

MIT Open Access Articles

Diffuse imaging: Replacing lenses and mirrors with omnitemporal cameras

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Kirmani, Ahmed et al. "Diffuse Imaging: Replacing Lenses and Mirrors with Omnitemporal Cameras." Wavelets and Sparsity XIV Conference, August 21, 2011, San Diego, California, USA (Proc. SPIE 8138), 2011. 813800–813800–6. Web.

As Published: <http://dx.doi.org/10.1117/12.896718>

Publisher: Society of Photo-optical Instrumentation Engineers

Persistent URL: <http://hdl.handle.net/1721.1/71839>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of Use: Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



Diffuse imaging: Replacing lenses and mirrors with omnitemporal cameras

Ahmed Kirmani,^a Haris Jeelani, Vahid Montazerhodjat,^a and Vivek K Goyal^a

^aMassachusetts Institute of Technology, Cambridge, MA 02139, USA

ABSTRACT

Conventional imaging uses steady-state illumination and light sensing with focusing optics; variations of the light field with time are not exploited. We develop a signal processing framework for estimating the reflectance f of a Lambertian planar surface in a known position using omnidirectional, time-varying illumination and unfocused, time-resolved sensing in place of traditional optical elements such as lenses and mirrors. Our model associates time sampling of the intensity of light incident at each sensor with a linear functional of f . The discrete-time samples are processed to obtain ℓ^2 -regularized estimates of f . Using non-impulsive, bandlimited light sources instead of impulsive illumination significantly improves signal-to-noise ratio (SNR) and reconstruction quality.

Keywords: light transport, optical imaging, reflectance, resolution, quantization, sampling, time of flight

1. INTRODUCTION

Imaging is distinguished from other forms of sensing by having the goal of producing an *image*—a representation in one-to-one spatial correspondence with an object or scene. Traditional imaging involves illumination of the scene with a light source whose intensity does not vary with time. The image, or the reflectance pattern on scene surfaces, is acquired by using a lens to focus the reflected light on a two-dimensional (2D) array of light meters. This paper expounds a *diffuse imaging* framework through which a focused image is produced without optical focus; instead, the image is reconstructed from time-resolved measurements of light intensity in response to time-varying scene illumination. Diffuse imaging opens up possibilities for forming images without lenses and mirrors and enables practical implementations of challenging new applications such as imaging occluded scenes.¹

The use of high-speed sensing in photography is associated with effectively stopping motion.^{2,3} We use time-resolved sensing differently: to differentiate among paths of different lengths from light source to scene to sensor. Thus, like in time-of-flight range measurement systems,^{4,5} we are exploiting that the speed of light is finite. However, rather than using the duration of delay to infer distances, we use temporal variation in the *intensity* of measured light to infer scene reflectance. This is achieved through a computational unmixing of the reflectance information that is linearly combined at the sensor because distinct optical paths may have equal path length.

Diffuse imaging involves only the use of diffuse or omnidirectional light sources and omnidirectional light sensors. In conventional photography, the light source is omnidirectional but the sensor array is focused at the scene. A dual configuration is also possible in which the sensing is omnidirectional and the light source is directed, with optical focusing.⁶ Optical focusing enables imaging without the use of time, while in diffuse imaging we create focus computationally by exploiting time variations in the incident light field.

The paper is organized as follows: Sections 2, 3, and 4 expand on earlier work of Kirmani *et al.*,¹ especially by providing a precise signal processing model. Sections 5 and Section 6 show how to improve upon the theoretical and practical limitations of using impulsive illumination and ultrafast sensing as proposed in the earlier work.¹ In particular, simulations suggest that diffuse imaging can be applied to build practical imaging systems using non-ultrafast, opto-electronic hardware used in standard optical communication systems.

