Neural Density Field 上での光子追跡による 高時間分解画像の解析手法

藤村 友貴^{1,a)} 櫛田 貴弘^{1,b)} 舩冨 卓哉^{1,c)} 向川 康博^{1,d)}

概要

本研究では高時間分解画像の新たな解析手法を提案す る.近年, neural radiance field (NeRF) に代表される多 層パーセプトロン (MLP) を用いたシーンの表現方法が多 く提案されている.本研究では neural density field で表 現されたシーンにおいて,光子一つ一つを追跡し, single photon avalanche diode (SPAD) カメラで計測された高時 間分解画像をレンダリングするように MLP を学習する手 法を提案する.これにより,光子一つ一つがどのような経 路で SPAD に到達したかを可視化することができる.シ ミューレションデータによる実験で,シーンの密度推定と 2次の相互反射の分離が可能であることを示す.

1. はじめに

カメラで撮影された画像から,光の反射といったシーン 中の光の物理現象を解析する研究が行われている.しかし ながら,2次元に縮退した通常の画像から相互反射といっ た複雑な物理現象を解析することは困難である.これに対 し, single photon avalanche diode (SPAD) による時間相 関単一光子計数 (TCSPC) を利用した研究が行われている [6],[7]. SPAD と高速な時間計測が可能な time-to-digital converter (TDC) 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.近々 SPAD カメラが 市場に普及することが予想されており,高時間分解画像の 新たな解析手法の検討は非常に重要である.

これに対し,近年深層学習の枠組みによる新たなシーン の表現方法が提案されている. Mildenhall ら [2] は, neural radiance field (NeRF) と呼ばれる多層パーセプトロン (MLP) によるシーンの表現方法を提案している. MLP が シーンの各点における密度と色を推定し,ボリュームレン ダリングによって観測画像を再現するように MLP を学習 する. 本研究は、高時間分解計測の深層学習を用いた新たな解 析手法を提案するものである.具体的には、NeRFのよう な MLP によるシーン表現を SPAD カメラ による高時間 分解計測へ拡張することを試みる.ボリュームレンダリ ングベースの従来手法 [4] とは異なり、フォトンマッピン グに近いアプローチを取る.MLP から出力される密度場 (density field)上で光子一つ一つを追跡し、SPAD カメラ の観測を擬似的に再現する.上記の過程を微分可能な形で 実装することで MLP の学習が可能になる.学習した結果 を用いることでシーンの密度が得られるとともに、光子一 つ一つが相互反射も含めてどのような経路で SPAD に到 達したかを可視化することができる.

NeRF ベースの相互反射のモデル化については,1次の 相互反射のみを考慮した従来手法 [5] が存在する.一方で, 光子を追跡する本研究の定式化では,相互反射の回数に よって計算量が指数関数的に増大することはなく,原理的 には任意のn次の相互反射をモデル化することが可能であ る.今回実験で用いたデータでは3次以上の相互反射の影 響は小さかったため,本研究では2次の相互反射までをモ デル化し,それらの分離と可視化を行なった.

2. 提案手法

2.1 SPAD カメラによる高時間分解計測

SPAD と TDC を用いた計測システムでは、光源から照 射された光子一つ一つが確率的に観測される. この観測を 複数回繰り返すことで、各時刻 $t \in \{0, \dots, T-1\}$ に到達し た光子の数を計測し、ビン数が T のヒストグラム $\mathbf{h} \in \mathbb{R}^T$ を取得できる. SPAD が一般的なカメラのような $H \times W$ の二次元配列である場合は、時間を加えた $H \times W \times T$ の 三次元配列が取得できる.

本研究では,確率的な光子の観測を繰り返すというこの プロセスを MLP が出力した密度場上でモデル化し,観測 を再現するように MLP を学習する.すべての過程を微分 可能な形で実装することで,MLP の学習が可能となる.

2.2 手法の概要

図1に本手法の概要を示す.なお、本研究では対象物体

¹ 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.ip

c) funatomi@is.naist.jp

 $^{^{\}rm d)} \quad mukaigawa@is.naist.jp$



図 1 手法の概要. MLP f_{θ} に三次元点 x を入力し密度 σ を求める. 密度場上で光子を追跡 して経路長を計算し,経路長に相当するビンの値が 1 である one-hot ベクトルを生成す る. この操作を繰り返し足し合わせて SPAD カメラの観測を再現し, f_{θ} の学習を行う.

は反射率が1の完全拡散反射面とし,2次の相互反射まで をモデル化する.

