

# ToF カメラを用いた散乱媒体下での 物体領域と奥行きの同時推定

藤村 友貴<sup>1,a)</sup> 藪頭 元春<sup>2</sup> 飯山 将晃<sup>2</sup>

**概要:** 本研究では霧や煙が充満した環境（散乱媒体）下で Time-of-Flight (ToF) カメラを用い、観測した画像中の物体領域とその奥行きを同時に推定する手法を提案する。ToF カメラはシーンに光を照射し物体表面で反射して受光した光から、物体までの奥行きを計測することができる。しかし、散乱媒体下では物体表面で直接反射した光のほかに散乱光が観測されてしまうため、奥行き計測精度が低下してしまう。これに対し、観測される散乱成分はカメラから比較的短い距離で飽和し、加えて十分離れた物体を観測した場合、散乱媒体下での減衰により物体表面での反射成分は無視することができる。したがって、ある程度離れていれば散乱成分のみが観測された領域が存在し、それを用いて物体領域における散乱成分を推定することができる。本研究では、これらの領域と散乱成分を iteratively reweighted least squares (IRLS) の枠組みで同時に推定することにより、シーン中の物体の奥行きを復元する。人工的に霧を発生させた環境下で Kinect v2 で撮影を行い、提案手法を適用することで物体領域と奥行きが推定できることを示す。

**キーワード:** 散乱光, Time-of-Flight, iteratively reweighted least squares

## 1. はじめに

二次元の画像から物体の三次元形状を計測したり、シーンの奥行きを復元したりする三次元復元は、コンピュータビジョンの分野における重要なテーマである。このような技術は、ロボットの三次元認識や自動運転へ応用されることが期待されている。しかしながら、例えばシーン中で霧や煙などが充満している場合、散乱光の影響により精度が低下してしまう。図1左下に、霧が充満した環境で撮影されたRGB画像を示す。このような環境では光が減衰することに加え、物体表面で反射した光のほかに空間で散乱した光も観測されてしまうため、画像のコントラストが低下してしまう。したがって、structure from motion や shape-from-shading などの、画像中の特徴点や輝度値を手掛かりとする手法は三次元復元の精度が低下する。

本研究では、このような散乱光が生じる環境（散乱媒体）においてシーンの三次元情報を取得するための手法を提案する。具体的には、Time-of-Flight (ToF) カメラから取得された画像を用いて、画像中の物体領域とその物体の奥行きを同時に推定する（図1）。ToF カメラとは、照射した光

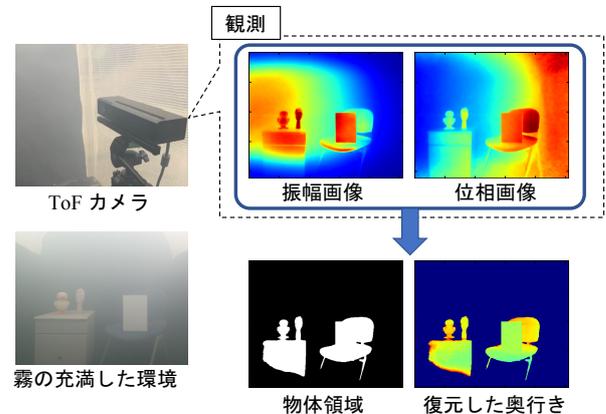


図1: 提案手法の概要。散乱媒体下において ToF カメラで撮影された振幅画像と位相画像を入力し、画像中の物体領域と奥行きを推定する。

と物体で反射して受光した光から、物体の奥行きを計測するカメラのことである。提案手法により、例えば火災現場等の煙が充満した環境下において、障害物の自動検知とその距離の推定が可能になる。

ToF カメラにはいくつかの方式が存在し、Microsoft の Kinect v2 等は正弦波変調方式の ToF カメラである。正弦波変調方式では、照射する光を強度変調させることで正弦波状の光を照射する。物体表面で反射した光は物体の反

<sup>1</sup> 京都大学大学院情報学研究科  
Graduate School of Informatics, Kyoto University

<sup>2</sup> 京都大学学術情報メディアセンター  
ACCMS, Kyoto University

<sup>a)</sup> fujimura@mm.media.kyoto-u.ac.jp

射率の影響で振幅が減衰し、距離に応じて生じた位相差を伴ってカメラで受光される。したがって、カメラでは振幅と位相差を記録した振幅画像と位相画像がそれぞれ得られ、位相差から物体までの奥行きを計測することができる。しかしながら、この方式は本来カメラの各ピクセルにおいて、三次元空間中のある一点を観測することを想定している。一方で散乱媒体下では、先ほど述べたように物体表面で直接反射した光のほかに散乱光が受光されるため、振幅と位相差もその影響を受けてしまい奥行き計測に誤りが生じる。本研究では散乱媒体下において、正弦波変調方式のToFカメラであるKinect v2から得られた振幅画像と位相画像を物理現象に基づいてモデル化し散乱の影響に対処する。