2. SCENE RESPONSE TO IMPULSIVE ILLUMINATION

Consider the imaging scenario depicted in Fig. 1, with a planar surface to be imaged, a single, time-varying, monochromatic, omnidirectional illumination source, and omnidirectional time-resolved sensors indexed by $k \in \{1, 2, \dots, K\}$. We assume that the position, orientation, and dimensions (L -by- L) of the planar surface are known. Estimation of these geometric parameters using diffuse illumination and time-resolved sensing was

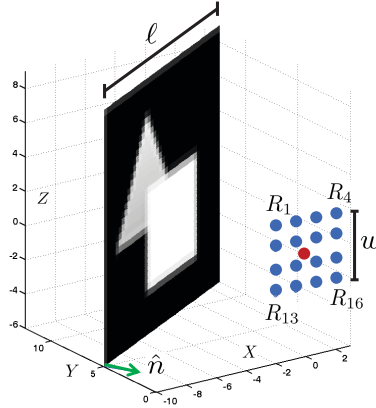


Figure 1. An omnidirectional source (red point) illuminates a planar scene, and the reflected light is measured at a 4-by-4 array of sensors (blue points).

recently demonstrated.⁷ Formation of an ideal gray scale image is the recovery of the reflectance pattern on the surface, which can be modeled as a 2D function $f : [0, L]^2 \rightarrow [0, 1]$. We assume the surface to be *Lambertian* so that its perceived brightness is invariant to the angle of observation;⁸ incorporation of any known bidirectional reflectance distribution function would not add insight.

The light incident at Sensor k is a combination of the time-delayed reflections from all points on the planar surface. For any one point $\mathbf{x} = (x_1, x_2) \in [0, L]^2$, the contribution is delayed based on the total distance $d_k(\mathbf{x})$ traveled by the light, and it is attenuated by both the reflectance $f(\mathbf{x})$ and a geometric factor related to radial fall-off and foreshortening.⁸ When the intensity of the omnidirectional illumination is a unit impulse at time 0, denoted $s(t) = \delta(t)$, the contribution from point \mathbf{x} is the light signal

$$a_k(\mathbf{x}) f(\mathbf{x}) \delta(t - d_k(\mathbf{x}))$$

for a known attenuation function $a_k(\mathbf{x})$, where we have normalized to unit speed of light.

Combining contributions over the plane, the total light incident at Sensor k is

$$g_k(t) = \int_0^L \int_0^L a_k(\mathbf{x}(s)) f(\mathbf{x}) \delta(t - d_k(\mathbf{x})) dx_1 dx_2. \quad (1)$$

Thus, evaluating $g_k(t)$ at a fixed t amounts to integrating over $\mathbf{x} \in \mathbb{R}^2$ with $d_k(\mathbf{x}) = t$. Define the isochronal curve

$$C_k^t = \{\mathbf{x} : d_k(\mathbf{x}) = t\}.$$

Then

$$g_k(t) = \int a_k(\mathbf{x}(s)) f(\mathbf{x}(k, s)) ds \quad (2)$$

where $\mathbf{x}(k, s)$ is a parameterization of C_k^t with unit speed. The intensity $g_k(t)$ thus contains the contour integrals over C_k^t 's of the desired function f . Each C_k^t is a level curve of $d_k(\mathbf{x})$; these are ellipses.

3. SAMPLING THE SCENE RESPONSE

A digital system can use only samples of $g_k(t)$ rather than the continuous-time function itself. We now see how uniform sampling of $g_k(t)$ with a linear time-invariant (LTI) prefilter relates to linear functional measurements of f . This establishes the foundations of a Hilbert space view of diffuse imaging.

Suppose discrete samples are obtained at Sensor k with sampling prefilter $h_k(t)$ and sampling interval T_k :

$$y_k[n] = (g_k(t) * h_k(t))|_{t=nT_k}, \quad n = 1, 2, \dots, N. \quad (3)$$

A sample $y_k[n]$ can be seen as a standard $\mathcal{L}^2(\mathbb{R})$ inner product between g_k and a time-reversed and shifted h_k :⁹

$$y_k[n] = \langle g_k(t), h_k(nT_k - t) \rangle. \quad (4)$$

Using (1), we can express (4) in terms of f using the standard $\mathcal{L}^2([0, L]^2)$ inner product:

$$y_k[n] = \langle f, \varphi_{k,n} \rangle \quad \text{where} \quad (5a)$$

$$\varphi_{k,n}(\mathbf{x}) = a_k(\mathbf{x}) h_k(nT_k - d_k(\mathbf{x})). \quad (5b)$$

Over a set of sensors and sample times, $\{\varphi_{k,n}\}$ will span a subspace of $\mathcal{L}^2([0, L]^2)$, and a sensible goal is to form a good approximation of f in that subspace.

