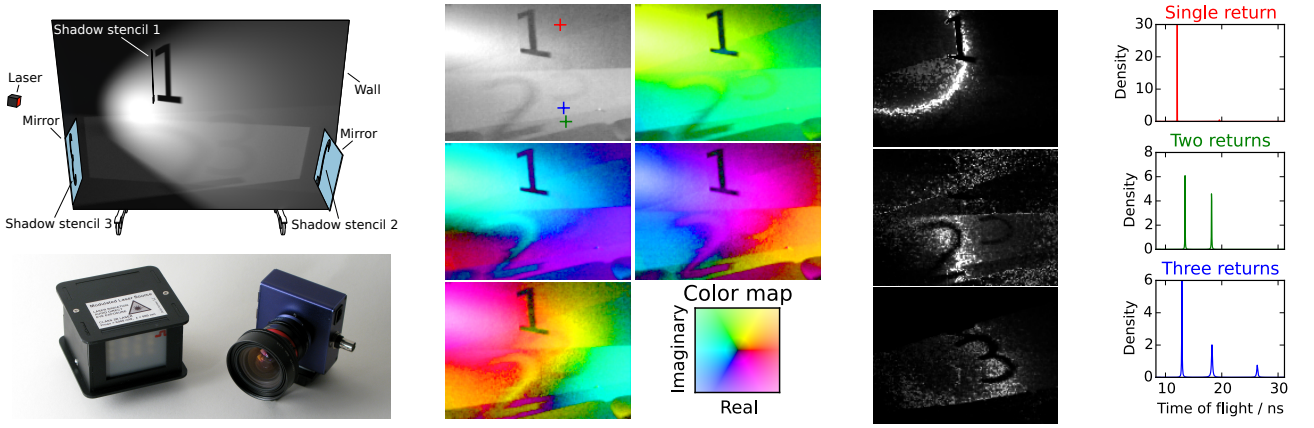
Solving Trigonometric Moment Problems for Fast Transient Imaging

Christoph Peters[1] Jonathan Klein[2,1] Matthias B. Hullin[1] Reinhard Klein[1]
[1]University of Bonn [2]French-German Research Institute of Saint-Louis



Scene, laser and AMCW lidar imager. Trigonometric moment images. Transient frames. Transient pixels.

Figure 1: *We use an AMCW lidar system to capture transient images at interactive frame rates. The shown scene is designed to provide a challenging test case. A wall is lit directly and indirectly via two mirrors such that parts of it return light after three different times of flight. Our customized AMCW lidar system captures all shown trigonometric moment images within* 113*ms. This is all data that enters our closed-form reconstruction. Using a GPU implementation we can reconstruct* $2.2 \cdot 10^5$ *transient frames per second. The reconstruction successfully separates the three returns due to the two specular interactions and direct illumination.*

Abstract

Transient images help to analyze light transport in scenes. Besides two spatial dimensions, they are resolved in time of flight. Cost-efficient approaches for their capture use amplitude modulated continuous wave lidar systems but typically take more than a minute of capture time. We propose new techniques for measurement and reconstruction of transient images, which drastically reduce this capture time. To this end, we pose the problem of reconstruction as a trigonometric moment problem. A vast body of mathematical literature provides powerful solutions to such problems. In particular, the maximum entropy spectral estimate and the Pisarenko estimate provide two closed-form solutions for reconstruction using continuous densities or sparse distributions, respectively. Both methods can separate m distinct returns using measurements at m modulation frequencies. For $m = 3$ our experiments with measured data confirm this. Our GPU-accelerated implementation can reconstruct more than 100000 frames of a transient image per second. Additionally, we propose modifications of the capture routine to achieve the required sinusoidal modulation without increasing the capture time. This allows us to capture up to 18.6 transient images per second, leading to transient video. An important byproduct is a method for removal of multipath interference in range imaging.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Digitizing and scanning

Keywords: AMCW lidar systems, transient imaging, range imaging, closed-form solution, trigonometric moment problem

1 Introduction

A wide range of consumers has access to amplitude modulated continuous wave (AMCW) lidar systems such as Microsoft Kinect for Xbox One. Typically, these cameras are used for range imaging. In this application the time of flight light takes from an active illumination into the scene and back to the camera is measured indirectly. The active illumination is modulated with a high-frequent periodic signal and the phase shift of the signal scattered back to the sensor is measured.

In the presence of global illumination effects, the assumption of a unique time of flight no longer holds, since light may reach points in the scene on many different paths of different length. Assuming existence of a unique phase shift during reconstruction leads to systematic errors in range images, which are often far greater than precision errors due to sensor noise. This is known as multipath interference.

Transient images model this complex behavior more completely. In such an image each pixel stores a time-dependent impulse response indicating how much light returned after a particular time of flight. This enables applications such as separation of direct and indirect illumination [Wu et al. 2014] or non-line-of-sight imaging [Velten et al. 2012]. It has been shown that an AMCW lidar system can

be used to estimate a transient image using measurements at many modulation frequencies [Heide et al. 2013]. While this is among the most cost-efficient approaches for measuring transient images, it has difficulties reconstructing complex impulse responses, measurement takes a minute and reconstruction takes even longer.

We propose a novel method for reconstructing transient images from AMCW lidar measurements. Our approach drastically reduces the number of required measurements and still successfully reconstructs complex impulse responses. Reconstruction is done through highly efficient closed-form solutions. The technique scales well from measurement of high quality transient images to quick heuristic measurements. We use the latter for reduction of multipath interference in range imaging.

At the core of our work is the finding that a specific measurement procedure turns reconstruction of impulse responses into so-called trigonometric moment problems. This kind of inverse problems is extremely well-studied and a body of literature provides various efficient solutions. In particular, we use the *maximum entropy spectral estimate* introduced in Section 3.2 and the *Pisarenko estimate* described in Section 5.1.

The maximum entropy spectral estimate reconstructs responses as continuous densities via a closed-form solution. The Pisarenko estimate is the limit case of the maximum entropy spectral estimate for sparse responses. If a sparse response with $m \in \mathbb{N}$ returns is present, this can be detected using measurements at $m+1$ frequencies and the Pisarenko estimate can provide a perfect reconstruction. Otherwise, the maximum entropy spectral estimate can provide a continuous reconstruction, reflecting remaining uncertainty through the smoothness of the peaks.

Both techniques require sinusoidal modulation. We introduce a novel method to accomplish this in Section 4.1. This way, our prototype hardware is capable of measuring up to 18.6 transient images of reasonable quality per second as demonstrated in Section 6.3. Such images can also be used for improved range imaging as demonstrated in Section 6.2. Higher quality measurements can take a few seconds because it is advisable to average many captures to improve the signal to noise ratio.

2 Related Work

In recent years transient imaging has been introduced as an exciting new imaging modality. Such images can be understood as video recording the return of light to a camera at an extreme frame rate when the scene is lit by an infinitesimally short light pulse. The first general hardware setup for their measurement uses a femtosecond laser and a streak camera [Velten et al. 2011; Velten et al. 2013]. The laser sends repeated short light pulses into the scene while the streak camera directs light returning at different times to different rows of the image sensor. This way a transient image can be captured one row at a time with a temporal resolution around 2ps. Capture takes roughly one hour. Later work uses interferometry and several hours of capture time to push temporal resolution to 33fs within a small capture volume [Gkioulekas et al. 2015].

Transient images add a fundamentally new dimension to images, thus enabling new applications. Geometry can be reconstructed solely by analyzing the light it reflects onto a diffuse wall [Velten et al. 2012]. Similarly, surface reflectance can be reconstructed around a corner [Naik et al. 2011]. Separation of images into direct illumination, subsurface scattering and indirect illumination can be performed by analyzing impulse responses [Wu et al. 2014].

While these applications demonstrate the usefulness of transient images, they are limited by the high cost and long measurement times of the involved hardware. A drastically faster and more cost-efficient approach uses AMCW lidar systems [Heide et al. 2013]. These cameras apply a modulation signal at the light source and the sensor. Effectively this means that they measure the correlation of a transient image with a time-dependent, periodic signal. Using measurements at many different modulation frequencies, the authors reconstruct the transient image by solving an inverse problem with soft, linear constraints enforcing compatibility with the measurements and additional temporal and spatial regularization priors. This way, transient images can be captured within a minute, but the regularization priors tend to lose high-frequency temporal details and reconstruction takes several minutes.