散乱媒体下で得られた振幅画像と位相画像からシーン中の物体の奥行きを計測するには、それぞれの画像中で散乱光の影響により観測した散乱成分を推定し、物体表面で反射した直接反射成分を復元する必要がある。しかしながら、これは一つの観測から二つの成分を分離する不良設定問題である。これに対し、散乱媒体下での散乱成分の飽和と光の減衰を利用することができる。散乱媒体下では近接光源を仮定した場合、観測される散乱成分はカメラから短い距離で飽和することが知られている[20], [21]。つまり、散乱成分は物体までの奥行きには依存しないと仮定できる。加えて、カメラから十分離れた物体からの反射光は散乱媒体下での減衰により無視できると考えられ、このような領域では散乱成分のみが観測されると考えられる。

そこで本研究では、対象とするシーンはカメラからある程度離れており、物体領域と散乱成分のみが観測される領域に分割できると仮定する。散乱成分は物体に依存しないため、散乱成分のみが観測された領域から物体領域上の散乱成分を推定する。これらの仮定に加え、観測される散乱成分の強度は画像上で局所的に二次関数で近似できるという局所的な事前知識と、画像上で対称性を持つという大域的な事前知識を導入し、iteratively reweighted least squares (IRLS) [4], [12]の枠組みで、領域の分割と散乱成分の同時推定を実現する。

## 2. 関連研究

散乱媒体下では光の減衰と散乱光によって観測される画像のコントラストが低下する。コンピュータビジョンや画像処理の分野では、このような環境下で撮影された画像を入力し、劣化する前の画像を復元する研究が行われている[2], [5], [9], [16]。これらの手法はdehazingと呼ばれており、例えばHeら[9]やBermanら[2]は、dark channel priorやhaze-line priorなどのpriorベースのdehazing手法を提案している。また、近年はdeep learningを用いた学習ベースの手法も多く提案されている[3], [13], [17]。

散乱媒体下で三次元復元を行う手法もいくつか提案され

ている(multi-view stereo [14], structured light [7], [15], photometric stereo [6], [21])。これらの手法は、観測される散乱光を物理現象に基づいてモデル化し、従来の三次元復元手法を直接適用する。

散乱媒体下でToFを用いた奥行き計測を行う研究としては[11], [18]がある。Heideら[11]は、photonic mixer device (PMD)を用いたtransient imaging [10]を散乱媒体に対して適用した際、観測を指数ガウス分布の和でモデル化する手法を提案した。Satatら[18]は、single photon avalanche diode (SPAD)を用い、散乱してセンサに到達した光子がガンマ分布でモデル化できることを利用して、散乱成分を分離する手法を提案した。

本研究はこれらの手法と異なり、市販のKinect v2をハードウェアを変更することなく用いるため、安価に実現が可能である。本研究同様、正弦波変調方式のToFカメラを用いた研究としては鈴木ら[23]の研究がある。鈴木ら[23]は単一散乱による影響をモデル化し、画像中で観測される散乱光は飽和するため一様であると仮定している。この場合、画像中で同じ奥行きであり且つ反射率の異なる二つのピクセルを2ペア見つけることで、散乱成分の推定が可能である。これに対して本研究は、カメラと光源に依存した不均一な散乱光をモデル化している点で異なる。加えて、散乱成分の推定のために必要な画像中の領域の推定も同時に行う。

## 3. 散乱媒体下でのToFカメラの観測

最初にToFカメラでの散乱光の観測について述べる。なお、本研究では鈴木ら[23]同様、単一散乱を仮定し、前方散乱の影響は無視できるとする。また、シーンの散乱媒体の濃度は均一であると仮定する。

最初に、正弦波変調方式のToFカメラの観測と奥行き計測方法について簡単に述べる。正弦波変調方式のToFカメラでは、照射した光と物体表面で反射して受光した光との位相差、そして受光した光の振幅を計測する。したがって、ToFカメラの観測は位相差 $\varphi$ と振幅 $\alpha$ を用いて以下のように複素数で記述できる。

$$\alpha e^{j\varphi} \in \mathbb{C} \quad (1)$$

これはフェーザ表示[8]と呼ばれる。位相差は物体までの奥行きにより生じるため、カメラから物体までの奥行き $z$ は以下の式で計算することができる。

$$z = \frac{c\varphi}{4\pi f} \quad (2)$$

ここで、 $c$ は光の速さ、 $f$ は照射した光の変調周波数である。

散乱媒体下では、物体表面で直接反射した光のほかに、散乱光が観測される。散乱光は図2に示すようにカメラと物体を結ぶ光路上で生じ、それらの総和がカメラでの観測となる。したがって、いまカメラを中心とした三次元座標系

