

Demultiplexing Illumination via Low Cost Sensing and Nanosecond Coding

Achuta Kadambi^{1*}, Ayush Bhandari¹, Refael Whyte², Adrian Dorrington², Ramesh Raskar¹
¹Massachusetts Institute of Technology ²University of Waikato

Abstract

Several computer vision algorithms require a sequence of photographs taken in different illumination conditions, which has spurred development in the area of illumination multiplexing. Various techniques for optimizing the multiplexing process already exist, but are geared toward regular or high speed cameras. Such cameras are fast, but code on the order of milliseconds. In this paper we propose a fusion of two popular contexts, time of flight range cameras and illumination multiplexing. Time of flight cameras are a low cost, consumer-oriented technology capable of acquiring range maps at 30 frames per second. Such cameras have a natural connection to conventional illumination multiplexing strategies as both paradigms rely on the capture of multiple shots and synchronized illumination. While previous work on illumination multiplexing has exploited coding at millisecond intervals, we repurpose sensors that are ordinarily used in time of flight imaging to demultiplex via nanosecond coding strategies.

1. Introduction

Capturing a scene under a variety of lighting conditions provides a rich collection of information that is used in several practical computer vision contexts. Such information, when combined with computational processing, enables techniques that reconstruct surface normals [36], recover depth edges [25], compute light transport parameters [21], or simply relight a scene from an illumination basis [17]. For static scenes, the capture process is as straightforward as capturing different illuminations in a time-sequential manner, but unfortunately, such an approach may not scale to dynamic scenes. A valid option is to use a high speed camera setup [34], but cost and synchronization requirements limit the broader applicability.

In this paper, we use time of flight cameras to time multiplex illumination sources. Time of flight cameras represent an increasingly popular method to acquire 3D maps in real-time and unlike high speed cameras, are geared toward low-cost consumer applications. They measure depth

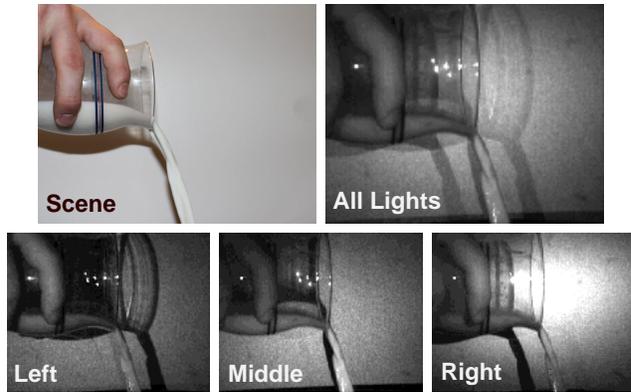


Figure 1: Time multiplexing a dynamic scene can be challenging. Here milk is being poured from a cup and the measurement (upper right) includes contributions from all three lights illuminating the scene. Coded demodulation allows us to demultiplex the scene into the individual illuminants. In contrast to previous approaches this technique exploits low cost time of flight sensors which allow nanosecond coding within the millisecond exposure.

by sending an optical signal to the scene and measuring the time it takes to arrive back at the sensor. Because they operate at a fixed viewpoint and probe the scene directly they avoid some of the pitfalls of stereo-based methods caused by occlusions or lack of texture. Recent work in computer graphics has been successful at repurposing conventional ToF technology to acquire 3D models of translucent objects, or create visualizations of light transport [13, 11]. In this paper, we extend the scope of ToF technology to the illumination multiplexing problem in computer vision. We note that the connection seems natural as such cameras already require multiple shots and time-synchronized illumination to obtain depth.

Our key contribution:

- Demultiplexing individual light sources by repurposing the system on chip hardware used for time of flight cameras.

Secondary technical contributions:

- Analysis of different strategies for ToF illumination multiplexing.
- Demonstration of a real time multispectral depth camera that demultiplexes lights of different wavelengths.

*achoo@mit.edu

2. Related Work

Dynamic relighting of video is of great interest to the motion picture and performance capture industries. In a conventional setup, time sequential multiplexing is only suitable for static scenes, but Wenger *et al.* [34] construct a synchronized high speed camera and light rig. The results from [34] also considered different, well-known bases for time multiplexing, but the focus of the paper is on production-quality relighting. As an extension, De Decker *et al.* [6] propose coding in both time and color to relight dynamic scenes. Such an approach relaxes the framerate requirements of the camera by, approximately, a factor of three [14].

Multiplexing illumination has been a well explored problem in the vision and graphics communities. In Schechner *et al.* [30], a detailed analysis of optimal multiplexing matrices was performed in the context of additive noise. In comparison to the naive case, where a single light is on at a given time, the authors suggest using a Hadamard matrix as the multiplexing matrix, which is a concept similar to earlier work in spectroscopy [10]. In more recent literature, Ratner *et al.* [28, 26] formulate a constrained optimization program to find time multiplexing codes for a more general setting. Recently, Chan *et al.* [5], extend this technique by setting an optimization program for multiplexing in both time and color. In [1], they develop a new theory of multiplexing in the context of fluorescence imaging. Illumination multiplexing has also been explored in structured light systems. In [9] spatial frequency multiplexing of illumination patterns has been explored to separate direct and global light paths for different illumination sources. This technique is used primarily in the context of performing direct and global separation on static scenes, and has been shown to achieve results using fewer measurements than the checkerboard method in [20]. In [22], a monochrome video camera is combined with multiplexing of illumination sources to capture an RGB image of the scene.

