

Towards Photography Through Realistic Fog

Guy Satat, Matthew Tancik, Ramesh Raskar

MIT Media Lab

guysatat@mit.edu

Abstract

Imaging through fog has important applications in industries such as self-driving cars, augmented driving, airplanes, helicopters, drones and trains. Here we show that time profiles of light reflected from fog have a distribution (Gamma) that is different from light reflected from objects occluded by fog (Gaussian). This helps to distinguish between background photons reflected from the fog and signal photons reflected from the occluded object. Based on this observation, we recover reflectance and depth of a scene obstructed by dense, dynamic, and heterogeneous fog. For practical use cases, the imaging system is designed in optical reflection mode with minimal footprint and is based on LIDAR hardware. Specifically, we use a single photon avalanche diode (SPAD) camera that time-tags individual detected photons. A probabilistic computational framework is developed to estimate the fog properties from the measurement itself, without prior knowledge. Other solutions are based on radar that suffers from poor resolution (due to the long wavelength), or on time gating that suffers from low signal-to-noise ratio. The suggested technique is experimentally evaluated in a wide range of fog densities created in a fog chamber. It demonstrates recovering objects 57cm away from the camera when the visibility is 37cm. In that case it recovers depth with a resolution of 5cm and scene reflectance with an improvement of 4dB in PSNR and 3.4× reconstruction quality in SSIM over time gating techniques.

1. Introduction

The ability to see through dense fog is exciting for photography and essential for safer transportation systems. Examples include self-driving cars that can operate in challenging weather; augmenting a driver with a heads-up-display to alert of occluded objects and read road signs; providing a clear flight path for drones, airplanes, and helicopters at low-level flight; and allowing trains to maintain speed in foggy weather.

Here, we show that the time profiles of light scattered from fog and occluded objects have different distributions

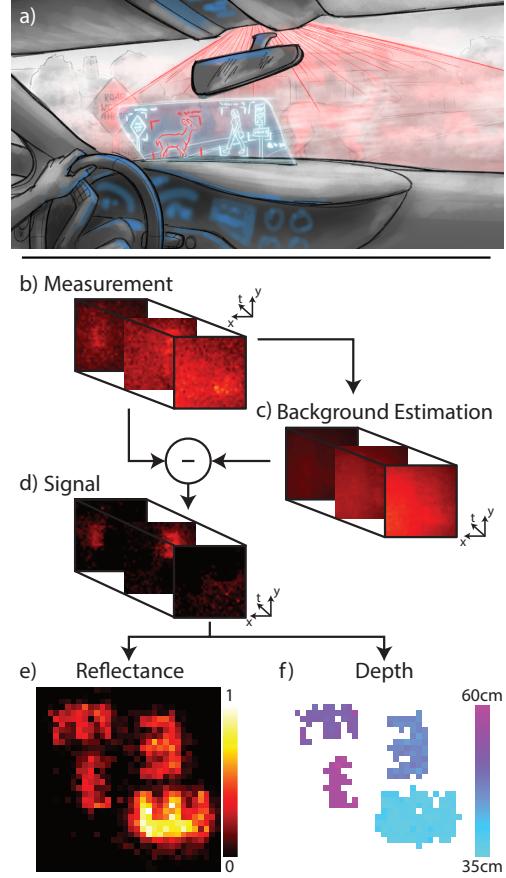


Figure 1. Seeing through dense, dynamic, and heterogeneous fog. a) A concept sketch of LIDAR sensing system in heads-up-display to show reflectance modulated by depth of objects obscured by fog. Other techniques or EM wavelengths return geometry but not (near) visible light reflectance. Imaging pipeline: b) A SPAD camera captures time-resolved single photon counts. c) A probabilistic framework is used to estimate the properties of the back reflection from the fog. d) The back reflection is subtracted from the measurement to recover the signal and is processed to produce the target e) reflectance and f) depth map.

(Gamma and Gaussian respectively). A probabilistic framework is used to evaluate the fog properties from the measurement itself, and separate between photons reflected from the fog and those reflected from the object. The suggested system relies on time-resolved sensing, similar to the technologies used in LIDAR. Our solution is based on a single photon avalanche diode (SPAD) camera that detects and time-tags photons as they arrive at different pixels on the sensor. Since the system is based on time-resolved sensing, it is capable of recovering the occluded scene reflectance and depth maps.

The main industry solution for imaging through fog is based on radio waves, e.g. in 94GHz [2] where fog is transparent. There are several challenges in using radar for imaging, including: 1) resolution – due to the long wavelength it is hard or even impossible to classify objects, and the use cases are limited to detection; and 2) optical contrast – at such long wavelengths it is only possible to measure bulk properties of materials [1], and impossible for example to identify road lane marks and read road signs. Techniques for long range and large field of view imaging through volumetric scattering in the visible range are usually limited to time gating, which requires long integration times and prior knowledge of the scene depth map. Furthermore, time gating and many other techniques to image through scattering media, such as phase conjugation and acousto-optics, reject scattered photons during the measurement process. Because these methods reject a substantial amount of the optical signal, they operate at a low signal-to-noise ratio (SNR).

Instead of rejecting the scattered photons during the measurement process, we measure both scattered and unscattered photons and computationally use all of the optical signal to reject the fog. The motivation for our approach is the fact that the scattered photons hold substantial information about the target and the scattering medium, and that information is useful for improving imaging capabilities.

1.1. Main Contributions

The main contributions of this paper are:

1. A time-domain technique for seeing through fog, demonstrating the ability to recover the scene image and depth.
2. An experimental demonstration of the technique in dense, dynamic, and heterogeneous fog conditions in a fog chamber (as opposed to milky water or other phantoms) for a wide range of fog densities (visibilities) and static scenes.
3. A probabilistic method to estimate pixel-wise fog parameters from the measurement itself without any calibration or prior knowledge.