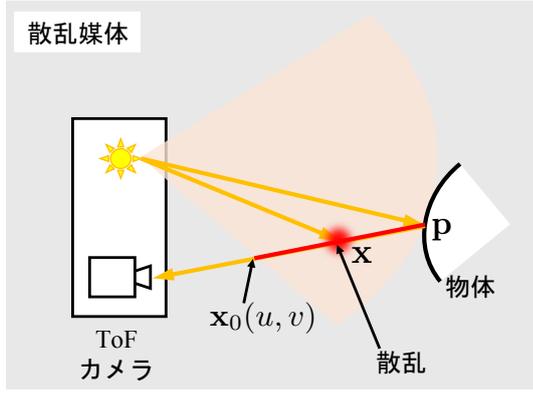


図 2: 散乱媒体下での光の観測。散乱光はカメラと物体を結ぶ光路上で生じる。理想的な点光源は光源中心から任意の方向に向かって同じ強度で光が照射されるが、実際はスポットライトのように強度が均一ではなく光の照射範囲が決まっているため、図中の赤線上で生じた散乱光の総和が観測となる。

を考え、物体表面上の一点  $\mathbf{p}$  をカメラのピクセル  $(u, v)$  で観測したとすると、ToF カメラの観測  $\tilde{\alpha}(u, v; \mathbf{p})e^{j\tilde{\varphi}(u, v; \mathbf{p})}$  は以下の式で書ける。

$$\tilde{\alpha}(u, v; \mathbf{p})e^{j\tilde{\varphi}(u, v; \mathbf{p})} = \alpha_d(u, v; \mathbf{p})e^{j\varphi_d(u, v; \mathbf{p})} + \int_{\|\mathbf{x}\|=\|\mathbf{x}_0(u, v)\|}^{\|\mathbf{p}\|} \alpha(u, v; \mathbf{x})e^{j\varphi(u, v; \mathbf{x})} d\|\mathbf{x}\| \quad (3)$$

$\alpha_d(u, v; \mathbf{p})$  と  $\varphi_d(u, v; \mathbf{p})$  は物体表面  $\mathbf{p}$  で直接反射した光の観測である。ここで、散乱成分は光路上の積分で書くことができるが、積分の始点は  $\|\mathbf{x}\| = \|\mathbf{x}_0(u, v)\|$  でありカメラ中心ではないことに注意が必要である。また、この始点位置は光源とカメラのピクセルの相対位置に依存している。これは、理想的な点光源は光源中心から任意の方向に向かって同じ強度で光が照射されるが、ToF カメラで用いられるものも含む多くの光源は、スポットライトのように強度が均一ではなく、光の照射範囲が決まっているためである [21]。

近接光源を仮定した場合、観測される散乱光の強度はカメラから短い距離で飽和することが知られている [20], [21]。これはすなわち、ある  $\mathbf{x}_{\text{saturation}}$  が存在して、

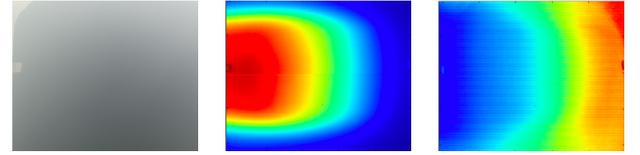
$$\|\mathbf{x}\| \geq \|\mathbf{x}_{\text{saturation}}\| \Rightarrow \alpha(u, v; \mathbf{x}) = 0 \quad (4)$$

となることを意味する。したがって式 (3) は

$$\begin{aligned} \tilde{\alpha}(u, v; \mathbf{p})e^{j\tilde{\varphi}(u, v; \mathbf{p})} &= \alpha_d(u, v; \mathbf{p})e^{j\varphi_d(u, v; \mathbf{p})} \\ &+ \int_{\|\mathbf{x}\|=\|\mathbf{x}_0(u, v)\|}^{\|\mathbf{x}_{\text{saturation}}\|} \alpha(u, v; \mathbf{x})e^{j\varphi(u, v; \mathbf{x})} d\|\mathbf{x}\| \\ &= \alpha_d(u, v; \mathbf{p})e^{j\varphi_d(u, v; \mathbf{p})} + \alpha_s(u, v)e^{j\varphi_s(u, v)} \end{aligned} \quad (5)$$

