

# NLOS-NeuS: 非視線方向撮影における ニューラル陰関数表面

藤村 友貴<sup>1,a)</sup> 櫛田 貴弘<sup>1,b)</sup> 船富 卓哉<sup>1,c)</sup> 向川 康博<sup>1,d)</sup>

## 概要

非視線方向撮影 (Non-line-of-sight (NLOS) imaging) とは、カメラから見えないシーンを知覚する技術である。本研究では高時間分解データを入力とし、ニューラル陰関数表面による NLOS シーンの 3 次元形状復元手法を提案する。ニューラル陰関数表面として符号付き距離関数 (Signed distance function (SDF)) を多層パーセプトロンで表現し、NLOS シーンにおいて SDF を正しく学習するための制約を導入する。ボクセル表現を用いた従来手法に比べて、詳細かつ滑らかな形状復元が可能であることを示す。

## 1. はじめに

近年コンピュータビジョンの分野では、多くのアプリケーションが社会実装されている。しかしながら、これらの多くはセンサからシーンが見えていることを前提としている。これに対し、非視線方向撮影 (Non-line-of-sight (NLOS) imaging) [6], [18] とは、センサから見えないシーンを再構成する技術である。本研究は、NLOS シーンにおける 3 次元形状復元を目的としたものである [4], [14], [16], [17], [20], [23].

一般的な NLOS 撮影の設定を図 1(a) に示す。光源とセンサは同軸に配置されているが、センサからは壁のみが見えており NLOS シーンは見えない。光源から照射されたパルス光は壁での反射を経由し、NLOS シーンに到達する。NLOS シーンで反射した光は再び壁を経由し、最終的にセンサで観測される。NLOS 撮影で用いるセンサの一つに Single photon avalanche diode (SPAD) がある (図 1(b) 上)。SPAD と高速な時間計測が可能な Time-to-digital converter を組み合わせたシステムでは、ピコ秒単位の光伝播を計測することができる。図 1(b) 下は SPAD で計測された高時間分解データであり、横軸は時刻、縦軸は各時刻に SPAD に到達した光子の数を表す。本研究は、見えている壁で計測された高時間分解データを入力とし、

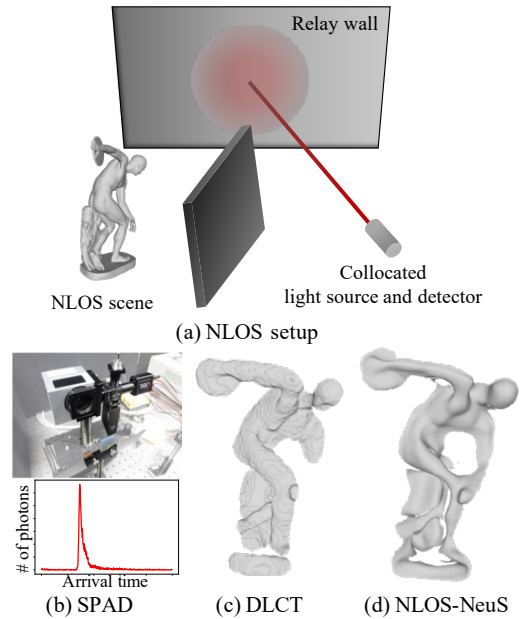


図 1 (a) 一般的な NLOS 撮影の設定, (b) SPAD で計測された高時間分解データ, (c) DLCT [23] による 3 次元形状復元, (d) 提案手法.

NLOS シーンの 3 次元形状を復元する。

近年, Neural radiance field (NeRF) [10] をはじめとした多層パーセプトロン (Multi-layer perceptron (MLP)) によるシーン表現が、従来の離散的な表現に対しコンパクトかつ詳細なシーン表現として成果をあげている。NLOS 撮影においても、MLP による radiance field を用いた Neural transient field (NeTF) [15] が提案されている。本研究では、符号付き距離関数 (Signed distance function (SDF)) によるニューラル陰関数表面 (Neural implicit surface (NeuS) [19]) を用いて NeTF を拡張し、NLOS シーンでの詳細な 3 次元復元を実現する。SDF は物体表面への最小距離を表し、符号が正ならば物体外、負ならば物体内部を表す。したがって、符号付き距離が 0 の点のみを抽出することで、物体表面を復元できる。図 1(c) と (d) に、従来のボクセル表現による 3 次元復元 [23] と提案手法による NLOS シーンの 3 次元復元の結果を示す。ニューラル陰関数表面を用いることで、より詳細な形状の復元が可能となる。

