Compressive Structured Light for Recovering Inhomogeneous Participating Media

Jinwei Gu, Shree Nayar, Eitan Grinspun,
Peter Belhumeur, and Ravi Ramamoorthi

Columbia University, New York, NY 10027, USA
jwgu@cs.columbia.edu

Abstract. We propose a new method named compressive structured light for recovering inhomogeneous participating media. Whereas conventional structured light methods emit coded light patterns onto the surface of an opaque object to establish correspondence for triangulation, compressive structured light projects patterns into a *volume* of participating medium to produce images which are *integral measurements* of the volume density along the line of sight. For a typical participating medium encountered in the real world, the integral nature of the acquired images enables the use of *compressive sensing* techniques that can recover the entire volume density from only a few measurements. This makes the acquisition process more efficient and enables reconstruction of dynamic volumetric phenomena. Moreover, our method requires the projection of multiplexed coded illumination, which has the added advantage of increasing the signal-to-noise ratio of the acquisition. Finally, we propose an iterative algorithm to correct for the attenuation of the participating medium during the reconstruction process. We show the effectiveness of our method with simulations as well as experiments on the volumetric recovery of multiple translucent layers, 3D point clouds etched in glass, and the dynamic process of milk drops dissolving in water.

1 Introduction

Structured light has a long history in the computer vision community [1]. It has matured into a robust and efficient method for recovering the surfaces of objects. By projecting coded light patterns on the scene, and observing it using a camera, correspondences are established and the 3D structure of the scene is recovered by triangulation. Over the years, researchers have developed various types of coding strategies, such as binary codes, phase shifting, spatial neighborhood coding, etc. All structured light range finding approaches are based on a common assumption: Each point in the camera image receives light reflected from a single *surface* point in the scene.

However, many real-world phenomena can only be described by volume densities rather than boundary surfaces. Such phenomena are often referred to as *participating media*. Examples include translucent objects, smoke, clouds, mixing fluids, and biological tissues. Consider an image acquired by photographing

a volume of a participating medium. Unlike in the case of an opaque object, here each pixel receives scattered light from *all points along the line of sight* within the volume. Narasimhan et al. [2] have addressed the problem of recovering opaque surfaces immersed in a participating medium (rather than a clear medium) using structured light range finding. The goal in this case is to make existing structured light range finding methods robust to light scattering by the medium, rather than recover the medium itself.

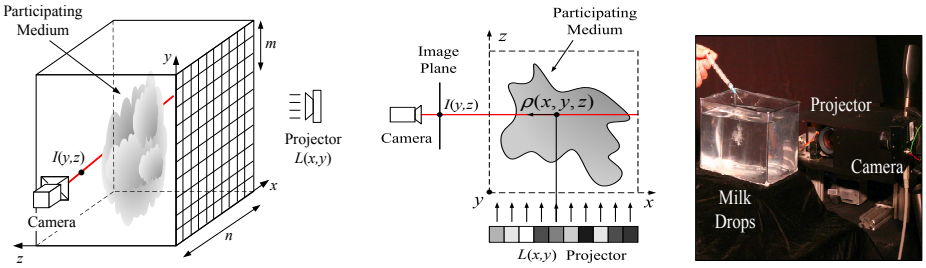
The problem of recovering the volume density of a participating medium (in particular, smoke), was addressed by Hawkins et al. [3]. They used a high-powered laser sheet and a high-speed camera (5000fps) to measure thin slices of a smoke density field via *scanning*, which is similar to a technique termed laser-induced fluorescence (LIF) in the fluid imaging community [4]. Fuchs et al. [5] proposed the idea of shooting a set of static laser rays into the volume and using spatial *interpolation* to reconstruct the volume. However, the measurements are inherently sparse in this case and hence the recovered density is low in resolution.

In this paper, we show that by using coded light patterns, one can make the measurement of a participating medium highly efficient in terms of acquisition time as well as illumination power. In particular, we exploit the fact that the brightness measurements made at image pixels correspond to true line-integrals through the medium (see Fig. 1a), and then solve for its volumetric density. We consider both spatially- and temporally-coded light patterns. Because the patterns are predetermined, measurement and reconstruction time are decoupled. We target low-density inhomogeneous media, for which the density function is *sparse* in an appropriately-chosen basis[1]; this allows us to harness compressive sensing techniques [6, 7] that accurately reconstruct a signal from only a few measurements. We refer to our approach as *compressive structured light*.

We show that compressive structured light is more economical than a straightforward sequential scanning of a volume. Whereas the sampling rate of the latter is limited by the desired resolution, the sampling rate of the former is restricted by the sparsity of the data–a considerably more relaxed constraint for low-density phenomena. Since our approach requires fewer measurements, it naturally enables the recovery of dynamic participating media. An added advantage of compressive structured light, is that it requires the projection of multiplexed coded illumination which results in measurements with higher signal-to-noise ratio [8]. An important practical consequence is that light sources of significantly lower power than in the case of sequential scanning can be used.

We have implemented our approach using a digital projector and a camera as shown in Fig. 1c. The projector and the camera are synchronized and both operate at 360fps. Using 24 coded light patterns, we are able to recover a 128^3 volume at 15fps. Using this system, we have recovered various types of inhomogeneous participating media, as shown in §7.