となる。ここで、 $\alpha_s(u, v)$  と  $\varphi_s(u, v)$  は散乱成分であり、物体上の点  $\mathbf{p}$  には依存せず、ピクセル  $(u, v)$  に依存している。

散乱媒体下では光は媒体を伝達する間に減衰する。よっ



(a) (b) (c)

図 3: 霧を充満させた環境下で黒い背景を撮影したときの画像。(a) RGB 画像。(b) 振幅画像。(c) 位相画像。振幅画像と位相画像において、不均一な散乱成分が観測されている。

て、カメラから十分離れた距離にある物体上の点  $\mathbf{p}_{\text{far}}$  から反射した光は無視できると考えられる。

$$\alpha_d(u, v; \mathbf{p}_{\text{far}}) = 0 \quad (6)$$

したがって、点  $\mathbf{p}_{\text{far}}$  を観測すると散乱成分のみが観測される。

$$\tilde{\alpha}(u, v; \mathbf{p}_{\text{far}})e^{j\tilde{\varphi}(u, v; \mathbf{p}_{\text{far}})} = \alpha_s(u, v)e^{j\varphi_s(u, v)} \quad (7)$$

図 3 は実際に霧を充満させた環境において、ToF カメラで黒い背景を観測したときに得られた振幅画像と位相画像である。黒い背景は十分離れた点を観測している場合と等価と考えられる。したがって、図 3 は散乱成分のみを観測しており、先ほど述べたようにこの散乱成分は光源からの光の強度が均一でなく照射範囲が決まっているために、不均一な値が観測されていることがわかる。

## 4. 物体領域と散乱成分の同時推定

3 章で、散乱成分は物体には依存しない一方で、光源とピクセルの相対位置に依存することを述べた。また、物体までの距離が十分離れたピクセルでは散乱成分のみが観測されることを示した。本研究の目的は、このような散乱成分のみが観測された領域から、物体領域における散乱成分を推定することである。本章では、これらの領域と散乱成分を同時に推定する方法について述べる。

### 4.1 散乱成分の推定

散乱成分は物体に依存しないため、散乱成分のみが観測されたピクセルから推定することが可能である。Tsotsios ら [21] は、通常のカメラとスポットライトを用いたときに観測される後方散乱成分を二次関数で近似する手法を提案した。本研究でも Tsotsios ら [21] 同様、散乱成分に事前知識を導入することで推定を可能にする。

本研究の ToF カメラを用いた問題設定では、Tsotsios ら [21] のように散乱成分を単純な関数で大域的に近似できないことが経験的にわかった。そこで、散乱成分を局所的に二次関数で近似することを考える (図 4)。

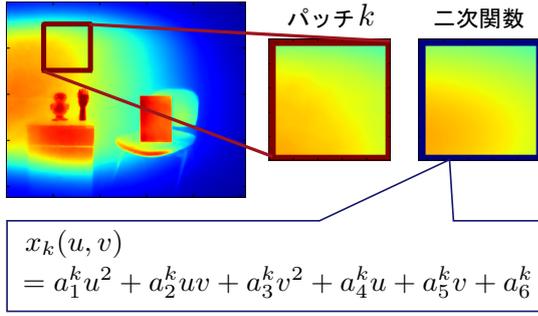


図 4: 散乱成分に対する局所的な事前知識. 散乱成分は局所的には二次関数で近似できると仮定する.

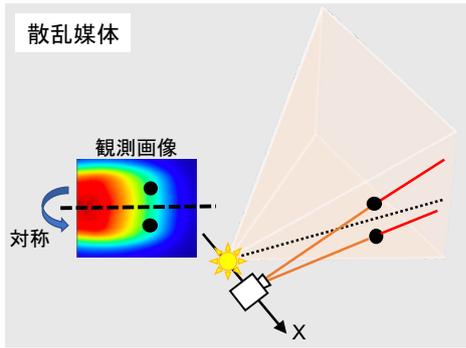


図 5: 大域的な事前知識. カメラと光源が画像の中心軸上に設置されていれば, 散乱成分は中心軸に対して対称になる.

$$x_k(u, v) = a_1^k u^2 + a_2^k uv + a_3^k v^2 + a_4^k u + a_5^k v + a_6^k = \mathbf{a}_k^T \mathbf{u} \quad (8)$$

$x_k(u, v)$  は観測画像 (以下, 本稿では単純に画像と記述した場合, 位相画像  $\varphi$  または振幅画像  $\alpha$  を指す) 上の局所的なパッチ  $k$  における, ピクセル  $(u, v)$  での観測値である.  $\mathbf{u}$  は  $\mathbf{u} = [u^2 \ uv \ v^2 \ u \ v \ 1]^T$  で,  $\mathbf{a}_k = [a_1^k \ a_2^k \ a_3^k \ a_4^k \ a_5^k \ a_6^k]^T$  はパッチ  $k$  の二次関数の係数である.