Subsequent works explore various measurement procedures, priors and reconstruction algorithms. Kadambi et al. [2013] use a broadband modulation and sample it at many phase shifts. The arising inverse problem is solved with various linear and non-linear priors. Kirmani et al. [2013] assume sinusoidal modulation at multiple frequencies with a common base frequency. These measurements are used as soft constraint to reconstruct a combination of two Dirac-δ pulses. Lin et al. [2014] use a similar input but employ an inverse Fourier transform with subsequent corrective post-processing. Bhandari et al. [2014b] use measurements at many frequencies and orthogonal matching pursuit to estimate a few Dirac-δ pulses. Qiao et al. [2015] use a logarithmic prior to reward sparsity. Kadambi et al. [2015] consider measurements as a function of frequency and derive times of flight from the frequencies in this signal. Bhandari et al. [2014a] present a method using sinusoidal modulation at $2 \cdot m + 1$ frequencies to reconstruct a linear combination of $m \in \mathbb{N}$ Dirac-δ distributions in closed-form. With measurements from Microsoft Kinect for Xbox One they successfully separate two returns using measurements at 21 frequencies.

At the other end of the spectrum there are works using far fewer measurements to reconstruct range images. In this context reconstruction of impulse responses only serves as intermediate step to model multipath interference. Using measurements at two modulation frequencies, two Dirac-δ distributions can be fitted to the measurements using a non-linear optimization [Dorrington et al. 2011]. Alternatively, measurements at four frequencies can be used to estimate parameters of a similar model in closed form [Godbaz et al. 2012]. It is also possible to use a more sophisticated L_1-regularization on a general impulse response [Freedman et al. 2014]. For application in real time the authors store the results in a four-dimensional look-up table to process three frequency measurements quickly. Gupta et al. [2015] observe that diffuse multipath interference tends to cancel out at high frequencies and thus propose to reconstruct range from few measurements at high frequency.

The above works can only capture transient images using an active illumination to generate a repetitive event. Transient images of non-repetitive events have been recorded in a single capture using compressed sensing, although this approach sacrifices spatial resolution for temporal resolution [Gao et al. 2014]. When range imaging is the primary concern, it is also possible to reduce multipath interference without additional measurements. To this end, diffuse interreflections in the scene can be modeled explicitly and the estimated multipath interference can be subtracted from the measurement [Fuchs 2010; Jimenez et al. 2012]. While this saves measurement time, it requires substantial post-processing time. Ground truth data for transient imaging can be generated with specialized Monte Carlo renderers [Jarabo et al. 2014].

3 Reconstruction of Impulse Responses

In the present section we demonstrate how to cast the inverse problem encountered in transient imaging with AMCW lidar systems into a trigonometric moment problem. This allows us to introduce our highly efficient closed-form solution and to understand some of its properties. Though, before dealing with the inverse problem we need to describe the forward model.

3.1 Signal Formation Model

Suppose $g(\tau)$ is the impulse response for a single pixel in a transient image, mapping $\tau \in \mathbb{R}$ to the density of light returning after this time of flight (see Figure 2a, 2b). The integral $\int_a^b g(\tau)\mathrm{d}\tau$ tells how much light returns in the time interval $[a, b]$. Measurements of an AMCW lidar system for a single pixel can be computed directly from the information of the corresponding transient pixel g.

The active illumination is modulated by a T-periodic signal $s_i(\tau)$ and the pixel receives the convolved signal $g * s_i$. The sensor is modulated with another T-periodic signal $s_s(\tau)$. Throughout the exposure time it integrates over the resulting signal $g * s_i \cdot s_s$. For simplicity we assume that it integrates exactly one period. Thus the measurement at the pixel is

$$\frac{1}{T} \cdot \int_0^T (g * s_i)(\tau) \cdot s_s(\tau)\mathrm{d}\tau = \int_0^\infty g(\tau) \cdot (s_i \star s_s)(\tau)\mathrm{d}\tau$$

where $s_i \star s_s$ denotes periodic cross-correlation.

This means that the sensor measures the correlation between the impulse response g and the effective modulation $s_e := s_i \star s_s$. At this point, it is interesting to note that all information that can possibly be captured by a pixel of an AMCW lidar system is part of the transient pixel g. Vice versa, many AMCW lidar systems allow customization of the modulation and thus a lot of information about g can be captured by using many modulation functions. This shows how closely these two imaging modalities are linked.

Our method assumes measurements with a specific set of modulation functions. We now introduce these assumptions and later demonstrate their practical implementation in Section 4. First we fix a base frequency $f \in \mathbb{R}$. In most of our experiments this is 23MHz. Furthermore, we fix the number of non-zero frequencies $m \in \mathbb{N}$ we want to measure. This is one of two major parameters allowing trade-offs between capture time and quality. Our experiments use $m \in \{3, \dots, 8\}$. Now for all $j \in \{0, \dots, m\}$ we sequentially use the effective modulation functions

$$s_e(\tau) = \cos(j \cdot 2 \cdot \pi \cdot f \cdot \tau) \quad \text{and} \quad s_e(\tau) = \sin(j \cdot 2 \cdot \pi \cdot f \cdot \tau).$$

For convenience let $\varphi := 2 \cdot \pi \cdot f \cdot \tau$ and $s(\varphi) := s_e(\frac{\varphi}{2 \cdot \pi \cdot f})$. The base frequency should be chosen such that its wavelength is longer than all interesting light paths. Otherwise phase ambiguity arises which we model by defining

$$h(\varphi) := \sum_{l=-\infty}^{\infty} g\left(\frac{\varphi + l \cdot 2 \cdot \pi}{2 \cdot \pi \cdot f}\right).$$

We can only reconstruct the phase-dependent, $2 \cdot \pi$-periodic function h, not the time-dependent, aperiodic function g. To further simplify notions we combine two real measurements into a single complex phasor (see Figure 2b):

$$\begin{aligned} b_j &:= \int_0^{2\cdot\pi} h(\varphi) \cdot \cos(j \cdot \varphi)\mathrm{d}\varphi + i \cdot \int_0^{2\cdot\pi} h(\varphi) \cdot \sin(j \cdot \varphi)\mathrm{d}\varphi \\ &= \int_0^{2\cdot\pi} h(\varphi) \cdot \exp(i \cdot j \cdot \varphi)\mathrm{d}\varphi \in \mathbb{C} \end{aligned}$$

The measurements $b_0, \dots, b_m$ are essentially Fourier coefficients of the $2 \cdot \pi$-periodic function h. However, since h is a density function, the preferred term in probability theory is *trigonometric moments*. Figure 2c visualizes them per pixel.

From this point onward our main concern is reconstruction of h from $b_0, \dots, b_m$. This inverse problem is known as trigonometric moment problem and it is very well-studied [Karlin and Studden 1966; Kreĭn and Nudel'man 1977]. To present our solution we require a few additional definitions.

Definition 1. Let $\mathbf{b} := (b_0, \dots, b_m)^\mathsf{T} \in \mathbb{C}^{m+1}$ be the vector of trigonometric moments. It is generated by the vector of basis functions $\mathbf{s} : \mathbb{R} \to \mathbb{C}^{m+1}$ with $\mathbf{s}_j(\varphi) := \exp(i \cdot j \cdot \varphi)$, i.e.

$$\mathbf{b} = \int_0^{2\cdot\pi} h(\varphi) \cdot \mathbf{s}(\varphi)\mathrm{d}\varphi. \tag{1}$$

Since h is real, we define moments with negative index using the complex conjugate $b_{-j} := \overline{b_j}$. Finally, let $B \in \mathbb{C}^{(m+1)\times(m+1)}$ denote the Toeplitz matrix defined by

$$B := (b_{j-k})_{j,k=0}^{m} = \begin{pmatrix} b_0 & b_{-1} & \cdots & b_{-m} \\ b_1 & b_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & b_{-1} \\ b_m & \cdots & b_1 & b_0 \end{pmatrix}.$$

3.2 Reconstruction via Trigonometric Moments

To make the most of our measurements we incorporate them into the reconstruction as a hard constraint; that is, we only admit densities fulfilling Equation (1). Unfortunately, this generally does not determine the solution uniquely. The data leaves us with uncertainty about the precise shape of the density function. A good reconstruction should reflect this uncertainty. It should not localize density unless the data enforces such a localization. Any other behavior would be arbitrary and could lead to wrong conclusions.

We implement this requirement by asking for the distribution of minimal Burg entropy [Burg 1975]. Burg entropy of a density h on $[0, 2 \cdot \pi]$ is defined as

$$H(h) := \int_0^{2\cdot\pi} -\log h(\varphi)\mathrm{d}\varphi.$$

This is not to be confused with the more common Boltzmann-Shannon entropy, which integrates $-h(\varphi) \cdot \log h(\varphi)$. By minimizing Burg entropy we punish small densities heavily because $-\log h \to \infty$ as $h \to 0$. On the other hand very big densities are rewarded only slightly because $\log h$ grows slowly. In terms of minimal Burg entropy, a density is optimal if it achieves moderate densities over large intervals. In this sense uncertainty is rewarded.