[1] "sparse" does *not* necessarily imply that the volume density must be sparsely distributed in space. It means that the density can be represented with a few non-zero coefficients in an appropriately-chosen basis, such as, wavelets, gradients, principal components, etc.



(a) (b) (c)

Fig. 1. (a) Compressive structured light for recovering inhomogeneous participating media. Coded light is emitted along the z-axis to the volume while the camera acquires images as line-integrated measurements of the volume density along the x-axis. The light is coded in either the spatial domain or temporal domain with a predetermined sequence. We reconstruct the volume density from the measurements by using *compressive sensing* techniques. (b) Image formation model for participating medium under single scattering. The image irradiance at one pixel, $I(y, z)$, depends on the integral along the x-axis of the projector's light, $L(x, y)$, and the medium density, $\rho(x, y, z)$, along a ray through the camera center; refer to (1)–(2). (c) Experiment setup.

2 Related Work

Compressive Sensing Compressive sensing [6, 7] is a nascent field of applied mathematics with a variety of successful applications including imaging [9], medical visualization [10], and face recognition [11]. It offers a theoretical framework to reconstruct "sparse" signals from far fewer samples than required by the conventional Shannon sampling theorem. Our work builds on the basic formulation of compressive sensing, which we augment with auxiliary terms specific to the reconstruction of volume density.

Reconstruction of Volumetric Phenomena There are several recent works in reconstruction of volumetric phenomena from multiple views. Hasinoff et al. [12] used two views to reconstruct flames by assuming flames are surfaces in the volume. Based on tomographic algorithms, Ihrke et al. [13, 14] and Trifonov et al. [15] used eight views and 72–360 views, respectively, for recovering flames (and smoke) as well as transparent objects. We mentioned light-based methods [3–5] earlier in the previous section. For a comprehensive survey of works in this area, see Ihrke et al. [16].

Multiplexed Illumination Our work is also related to multiplexed illumination [8] in that both use coded light as illumination. However, there is a fundamental difference: Whereas the conventional multiplexing aims at increasing signal-to-noise ratio of the measurements, our work aims at increasing the efficiency of the acquisition process, i.e., to reconstruct high dimensional signals from a few measurements. In summary, both the coding strategies and the reconstruction algorithms are different.

3 Background on Compressive Sensing

In its simplest form, compressive sensing seeks a solution of the underdetermined linear system $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{x} \in \mathbb{R}^n$ is a sparse signal, $\mathbf{A}$ is an $m \times n$ matrix ("measurement ensemble"), and $\mathbf{b}$ is the vector of m measurements, with $m < n$.

Compressive sensing theory asserts that one can recover the signal from far fewer measurements than the dimension of the signal, if the signal is *sparse*—it is represented with few non-zero coefficients in a suitable basis—and the measurements are uncorrelated, in the sense that each measurement is an inner product of the signal with a *test function* that has a necessarily *dense* representation in the chosen basis. Given a measurement ensemble matrix $\mathbf{A}$, compressive sampling theory predicts that $\mathbf{x}$ is the minimizer of $||\mathbf{x}||_1$, subject to $\mathbf{Ax} = \mathbf{b}$.

The above reconstruction strategy has been shown to work well for sparse signal estimation, even from a noisy measurement [17]. In our work, we augment the basic problem above with auxiliary terms that enforce the nonnegative constraint for the signal, and that exploit the sparsity not only of the signal value but also its gradient. Indeed, our evaluation (§6, §7) indicates that using the sparsity of the gradient is important for accurate and efficient reconstruction.

4 Image Formation Model

In this section, we derive the relationship between the volume density ρ and the image irradiance I of the camera under our camera/projector setting. We focus on non-emissive participating media with relatively low density in which multiple scattering is assumed to be negligible.

As shown in Fig. 1b, each camera pixel receives light scattered from a row of voxels along the line of sight in the volume (i.e., the red line in Fig. 1b). Consider one such voxel. Before the light scattered by this voxel reaches the camera, it is first *attenuated* as it travels from the projector to the voxel, *scattered* at the voxel, and then *attenuated* as it travels from the voxel to the camera. Under the assumption of single scattering, the final radiance sensed by the camera from this particular voxel is [18]: $L(x,y) \cdot \exp(-\tau_1) \cdot \sigma_s \cdot \rho(x,y,z) \cdot p(\theta) \cdot \exp(-\tau_2)$, where $\rho(x,y,z)$ is the participating medium's volume density at the voxel, $p(\theta)$ is the phase function, and $\tau_1 = \sigma_t \int_{s_1} \rho ds_1$ and $\tau_2 = \sigma_t \int_{s_2} \rho ds_2$ are the "optical distance" from the projector to the voxel and from the voxel to the camera, respectively; σ_s and σ_t are the scattering coefficient and the attenuation coefficient of the participating medium [18]. Since $p(\theta)$ is the same for all voxels under orthographic projection for both the camera and the projector, the above formula can be simplified to (up to a scale related to $p(\theta)$ and σ_s): $L(x,y) \cdot \exp\left(-(\tau_1 + \tau_2)\right) \cdot \rho(x,y,z)$. The image irradiance, $I(y,z)$, which is the integral of the scattered light from all the voxels along the line, is therefore

$$I(y,z) = \int_x L(x,y) \cdot \exp\left(-(\tau_1 + \tau_2)\right) \cdot \rho(x,y,z) dx \ . \tag{1}$$