MLP f_{θ} : $\mathbb{R}^3 \to \mathbb{R}$ は、シーン中の三次元座標 $\mathbf{x} \in \mathbb{R}^3$ を 入力として、その点における密度 $\sigma(\mathbf{x}) \in \mathbb{R}$ を出力する. こ こで、 θ は MLP の学習パラメータである.本研究では計 算量の削減のため、シーンを $81 \times 81 \times 81$ のグリッド状に 分割してグリッドの各点を f_{θ} に入力し、 $\mathcal{V}_{\sigma} \in \mathbb{R}^{81 \times 81 \times 81}$ のボリュームを得る. 任意の点の座標、密度はバイリニア 補間により求める.

出力された \mathcal{V}_{σ} は $g: \mathbb{R}^{81 \times 81 \times 81} \to \mathbb{R}^{T}$ に入力され,ヒス トグラム h を得る.関数gは密度場上で光子を追跡するこ とにより SPAD カメラの観測を再現する. f_{θ} の学習のた め、g内のすべての処理は微分可能な形で実装する必要が ある.

点光源から照射された光子は,密度場上で確率的に 計算された点 \mathbf{p}_1 において反射する.その後, \mathbf{p}_1 にお いて次の反射方向をサンプリングし,その反射方向上 で次の反射点 \mathbf{p}_2 を計算する.同様の方法でさらに光子 を追跡し,3回目の反射点 \mathbf{p}_3 も計算する.そして,光 源 \mathbf{s} から SPAD カメラ \mathbf{c} までの経路長 $l_1 = l(\mathbf{s}, \mathbf{p}_1, \mathbf{c}),$ $l_2 = l(\mathbf{s}, \mathbf{p}_1, \mathbf{p}_2, \mathbf{c}), l_3 = l(\mathbf{s}, \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \mathbf{c})$ を計算する.こ こで, $l(\mathbf{x}_1, \dots, \mathbf{x}_N) = \sum_{i=1}^{N-1} ||\mathbf{x}_i - \mathbf{x}_{i+1}||$ である.これら の経路長から,各経路長に相当するビンの値が1である*T* 次元 one-hot ベクトル $\mathbf{v}(l_1), \mathbf{v}(l_2), \mathbf{v}(l_3)$ を微分可能な形 で生成する (式 (4)).

SPAD カメラと同様,上記のプロセスを複数回繰り返し て one-hot ベクトルを足し合わせることで,ヒストグラム h を生成することができる.最後に,生成したヒストグラ ムと実際に SPAD で観測したヒストグラム h* が一致する ように f_{θ} を学習する.

$$\theta^* = \arg\min_{\rho} \mathbb{E}_{(y,x)\sim H\times W} \left[\rho(\mathbf{h}_{y,x}, \mathbf{h}_{y,x}^*) \right]$$
(1)

ここで、 $\mathbf{h}_{y,x}$ 、 $\mathbf{h}_{y,x}^*$ はそれぞれピクセル (y,x)におけるヒストグラムであり、 $\rho(\cdot, \cdot)$ はヒストグラム間の誤差関数である.本研究ではヒストグラム間のスケールを調整したのち、それらの L1 誤差を計算する.以降、各ステップの詳

細について説明を行う.

2.3 密度場の生成

NeRF [2] と同様,ある三次元点 x における密度 $\sigma(\mathbf{x})$ は,MLP の出力 $\sigma = f_{\theta}(\mathbf{x})$ として与えられる.入力 x は positional encoding [2] により,24 次元の空間にあらかじ め埋め込む.その後,出力の次元数が 64 である全結合層 5 層ののち,最終層で1次元の σ を出力する.活性化関数に は最終層含めた全ての層で ReLU を用いる.

 f_{θ} から直接出力された $\sigma(\mathbf{x})$ に対して、物体表面の後ろ 側は視認できないことを考慮するため、ある任意の点 \mathbf{x}_0 から見た方向 d 上の点 $\mathbf{x}_i = \mathbf{x}_0 + i\mathbf{d}$ の密度は以下のよう に計算する.

$$\sigma_{\mathbf{x}_0,\mathbf{d}}(\mathbf{x}_i) = T_i \sigma(\mathbf{x}_i), \quad T_i = \exp\left(-\sum_{j=0}^i \sigma(\mathbf{x}_j)\right) \quad (2)$$

 T_i は一般的なボリュームレンダリング同様, \mathbf{x}_0 から \mathbf{x}_i の間の透過率を表す.