<sup>1</sup> 奈良先端科学技術大学院大学

a) fujimura.yuki@is.naist.jp

b) kushida.takahiro.kh3@is.naist.jp

c) funatomi@is.naist.jp

d) mukaigawa@is.naist.jp

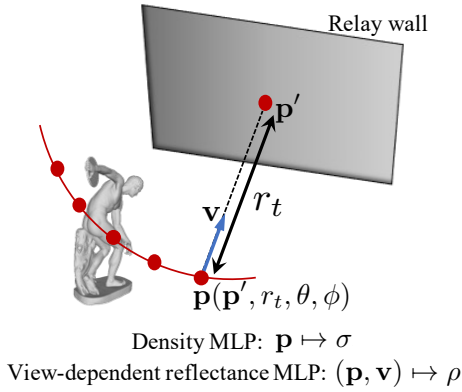


図 2 NeTF [15] における球面上の点のサンプリング

## 2. 提案手法

### 2.1 問題設定

本研究の目的は壁で計測された高時間分解データから、2つの MLP,  $d: \mathbb{R}^3 \rightarrow \mathbb{R}$  と  $\rho: \mathbb{R}^6 \rightarrow \mathbb{R}$  を学習することである。  $d$  は SDF を MLP として表現したものであり、入力 は NLOS シーン中の点  $\mathbf{p} = (x, y, z) \in \mathbb{R}^3$ , 出力はその点における符号付き距離である。  $\rho$  の入力 は NLOS シーン中の点  $\mathbf{p}$  と、その点から壁上のある一点  $\mathbf{p}' = (x', y', z') \in \mathbb{R}^3$  への方向  $\mathbf{v}$  であり、出力は  $\mathbf{p}$  における方向に依存した反射率<sup>\*1</sup>である。計測した壁上の高時間分解データは  $\tau(\mathbf{p}', t) \in \mathbb{R}^B$  で表し、ここで  $B$  は時刻のビン数である。例えば SPAD による計測の場合、 $\tau(\mathbf{p}', t)$  は時刻  $t$  に到達したフォトン の数を表す。

### 2.2 NeTF [15]

最初に、本研究の元となる NeTF [15] について説明を行う。NeTF でも 2 つの MLP の学習を行う。1 つは  $\sigma: \mathbb{R}^3 \rightarrow \mathbb{R}$  であり、NLOS シーン中の点  $(x, y, z)$  を入力とし、その点における密度を出力する。もう一つは  $\rho$  であり、これは提案手法と同様 NLOS シーン中の各点における方向に依存する反射率を出力する。カメラからの視線方向上で点をサンプルする NeRF [10] とは異なり、NeTF では壁上の一点を中心とした球面上で点をサンプルする (図 2)。

$$\mathbf{p}(\mathbf{p}', r_t, \theta, \phi) = \begin{bmatrix} r_t \sin \theta \cos \phi + x' \\ r_t \sin \theta \sin \phi + y' \\ r_t \cos \theta + z' \end{bmatrix} \quad (1)$$

ここで、 $\theta$  と  $\phi$  は仰角と方位角であり、 $r_t$  はビン  $t$  に対応する半径である。例えば、壁上の点において  $t = 0$  ならば、 $c$  を光速として  $r_t = ct/2$  である。壁上の点  $\mathbf{p}'$  での時刻  $t$  における観測はこの球面上での積分で計算できる。

$$\tau(\mathbf{p}', t) = \iint A(r_t, \theta) T(\mathbf{p}, \mathbf{v}) \sigma(\mathbf{p}) \rho(\mathbf{p}, \mathbf{v}) d\theta d\phi \quad (2)$$

<sup>\*1</sup> 物体表面以外については、式 (6) の  $w$  が小さくなるように学習することで関与しなくなる。

ここで、 $A(r_t, \theta) = \sin \theta / r_t^2$  は減衰係数であり、 $T(\mathbf{p}, \mathbf{v})$  は  $\mathbf{p}$  と  $\mathbf{p}'$  の間の透過率である。なお、 $\mathbf{p}$  の  $\mathbf{v}$  の引数は省略してある。(つまり、 $\mathbf{p}(\mathbf{p}', r_t, \theta, \phi)$  と  $\mathbf{v}(\mathbf{p}(\mathbf{p}', r_t, \theta, \phi), \mathbf{p}')$ .)