In the low density case, or when σ_t is relatively small compared with the scattering, the effect of attenuation usually can be ignored [3, 5], i.e., the exponential

term in the above equation is equal to 1. Equation (1) thus can be reduced to a linear projection of the light and the volume density,

$$I(y,z) = \int_x \rho(x,y,z) \cdot L(x,y)\, dx \ . \tag{2}$$

For media where the attenuation cannot be ignored, we present a simple, iterative method based on iterative relinearization (see §5.3).

5 Compressive Structured Light

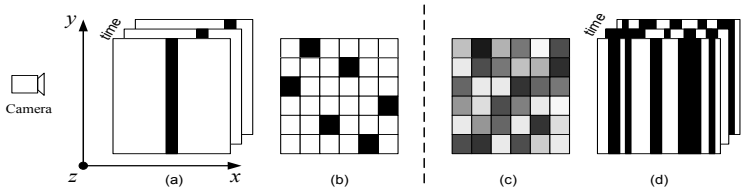


Fig. 2. Different coding strategies of the light $L(x,y)$ at time t for recovering inhomogeneous participating media: (a) scan (one stripe turned on) [4, 3]; (b) laser-lines interpolation (one pixel turned on per one row) [5]; (c) Spatial coding of compressive structured light (all pixels are turned on with random values per time frame); (d) Temporal coding of compressive structured light (random binary stripes are turned on per time frame). Compressive structured light, shown in (c) and (d), recovers the volume by reconstructing the 1D signal along x-axis from a few integral measurements.

In this section, we explain the idea of compressive structured light for recovering inhomogeneous participating media. For participating media, each camera pixel receives light from all points along the line of sight within the volume. Thus each camera pixel is an integral measurement of one row of the volume density. Whereas conventional structured light range finding methods seek to triangulate the position of a single point, compressed structured light seeks to reconstruct the 1D density "signal" from a few measured integrals of this signal.

This is clearly a more difficult problem. One way to avoid this problem is to break the integrals into pieces which can be measured directly. The price, however, is the deterioration of either spatial resolution or temporal resolution of the acquisition. Existing methods either illuminate a single slice at a time and scan the volume (see Fig. 2a and [4, 3]), thus sacrificing temporal resolution, or they illuminate a single pixel per row and use interpolation to reconstruct the volume (e.g., Fig. 2b and [5]), sacrificing spatial resolution.

In contrast, the proposed compressive structured light method uses the light much more efficiently, projecting coded light patterns that yield "signatures," or integral measurements, of the unknown volume density function.

The didactic illustration in Fig. 1a depicts a simple lighting/viewpoint geometry under orthographic projection, with the camera viewpoint along the x-axis, and the projector emitting along the z-axis. Consider various coding strategies of the 3D light function $L(x, y, t)$: *Spatial* codes (Fig. 2c) recover the volume from a single image by trading spatial resolution along one dimension; *Temporal* codes (Fig. 2d) trade temporal resolution by emitting a sequence of vertical binary stripes (with no coding along y-axis), so that full spatial resolution is retained.[2]

We will see that these compressive structured light codes yield high efficiency both in acquisition time and illumination power; this comes at the cost of a more sophisticated reconstruction process, to which we now turn our attention.

5.1 Formulation

To better visualize our formulation, consider first the case of spatial coding. Suppose we want to reconstruct a volume at the resolution $n \times n \times n$ (e.g., $n = 100$). The camera and the projector have the resolution of $M \times M$ pixels (e.g., $M = 1024$). Therefore, one row of voxels along the x-axis (refer to the red line in Fig. 1a) will receive light from $m = M/n$ (e.g., $m = 1024/100 \approx 10$) rows of the projector's pixels. The light scattered by these voxels in the viewing direction will then be measured, at each z-coordinate, by a vertical column of m camera pixels. Without loss of generality, we use $\mathbf{l}_1 = L(x, 1), \cdots, \mathbf{l}_m = L(x, m)$ to denote the m rows of pixels from the projector, and $b_1 = I(1, z), \cdots, b_m = I(m, z)$ to denote the image irradiance of the m pixels in the camera image. Let $\mathbf{x} = [\rho_1, \cdots, \rho_n]^T$ be the vector of the voxel densities along the row. Assuming no attenuation, the image irradiance for each of these m pixels is a linear projection of the light and the voxels' density from (2): $b_i = \mathbf{l}_i^T \mathbf{x}$, $i = 1, \cdots, m$. Rewriting these m equations in matrix form, we have: $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{A} = [\mathbf{l}_1, \cdots, \mathbf{l}_m]^T$ is a $m \times n$ matrix, $\mathbf{b} = [b_1, \cdots, b_m]^T$ is a $m \times 1$ vector.

