Analysis of the C-NLOS Approach

Caio da Silva, Daniel Sawyer, Hunter Morera, Palak Dave
University of South Florida
Computer Science and Engineering
caiodasilva@usf.edu, danielsawyer@usf.edu, hmorera@usf.edu, palakdave@usf.edu

Abstract

Non-Line-Of-Sight (NLOS) approaches reconstruct the hidden object's shape and albedo from indirect reflection, but the current approaches are expensive in terms of memory requirements, computational complexity, and weak signal of multiply scattered light. Confocal NLOS (C-NLOS) imaging method was designed to overcome those shortcomings, illuminating and imaging the same point of the scene. This report performs an analysis on the C-NLOS method, including the implementation and reproducing the results. It was possible to achieve the same results, and we also implemented a pre- and post-processing Gaussian filter to reduce the noise and increase the quality of the reconstructed images.

1. Introduction

Finding a way to image an object not visible to the camera is gaining more and more research interest in current time. Such Non-Line-Of-Sight imaging capability have potential applications in the fields of autonomous driving, remote sensing, medical imaging, defence etc. NLOS approaches reconstruct the hidden object's shape and albedo from indirect reflection (multiply scattered light) in-contrast to the LIDAR systems which reconstructs shape of visible objects by measuring time of direct reflection of incident light. The current approaches of NLOS is expensive in terms of memory requirements, computational complexity, and extremely weak signal of multiply scattered light. Moreover, current approaches needs a custom hardware system equipped with costly components. All these limitations of the current approaches make them impractical for widespread use.

This report is about analysis of a Confocal NLOS (C-NLOS) imaging method [2] which is designed to overcome the shortcomings of the state-of-the-art methods as described in the previous paragraph. A major difference between C-NLOS and previous method is that the later exhaustively illuminate and measure a pair of distinct points on a visible surface, whereas the C-NLOS method reviewed here illuminates and images the same point. The visible surface is raster-scanned to obtain a time-resolved image.

The major contributions of the C-NLOS method studied here are, 1. It offers a closed-form solution for the NLOS problem which leads to a faster, more memory efficient, and high quality reconstruction. 2. C-NLOS method is capable of exploiting the retroreflectance property of retroreflective objects. The retroreflectance property of an object allows it to reflect large portion of the incident light back to its source with minimal scattering. This property can be exploited by simultaneous illuminating and imaging of a common point which is done by C-NLOS. 3. Existing LIDAR systems, which work on confocal scanning in direct light path, may be capable of supporting C-NLOS algorithm with minimal hardware change 4. It works reasonablly well outdoors, underindirect sunlight.

The assumptions in the reviewed method are 1. there is no inter-reflection behind the wall (that is only single scattering) 2. light scattering is isotropic 3. no occlusions within the hidden scene.

The remaining report is organized as follows: The hardware setup for the experiments is described in section 2. That is followed by investigating the forward model and the inverse problem formulation in the given method in sections 3 and 4, respectively. Next, section 5 is a discussion about the results obtained by recreating the original paper and a few modifications tried to improve the reconstruction. Finally, the study is concluded in section 6.

2. Experimental Setup

The prototype system for the experiments were created from the ground up bu the authors. However, it might be possible for the exsiting LIDAR systems to support the reviewed method by a minimal hardware change, as mentioned in the previous section.

A photograph of the experimental setup is presented in Fig. 1. The hidden object 'S' is the object of interest for reconstruction. The black cloth is acting as an occluder and the camera system is depicted in detail in Fig. 2.

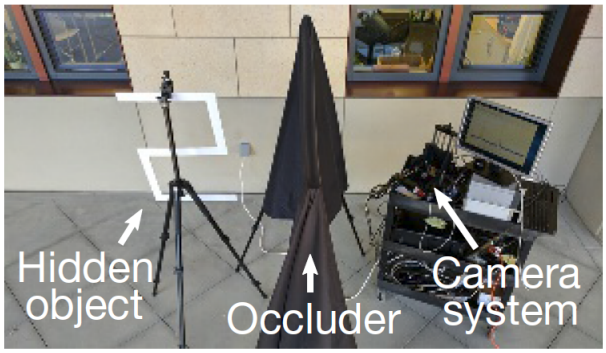

Figure 1. A picture of the experimental setup. Hidden 'S' is the object of interest, black cloth is an occluder, and the camera system is depicted in detail in 2.