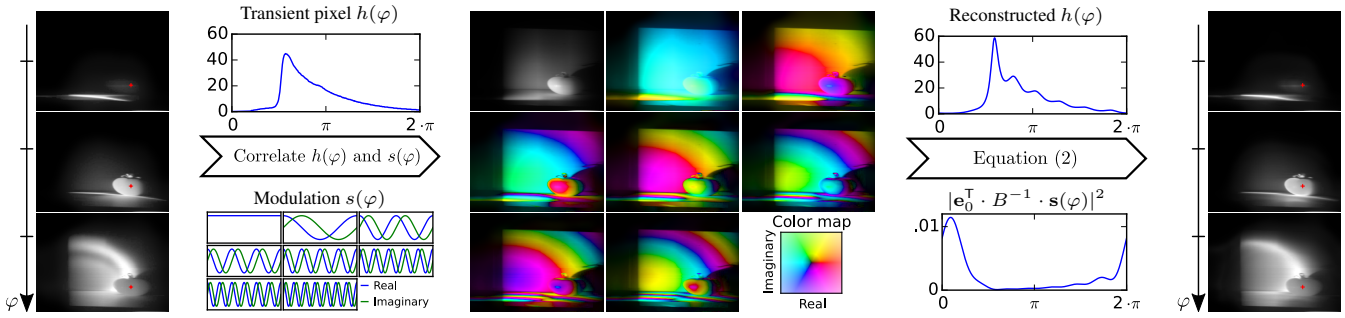
This prior is of particular interest to us because it admits a closed-form solution of the trigonometric moment problem. It is known as maximum entropy spectral estimate [Burg 1975].

Theorem 1. *Suppose that B is positive-definite. Let*

$$h(\varphi) := \frac{1}{2 \cdot \pi} \cdot \frac{\mathbf{e}_0^\mathsf{T} \cdot B^{-1} \cdot \mathbf{e}_0}{|\mathbf{e}_0^\mathsf{T} \cdot B^{-1} \cdot \mathbf{s}(\varphi)|^2} \tag{2}$$

where $\mathbf{e}_0 := (1, 0, \dots, 0)^\mathsf{T} \in \mathbb{R}^{m+1}$ denotes the first canonical basis vector. Then h is a positive density fulfilling Equation (1) and among all such densities it has minimal Burg entropy $H(h)$.

Proof. See [Burg 1975, p. 8 ff.] or supplementary. □



(a) *Transient image.* **(b)** *Signal formation.* **(c)** *Trigonometric moment images.* **(d)** *Reconstruction.* **(e)** *Result.*

Figure 2: *A schematic visualization of signal formation and our signal reconstruction. A lit scene implicitly defines a transient image resolved in time of flight τ or equivalently phase $\varphi = 2\cdot\pi\cdot f\cdot\tau$ (2a [Velten et al. 2013]). Per pixel, an AMCW lidar system correlates this signal with $m+1$ periodic modulation functions (2b). This yields $m+1=8$ images holding complex trigonometric moments per pixel (2c). These images are the only input of our closed-form reconstruction. The signal is reconstructed as continuous density (2d top) which is the reciprocal of a Fourier series (2d bottom). Reconstruction per pixel yields the full time-resolved transient image (2e).*

In spite of its remarkable properties of matching all measurements exactly while minimizing the prior, Equation (2) can be evaluated easily. The term mostly consists of dot products and basic arithmetic operations. To compute $\mathbf{e}_0^\mathsf{T}\cdot B^{-1}$ we need to solve a system of linear equations. This system has a very special structure which can be exploited by fast algorithms solving it in time $O(m^2)$ or even superfast algorithms solving it in $O(m\cdot\log^2 m)$ [Ammar and Gragg 1988]. Details are given in the supplementary. The result is a smooth reconstruction as shown in Figure 2d and 2e.

3.3 Properties of the Reconstruction

At a very general level Theorem 1 provides an alternative to a common inverse Fourier transform for a truncated series of Fourier coefficients. While a common inverse Fourier transform would simply set all unknown Fourier coefficients to zero, this solution chooses them to minimize Burg entropy. Still, it matches the given Fourier coefficients exactly. The major advantage is that the reconstruction is known to be a positive density. Therefore, when the application provides this prior knowledge, the maximum entropy spectral estimate should be preferred over a common inverse Fourier transform.

Effectively, Equation (2) defines the reciprocal of a positive Fourier series. Everything except for the term including $\mathbf{s}(\varphi)$ is a constant. Using $|z|^2 = z\cdot\bar{z}$ for $z\in\mathbb{C}$, this last expression can be rewritten as

$$|\mathbf{e}_0^\mathsf{T}\cdot B^{-1}\cdot\mathbf{s}(\varphi)|^2 = \mathbf{e}_0^\mathsf{T}\cdot B^{-1}\cdot\mathbf{s}(\varphi)\cdot\mathbf{s}^*(\varphi)\cdot B^{-1}\cdot\mathbf{e}_0,$$

where $\mathbf{s}^*$ denotes the conjugate transpose. The product $\mathbf{s}\cdot\mathbf{s}^*$ is an $(m+1)\times(m+1)$ matrix with entries $\exp(-i\cdot m\cdot\varphi),\ldots,\exp(i\cdot m\cdot\varphi)$ on its diagonals. Thus, the expression is a positive Fourier series with frequency components ranging from $-m$ to m (see Figure 2d).

It can be regarded as the Fourier series of minimal degree such that its reciprocal produces the prescribed trigonometric moments. If B is positive-definite, this Fourier series has no root. Still, it can be close to zero. Whenever this happens, the reconstructed density exhibits a sharp peak. In practice, this is very common. The limit case is a perfectly sparse distribution (see Section 5.1).

The assumption of positive-definite B may seem like a limitation of the reconstruction, but in fact it is not. Physically meaningful measurements allow nothing else.

Proposition 1 ([Kreĭn and Nudel'man 1977, p. 65, p. 78])**.** *Let h be non-negative and let B be the Toeplitz matrix as in Definition 1. If $B\neq 0$, B is positive-definite.*

This is a very elegant characterization of all valid measurements. If a measurement induces a matrix B with a negative eigenvalue, it has to be faulty because no impulse response h could have generated it. Otherwise, Theorem 1 (or Theorem 2 below) guarantees a reconstruction compatible with all measured data. Being able to judge the validity of measurements in such a way is a very useful tool for validation of measurement procedures. Alternatively, this result can be used to estimate the zeroth moment b_0 as demonstrated in Section 4.3.

4 Measurement Procedure

To use the maximum entropy spectral estimate with measured data we need to acquire measurements at specific frequencies with sinusoidal modulation as explained in Section 3.1. In the following, we present methods to achieve this robustly. Our experiments use a modified version of the hardware setup used in [Heide et al. 2013], but we believe that the proposed methods can be employed to a wide range of hardware including Microsoft Kinect for Xbox One [Bhandari et al. 2014a].

4.1 Achieving Sinusoidal Modulation

It is difficult to achieve exact sinusoidal modulation by adjusting the electronics providing the modulation signal. Fortunately, there is a robust workaround leading to a modulation that is arbitrarily close to a sinusoidal. For this to work, it has to be possible to adjust the phase shift between the light modulation s_i and the sensor modulation s_s. This shifts the effective modulation s.

Harmonic cancellation [Payne et al. 2010] uses $n\in\mathbb{N}$ equidistant phase shifts and builds a linear combination of the resulting measurements. This is equivalent to generating a linear combination of the phase-shifted modulations. The specific linear combination is

$$\sum_{k=0}^{n-1}\sin\left((k+1)\cdot\frac{\pi}{n+1}\right)\cdot s\left(\varphi-k\cdot\frac{\pi}{n+1}\right).$$

This new effective modulation is free of harmonic frequencies up to harmonic $2\cdot n-1$. Use of harmonic cancellation does not increase measurement times because each phase shift can be used for a fraction of the exposure time that reflects the weight in the linear combination [Payne et al. 2010].