Thus, if attenuation is not considered, the problem of recovering the volume is formulated as the problem of reconstructing the 1D signal $\mathbf{x}$ given the constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$. To retain high spatial and temporal resolution, we often can only afford far fewer measurements than the number of unknowns, i.e., $m < n$, which means the above equation is an underdetermined linear system and optimization is required to solve for the best $\mathbf{x}$ according to certain priors.

One benefit of this optimization-based reconstruction is high efficiency in acquisition, which we quantify using the *measurement cost*, m/n, where m is the number of the measurements and n is the number of unknowns (i.e., the dimension of the signal). For example, the measurement cost of the scanning method [4, 3] is one. We show that by exploiting the sparsity of the signal, we can reconstruct the volume with much lower measurement cost (about $\frac{1}{8}$ to $\frac{1}{4}$).

5.2 Reconstruction via Optimization

Formulation Solving the underdetermined linear system requires some prior (assumed) knowledge of the unknown signal, which can be represented as optimiza-

[2] All of the 4 methods shown in Fig. 2 can be equally improved using color channels.

Table 1. Different norms used for reconstruction

Method	Optimization Functional	Constraints
Least Square (LS)	$\|\|\mathbf{Ax}-\mathbf{b}\|\|_2$	
Nonnegative Least Square (NLS)	$\|\|\mathbf{Ax}-\mathbf{b}\|\|_2$	$\mathbf{x} \geq 0$
CS-Value	$\|\|\mathbf{x}\|\|_1$	$\mathbf{Ax}=\mathbf{b},\ \mathbf{x} \geq 0$
CS-Gradient	$\|\|\mathbf{x}'\|\|_1$	$\mathbf{Ax}=\mathbf{b},\ \mathbf{x} \geq 0$
CS-Both	$\|\|\mathbf{x}\|\|_1 + \|\|\mathbf{x}'\|\|_1$	$\mathbf{Ax}=\mathbf{b},\ \mathbf{x} \geq 0$

tion functionals or constraints on the data. We consider several alternatives, as listed in Table 1. Besides the commonly-used Least Square (LS) and Nonnegative Least Square (NLS), we consider functionals using ℓ_1-norms, as these bias toward sparse representations:[3]

First, we observe that for many natural volumetric phenomena, often only a small portion of the entire volume is occupied by the participating media. For example, consider the beautiful ribbon patterns generated by smoke; similarly, sparsity was implicitly used to reconstruct (surface-like) flames [12]). This suggests the use of the ℓ_1-norm of the signal value (CS-Value).

Furthermore, the sparsity of *gradients* of natural images is well studied [20, 21]. Related work in image restoration [22] uses nonlinear optimization to minimize "total variation," i.e., the sum of ℓ_2-norm of image gradient. In this vein, we consider the use of ℓ_1-norm on the signal's gradient (CS-Gradient).

Finally, consider a dynamic process, such as milk dissolving in water: here diffusion decreases the signal value's sparsity over time, but it increases the gradient sparsity. Motivated by this observation, we consider the sum of ℓ_1-norms of both the value and the gradient (CS-Both), so that the algorithm has the ability to "adapt" for the sparsity.

Analysis Comparison of these reconstruction methods is first performed on 1D synthetic signals. These signals are randomly sampled rows from the volume density of smoke acquired in Hawkins et al. [3]. We restrict the measurement cost, m/n, to be $1/4$. The measurement ensemble, $\mathbf{A}$, is generated in a way that each element is drawn independently from a normal distribution and each column is normalized to 1, which is effectively a white noise matrix and is known to be good for compressive sensing [7]. NRMSE (normalized root mean squared error) is used as the measure of error.

The reconstruction results are shown in Fig. 3. The commonly-used LS performs the worst, since it merely minimizes the errors without using any prior on the data. With the nonnegative constraint added, NLS has better performance. CS-Value and CS-Gradient are better than NLS given that both use one more prior—the sparsity on the signal value or on the signal gradient. The fact that CS-Gradient is better than CS-Value indicates that the sparsity on the signal gradient holds better than the sparsity on the signal value. Finally, as expected,

[3] LS and NLS are solved with SVD and Levenberg-Marquardt, respectively. The other functionals are formulated as Linear Programming (LP) and solved with GLPK [19].

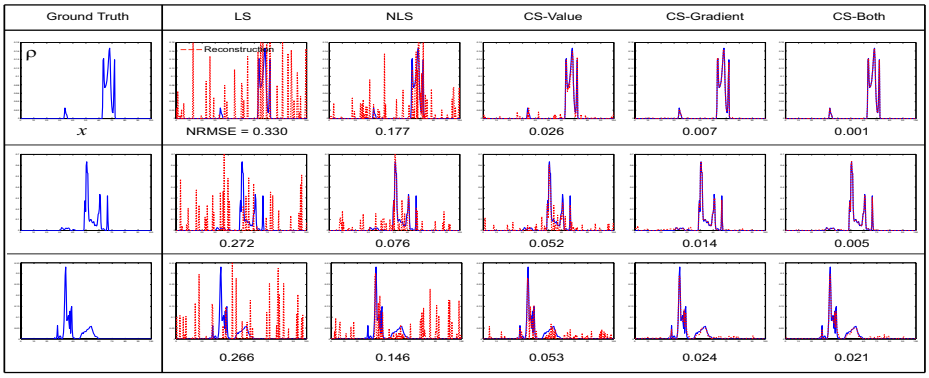


Fig. 3. Comparison of different reconstruction methods. The first column is the original signal. The remaining columns show reconstruction results (red dashed lines) for different methods, given the measurement cost, m/n, is equal to 1/4. The value below each plot is the NRMSE(normalized root mean squared error) of reconstruction.

