently some interesting research has been done in the area we explore. In [32, 31], Raskar and Davis proposed inverse analysis using a 5D time-light transport matrix to recover geometric and photometric scene parameters. By careful calibration

of scene, camera and surrounding world, it is possible to image hidden aspects of scene by switching lighting and cameras [35]. This approach can be used to get a faster 6D representation of scene as well. This approach requires a high resolution, highly versatile light source like a projector focusing light at hidden object and therefore is not of much use in problem we describe. Highly sophisticated techniques such as LIDAR (LIght Detection and Ranging) [16] and two dimensional gated viewing [8] can be used to obtain low/high scale 3D representation of objects. Stealth technologies such as TADAR [7], infrared heat signatures can be used to look through diffused occluders such as walls. But all these technologies ignore the multiple reflections as noise and try to use filters to eliminate this information rather than use it for good. Can we exploit the time taken by photons to arrive to infer hidden 3D shapes ? In paper [17] an interesting technique is evolved which uses time of arrival of photons at high resolution to infer trivial information about scene like two dark lines separated by white line. Our work builds on some of this work and develops a full conceptual model of sensing photons after multiple reflections. Then we use this model to develop robust tools and infer shapes of 3D objects of varying complexity.

One may argue that TADAR is a good alternative to our strategies. In fact hidden imaging has always been approached as a problem involving sensing light transmitting through a diffuser which allows a few wavelengths which is a very strong assumption. But our approach provides a workaround to this obvious flaw, and makes no such assumption. Our final goal is to use photons which came after multiple reflections and determine the hidden objects. We believe that such an invention can revolutionize fields such as endoscopy, fast robot navigation, ultrafast reflectance capture and ultrafast geometry acquisition without worrying about occlusions.

1.2 Problem Description

We aim at exploring the relationship between hidden 3D structure of objects and the associated high dimensional light transport (in space and time). Our main emphasis will be on rigorously formulating the problem and developing methods to invert tran-

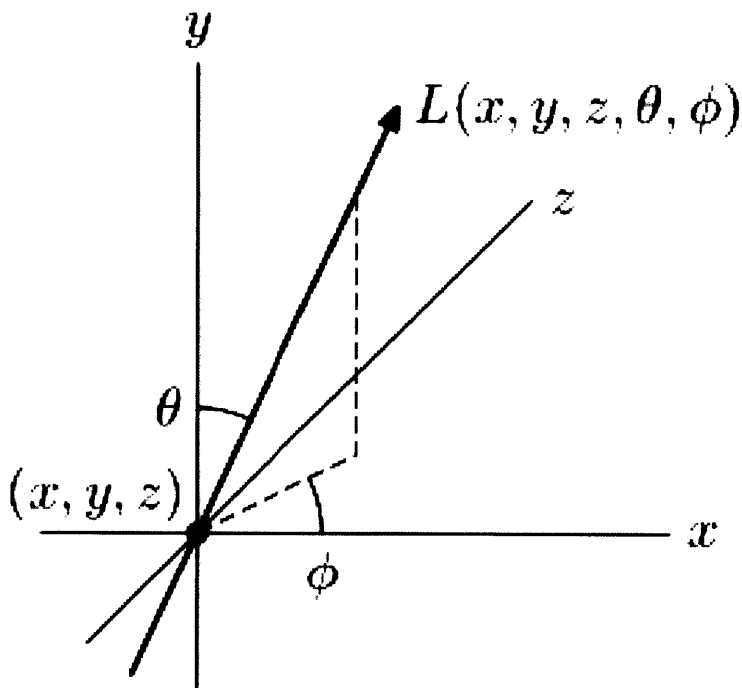


Figure 1-1: Plenoptic function parametrizes world pixels by location, wavelength and orientation. (Source: Wikipedia)

sient light transport and recover hidden 3D shapes. In order to successfully invert the problem we will model the propagation of light in the scene, and its capture by camera. We will also focus at studying issues like existence and uniqueness of solution, and stability of solution algorithm.

We will exploit hardware such as streak cameras to image light with picoseconds level resolution in time. For determining accurate starting time of photons we will illuminate real world scenes using femtosecond duration pulses. Using this we can calculate distances travelled by photons with sub-millimeter level of accuracy.

While developing algorithms we aim at exploiting sparsity based methods to increase robustness and generate reconstruction for challenging inputs. We will develop models for various kinds of noises that might present in the process of capture, and observe how that might affect the results from reconstruction algorithm. We will validate the reconstruction algorithm using data collected in real world scenarios.

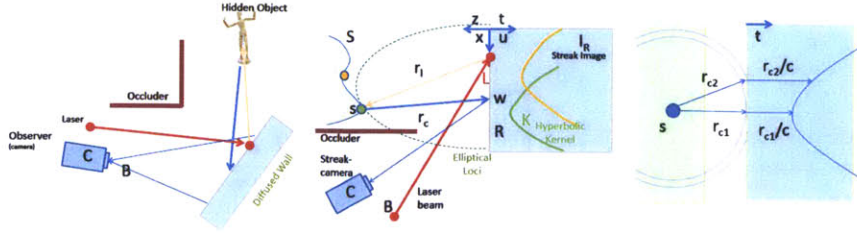


Figure 1-2: Setup (Left). Forward model (Center) The laser illuminates the surface S and each point $s \in S$ generates an energy front. The spherical energy front contributes to a hyperbola in the space-time streak photo, I_R . (Right) Spherical energy fronts propagating from a point create a hyperbolic space-time curve in streak photo.

1.3 Related Work

Recent work in computer graphics and computer vision has shown surprising results by inverting light transport. Our work is influenced by this pioneering work [34, 25, 35, 17, 21]. We also inspire from research in ultra fast imaging as discussed in [24, 40, 45, 27, 31]. We will supplement the framework provided by these methods by analysing and measuring the temporal evolution of light transport. Analysing the energyfront propagation, we will exploit the added fine-scale time dimension to deal with and recover the geometry of moderately complex hidden objects.

Comparison with Dual Photography and Transient Imaging : Light transport acquisition, analysis and inversion is recently emerging to be a powerful tool to extract intrinsic information about the scene even when the scene is hidden either physically or because of global light transport effects. Sen et. al. exploited Helmholtz reciprocity to recover view of a playing card visible to a light source but not to the primal camera [35] at 66x87 pixel resolution. This does require a projector to be in the line of sight. See figure 1-3.

Kirmani et. al. showed an intriguing idea of a time of flight camera to look around corners. The results show how to read a mirror-based barcode beyond the line of sight [17]. Given the challenging nature of the problem, they made strong

assumptions about the hidden points: they lie on a plane, the correspondence between time profiles of each point is recoverable and hidden points are not on a continuous surface (i.e. patches are discretized in position with gaps between the patches). This aids in marking contribution from each patch. This would be impractical in presence of real world continuous objects with unknown depth. Solving the problem in the *absence of correspondences* on three dimensional objects will be the key aspect of our work.

Range estimation using time of flight : 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 - see 1-5), and acoustic (SONAR - refer to 1-4) [36, 15, 13]. However these techniques are limited by penetrability and the assumption of single path and one-bounce reflections. Multibounce RADAR has also been investigated, for example for motion detection [38]. Due to the longer wavelength of RADAR, bounces are always specular and reconstruction is trivial. The resolution and capabilities are very limited.

Time of Flight and Scattering : In radio-location , geophysical surveys, synthetic aperture radar and medical imaging, problems involve determining geometric information from scattered emission, reflection or transmission [23, 10, 28, 30, 42, 44]. For example, in radio-location, one can conveniently query the active emitter while in borehole tomography one can use one explosive source at a time to generate signals which are picked up by receivers at the surface and in boreholes. The inversion is challenging because the emitters and receivers are sparse. Most problems also involve nonhomogeneous medium with varying refractive index. Kirchhoff integral techniques represent and solve the propagation with partial differential equations [20, 18, 9, 37, 43]. The geometric solutions are relatively coarse and low resolution and exploit transmission or one bounce reflection.

Apart from these the theory of inverse light transport [34] can be used to eliminate inter-reflections from real scenes. The frequency domain properties of direct and global components of scattered light can be exploited to recover image of objects

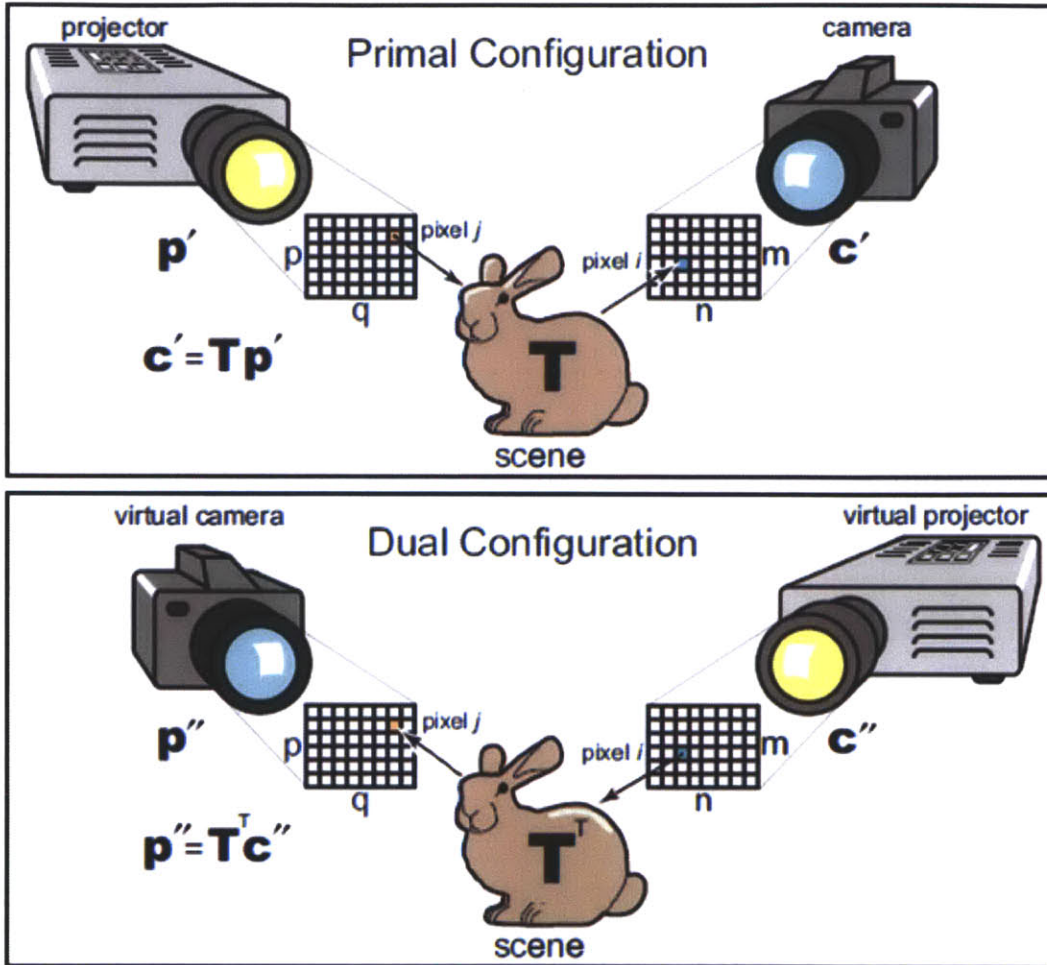


Figure 1-3: Dual Photography: *Helmholtz Reciprocity*. (Please refer to [35])

behind a shower curtain [25].

1.4 Contribution

We will explore the relationship between hidden 3D structure of objects and the associated high dimensional light transport (space and time). We show that using multi-bounce energyfront propagation based analysis one can recover hidden 3D geometry. We rigorously formulate the problem, elicit the relationships between geometry and acquired light transport and also develop a practical and robust framework for inversion. The specific technical contributions of the thesis are



Figure 1-4: TADAR : *millimeter wavelength imaging*. Source [2]

- An algorithm for backpropagating the acquired space-time light transport to overcomes the lack of correspondences and recover 3D shapes
- Exploiting sparsity based methods to reconstruct hidden 3D shapes
- An analysis of the invertibility and dependence on resolution of space-time dimensions
- Several synthetic and physical experiments to validate the concepts

We will show that using multi-bounce energy front propagation based analysis one can recover hidden 3D geometry. We analyze the problem of recovering a 3D shape from its tertiary diffuse reflections. If there was only a single hidden point, the reflected energy front directly encodes the position of that point in 3D. We rigorously formulate the problem, elicit the relationships between geometry and acquired light transport and also develop a practical and robust framework for inversion. We show that it can be cast as a very peculiar type of tomographic reconstruction problem. We call the associated imaging process *elliptic tomography*. The inverse problem, i.e., the recovery of the unknown scene from the measurements, is challenging. We provide a fast algorithm, which is essentially the analogue of the filtered backprojec-

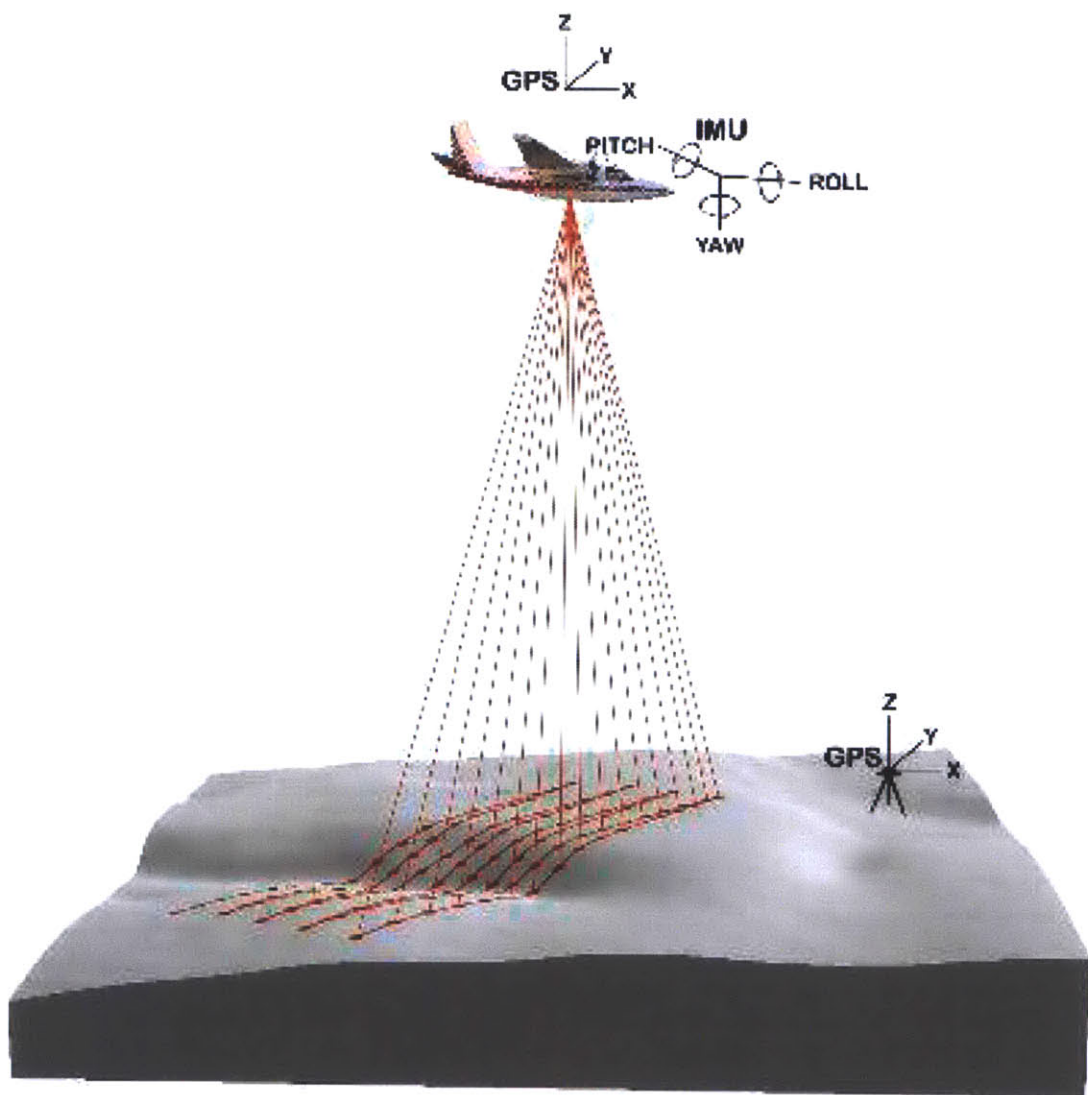


Figure 1-5: LIDAR *Time of flight of photons.*(Source [3])

tion algorithm in traditional tomography. We perform several synthetic and physical experiments to validate the concepts.