4. IMAGE RECOVERY USING LINEAR BACKPROJECTION

To express an estimate \hat{f} of the reflectance f , it is convenient to fix an orthonormal basis for a subspace of $\mathcal{L}^2([0, L]^2)$ and estimate the expansion coefficients in that basis. For an M -by- M pixel representation, let

$$\psi_{i,j}(\mathbf{x}) = \begin{cases} M/L, & \text{for } (i-1)L/M \leq x_1 < iL/M, \\ & (j-1)L/M \leq x_2 < jL/M; \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

so that

$$\hat{f} = \sum_{i=1}^M \sum_{j=1}^M c_{i,j} \psi_{i,j}$$

in the span of $\{\psi_{i,j}\}$ is constant on Δ -by- Δ patches, where $\Delta = L/M$. We will form a system of linear equations to find coefficients $\{c_{i,j}\}$.

For \hat{f} to be consistent with the value measured by Sensor k at time n , we must have

$$y_k[n] = \langle \hat{f}, \varphi_{k,n} \rangle = \sum_{i=1}^M \sum_{j=1}^M c_{i,j} \langle \psi_{i,j}, \varphi_{k,n} \rangle. \quad (7)$$

Note that the inner products $\{\langle \psi_{i,j}, \varphi_{k,n} \rangle\}$ exclusively depend on Δ , the positions of illumination and sensors, the plane geometry, the sampling prefilters $\{h_k\}_{k=1}^K$, and the sampling intervals $\{T_k\}_{k=1}^K$ —not on the unknown reflectance of interest f . Hence, we have a system of linear equations to solve for the coefficients $\{c_{i,j}\}$. (In the case of basis (6), the coefficients are directly the pixel values.)

When we specialize to a box sensor impulse response and basis (6), many inner products $\langle \psi_{i,j}, \varphi_{k,n} \rangle$ are zero so the linear system is sparse. The inner product $\langle \psi_{i,j}, \varphi_{k,n} \rangle$ is nonzero when reflection from the (i, j) pixel affects the light intensity at Sensor k within time interval $[(n-1)T, nT]$. Thus, for a nonzero inner product the (i, j) pixel must intersect the elliptical annulus between $C_k^{(n-1)T}$ and C_k^{nT} . This occurs for a small number of (i, j) pairs unless M is small or T is large. The value of a nonzero inner product depends on the fraction of the square pixel that overlaps with the elliptical annulus and the attenuation factor $a_k(\mathbf{x})$.

To express (7) with a matrix multiplication, replace double indexes with single indexes (i.e., vectorize, or reshape) as

$$\mathbf{y} = A \mathbf{c} \quad (8)$$

where $\mathbf{y} \in \mathbb{R}^{KN}$ contains the data samples $\{y_k[n]\}$, the first N from Sensor 1, the next N from Sensor 2, etc.; and $\mathbf{c} \in \mathbb{R}^{M^2}$ contains the coefficients $\{c_{i,j}\}$, varying i first and then j . Then $\langle \psi_{i,j}, \varphi_{k,n} \rangle$ appears in row $(k-1)N + n$, column $(i-1)M + j$ of $A \in \mathbb{R}^{KN \times M^2}$.

Assuming that A has a left inverse (i.e., $\text{rank}(A) = M^2$), one can form an image by solving (8). The portion of A from one sensor cannot have full column rank because of the collapse of information along elliptical annuli. Full rank can be achieved with an adequate number of sensors, noting that sensor positions must differ to increase rank, and—as demonstrated in Section 6—greater distance between sensor positions improves conditioning.

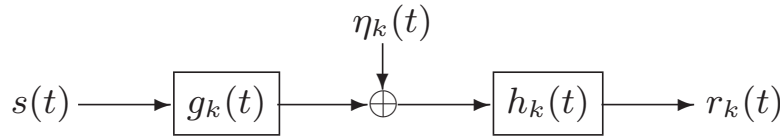


Figure 2. Block diagram abstraction for signal sampled at Sensor k .