CS-Both outperforms other methods due to its adaptive ability. In our trials, the favorable performance of CS-Both was not sensitive to changes of the relative weighting of the value and gradient terms. These observations carry over to the 3D setting (see Fig. 4), where we reconstruct a 128^3 volume; note that this requires 128×128 independent 1D reconstructions.

5.3 Iterative Attenuation Correction

Until now, we have not considered the attenuation in the image formation model in (1) yet. To take into account attenuation, we use a simple iterative relinearization algorithm as follows:

1. Assume no attenuation, solve the optimization problem with techniques from §5.2 to get the initial reconstruction of the volume density $\rho^{(0)}$.
2. At iteration k, assuming σ_t is known[4], compute the attenuated light as: $L^{(k)}(x, y, z) = \exp\left(-\left(\tau_1 + \tau_2\right)\right) \cdot L(x, y)$, where τ_1 and τ_2 are computed using $\rho^{(k-1)}$ as shown in §4.
3. With the attenuated light $L^{(k)}(x, y, z)$, (1) becomes a linear equation. We solve for $\rho^{(k)}$ and go to next iteration until it converges.[5]

Since our overall framework accommodates the scanning method [4, 3] and the interpolation method [5] as special cases, the iterative algorithm could be directly applied to these prior methods as well.

[4] The attenuation coefficient, σ_t, of the participating medium can be obtained from literature, specified by a user, or be measured by a second camera taking the shadowgram of the volume.

[5] In practice, we found that the algorithm usually converges within 3-4 iterations.

6 Validation via Simulation

To further validate our method, we perform simulations on a synthetic volume. The volume is generated from a triangular mesh of a horse and it is discretized into 128^3 voxels. For each voxel, if it is inside the mesh, the density is designed to be proportional to the distance from the center of the voxel to the center of the mesh, otherwise the density is 0. Fig. 4a shows the volume where blue corresponds to the lowest density while yellow corresponds to the highest density. A slice of the volume is shown in Fig. 4b.

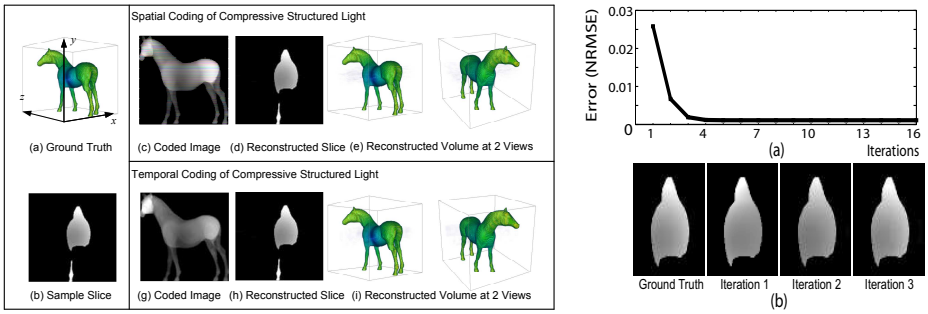


Fig. 4. Simulation results of volume reconstruction using compressive structured light. **LEFT**: (a) The original volume where blue means the lowest density and yellow means the highest density. (b) A slice of the volume. The top and the bottom row on the right shows the reconstruction results for spatial coding and temporal coding, respectively. For each row, from left to right are the coded image acquired by the camera, the reconstruction of the slice, and the reconstructed volume under two different views. **RIGHT**: (a) Reconstruction errors and (b) slices with iterative attenuation correction.

Both spatial coding and temporal coding of compressive structured light are tested. The measurement cost, m/n, is fixed to 1/4. For spatial coding, we use a random color image with resolution of 1280×1280 as the coded light from the projector. This gives us $m = 1280/128 \times 3 = 30$ measurements to recover densities of 128 voxels on one row of the volume. Based on (1), a single image (shown in Fig. 4c) is generated from the camera view and used for reconstruction. For temporal coding, we use random binary stripes as illumination and generate 32 images for reconstruction. One of these images is shown in Fig. 4g. CS-Both is used to reconstruct the volume for both cases. As shown in Fig. 4, both methods accurately reconstruct the volume. Moreover, Fig. 4(right) shows the reconstruction errors and reconstructed slices at different iterations of attenuation correction, which demonstrates the effectiveness of the iterative algorithm.