2. Related Works

Computer vision has been tackling imaging through scattering along different axes. Significant efforts have recently been dedicated to discrete scattering events such as looking around corners and single scatter diffusers [43, 15, 20, 42, 33, 7, 31, 22, 5]. Other examples include vision through the atmosphere [30, 29], seeing through translucent objects [10], underwater imaging [36], 3D reconstruction [40], color recovery [38], and controlling light transport [13]. Structured light in scattering media was demonstrated [28] by spatially decoupling the back reflectance and signal. Here we achieve this with a per-pixel time profile.

In the context of volumetric scattering, significant effort has been dedicated to improving contrast in haze [9]. These cases handle very few scattering events, and it is common to model the scattering with a single scatter event [38]. Some techniques to overcome such atmospheric scattering include polarization [34, 41], patch base recurrence [3], haze line estimation [4], and lightfield imaging [39]. Data driven techniques for dehazing have been explored [6] along with rendering techniques [26, 35]. While most work is dedicated to reflectance recovery, some recover depth of objects through scattering media [16, 14, 19, 39].

Works that explicitly handle the case of dense scattering materials similar to the one discussed here include depth sensing [16, 14, 20], cloud tomography [17, 24, 25], scattering parameters estimation for computer graphics [11, 12], and structured light [28]. Our goal is different since we aim to perform both reflectance and depth recovery of a scene occluded by highly scattering media. We demonstrate our recovery technique in a fog chamber that generates dense, dynamic, and heterogeneous fog. In our demonstration, the camera and illumination are adjacent to the fog chamber to further emulate realistic imaging scenarios.

Other techniques for large scale imaging through dense fog include time gating [8, 23] (a recent survey on transient imaging is provided in [18]). These methods work when the object is far away (less coupling with the back reflection from the fog), but are limited by signal-to-noise ratio which requires long integration times and a stationary scene. Another limitation of time gating is the need to manually select the time-gate bin or to rely on prior depth map knowledge. The suggested method uses the scattered photons to estimate the fog properties and computationally reject it from the measurement. Thus, it works at higher SNR and does not depend on prior knowledge of the scene or scatterer.

In this work we use single photon counts along with a probabilistic model to reject the fog. The use of probabilistic frameworks for imaging through scattering media has been previously suggested in [32]; however it was limited to transmission mode which is fundamentally different than the reflection problem discussed here (in transmission mode there is no need to handle back reflection). Other examples

of using probabilistic models for imaging with a few photons [21, 37] did not consider imaging through scattering media.

3. Parametric Model for Imaging Through Fog

Consider a pulsed light source emitting photons into a foggy scene. Adjacent to the light source is a time-resolved camera. Each pixel in the camera time-tags individual photons as they arrive to the sensor. Each pixel in a measurement frame holds the arrival time of the first detected photon during the frame acquisition time. A measured photon can be classified as:

- **Background photon** – a photon that did not interact with the target, thus it only holds information about the fog. Due to the scattering dynamics, background photons arrive at different times.
- **Signal photon** – a photon that interacted with the target, thus it holds information about target reflectivity and depth.
- **Dark counts (noise)** – these false detections are uniformly distributed in time.

The dark count rate in our detector is below 30Hz which is an order of magnitude less than the background and signal counts in our measurements, thus we neglect it from our model.

The parametric model described below is pixel-wise (see further discussion in Sec. 7). First, we model the background photons. It is common to model scattering with diffusion theory. In this case, the distance a photon propagates between consecutive scattering events is exponentially distributed with a mean of $1/\lambda$, where λ is known as the mean free path, the mean distance a photon propagates between scattering events. Equivalently the time between scattering events is also exponentially distributed $\tau_i \sim EXP(\mu_s)$, where $1/\mu_s$ is the mean time between scattering events. We note that since the scattering events are independent, so are the different times τ_i . A detected photon undergoes multiple scattering events such that the detection time is $T = \sum_{i=1}^k \tau_i$. The sum of independent exponential random variables is Gamma distributed $T \sim GAMMA(k, \mu_s)$, where k and μ_s are the shape and rate parameters. Thus, we can model the probability of measuring a background photon at time t , denoted as $f_T(t|B)$, with the parameters k and μ_s encoding the physical properties of the fog (see further analysis on k in the Discussion):

$$f_T(t|B) = \frac{\mu_s^k}{\Gamma(k)} t^{(k-1)} \exp \{-\mu_s t\} \quad (1)$$

where $\Gamma(k)$ is the Gamma function. Figure 2 shows time-resolved measurements at different concentrations of fog

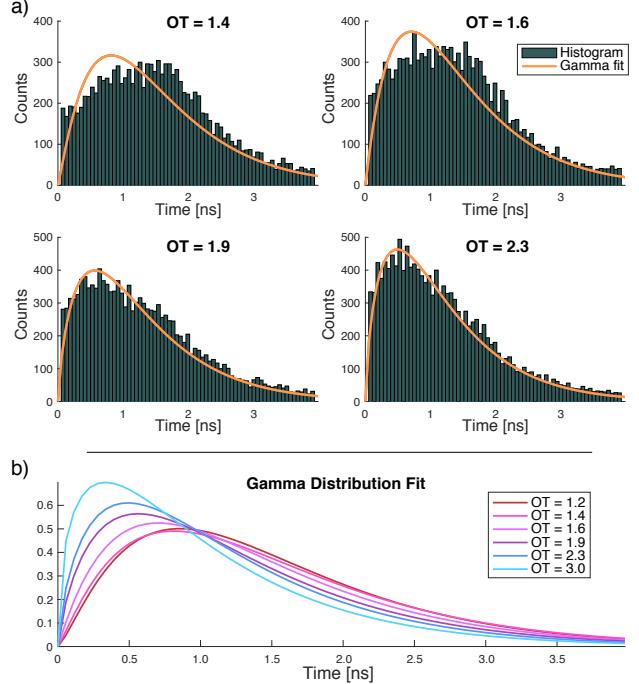


Figure 2. Back reflection model. a) Experimental time-resolved measured histograms along with fitted Gamma distributions. The panels correspond to different optical thicknesses (OT) of fog. The plots show that a Gamma distribution captures well the dynamics of time-resolved scattering in fog, especially at high densities. b) Fitted Gamma distributions for a wide range of fog densities. The plots show that different fog densities result in different time profiles.

and the corresponding Gamma distribution fits. We measure fog densities with optical thickness (OT) where $OT=0$ is clear visibility.