$\sigma$  と  $\rho$  のパラメータは、計測したデータと式 (2) の間の誤差  $\mathcal{L}_\tau$  を最小化することで最適化する。

$$\mathcal{L}_\tau = \frac{1}{MB} \sum_{\mathbf{p}', t} (\tau_m(\mathbf{p}', t) - \tau(\mathbf{p}', t))^2 \quad (3)$$

ここで、 $\tau_m$  は計測したデータであり、 $M$  は計測した壁上の点の数である。

### 2.3 NLOS-NeuS

提案手法では、まず最初に MLP  $d$  に  $\mathbf{p}$  を入力しその点における符号付き距離を出力する。その後、 $\rho$  に  $\mathbf{p}$  と  $\mathbf{v}$  を入力し、方向に依存する反射率を出力する。式 (2) を用いるため、出力された符号付き距離は以下の式に従って密度に変換する [13]。

$$\sigma(\mathbf{p}) = \frac{1}{\alpha} \text{Sigmoid}\left(-\frac{d(\mathbf{p})}{\alpha}\right) \quad (4)$$

ここで、 $\text{Sigmoid}(\cdot)$  はシグモイド関数であり、 $\alpha > 0$  は学習パラメータである。 $\alpha$  が 0 に十分近い場合は、符号付き距離が 0 のときのみ密度が値を持つため、完全に不透明な物体表面となる。密度が得られたあとは、式 (2) を用いて  $\tau(\mathbf{p}', t)$  を計算する。なお、実装では以下の離散化したものを用いる [9], [10], [19]。

$$\tau(\mathbf{p}', t) = \sum_{\theta, \phi} A(r_t, \theta) w(\mathbf{p}, \mathbf{v}) \rho(\mathbf{p}, \mathbf{v}) \Delta\theta \Delta\phi \quad (5)$$

$$w(\mathbf{p}, \mathbf{v}) = \sum_{s=t_{\min}}^t T_s \left(1 - \exp\{-\sigma(\mathbf{p}(\mathbf{p}', r_s, \theta, \phi)) \Delta r_s\}\right) \quad (6)$$

$$T_s = \exp\left\{-\sum_{u=t_{\min}}^{s-1} \sigma(\mathbf{p}(\mathbf{p}', r_u, \theta, \phi)) \Delta r_u\right\} \quad (7)$$

ここで、 $w(\mathbf{p}, \mathbf{v})$  は壁からの視線方向上における反射率の重みである。

### 2.4 NLOS における SDF の学習の制約

多視点で撮影された画像から SDF を学習する手法は多く提案されているが [11], [12], [19], [21], [22], このような多視点の観測とは異なり、NLOS の問題設定では物体の全周を観測できないため、SDF の学習が難しい。図 3(a) に示すように、壁側からしか観測できないため物体の内外の学習に失敗し、符号付き距離が 0 ではない点で重み (式 (6) の  $w$ ) が最大となってしまうことがある。図 3(b) は合成データで実際に学習した SDF であり、緑線で囲まれた領域で SDF の学習に失敗している。