Chapter 2

Theoretical Foundations

2.1 Modelling Propagation of a Light Pulse for Multiple Bounces

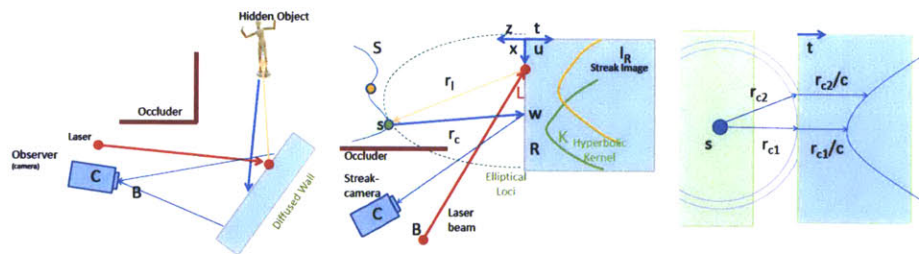


Figure 2-1: Forward Model. (Left) The laser illuminates the surface S and each point $s \in S$ generates a energy front. The spherical energy front contributes to a hyperbola in the space-time streak photo, I_R . (Right) Spherical energy fronts propagating from a point create a hyperbolic space-time curve in streak photo.

Consider a scene, as shown in Figure 2-1, which contains a hidden object (whose surface we are interested in estimating) in front of a visible surface or 'first surface' (this can be a wall or the floor). For simplicity we will assume that the visible surface (hereafter called wall) is planar and diffuse, but we will relax this assumption later. A laser beam(B) pointed towards the first visible surface (wall) to form a laser

spot L emits a very short (few picoseconds) light pulse. The light reflected by the wall reaches the hidden surface, is reflected and returns back to the wall. A streak camera pointed towards the wall records the time varying image on the wall at a very high temporal resolution of around 2 picoseconds. Our goal is to understand the relationship between the geometry of the hidden scene and the observed intensities at each streak camera pixel and time bin. (This work was done in collaboration with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41].)

For each location of the laser spot L , a 3D image (2 spatial and 1 temporal dimension) is recorded. The laser spot is moved to multiple locations on the wall ($2D$) to record $5D$ light transport data. The pulse return time at each location on the wall depends upon several known parameters such as the location of the laser spot and the unknown surface profile. So one can use the observed $5D$ light transport data to infer the hidden surface shape. In order to understand the basic intuition, consider the hidden scene to be a single point, as shown in Figure 2-1. The reflected spherical wavefront propagating from that scene point reaches the different pixels on the wall surface at different times creating a streak image which is a hyperbolic curve. There is a one-one invertible mapping between the parameters of the hyperbola and the $3D$ location of the hidden scene point.

When the hidden scene contains a surface instead of individual scene points, the space-time hyperbolas corresponding to the different surface points are added together to produce the captured streak images and so we need to demultiplex or deconvolve these signals. In general, we could use a captured $5D$ light transport data but in our experiments, we are restricted to a 1D streak camera. While the information from a 1D streak camera is sufficient for 3D reconstruction the spatial resolution in the axis perpendicular to the sensor axis is heavily compromised.

It is convenient to move to the coordinate system of the wall. For this we apply a space-time transform. We need to consider four optical path segments: beam B to laser spot L on first surface, spot L to a hidden surface points $s \in S$, s to points on the third surface $r \in R$ and finally r to the camera. We have folded first and third

surfaces into a single 'wall' for compactness. With respect to the wall, the shape of the objects can be recovered as a 2D height field. By shining a short pulse duration laser and recovering a full time profile of the returned signal, we capture a 5D light transport, 2D for camera pixels, 2D for laser projector positions and time. Thus, our inversion method needs to recover a 2D height map manifold from a 5D light transport data. As we show later, recovering this hidden 2D manifold with a 4D camera-projector light transport and without time-dimension is impractical due to ill-conditioning of the multiplexing matrix.

2.1.1 Space-Time Warping for Bounce Reduction

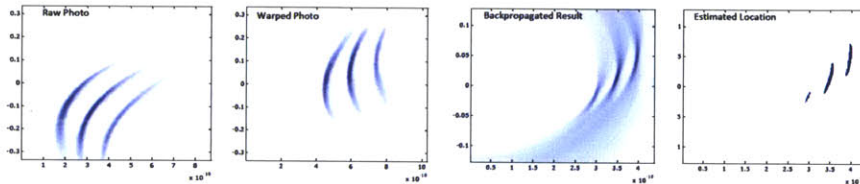


Figure 2-2: A space time transform on a raw streak photo allows us to convert 4 segment problem into a sequence of 2 segment problems. The toy scene is a small $1\text{cm} \times 1\text{cm}$ patch creating a prominent (blurred) hyperbola in warped photo. Back-propagation creates low frequency residual but simple thresholding recovers the patch geometry.

For the looking around corner setup light path can be divided into 4 straight line segments with 3 bounces in between. The first segment involves collimated laser beam travelling from laser source to wall. In the second segment the laser spot on wall behaves as a point light source. The third segment(s) involve scattering from hidden object. For the fourth segment light travels from wall to camera *which has wall in focus*. The data received on camera $I_c(p, t)$ has two degrees of freedom - space and time. Since first and fourth segments are focused, we can apply transforms to streak image to eliminate their effects. The effect of first segment can be removed by shifting streak image in time which is constant for each location on wall. To remove effect of fourth segment we use the distances between wall pixels and camera sensors ($\|C - w\|$). We assume that we know the geometry of R to calculate camera

homography and correspondence between camera sensor and wall pixels. The following mathematical formulation is concise representation of this concept. Here H is the projective transformation (homography) mapping coordinates on R to camera coordinates. The time shift $\|C - w\|$ by the distance from camera to screen varies hyperbolically with the pixel coordinate w . Note that we don't require to adjust for $\cos(\theta)$ factor or $1/r^2$ fall off because camera sensor integrates for more wall pixels if they are farther away.

We can simplify the task of inference from four segment reflection paths to a sequence of two segment paths. First and fourth paths are directional and focused and do not involve scattering. The second segment can be treated as originating from unfocussed sensors while the third segment can be treated as being recorded by unfocused sensors. Our approach simplifies this three bounce problem to one bounce problem, by introducing corresponding space and time warps due to first and fourth paths. For the first path we calculate the time taken by laser to reach the wall and offset everything to include this. For the fourth path we compute the homography for streak camera imaging system and map the 1D image created on streak sensor to the wall in front. Then the pixels are offset in time to include the variation in time for light to reach the streak camera sensor. After discretization this situation is analogous to an unfocussed array of emitters and an unfocussed array of receivers. Unfortunately each of the emitters has a different phase depending on its position created due to its distance to laser spot and distance from the sensor. In addition the intensities are also impacted by aforementioned distances. Our goal is to recover from this multiplexed readings the location of each of transmitters.

2.1.2 Scattering of a pulse

Consider the impact of any scene point $s \in S$ on the recorded 5D light transport. The laser illuminates a point L on the visible surface of the wall and this in turn creates a virtual unfocused impulse light source. The light reflected from each scene point s is recorded by unfocused virtual sensors $w \in R$. We model the impulse wave propagation from each scene point using the well-known Huygens–Fresnel principle. Phase changes

hyperbolically and intensity changes as r^2 across the planar section for the wavefront. This hyperbolic streak, or 'streak' for short is captured by a streak camera.

Generating Streak Photos

Let's analyze the scattering of light in second and third segments. For simplicity we model the hidden object as a collection of unfocussed emitters sending impulses $I_s(s, \tau)$ at τ times. We model the receiver as an array of unfocussed receivers which capture photons at picosecond resolution. We can achieve this configuration experimentally by using a picosecond resolution streak camera focused on wall pixels. To simplify further we can assume accurately that speed of light is constant and express everything in distance units. The following mathematical equation represents the data recorded by streak camera after making the mentioned assumptions. Note that speed of light is set $c=1$ for simplicity. We also ignore the local changes in normals for sender surface and receiver surface. (Algorithm programming with Thomas Willwacher and data collection with Andreas Velten, results in paper [41].)

$$I_R(w, t) = \int_S \int_{\tau} \frac{1}{\pi r_c^2} \delta(r_c - t + \tau) I_S(s, \tau) d\tau d^2s \quad (2.1)$$

where $w \in R$, $s \in S$, $t, \tau \in \mathbb{R}$ and $r_c = ||w - s||$. and $I_R(w, t)$ is the intensity observed at $w \in R$ at time t . After removing the time shifts as described in previous section and applying transforms from calculated homography we can further simplify the equation and remove the δ required to adjust for receiver camera distances. The following equation provides the mathematical summary of this analysis. Note that overall these equations we assume that receiver and senders are perfectly lambertian and ignore the local superficial variation in Normal vectors. Equation (2.1) hence becomes

$$I_R(w, t) = \int_S I \frac{1}{\pi r_c^2} \frac{1}{\pi r_l^2} \delta(t - r_c - r_l) d^2s \quad (2.2)$$

2.1.3 Hyperbolic Contribution

Hyperbolic Contribution

Lets analyze the relationship between time when a sender emits a pulse to time and location of a receiver. For a fixed sender the response function is a hyperboloid in space and time given by following mathematical equation. The parameters of the hyperboloid depend on location of sender, a lateral displacement leads to shifts, while a displacement in depth corresponds to flattening. Change in sender time equates to a constant time shift for responses to any of receivers.

$$t - r_t = r_c = \sqrt{(x - u)^2 + (y - v)^2 + z(x, y)^2} \quad (2.3)$$

where u, v are the two coordinates of w in the receiver plane. Careful observation shows that this equation is an ellipsoid in sender location if we fix the laser and receiver location. Ellipse parameters depend on time when a receiver receives a impulse. The laser and receiver lie on the two foci of this ellipse. The eccentricity depends on the time when impulse is received.

2.2 Forward model: Elliptical Tomographic Projection

In this section we rephrase the above approximation to the forward light transport using notions from tomography. In an idealized case, the inverse problem of recovering the hidden shape can be solved explicitly. We use this explicit solution to inspire our algorithm for real world scenarios. This work was done in collaboration with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mouni G. Bawendi and Dr. Ramesh Raskar [41].

2.2.1 Elliptical tomography problem description

Our problem has similarities to tomographic projection. Let us rewrite equation (2.2) in the following form

$$\begin{aligned} I_R(w, t, L) &= \int_S I \frac{1}{\pi r_c^2} \frac{1}{\pi r_l^2} \delta(t - r_c - r_l) d^2 s = \int_{\mathbb{R}^3} \frac{1}{\pi r_c^2} \frac{1}{\pi r_l^2} \delta(t - r_c - r_l) I \delta_S(x) d^3 x \\ &= \int_{\mathbb{R}^3} \frac{1}{\pi r_c^2} \frac{1}{\pi r_l^2} \delta(t - r_c - r_l) W(x) d^3 x \end{aligned}$$

where the unknown world volume $W(x) = I \delta_S(x)$ is a delta function with support on the surface S to be reconstructed. Apart from the $1/r^2$ factors, which typically vary slowly over the region of interest, we hence see that individual streak image pixels measure elliptical projections of the world volume $W(x)$. Due to its similarity with traditional tomography, we call this problem *elliptical tomography*. Note however that there are also key differences to traditional tomography: (i) The recorded projections live in a higher dimensional (5D) space than the world (3D). (ii) The projections are along (2D) ellipsoids instead of (1D) lines. This makes the analysis much more complicated.

2.2.2 Challenges and missing cones

It is instructive to consider the above tomography problem in the limit when the object is small compared to the distance to the diffuser wall. In this case the elliptic tomography problem reduces to a planar tomography problem, see Figure 3-1. Each pair of a camera point and a laser position on the diffuser wall (approximately) measures intersections of the target object with planes whose normals agree with the normals to the ellipsoids. By the standard Fourier slice theorem, each line of each streak image will hence measure one line of the Fourier transform of the object, in the direction of the normal. Unfortunately, in our situation these normals cover a limited region of the unit sphere. Hence without additional priors it is not possible to reconstruct the Fourier transform of the target object in the missing directions. This is the missing cones problem well known from traditional tomography. Experimentally

we get a very good resolution in the depth (orthogonal to the wall) direction, while transverse (parallel to the wall) high frequency features tend to get lost.

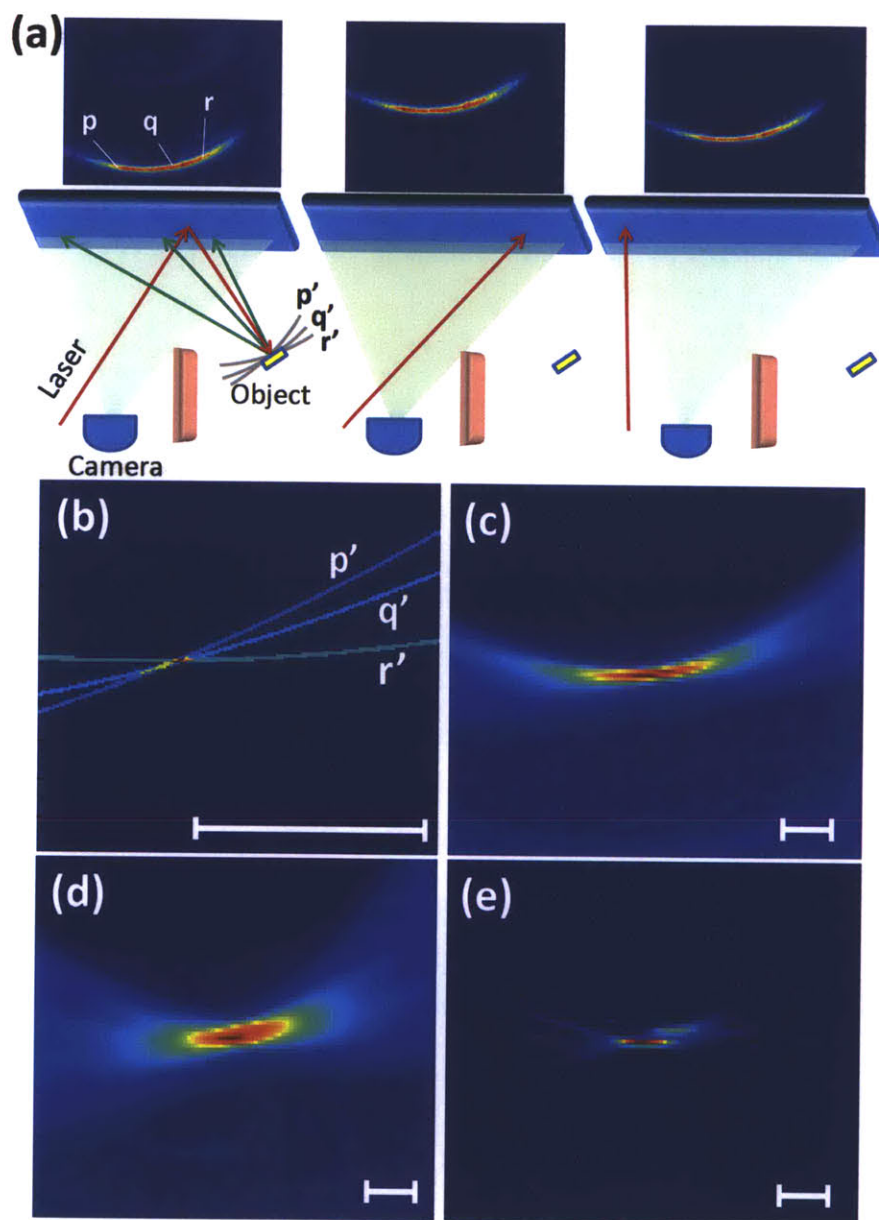


Figure 2-3: **Reconstruction Algorithm** An illustrative example of geometric reconstruction using streak images. (a) *Data capture*. The hidden object consists of a 2 cm \times 2 cm square white patch. The captured streak images are displayed in the top row. (b) *Contributing voxels in Cartesian space*. For recovery of hidden position the possible locations in Cartesian space that could have contributed to the streak image pixels p, q, r are ellipsoids in 3D. If there is a single world point contributing intensity to all 3 pixels, the corresponding ellipses intersect. The white bar corresponds to 2 centimetres. (c) *Backprojection and heatmap*. We use a back-projection algorithm that finds overlaid ellipses corresponding to all pixels. (d) Backprojection using all pixels in a set of 59 streak images. (e) *Filtering*. After filtering with a second derivative, the patch location and 2 centimeter lateral size are recovered. Joint work with Andreas Velten, Thomas Willwacher, Ashok Veeraraghavan, Mounqi G. Bawendi and Ramesh Raskar [41].

Chapter 3

Inversion Analysis

3.1 Inverse Algorithm: Filtered Back Projection

In this section we give a detailed description of our reconstruction algorithm. (Algorithm programming with Thomas Willwacher, Andreas Velten and data collection with Andreas Velten, in collaboration with Dr. Ashok Veeraraghavan, Dr. Mouni G. Bawendi and Dr. Ramesh Raskar [41].)

3.1.1 Overview of the algorithm

The imaging and reconstruction process consists of 3 phases:

- **Phase 1: Data Acquisition.** We direct the laser to 60 different positions on the diffuser wall and capture the corresponding streak images. For each of the 60 positions more than one images are taken and overlaid to reduce noise.
- **Phase 2: Data Preprocessing.** The streak images are loaded, intensity corrected and shifted to adjust for spatiotemporal jitter.
- **Phase 3: 3D Reconstruction.** The clean streak images are used to reconstruct the unknown shape using our backprojection-type algorithm.

The first of the three phases has been described in chapter 1 and 2. Let us focus on Phases 2 and 3.