Next we model the time of arrival of photons that interacted with the target as a Gamma distributed random variable, given that the photon interacted with the target (with similar arguments to the background model). In practice we find that we can use a Gaussian model for this distribution. This can be justified since in this case the number of scattering events is large, and when the shape parameter, k , of a Gamma distribution is large it resembles a Gaussian distribution. The distribution mean, μ , corresponds to the depth of the object, and the variance, σ^2 , encodes the scattering these photons undergo, such that the probability of measuring a signal photon at time t is:

$$f_T(t|S) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\left(\frac{t-\mu}{\sigma} \right)^2 \right\} \quad (2)$$

Finally, the models are combined with the law of total probability. The overall probability of detecting a photon at

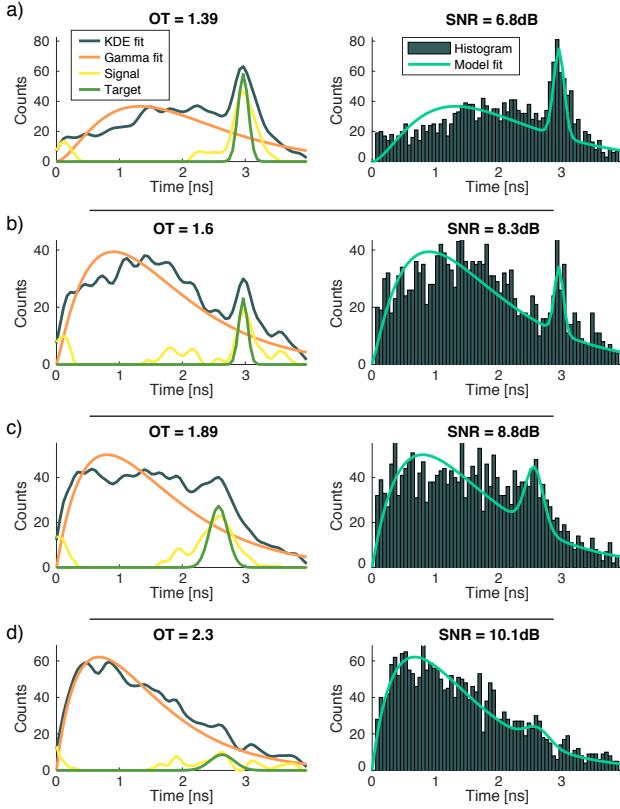


Figure 3. Rejecting back reflection and signal recovery. The left column shows the recovered KDE and Gamma distributions (from the raw photon counts), the estimated signals by subtracting the two, and the estimated target distributions. The right column shows the histogram generated by the raw photon counts and the fitted model by Eq. 5 including the SNR between the two. Rows show different cases of fog concentrations, optical thicknesses of $OT = 1.4, 1.6, 1.9, 2.3$ for panels a-d respectively. The target in panels a+b is at a depth that corresponds to 3.02ns, and the target in panels c+d is at 2.58ns. Note that in all cases there are substantially more background photons than signal photons.

time t is:

$$f_T(t) = f_T(t|S)P(S) + f_T(t|B)P(B) \quad (3)$$

We define $P(B) = b$, and $P(S) = r$. Here b and r include the optical losses of the system (such as absorption and detector efficiency). The ratio between b and r captures the probability of measuring a background vs. signal photon. The target reflectivity is captured by r .

4. Imaging Through Dense Fog

Given the pixel-wise time-resolved photon counts, our goal is to decouple the back reflectance and the signal. The estimation pipeline is demonstrated in Fig. 1 and is com-

posed of four steps: 1) time profile estimation, 2) background distribution estimation, 3) signal estimation, and 4) target recovery. Fig. 3 shows the estimation results for different levels of fog and for targets at different depths.

4.1. Time Profile Estimation

The individual photon detection events are mapped to the distribution $f_T(t)$ using a kernel density estimator (KDE) denoted by $\widehat{f}_T(t)$. KDE has several advantages as opposed to a traditional histogram, primarily since there is no need to specify a bin size, and it performs well in cases of a few sampling points such as here.

4.2. Background Estimation

The background is governing the signal in the cases discussed here. Thus, the raw photon counts are used to estimate the parameters in $f_T(t|B)$. As mentioned above, the physical model describing the background leads to a Gamma distribution. The distribution parameters are estimated using maximum likelihood. The estimated distribution is denoted by $\widehat{f}_T(t|B)$. Since the majority of the photons are background photons, we use all measured photons in this step and effectively treat the signal photons as noise.

4.3. Signal Estimation

With the probability functions for the total time profile and background, it is possible to subtract the two curves and isolate a proxy to the probability distribution of the signal $f_T(t|S) \sim f_T(t) - f_T(t|B)$. To that end we fit $\widehat{f}_T(t) - \widehat{f}_T(t|B)$ to a Gaussian that estimates the signal $\widehat{f}_T(t|S)$.

4.4. Scene Reconstruction

With the background and signal estimated distributions, the parameters r and b are estimated by solving:

$$\left[\hat{r}, \hat{b} \right] = \arg \min_{[r,b]} \sum_t \left(r \widehat{f}_T(t|S) + b \widehat{f}_T(t|B) - \widehat{f}_T(t) \right)^2 \quad (4)$$

This is where edge cases in which there is no fog or no target are accounted for – the solution will be $b \rightarrow 0$ or $r \rightarrow 0$ accordingly.