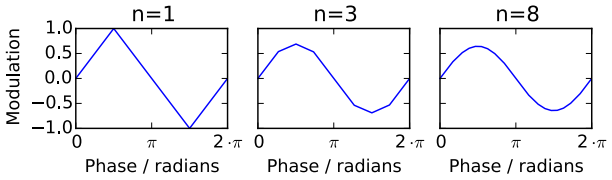


Figure 3: *The modulation arising from Equation (3) for a triangular original modulation s. For $n = 8$ it is close to a sinusoidal.*

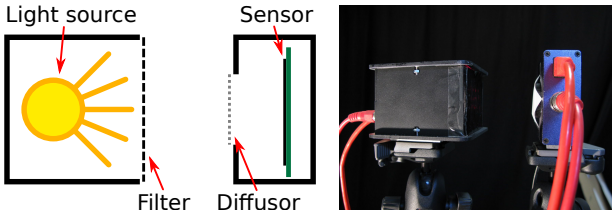


Figure 4: *We point the laser at the sensor at short range using a neutral density filter with $5.8‰$ translucency to avoid overexposure and a diffusor to ensure a uniform light distribution on the sensor. The filter has low reflectivity to avoid undesired interreflections.*

We have adopted harmonic cancellation in our prototype hardware but only get robust results up to $n = 3$ due to timing issues. Therefore, we propose an alternate scheme. We do not use equidistant phase shifts but split up the exposure time evenly. The k-th interval of the exposure time uses phase shift $\arccos\left(1 - \frac{2\cdot k+1}{n}\right)$. Thus, the effective modulation becomes

$$\frac{1}{n} \cdot \sum_{k=0}^{n-1} s\left(\varphi - \arccos\left(1 - \frac{2 \cdot k + 1}{n}\right)\right). \tag{3}$$

In the supplementary material we show that this sum approaches a perfect sinusoidal for $n \to \infty$. For this to work it should be used in combination with four-bucket sampling. In practice, we use this $\arccos$-phase sampling with $n = 8$. This yields a satisfactory approximation to a sinusoidal as shown in Figure 3.

4.2 Calibration

The above methods make the effective modulation sinusoidal. Since we use direct digital synthesis for generation of the modulation signal, we can also rely on the accuracy of the frequency ratios. Thus, the only remaining degrees of freedom for the modulation signal are the phase shift and the amplitude. Measurements show that these need to be calibrated per pixel.

To this end, we point the light source at the sensor as shown in Figure 4. This whole setup is designed to avoid multipath interference. We expect that most light only passes through the filter once before it reaches the sensor. Other light paths should be attenuated due to the low reflectivity of the filter and either way they should be rather short.

We model the impulse response at each pixel in this setup by a scaled Dirac-δ pulse at phase zero. Thus, in a properly calibrated setup we would expect to obtain moments

$$b_j = b_0 \cdot \exp(i \cdot j \cdot 0) = b_0 \text{ for } j \in \{0, \ldots, m\}.$$

The actual measured values will deviate from this. To apply a calibration to a measurement we simply divide the moments computed from the measurement by the moments computed from the calibration. This way, we simultaneously compensate varying phase shifts and amplitudes. To avoid changing overall brightness we normalize vectors of moments from the calibration such that $b_0 = 1$ prior to this step. For optimal results, calibration measurements should be performed with the same parameters as actual measurements.

Note that the setup in Figure 4 differs from the more common calibration setup of pointing the light source and the camera at a white wall [Heide et al. 2013; Lin et al. 2014]. We also experimented with this setup but found that it makes it hard to avoid multipath interference leading to systematic errors. If the impulse response in the calibration is not a Dirac-δ, calibration procedures effectively perform a deconvolution with the actual impulse response leading to systematic distortions of the reconstruction.

4.3 The Zeroth Moment

The zeroth moment is defined by $b_0 = \int_0^{2\cdot\pi} h(\varphi)\mathrm{d}\varphi$, so it captures total brightness due to the active illumination without any modulation. All related work using AMCW lidar systems, except for Godbaz et al. [2012], only incorporates measurements with zero-mean modulation, meaning that the captured data is literally orthogonal to the zeroth moment. This misses important information. Consider the uniform density $h(\varphi) := \frac{1}{2\cdot\pi}$. Its zeroth moment is one, whereas all other moments are zero. If the zeroth moment is not measured, an arbitrarily strong uniform component can be added to the impulse response without changing the data. In this sense, the zeroth moment governs sparsity of the distribution, as demonstrated in Figure 5. In Section 5.1 we find that a minimal choice of the zeroth moment can be used to enforce a sparse reconstruction for any measurement.

The best practice for capturing the zeroth moment is to capture two images without sensor modulation, one with and one without active illumination. Their difference provides the zeroth moment. Since our prototype hardware cannot measure without modulation, we instead perform measurements at 900kHz. This corresponds to a wavelength of 333.1m so the sinusoidal modulation wave should be nearly constant across relevant lengths of light paths.

Alternatively, we can exploit Proposition 1 to estimate b_0 based on prior knowledge about the sparsity of impulse responses. The zeroth moment b_0 constitutes the main diagonal of the Toeplitz matrix B. If we have not measured b_0 yet, we can set the main diagonal of B to zero and compute its smallest eigenvalue λ_m which will be negative. We can fix the estimated uniform component $\varepsilon > 0$ and set $b_0 := \varepsilon - \lambda_m$ to ensure that the Toeplitz matrix is positive definite with smallest eigenvalue ε. Smaller values of ε lead to sparser reconstructions.

This method is also suited for correcting invalid measurements if we have measured the zeroth moment. Whenever we encounter a Toeplitz matrix with a smallest eigenvalue less than ε, we replace b_0 by $\varepsilon - \lambda_m$ as defined above. This changes the measurement in a minimal way to make it valid. Alternatively, $b_1, \ldots, b_m$ can be scaled down to avoid changing overall brightness. We refer to this procedure as biasing. For scenes with sparse impulse responses sensor noise makes it indispensable.

5 Analysis of Transient Images

The maximum entropy spectral estimate gives us efficient random access to a transient image. Though, in many application scenarios we would like to infer other information immediately. In this section we present various efficient methods to infer information about a transient image without computing it completely. We also present upper bounds for the error of the reconstruction.

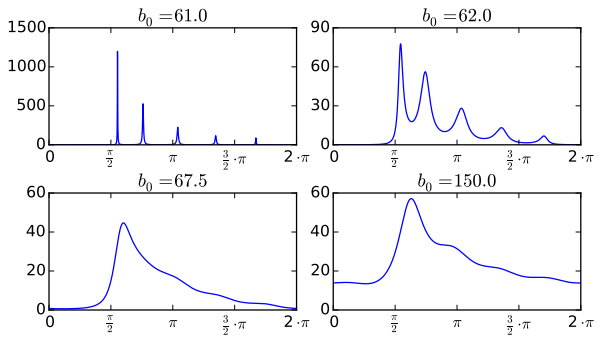


Figure 5: *Various reconstructions arising from Equation (2) for the transient pixel from Figure 2b with $m = 5$ and different values of the zeroth moment b_0. The x-axis shows φ, the y-axis $h(\varphi)$. The zeroth moment governs sparsity of the reconstruction. The best result is obtained with the ground truth $b_0 = 67.5$. For $b_0 = 61$ the Toeplitz matrix B is nearly singular and the reconstruction is nearly sparse. In between it changes continuously. The strong changes illustrate the importance of measuring b_0 accurately.*

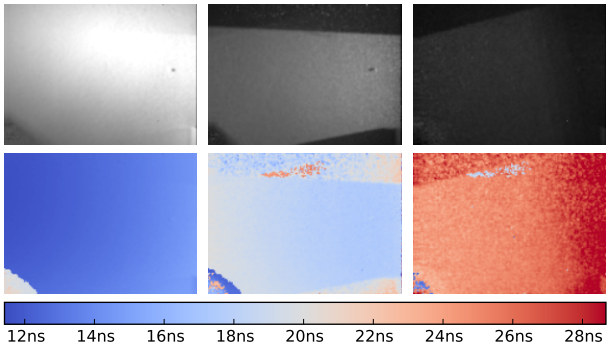


Figure 6: *A scene where direct illumination and two mirrors are used to illuminate a wall three times (cf. Figure 1). Four moments are measured by averaging 30 takes to improve signal to noise ratio. Top: The strength of the strongest three returns computed with the Pisarenko estimate. Bottom: The corresponding time of flight. The different light paths through the mirrors are separated clearly. Phase noise increases as the strength of the return weakens.*

5.1 Perfect Reconstruction of a Sparse Response

In Section 4.3 we observed that the smallest eigenvalue of the Toeplitz matrix B can be seen as a measure of the strength of uniform components in the impulse response (see Figure 5). From Proposition 1 we know that this eigenvalue can never be negative. A natural question to ask is what happens when the eigenvalue is zero. It turns out that this corresponds to a perfectly sparse impulse response with no more than m returns.