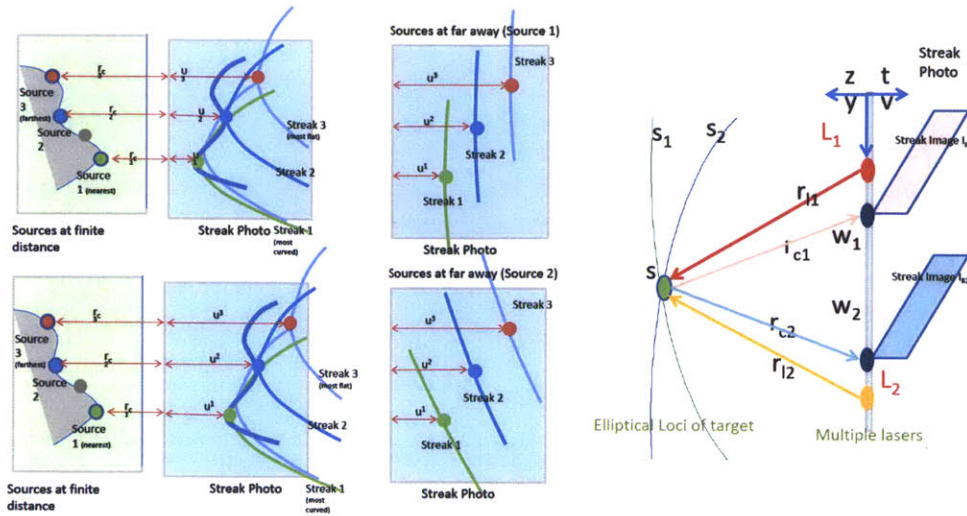


Figure 3-1: The top left figure shows streak images being generated by near field sources. On bottom left we see effect when this sources travel farther away, The rightmost figure depicts how we can analytically predict single sources using multiple sensor laser combinations. Notice how the accuracy is affected if lasers shift.

3.1.2 Phase 2: Data Preprocessing

1. **Timing correction.** To correct for drift in camera timing synchronization (jitter) both in space and time we direct part of the laser directly to the diffuser wall. This produces a sharp “calibration spot” in each streak image. The calibration spot is detected in each image, and the image is subsequently shifted in order to align the reference spot at the same pixel location in all streak images. The severity of the jitter is monitored in order to detect outliers or broken datasets. (Data collection with Andreas Velten, results in paper [41].)
2. **Intensity correction.** To remove a common bias in the streak images we subtract a reference (background) image taken without the target object being present in the setup.
3. **Gain correction.** We correct for non-uniform gain of the streak camera’s CCD sensor by dividing by a white light image taken beforehand.

3.1.3 Phase 3: 3D Reconstruction

1. **Voxel Grid Setup.** We estimate an oriented bounding box for the working volume to set up a voxel grid (see below).
2. **Downsampling (optional).** In order to improve speed the data may be down-sampled by discarding a fraction of the cameras pixels for each streak image and/or entire streak images. Experiments showed that every second camera pixel may be discarded without losing much reconstruction accuracy. When discarding entire streak images, it is important that the laser positions on the diffuser wall corresponding to the remaining images still cover a large area.
3. **Backprojection.** For each voxel in the working volume and for each streak image, we compute the streak image pixels that the voxel under consideration might have contributed. Concretely, the voxel at location v can contribute to a pixel corresponding to a point w on the wall at time t if

$$ct = |v - L| + |v - w| + |w - C|.$$

Here C is the camera's center of projection and L is the laser position as above. Let us call the streak image pixels satisfying this condition the *contributing pixels*. We compute a function on voxel space, the *heatmap* H . For the voxel v under consideration we assign the value

$$H(v) = \sum_p (|v - w||v - L|)^\alpha I_p.$$

Here the sum is over all contributing pixels p , and I_p is the intensity measured at that pixel. The prefactor corrects for the distance attenuation, with α being some constant. We use $\alpha = 1$.

4. **Filtering.** The heatmap H now is a function on our 3 dimensional voxel grid. We assume that the axis of that grid are ordered such that the third axis faces away from the diffuser wall. We compute the *filtered heatmap* H_f as the second

derivative of the heatmap along that third axis of the voxel grid.

$$H_f = -(\partial_3)^2 H.$$

The filtered heatmap measures the confidence we have that at a specific voxel location there is a surface patch of the hidden object.

5. **Thresholding.** We compute a sliding window maximum M_{loc} of the filtered heatmap H_f . Typically we use a 20x20x20 voxel window. Our estimate of the 3D shape to be reconstructed consists of those voxels that satisfy the condition

$$H_f > \lambda_{loc} M_{loc} + \lambda_{glob} M_{glob}$$

where $M_{glob} = \max(H_f)$ is the global maximum of the filtered heatmap and $\lambda_{loc}, \lambda_{glob}$ are constants. Typically we use $\lambda_{loc} = 0.45, \lambda_{glob} = 0.15$

6. Compressive Reconstruction

We use techniques like SPGL1, and CoSAMP as an alternative to back projection and filtering. We rely on the fact that the 3D voxel grid is sparsely filled, containing surfaces which can occupy only one voxel in depth.

7. **Rendering.** The result of thresholding step is a 3D point cloud. We use the Chimera rendering software to visualize this point cloud.

In order to set up the voxel grid in the first step we run a low resolution version of the above algorithm with a conservative initial estimate of the region of interest. In particular, we greatly down sample the input data. By this we obtain a low resolution point cloud which is a coarse approximation to the object to be reconstructed. We compute the center of mass and principal axis of this point cloud. The voxel grid is then fit so as to align with these axis.

3.2 Inversion using parameter estimation

Another good approach that we could use for reconstruction involves modeling captured data as collection of hyperbolic streaks with unknown parameters. We can formulate this problem as optimization problem where we are trying to fit a conic section (hyperboloid) to a streak, minimizing total number of hyperbolas fitted. As discussed in next chapter this can be achieved by imposing a sparsity constraint on reconstructed world. In a very simple model each active pixel corresponds to a unique hyperboloid and we are trying to minimize number of active pixels while reducing residual errors. We can describe the residual error by following equation:

$$\sum_{(x,t) \in \text{StreakImages}} \left\| I(x,t) - \sum_{i \in \text{ActivePixels}} S_i * \delta(x^2/a_i^2 - t^2/b_i^2 - 1) \right\|^2 \quad (3.1)$$

Here $I(x,t)$ is the recorded streak image, a_i and b_i are parameters for streak corresponding to active pixel and S_i is corresponding intensity. We can then include the sum of all active voxels or number of active voxels as a regularization parameter to complete the formulation. The main issue with this formulation is that actual streaks have a lot of blur and noise, and it can lead to a lot of computation overhead because of large parameters. Also because of systematic noise we cannot be sure that hyperbolas will overlap exactly. To overcome this issue, instead of fitting the pixels exactly we can minimize their from distance from the hyperbolas.

There are two interesting observations to make about this technique. For scenes with isolated patches, we could reduce our search space to a few parameters. For scenes with continuous random surfaces, this search space grows infinite and needs to be discretized. A robust way to estimate the active voxels can be to use hough transform. Very simply, we can overlay the streak image due to each voxel and integrate to get a likelihood estimate for that voxel being active. We can use appropriate threshold to decide which voxels to choose. This technique is exactly same as backprojection technique we discussed earlier.

3.2.1 A remark about the filtering step

To motivate our choice of filter by taking the second derivative, let us consider again the planar ("far field") approximation to the elliptical tomography problem as discussed in section 2.2.2. In this setting, at least for full and uniform coverage of the sphere with plane directions, the theoretically correct filtering in the filtered back-projection algorithm is a $|k|^2$ filter in Fourier space. In image space, this amounts to taking a second derivative along each scanline. This motivates our choice of filter above. We tested both filtering by the second derivative in image and world space and found that taking the derivative approximately along the target surfaces normal yielded the best results.

Chapter 4

Sparsity Based Reconstruction

In this chapter we show how we can use SPGL1 and other compressive sensing techniques to reconstruct hidden volumes. Because we aim at reconstructing hidden object's 2D manifold, we can consider the solution as a sparse set of points inside the voxelized world volume. This approach has many advantages over the simple back-projection based techniques. This approach actually inverts the system in front of us instead of using a voting technique like backprojection or carving. This approach will also minimize noise by choosing accurate parameters. Further as results show, this technique can yield much better results than ad-hoc approaches like backprojection.

4.1 Linearizing The system

Consider a isolated voxel in world, and focus on response it produces in streak images. The response changes orientation in time and space based on the laser location and wall pixels being observed. The response observed at each pixel on wall is a linear time invariant sum of time signals sent by all active world voxels. In essence to get any streak image we can add streak images from all world voxels provided that occlusion doesn't occur. We can construct a large matrix to represent this system, we denote the corresponding linear transform by matrix A . Each column of this matrix represents the streak image from corresponding world voxel. Each row in corresponding linear transform matrix A identifies a unique pixel,laser-location pair. This matrix can be

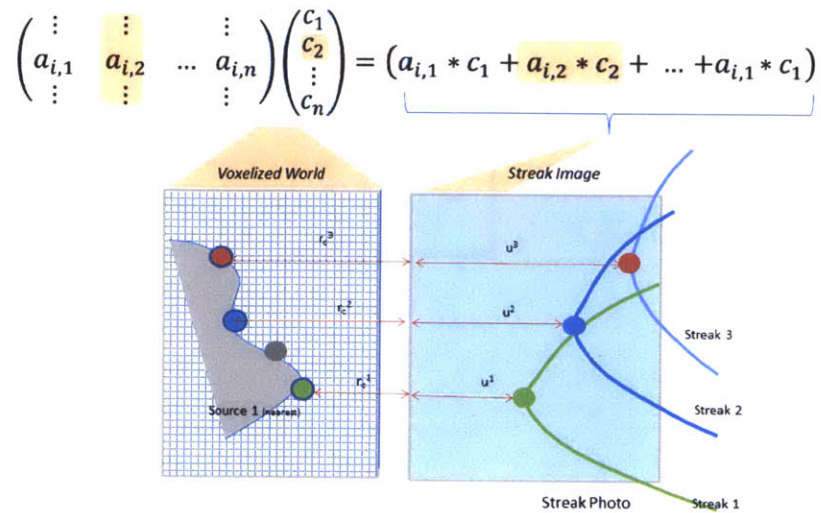


Figure 4-1: How each world voxel contributes to a streak inside data. The entire streak image is composed by linear sum of all the streaks. If there is no occlusion we can turn the forward model into a linear system and can generate a corresponding matrix representation. Vectorized world voxels (denoted by c_i) on multiplication with linear transform matrix A yield streak image. The right hand is sum of streak images A_j corresponding to each voxel.

used to quickly draw streak images and backproject. The matrix comes to roughly 1 billion entries, and can easily takes Gigabytes of space. The matrix is practical only if the voxelized world is small in size or we are willing to sacrifice resolution. (please refer to 4-1)

Since number of wall pixels and laser locations is considerably more than world voxels, size of corresponding linear transform matrix A can be reduce by sub sampling world voxels and selecting only a few lasers. This too leads to poorer reconstruction (4-4) but the results are a lot better than slashing down resolution.

Once we have a matrix to represent our system, we can try to invert the system by solving linear system $Ax = B$. However using least square or inverting the system is a naive approach and can be computationally heavy. A easy solution is to reduce number of columns of the linear system by multiplying by a well suited matrix M on both sides ($MAX = Mb$). In our problems we used A^T to conserve rank and create

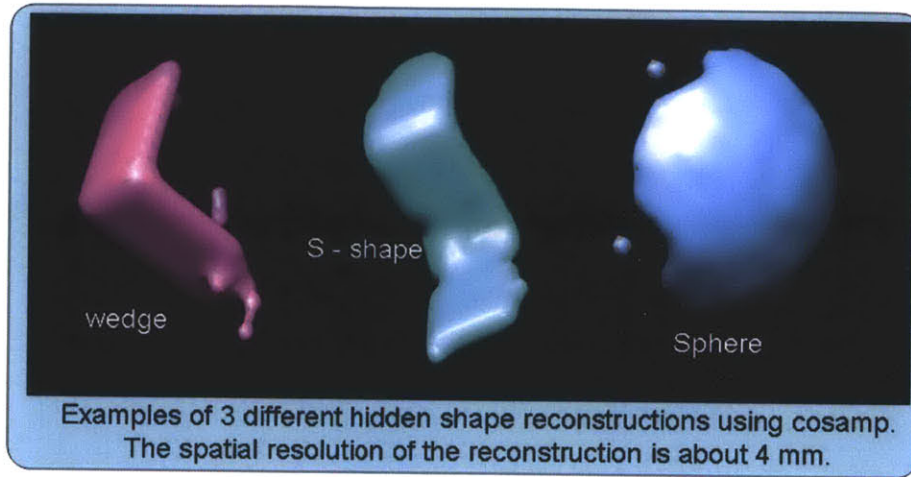


Figure 4-2: Our approach can handle steep angles, sharp corners, concave as well as convex features. Reconstruction of (a) Wedge (b) S shape and a (c) Sphere using 8 streak images and COSaMP matching pursuit algorithm. Referenced from joint work with Ashok Veeraraghvan in [41].

a system easier to solve. Figure 4-3 shows the reconstruction from this approach.

4.2 Methods

By orienting the world volume appropriately we can even put an upper bound on the number of "active" points by counting number of voxels in a plane parallel to observed wall. Alternatively this number can be set to number of voxels in a plane which cuts through the world and maximizes this number. Once we have this number we can use techniques like CoSAMP [26] to reconstruct the result. The figure 4-2 shows simulated reconstruction on 3 different volumetric objects using CoSAMP.

Because it's not very accurate to guess world voxels, a better approach is to guess the sum of intensity values of all voxels. We can get a fair estimate of sum of all voxels by backprojecting and summing values above a certain threshold. We use sparsity based methods to reconstruct the 3D object once this value is known. Sparse approximation algorithms have been studied in detail and applied in fields of image reconstruction, signal processing and statistics. Techniques like *l1ls* are being applied successfully in areas of MRI and seismic imaging. We intend to exploit these

techniques in our problem by posing it as a tomographic image reconstruction.

4.2.1 Sparse Formulation

In previous section we showed that after fixing few parameters, our forward model is a linear time-invariant system. We also showed how to construct a matrix representation of this linear system. In following section we describe a few formulations which impose sparseness on solution. We can achieve that by solving $Ax = b$ while imposing $\|x\|_1$ in an under constrained system. Greedy strategies and matching pursuit algorithms like OMP [39] and StOMP [22] can be used to solve this problem but they don't ensure convergence. We can also pose this as a pure signal processing and recovering problem. The following equation shows the problem formulation for noiseless cases.

$$\text{minimize } \|x\|_1 \text{ subject to } Ax = b \quad (4.1)$$

4.2.2 Matching Pursuit Algorithms

We can use algorithm CoSAMP [26] to invert the problem in presence of noise. We can regard streak images as projections of unknown signal x into a high dimensional space using operator A that we developed in earlier sections. The operator A can be used to generate a proxy for hidden volume x using $\bar{x} = A^T Ax$. We cannot invert this system directly because of extremely large size of A . After we have an initial support we can apply the same operator to residual signal to get more support points and merge them with the existing support points. The following equation concisely describes this concept.

$$\bar{x} = A_S^{-1} * b \text{ where } S_n = \{A_i | (A^T * (b - Ax))_i \geq n^{th} \max(A^T (b - Ax))\} \quad (4.2)$$

This is same as back-projecting the streak image residuals without applying the filter ($\bar{x} = A^T b$). Once we have a *proxy* for signal, we can identify the larger components

and invert linear system $Ax = b$ just for the support vectors of larger components, to get better estimates of x . Since number of support points is usually much smaller than dimension of x this problem becomes computationally feasible. Pseudocode for CoSAMP is given below:

Algorithm 1 CoSAMP(A,b,n)

```

i ← 0
x0 ←  $\bar{0}$ 
r ← b
repeat
  i ← i + 1
  y ←  $A^T * r$ 
   $\Phi \leftarrow \text{supp}(y_{2n})$ 
   $S_i \leftarrow \Phi \cup S_{i-1}$ 
   $x_{tot} = A_S^{-1} * b$ 
   $x_i \leftarrow x_S^{tot}$ 
  r ←  $b - A * x_i$ 
until  $\|r\| \geq \epsilon$ 

```

4.2.3 Fixed Point Methods

This approach discussed previously is idealistic but fails because of systematic noise, and because linear transform matrix A may not have restricted isometry constant $\delta_s < 1$. A better approach will be to formulate it as one norm regularized problem discussed in [11]. In following equation we show how both residual error and one norm of x can be minimized with a single parameter λ to control the sparseness.

$$\arg \min \|Ax - b\|_2^2 + \lambda \|x\|_1 \quad (4.3)$$

We can solve this problem using fixed point methods. Hale, Yin and Zhang in [11] come up with an algorithmic framework to solve this problem using fixed point iterative procedure. Let $T_1(x) = \lambda \partial \|x\|_1$ and $T_2(x) = \nabla f(x)$, the following composite operator can be used to generate better approximations of x

$$x_{k+1} = (1 + \tau T_1)^{-1} \circ (1 - \tau T_2)(x_k) \quad (4.4)$$

where τ is any scalar. We can derive the following result (Refer to [11]):

$$s_v = (1 + \tau T_1)^{-1}, h = (1 - \tau T_2) \quad (4.5)$$

Please refer to [11] for detailed description and analysis. The downside of this approach 2 is that it makes a lot of convexity related assumptions on the function. Because of large size of matrix A, the operator generation can become computationally infeasible. The next section discusses a better approach which can provably yield same results.