Time of flight technology is primed to be a significant RGBD sensor for the robotics and vision communities (it forms the basis for the new Microsoft Kinect). Significant advantages over structured light version include fewer moving parts, resistance to occlusions and ambient lighting, and increased range accuracy. In the ToF literature, [35] has explored allocating coding strategies to different cameras so as to avoid interference. In [4] they use ToF cameras in a stereo configuration, at different frequencies to increase the accuracy of range maps. Multiplexing illumination for ToF cameras—where one ToF camera captures and demodulates different light sources—has yet to be studied. In terms of color, the ToF camera in [18] time multiplexes red green and blue LEDs, but the technique is for static scenes. Recently, Kim *et al.* [15] report the first real-time color RGBD sensor, by redesigning the silicon sensor to multiplex both

color and depth pixels (this decreases the depth pixel count by a factor of 4). Time of flight technology is of interest to the broader community beyond range imaging; see recent work in ultrafast imaging [24, 13, 11, 33, 12], BRDF acquisition [19], multi-path imaging [13], [3], [2], and looking around corners [32].

3. Theory

3.1. The Multiplexing Problem

The problem of time multiplexing illumination is cast as a pixel-wise linear inverse problem

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta} \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ is a vector of unknowns, $\mathbf{y} \in \mathbb{R}^m$ is a measurement vector, $\mathbf{H} \in \mathbb{R}^{m \times n}$ is a known matrix, and $\boldsymbol{\eta} \in \mathbb{R}^m$ is an additive noise vector with zero mean and variance σ^2 . For the specific context of illumination multiplexing, the vector \mathbf{x} represents the amplitudes of n light sources, the vector \mathbf{y} represents the captured images at m time points, and each column vector of the multiplexing matrix, \mathbf{H} , represents the intensity modulation of the light source in time. We have simplified (1) to the case of real-valued parameters with the additional constraint that elements of \mathbf{H} are between 0 and 1, corresponding to the illumination sources being fully off and on. As an example, consider sequential time multiplexing, which describes a system where only one light is on at a given time. The inverse problem is then

$$\mathbf{y} = \mathbf{I}_n \mathbf{x} + \boldsymbol{\eta} \quad (2)$$

where the multiplexing matrix is now $\mathbf{H} = \mathbf{I}_n$. To recover the intensities of the light sources from (2) we simply invert the system: $\hat{\mathbf{x}} = \mathbf{I}_n^{-1} \mathbf{y}$. In the multiplexing lexicon, such an attempt is labelled the *trivial* formulation. For the more general case in (1), we can compute the mean squared error of \mathbf{x} as

$$\text{MSE}_{\hat{\mathbf{x}}} = \frac{1}{n} \sigma^2 \text{tr} \left[(\mathbf{H}^\top \mathbf{H})^{-1} \right] \quad (3)$$

where tr represents the trace operator and n the number of light sources that we desire to multiplex. Recall that the MSE corresponds to the variance for an unbiased estimator, which allows a relation between mean squared error to a gain in SNR:

$$\text{Gain} = \sqrt{\frac{\sigma^2}{\text{MSE}_{\hat{\mathbf{x}}}}} \quad (4)$$

The goal is then to find a multiplexing matrix, which is optimal in the sense of maximizing SNR, or equivalently, minimizing $\text{MSE}_{\hat{\mathbf{x}}}$:

$$\arg \min_{\mathbf{H}} \text{MSE}_{\hat{\mathbf{x}}}, \quad \text{s.t. } \mathbf{H} \in \mathcal{H} \quad (5)$$

where \mathcal{H} is a set of domain-specific constraints. Design of the optimal multiplexing matrix for the i.i.d. case in (3) is not a new problem (c.f. [29, 34, 26, 23, 10]).

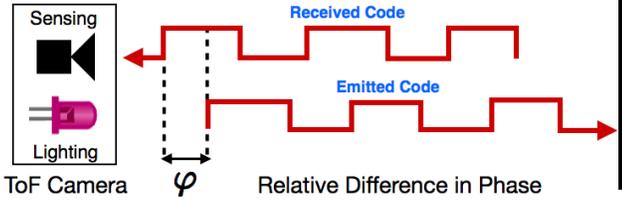


Figure 2: Operation of ToF range cameras. The time difference of arrival between the emitted and received code is encoded in the phase offset.

3.2. Overview of Time of Flight Operation

A ToF camera is for the most part a regular camera with a timecoded illumination circuit. ToF cameras are usually used to estimate depth, where the process is as follows. The camera strobes a light source using a discrete time illumination control signal $i_\omega[t]$, $t = 0, \dots, n$, and receives, from the scene environment, a discrete time optical signal $\xi_\omega[t]$. The subscript ω represents the fundamental/strobing frequency of the signal. The relative difference in phase, φ between $\xi_\omega[t]$ and a discrete time reference signal $r_\omega[t]$ is used to calculate the physical distance of objects:

$$d = \frac{c\varphi}{4\pi\omega} \quad (6)$$

where c is the speed of light and d is the distance. A graphical overview of this process is illustrated in Figure 2. We will denote discrete signals using vector notation,

$$\mathbf{i}_\omega = \text{vec}(i_\omega[t]), \quad \mathbf{i}_\omega \in \mathbb{R}^m. \quad (7)$$

The dimensions of ToF signals are as follows; $\mathbf{r}_\omega \in \mathbb{R}^m$, $\xi_\omega \in \mathbb{R}^m$ and $\mathbf{i}_\omega \in \mathbb{R}^m$.