Theorem 2 ([Kreĭn and Nudel'man 1977, p. 65, p. 78])**.** *Suppose B is positive semi-definite but singular. Then there exists $\varphi_0, \dots, \varphi_{m-1} \in [0, 2 \cdot \pi]$ and $w_0, \dots, w_{m-1} \in \mathbb{R}_+$ such that*

$$\mathbf{b} = \sum_{k=0}^{m-1} w_k \cdot \mathbf{s}(\varphi_k). \tag{4}$$

In consequence, a valid reconstruction for the trigonometric moments $b_0, \dots, b_m$ is given by $\sum_{k=0}^{m-1} w_k \cdot \delta_{\varphi_k}$ where δ_{φ_k} denotes a Dirac-δ distribution at φ_k.

There is no other valid reconstruction. Vice versa B is known to be positive semi-definite and singular if it arises from an impulse response consisting of no more than m Dirac-δ distributions.

Let $\mathbf{c} \in \mathbb{C}^{m+1}$ with $\mathbf{c} \neq 0$ and $B \cdot \mathbf{c} = 0$. Then $\varphi_0, \dots, \varphi_{m-1}$ are roots of the function $\mathbf{c}^* \cdot \mathbf{s}(\varphi)$.

The fact that this reconstruction is the only one providing a valid explanation for the measurements is remarkable. It means that we can achieve a perfect reconstruction whenever the ground truth consists of m Dirac-δ pulses or less. We can also compute this unique reconstruction efficiently. We only need to compute the kernel of B to obtain $\mathbf{c}$. Computing the roots of $\mathbf{c}^* \cdot \mathbf{s}(\varphi)$ yields the location of the Dirac-δ pulses. We can substitute $z := \exp(i \cdot \varphi)$ to turn this into a polynomial equation of degree m. Finally, the weights $w_0, \dots, w_{m-1}$ can be computed by solving the system of linear equations (4).

While we have formulated this result for the case that B is singular, it can also be used in the general case. We can simply separate the distribution into a uniform component and a sparse component. To this end we compute the smallest eigenvalue λ_m of B and a corresponding eigenvector $\mathbf{c}$. Then λ_m gives the strength of the uniform component and the sparse component can be computed from $\mathbf{c}$ as above but using $b_0 - \lambda_m$ in place of b_0.

If specular interactions dominate, the uniform component becomes small and measurement of the zeroth moment may be skipped. This method is known as Pisarenko estimate and optimized algorithms exist for its computation [Cybenko and Loan 1986]. It is closely related to the work by Bhandari et al. [2014a], except that their method requires more than twice as many measurements and does not necessarily find a distribution compatible with all of them. It can also be understood as closed-form implementation of the work by Freedman et al. [2014] without error tolerance because their technique minimizes b_0. The Pisarenko estimate realizes the theoretical best case of reconstructing the $2 \cdot m$ real parameters describing m Dirac-δ distributions from m complex phasors.

While the Pisarenko estimate does not reflect uncertainty as reasonably as the maximum entropy spectral estimate, it provides a more explicit reconstruction. This eases analysis of transient images and provides excellent results if specular interactions dominate as demonstrated in Figure 6. The reconstructed data directly provides insight into the strength and time of returns per pixel.

A possible application of this is fast separation of illumination into direct and indirect illumination. The direct component can be identified as first return with a weight above a relative threshold. Its weight provides the strength of the direct return. The sum of the other weights provides indirect returns.

5.2 Error Bounds

From Theorem 2 we know that we can obtain a perfect reconstruction if the ground truth is perfectly sparse. It is natural to expect that the reconstruction is still very close to the ground truth when the ground truth has a small uniform component. This motivates the search for bounds on the error of the reconstruction. We suppose that our measurements are correct and ask for the maximal possible distance between the unknown ground truth and our reconstruction.

For the trigonometric moment problem this question is solved [Karlsson and Georgiou 2013]. The authors observe that no meaningful statements can be obtained if densities are considered directly. At any phase, sparse impulse responses may have zero density or a small Dirac-δ pulse corresponding to infinite density. Thus, no useful bounds can be obtained. To get a meaningful result, den-

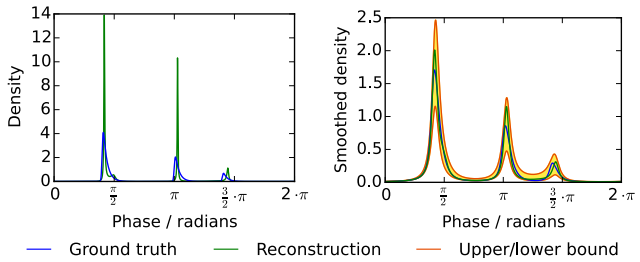


Figure 7: *Left: A synthetic ground truth (cf. Figure 11b) and the corresponding maximum entropy spectral estimate for $m = 5$. Right: The upper and lower bound for the smoothed density and the smoothed densities themselves. Any density compatible with the trigonometric moments used in the reconstruction is known to lie within these bounds after smoothing. These bounds use $r = 0.93$.*

sities are smoothed before analysis. This is done using the Poisson kernel

$$P_r(\varphi) := \frac{1}{2 \cdot \pi} \cdot \frac{1 - r^2}{|1 - r \cdot \exp(i \cdot \varphi)|^2},$$

where $r \in [0, 1)$ governs the sharpness of the kernel. For $r = 0$ it is constant, for $r \to 1$ it converges to δ_0.

The authors proceed to derive sharp upper and lower bounds for the smoothed density $h * P_r(\varphi)$ where $\varphi \in \mathbb{R}$ is arbitrary. The bounds rely solely on the knowledge that h fulfills Equation (1). They can be evaluated in closed form for any given $\mathbf{b} \in \mathbb{C}^{m+1}$, $r \in (0, 1)$ and $\varphi \in \mathbb{R}$. For details please refer to the supplementary material.

Assuming correct measurements, the ground truth is known to fulfill Equation (1). Thus, we can compute an area containing its smoothed density. The same holds for our reconstruction. If the ground truth is reasonably close to a sparse distribution, this area is pleasantly small as demonstrated in Figure 7. In this case, we can be certain that the reconstruction is close to the ground truth and we can give specific bounds on possible locations of local maxima in the unknown ground truth.

5.3 Estimating Range

Range imaging with AMCW lidar systems is typically done in real-time with a limited computing time budget. Therefore, it is worthwhile to implement highly optimized methods. The natural candidates for estimates of range are local maxima of the reconstructed density in Equation (2). Critical points of this function fulfill

$$\sum_{j=0}^{m} \sum_{k=0}^{m} \overline{(B^{-1} \cdot \mathbf{e}_0)_j} \cdot (B^{-1} \cdot \mathbf{e}_0)_k \cdot (j - k) \cdot z^{m+j-k} = 0,$$

where $z := \exp(i \cdot \varphi)$. This is a polynomial equation of degree $2 \cdot m$ so it can have up to $2 \cdot m$ roots on the unit circle which can be computed efficiently. Since the reconstructed density is a smooth, periodic function, at most half of the critical points can correspond to local maxima. A reasonable estimate for range is that it corresponds to the first local maximum above a threshold.

If specular interactions are known to be the primary cause of multipath interference, an even faster approach uses the Pisarenko estimate introduced in Section 5.1. This way, the degree of the arising polynomial equation is halved. It is interesting to note that both methods can reconstruct m distinct peaks. The maximum entropy spectral estimate only requires the additional measurement of the zeroth moment to estimate smoothness of the impulse response and it only constructs m peaks when the data demands it (see Figure 5).

6 Results

In the following we present results measured with our prototype hardware. The setup is similar to the one presented by Heide et al. [2013] but additionally features the arccos-phase sampling described in Section 4.1. The sensor is a CamBoard nano by PMDTechnologies with a resolution of $163 \cdot 120$ pixels.

For the results in Figure 6, 8b, 8c, 9b, 9c, 10, 12 and 15 we average multiple takes to improve the signal to noise ratio of the data. Note that our prototype hardware suffers from some systematic outliers due to synchronization issues. $5 - 10\%$ of all captured images differ significantly from other images captured with the same configuration. When multiple takes are given for averaging, we automatically discard such outliers before averaging. In videos we fill in data missing due to outliers using data from the previous frame.