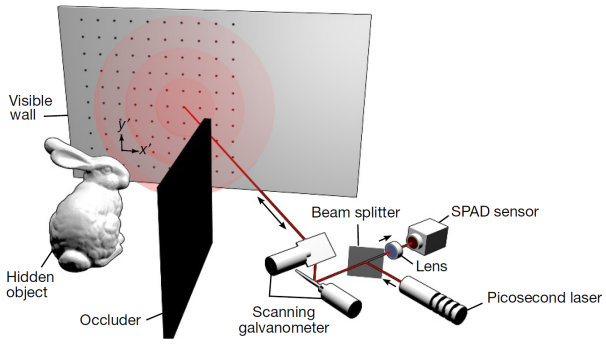

Figure 2. An illustration of the prototype camera system used. A pulsed laser and a time-resolved detector is used to raster-scan a visible wall to collect measurements of both the direct light reflecting off the wall and indirect light from a hidden object.

In the camera system (Fig. 2), a pulsed laser and a time-resolved detector is utilized to raster scan the visible wall in order to record measurements of direct reflection from the wall and indirect reflection from the hidden object (which is a bunny model).

3. Forward Model

Typical NLOS systems require very exepensive custom hardware, reconstruction is prohibitively computationally demanding in both memory and cpu, and the flux of multiply scattered light is extremely low. Confocal NLOS (C-NLOS) imaging overcomes a lot of these problems. It does this by using raster-scans to acquire a 3D transient and time resolved image instead of an exhaustive one. It facilitates the derivation of a closed-form solutionto the NLOS problem. Taking the measurements in this way, the reconstruction is orders of magnitude faster and much better on memory than previous NLOS systems with much higher resolution. Next it uses retroreflectance which can only be exploited by confocal systems that simultaneously illuminate and image a common point. This system was built from the ground up and differs slightly from LIDAR but LIDAR systems do have the ability to do this will little hardware modifications.

This systems image formation model makes a few assumptions. There is only a single scattering behind the wall so that there are no inter-reflections, the light scatters isotropically, and no occlusions occur within the hidden object. C-NLOS measurements consist of a 2D set of temporal histograms that are obtained by confocal scanning of points x' and y' on a plane, which in this case is a wall, at position z'. This generates a 3D volume of measurements over a temporal windows intervaled in picoseconds. The following equation shows how these measurements are obtained and calculated before reconstruction.

$$\tau(x',y',t) = \iiint_{\Omega} \frac{1}{r^4}\,\rho(x,y,z)\,\delta(2\sqrt{(x'-x)^2+(y'-y)^2+z^2}-tc)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z \quad [2]$$

In this formula, $\tau(x',y',t)$ is the photon flux and time t captured. The $p(x,y,z)$ function gives the albedo (whiteness) of the hidden object at the point (x,y,z) in 3D half-space Ω. The Dirac delta functions represents the surface of the 4D hypercone which allows them to model light propagation from the wall to the object and back. It is closely related to Minkowski's light cone. Also, c is obviously the speed of light. It is also shift-invariant on the x and y axes but not on the depth z. This can then be easily transformed into an equation that can be expressed as 3D convolution that is also shift-invariant.

$$\underbrace{v^{3/2}\tau(x',y',2\sqrt{v}/c)}_{\mathcal{R}_t\{\tau\}\,(x',y',v)} = \iiint_{\Omega} \underbrace{\frac{1}{2\sqrt{u}}\,\rho(x,y,\sqrt{u})}_{\mathcal{R}_z\{\rho\}\,(x,y,u)}\,\underbrace{\delta((x'-x)^2+(y'-y)^2+u-v)}_{h(x'-x,y'-y,v-u)}\,\mathrm{d}x\mathrm{d}y\mathrm{d}u \quad [2]$$

So basically, the forward problem in this paper is data/measurement collection but the more involved part of the algorithm and where most of its novelty come from are in the inverse problem in the next section and when things really start to get interesting.