一方で, 局所的な関数のフィッティングのみでは, 例えば画像上の物体領域が大きい場合, あるいは物体領域上の観測値そのものが二次関数で近似できてしまった場合は正しく推定することができない. そこで, 本研究では局所的な事前知識に加えて大域的な事前知識も導入する.

3章で述べたように, 散乱成分は光源とピクセルの相対位置に依存する. これは, 式 (3) の積分の始点位置がピクセルごとに異なっていることに加え, 光源も方向について強度分布が一樣ではないためである. これに対し, いま図 5 のように画像の中心軸上にカメラと光源が設置されているものとする. この場合, 画像上の中心軸に対して対称なピクセルは積分範囲が同じとなり, 光源の強度分布も対称であるならば, 観測される散乱成分も対称となる. そこで, この対称性を大域的な事前知識として導入する.

いま, 画像上の物体領域と散乱のみが観測された領域が

既知であるとする, これら二つの事前知識を用いて, 以下の最適化問題を解くことで散乱成分を推定する.

$$\min_{\mathbf{x}, \mathbf{a}_1, \dots, \mathbf{a}_K} (\mathbf{x} - \tilde{\mathbf{x}})^T \mathbf{M} (\mathbf{x} - \tilde{\mathbf{x}}) + \lambda_1 \sum_{k=1}^K \|\mathbf{U} \mathbf{a}_k - \mathbf{x}_k\|^2 + \lambda_2 \|\mathbf{F} \mathbf{x} - \mathbf{x}\|^2 + \lambda_3 \|\nabla \mathbf{x}\|^2 \quad (9)$$

ここで,  $\tilde{\mathbf{x}}$  は観測画像,  $\mathbf{x}$  は散乱成分である.  $\mathbf{M}$  は  $\mathbf{M} = \text{diag}(\mathbf{m})$  で,  $\mathbf{m}$  は散乱のみが観測された領域のマスク画像である. すなわち, 各要素  $m_i \in \{0, 1\}$  で,  $m_i = 1$  ならばピクセル  $i$  は散乱のみが観測されたピクセルで,  $m_i = 0$  ならば物体領域のピクセルである. 式 (9) の二項目が局所的な事前知識に基づく項であり, 各パッチ  $k$  に対して二次関数のフィッティングを行う. 三項目は大域的な事前知識に基づく項であり, 行列  $\mathbf{F}$  は画像を中心軸周りで反転させる行列である. そして, 最後の項は平滑項であり,  $\lambda_1, \lambda_2, \lambda_3$  は各項の重みである. この目的関数を, 散乱成分  $\mathbf{x}$  と各パッチの二次関数の係数  $\mathbf{a}_1, \dots, \mathbf{a}_K$  について最小化する. この最適解は以下のように閉形式で得ることができる.

$$\mathbf{x} = (\mathbf{M} + \lambda_1 \mathbf{I} + \lambda_2 (\mathbf{F} - \mathbf{I})^T (\mathbf{F} - \mathbf{I}) + \lambda_3 \nabla^T \nabla)^{-1} (\mathbf{M} \tilde{\mathbf{x}} + \lambda_1 \mathbf{y}(\mathbf{a}_1, \dots, \mathbf{a}_K)) \quad (10)$$

$$\mathbf{a}_k = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \mathbf{x}_k \quad (11)$$

ここで,  $\mathbf{y}(\mathbf{a}_1, \dots, \mathbf{a}_K)$  は二次関数のフィッティングで得られた散乱成分の画像である.

## 4.2 IRLS を用いたマスクの推定

4.1 節で, 散乱のみが観測された領域かどうかを表すマスク  $\mathbf{m}$  が与えられたときの, 散乱成分の推定方法について述べた. 本研究では, 現在の推定結果を用いて IRLS の枠組みでマスク  $\mathbf{m}$  を更新することで, 散乱成分と同時に推定する方法を提案する.

IRLS とは重み付き二乗誤差を, 重みを反復的に更新することで最小化するアルゴリズムであり,  $L_p$  ノルムの近似として圧縮センシング [4], [22] や, ロバスト推定 [12] に用いられる. 本研究では, マスク  $\mathbf{m}$  を IRLS における重みと同様の方法で更新する.