We also evaluate different reconstruction methods at various measurement costs from 1/16 to 1. The results are shown as a table in Fig. 5. Conclusions similar to the ones from the previous 1D signal simulation (Fig. 3) can be drawn from these results: (1) As expected, all methods have improvements as the mea-

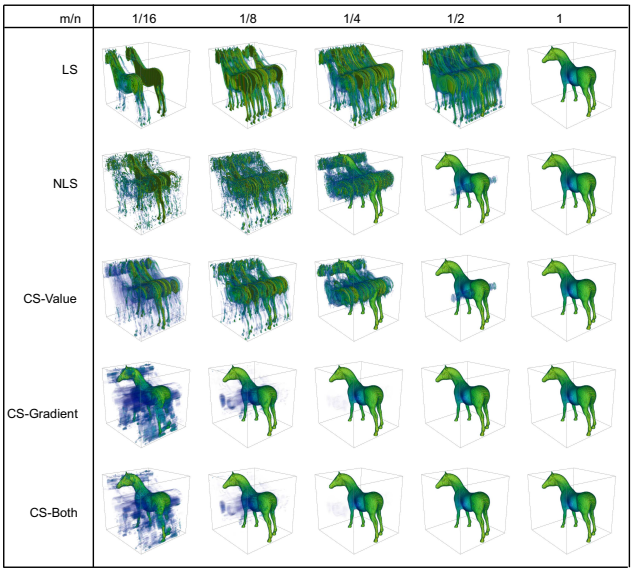


Fig. 5. Comparison of different reconstruction methods at different measurement costs, m/n. CS-Both outperforms other methods.

surement cost increases. (2) Without using any prior of the data, LS is the worst for reconstruction with insufficient measurements. (3) CS-Gradient and CS-Both largely outperform other methods, especially for low measurement cost, which indicating strong sparsity in the signal's gradient. (4) CS-Both is better than CS-Gradient, especially at low measurement cost (e.g., as shown in Fig. 5 at $m/n = 1/16$). Based on these preliminary simulations, we chose to run our actual acquisition experiments with a measurement cost of 1/4 and the CS-Both optimization functional.

7 Experimental Results

We have implemented the temporal coding of compressive structured light for recovering inhomogeneous participating media. As shown in Fig. 1c, our system consists of a 1024×768 DLP projector and a 640×480 Dragonfly Express 8-bit camera, positioned at right angles, both viewing the inhomogeneous participating medium (milk drops in water). The projector and the camera are synchronized and both operate at 360fps.[6] Using 24 coded light patterns, we are able to recover a 128^3 volume at 15fps. These light patterns consist of 128 vertical stripes. Each stripe is assigned 0 or 1 randomly with the probability of 0.5. In this way, about half amount of the light is turned on for each measurement. We also tried alternative light patterns such as Hadamard codes , and found the random binary codes have better performance.

[6] The camera's resolution is set to 320×140 in order to achieve 360fps.

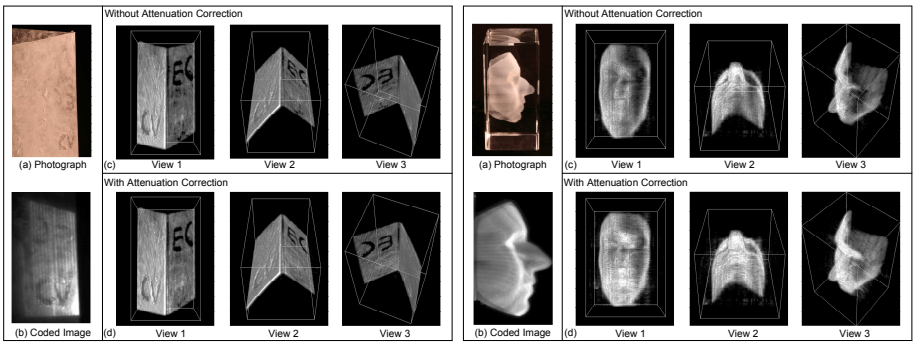


Fig. 6. Reconstruction results of **LEFT**: an object consisting of two glass slabs with powder where the letters "EC" are on the back slab and "CV" on the front slab, and **RIGHT**: point cloud of a face etched in a glass cube. Both examples show: (a) a photograph of the objects, (b) one of the 24 images captured by the camera, and reconstructed volumes at different views with (c) and without (d) attenuation correction.

We used this system to recover several types of inhomogeneous participating media, including, multiple translucent layers, a 3D point cloud of a face etched in a glass cube, and the dynamic process of milk mixing with water. The reconstructed volumes are visualized with the ray casting algorithm [23] in which the opacity function is set to the volume density.

We first perform reconstruction on static volumes. Fig. 6(left) shows the results of an object consisting of two glass slabs with powder on both. The letters "EC" are drawn manually on the back plane and "CV" on the front plane by removing the powder. Thus we create a volume in which only two planes have non-zero density. A photograph of the object is shown in Fig. 6a. We then reconstruct the volume using the proposed method. Fig. 6 shows one of the 24 captured images as well as the reconstructed volume at different views with and without attenuation correction. It shows that attenuation correction improves the results by increasing the density on the back plane.

Similarly, Fig. 6(right) show the reconstruction for a 3D point cloud of a face etched in a glass cube. As shows, our method also achieved good reconstruction of the volume. In this example, multiple scattering and attenuation within the point cloud are much stronger than the previous example. In result, in the reconstructed volume, the half of the face not directly visible to the camera has a lower estimated density (e.g., the relative darker area of the right eye in Fig. 6).