Algorithm 2 Fixed Point Continuation Algorithm, FPC(A,b,x⁰, \bar{u})

```

Select  $0 < u_1 < u_2 < \dots < u_L = \bar{u}$ . Set  $x = x^0$ 
for  $u = u_1, u_2, \dots, u_L$  do
  while "not converged" do
    Select  $\tau$  and set  $v = \tau/u$ 
     $x \leftarrow s_v \circ h(x)$ 
  end while
end for

```

4.2.4 Projective Gradient with Lasso

A very similar formulation is called basis pursuit denoise and involves minimizing one norm with a restriction on observed noise. The equation below summarizes how this can be done.

$$\arg \min \|x\|_1 \text{ subject to } \|Ax - b\|_2 \leq \sigma \quad (4.6)$$

The parameter σ is related to λ as shown by Berg and Friedlander in [6]. While these formulation enforce sparseness as a primary constraint, we found that a more practical strategy was to minimize the residual error while enforcing one norm of x below a certain threshold τ . We adopted the Lasso formulation by Chen discussed [33].

$$\arg \min \|Ax - b\|_2^2 \text{ subject to } \|x\|_1 \leq \tau \quad (4.7)$$

The parameters τ , σ and λ are interconnected, and can lead to same solutions[6]. λ is connected with Lagrange multiplier of constraint in lasso formulation[6]. Identifying the parameters is really the key to getting to right solution. Since backprojection gives us a robust estimate of τ we can use the third formulation given below with ease. Once we know the formulation we can pose basis pursuit as a linear program and use packages like CVX and CPLEX to solve the problem. Unfortunately this approach is computationally intensive and impractical in dealing with large amounts of data that we encounter.

In [6], Berg and Friedlander show how a spectral projected gradient algorithm can be used for lasso LS_τ formulation of our problem. We can re-cast the Lasso problem as a trade off between one norm of solution and two norm of residual (refer to 4-6, 4-7). The pareto curve between these two is described by following equation[6]. It can be shown that this curve is convex, strictly decreasing and differentiable for all points of interest [6] - this is an important property.

$$\phi(\tau) = \min \|r\|_2 \text{ subject to } Ax + r = b, \|x\|_1 \leq \tau \quad (4.8)$$

Because of convex properties of aforementioned function we can use newton's method to solve for roots of $\phi(\tau) = \sigma$.

The figure 4-13 shows some exciting reconstructions using SPGL.

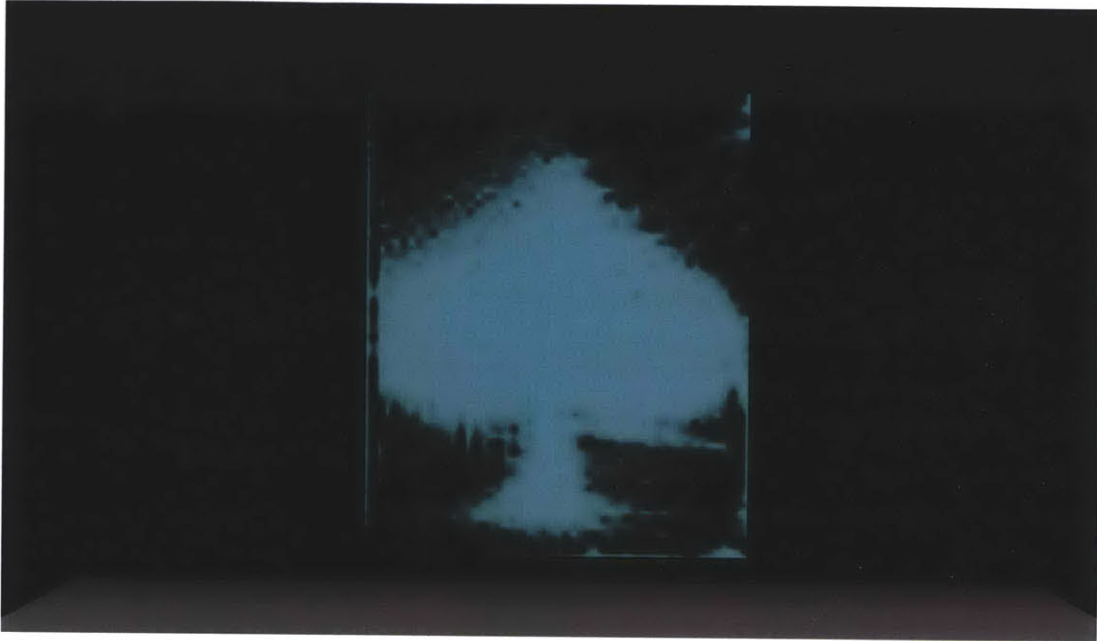


Figure 4-3: Reconstruction of a spade using plain linear inversion techniques. The matrix A was generated using methods described in section 4.1. We multiplied both sides with A^T to make problem computationally feasible. We used pseudo inverse implementation in MATLAB to find the solution.

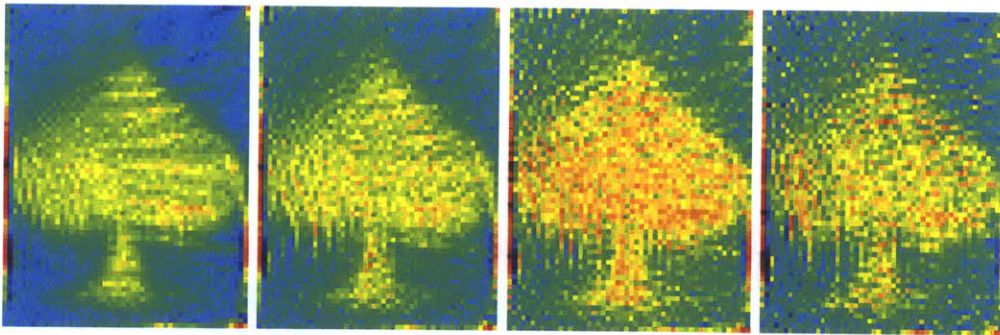


Figure 4-4: This figure shows how the reconstruction quality deteriorates by down sampling world voxels. (Left) Zero down-sampling (Middle) one third down-sampling (Right) One fifth down-sampling.

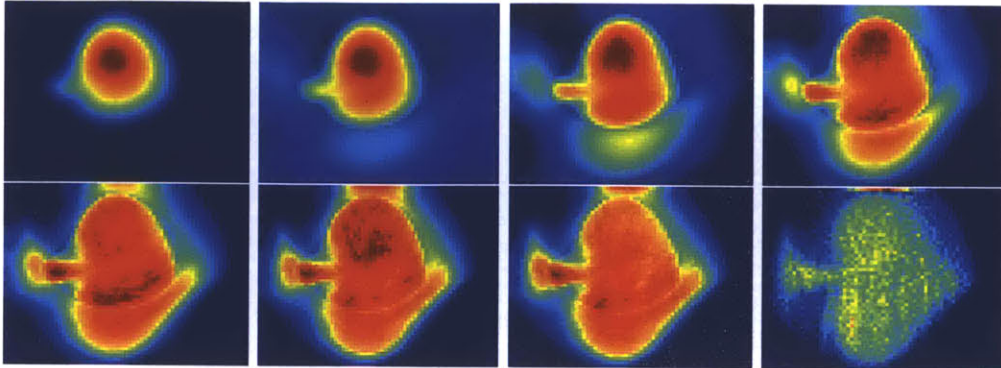


Figure 4-5: Above figure shows how spade evolves with every few iteration of SPGL. Notice that the interior points get picked first. The last image is result from 30 iterations.

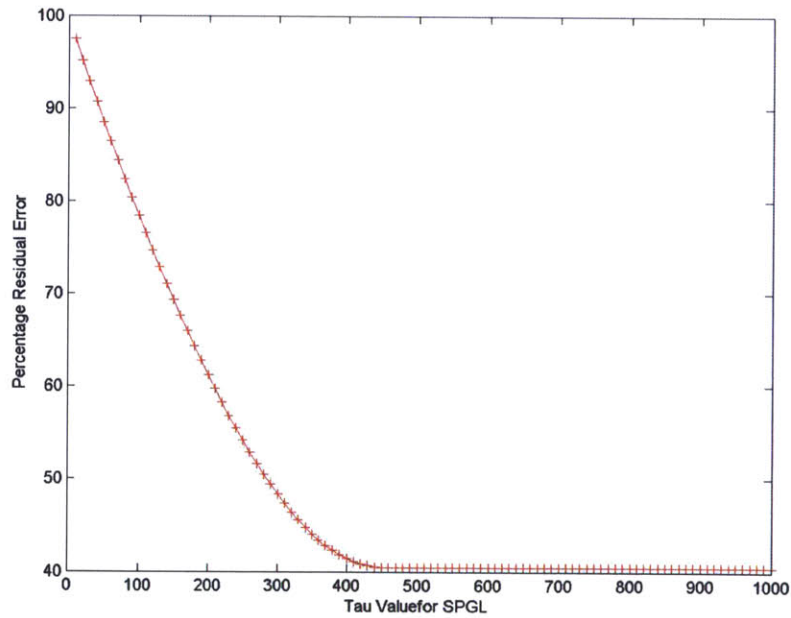


Figure 4-6: Methods like SPGL requiring only single parameter have an optimal parameter value which results in correct reconstruction. The above figure shows the variation in error with variation in tau. This is in fact the pareto curve described in [6] between one norm of solution and two norm of residual.

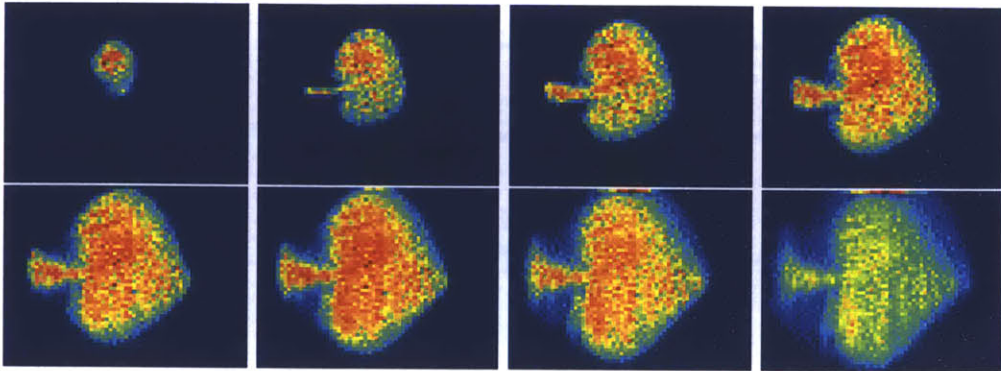


Figure 4-7: This figure shows how spade reconstruction changes with increasing tau. For lesser values of tau only interior points get selected and rest of shape remains unknown. As tau is increased the solution converges to a spade.

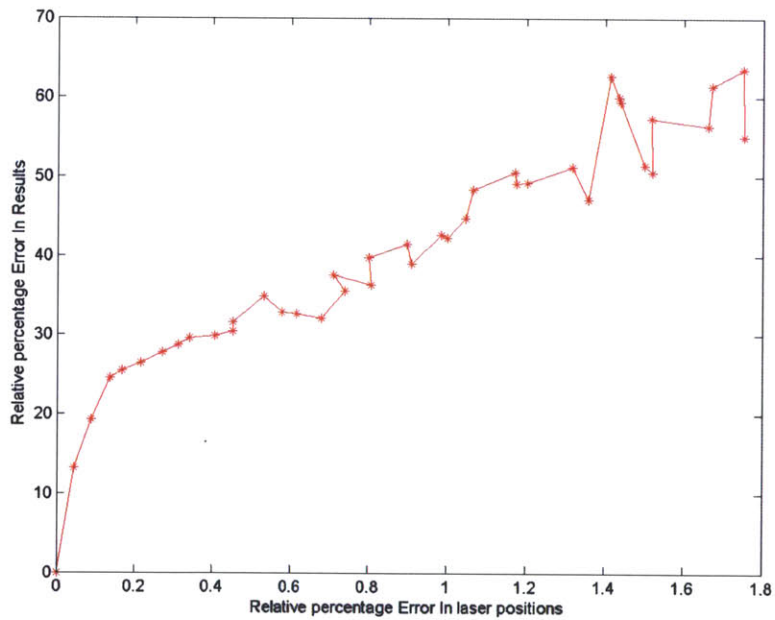


Figure 4-8: . Analysis of reconstruction if random noise is added to parameters. The reconstruction is pretty robust to random salt and pepper noise in laser locations.

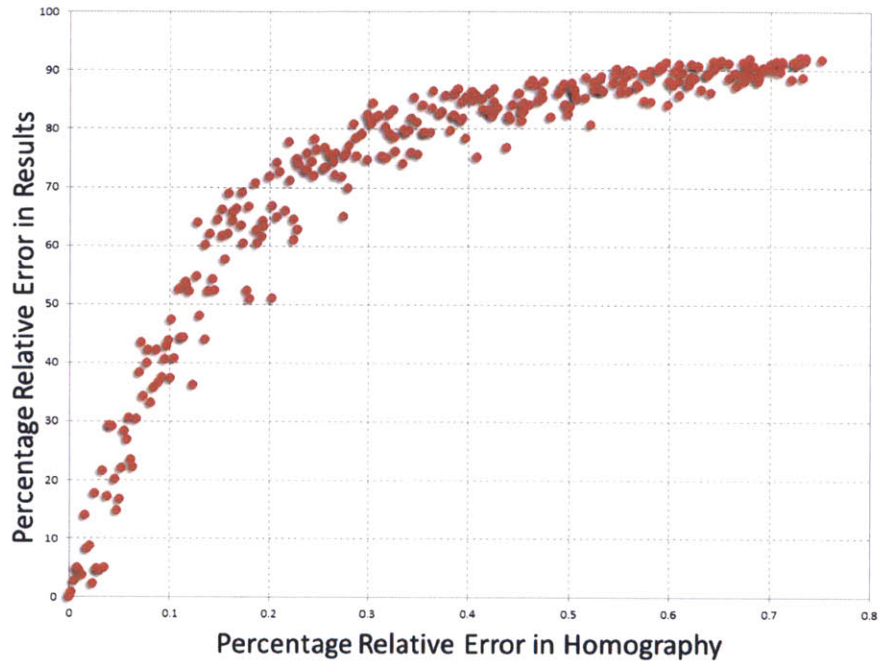


Figure 4-9: Random Salt and pepper noise in homography and its effect on reconstruction.

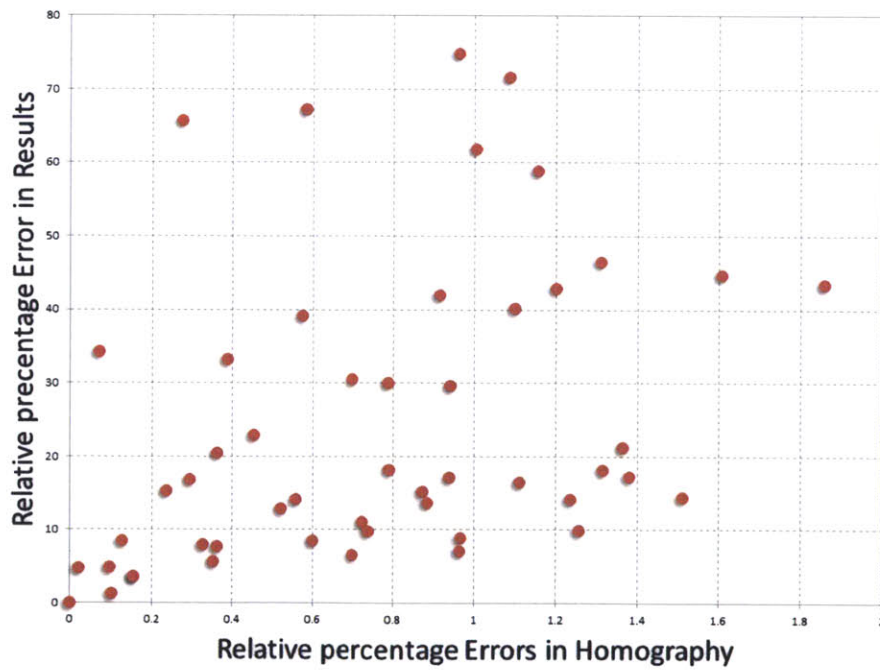


Figure 4-10: Effect on reconstruction with systemic noise.