Calculating phase: To calculate phase the measured optical signal is crosscorrelated with the reference code:

$$\mathbf{c}_\omega = \xi_\omega \otimes \mathbf{r}_\omega = \xi_\omega * \bar{\mathbf{r}}_\omega \quad (8)$$

where \otimes denotes the cross correlation operator and $*$ the convolution operator, \mathbf{c}_ω denotes the cross-correlation vector, and $\bar{(\cdot)}$ denotes a flip of the vector, that is $\bar{f}[t] = f[-t]$. The phase offset, φ , corresponds to the location of the maximum in \mathbf{c}_ω , or in Fourier domain, the phase of the fundamental frequency, ω . Equation (8) can also be formulated in frequency domain using the Fourier duality,

$$\mathbf{c}_\omega = F^{-1} \{F \{\mathbf{c}_\omega\}\} \Leftrightarrow F^{-1} \{F \{\xi_\omega\} \cdot F \{\bar{\mathbf{r}}_\omega\}\} \quad (9)$$

where $F \{\cdot\}$ denotes the Fourier transform. We have discussed the case when a single light is strobed. We can have n number of illumination control vectors $\mathbf{i}_{\omega,1}, \dots, \mathbf{i}_{\omega,n}$, but only one of the reference and optical signals.

The cross-correlation occurs at the sensor-level of the ToF camera, which cements an association between such

cameras and cross-correlation. In particular, to obtain m samples of the correlation vector, the ToF camera acquires m pictures of the scene, where in each picture, the signals, ξ_ω and \mathbf{r}_ω are at a different phase offset. In that sense, the time of flight camera is a multi-shot imaging technique with the additional caveat of a hardware level cross-correlation.

Typical Codes: It is possible to set both the reference signal and control signal of the light sources, i.e., we have direct control over \mathbf{r}_ω and \mathbf{i}_ω and, by extension control of ξ_ω . For the purposes of ToF ranging, conventional cameras use a square wave for the reference and illumination signals Figure 2. The period of these signals is on the order of nanoseconds.

Restricted Codes: We now remark that the reference code cannot be a delta function, i.e., it cannot have a support of only one bit. There are two reasons for this. First, in the hardware a bit is sent for a finite amount of time, which turns the delta code into a very narrow square code. Second, by only having a support of one bit the SNR decreases. This is an important point to make. A similar remark can be found in [13].

4. Time of Flight Multiplexing

4.1. Conventional Multiplexing is Millisecond Coding

As illustrated in Figure 3, the forward model for modulated sensors, such as in time of flight, includes an additional matrix multiplication with a heavily underdetermined decimation matrix. The support of each row within this matrix corresponds to a shifted version of the reference code. Intuitively, the support of each row corresponds to the *coding* at the sensor level. In a regular high speed camera operating at millisecond framerates, all the light during the exposure of one frame contributes to the image and hence there is no coding within the exposure. For multiplexing, the illumination coding matches the exposure time, and thus operates at millisecond time scales.

4.2. Nanosecond Coding within the Exposure

Now, we turn to coding within each millisecond exposure using a sensor from time of flight camera technology. As before, the measurements are still taken at millisecond intervals — the camera is not any faster — but the clock on the sensor supports nanosecond synchronization. This allows for additional coding *within* each one millisecond exposure. In particular, as illustrated in cases two and three of Figure 3 sub millisecond coding can occur at the sensor level and at the illumination level. Case two reflects conventional time of flight camera operation where a single frequency square wave codes the exposure. These codes are designed to lock on to strobed illumination at a single frequency and cannot multiplex lights at different frequencies.

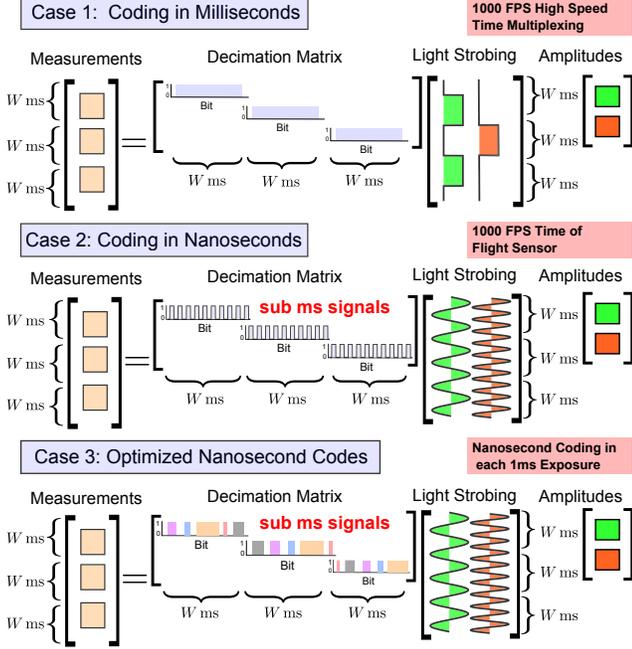


Figure 3: Comparison a ToF camera with a high speed camera for the same duration of exposure time $\sim W$ milliseconds. In all cases, the Decimation matrix is a block-diagonal matrix where block-wise elements are circularly shifted versions of previous diagonal blocks. This leads to an analogy with Toeplitz matrices and convolution. **Case 1:** The forward model is presented for a fast videocamera where measurements are taken each W milliseconds. For a single measurement, the camera integrates all light equally over the exposure. Hence there is no coding of exposure in this case. **Case 2:** Considers a ToF camera-like sensor where each W millisecond exposure is coded. Hence it is able to achieve nanosecond coding by frequency locking a reference clock signal with the illumination. Because the frequencies are locked it is biased to the contribution from the similar strobing frequency. **Case 3:** By carefully coding the W millisecond, exposure this sensor can demultiplex illumination using nanosecond scale codes.