本研究では正しく SDF を学習するため、以下の二つの制約を満たすような誤差関数を導入する。(1) 物体表面で

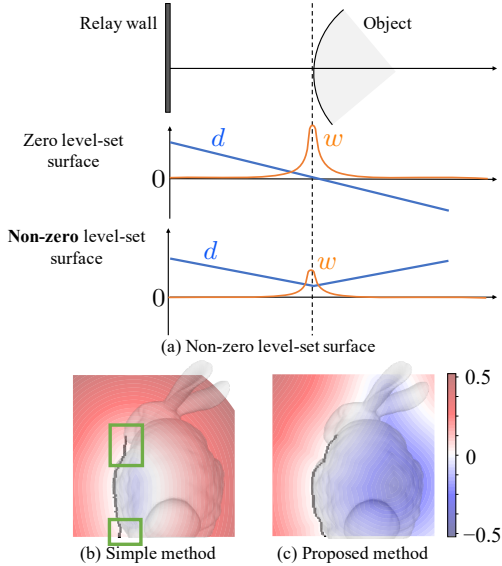


図 3 (a) NLOS シーンでは壁側からしか物体を観測できないため物体の内外の学習に失敗し、SDF が 0 ではない点において重みが最大になってしまうことがある。(b) 壁からの視線方向上で重みが最大である点から推定した物体表面 (黒線) と、SDF が 0 である位置が一致していない。(c) 重みが最大である位置と SDF が 0 である位置が一致するように学習を行う。

は SDF が 0 となる。(2) SDF が 0 ではない点は物体表面に寄与しない。

### 2.4.1 物体表面で SDF を 0 とする誤差関数

SDF を学習する際、物体表面である点を推定し、その点において SDF が 0 となるような誤差関数を導入する。式 (5) の離散化した積分を行う各球面上において、 $w(\mathbf{p}, \mathbf{v})\rho(\mathbf{p}, \mathbf{v})$  を正規化して確率密度関数とし、各球面上で  $N_z$  個の点をサンプリングする。これらの点は式 (5) の積分に物体表面として寄与している確率が高いため、これらの点において SDF が 0 となるように以下の誤差関数を導入する。

$$\mathcal{L}_z = \frac{1}{MBN_z} \sum_{t, \mathbf{p}'} \sum_{n=1}^{N_z} m(\mathbf{p}', t) |d(\mathbf{p}(t, \theta_n, \phi_n))| \quad (8)$$

ここで、 $m(\mathbf{p}', t) \in \{0, 1\}$  は点  $\mathbf{p}'$  を中心とした半径  $r_t$  の球面上に物体が存在する場合は 1、そうでない場合は 0 の値をとる。この  $m$  は観測した高時間分解データ  $\tau_m$  から単純な閾値処理によって計算できる。

### 2.4.2 重み $w$ に対する誤差関数

2.3 で述べたように、SDF を密度に変換する関数 (式 (4)) において  $\alpha \rightarrow 0$  は完全な不透明物体を表す。したがって、制約 (2) を満たすためには、学習過程において  $\alpha$  を正しく 0 に収束させる必要がある。本研究では以下の誤差関数を導入することによりこれを実現する。

$$\mathcal{L}_{en} = \frac{1}{M|\theta||\phi|} \sum_{\mathbf{p}', \theta, \phi} -\hat{\delta} \log_2 \hat{\delta} - (1 - \hat{\delta}) \log_2 (1 - \hat{\delta}) \quad (9)$$

ここで、 $\hat{\delta} = \sum_{t=t_{min}}^{t_{max}} w(\mathbf{p}, \mathbf{v})$  は壁からの視線方向上での