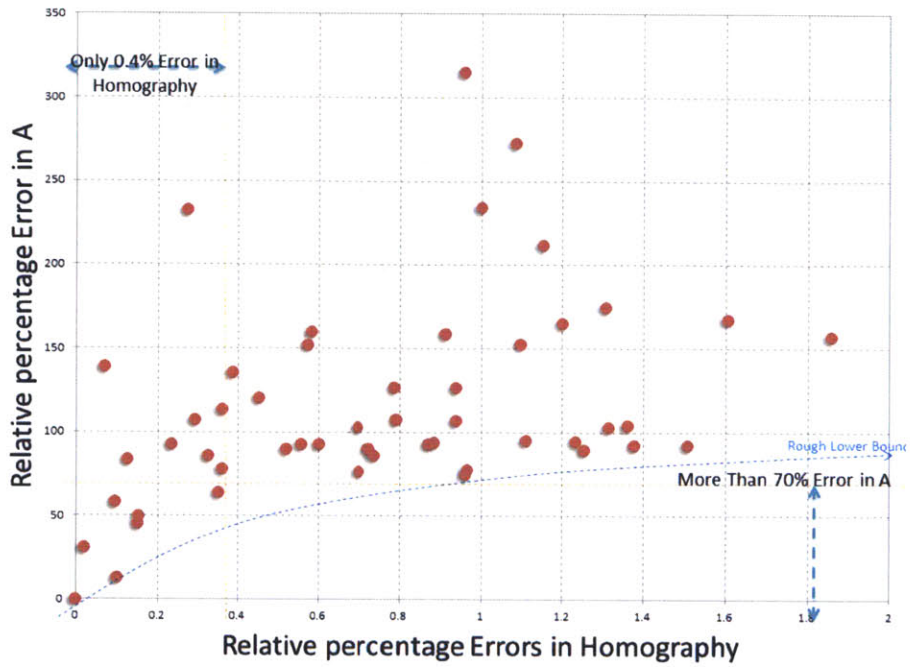


Figure 4-11: Variation in A with systemic noise. As we can see even small amount of homography error can change A significantly.

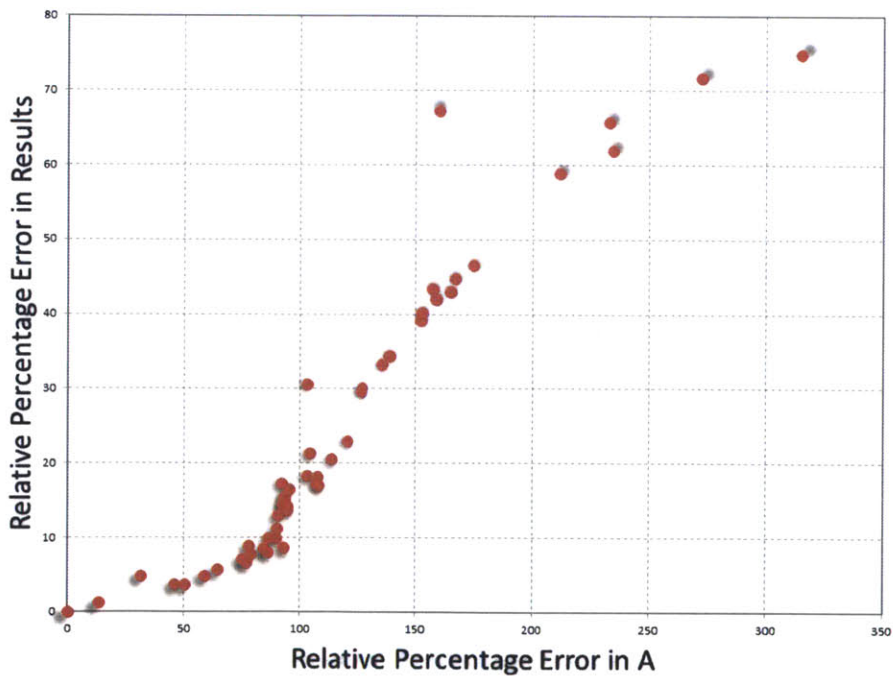


Figure 4-12: Effect on reconstruction with systemic noise in A.

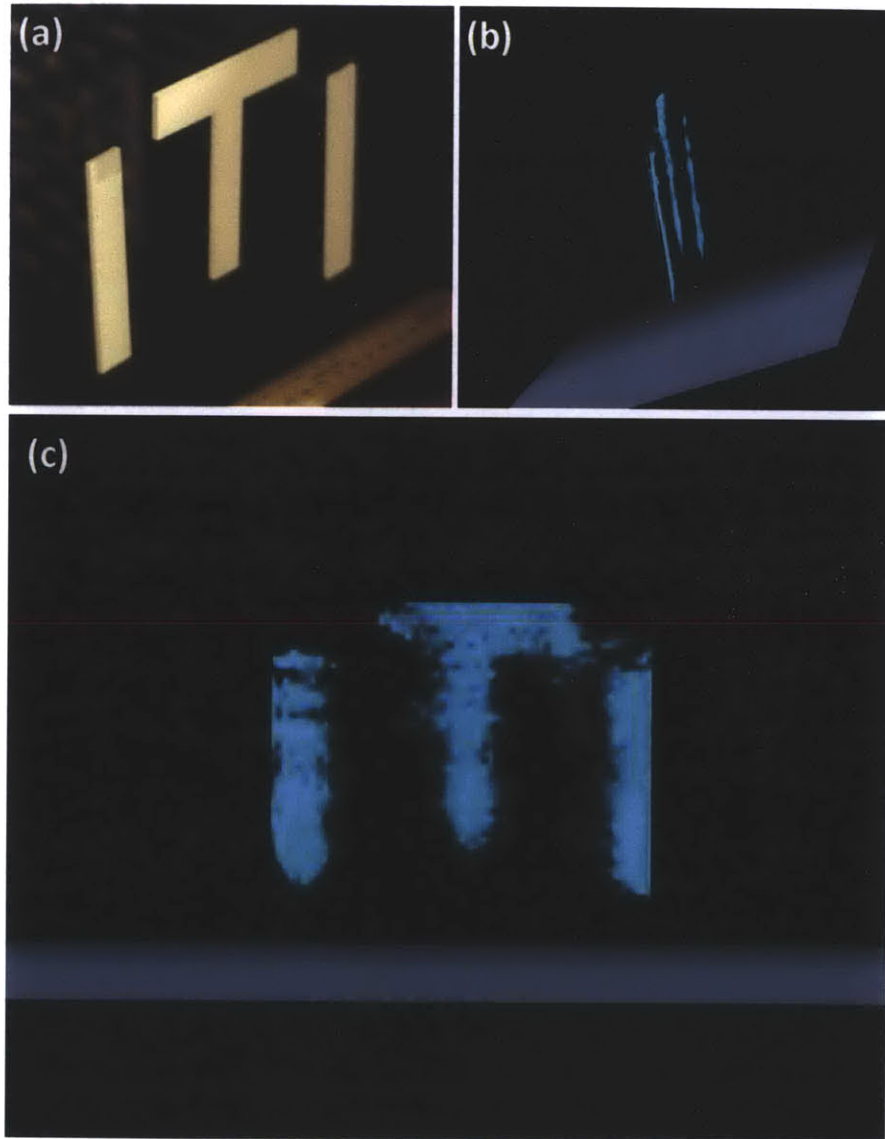


Figure 4-13: (Above) Hidden MIT around the corner (Below) Reconstruction using SPGL

Chapter 5

Experiments and Results

5.1 Hardware assembly

High quality time of flight cameras are now widely available [19, 1], but unfortunately a programmable capture experiment is still extremely challenging, very expensive and cumbersome. Below we describe our ambitious effort but also emphasize that commercial Lidar's can be repurposed for the same goal. For example, aerial Lidar systems use sub-nanonsecond pulses and use clever time-wavelength filtering to image over several kilometers. Unfortunately, it is difficult to program and get time profile out of them because they are geared for recording only the first bounce. Our streak camera is a Hamamatsu C5680, with an internal time resolution of 2 picoseconds. We use a mode-locked Ti:Sapphire laser to generate pulses at 795 nm wavelength with about 50 femtosecond pulse duration at a repetition rate of 75 MHz. The laser's average output power is about 500 mW. The streak camera has a one dimensional field of view, imaging a line in the scene. It provides a two dimensional image in which one dimension corresponds to space and the other to time [14]. (Joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41].)

We use a Faro Gauge measurement arm to calibrate the laser and camera. We treat the camera and laser as a rigid pair with known intrinsic and extrinsic parameter [12]. The visible parts of the geometry could be measured with the laser directly using

time of flight with micrometer precision. Methods like LiDAR and OCT can achieve this and are well understood. In the interest of focusing on the novel components of our system we instead measure the visible parts of our scene with the Faro Gauge. We also collect ground truth data to validate our reconstructions.

Synchronization of camera and laser time is challenging as it needs to be accurate to about 2 ps (≈ 0.6 mm path length). The camera shows drift, time scaling, time jitter and intensity variation intra and inter exposure time of the camera. This time-noise (location and intensity uncertainty along time dimension in streak photo) affects the point spread function and is difficult to represent in our forward model. We address this using a delayed reference beam and recording the first bounce on the wall (see Figure 5-11). We weaken this reference beam significantly via polarizers and neutral density filters as the camera dynamic range is chosen suitable for third bounces and the reference beam will easily saturate. The streak camera's photocathode tube, much like an oscilloscope, has time decayed burn out and local gain variations. We use a reference background photo to divide and compensate. Furthermore free space lasers have relatively poor pointing stability. Whenever the direction of the laser beam jumps, we have to recalibrate the system. (Data collection in collaboration with Andreas Velten, results in paper [41].)

Improving SNR: The laser emits a pulse every 13.3 nanoseconds (75 MHz) and consequently the reflected signal repeats at the same rate. We average 7.5 million such 13.3 nanosecond windows in a 100 ms exposure time on our streak camera readout camera. We add 50 to 200 such images to minimize noise from the readout camera. The light returned from a single hidden patch is attenuated only in the second and third path segment. In our setups this attenuation factor is about 10^{-9} corresponding to about 10 transmitted photons per pulse. The quantum efficiency of our camera is about 3 %. Like many LIDAR systems our camera thus receives less than 1 photon per pulse.

Limiting Intensity: 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. A

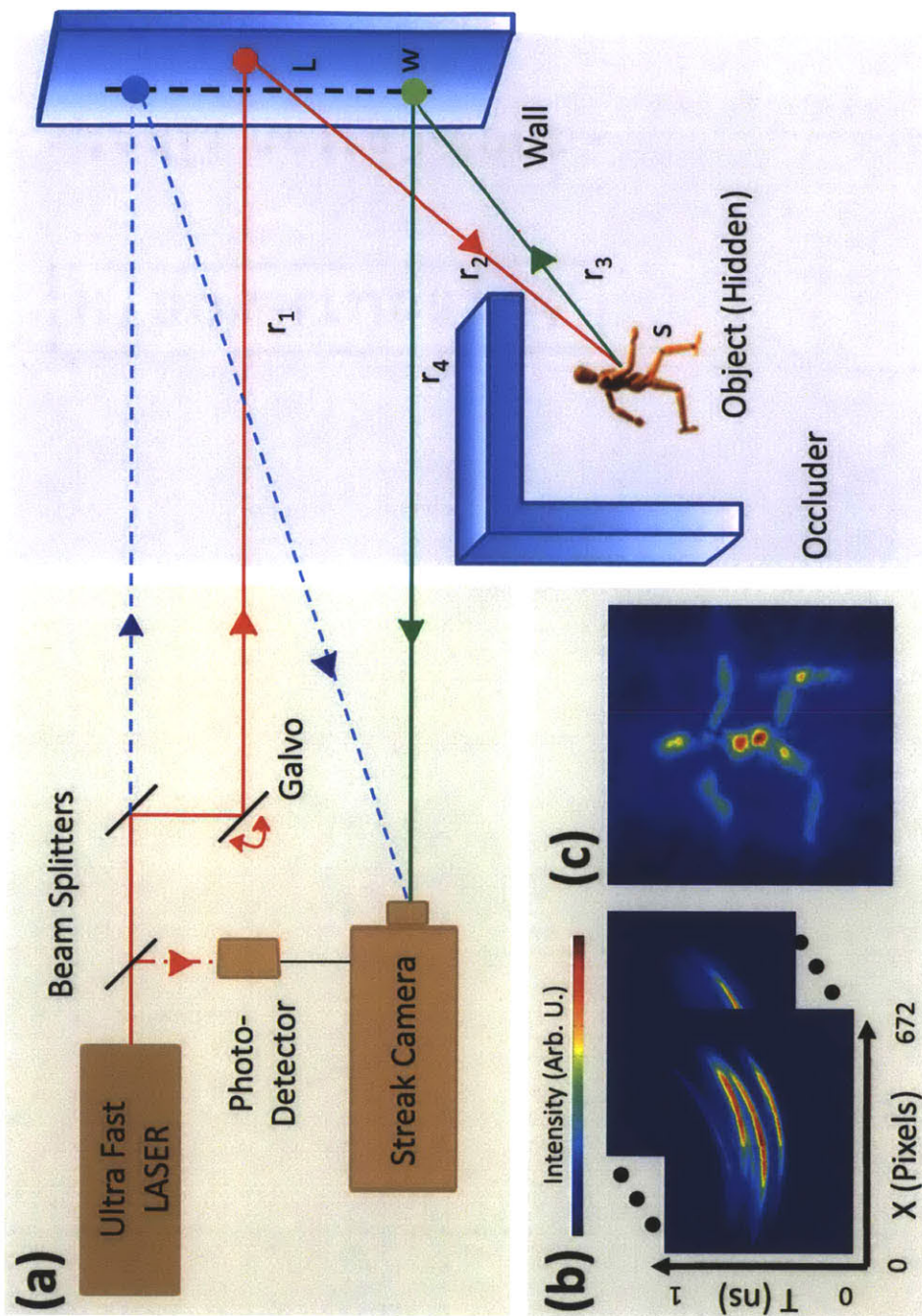


Figure 5-1: The experimental setup. (B). The laser pulses strike the wall at a point l and some of the scattered light light strikes the hidden object (e. g. at s), returns to the wall (w) and is collected by the camera (c). Both the galvo scanner and the camera are controlled by a computer. Done in collaboration with Andreas Velten, Thomas Willwacher, Ashok Veeraraghavan, Mounqi G. Bawendi and Ramesh Raskar [41].

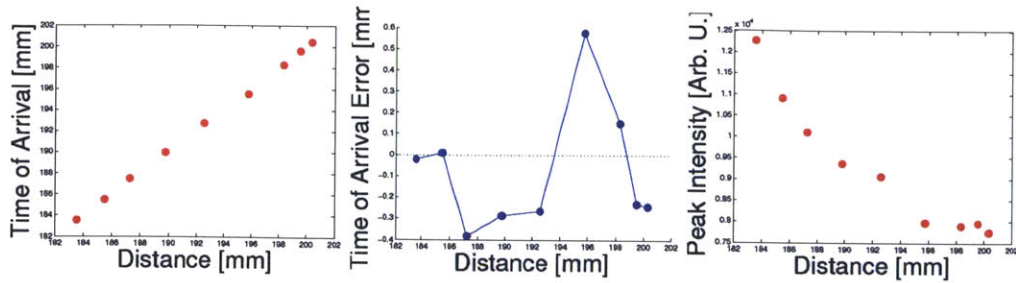


Figure 5-2: Resolution in depth. Left: Distance estimation. Time here is measured in mm of traveled distance at the speed of light $1 \text{ mm} \approx 0.3 \text{ ps}$. Middle: Error is less than 1 mm. Right: Plot of intensity as a small patch is moved perpendicular to the first surface.



Figure 5-3: Resolution in lateral dimension measured with a chirp pattern. (Left) Setup with chirp pattern (occluder removed in this photo) (Middle) Raw streak photo from streak camera (Right) The blue curve shows reconstruction of the geometry and indicates that we can recover features with 0.5 cm in lateral dimensions in the given scenario. The curve shows the confidence in the backprojected values.

large diameter however limits the time resolution of our technique. Using techniques in computational photography such as coded illumination, one can project a sequence of basis coded pattern that has reduced energy per unit area and yet sufficient SNR for received photo.

Effect of time jitter: One of the major challenges we face in capturing real datasets is temporal jitter and synchronization across different streak images. Recall that for each location of the laser spot, we record an independent streak image and the all the captured streak images are used together during reconstruction. Unfortunately, obtaining exact time synchronization (of the order of picoseconds) is extremely challenging in real data capture. Even with very careful calibration and synchronization, there was about 3-4 voxels (few nanoseconds) of temporal jitter across the

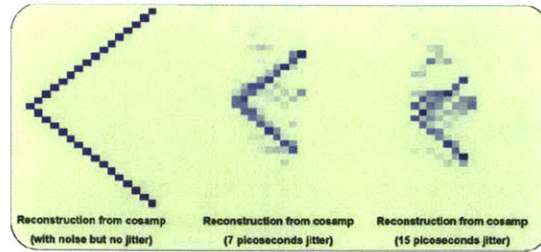


Figure 5-4: Performance degradation due to temporal jitter.

captured streak images. This temporal jitter implies that the captured streak images are not consistent with the scene geometry and each of the streak images have performed a different warping of the true underlying geometry. This significantly degrades performance of reconstruction and leads to severe artifacts and distortions in the reconstructions. This is the primary challenge in obtaining real-world results. Shown in Figure 5-4 are examples of a reconstruction of a wedge with and without jitter. Notice that while intensity noise does not affect performance significantly, even a small amount of temporal jitter causes significant performance degradation.

Limitation of 1D streak camera: Our experimental prototype system uses a 1D streak camera and the camera pixels are aligned with the x-axis of the world coordinate system. This effectively means that our resolution in the y-axis is severely compromised. As shown in Figure 5-7, the y resolution of a 1D streak camera is much lower than the x and z resolutions. This limitation manifests itself in the form of low y resolution in our real world reconstructions.

Validation: We validated the robustness of our approach for different linear distortions such as blurring, time jitter, saturation and partial occlusion. In synthetic experiments, we noticed that when SNR reduces below 10dB the reconstruction performance suffers.