4. Inverse Problem

From the previous mathematical formula, you can see that the function can be expressed as a 3D convolution $R_t(\tau) = h * R_z(p)$ where the function h is the shift-invariant 3D convolutional kernel or mask and the transform R_z non-uniformly re-samples and attenuates the elements of the volume collected p along the z axis. The R_t, hopefully obviously, re-samples the measurements along τ as usual, along the time axis which makes this a good fit for the algorithm due to its temporal nature in measurement collection. The inverse of R_t and R_z have a closed-form expressions that they refer to as light-cone transform (LCT). It is then discretized with a vectorized representation of the

measurements and volume of the albedos from the hidden object at twice the distance from the scanned point. It uses FWHM, full width at half maximum, and scanning a sequence of points along the diffuse reflector, or wall, produces a streak image that captures the spatial and temporal geometry of that indirect light. This can then be computed efficiently in the Fourier domain and represents the discrete LCT.

$$\boldsymbol{\rho}_* = R_z^{-1} F^{-1} \left(\frac{1}{\widehat{H}} \frac{|\widehat{H}|^2}{|\widehat{H}|^2 + \frac{1}{\alpha}} \right) F R_t \boldsymbol{\tau} \quad [2]$$

By treating this problem as a spatially invariant 3D deconvolution problem, the convolution operation can be expressed as an element-wise multiplication in the Fourier domain and inverted according to the equation above. This equation is, basically, the Tikhonov regularization we learned about in class and helps to regularize the output. Again, p is the estimated volume of the albedos. H is a diagonal matrix containing the Fourier coefficients of the 3D convolutional kernel, α is the frequency dependent signal to noise ratio, and is based off Wiener filtering and slightly different than Tikhonov. As α reaches infinity, it becomes an inverse filter. They also derive an iterative reconstruction procedure that combines the LCT with a Poisson noise model.

All these methods allow for a much more computationally efficient algorithm that could be optimized to run in real-time one day and could actually be used on practical applications like remote sensing, robitics, and autonomous vehicles. It also has the ability to work outside in indirect sunlight, which would be required for those practical applications. The fundamental bounds for the resolution can be seen below.

$$\Delta z \geq \frac{c\gamma}{2} \quad \text{and} \quad \Delta x \geq \frac{c\sqrt{w^2 + z^2}}{2w}\gamma$$

5. Results

All C-NLOS images were taken of objects with reflective tape or paint. This allows more light from the projected laser to be reflected back to the diffuser wall for measurement. In our experiments we only used data provided by the original authors and did not attempt to create any synthetic data. We used the authors code as a template to write our own version in Matlab. Our first goal was to achieve the baseline performance as shown in the paper and then add additional post processing to increase the image quality.

5.1. Original Implementation

We began with the code as a template to write our own version to better understand the implementation. We were able to achieve results identical to the original paper. Due to the fact that the measurements are taken over a third dimension of time, it is possible to recover an estimate of how the object looks at the front, side, and top viewing angles. Our version of the original implementation was shown to be identical to the authors implementation using a subtraction of resulting images. As an empirical example please see the figure 3 below which compares our output to the output of the code linked in the paper. The original output is on the left and our output is on the right. It is clear that our original implementation exactly reproduces the authors results. These results are very shocking in their quality considering these objects were completely occluded from the sight of the sensor. It is very clear that the letters can be read and the orientation of the letters were captured to a surprising level of accuracy. After we implemented this we began to consider if the results could be improved with some filtering to reduce the aliasing present in the final images.