5. ILLUMINATION DESIGN

A Dirac impulse illumination is an abstraction that cannot be realized in practice. One can use expensive, ultrafast optical lasers that achieve Terahertz bandwidth as an approximation to impulsive illumination, as in experimental verification of our framework.¹ We show that non-impulsive illuminations are not only more practical for implementation, but they improve upon impulsive illumination for typical scenes and sensors.

Light transport is linear and time invariant. Hence, one can consider the effect of a general illumination $s(t)$ as the superposition of effects of constant illuminations over infinitesimal intervals. The light incident at Sensor k is changed from (1) by a convolution with $s(t)$. Thus, the block diagram in Fig. 2 represents the signal at Sensor k , including its sampling prefilter and photodetector noise represented by $\eta_k(t)$. Except at very low flux, $\eta_k(t)$ is modeled well as single-independent, zero-mean, white and Gaussian. The noise variance σ^2 depends on the device physics and assembly; our simulations use $\sigma = 0.1$.

A typical natural scene f has a good bandlimited approximation. Integration over elliptical contours C_k^t further smooths the signal. Thus, $s(t)$ is best chosen to be lowpass to put signal energy at frequencies most present in $g_k(t)$.

6. SIMULATIONS

In earlier experimental work,¹ high-bandwidth illumination and sampling were used under the assumption that these would lead to higher reconstruction resolution. However, impulsive illumination severely limits the illumination energy, leading to poor SNR, especially due to the radial fall-off attenuations.

Here we compare impulsive and lowpass illumination. All image reconstructions are obtained with a modification of (8) using regularization by the ℓ^2 -norm of the Laplacian, a conventional technique for backprojection.¹⁰ The regularization is optimized for ℓ^2 error. Results are for $M = 40$ and several values of sampling period T and sensor array extent W .

In direct analogy with the earlier experimental work,¹ we simulated short-pulsed, high-bandwidth illumination using a square wave source with unit amplitude and time width equal to one-fifth of T . Fig. 3 shows the results. Reconstruction with good spatial resolution is indeed possible, and the conditioning improves as W increases and as T decreases.

Using non-impulsive, low-bandwidth illumination, we show that high SNR can be achieved while improving the reconstruction resolution. We chose $s(t)$ to be the truncated impulse response of a third-order Butterworth filter, with again a unit amplitude. This choice of $s(t)$ produces a much stronger scene reflection and hence improves the SNR at the detector without any loss of matrix conditioning. Fig. 4 shows the resulting improvements in reconstructed images. Hence, the choice of a non-impulsive illumination is not only practical but demonstrably better in terms of image reconstruction quality.

7. SUMMARY

Earlier proof-of-concept experiments show that diffuse imaging can succeed in forming image reconstructions.¹ In this paper we have used signal processing abstractions to show that using lowpass time-varying illumination instead of impulsive sources improves the SNR and reconstruction quality. Our work inspires detailed study of how pixel size, sensor locations, sampling prefilters, and sampling rates affect the conditioning of the inverse problem.

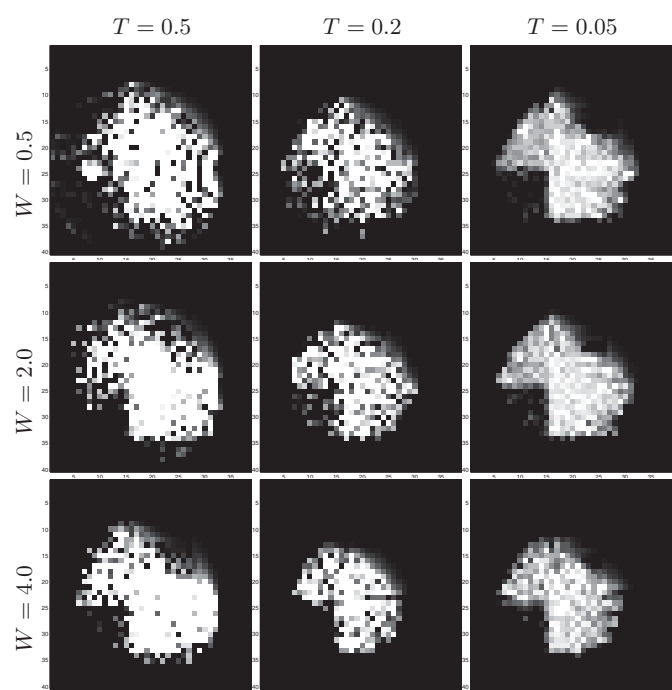


Figure 3. Reconstructions with impulsive illumination and 4-by-4 array of sensors as in Fig. 1.