	Camera Exposure	Timescale of Light Strobing	Timescale of Sensor Codes	Demultiplexed Lighting	Demultiplexed Depth
Time Multiplexing	W_{ms}	Millisecond	N/A	Yes	No
Time of Flight Coding	W_{ms}	Nanosecond	Nanosecond	No	No
ToF Optimal	W_{ms}	Nanosecond	Nanosecond	Yes	Yes

Figure 4: Comparing the techniques illustrated in Figure 3.

By modifying the time of flight sensor we are allotted control over the codes used in the decimation matrix. Consequently, the third case in Figure 3 represents one possible solution for the decimation and strobing matrices. Optimizing these codes is a challenging problem that involves proper consideration of camera noise sources and other system parameters (See Appendix).¹ In our approach we use a simple argument about preconditioners to select the codes

¹Previous papers on multiplexing illumination involve an optimization. To tailor such work to ToF, equation 3 would have to be modified; we leave this discussion to the appendix.

(Section 5.2).

We now write the forward and inverse model. As illustrated in Figure 3, the measurement at each 1ms step represents the inner product of the exposure code and the corresponding 1ms of the strobing signal. To simplify the model we will consider *periodic* forms for the strobing signal where the period is much less than the exposure time, e.g., on the order of nanoseconds. Then, each 1ms block of the strobing signal is assumed to be identical and we can recast the problem in terms of convolution. In particular, the sensor used on ToF cameras measures a linear combination of the different convolutions between the environment vectors $\xi_{\omega,1}, \dots, \xi_{\omega,n}$ and the reference code r_{ω} . The resulting system is

$$\mathbf{c}_{\omega} = \begin{bmatrix} \uparrow & & \uparrow \\ \xi_{\omega,1} \otimes \mathbf{r}_{\omega} & \cdots & \xi_{\omega,n} \otimes \mathbf{r}_{\omega} \\ \downarrow & & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}. \quad (10)$$

Accounting for noise, and simplifying the system compactly, we have

$$\mathbf{c}_{\omega} = \mathbf{T}\mathbf{L}\mathbf{x} + \boldsymbol{\eta} \quad (11)$$

where $\mathbf{T} \in \mathbb{R}^{m \times p}$ is a circulant Toeplitz matrix formed from the p length reference code, and the k^{th} column of $\mathbf{L} \in \mathbb{R}^{p \times n}$ is the optical signal $\xi_{\omega,k}$, $\mathbf{x} \in \mathbb{R}^{n \times 1}$, and $\mathbf{c}_{\omega} \in \mathbb{R}^{m \times 1}$. Additive noise characterization in ToF cameras is expressed as covariance matrix, $\boldsymbol{\Sigma}$ of the noise. To estimate $\hat{\mathbf{x}}$, consider the best linear unbiased estimator (BLUE) for (11), that is,

$$\hat{\mathbf{x}}_{\text{BLUE}} = (\mathbf{L}^{\top} \mathbf{T}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{T} \mathbf{L})^{-1} \mathbf{L}^{\top} \mathbf{T}^{\top} \boldsymbol{\Sigma}^{-1} \mathbf{c}_{\omega} \quad (12)$$

which is a weighted least squares on the positive semidefinite matrix $\boldsymbol{\Sigma}$. Our task is to find the optimal codes, i.e., the single reference vector that determines \mathbf{T} and the n column vectors $\xi_{\omega,1}, \dots, \xi_{\omega,n}$ that comprise \mathbf{L} .

5. Selecting Reference and Illumination Codes

5.1. Possibility of Convex Optimization

A reasonable option is to write an optimization program for Equation 11 for the decision variables \mathbf{T} and \mathbf{L} . The optimization strategy we sketch in the appendix involves characterizing all of the noise sources in the model, then solving a semidefinite program, and finally another optimization for non-negative matrix factorization. However, we realize that it may not always be scalable to write such an optimization. Readers are directed to the appendix for the complete sketch of an optimization strategy.

5.2. Shortcut via Preconditioning

In this section we propose a simple rule of thumb that *completely avoids* using an optimization and does not require knowledge of the noise sources in the scene. There

are trade-offs, e.g., the selected codes may not be optimal in the sense of SNR, but the approach is much simpler and still allows multiplexing of RGBD.

Connection to Preconditioning: The linear inverse problem for ToF cameras (11) can be viewed as a preconditioned version of the classic time multiplexed inverse problem. For noiseless measurements, we can rewrite (1) in terms of the ToF signals:

$$\mathbf{y} = \mathbf{H}\mathbf{x} \mapsto \mathbf{y} = \sum_{k=1}^n \xi_{\omega,k} \equiv \mathbf{y} = \mathbf{L}\mathbf{x}. \quad (13)$$

Since $\mathbf{c}_\omega = \mathbf{T}\mathbf{L}\mathbf{x} = \mathbf{T}\mathbf{y}$, the preconditioning step is:

$$\mathbf{y} = \mathbf{L}\mathbf{x} \xrightarrow{T(\cdot)} \mathbf{T}\mathbf{y} = \mathbf{T}\mathbf{L}\mathbf{x}. \quad (14)$$

In crux, circulant matrices offer an intuitive viewpoint of preconditioning: the Fourier spectrum of the vector code corresponds to the eigenvalues of the matrix. Choosing the correct reference code allows the preconditioner operator to stabilize the eigenvalues of the inversion, which makes the overall problem better conditioned. Of course, choosing the wrong matrix to precondition will only make the problem much worse. For instance, the classic form for \mathbf{r}_ω in ToF cameras is a box function, which, in its Toeplitz form \mathbf{T} is characterized by a very high condition number. Finding the optimal forms of \mathbf{T} and \mathbf{L} is therefore an important design consideration.