Finally, we use our system to reconstruct time-varying volumes. We take the dynamic process of milk drops dissolving in water as an example. We use a syringe to drip milk drops into a water tank as shown in the adjacent figure. With the proposed method, we are able to reconstruct time-varying volumes with high spatial resolution ($128 \times 128 \times 250$) at 15fps, which recovers the interesting patterns of the dynamic process (see Fig. 7).

8 Limitations

Multiple Scattering. Although utilizing more light elements increases the efficiency of the acquisition, it will increase multiple scattering as well, which will cause biased reconstruction, as the artifacts shown in Fig. 6. One potential way to alleviate this problem is to separate multiple/single scattering by using more complex light codes in a similar way to Nayar et al. [24].
Calibration for the Spatial Coding Method. The spatial coding seems more desirable than the temporal coding due to its high temporal resolution (i.e., volume reconstruction from one single image) and the easy access of high spatial resolution devices. However, it requires highly accurate calibration both geometrically and radiometrically. The defocus of both the projector and the camera needs to be considered as well. In contrast, the temporal coding method is more robust to noise and defocus and easy to calibrate.

9 Conclusions

We proposed compressive structured light for recovering the volume densities of inhomogeneous participating media. Unlike conventional structured light range finding methods where coded light patterns are used to establish correspondence for triangulation, compressive structured light uses coded light as a way to generate measurements which are line-integrals of volume density. By exploiting the sparsity of the volume density, the volume can be accurately reconstructed from a few measurements. This makes the acquisition highly efficient both in acquisition time and illumination power, and thus enables the recovery of time-varying volumetric phenomena.

We view compressive structured light as a general framework for coding the 3D light function $L(x, y, t)$ for reconstruction of signals from line-integral measurements. In this light, existing methods such as laser sheet scanning and laser line interpolation, as well as the spatial coding and temporal coding discussed in this paper, can be considered as special cases. One interesting future direction is to design more complex coding strategies to improve the performance or apply the method to new problems.

Acknowledgments: This work was supported in part by the NSF (ITR-03-25867, CCF-05-41259, IIS-04-12759, IIS-05-28402, CNS-06-14770, and CCF-06-43268), a Sloan Research Fellowship BR-4485, and an ONR Young Investigator award N00014-07-1-0900.

References

1. Salvi, J., Pages, J., Batlle, J.: Pattern codification strategies in structured light systems. Pattern Recognition **37** (2004) 827–849
2. Narasimhan, S., Nayar, S., Sun, B., Koppal, S.: Structured light in scattering media. In: ICCV'05. (2005) 420–427

3. Hawkins, T., Einarsson, P., Debevec, P.: Acquisition of time-varying participating media. In: SIGGRAPH'05. (2005) 812–815
4. Deusch, S., Dracos, T.: Time resolved 3d passive scalar concentration-field imaging by induced fluorescence (LIF) in moving liquids. Measurement Science and Technology **12**(2) (2001) 188–200
5. Fuchs, C., Chen, T., Goesele, M., Theisel, H., Seidel, H.: Density estimation for dynamic volumes. Computers and Graphics **31**(2) (2007) 205–211
6. Candes, E.J., Romberg, J.: Sparsity and incoherence in compressive sampling. Inverse Problems **23**(3) (2007) 969–985
7. Donoho, D.: Compressed sensing. IEEE Trans. on Information Theory **52**(4) (2006) 1289–1306
8. Schechner, Y.Y., Nayar, S.K., Belhumeur, P.N.: A theory of multiplexed illumination. In: ICCV'03. (2003) 808–815
9. Willett, R., Gehm, M., Brady, D.: Multiscale reconstruction for computational spectral imaging. In: Computational Imaging V at SPIE Electronic Imaging. (2007)
10. Lustig, M., Donoho, D., Pauly, J.M.: Sparse MRI: The application of compressed sensing for rapid MRI imaging. Magnetic Resonance in Medicine **58**(6) (2007) 1182–1195
11. Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y.: Robust face recognition via sparse representation (2008)
12. Hasinoff, S., Kutulakos, K.: Photo-consistent reconstruction of semi-transparent scenes by density sheet decomposition. IEEE Trans. on PAMI **29**(5) (2007) 870–885
13. Ihrke, I., Magnor, M.: Image-based tomographic reconstruction of flames. In: SCA. (2004) 36l–375
14. Ihrke, I., Magnor, M.: Adaptive grid optical tomography. Graphical Models **68**(5) (2006) 484–495
15. Trifonov, B., Bradley, D., Heidrich, W.: Tomographic reconstruction of transparent objects. In: EGSR. (2006) 51–60
16. Ihrke, I., Kutulakos, K.N., Lensch, H.P.A., Magnor, M., Heidrich, W.: State of the art in transparent and specular object reconstruction. In: STAR Proc. of Eurographics. (2008)
17. Candes, E.J., Romberg, J., Tao, T.: Stable signal recovery from incomplete and inaccurate measurements. Communications on Pure and Applied Mathematics **59**(8) (2006) 1207–1223
18. Ishimaru, A.: Wave Propagation and Scattering in Random Media. IEEE Press, New York (1978)
19. GLPK: The Gnu Linear Programming Kit. http://www.gnu.org/software/glpk
20. Olshausen, B.A., Field, D.J.: Emergence of simple-cell receptive field properties by learing a sparse code for natural images. Nature **381** (1996) 607–608
21. Simoncelli, E.P.: Statistical models for images: compression restoration and synthesis. In: Proc Asilomar Conference on Signals, Systems and Computers. (1997) 673–678
22. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation noise removal algorithm. Physica D **60** (1992) 259–268
23. Schroeder, W., Martin, K., Lorensen, B.: The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics. 4th edn. Pearson Eduction, Inc. (2006)
24. Nayar, S.K., Krishnan, G., Grossberg, M.D., Raskar, R.: Fast separation of direct and global components of a scene using high frequency illumination. In: SIGGRAPH'06. (2006) 935–944