So far, all estimators were probability functions. These are mapped to actual photon counts $N(t)$, the number of photons measured at time bin t , by computing a normalization factor $\hat{\alpha}$ such that $\sum_t N(t) = \hat{\alpha} \sum_t [\hat{r} \widehat{f}_T(t|S) + \hat{b} \widehat{f}_T(t|B)]$. This step is necessary for consistent results across pixels that receive a different number of photons. The final estimated model is:

$$\hat{N}(t) = \hat{\alpha} \left[\hat{r} \widehat{f}_T(t|S) + \hat{b} \widehat{f}_T(t|B) \right] \quad (5)$$

Figure. 3 shows the different estimations in different cases of fog and targets. It also provides the SNR between the estimated model and the measured histogram. The supplement provides further examples.

Lastly, the target parameters are extracted:

Reflectance Estimation: The reflectance value per pixel is captured by $\hat{r} \hat{f}_T(t|S)$. We find that using the peak of this function provides good reflectance recovery. Another option uses \hat{r} which proved noisier in practice. The estimated depth is then used to account for depth dependent signal attenuation (see below).

Depth Estimation: The target depth is encoded in the early part of the Gaussian distribution (corresponds to ballistic photons). In practice, the estimated Gaussian variances in our experiments are in the order of 1 – 2 time bins. This is due to both scattering and measurement jitter. We found that using $\hat{\mu}$ from the Gaussian estimation is a robust estimate of the target depth. Since a Gaussian distribution is fitted regardless to whether there is an object in the pixel or not, we use the reflectance estimation as a confidence map to reject pixels where no target exists.

4.5. Implementation Details

For numerical stability in parameter estimation, the photons arrival times are provided in picoseconds. The KDE estimator uses a standard Gaussian kernel with a bandwidth of 80ps. Negative values in the subtraction of the measured time profile and background are set to zero. To further improve the reconstruction quality of the reflectance map we multiply it by μ (to amplify points that are further away) and threshold it to 0.2 of the peak. The threshold used as a confidence map for the depth estimation is 0.2 of the final reflectance map.

The computation time to recover the reflectance and depth per pixel is 0.03sec using an unoptimized Matlab code on a 3rd generation i7 desktop.

5. Experimental Setup

The experimental setup is shown in Fig. 4. The detector is a PhotonForce PF32 SPAD camera that is composed of 32×32 pixels. Each pixel is single photon sensitive and time-tags measured photons with a nominal time resolution of 56ps. The camera exposure time is set to $100\mu\text{s}$ (the PF32 measures the arrival time of the first detected photon per-pixel per-exposure). Each reconstruction is composed of 20,000 frames. We can produce a new reconstruction every $100\mu\text{s}$, while using a sliding window with a history of 2sec (see further analysis in the Discussion). For illumination, a SuperK pulsed super-continuum laser is spectrally filtered to a narrow band around 580nm (the camera is equipped with a similar filter to reject background). The laser repetition rate is 80MHz with a pulse duration of 5ps

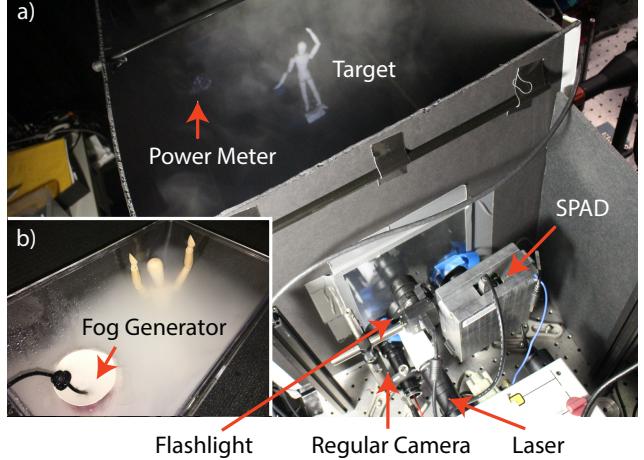


Figure 4. Experimental setup photograph. a) The fog chamber with a mannequin inside. This photograph was taken with minimal fog density and shows the SPAD, pulsed laser, traditional camera, and flashlight. Illumination and measurement are performed through a glass window on the chamber. A power meter is placed inside the fog chamber to quantify the optical thickness. The fog generator is composed of an ultrasonic transducer in water and a fan placed on the far side of the chamber (not visible). b) Example of the fog generator inside a small open aquarium. In this case the fan is off which results in a low concentration.

with an average laser's optical power of 0.15W. The laser is diffused before entering the chamber and it flood illuminates the scene (without fog). The camera and laser are positioned in a reflection mode.

To evaluate and compare our method we placed a regular monochromatic camera (Point Grey Chameleon) along with an independent CW flashlight at a wavelength of 850nm. The flashlight average optical power is 1W and the camera quantum efficiency at this wavelength is 15%. The different wavelengths were chosen to make sure that the two imaging systems can operate simultaneously without affecting one another. Spectral filters are placed on the SPAD and regular cameras to ensure that each camera measures contributions only from its dedicated light source. The flashlight floodlit the tank and was positioned such that it did not illuminate the target directly to reduce the glare from the fog. The camera integration time is 100ms. In all of the reported results, the background lights are turned off so that both imaging systems equally benefit from a directional illumination source.

The cameras and illumination sources are placed adjacent to the fog chamber. The chamber dimensions are $0.5 \times 0.5 \times 1\text{m}^3$. To generate fog, a powerful ultrasonic transducer is placed in water along with a fan (similar to a cold mist humidifier). The fog generator is placed at the far side of the fog chamber. This configuration is capable

of producing dense fog with visibility of a few centimeters into the chamber. A supplemental video taken with the regular camera shows the chamber as it is being filled with fog. To quantitatively evaluate the level of the fog, the far side of the chamber also holds an optical power meter to measure the optical thickness of the fog. Optical thickness at time t is calculated by $-\log(P_t/P_0)$, where P_0 is the power measured when there is no fog, and P_t is the power at time t .