To further reduce noise we smooth moment images using a Gaussian filter with a standard deviation of 0.6 pixels. For the maximum entropy spectral estimate we use biasing as described in Section 4.3 to ensure that $\lambda_m \geq 4 \cdot 10^{-3} \cdot b_0$. Unless otherwise noted, this only affects few pixels.

We have implemented evaluation of Equation (2) on the GPU in a pixel shader and measure the run time on an nVidia Geforce GTX 780. The shader reconstructs $2.9 \cdot 10^5$, $2.2 \cdot 10^5$ and $1.1 \cdot 10^5$ transient frames per second for $m = 3$, $m = 4$ and $m = 8$, respectively. This includes repeated computation of $B^{-1} \cdot \mathbf{e}_0$, although this vector could be precomputed. This means that Equation (2) can be evaluated $163 \cdot 120 \cdot 1.1 \cdot 10^5 = 2.2 \cdot 10^9$ times per second for $m = 8$. In fact, the algorithm is likely bandwidth limited so the arithmetic operations can be completed even faster.

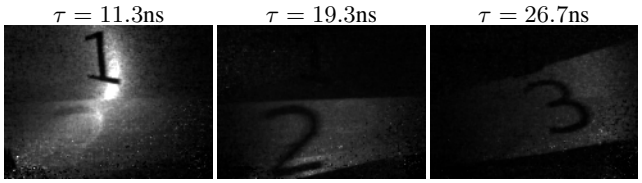
6.1 Transient Imaging

Our first experiment uses the scene shown in Figure 1 to provide a challenging test case with complex specular interactions. The impulse responses encountered for a single pixel can consist of three distinct returns with high dynamic range. Light initially sweeps from left to right, then the right mirror reflects it to the left and finally the left mirror reflects it to the right. These three returns have different times of flight and the reconstruction has to separate them.

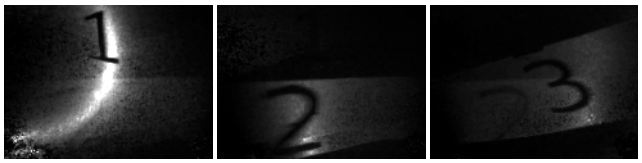
Figure 8 shows results obtained with the maximum entropy spectral estimate for different measurement times using an exposure time of 1.92ms. With $m = 3$ (i.e. measurements at four frequencies including the measurement for the zeroth moment at 900kHz) and a single take we already successfully separate the three distinct returns. However, the reconstruction includes significant uncertainty expressed by means of broad peaks with low density. This behavior depends on the shape of the impulse response. Therefore, the number two is slightly visible in Figure 8a for $\tau = 11.3$ns.

Additional frequency measurements yield sharper peaks. At $m = 4$ some visible artifacts remain but all important features are reconstructed. Using $m = 8$ further reduces these artifacts, leading to a reconstruction with sharp peaks. On the other hand, the additional measurements also introduce additional noise and potential contradictions in the sense of Proposition 1, thus making biasing mandatory for 90% of all pixels.

The additional takes used for Figure 8b and 8c reduce the noise in the input and the output alike. Figure 1 and Figure 8b show the same measurement with one and 20 averaged takes, respectively. While the reconstructed features are essentially identical, the averaging leads to a result with substantially less noise. In spite of the single take, Figure 8a appears less noisy than Figure 1 because the greater uncertainty causes smoothing of impulse responses.



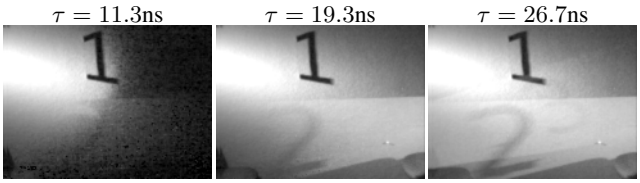
(a) $m = 3$, $f = 23$*MHz, no averaging,* 11*Hz capture rate.*



(b) $m = 4$, $f = 23$*MHz, 20 averaged takes,* 0.44*Hz capture rate.*

(c) $m = 8$, $f = 11.5$*MHz, 40 averaged takes,* 0.12*Hz capture rate.*

Figure 8: *A transient image of the scene shown in Figure 1 captured with different trade-offs between capture time and quality. The images show the maximum entropy spectral estimate at different times. Note how the three returns are separated.*



(a) $m = 3$, $f = 23$*MHz, no averaging,* 11*Hz capture rate.*

(b) $m = 4$, $f = 23$*MHz, 20 averaged takes,* 0.44*Hz capture rate.*

(c) $m = 8$, $f = 11.5$*MHz, 40 averaged takes,* 0.12*Hz capture rate.*

Figure 9: *Frames of the cumulative transient images corresponding to Figure 8. Each image is obtained by numerical quadrature of the maximum entropy spectral estimate up to the time specified.*

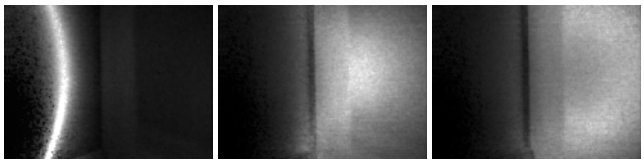
Density images such as those shown in Figure 8 generally exhibit rather strong noise because slight changes in the sharpness or phase of a peak can lead to strong changes in density at a fixed point in time. To analyze whether this noise is systematic, we consider the total density integrated up to a point in time. We refer to the resulting images as cumulative transient images and compute them by numerical quadrature with 10^5 equidistant samples.

Figure 9 shows such cumulative transient images for the above experiment. We note that uncertainty in the reconstruction translates to smeared out or misshaped wave fronts. However, the total brightness contributed by the waves is always reconstructed correctly.

For an example with diffuse interactions we point the camera at a corner but only illuminate the left wall of this corner directly. Most of the right wall is lit indirectly. As shown in Figure 10, the measurement of the zeroth moment helps us to adequately reconstruct the corresponding transient image. While the wave on the left wall is very sharp, the right wall receives a soft wave due to diffuse interreflections.

The comparison of our proposed methods to related work in Figure 11 uses synthetic data. For visualization we once more use cumulative density because sparse reconstructions cannot be represented by a common density function. The first example in Figure 11a constitutes an ideal case for all techniques and thus all techniques obtain an excellent reconstruction without noise. However, the Dirac-δ model [Godbaz et al. 2012] and SPUMIC [Kirmani et al. 2013] are quite sensitive to noise.

The second example in Figure 11b provides a more realistic test case consisting of three continuous returns modeled by exponentially modified Gaussians [Heide et al. 2014]. The Dirac-δ model and SPUMIC, both targeted at two sparse returns, only capture the first return adequately and become even more sensitive to noise. SRA [Freedman et al. 2014] has a bias towards stronger sparsity and less overall brightness. Therefore, it loses the third return but successfully reduces the impact of noise. The Pisarenko estimate provides a better reconstruction but is more sensitive to noise. The maximum entropy spectral estimate adequately reconstructs the continuous return. Noise mostly affects sharpness of the peaks.

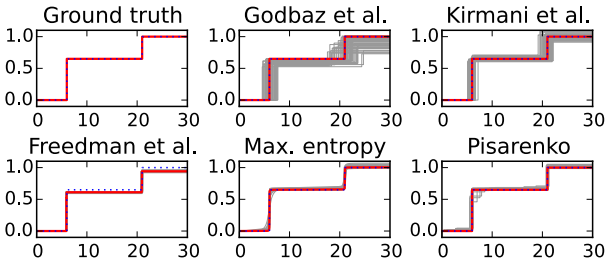


(a) $\tau = 13.1$*ns* **(b)** $\tau = 15.5$*ns* **(c)** $\tau = 16.1$*ns*

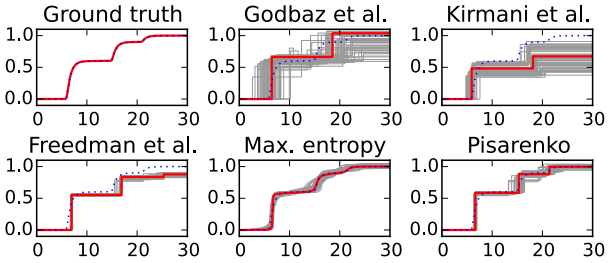
Figure 10: *Frames of a transient image showing a corner where only the left wall is lit directly. Thanks to the measurement of the zeroth moment the maximum entropy spectral estimate reconstructs a sharp peak for directly lit parts and a smooth peak for indirectly lit parts. This measurement uses* $f = 23$*MHz,* $m = 4$ *and 40 takes. The capture rate is* 0.22*Hz.*

6.2 Range Imaging

As benchmark for range imaging we place camera and light source next to each other and capture a diffuse corner. This is a prime example of diffuse multipath interference. Figure 12 shows our results. A naive reconstruction using measurements at a single frequency exhibits severe distortions. Range is overestimated because long indirect paths contribute to the estimate. Using the Dirac-δ model [Godbaz et al. 2012] reduces these systematic distortions but does not behave robustly. The maximum entropy spectral estimate (see Section 5.3) provides robust results and reduces distortions due to multipath interference heavily. The Pisarenko estimate suffers from severe outliers. This is understandable because its inherent assumption of a sparse impulse response is inadequate for diffuse multipath interference. This demonstrates the benefit of including the zeroth moment in the reconstruction.