5.1.1 Results

We recorded series of streak images for several simple 3D scenes, comprised of white Lambertian objects. We used 30-60 laser positions, spread over a 20 x 40 cm wall. The reconstructed surfaces are displayed in Figures 5-5, 5-9, 5-10. To produce the

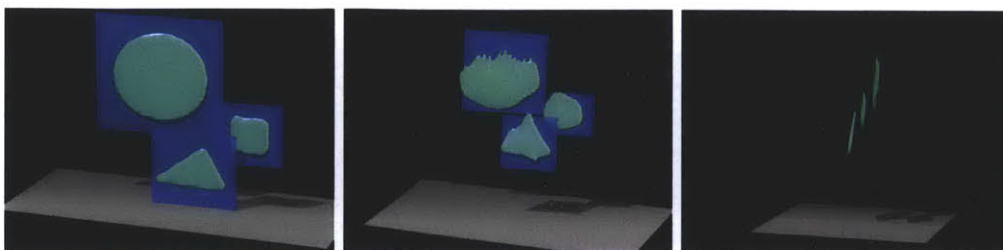


Figure 5-5: Reconstruction of a scene consisting of a big disk, a triangle and a square at different depth. (Left) Ground truth. (Middle) Reconstruction, front view. (Right) Reconstruction, side view. Note that the disk is only partially reconstructed, and the square is rounded off, while the triangle is recovered very well. This illustrates the diminishing resolution in directions parallel to the receiver plane towards the borders of the field of view. The blue planes indicate the ground truth. The gray ground planes and shadows have been added to help visualization. Referenced from joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounqi G. Bawendi and Dr. Ramesh Raskar [41].

3D pictures from the volumetric data, we use the Chimera visualization software [29] for thresholding and rendering. (Joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounqi G. Bawendi and Dr. Ramesh Raskar [41].)

5.1.2 Performance Evaluation

We conducted several experiments to test the performance of our experimental setup. This includes verifying the spatial and temporal resolution of the camera and the resolution obtained in a simple hidden scene (Fig. 5-3). (Data collection in collaboration with Andreas Velten, results in paper [41].)

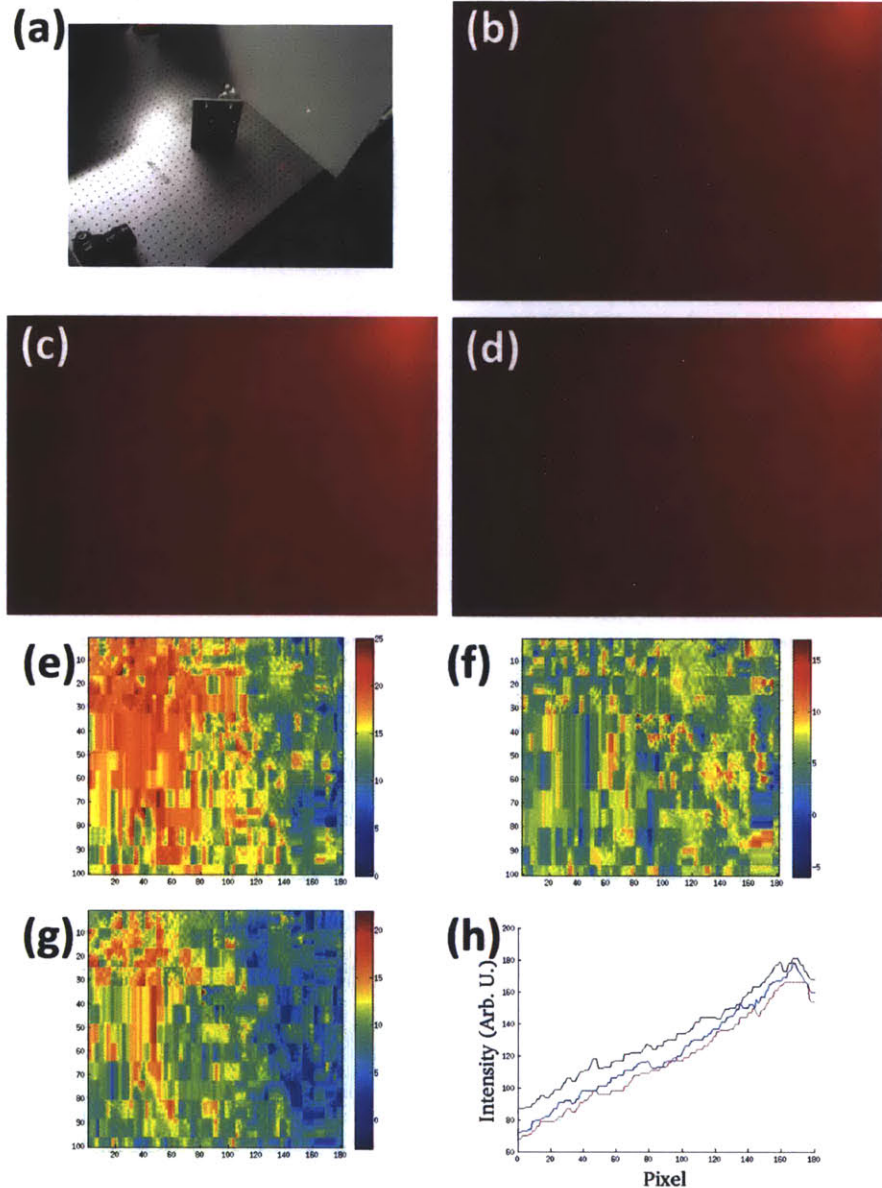


Figure 5-6: Challenges in imaging around the corner with a conventional, low temporal resolution laser and camera. (a) A setup with hidden mannequin but using a red continuous laser and a Canon 5D camera. (b) An image of the wall recorded with the Canon 5D camera with the room lights turned off and no hidden object present. (The recorded light is due to the reflections from walls behind the laser and camera.) (c) An image recorded with the hidden mannequin present. The increased light level on the wall is marginal, is low spatial frequency and shows no noticeable high frequency structure. (d) An image of the wall with the hidden mannequin moved away from the wall by 10 cm. The reduction in light level on the wall has no visible structure. (e) The difference between image in (b) and (c) using a false color map. (f) The difference between (b) and (d). (g) The difference between (c) and (d). (h) The plot of intensities along the centered horizontal scanline of each of the images (b=red, c=black, d=blue).

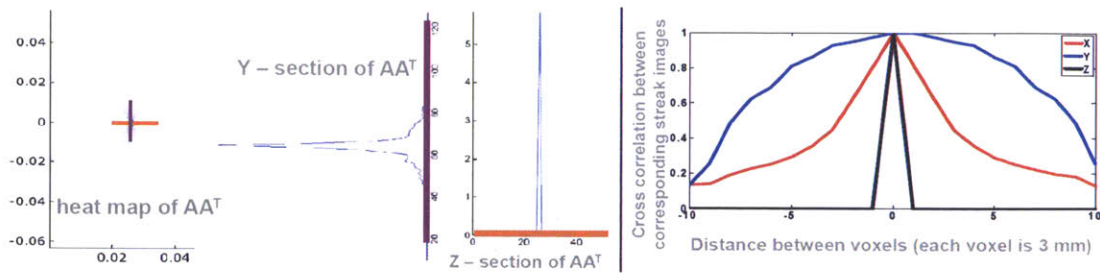


Figure 5-7: Limitations of heatmap after the backprojection algorithm. (Left) Propagation of a single point and its backpropagation for a flatland case. Reconstruction using backpropagation shows that one can recover a sharp peak, but it is surrounded by a low frequency residual. Y and Z sections are shown. The resolution in Y is lower than the resolution in Z. (Right) Cross Correlation of the streak images corresponding to nearby voxels (3D case with 1D streak image). Notice that the Y-resolution is worse because the sensor is 1D along the x-axis. Results from joint work with Andreas Velten in [41].

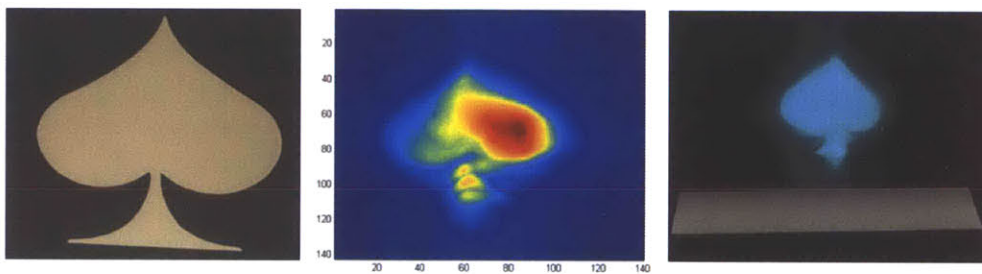


Figure 5-8: Reconstruction of a planar object in an unknown plane in 3D. (Left) The object. (Middle Left) 2D Projection of the filtered heatmap. (Middle Right) A 3D visualization of the filtered heatmap. (Right) Reconstruction using sparsity based methods. The gray ground plane has been added to aid visualization. Referenced from joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41].



Figure 5-9: Reconstruction of a wooden man, painted white. Center - reconstruction using simple back projection based methods. Right - reconstruction using sparse reconstruction methods. Results from joint work with Andreas Velten in [41].

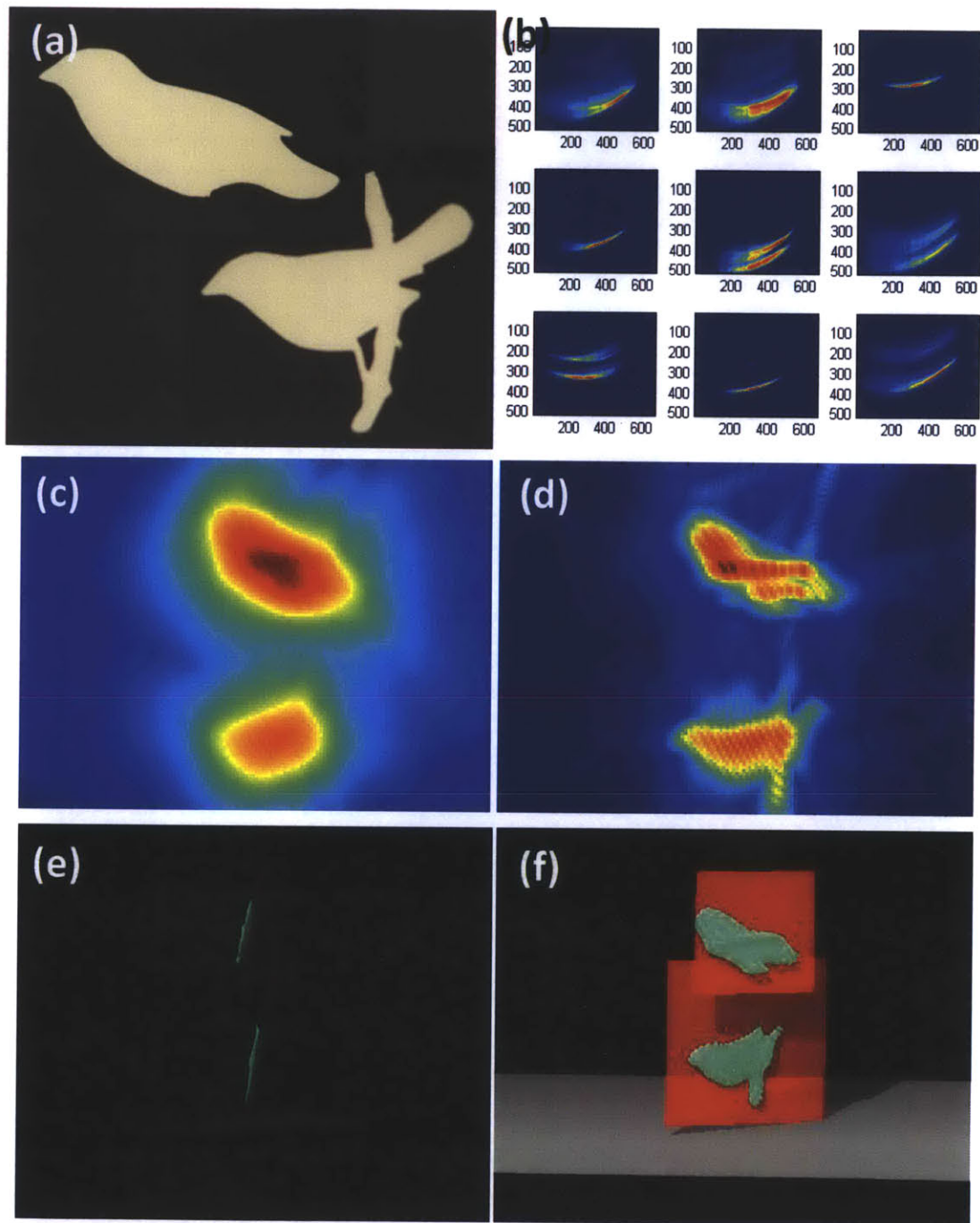


Figure 5-10: Depiction of our reconstruction algorithm for a scene consisting of two birds in different planes. (a) Photographs of the input models. (b) 9 out of 33 *streak images* used for reconstruction. (c) The raw (unfiltered) *backprojection*. (d) The *filtered* backprojection, after taking a second derivative.(e,f) 3D renderings in Chimera.

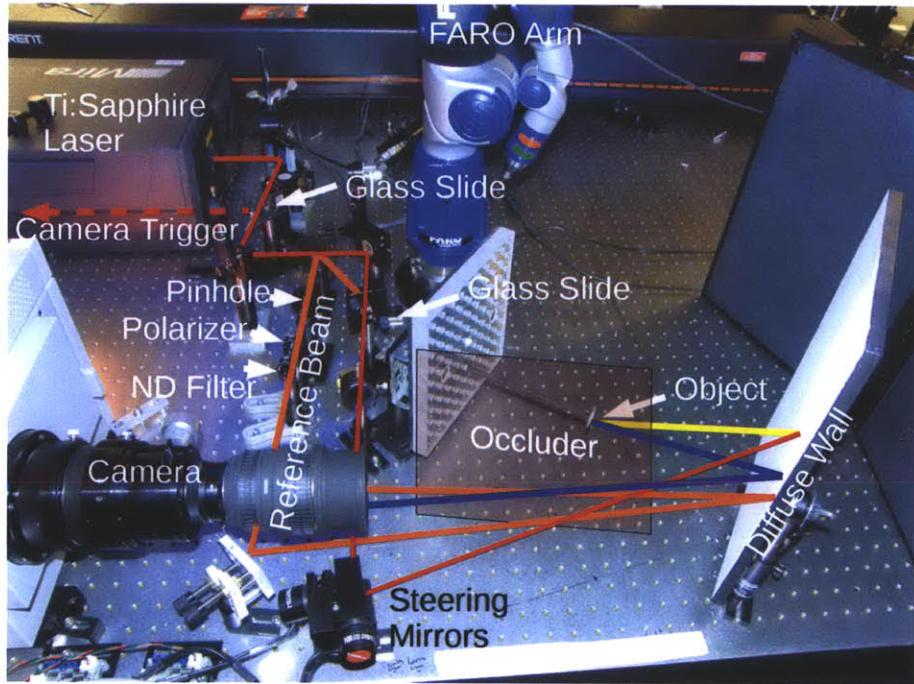


Figure 5-11: The laser beam (red) is split to provide a synchronization signal for the camera (dotted red) and an attenuated reference pulse (orange) to compensate for synchronization drifts and laser intensity fluctuations. The main laser beam is directed to a wall with a steering mirror and the returned third bounce is captured by the streak camera. An Occluder inserted at the indicated position does not significantly change the collected image. Referenced from joint work with Andreas Velten in [41].