まず最初に, 式 (9) の最適化についてロバスト推定を考える. ロバスト推定では各誤差項  $r_i$  に対して二乗誤差  $r_i^2$  ではなく, 代わりに微分可能である非線形な損失関数  $\rho(x)$  を用いることで外れ値にロバストな推定を可能にする. したがって, 式 (9) の一項目についてロバスト推定を考えると,

$$\min_{\mathbf{x}} \sum_{i=1}^N \rho(r_i(\mathbf{x})) \quad (12)$$

## アルゴリズム 1 Estimate $\mathbf{x}$ and $\mathbf{m}$

- 1: Initialization:  $\mathbf{M} \leftarrow \mathbf{I}$ ,  $\mathbf{a}_k \leftarrow \underset{\mathbf{a}_k}{\operatorname{argmin}} \|\mathbf{U}\mathbf{a}_k - \tilde{\mathbf{x}}_k\|^2$
- 2: Update  $\mathbf{x}$  using Eq. (10)
- 3: Update  $\mathbf{a}_1, \dots, \mathbf{a}_K$  using Eq. (11)
- 4: Update  $\mathbf{m}$  with IRLS
  - 5: Compute  $r_i$  using Eq. (13), then normalize it
  - 6: Update  $m_i$  using Eq. (17)
- 7: Back to 2 until convergence
- 8: Binarize  $\mathbf{m}$

となる, ここで,  $N$  は画像のピクセル数で, 誤差  $r_i$  は

$$r_i = |x_i - \tilde{x}_i| \quad (13)$$

である. なお, 一般に誤差  $r_i$  は正規化された値が用いられる. 式 (9) では, 物体領域のピクセルをマスク  $m_i = 0$  として明示的に与えていたが, この式では外れ値として扱っていると解釈できる.

IRLS ではロバスト推定を重み付き二乗誤差の最小化に置き換える. すなわち, 式 (12) を

$$\min_{\mathbf{x}} \sum_{i=1}^N m_i r_i(\mathbf{x})^2 \quad (14)$$

とする. ここで,  $m_i$  が各誤差  $r_i$  に対応する重みであり, 現在の推定結果を用いて

$$m_i = \frac{\rho'(r_i)}{r_i} \quad (15)$$

として更新される. この更新式により, 誤差の大きい項に対応する重み  $m_i$  の値を減少させることでロバスト推定を実現する. ここで, 式 (14) は式 (9) の一項目と等価である. また, 重み  $m_i$  の性質は本研究のマスクと一致する. つまり, 誤差  $r_i$  が大きい場合は物体領域である可能性が高いため, マスク  $m_i$  の値は小さくなるように更新される. 以上により, 本研究のマスクは IRLS における重みと同様の枠組みで更新できる.

具体的にどのように更新を行うかは, 損失関数  $\rho(x)$  として何を用いるかに依存する. 本研究では以下の損失関数を用いた.

$$\rho(x) = \begin{cases} \frac{c^2}{6} \left[ 1 - \left\{ 1 - \left( \frac{x}{c} \right)^2 \right\}^3 \right] & \text{if } |x| \leq c \\ \frac{c^2}{6} & \text{otherwise} \end{cases} \quad (16)$$

この場合, 重みの更新式は

$$m_i = \begin{cases} \left\{ 1 - \left( \frac{\hat{r}_i}{c} \right)^2 \right\}^2 & \text{if } \hat{r}_i \leq c \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

となる. これは Tukey's biweight [1] と呼ばれ,  $0 \leq m_i \leq 1$  である. ここで, 誤差  $r_i$  は標準偏差  $\sigma$  で正規化した値  $\hat{r}_i = r_i/\sigma$  を用いた.

散乱成分とマスクの同時推定を行う最終的なアルゴリズム



図 6: 実験環境

ムをアルゴリズム 1 に示す. 初期値は  $\mathbf{M} = \mathbf{I}$  で, すなわち全てのピクセルが散乱のみが観測されたピクセルであるとする. なお, 本研究の  $m_i$  は  $m_i \in \{0, 1\}$  であるので, 推定の最後に二値化を行う.

### 4.3 奥行きへの復元と物体領域

各ピクセルで散乱成分  $\alpha_s(u, v)$  と  $\varphi_s(u, v)$  が推定できれば, 式 (5) から以下のように直接反射成分の位相差を計算できる.

$$\begin{aligned} \varphi_d(u, v; \mathbf{p}) &= \arctan \frac{\tilde{\alpha}(u, v; \mathbf{p}) \sin \tilde{\varphi}(u, v; \mathbf{p}) - \alpha_s(u, v) \sin \varphi_s(u, v)}{\tilde{\alpha}(u, v; \mathbf{p}) \cos \tilde{\varphi}(u, v; \mathbf{p}) - \alpha_s(u, v) \cos \varphi_s(u, v)} \end{aligned} \quad (18)$$

その後, 式 (2) を用いて奥行きを復元する.