(a) *Reconstruction of a sparse distribution.*



(b) *Reconstruction of a continuous distribution.*

Figure 11: *Reconstruction results of various techniques using synthetic data with sinusoidal modulation. The x-axis shows time of flight in nanoseconds, the y-axis cumulative density of the reconstruction. Red graphs use measurements without noise, gray graphs originate from measurements with a simulated signal to noise ratio of $70 : 1$ due to Gaussian noise. Each plot contains the ground truth as dotted blue line. The Dirac-δ model [Godbaz et al. 2012] uses measurements at 11, 22, 33, 44MHz, Kirmani et al. [2013] uses 11, 22, ..., 66MHz and Freedman et al. [2014] uses 23, 46, 69MHz and $\varepsilon = 0.05$. Our proposed techniques use measurements at 0 (maximum entropy spectral estimate only) 23, 46, 69MHz.*

6.3 Transient Video

Our hardware is capable of capturing transient images at video frame rate. If we choose $m = 3$ and use a Pisarenko estimate or a maximum entropy spectral estimate with biasing of the zeroth moment, we require measurements at three frequencies. With four-bucket sampling this amounts to twelve images per transient image. Using an exposure of 0.5ms and $f = 23$MHz we can capture such sets of images at 18.6Hz. The result is a transient video, i.e. a four-dimensional function $g(\tau, t, x, y)$ mapping time of flight $\tau \in \mathbb{R}$, time $t \in \mathbb{R}$ and pixel coordinates $x, y \in \mathbb{R}$ to density of brightness.

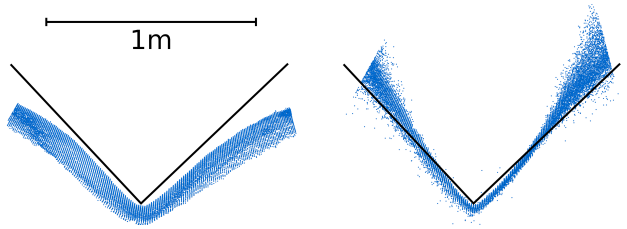
In our experimental setup a mirror is placed in front of a lit wall to reflect part of the light away from the wall and into another mirror. The latter mirror reflects the light back onto the wall such that part of it receives light at two different times of flight. An actor enters from the left, waves and leaves to the right. Therefore, the scene exhibits interesting features in the dimension of time of flight τ as well as common time t.

Figure 13 shows three frames of the transient images for three frames of the transient video. Although the images are quite noisy due to the short exposure time, all important features are reconstructed correctly. The light first returns from the actor, then from the direct interaction with the wall and finally it returns after being reflected by both mirrors and the wall.

We can use the Pisarenko estimate for separation of direct and indirect illumination as proposed in Section 5.1. Figure 14 shows the same frames as Figure 13 but this time direct and indirect illumination are separated. The indirect component exhibits a few outliers but generally separates the lighting due to the mirrors correctly from other lighting. Thanks to this correct separation, the time of flight for the direct return is free of multipath interference.

The sum of both components is the zeroth moment shown in Figure 14a. This image has not been measured directly but has been computed from three frequency measurements as described in Section 4.3. This is of interest by itself because it provides a method to compute images including solely active illumination.

More results can be found in the supplementary video.



(a) *Single frequency ($f = 23$MHz).* **(b)** *[Godbaz et al. 2012].*

(c) *Maximum entropy spectral estimate ($m = 3$, proposed method).* **(d)** *Pisarenko estimate ($m = 4$, proposed method).*

Figure 12: *Point clouds of a corner reconstructed with various techniques. All reconstructions use the same data set with $f = 23$MHz and 4 averaged takes (without outliers). The capture rate is 2.2Hz. The black line is the ground truth.*

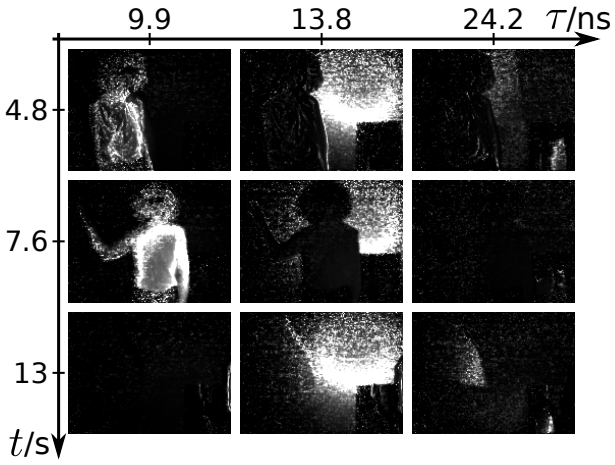
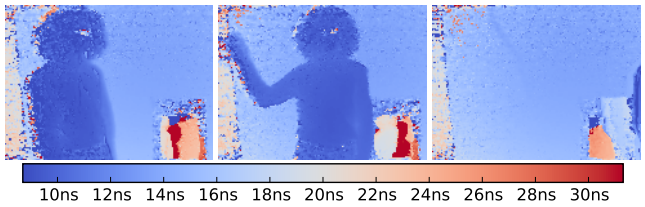


Figure 13: *A transient video is a four-dimensional image parameterized over time of flight, time and two spatial dimensions. The light wave progresses through the scene as time of flight increases whereas the actor moves through the scene as time increases.*

(a) *Direct and indirect illumination.*

(b) *Direct illumination.*

(c) *Indirect illumination.*



(d) *Time of flight for the direct return.*

Figure 14: *Separation of direct and indirect illumination in a transient video. Note that the lighting due to the mirror is separated from direct illumination and does not distort the time of flight.*

6.4 Limitations

Our proposed methods are fundamentally limited by the small amount of input data they require. While a continuous impulse response has infinite degrees of freedom our methods only get $2 \cdot m+1$ real quantities to reconstruct it. An exact reconstruction is only guaranteed in the case of a sparse response with m returns or less (see Theorem 2). In absence of this special case the prior in the maximum entropy spectral estimate reflects uncertainty in the reconstruction by smoothing of the peaks. This behavior depends on the form of the impulse response, which is why the number 2 is visible in Figure 8a for $\tau = 11.3$ns. The smoothing can provide a plausible reconstruction for diffuse multipath interference (see Figure 10) but accuracy is limited by the lack of data.

If two returns are temporally close, uncertainty can cause them to merge into one (see Figure 15). It is not possible to specify generic lower or upper bounds on the required distance of returns for successful separation. It rather depends on the complexity of the impulse response and the amount and quality of input data. Under the assumption of perfect data and perfect sparsity, returns can be arbitrarily close, but in presence of uniform components and noise they have to be farther apart.

Finally, sensor noise and calibration can be problematic. This is particularly true for our prototype hardware, which lacks active cooling. Since AMCW lidar systems exhibit thermal drift, we choose to perform calibration and all measurements after some warm-up time. Otherwise thermal drift may invalidate the calibration. However, these increased temperatures diminish the signal to noise ratio. Besides slight miscalibration due to thermal drift can still be an issue. This can be detected by observing that the Toeplitz matrix has negative eigenvalues for many pixels (see Proposition 1).

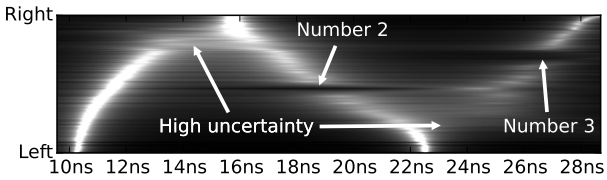


Figure 15: *A streak image showing density for the dataset in Figure 9b. The x-axis corresponds to time of flight, the y-axis to coordinates along the scanline highlighted in Figure 9b. Note the smoothed densities in regions of high uncertainty.*

Noise affects reconstructed density profiles by shifting peaks and changing their sharpness. Since these effects are independent for adjacent pixels, density images with sharp returns can appear very noisy. The cumulative density within peaks is quite robust to noise (see Figure 9). Averaging provides a reliable way to counteract noise at the cost of increased capture times of several seconds. Better hardware can speed this up.