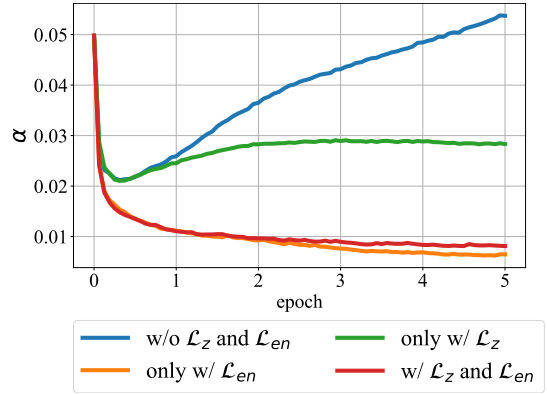


図 4 学習中の  $\alpha$  の推移

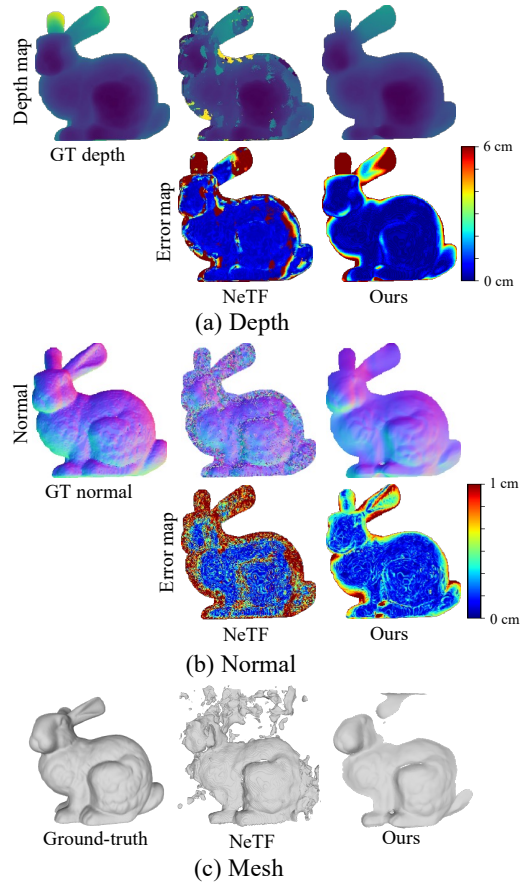


図 5 NeTF [15] と提案手法による合成データ [1] での比較。(a) 深度、(b) 法線、(c) メッシュの復元結果。法線は End-point-error にて誤差を計算。

重みの和である。この誤差関数を用いることにより  $\alpha$  を 0 に収束させることができる。直感的には物体が存在しない点において重み  $w$  を 0 とすることでこれを実現する。数学的な証明は追加資料に記載する。

図 4 に各誤差関数を用いた場合の学習中の  $\alpha$  の推移を示す。 $\mathcal{L}_{en}$  が  $\alpha$  の 0 への収束に寄与することがわかる。

## 2.5 誤差関数

まとめると、SDF の学習に用いる誤差関数は以下となる。

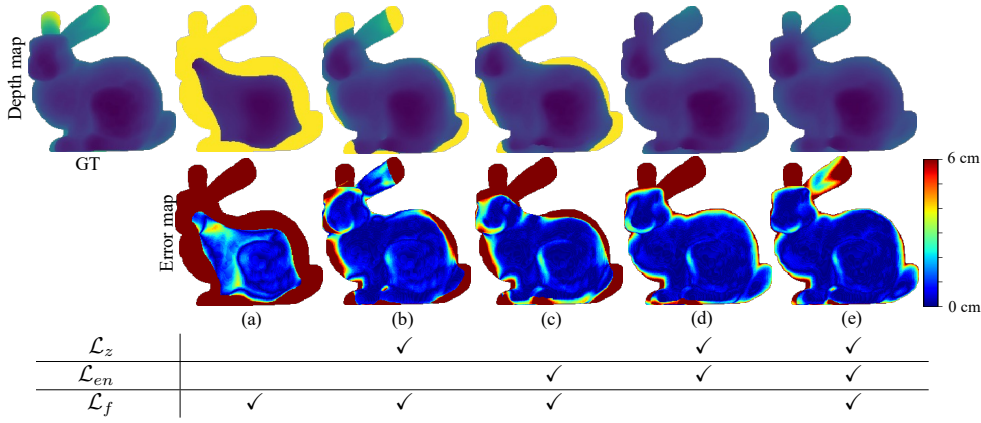


図 6 誤差関数に関する Ablation study

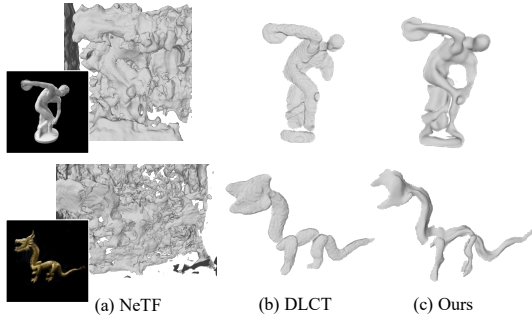


図 7 SPAD で計測した実データ [8] でのメッシュ復元. (a) NeTF [15], (b) DLCT [23], (c) 提案手法.

$$\mathcal{L} = \lambda_\tau \mathcal{L}_\tau + \lambda_{ei} \mathcal{L}_{ei} + \lambda_z \mathcal{L}_z + \lambda_{en} \mathcal{L}_{en} + \lambda_f \mathcal{L}_f \quad (10)$$