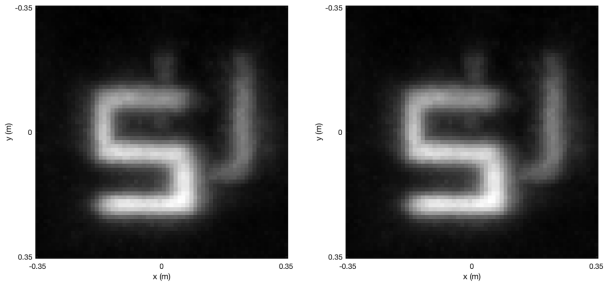

Figure 3. Original output vs our code output

5.2. Modifications

The original results of the paper are very impressive given the setup, however, we were interested in improving upon this using different pre and post processing techniques. These images suffer from obvious aliasing due to the resolution of the device. We experimented with applying noise reducing filters to both the raw measurements, as well as the inversed images. These techniques could be used to increase the quality of the final output and aid in further processing such as segmentation. The most obvious technique to apply was Guassian filtering on the final inversed image to reduce noise and provide an overall smoother image, as seen on [1]. We applied a 2D Guassian filter with varying values for σ, we found a value of 0.75 to provide the most visually appealing results. You can see the comparison between the original authors output and the post processed image in figure 4 below. The raw image is on the left and the post processed image is on the right. As expected it cleaned

up the noise in the image and provides a smoother image.

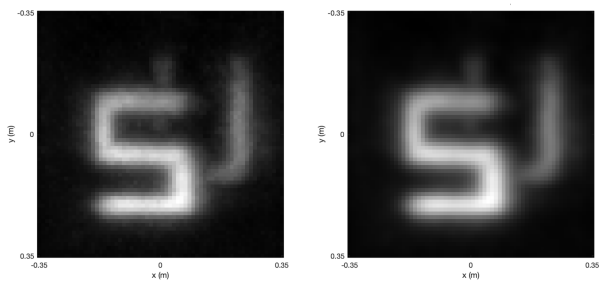

Figure 4. Original output vs Guassian post-processing

The post processing did perform well but we began to wonder if smoothing the measurements prior to inversing would provide a similar or better result. Because the measurements are taken out in the natural sunlight we supposed there may be some high frequency components that skew the results. We decided to apply the same Guassian smoothing to the measurements as a pre-processing step. These results are very similar to the post processed results. In figure 5 you can see the original image on the left and the pre-processed version on the right. We used the same sigma value as in post processing.

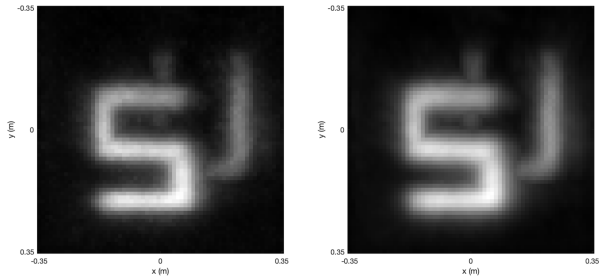

Figure 5. Original output vs Guassian pre-processing

Both pre and post processing do enhance the results over the raw output. This processing would increase the ability to perform segmentation which is very important for practical uses of this type of imaging in applications like driverless vehicles.

6. Conclusion

This project implementation was capable of reproducing the same results of the original paper, following the conclusion that the proposed technique enables NLOS imaging using the specified hardware with a high speed, under ambient lighting and at higher resolution than other existing approaches during the paper's publication.

Alongside the understanding of the forward and inverse problem of this method, and it's correct implementation in MATLAB, it was possible for us to observe a gap and implement Gaussian filters to solve the noise and quality level presented in the results. This represents a clear understanding of the concepts learned in the class, theirs functionality, how to implement and how to combine them.

As future works, we could modify this approach using Deep Learning methods to improve accuracy and resolution. Using a combination of Temporal Convolutional Networks (TCN) along with Generative Adversarial Networks (GAN) we should be able to reconstruct the occluded objects with better accuracy and higher resolution. Another possibility would be to apply this method to objects with different materials and colors.

References

[1] M. Bertero and P. Boccacci. *Introduction to inverse problems in imaging*. CRC press, 2020.

[2] L. D. . W. G. O'Toole, M. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature555*, pages 338—-341, 2018.