It should also be noted that current AMCW lidar systems do not support sufficiently high modulation frequencies to reach the temporal resolution of setups with more sophisticated hardware [Velten et al. 2013; Gkioulekas et al. 2015]. This is problematic for applications such as separation of subsurface scattering [Wu et al. 2014]. Future hardware developments may change this [Gupta et al. 2015].

7 Conclusion

We have presented a novel method to transfer AMCW lidar measurements from the frequency domain to the time domain. At capture times of several seconds the quality of the results is competitive with regard to previous works using AMCW lidar systems. It realizes the theoretical best case where m peaks can be reconstructed from measurements at m frequencies. The technique scales well to very short capture times. In fact measurement times for a whole transient image become short enough to capture transient video at 18.6Hz where each frame is a transient image of lower quality. As an important byproduct we obtain a robust and fast method to remove multipath interference for range imaging.

We believe that this method can become the default way to post-process AMCW lidar measurements. Given access to a firmware admitting customization of modulation frequencies, it should be compatible with Microsoft Kinect for Xbox One [Bhandari et al. 2014a]. This will make fast transient imaging available to a wide audience and may very well be a breakthrough for this novel imaging modality. Besides, future software updates for shipped range sensors may implement our technique to get improved range imaging at no additional cost.

Acknowledgments

We would like to thank Jaime Martín and Ralf Sarlette for support with the measurements, Dominik Michels, Paul Müller, Andreas Kolb, Damien Lefloch and Hamed Sarbolandi for fruitful discussions and all reviewers for their helpful comments. This work was supported by the German Research Foundation (HU 2273/2-1) and by X-Rite Inc. through the Chair for Digital Material Appearance.

References

AMMAR, G. S., AND GRAGG, W. B. 1988. Superfast solution of real positive definite Toeplitz systems. *SIAM Journal on Matrix Analysis and Applications 9*, 1, 61–76.

BHANDARI, A., FEIGIN, M., IZADI, S., RHEMANN, C., SCHMIDT, M., AND RASKAR, R. 2014. Resolving multipath interference in Kinect: An inverse problem approach. In *SENSORS, 2014 IEEE*, 614–617.

BHANDARI, A., KADAMBI, A., WHYTE, R., BARSI, C., FEIGIN, M., DORRINGTON, A., AND RASKAR, R. 2014. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Opt. Lett. 39*, 6 (Mar), 1705–1708.

BURG, J. P. 1975. *Maximum Entropy Spectral Analysis*. Ph.d. dissertation, Stanford University, Department of Geophysics.

CYBENKO, G., AND LOAN, C. V. 1986. Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. *SIAM Journal on Scientific and Statistical Computing 7*, 1, 123–131.

DORRINGTON, A. A., GODBAZ, J. P., CREE, M. J., PAYNE, A. D., AND STREETER, L. V. 2011. Separating true range measurements from multi-path and scattering interference in commercial range cameras. *Proc. SPIE 7864*, 786404-1–786404-10.

FREEDMAN, D., SMOLIN, Y., KRUPKA, E., LEICHTER, I., AND SCHMIDT, M. 2014. SRA: Fast removal of general multipath for ToF sensors. In *Computer Vision - ECCV 2014*, vol. 8689 of *Lecture Notes in Computer Science*. Springer International Publishing, 234–249.

FUCHS, S. 2010. Multipath interference compensation in time-of-flight camera images. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, 3583–3586.

GAO, L., LIANG, J., LI, C., AND WANG, L. V. 2014. Single-shot compressed ultrafast photography at one hundred billion frames per second. *Nature 516* (April).

GKIOULEKAS, I., LEVIN, A., DURAND, F., AND ZICKLER, T. 2015. Micron-scale light transport decomposition using interferometry. *ACM Trans. Graph. (Proc. SIGGRAPH 2015) 34*, 4 (July), 37:1–37:14.

GODBAZ, J. P., CREE, M. J., AND DORRINGTON, A. A. 2012. Closed-form inverses for the mixed pixel/multipath interference problem in AMCW lidar. *Proc. SPIE 8296*, 829618-1–829618-15.

GUPTA, M., NAYAR, S. K., HULLIN, M. B., AND MARTÍN, J. 2015. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Trans. Graph. (to appear)*.

HEIDE, F., HULLIN, M. B., GREGSON, J., AND HEIDRICH, W. 2013. Low-budget transient imaging using photonic mixer devices. *ACM Trans. Graph. (Proc. SIGGRAPH 2013) 32*, 4 (July), 45:1–45:10.

HEIDE, F., XIAO, L., KOLB, A., HULLIN, M. B., AND HEIDRICH, W. 2014. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Opt. Express 22*, 21 (Oct), 26338–26350.

JARABO, A., MARCO, J., MUÑOZ, A., BUISAN, R., JAROSZ, W., AND GUTIERREZ, D. 2014. A framework for transient rendering. *ACM Trans. Graph. (Proc. Siggraph Asia 2014) 33*, 6 (Nov.), 177:1–177:10.

JIMENEZ, D., PIZARRO, D., MAZO, M., AND PALAZUELOS, S. 2012. Modelling and correction of multipath interference in time of flight cameras. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 893–900.

KADAMBI, A., WHYTE, R., BHANDARI, A., STREETER, L., BARSI, C., DORRINGTON, A., AND RASKAR, R. 2013. Coded time of flight cameras: Sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. Graph. (Proc. SIGGRAPH Asia 2013) 32*, 6 (Nov.), 167:1–167:10.

KADAMBI, A., TAAMAZYAN, V., JAYASURIYA, S., AND RASKAR, R. 2015. Frequency domain ToF: encoding object depth in modulation frequency. *CoRR abs/1503.01804*.

KARLIN, S., AND STUDDEN, W. J. 1966. *Tchebycheff systems: with applications in analysis and statistics*. Pure and applied mathematics. Interscience Publishers.

KARLSSON, J., AND GEORGIOU, T. 2013. Uncertainty bounds for spectral estimation. *Automatic Control, IEEE Transactions on 58*, 7 (July), 1659–1673.

KIRMANI, A., BENEDETTI, A., AND CHOU, P. 2013. SPUMIC: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods. In *Multimedia and Expo (ICME), 2013 IEEE International Conference on*, 1–6.

KREĬN, M. G., AND NUDEL'MAN, A. A. 1977. *The Markov Moment Problem and Extremal Problems*, vol. 50 of *Translations of Mathematical Monographs*. American Mathematical Society.

LIN, J., LIU, Y., HULLIN, M. B., AND DAI, Q. 2014. Fourier analysis on transient imaging with a multifrequency time-of-flight camera. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 3230–3237.

NAIK, N., ZHAO, S., VELTEN, A., RASKAR, R., AND BALA, K. 2011. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. *ACM Trans. Graph. (Proc. SIGGRAPH Asia 2011) 30*, 6 (Dec.), 171:1–171:10.

PAYNE, A. D., DORRINGTON, A. A., CREE, M. J., AND CARNEGIE, D. A. 2010. Improved measurement linearity and precision for AMCW time-of-flight range imaging cameras. *Appl. Opt. 49*, 23 (Aug), 4392–4403.

QIAO, H., LIN, J., LIU, Y., HULLIN, M. B., AND DAI, Q. 2015. Resolving transient time profile in ToF imaging via log-sum sparse regularization. *Opt. Lett. 40*, 6 (Mar), 918–921.

VELTEN, A., LAWSON, E., BARDAGJY, A., BAWENDI, M., AND RASKAR, R. 2011. Slow art with a trillion frames per second camera. In *ACM SIGGRAPH 2011 Posters*, 13:1–13:1.

VELTEN, A., WILLWACHER, T., GUPTA, O., VEERARAGHAVAN, A., BAWENDI, M. G., AND RASKAR, R. 2012. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications 3* (March).

VELTEN, A., WU, D., JARABO, A., MASIA, B., BARSI, C., JOSHI, C., LAWSON, E., BAWENDI, M., GUTIERREZ, D., AND RASKAR, R. 2013. Femto-photography: Capturing and visualizing the propagation of light. *ACM Trans. Graph. (Proc. SIGGRAPH 2013) 32*, 4 (July), 44:1–44:8.

WU, D., VELTEN, A., O'TOOLE, M., MASIA, B., AGRAWAL, A., DAI, Q., AND RASKAR, R. 2014. Decomposing global light transport using time of flight imaging. *International Journal of Computer Vision 107*, 2, 123–138.