2.4 密度場からの反射点の推定

ある点 \mathbf{x}_0 から方向 \mathbf{d} へ光子が放射されたとき,式 (2) で求めた \mathbf{d} 上の密度をサンプリングし,それらから光子の 反射点 \mathbf{p} を推定する.いま,サンプリングされた密度が $\{\sigma_{\mathbf{x}_0,\mathbf{v}}(\mathbf{x}_i)\}_{i=1}^N$ であるとする.本研究では点 \mathbf{x}_i で光子が 反射する確率をソフトマックス関数を用いて以下のように 計算する.

$$q_i = \frac{\exp\left(\sigma_{\mathbf{x}_0,\mathbf{d}}(\mathbf{x}_i)\right)}{\sum_j \exp\left(\sigma_{\mathbf{x}_0,\mathbf{d}}(\mathbf{x}_j)\right)} \tag{3}$$

次の反射点はこれらの確率を用いて $\mathbf{p} = \sum_{i=1}^{N} q_i \mathbf{x}_i$ として 計算する.

しかしながら,決定的に反射点を計算する上記の方法で は,確率的な SPAD カメラの観測をモデル化することがで きない.そこで本研究では,光源から放射された最初の光子 の反射点を求める際,ソフトマックスの代わりに確率的に出 力を行う Gumbel-softmax [1] を用いる.Gumbel-softmax では,温度パラメータτを小さくするにつれて出力される



図2 法線ベクトルの推定.光源からの各方向上で反射点(黄点)を 平均し、物体の三次元点群を求める(青点).各平均点におい て、近傍の平均点との外積を計算して法線ベクトルを計算す る.各方向上の反射点の法線ベクトルは、それぞれの方向上の 平均点における法線ベクトルを用いる.

値が one-hot ベクトルに近くなる.本研究では $\tau = 0.1$ と して SPAD カメラ の確率的な観測を擬似的に再現する.

2.5 反射方向のサンプリング

本研究では完全拡散反射面を仮定しているため,次の反 射方向は反射点における法線ベクトルを中心とした半球面 上からサンプリングを行う.したがって,次の反射方向を 決定するためには,反射点における法線ベクトルを求める 必要がある.

本研究では図2に示す方法で法線ベクトルを求める. 3.2 で述べるように,光源からはある方向に対して複数の光子 を照射する.各光子の反射点は2.4 で説明したように確率 的に決定される(図中黄点).各方向でこれらの反射点を 平均し,物体の三次元点群を計算する(図中青点).そし て,三次元点群の各点において,その近傍の点を用いて外 積により法線ベクトルを求める.各反射点の法線ベクトル には,それぞれの方向上で求められた法線ベクトルを用い る.2回目以降の反射点における法線ベクトルは,三次元 点群の中から距離が最小の点を計算し,その点における法 線ベクトルを用いる.

2.6 経路長の one-hot ベクトルへの変換

本研究では 2 次の相互反射をモデル化するため,光源か ら最初の反射点 \mathbf{p}_1 を経由して SPAD に到達する経路の長 さ l_1 と,反射点 \mathbf{p}_1 , \mathbf{p}_2 ,及び反射点 \mathbf{p}_1 , \mathbf{p}_2 , \mathbf{p}_3 を経由 して SPAD に到達する経路の長さ l_2 , l_3 を計算する.そし て,SPAD の観測であるヒストグラムを生成するため,こ れらの経路長を one-hot ベクトル v に変換する.本研究で は f_{θ} の学習のため,このプロセスを以下の式に従って微 分可能な形で擬似的に実装する.

$$\mathbf{v}(l) = \frac{\tilde{\mathbf{v}}(l)}{\|\tilde{\mathbf{v}}(l)\|}, \quad \tilde{v}_i(l) = \exp\left(-\frac{(l-l_i)^2}{\gamma}\right) \tag{4}$$

ここで, l_i はヒストグラムのビンiに対応する経路長である. γ はハイパパラメータであり,本研究では $\gamma = 0.5$ とした.

最終的なヒストグラムhは、照射した全ての光子の経路



図3 対象のシーン



図 4 MLP が推定した密度

についてこれらの one-hot ベクトルを計算し,それら全て を足し合わせて生成する.