A justifiable choice for \mathbf{M} is an orthogonal matrix [9, 30] to exploit, among other properties, the fact that an orthogonal matrix has a condition number of 1.²

Proposition 5.1. *Suppose \mathbf{M} is orthogonal. Then, \mathbf{L} is an orthogonal matrix and the optimal code \mathbf{r}_ω is a code that is broadband in frequency domain.*

Proof. To prove the first part note that $\mathbf{M}^\top\mathbf{M} = \mathbf{I}$ and therefore $\mathbf{L}_i^\top\mathbf{T}^\top\mathbf{T}\mathbf{L}_j = 0 \forall i \neq j$. The second follows from the properties of condition numbers. If \mathbf{L} is orthogonal, then the optimal preconditioner matrix has all eigenvalues equal. Since \mathbf{T} is a circulant matrix, the spectrum of \mathbf{r}_ω corresponds to the eigenvalues and must be spectrally flat. \square

In summary, the rule of thumb for codes: (i) pick a reference code that is broadband (ii) pick codes for lights that are orthogonal. As a concrete example, one option would be to strobe light sources at harmonic sine functions and use an m-sequence as the reference code. Proposition 5.1 underscores the importance of having a spectrally flat code, but recall from section 3.2 that the code cannot be a delta function. Therefore, if solving an optimization is out of

²The orthogonal Hadamard matrix in [30] achieves multiplexed gain boost when photon shot noise is negligible

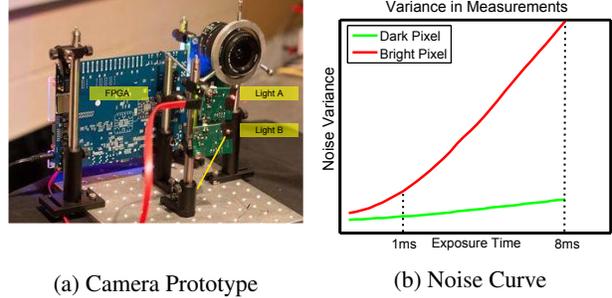


Figure 5: (a) A prototype of the camera with different light sources labeled. This is a conventional time of flight camera, but the readout and modulation is handled by an FPGA, which allows customized reference and strobing codes. (b) Example plot of variance vs integration (i.e., exposure) time for two pixels. The y-intercept of the curve corresponds to signal independent noise sources.



Figure 6: From left: the measurement, the right illumination, the left illumination. These results match ground truth (not shown).

scope, we recommend a form similar to case three in Figure 3, where the strobing signals are harmonic sine waves and the reference code is broadband.³

6. Implementation

The technique is validated using a custom built Time-of-Flight range camera (Figure 5a) that allows the design to be modular and customizable. The camera is designed around a PMD19k-3 sensor which has an array size of 120 by 160 pixels. The sensor interfaces to a Cyclone IV E FPGA. The FPGA controls the modulation signal (frequency and binary sequence), integration period, readout timing and the UDP interface to the host computer. The pixel values are converted to 16bit unsigned values by an Analog Devices AD9826 ADC. Multiple solid state light sources are used: to reproduce your own design please ensure that the illumination supports high frequency strobing.

7. Results

Figure 7 shows a potential consumer application of lighting. Here, the scene is lit from two illumination sources and can be demultiplexed to reflect contributions from only one of the sources. Note the two shadows in the composite image. Observe that in the demultiplexed version, the left and right light illumination sources are isolated, and as expected, there is only one shadow present in each image. Note that in these two examples, no significant channel

³For case 3 in Fig. 3 the following is a reasonable broadband code: 1010011010000110001010011010000000010000000101000000100000011001



Figure 7: Because the technique is inherently low cost, consumer applications, such as this relighting may be possible. A composite measurement with two lights and the demultiplexed results.

cross-talk is observed in the images.

As illustrated in Figure 9 multiplexing three RGB illumination sources does not impact framerate and we can capture video. Using custom codes does not impact the ability to acquire depth. In particular, by combining color and depth information we are able to generate combined color-depth maps (8). These images are unprocessed, and of course, color balancing techniques can be applied for the desired color reproduction. An imperfection in our technique is due to the geometry of our lights. For example, there is an artifact in Figure 8 due to the corner of the red object receiving a large amount of light. In addition, because the light sources are not co-located, shadows and occlusions cause color imbalances (green shadow around blue box in Figure 8). Multiplexing color still allows the depth information to be preserved, e.g., Figure 8 shows the collection of a depth map without red objects.

Multiplexing with ToF provides no benefit in the number of light sources that can be multiplexed, i.e., the calculation still depends on the number of measurements. As our prototype camera requires, at a minimum 2ms—including readout and exposure—to sample one step of the correlation waveform it provides—in the best case—500 measurements per second. Therefore we can multiplex up to about $n \approx 15$ lights at 30 fps. Since our prototype costs around 500 dollars, this is a very good value in the context of comparable high speed cameras.⁴ Because a ToF camera inherently requires multiple shots for one depth frame, multiplexing up to 10 illumination sources is possible without sacrificing temporal resolution. This enables dynamic relighting as illustrated in Figure 1.