6. Experimental Results

The experimental system described above was used to evaluate the suggested approach. Several targets are placed in the fog chamber. The fog generator is turned on, and a continuous capture of SPAD frames is performed until the fog density in the chamber saturates (roughly 15 minutes).

To evaluate the results we compare our reflectance reconstruction to the measurement taken with the regular camera. Note that the different wavelength used for the regular camera undergoes less scattering which results in a sharper image, especially in low fog densities. Because of the different perspective and acquisition properties, we use the regular camera only for qualitative comparison. The second comparison is to photon counting from the SPAD camera. In this mode the camera simply accumulates individual detected photon events. The third comparison is to time gating. In this mode the time bin was selected manually to the first time bin that holds information about the target. We compute peak SNR (PSNR) and structural similarity (SSIM, ranges in $[0, 1]$, higher is better) to quantitatively compare our reflectance recovery method to photon counting and time gating (the chosen ground truth is taken from a photon counting measurement without fog). Similarly, the regular camera results are compared (with SSIM and PSNR) to a regular photo taken without fog (not shown).

Figure 5 shows results for two sets of targets at different fog densities. The first target is composed of a set of ‘E’ shapes ($3 \times 5\text{cm}^2$) at different orientations and depths (equal spacings in the range of 37-57cm from the camera). As can be seen from the results, our method is able to reject the significant backscatter that governs the regular camera and photon counting results. In comparison to time gating we note that time gating is much noisier, requires one to manually select the correct time bin, and recovers only one depth. The suggested method outperforms these techniques in both SSIM and PSNR, and degrades much slower with increasing fog levels. Furthermore, the method accurately recovers the depth of the different targets up to an optical thickness of 2.1, after which it loses the far target while still recovering the closer ones. The average depth reconstruction error (for all pixels, depths, and optical thicknesses up to 2.1) is $0.08 \pm 0.3\text{cm}$. The second result in Fig. 5 is of a mannequin ($\sim 20\text{cm}$ tall) placed at the depth of 35cm. This result further demonstrates the above conclusions. The sup-

plemental videos show additional results over time as fog fills the chamber.

7. Discussion

7.1. Limitations

The key limitations of the suggested approach are listed below; the remaining discussion section further analyzes these limitations and suggests future work to alleviate them:

1. Our approach is pixel-wise and neglects the spatial nature of scattering. While this is enough to reject the background from the measurement, it is not able to spatially deblur the signal. Spatial blurring will be more prominent in large scale scenes and high resolution sensors.
2. The demonstrated results are captured with acquisition times of 2sec (with a mean of 2,440 photons per acquisition). While this is enough for dynamic fog with a stationary scene, it is not enough for dynamic scenes.

7.2. Sensitivity to Sensor Spatial Resolution

The sensor used here is composed of only 32×32 pixels. SPAD cameras with mega-pixel spatial resolution and nanosecond time resolution have already been demonstrated [27]. Such sensors with higher pixel count would be beneficial in utilizing spatial scene priors to denoise and smooth the result. Furthermore, a higher resolution sensor would also be useful as part of an imaging framework that accounts for the complete space-time scattering profile similar to the one suggested in [32]. This would potentially sharpen the results further.

7.3. Sensitivity to Sensor Time Resolution

The suggested imaging method is based on the notion that background and signal photons have different statistics in time. This allows one to distinguish between them and reject the back reflectance from the fog. As the detector time resolution reduces, this ability diminishes. The relevant time scales to consider are the standard deviation of the background Gamma distribution and Gaussian target distribution. The sensor time resolution should be smaller than both. Another aspect of time resolution is its mapping to depth resolution and accuracy, which is potentially stricter in real world situations.

7.4. How Many Photons are Needed

In our reconstructions we used 20,000 frames. The actual number of measured photons varies as a function of the scene and fog density (brighter target and denser fog result in more photons). Fig. 6a shows the number of measured photons from pixels that correspond to targets at different