$\mathcal{L}_\tau$ ,  $\mathcal{L}_z$ ,  $\mathcal{L}_{en}$  はここまでで述べた誤差関数であり,  $\lambda_*$  は各誤差関数に対する重みである.  $\mathcal{L}_{ei}$  は Eikonal loss [2] であり, SDF の空間的な微分のノルムを 1 とする制約である.

$$\mathcal{L}_{ei} = \frac{1}{|\Omega_{nlos}|} \sum_{\mathbf{p} \in \Omega_{nlos}} (\|\nabla d(\mathbf{p})\| - 1)^2, \quad (11)$$

ここで,  $\Omega_{nlos}$  は対象とする NLOS の空間である.

本研究ではさらに, 高時間分解データを用いた space carving [16] によってあらかじめ大まかな形状と SDF を計算し, これを SDF の真値の下限 [7] として以下の誤差関数を導入する.

$$\mathcal{L}_f = \frac{1}{|\Omega_{free}|} \sum_{\mathbf{p} \in \Omega_{free}} \max(0, b(\mathbf{p}) - d(\mathbf{p})) \quad (12)$$

ここで,  $\Omega_{free}$  は space carving で得られた物体が存在しない空間であり,  $b(\mathbf{p})$  は  $\mathbf{p}$  における SDF の下限である.

### 3. 実験

本研究では高時間分解データとして, 合成データである ZNLOS データセット [1] と SPAD で計測された実データである f-k データセット [8] を用いた. MLP は NeuS [19] と同じものを用いた. 推定した SDF に対して Sphere tracing [3] と Poisson surface reconstruction [5] を用いて 3 次元メッシュを復元した.

### 3.1 合成データ [1] での実験結果

図 5 に, NeTF [15] と提案手法の合成データにおける深度, 法線, メッシュの復元結果を示す. Radiance field としてシーンを表現する NeTF は復元した形状に大きな誤差が含まれる. 一方で, 提案手法は SDF を用いて物体表面をモデル化しているため, 滑らかな形状復元が可能である.

図 6 に誤差関数に関する ablation study を示す.  $\mathcal{L}_z$  と  $\mathcal{L}_{en}$  を用いない場合はほとんどの箇所で形状復元に失敗する (図 (a)). また, それぞれを用いると大きく精度が向上するものの, 物体境界が失われてしまっている (図 (b)(c)). これらに対し, 両方の誤差関数を用いることで形状復元が可能となる (図 (e)). 加えて, SDF の下限を用いた  $\mathcal{L}_f$  によって全体的に精度が向上することも確認できる (図 (d)(e)).

### 3.2 SPAD で計測した実データ [8] での実験結果

図 7 に SPAD で計測した実データでのメッシュの復元結果を示す. 合成データでの実験と同様, NeTF [15] は形状復元に適していない. NLOS シーンでの形状復元手法である DLCT [23] では, 形状表現として離散的なボクセルを用いているため細部の復元が難しい. これに対し, SDF を用いた提案手法では, 滑らかかつ詳細な形状復元が可能である.

## 4. まとめ

本研究ではニューラル陰関数表面を用いて, NLOS シーンにおける 3 次元形状復元手法を提案した. 離散的なボクセル表現による従来手法と比較し, 滑らかかつ詳細な形状復元を可能とした. 今後は非同軸系への拡張や, 他のニューラル形状表現を用いた定式化について研究を進める予定である.

### 謝辞

本研究は JSPS 科研費 21K21317, 20K20629 の助成を受けたものである.