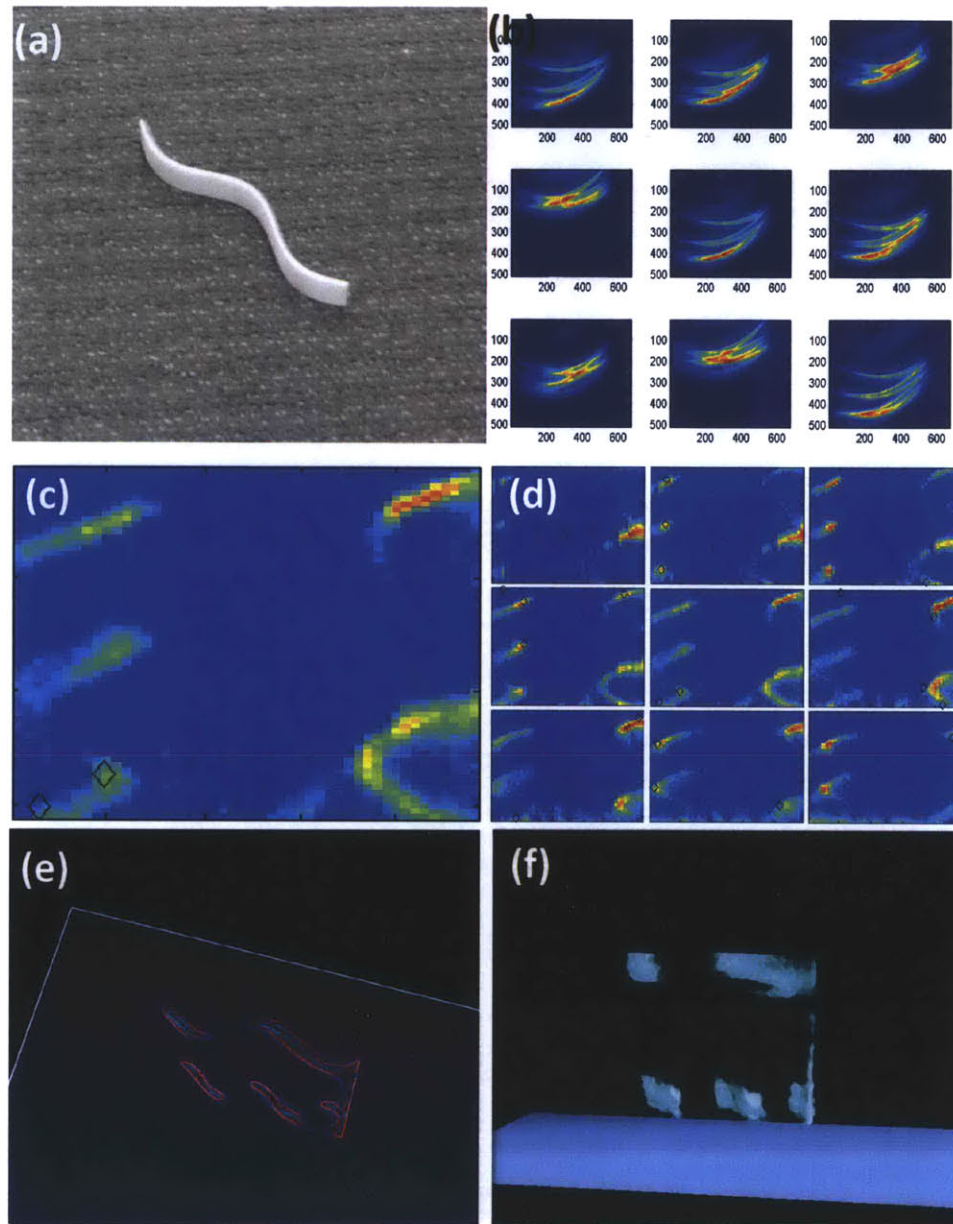


Figure 5-12: **Smooth Object Reconstruction** Reconstruction of two sinusoids using SPGL1. (a) Photo of the object. The sinusoid is approximately 1.5 cm tall with total length of 5 cm. (b) Nine of the 60 raw streak images. (c) *Side View*. Visualization of one cross sections from reconstruction. (d) *3D cross-sections*. Chosen cross sections along depth (z) projected on the x - y plane reveals the hidden shape contour. (e) *Top View*. (f) *Confidence map*. A rendered point cloud of reconstruction values after soft threshold.

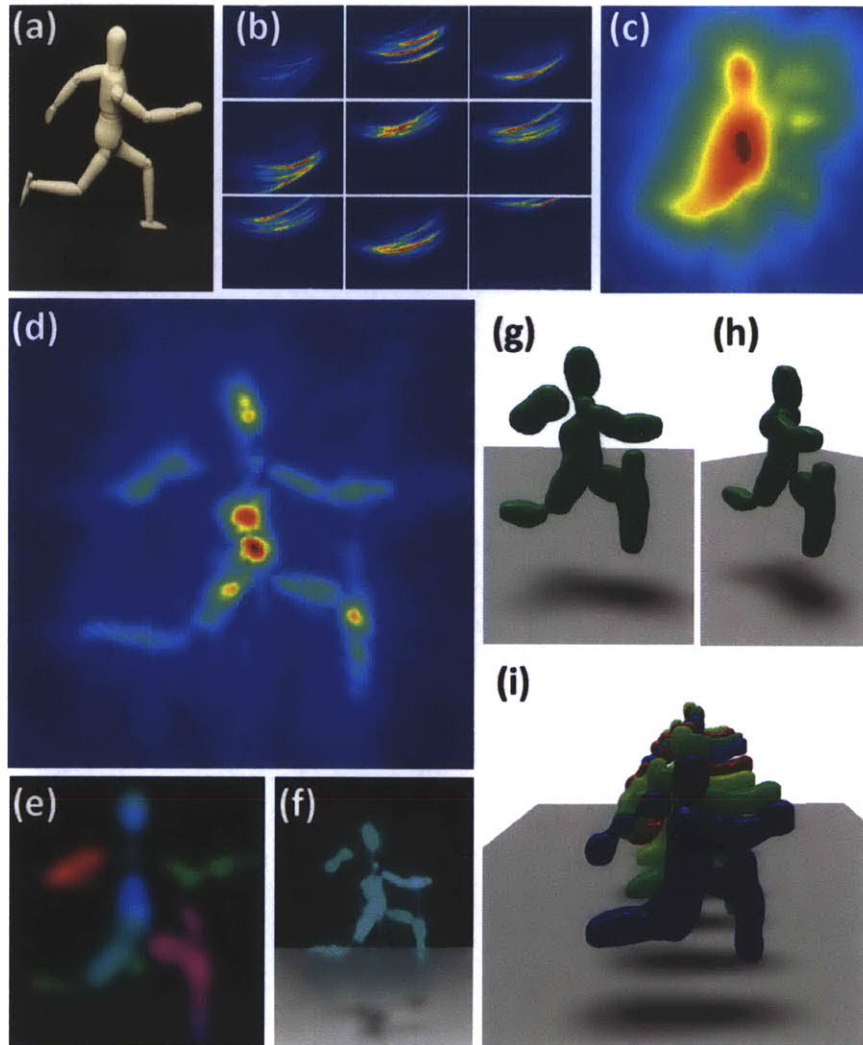


Figure 5-13: **Complex Object Reconstruction** in multiple poses. (a) Photo of the object. The mannequin is approximately 20 cm tall and is placed about 25 cm from the diffuser wall. (b) Nine of the 60 raw streak images. (c) *Heatmap*. Visualization of the heatmap after backprojection. The hidden shape is barely discernible. (d) *Filtering*. The second derivative of the heatmap along depth (z) projected on the x - y plane reveals the hidden shape contour. (e) *Depth map*. Color encoded depth shows the left leg and right arm closer in depth compared to the torso and other leg and arm. (f) *Confidence map*. A rendered point cloud of confidence values after soft threshold. (i) The stop-motion animation frames from multiple poses to demonstrate reproducibility. Shadows and the ground plane in images (f-i) have been added to aid visualization. Joint work with Andreas Velten, Thomas Willwacher, Ashok Veeraraghavan, Mounsi G. Bawendi and Ramesh Raskar [41].

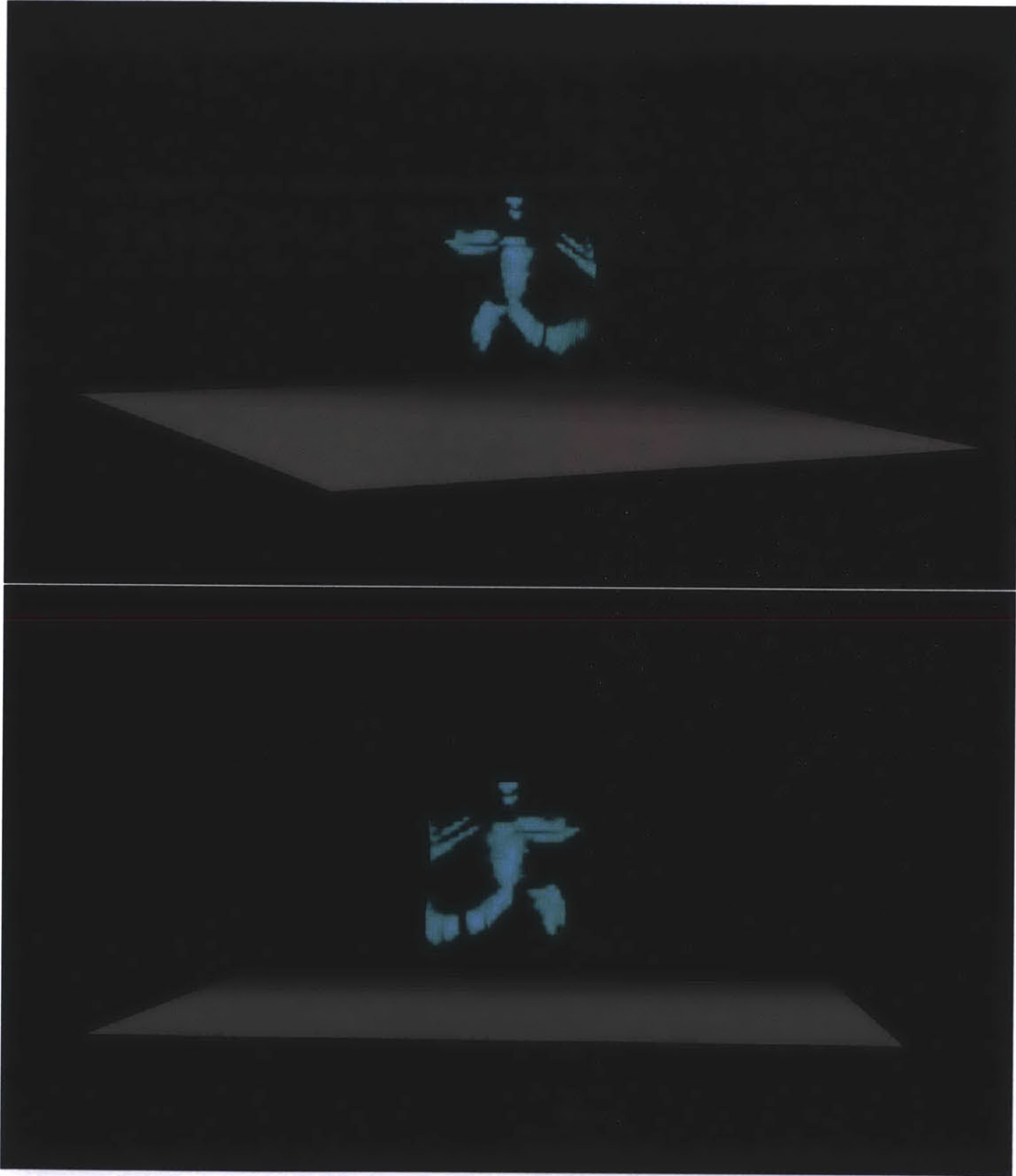


Figure 5-14: Reconstruction of a hidden wood mannequin using sparse reconstruction techniques

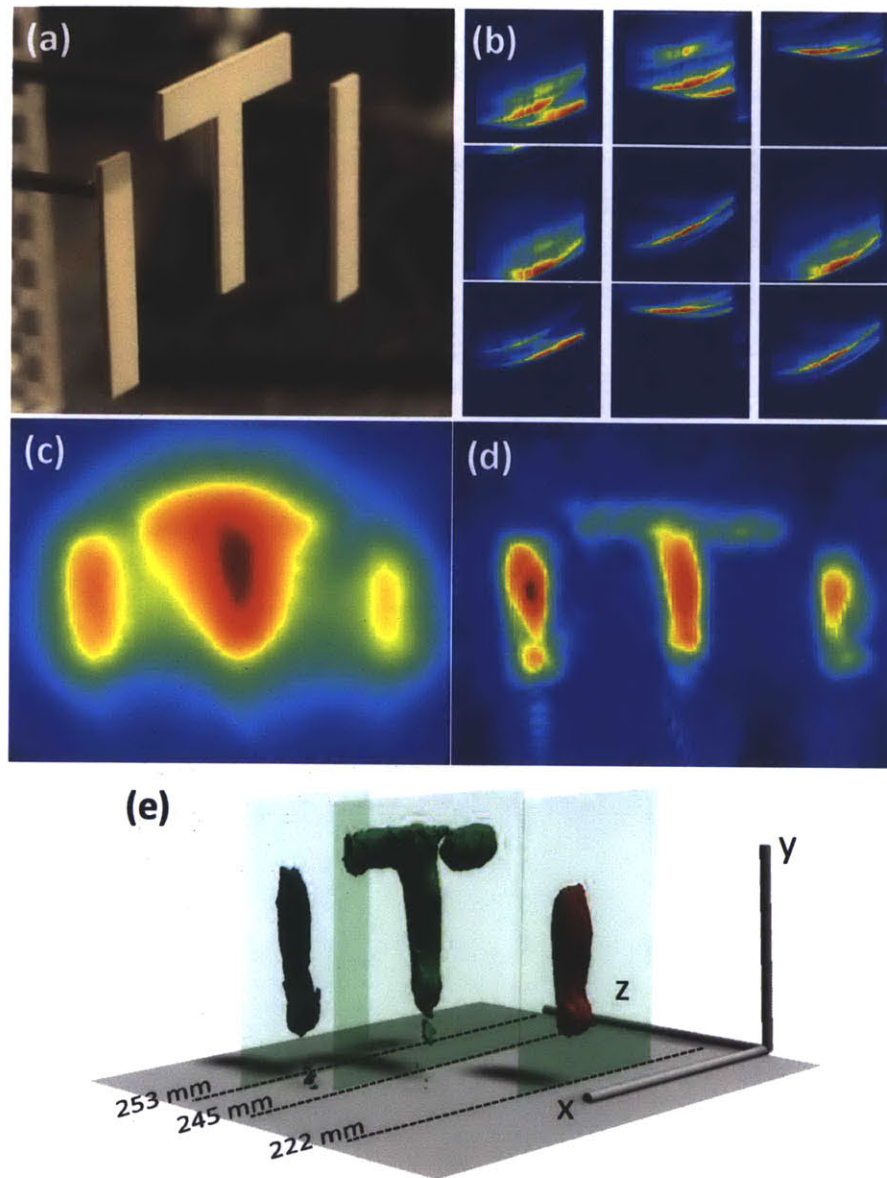


Figure 5-15: Demonstration of the depth and lateral resolution. (a) The hidden object to be recovered are three letters, I, T, I at varying depths. The "I" is 1.5 cm in wide and all letters are 8.2 cm high. (b) 9 of 60 images collected by the streak camera. (c) Projection of the heatmap on the x-y plane created by the back projection algorithm. (d) Filtering after computing second derivative along depth (z). The color in these images represents the confidence of finding an object at the pixel position. (e) A rendering of the reconstructed 3D shape. Depth is color coded and semi-transparent planes are inserted to indicate the ground truth. The depth axis is scaled to aid visualization of the depth resolution. Referenced from joint work with Dr. Andreas Velten, Dr. Thomas Willwacher, Dr. Ashok Veeraraghavan, Dr. Mounji G. Bawendi and Dr. Ramesh Raskar [41].

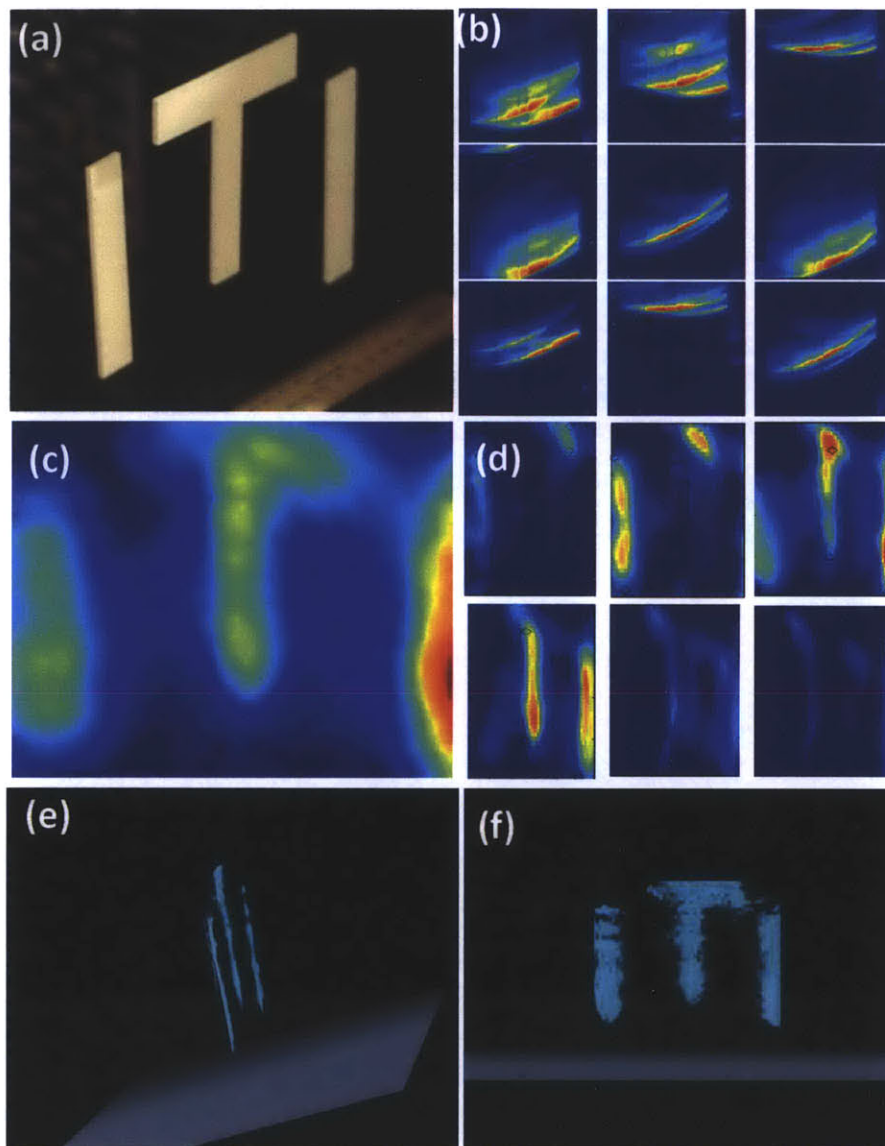


Figure 5-16: Reconstruction of hidden ITI(MIT) using sparse reconstruction techniques such as SPGL. Note that this result is from a technique entirely different from 5-15. (a) The hidden MIT. (b) Selected streak images. (c) Reconstructed volume with all sections merged into one image. (d) Volumetric slices of reconstructed hidden object. (e) The final result rendered in 3D (side view). Notice how accurately depth gets recovered. (f) Rendered MIT from front.