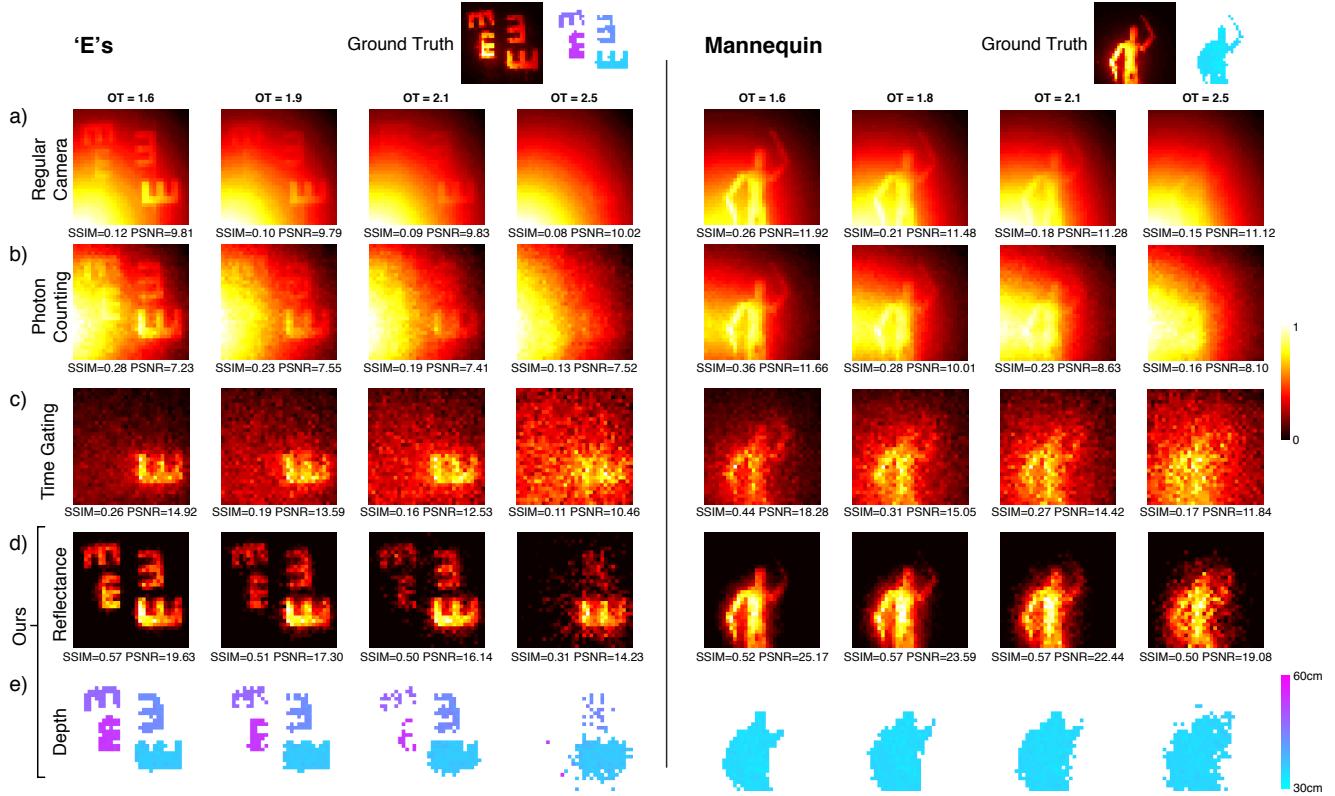


Figure 5. Results for imaging through fog. Different columns show cases of different levels of fog for a set of ‘E’ shapes at different orientations and a mannequin. Rows show different reconstructions including: a) Image taken with a regular camera (the longer wavelength used for this measurement undergoes less scattering, which results in less challenging imaging conditions). b) Result with SPAD camera photon counting mode. c) Result of time gating using the SPAD camera, where the time gate was selected manually to the first time bin with meaningful information. d) Reflectance reconstruction with the suggested technique. e) Depth reconstruction with the suggested technique. SSIM and PSNR metrics provide quantitative comparisons. The top row shows the ground truth reflectance and depth. This ground truth is based on the photon counting measurement without fog.

depths and a background pixel at different levels of fog. As predicted, when fog is added, the number of captured photons from targets reduces, whereas background photon counts increase, such that the number of photon counts from all pixels converge to similar values. At an optical thickness of 2, the number of photon counts from the further targets matches the number of background photon counts. We also note that for far targets, initially the number of photons is reduced and then, as more fog is added, it starts to rise and they behave more like background pixels.

The data for Fig. 6a is based on 10,000 frames and 100 million pulses which correspond to a mean acquisition efficiency of 2.5% at an optical thickness of 2.1. This number is a result of the small lens used in our experiment, low photo-sensitive fill factor in the SPAD camera (order of 1%), laser power, and laser repetition rate. Future hardware improvements along these axes may improve the measurement efficiency and reduce the acquisition time.

To further explore the dependency on the number of measured photons (directly related to acquisition time) we reconstructed the ‘E’s dataset using different numbers of measured frames; Fig. 6b,c shows the result. We note that for low fog densities it is possible to use less frames and recover with similar accuracy. However, as the concentration increases it is less likely to measure photons from the far targets and more photons are needed. Since part of the algorithm is estimating the properties of the fog, it is possible to adapt the acquisition time to work as fast as possible for a given level of fog. We leave this to a future work.

Another alternative is to use more photon efficient algorithms. We note that the model in Eq. 3 includes six parameters. To make the estimation process more robust, our pipeline decouples the estimation of these parameters to four steps based on the physical nature of the problem. In a future work, we plan to develop an expectation maximization algorithm to estimate these six parameters directly