物体領域は推定されたマスクを反転させることで得ることができる. 提案手法では, 振幅画像と位相画像のそれぞれで物体領域が得られるので, 最終的な物体領域はそれらの共通部分とした.

## 5. 実験

人工的に霧を発生させた環境で ToF カメラで画像を撮影し, 提案手法の有効性を評価した.

### 5.1 実験環境

実験環境を図 6 に示す. 186 cm  $\times$  161 cm の閉鎖空間内に Fogマシンと ToF カメラとして Kinect v2 を設置した. 閉鎖空間の壁面は黒くしてあり, 背景では散乱成分のみ観測される. Kinect v2 は, 120 MHz, 80 MHz, 16 MHz の強度変調光を照射することができる. 本研究では 16 MHz で取得された画像のみを用いた. 振幅画像と位相画像の取得には Tanaka ら [19] のコードを用いた. このコードでは複数フレームが平均されて出力されるが, 変更を加えて提案手法には 1 フレームで取得された画像のみを入力した. 高周波のノイズを取り除くため, 前処理としてバイラテラルフィルタを用いた.

Kinect v2 から取得された画像のサイズは 424  $\times$  512 ピクセルである. 本研究では画像を 3  $\times$  4 の 12 のパッチに分割し, それぞれにおいて二次関数のフィッティングを行っ

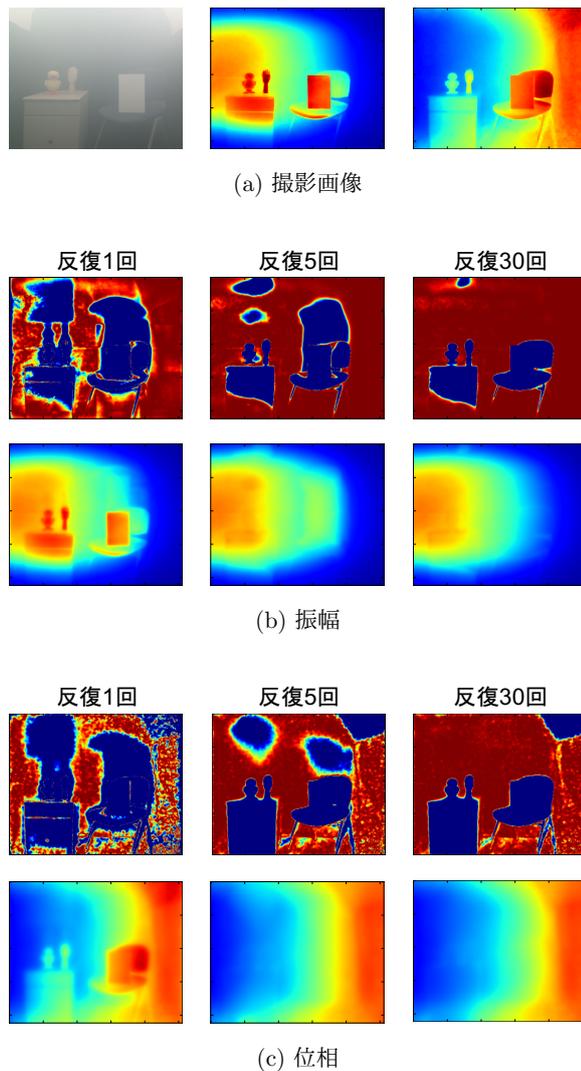


図 7: (a) 撮影画像. 左から RGB 画像, 振幅画像, 位相画像. (b) (c) 振幅画像と位相画像に提案手法を適用した結果. それぞれ, 各反復において推定されたマスクと散乱成分を示してある.

た. 提案手法では, 光源とカメラが画像の中心軸上に位置していることを仮定していたが, 実際は光源とカメラの軸に若干のずれがあり, この場合は対称軸がずれてしまうことに加え, 厳密には散乱成分は対称にはならない. 一方で, これにより生じる誤差は十分に小さいことがわかった. 本研究では行列  $\mathbf{F}$  として経験的に縦方向で 200 ピクセルの軸に沿って画像を反転させる行列を用いた. それに伴い式 (9) の第三項目では, 画像の下部 24 ピクセル分は最適化に含めないようにした.