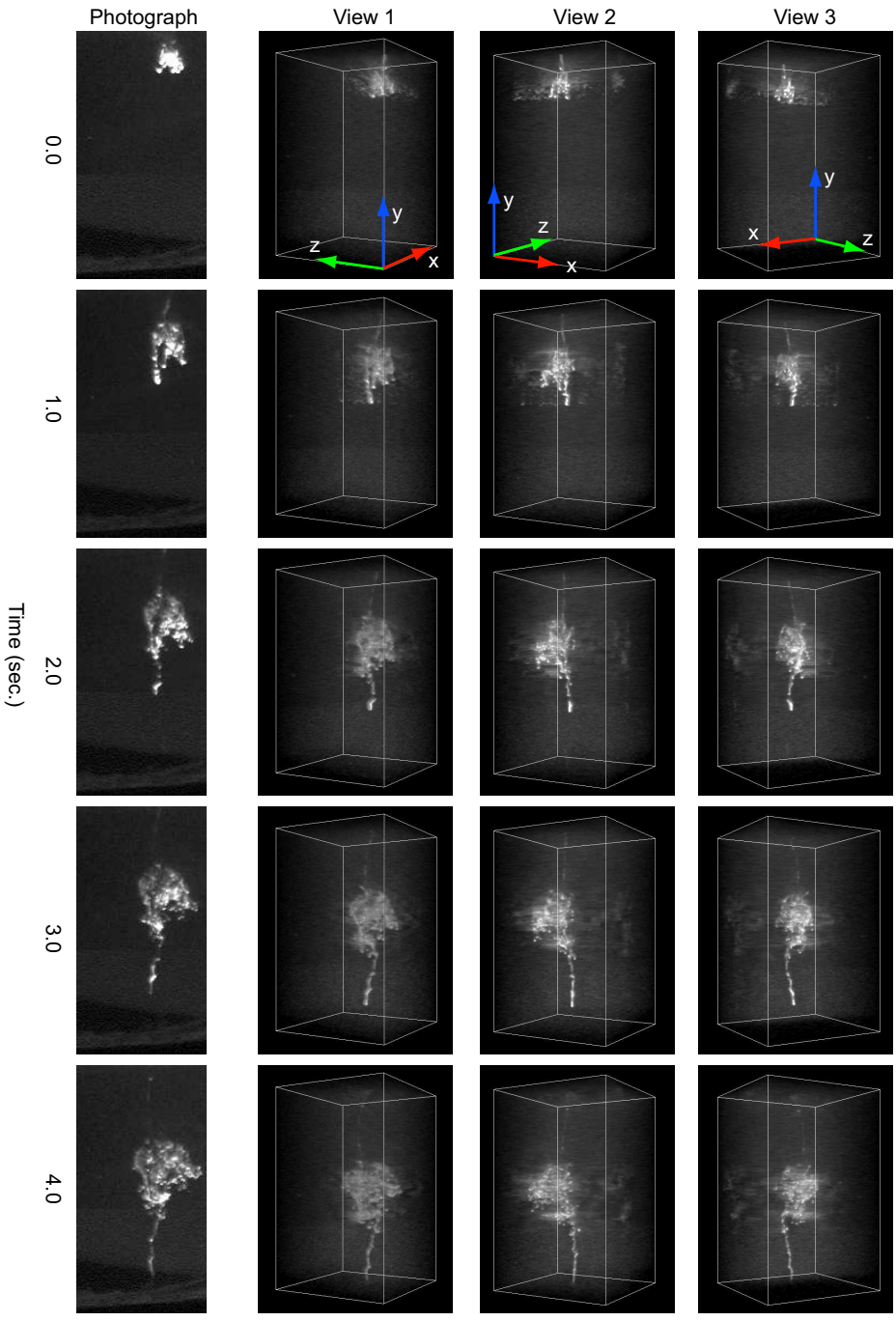


Fig. 7. Reconstruction results of milk drops dissolving in water. 24 images are used to reconstruct the volume at $128 \times 128 \times 250$ at 15fps. The reconstructed volumes are shown in three different views. Each row corresponds to one instance in time. The leftmost column shows the corresponding photograph (i.e., all projector pixels emit white) of the dynamic process.