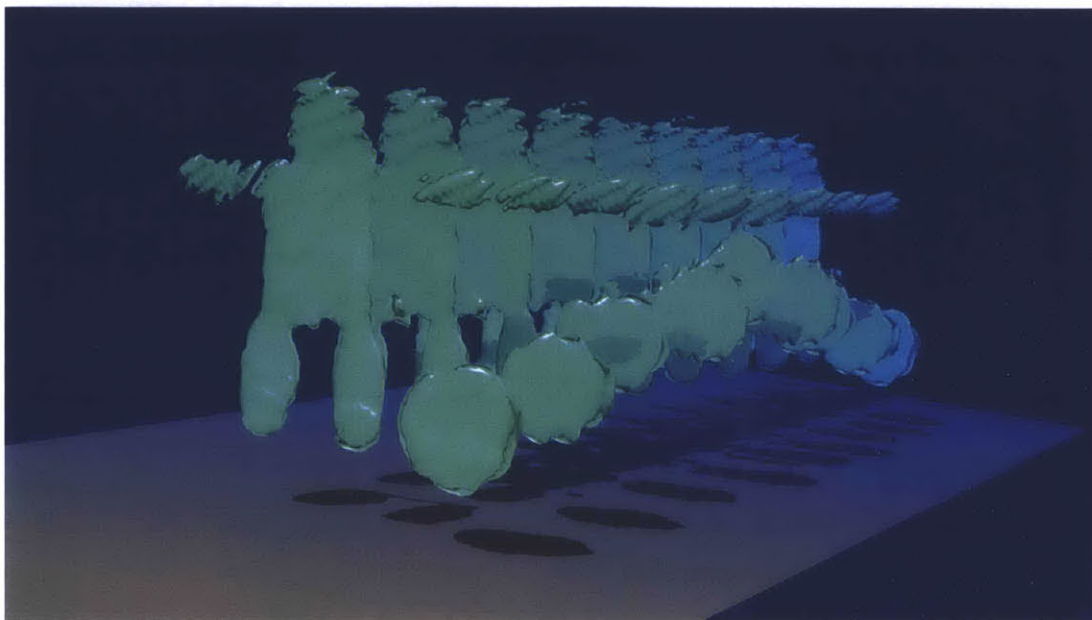


Figure 5-17: Results of a multi-pose stop motion animation dataset after filtered backprojection and soft-thresholding. A hidden model of a *man with a ball* is captured in various poses. The rendering shows the sequence of reconstructions created by our filtered backprojection algorithm and demonstrates the ability to remove low-frequency artifacts of backprojection. The mislabeled voxels remain consistent across different poses indicating stability of our capture and inversion process. Shadows are introduced to aid visualization. Results from joint work with Andreas Velten in [41].

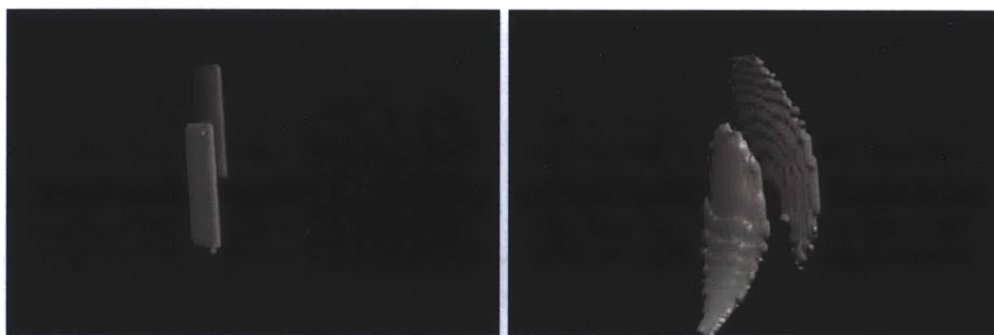


Figure 5-18: Results from reconstruction of two cards using simple backprojection techniques for real world data. (Left) The actual object. (Right) Reconstructed object.



Figure 5-19: Results from reconstruction of a wedge using simple backprojection techniques, and carving in real world data. (Left) The actual object. (Right) Reconstructed object.

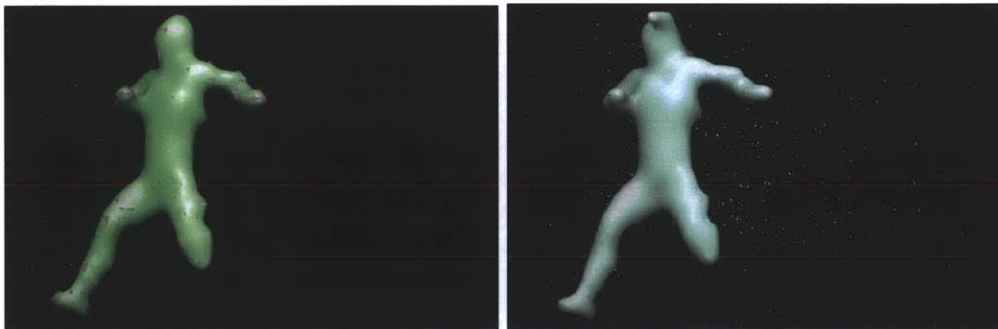


Figure 5-20: Results from reconstruction of a man using CoSAMP in simulation. (Left) The actual object. (Right) Reconstructed object.



Figure 5-21: Results from reconstruction of a man using CoSAMP in simulation. (Left) The actual object. (Right) Reconstructed object.

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that this is crucial for ensuring transparency and accountability in the organization's operations.

2. The second part of the document outlines the various methods and tools used to collect and analyze data. It highlights the need for consistent and reliable data collection processes to support effective decision-making.

3. The final part of the document provides a summary of the key findings and recommendations. It stresses the importance of ongoing monitoring and evaluation to ensure that the organization remains on track with its strategic goals.

Chapter 6

Future Directions

We have shown that the goal of recovering hidden shapes is only as challenging as the current hardware. The computational approaches show great promise. But, on the hardware front, emerging integrated solid state lasers, new sensors and non-linear optics will provide practical and portable imaging devices. Our formulation is also valid for shorter wavelengths (e.g., x-rays) or for ultrasound and sonar frequencies in large scenes where diffraction can be neglected. Beyond geometry, one maybe able to recover full light transport and bidirectional reflectance distribution function (BRDF) from a single viewpoint to eliminate encircling instrumentation. Our current method assumes friendly reflectances, i.e., a non-zero diffuse component towards the wall. Non-lambertian reflectance and partial occlusions will create non-uniform angular radiance. But one can improve the carving to model the non-uniformity and exploit the specular peaks which will actually have a higher SNR. The visible wall need not be planar and one can update the r_l and r_c distances from a known model of the visible parts. Supporting refraction involves multiple or continuous change in path vector. Acheson et. al. have shown a refraction tomography approach [5] which could be extended.

In the future, emerging integrated solid state lasers, new sensors and non-linear optics should provide practical and more sensitive imaging devices. Beyond 3D shape, new techniques should allow us to recover reflectance, refraction and scattering properties and achieve wavelength resolved spectroscopy beyond the line of sight. The

formulation could also be extended to shorter wavelengths (e.g., x-rays) or to ultrasound and sonar frequencies. The new goal of hidden 3D shape recovery may inspire new research in the design of future ultrafast imaging systems and novel algorithms for hidden scene reconstruction.

Initial applications may be in controlled settings like endoscopy, scientific imaging and industrial vision. This will require addressing more complex transport for volumetric scattering (e.g. for tissue) or refracting elements (e.g. for fluids). A very promising theoretical direction is in inference and inversion techniques that exploit scene priors, sparsity, rank, meaningful transforms and achieve bounded approximations. Adaptive sampling can decide the next-best laser direction based on current estimate of the carved hull. Future analysis will include coded sampling using compressive techniques and noise models for SNR and effective bandwidth.

We used the COSAMP [26] matching pursuit algorithm which allows us to explore the sparsity of the solution. We tested both the backprojection and linear equation based methods on both artificially generated and real data. It turned out that for artificial data the COSAMP based reconstruction algorithm was generally superior the backprojection algorithm. The backprojection algorithm can recover objects front-to-parallel to the wall quite well, but fails for highly sloped surfaces. However, for real data the linear equation based methods were very sensitive to calibration errors, i.e., errors in the matrix A above. One can obtain results for very good datasets, after changing A slightly to account for intricacies of our imaging system like vignetting and gain correction. On the other hand the backprojection algorithm turned out to be quite robust to calibration errors, though they can deteriorate the obtained resolution. Since we typically have to deal with some calibration error (see section 5.1), we used the backprojection algorithm to obtain the real data reconstruction results of this paper.

6.1 Conclusion

The ability to infer shapes of objects beyond the line of sight is an ambitious goal but it may transform recoding of visual information and will require a new set of algorithms for scene understanding, rendering and visualization. We have presented a new shape-from-x approach that goes beyond the abilities of today's model acquisition and scanning methods. Light transport with a time component presents a unique challenge due to the lack of correspondence but also provides a new opportunity. The emphasis in this paper is to present a forward model and novel inversion process. The nonlinear component due to jitter and system point spread function makes the ultrafast imaging equipment difficult to use. So the physical results can only be treated as a proof-of-concept.

One may wonder when ultrafast lasers and cameras will be broadly available to researchers in computer graphics and computational photography. Many lasers have transformed from their unsafe, bulky form factors to use in portable consumer devices. To further the research in this field by the community, we will make image datasets and matlab code freely available online.

The utility of ultrafast imagers like the streak camera has been limited to analysis of bio-chemical processes making them difficult to use for free space light transport. We hope our work will spur more applications of these imagers in computational photography and in turn the graphics research will influence the design and cost of future streak cameras.

Bibliography

- [1] Sick laser scanner: <http://www.sick.com/gus/products/new/s300/en.html>.
- [2] Tadar in action. <http://www.physorg.com/news7210.html>, October 2005.
- [3] Lidar overview. http://forsys.cfr.washington.edu/JFSP06/lidar_technology.htm, 2006.
- [4] E.H. Adelson and J.R. Bergen. The plenoptic function and the elements of early vision. *Computational Models of Visual Processing*, MIT Press, pages 3–20, 1991.
- [5] B. Atcheson, I. Ihrke, W. Heidrich, A. Tevs, D. Bradley, M. Magnor, and H.P. Seidel. Time-resolved 3d capture of non-stationary gas flows. *ACM Transactions on Graphics (TOG)*, 27(5):1–9, 2008.
- [6] Ewout Van Den Berg and Michael P. Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal of Scientific Computing*, 2008.
- [7] M. Bertero and P. Boccacci. Institute of Physics Publishing Ltd., 1998.
- [8] J. Busck and H. Heiselberg. Gated viewing and high-accuracy three-dimensional laser radar. *Applied Optics*, 43(24):4705–4710, 2004.
- [9] D. Colton and A. Kirsch. A simple method for solving inverse scattering problems in the resonance region. *Inverse problems*, 12:383, 1996.
- [10] D.L. Colton and R. Kress. *Inverse acoustic and electromagnetic scattering theory*. Springer Verlag, 1998.

- [11] Wotao Yin Elaine T. Hale and Yin Zhang. A fixed-point continuation method for l_1 -regularized minimization with applications to compressed sensing. *CAAM Technical Report TR07-07*, 2007.
- [12] D. Forsyth and J. Ponce. *Computer Vision, A Modern Approach*. Prentice Hall, 2002.
- [13] Dr Takashi Fujii and Tetsuo Fukuchi. Laser remote sensing. page 888, Jan 2005.
- [14] Hamamatsu. Hamamatsu Streak Camera, 2012.
- [15] Richard P. Hodges. Underwater acoustics: Analysis, design and performance of sonar. page 366, Jan 2010.
- [16] G. W. Kamerman. Laser Radar [M]. Chapter 1 of Active Electro-Optical System, Vol. 6, The Infrared and Electro-Optical System Handbook, 1993.
- [17] A. Kirmani, T. Hutchison, J. Davis, and R. Raskar. Looking around the corner using transient imaging. In *ICCV*, 2009.
- [18] A. Kirsch. Characterization of the shape of a scattering obstacle using the spectral data of the far field operator. *Inverse problems*, 14:1489, 1998.
- [19] A. Kolb, E. Barth, R. Koch, and R. Larsen. Time-of-flight sensors in computer graphics. *Eurographics State of the Art Reports*, pages 119–134, 2009.
- [20] R. Kress and L. Paivarinta. On the far field in obstacle scattering. *SIAM Journal on Applied Mathematics*, 59(4):1413–1426, 1999.
- [21] S. Liu, T.T. Ng, and Y. Matsushita. Shape from Second-Bounce of Light Transport. *ECCV 2010*, pages 280–293, 2010.
- [22] Santos J.M. M. Lustig, D. Donoho and J. Pauly. Compressed sensing mri. *IEEE Signal Processing MRI*, 2008.
- [23] R Mailloux. Phased array antenna handbook. *Boston*, Jan 1994.

- [24] Nikhil Naik, Shuang Zhao, Andreas Velten, Ramesh Raskar, and Kavita Bala. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. *ACM Trans. Graph.*, 30(6):171:1–171:10, December 2011.
- [25] Shree K. Nayar, Gurunandan Krishnan, Michael D. Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph.*, 25(3):935–944, July 2006.
- [26] D. Needell and J.A. Tropp. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, 26(3):301–321, 2009.
- [27] R. Pandharkar, A. Velten, A. Bardagjy, E. Lawson, M. Bawendi, and R. Raskar. Estimating motion and size of moving non-line-of-sight objects in cluttered environments. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 265 –272, june 2011.
- [28] YC Pati, R. Rezaifar, and PS Krishnaprasad. Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In *Signals, Systems and Computers*, pages 40–44. IEEE, 2002.
- [29] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, and T.E. Ferrin. UCSF Chimera-a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.
- [30] R. Potthast. A fast new method to solve inverse scattering problems. *Inverse Problems*, 12:731, 1996.
- [31] Ramesh Raskar. Looking around corners: New opportunities in femto-photography. *ICCP Conference at CMU, Invited Talk*, April 2011.
- [32] Ramesh Raskar and James Davis. 5d time-light transport matrix: What can we reason about scene properties? *MIT Technical Report*, 2008.

- [33] D. L. Donoho S. S. Chen and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal of Scientific Computing*, 1998.
- [34] S. M. Seitz, Y. Matsushita, and K. N. Kutulakos. A theory of inverse light transport. In *Proc. Tenth IEEE International Conference on Computer Vision ICCV 2005*, volume 2, pages 1440–1447, 17–21 Oct. 2005.
- [35] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch. Dual photography. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers*, pages 745–755, New York, NY, USA, 2005. ACM.
- [36] Merrill I. Skolnik. Introduction to radar systems. page 772, Jan 2002.
- [37] R.H. Stolt. A prestack residual time migration operator. *Geophysics*, 61:605, 1996.
- [38] A. Sume, M. Gustafsson, A. Jnis, S. Nilsson, J. Rahm, and A. rbom. Radar detection of moving objects around corners. In *Proc. of SPIE*, 2009.
- [39] J.A. Tropp and A.C. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 2007.
- [40] Andreas Velten, Everett Lawson, Andrew Bardagjy, Mounqi Bawendi, and Ramesh Raskar. Slow art with a trillion frames per second camera. In *ACM SIGGRAPH 2011 Posters*, SIGGRAPH '11, pages 13:1–13:1, New York, NY, USA, 2011. ACM.
- [41] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Mounqi G. Bawendi, and Ramesh Raskar. Recovering Three-Dimensional Shape around a Corner using Ultra-Fast Time-of-Flight Imaging. *Nature Communucations*, 2012.
- [42] Juefu Wang, Xishuo Wang, and Mike Perz. Structure preserving regularization for sparse deconvolution. *SEG Technical Program Expanded Abstracts*, 25(1):2072–2076, 2006.

- [43] L. Wang and G.K. Beare. Breaking the Optical Diffusion Limit: Photoacoustic Tomography. In *Frontiers in Optics*. Optical Society of America, 2010.
- [44] Wikipedia. Geophysical migration <http://en.wikipedia.org/wiki/geophysicalmigration>, 2010.
- [45] D. Wu, M. O'Toole, A. Velten, Agrawal A., and R. Raskar. Decomposing global light transport using time of flight imaging. In *Computer Vision and Pattern Recognition, 2012. CVPR 2012. IEEE Conference on*, pages 1–8. Ieee, 2012.