3. 実験

3.1 高時間分解画像の取得

本研究ではシミュレーションにより高時間分解画像を 生成し,提案手法の適用を試みた.シミュレーションには MitsubaToFRenderer [3] を用いた.図3は対象物体を正 面と上から見た図である.シーン中には反射率が1の拡散 反射面が二枚あり,その間で相互反射が生じる.生成した 高時間分解画像の画像サイズは64×64であり,ヒストグ ラムのビン数は400である.64×64ピクセルのうち,あ らかじめ物体領域をマスクし,289ピクセルのみを学習に 用いた.

3.2 光子の照射,ヒストグラム生成

光源はカメラの右手側にあり,学習時は一回のパラメー タ更新につき,1000方向をランダムにサンプリングし, 各方向に1000個の光子を照射する.したがって,一度に 1.0×10⁶個の光子を照射する.各光子がどのピクセルで観 測されるかは,反射点をカメラの内部パラメータで画像上 に投影して決定する.2回目以降の反射点については,物 体表面で反射したことを保証するため,2.5で求めた物体 の三次元点群との最小距離が0.1以下のものだけをヒスト グラムの生成に用いる.

3.3 学習の詳細

学習のバッチサイズは 128 とした.すなわち,対象とす る 289 ピクセルのうちランダムに 128 ピクセルを選択し, これらに対してヒストグラムの誤差を計算してパラメー タを一回更新する.合計で 3000 回パラメータの更新を行



図 5 ヒストグラムの生成結果と光子の経路の可視化. (a) 可視化したヒストグラムのピクセ ル位置を赤で示してある. (b) MitsubaToFRenderer (青)と提案手法(橙)が出力し たヒストグラム. (c) 反射回数での色分け. (d) 1 回目の反射点の可視化. (e,f) それぞ れ 2 次, 3 次の相互反射した光子の経路の可視化

なった. パラメータの更新には Adam を用いた.

3.4 実験結果

まず最初に、学習した MLP f_{θ} が出力したシーンの密度 を図 4 に示す. 81 × 81 × 81 のグリッドの各点について推 定を行い、それらを 2 方向から見た図を示してある. 物体 形状に対応して密度が高くなっており、正しく推定できて いることがわかる.

次に、ヒストグラムの生成結果と光子の経路の可視化を 図5に示す.(a)中に赤で示したピクセルについて、生成 したヒストグラムとこのピクセルで観測した光子の経路を 可視化している.(b)は MitsubaToFrenderer と提案手法 が生成したヒストグラムをそれぞれ青と橙で示してある. (c)では提案手法のヒストグラムについて、1回、2回、3 回反射した光子の観測による成分を、それぞれ緑、マゼン タ、赤で色分けしてある.この結果から、提案手法はヒス トグラムにおいて各反射成分が混合されていてもそれらの 分離が可能であることがわかる.(d)-(f)では1回目の反 射点1000個を青で示してある.また、(d)では各ピクセル で観測した1回目の反射点100個を緑で、(e,f)では1次と 2次の相互反射による光子の経路をそれぞれ20経路可視化 してある.提案手法により、光子一つ一つがどのような経 路で SPAD に到達したかを可視化することが可能である.

4. まとめ

本研究では,MLP が推定した密度場上で光子を追跡し, SPAD カメラの高時間分解計測を再現することで MLP を 学習する手法を提案した.実験では,シーンの密度推定と 2次までの相互反射の分離をシミュレーションデータ上で 示した.今後は実データによる実験で提案手法の評価を行 なっていく予定である.また,任意の BRDF のモデル化 や反射率の推定なども今後検討していく予定である.

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] Jang, E., Gu, S. and Poole, B.: Categorical Reparameterization with Gumbel-Softmax, *ICLR* (2017).
- [2] 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).
- [3] Pediredla, A., Veeraraghavan, A. and Gkioulekas, I.: Ellipsoidal path connections for time-gated rendering, ACM TOG (2019).
- [4] 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, *ICCP* (2021).
- [5] Srinivasan, P. P., Deng, B., Zhang, X., Tancik, M., Mildenhall, B. and Barron, J. T.: NeRV: Neural Reflectance and Visibility Fields for Relighting and View Synthesis, *CVPR* (2021).
- [6] 知念響紀,青山尚正,櫛田貴弘,田中賢一郎,舩冨卓哉, 向川康博:空間座標を時間軸に埋め込んだ高時間分解計測 による鏡面物体の形状推定,*MIRU* (2021).
- [7] 生坂優太,櫛田貴弘,田中賢一郎,舩冨卓哉,向川康博:時間分解ラジオシティを用いた反射率と形状の整合性評価, MIRU (2021).