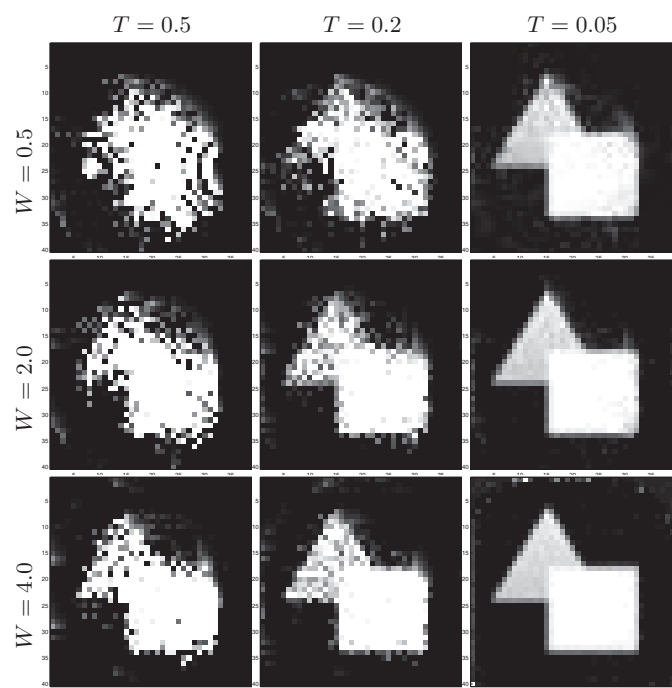


Figure 4. Reconstructions with lowpass illumination and 4-by-4 array of sensors as in Fig. 1.

ACKNOWLEDGMENTS

This material is based upon work supported in part by the National Science Foundation under Grant No. 0643836, the DARPA InPho program through the US Army Research Office award W911-NF-10-1-0404, the Texas Instruments Leadership University Program, and an Esther and Harold E. Edgerton Career Development Chair.

REFERENCES

1. A. Kirmani, A. Velten, T. Hutchison, M. E. Lawson, V. K. Goyal, M. Bawendi, and R. Raskar, "Reconstructing an image on a hidden plane using ultrafast imaging of diffuse reflections." Submitted, May 2011.
2. H. E. Edgerton, K. J. Germeshausen, and H. E. Grier, "High-speed photographic methods of measurements," *J. Appl. Physics* **8**, pp. 1–2, Jan. 1937.
3. H. E. Edgerton and J. R. Killian, Jr., *Flash! Seeing the Unseen by Ultra High-Speed Photography*, Hale, Cushman and Flint, Boston, MA, 1939.
4. B. Schwarz, "LIDAR: Mapping the world in 3D," *Nature Photonics* **4**, pp. 429–430, July 2010.
5. S. Foix, G. Alenyà, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sensors J.* **11**, pp. 1917–1926, Sept. 2011.
6. P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, "Dual photography," *ACM Trans. Graphics* **24**, pp. 745–755, July 2005.
7. A. Kirmani, T. Hutchison, J. Davis, and R. Raskar, "Looking around the corner using transient imaging," in *Proc. IEEE 12th Int. Conf. on Computer Vision*, (Kyoto, Japan), Sept.–Oct. 2009.
8. M. Oren and S. K. Nayar, "Generalization of the Lambertian model and implications for machine vision," *Int. J. Comput. Vis.* **14**, pp. 227–251, Apr. 1995.
9. M. Unser, "Sampling—50 years after Shannon," *Proc. IEEE* **88**, pp. 569–587, Apr. 2000.
10. M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy and undersampled measured images," *IEEE Trans. Image Process.* **6**, pp. 1646–1658, Dec. 1997.