Failure Cases: In Figure 10 multiple artifacts can be seen. These stem from a key source: motion artifacts. In the demultiplexed channels, we can see motion blur artifacts. This is inherent to even current state of the art time of flight cameras and multiplexing illumination sources has no positive or negative impact. However, the color image at right includes artifacts beyond motion blur. Concretely, the shad-

⁴The cost of our prototype vs low-end high speed is comparable, but the advantage of ToF is it includes the synchronization and modulated lighting. Production grade ToF cameras are now a consumer device for entertainment.

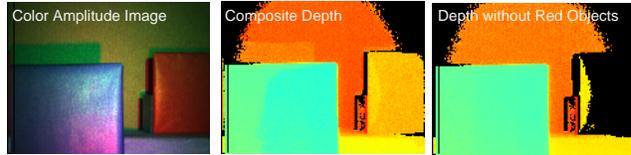


Figure 8: For ToF cameras multiplexing illumination also applies to depth maps. From left, the color ToF image of the scene, the composite depth map, and the depth map with reds subtracted.



Figure 10: A failure case of our technique. A Rubik's cube is rotated rapidly and motion artifacts are observed in the demultiplexed images as well as color artifacts in the color image. Please see text for explanation.

ows cast in the demultiplexed RGB channels lead to shadows in the color image. This is a problem that can easily be solved by locating the three color light sources at the same spatial location.

8. Discussion

8.1. Limitations

In this paper, we do not address the motion blur problem, which is an active and important topic in ToF range imaging. Using current time of flight system on chip architecture we can multiplex up to 15 light sources in real time. However, for each additional light it is necessary to fabricate and provide additional power for illumination. In our results we used laser diodes but using LEDs is both cheaper and more aesthetic (due to lack of speckle). In terms of multispectral imaging, we demonstrate a system that multiplexes only RGB illuminations, but the results generalize to more spectral bands. The rule of thumb we have proposed (Proposition 5.1) may not be useful for all noise models.

8.2. Comparisons

Nanosecond vs Millisecond Coding: Existing literature for time multiplexing in computer vision focuses on millisecond coding. Although a time of flight sensor has a similar capture speed, it is possible to perform nanosecond coding within each millisecond exposure. We believe there are some concrete benefits in the fusion of nanosecond coding and ToF implementation. First, the physical machinery already exists for synchronized, time-coded illumination using one SoC. This may lead to simpler and lower cost implementations compared to a high speed camera. Second, nanosecond coding preserves time of flight or depth information, which when combined with demultiplexing may impact techniques that are germane to contexts of both depth and illumination (for instance photometric stereo). Finally, the ToF camera fundamentally performs a matching between the nanosecond strobing and ref-



Figure 9: The demultiplexed frames are used to recover real-time single-chip RGB time of flight. This is the raw result without white balancing or image processing techniques.

erence codes, which mitigates ambient and fixed pattern noise sources.⁵ We think there is more to be explored with nanosecond coding.

Fusing Color and Depth: This paper presents the first *single-chip* RGBD time of flight sensor, which raises the question if this is a useful contribution over the conventional approach (to pair a depth camera with a video camera, e.g., Kinect). There are some reasonable arguments in favor of the single-chip sensor. For instance, using a single sensor avoids the need to register the two cameras or synchronize their framerates. In addition, specifically for time of flight, the images ignore ambient light. This may be important for machine vision applications, where ambient shadows are a known challenge. Finally, using multiple illumination sources can be used to statistically *correct* depth maps. In light of these benefits, research in single-chip RGBD is an emerging area. Our results compare well to the previous papers on single-sensor RGBD imaging, where in [15] spatial resolution is sacrificed and in [18] the acquisition is not real time.

8.3. Applications and Conclusion

The second generation Kinect (that uses ToF technology) is only one mark of the increasing popularity of time of flight sensors. It is speculated that similar sensors, once they adapt to low power scenarios, will be implemented in mobile platforms. In the consumer space often what distinguishes the quality of a photo is the lighting. Here, the relighting application may be useful where settings such as no-flash (ambient), light one, and light two can be offered on a device. Where a camera like Lytro offers the ability to “shoot now, focus later”, we consider the ability of “shoot now, relight later” to be just as important. For machine vision applications, where registration between color and depth images is critical, the single sensor RGBD imager may have scope. The key advantage is that our multiplexing method works in real-time and does not sacrifice spatial resolution (cf. [15]).

In summary, we have considered illumination multiplexing problem for ToF cameras. We remark that all results in

⁵Intuitively, if the signal we want is strobing in some nanosecond coded sequence, we ignore contributions from (noise and ambient) sources that do not meet this criteria.

the paper are demonstrated at real-time framerates.

Appendix

A1: Semidefinite Code Optimization

For completeness we sketch the approach for convex optimization of reference and illumination codes. We can avoid optimizing the gain directly for \mathbf{T} and \mathbf{L} by simplifying (11) as follows:

$$\mathbf{c}_\omega = \mathbf{M}\mathbf{x} + \boldsymbol{\eta} \quad (15)$$

where we have substituted \mathbf{M} for the product of \mathbf{T} and \mathbf{L} . This allows us to first derive the optimal \mathbf{M} , which is an easier problem to solve as there are no constraints on matrix structure. As before, $\boldsymbol{\Sigma}$ is the covariance model for the vector \mathbf{x} :

$$\boldsymbol{\Sigma} = \mathbb{E} \left[(\hat{\mathbf{x}} - \mathbb{E}[\hat{\mathbf{x}}]) (\hat{\mathbf{x}} - \mathbb{E}[\hat{\mathbf{x}}])^\top \right] \quad (16)$$

where \mathbb{E} denotes the expectation operator. The resulting mean squared error is

$$\text{MSE} = \frac{1}{n} \text{tr} \left[\left(\mathbf{M}^\top \boldsymbol{\Sigma}^{-1} \mathbf{M} \right)^{-1} \right] \quad (17)$$