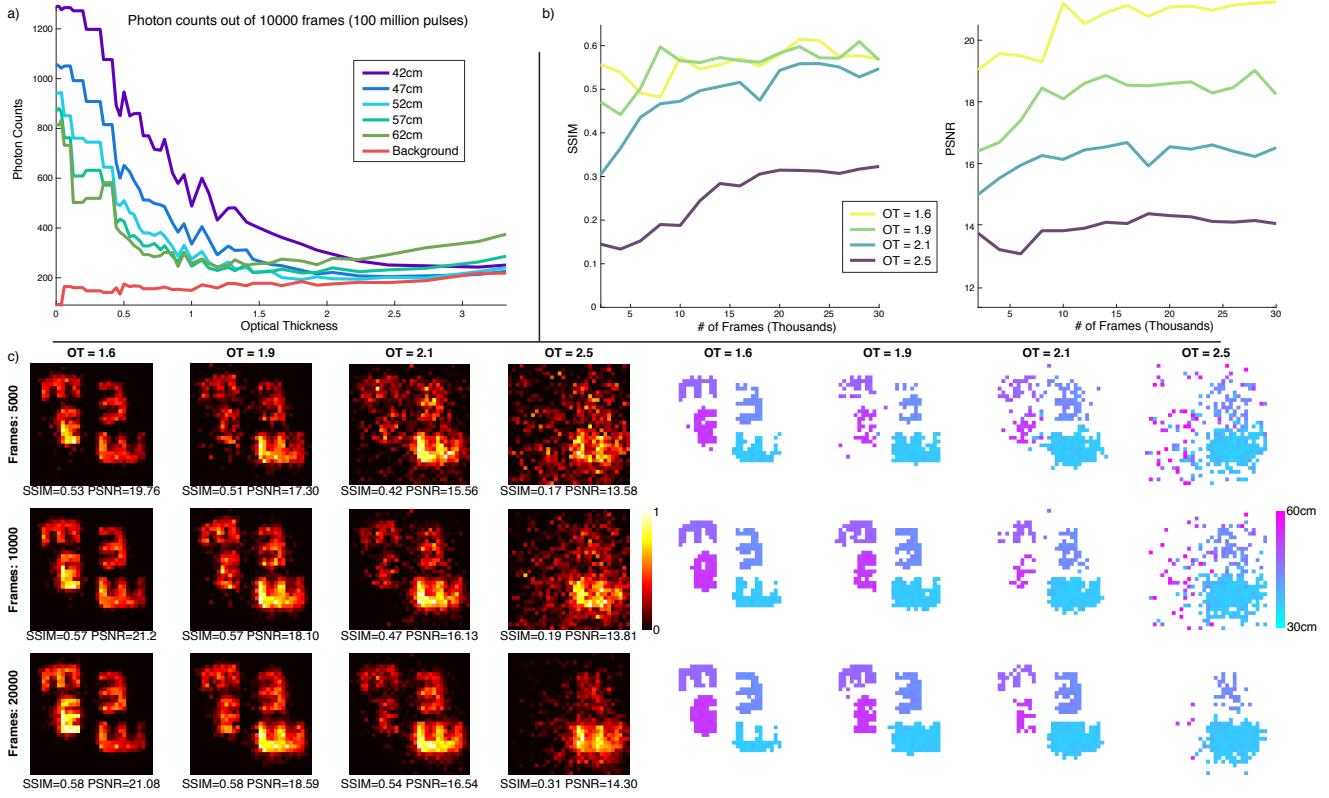


Figure 6. Effect of the number of measured photons on reconstruction quality. a) Number of measured photons as a function of optical thickness for targets at different depths between 42 to 62cm and a background region. b) Reconstruction quality (SSIM and PSNR) for a different number of frames, evaluated on the ‘E’s dataset. c) Several example reconstructions using a different number of frames (rows) and at different fog concentrations (columns). As the fog becomes denser, more photons are required to accurately recover the far objects.

from the photon counts along with the distribution to which each photon belongs to. This approach would potentially be more photon efficient. A maximum likelihood approach was used in [21] to reject background from signal with a few single photon measurements when the statistics of the background and signal were known a priori (this is similar to estimating r in our case with a priori knowledge of background and signal statistics).

7.5. Sources for Model Miss Match

Several aspects were neglected from our model:

Noise: As mentioned before, the dark counts in our experiments are negligible compared to the signal. Future work aiming at more photon efficient acquisition would have to take dark counts into account.

Absorption: The model essentially treats absorption as any other optical loss in the system. Since the model only takes into account measured photons, it is invariant to the number of actual photons sent to scene. Thus, absorption and other losses are irrelevant to the reconstruction procedure and would only affect the total acquisition time.

Gamma distribution: The model developed in section 3 assumes that k , the number of scattering events before detection, is a constant parameter similar to μ_s . In practice k is a random variable that is realized per photon detection. A more accurate approach would model the background as $f_T(t|B) = \int_{k=1}^{\infty} f_G(k, \mu_s) f_K(k) dk$ where $f_G(k, \mu_s)$ is a Gamma distribution and $f_K(k)$ would model the probability of a photon undergoing k scattering events before detection. In practice, we find that our simpler model is able to capture the time-resolved scattering accurately enough, especially in dense fog, as can be seen in Fig. 2.

7.6. Scaling to Real World Use Cases

Several aspects of the problem have to be taken into account when considering real world applications of the suggested approach:

Scene scale: Consider a specific concentration of fog in our chamber. When the chamber volume increases, the fog concentration reduces. Thus, the number of scattering events remains the same while the mean free path in-

creases. However, as the distance between scattering events increases, the spatial nature of scattering would become more significant. That would require detectors with higher spatial resolution and the consideration of the space-time scattering profile discussed above.

Time resolution: While the required spatial resolution for larger scale scenes is higher, the demand for time resolution decreases. As mentioned above, the required time resolution is a function of the scattering distribution standard deviation which would become larger. Thus the main constraint on time resolution would come from the requirement for accurate depth resolution.

Use of existing hardware: The development of LIDAR systems for self-driving cars is an ongoing process. Some of these systems are based on pulsed illumination with time gating for depth mapping and are usually based on raster scanning. This is very similar to a single pixel in the hardware used here.

8. Conclusions

The presented approach provides a computational imaging probabilistic framework for seeing through dense, dynamic, and heterogeneous fog. The technique provides an image and a depth map of an occluded scene and was demonstrated on a wide range of optical thicknesses, where the scattering mean free path is much smaller than the scene dimensions. The hardware requirements of the suggested system are similar to LIDAR, making it suitable for various imaging applications.

Acknowledgements

The authors thank professor Moungi Bawendi and his group for their support with experimental hardware.