## 参考文献

- [1] Galindo, M., Marco, J., O’Toole, M., Wetzstein, G., Gutierrez, D. and Jarabo, A.: A dataset for benchmarking time-resolved non-line-of-sight imaging, *ICCP* (2019).
- [2] Gropp, A., Yariv, L., Haim, N., Atzmon, M. and Lipman, Y.: Implicit Geometric Regularization for Learning Shapes, *ICML*, pp. 3569–3579 (2020).
- [3] Hart, J. C.: Sphere tracing: a geometric method for the antialiased ray tracing of implicit surfaces, *The Visual Computer*, Vol. 12, No. 10, pp. 527–545 (1996).
- [4] Iseringhausen, J. and Hullin, M. B.: Non-Line-of-Sight Reconstruction Using Efficient Transient Rendering, *ACM TOG*, Vol. 39, No. 1 (2020).
- [5] Kazhdan, M. and Hoppe, H.: Screened Poisson Surface Reconstruction, *ACM TOG*, Vol. 32, No. 3, p. 29 (2013).
- [6] Kirmani, A., Hutchison, T., Davis, J. and Raskar, R.: Looking Around the Corner using Ultrafast Transient Imaging, *IJCV*, Vol. 95, No. 1, pp. 13–28 (2011).
- [7] Lin, C.-H., Wang, C. and Lucey, S.: SDF-SRN: Learning Signed Distance 3D Object Reconstruction from Static Images, *NeurIPS* (2020).
- [8] Lindell, D. B., Wetzstein, G., Matthew O’ Toole: Wave-based non-line-of-sight imaging using fast f-k migration, *ACM TOG*, Vol. 38, No. 4, p. 116 (2019).
- [9] Max, N.: Optical models for direct volume rendering, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 1, No. 2, pp. 99–108 (1995).
- [10] Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R. and Ng, R.: NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, *ECCV* (2020).
- [11] Niemeyer, M., Mescheder, L., Oechsle, M. and Geiger, A.: Differentiable Volumetric Rendering: Learning Implicit 3D Representations Without 3D Supervision, *CVPR* (2020).
- [12] Oechsle, M., Peng, S. and Geiger, A.: UNISURF: Unifying Neural Implicit Surfaces and Radiance Fields for Multi-View Reconstruction, *ICCV*, pp. 5589–5599 (2021).
- [13] Or-El, R., Luo, X., Shan, M., Shechtman, E., Park, J. J. and Kemelmacher-Shlizerman, I.: StyleSDF: High-Resolution 3D-Consistent Image and Geometry Generation, *CVPR*, pp. 13503–13513 (2022).
- [14] Plack, M., Callenberg, C., Schneider, M. and Hullin, M. B.: Fast Differentiable Transient Rendering for Non-Line-of-Sight Reconstruction, *WACV*, pp. 3067–3076 (2023).
- [15] Shen, S., Wang, Z., Liu, P., Pan, Z., Li, R., Gao, T., Li, S. and Yu, J.: Non-line-of-Sight Imaging via Neural Transient Fields, *IEEE TPAMI*, Vol. 43, No. 7, pp. 2257–2268 (2021).
- [16] Tsai, C.-Y., Kutulakos, K. N., Narasimhan, S. G. and Sankaranarayanan, A. C.: The Geometry of First-Returning Photons for Non-Line-Of-Sight Imaging, *CVPR* (2017).
- [17] Tsai, C.-Y., Sankaranarayanan, A. C. and Gkioulekas, I.: Beyond Volumetric Albedo – A Surface Optimization Framework for Non-Line-Of-Sight Imaging, *CVPR* (2019).
- [18] Velten, A., Willwacher, T., Gupta, O., Veeraraghavan, A., Bawendi, M. G. and Raskar, R.: Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging, *Nature Communications*, Vol. 3, No. 1, p. 745 (2012).
- [19] Wang, P., Liu, L., Liu, Y., Theobalt, C., Komura, T. and Wang, W.: NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, *NeurIPS* (2021).
- [20] Xin, S., Nousias, S., Kutulakos, K. N., Sankaranarayanan, A. C., Narasimhan, S. G. and Gkioulekas, I.: A Theory of Fermat Paths for Non-Line-Of-Sight Shape Reconstruction, *CVPR* (2019).
- [21] Yariv, L., Gu, J., Kasten, Y. and Lipman, Y.: Volume rendering of neural implicit surfaces, *NeurIPS* (2021).
- [22] Yariv, L., Kasten, Y., Moran, D., Galun, M., Atzmon, M., Ronen, B. and Lipman, Y.: Multiview Neural Surface Reconstruction by Disentangling Geometry and Appearance, *NeurIPS* (2020).
- [23] Young, S. I., Lindell, D. B., Girod, B., Taubman, D. and Wetzstein, G.: Non-Line-of-Sight Surface Reconstruction Using the Directional Light-Cone Transform, *CVPR* (2020).