This is the more general form of (3), where the noise was i.i.d., which allowed factorization of the variance outside the trace. Here, $\boldsymbol{\Sigma}$ is independent (but not necessarily identically distributed), and so $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \dots, \sigma_m^2)$. As in [30, 9, 1, 27] the noise sources can be decomposed into various sources. For the time of flight camera, the variance in the k^{th} measurement is

$$\sigma_k^2 = \sigma_{\text{read}}^2 + \sigma_{\text{clock}}^2 + \sigma_{\text{dark}}^2 + \sigma_{\text{photon}}^2 \quad (18)$$

where photon noise is proportional to the incoming signal, and clock noise is proportional to the smoothness of the reference code \mathbf{r}_ω . Read noise and dark noise are signal independent. We assume an ideal system where the noise sources are solely due to the sensor and not due to discrete components, such as the analog to digital converter (ADC). In addition, we will ignore clock noise as it is usually negligible in context of read and shot noise. By taking into account the box constraint, (17) is cast as an optimization program:

$$\arg \min_{\mathbf{M}} \frac{1}{n} \text{tr} \left[\left(\mathbf{M}^\top \boldsymbol{\Sigma}^{-1} \mathbf{M} \right)^{-1} \right] \quad \text{s.t.} \quad \mathbf{1} \succeq \text{vec}(\mathbf{M}) \succeq \mathbf{0} \quad (19)$$

where \succeq denotes a component-wise inequality for vectors and linear matrix inequalities for matrices. We remark that this objective function arises in many domains and can be seen, more generally, as an estimate of the Cramér-Rao lower bound for (1). Note that (19) is more complicated to optimize when compared to (3) that has been studied in [30, 23, 27]. Recent literature [5] recasts this problem as a semidefinite program,

$$\begin{aligned} \mathbf{M}^* &= \arg \min_{\mathbf{Q}} \text{tr}(\mathbf{Q}) \quad \text{s.t.} \\ \mathbf{1} \succeq \text{vec}(\mathbf{M}) \succeq \mathbf{0}, \quad \mathbf{Q} &\succeq \left(\mathbf{M}^\top \boldsymbol{\Sigma}^{-1} \mathbf{M} \right)^{-1} \end{aligned} \quad (20)$$

where \mathbf{Q} is an auxiliary variable and $\mathbf{Q} - \left(\mathbf{M}^\top \boldsymbol{\Sigma}^{-1} \mathbf{M} \right)^{-1}$ is a positive semidefinite matrix. This program is not convex, but after linearization can be solved to yield the optimal matrix for time multiplexing \mathbf{M}^* .

Optimizing \mathbf{T} and \mathbf{L} can now be posed as a matrix factorization problem where

$$\mathbf{M}^* = \mathbf{T}^* \mathbf{L}^*. \quad (21)$$

Recall for the physical system, we must ensure that \mathbf{T}^* is a circulant Toeplitz matrix and values of \mathbf{T}^* and \mathbf{L}^* are constrained between 0 and 1. This problem can be posed as a nonnegative matrix factorization (NMF) with constraints

$$\begin{aligned} \{\mathbf{T}^*, \mathbf{L}^*\} = \arg \min_{\mathbf{T}, \mathbf{L}} \|\mathbf{M}^* - \mathbf{T}\mathbf{L}\|_F \quad \text{s.t.} \\ 1 \succcurlyeq \text{vec}(\mathbf{T}) \succcurlyeq 0, \quad 1 \succcurlyeq \text{vec}(\mathbf{L}) \succcurlyeq 0, \quad \mathbf{T} \in \mathcal{C} \end{aligned} \quad (22)$$

where the set $\mathcal{C} \subset \mathbb{R}^{m \times p}$ represents the subspace of circulant Toeplitz matrices and $\|\cdot\|_F$ denotes the Frobenius norm. Ignoring the circulant constraint for the moment, there are several valid techniques to solve (22), ranging from convex approaches [7, 16] to explicit update rules [31]. We use the former, which can be readily implemented in CVX [8]. In order to account for the last constraint we solve the following optimization program to find the closest circulant approximation:

$$\arg \min_{\mathbf{C} \in \mathcal{C}} \|\mathbf{T}^{(k)} - \mathbf{C}\|_F \quad (23)$$

where $\mathbf{T}^{(k)}$ is the k^{th} iterative update. This program can be solved in closed form, where graphically speaking, the p^{th} entry of the circulant matrix is the mean value of the p^{th} diagonal of $\mathbf{T}^{(k)}$.

A2: Implications of Noise on Coding

There are two likely noise-limiting scenarios for this camera: read or shot noise limited. Figure 5b shows the noise variance of the camera as a function of exposure time. The y-intercept of this curve corresponds to the signal-independent noise sources, e.g. read noise and dark noise. The slope of the curve illustrates the impact of photon shot noise. Concretely, it is observed that at the dim pixel the system is read noise limited and at the bright pixel shot noise can be a significant factor. The type of noise regime is to be incorporated into the design of optimal codes. In particular, if the system is shot noise limited, then illumination multiplexing, such as Hadamard multiplexing, might actually perform worse than trivial multiplexing (see [30] and the covariance matrix in the appendix for details).