References

- [1] F. Adib, C. Y. Hsu, H. Mao, D. Katabi, and F. Durand. Capturing the human figure through a wall. *ACM ToG*, 34, 2015.
- [2] R. Appleby and R. N. Anderton. Millimeter-wave and submillimeter-wave imaging for security and surveillance. *Proceedings of the IEEE*, 95, 2007.
- [3] Y. Bahat and M. Irani. Blind dehazing using internal patch recurrence. In *ICCP*, 2016.
- [4] D. Berman, S. Avidan, et al. Non-local image dehazing. In *CVPR*, 2016.
- [5] K. L. Bouman, V. Ye, A. B. Yedidia, F. Durand, G. W. Worrell, A. Torralba, and W. T. Freeman. Turning corners into cameras: Principles and methods. In *ICCV*, 2017.
- [6] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. DehazeNet: an end-to-end system for single image haze removal. *IEEE Trans. on Image Processing*, 2016.
- [7] P. Caramazza, A. Boccolini, D. Buschek, M. Hullin, C. Higham, R. Henderson, R. Murray-Smith, and D. Facio. Neural network identification of people hidden from view with a single-pixel, single-photon detector. *arXiv:1709.07244*, 2017.
- [8] O. David, N. S. Kopeika, and B. Weizer. Range gated active night vision system for automobiles. *App. Opt.*, 45, 2006.
- [9] R. Fattal. Single image dehazing. *ACM ToG*, 27, 2008.
- [10] C. Fuchs, M. Heinz, M. Levoy, H. P. Seidel, and H. P. A. Lensch. Combining confocal imaging and descattering. *Computer Graphics Forum*, 2008.
- [11] I. Gkioulekas, A. Levin, and T. Zickler. An evaluation of computational imaging techniques for heterogeneous inverse scattering. In *ECCV*, 2016.
- [12] I. Gkioulekas, S. Zhao, K. Bala, T. Zickler, and A. Levin. Inverse volume rendering with material dictionaries. *ACM ToG*, 32, 2013.
- [13] M. Gupta, S. G. Narasimhan, and Y. Y. Schechner. On controlling light transport in poor visibility environments. In *CVPR*, 2008.
- [14] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM ToG*, 34, 2015.
- [15] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin. Diffuse mirrors: 3d reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *CVPR*, 2014.
- [16] F. Heide, L. Xiao, A. Kolb, M. B. Hullin, and W. Heidrich. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Opt. Exp.*, 22, 2014.
- [17] V. Holodovsky, Y. Y. Schechner, A. Levin, A. Levis, and A. Aides. In-situ multi-view multi-scattering stochastic tomography. In *ICCP*, 2016.
- [18] A. Jarabo, B. Masia, J. Marco, and D. Gutierrez. Recent advances in transient imaging: A computer graphics and vision perspective. *Visual Informatics*, 1, 2017.
- [19] A. Kadambi, J. Schiel, and R. Raskar. Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight. In *CVPR*, 2016.
- [20] A. Kadambi, H. Zhao, B. Shi, and R. Raskar. Occluded imaging with time-of-flight sensors. *ACM ToG*, 35, 2016.
- [21] A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal. First-photon imaging. *Science*, 343, 2014.
- [22] J. Klein, C. Peters, J. Martín, M. Laurenzis, and M. B. Hullin. Tracking objects outside the line of sight using 2D intensity images. *Sci. Rep.*, 2016.
- [23] M. Laurenzis, F. Christnacher, E. Bacher, N. Metzger, S. Schertzer, and T. Scholz. New approaches of three-dimensional range-gated imaging in scattering environments. In *SPIE*, 2011.
- [24] A. Levis, Y. Y. Schechner, A. Aides, and A. B. Davis. Airborne three-dimensional cloud tomography. In *CVPR*, 2015.
- [25] A. Levis, Y. Y. Schechner, and A. B. Davis. Multiple-scattering microphysics tomography. In *CVPR*, 2017.
- [26] K. Li, Y. Li, S. You, and N. Barnes. Photo-realistic simulation of road scene for data-driven methods in bad weather. In *ICCV*, 2017.
- [27] R. M. Marino and W. Davis. Jigsaw: a foliage-penetrating 3d imaging laser radar system. *Lincoln Lab Journal*, 15, 2005.

- [28] S. Narasimhan, S. Nayar, Bo Sun, and S. Koppal. Structured light in scattering media. In *ICCV*, 2005.
- [29] S. G. Narasimhan and S. K. Nayar. Contrast restoration of weather degraded images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25, 2003.
- [30] S. K. Nayar and S. G. Narasimhan. Vision in bad weather. In *CVPR*, 1999.
- [31] A. K. Pediredla, M. Buttafava, A. Tosi, O. Cossairt, and A. Veeraraghavan. Reconstructing rooms using photon echoes: A plane based model and reconstruction algorithm for looking around the corner. In *ICCP*, 2017.
- [32] G. Satat, B. Heshmat, D. Raviv, and R. Raskar. All photons imaging through volumetric scattering. *Sci. Rep.*, 6, 2016.
- [33] G. Satat, M. Tancik, O. Gupta, B. Heshmat, and R. Raskar. Object classification through scattering media with deep learning on time resolved measurement. *Opt. Exp.*, 25, 2017.
- [34] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar. Instant dehazing of images using polarization. In *CVPR*, 2001.
- [35] A. Schwartzman, M. Alterman, R. Zamir, and Y. Y. Schechner. Turbulence-induced 2d correlated image distortion. In *ICCP*, 2017.
- [36] M. Sheinin and Y. Y. Schechner. The next best underwater view. In *CVPR*, 2016.
- [37] D. Shin, F. Xu, D. Venkatraman, R. Lussana, F. Villa, F. Zappa, V. K. Goyal, F. N. Wong, and J. H. Shapiro. Photon-efficient imaging with a single-photon camera. *Nat. Comm.*, 7, 2016.
- [38] O. Spier, T. Treibitz, and G. Gilboa. In situ target-less calibration of turbid media. *ICCP*, 2017.
- [39] J. Tian, Z. Murez, T. Cui, Z. Zhang, D. Kriegman, and R. Ramamoorthi. Depth and image restoration from light field in a scattering medium. In *ICCV*, 2017.
- [40] T. Treibitz and Y. Y. Schechner. Instant 3descatter. In *CVPR*, 2006.
- [41] T. Treibitz and Y. Y. Schechner. Active polarization descattering. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31, 2009.
- [42] C. Y. Tsai, K. N. Kutulakos, S. G. Narasimhan, and A. C. Sankaranarayanan. The geometry of first-returning photons for non-line-of-sight imaging. In *CVPR*, 2017.
- [43] 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. *Nat. Comm.*, 3, 2012.