References

- [1] M. Alterman, Y. Y. Schechner, and A. Weiss. Multiplexed fluorescence unmixing. In *Computational Photography (ICCP), 2010 IEEE International Conference on*, pages 1–8. IEEE, 2010.
- [2] A. Bhandari, A. Kadambi, and R. Raskar. Sparse linear operator identification without sparse regularization? applications to mixed pixel problem in time-of-flight / range imaging. In *Proc. of Intl. Conf. Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May 2014.
- [3] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorington, and R. Raskar. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Opt. Lett.*, 39(6):1705–1708, Mar 2014.
- [4] V. Castaneda, D. Mateus, and N. Navab. Stereo time-of-flight. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 1684–1691. IEEE, 2011.
- [5] T.-H. Chan, K. Jia, E. Wycoff, C.-Y. Chi, and Y. Ma. Towards optimal design of time and color multiplexing codes. In *Computer Vision—ECCV 2012*, pages 485–498. Springer, 2012.
- [6] B. De Decker, J. Kautz, T. Mertens, and P. Bekaert. Capturing multiple illumination conditions using time and color multiplexing. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 2536–2543. IEEE, 2009.
- [7] C. H. Ding, T. Li, and M. I. Jordan. Convex and semi-nonnegative matrix factorizations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32(1):45–55, 2010.
- [8] M. Grant, S. Boyd, and Y. Ye. Cvx: Matlab software for disciplined convex programming, 2008.
- [9] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar. Multiplexed illumination for scene recovery in the presence of global illumination. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 691–698. IEEE, 2011.
- [10] M. Harwit and N. Sloane. Hadamard transform optics. 1979.
- [11] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *Technical Paper to appear at SIGGRAPH 2013*, 2013.
- [12] A. Jarabo, J. Marco, A. Muñoz, R. Buisan, W. Jarosz, and D. Gutierrez. A framework for transient rendering. *ACM Transactions on Graphics (SIGGRAPH Asia 2014)*, 33(6), 2014.
- [13] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Transactions on Graphics (TOG)*, 32(6):167, 2013.
- [14] H. Kim, B. Wilburn, and M. Ben-Ezra. Photometric stereo for dynamic surface orientations. In *Computer Vision—ECCV 2010*, pages 59–72. Springer, 2010.
- [15] S.-J. Kim, B. Kang, J. D. Kim, K. Lee, C.-Y. Kim, and K. Kim. A 1920× 1080 3.65 μm-pixel 2d/3d image sensor with split and binning pixel structure in 0.11 pm standard cmos. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2012 IEEE International*, pages 396–398. IEEE, 2012.
- [16] C.-J. Lin. Projected gradient methods for nonnegative matrix factorization. *Neural computation*, 19(10):2756–2779, 2007.
- [17] F. Moreno-Noguer, S. K. Nayar, and P. N. Belhumeur. Optimal illumination for image and video relighting. In *Proc. IEE Conf. Visual Media Production*, pages 201–210, 2005.
- [18] A. Muttayane. *Towards Colour Imaging with the Image Ranger*. PhD thesis, The University of Waikato, 2006.
- [19] N. Naik, S. Zhao, A. Velten, R. Raskar, and K. Bala. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. In *ACM Transactions on Graphics (TOG)*, volume 30, page 171. ACM, 2011.
- [20] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Transactions on Graphics (TOG)*, 25(3):935–944, 2006.
- [21] M. O’Toole, R. Raskar, and K. N. Kutulakos. Primal-dual coding to probe light transport. *ACM Trans. Graph.*, 31(4):39, 2012.
- [22] J.-I. Park, M.-H. Lee, M. D. Grossberg, and S. K. Nayar. Multispectral imaging using multiplexed illumination. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–8. IEEE, 2007.
- [23] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 795–804. ACM, 2006.
- [24] R. Raskar and J. Davis. 5d time-light transport matrix: What can we reason about scene properties. *Int. Memo07*, 2, 2008.
- [25] R. Raskar, K.-H. Tan, R. Feris, J. Yu, and M. Turk. Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 679–688. ACM, 2004.
- [26] N. Ratner and Y. Y. Schechner. Illumination multiplexing within fundamental limits. In *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*, pages 1–8. IEEE, 2007.
- [27] N. Ratner and Y. Y. Schechner. Illumination multiplexing within fundamental limits. In *Computer Vision and Pattern Recognition, 2007. CVPR’07. IEEE Conference on*, pages 1–8. IEEE, 2007.
- [28] N. Ratner, Y. Y. Schechner, and F. Goldberg. Optimal multiplexed sensing: bounds, conditions and a graph theory link. *Opt. Express*, 15(25):17072–17092, 2007.
- [29] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. A theory of multiplexed illumination. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 808–815. IEEE, 2003.

- [30] Y. Y. Schechner, S. K. Nayar, and P. N. Belhumeur. Multiplexing for optimal lighting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(8):1339–1354, 2007.
- [31] D. Seung and L. Lee. Algorithms for non-negative matrix factorization. *Advances in neural information processing systems*, 13:556–562, 2001.
- [32] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications*, 3:745, 2012.
- [33] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar. Femto-photography: Capturing and visualizing the propagation of light. *ACM Transactions on Graphics (TOG)*, 32(4):44, 2013.
- [34] A. Wenger, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec. Performance relighting and reflectance transformation with time-multiplexed illumination. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 756–764. ACM, 2005.
- [35] R. Z. Whyte, A. D. Payne, A. A. Dorrington, and M. J. Cree. Multiple range imaging camera operation with minimal performance impact. In *IS&T/SPIE Electronic Imaging*, pages 75380I–75380I. International Society for Optics and Photonics, 2010.
- [36] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139–191139, 1980.