## 5.2 実験結果

図 7 と図 8 に実験結果を示す. 図 7(a) は Kinect v2 で撮影された画像で, 左から RGB 画像, 振幅画像, 位相画像である. 図 7(b)(c) は提案手法を適用した結果で, 振幅画像と位相画像に対して, それぞれ各反復で推定された散乱

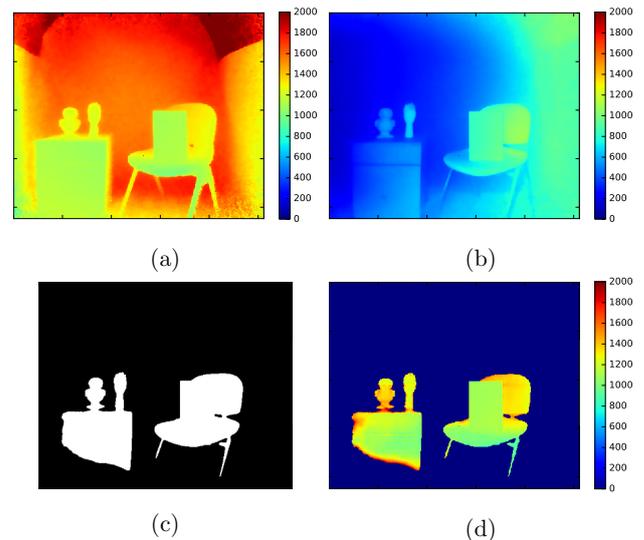


図 8: 奥行き復元の結果. (a) 霧が無い場合の奥行きの計測. (b) 霧がある場合の奥行きの計測. (c) 提案手法で出力した物体領域. (d) 提案手法で復元した奥行き. 物体領域でマスクしてある.

成分とマスク  $\mathbf{m}$  を示してある. なお, 図に示したマスクは二値化する前の結果である. 提案手法を用いることで, 物体領域の部分を抽出できていることがわかる. 特に, 椅子の足といった細い領域でも提案手法は抽出が可能である. また, 物体領域上の散乱成分についても正しく推定できていることがわかる. 図 8 は奥行きの復元結果である. (a) は霧が無い場合の奥行きの計測で, (b) は霧がある場合の奥行きの計測である. 霧がある場合は散乱光の影響で奥行きを正しく計測することができない. 対して, (c), (d) はそれぞれ提案手法で出力した物体領域と復元した奥行きであり, 霧で劣化した奥行きを正しく復元できていることがわかる.

霧の濃度を変えて実験した結果を図 9 と図 10 に示す. 図 9 は霧の濃度が低い場合で, 図 10 は霧の濃度が高い場合である. 霧の濃度が高い場合は, 散乱成分の影響が大きい箇所において奥行きの復元精度が低下してしまうものの, 全体的には提案手法を用いることで濃度に関係なく, 奥行きの精度の向上と物体領域の推定ができていることがわかる. 表 1 には濃度の各場合において, 散乱媒体が無い場合を真値として, 観測画像中の白い板の領域上で奥行きの平均誤差を計算したものである. 表中の“観測”は観測そのもので計算した誤差, “提案手法”は提案手法を適用して計算した誤差である. 散乱媒体下では散乱光の影響で数百 mm の誤差が生じてしまうものの, 提案手法を適用することで十数 mm まで減少させることができる.

## 6. まとめ

本研究では, 霧や煙などが充満した散乱光が生じる環境下において, ToF カメラを用いて物体領域とその奥行き

表 1: 白い板の奥行き平均誤差

	濃度	観測	提案手法
図 9(b)	低	111.39 mm	<b>14.13 mm</b>
図 8	中	244.94 mm	<b>14.50 mm</b>
図 10(b)	高	411.68 mm	<b>11.63 mm</b>

を推定する方法を提案した。ToF カメラでは、光源とピクセルの相対位置に依存する散乱成分が観測される。本研究では、散乱成分が局所的には二次関数で近似できるという局所的な事前知識と、散乱成分は画像の中心軸に対して対称となる大域的な事前知識を用いて、散乱成分の推定を行なった。また IRLS の枠組みで散乱のみが観測された領域をマスクとして抽出することで、物体領域と散乱成分の同時推定を可能にした。人工的に霧を発生させた環境で Kinect v2 を用いて撮影を行い、物体領域と奥行きが正しく推定できることを示した。

今後は、より現実的な環境で実験を行い提案手法の有効性を検証する予定である。また、提案手法により物体表面で直接反射した光の振幅画像が得られるため、これを用いて奥行き精度を向上させることも可能であると考えている。Kinect v2 では RGB 画像も同時に得られるため、得られた奥行きから RGB 画像中の散乱光を除去することも今後の方向性として考えられる。

#### 参考文献

[1] Beaton, A. E. and Tukey, J. W.: The Fitting of Power Series, Meaning Polynomials, Illustrated on Band-Spectroscopic Data, *Technometrics*, Vol. 16, No. 2, pp. 147–185 (1974).

[2] Berman, D., Treibitz, T. and Avidan, S.: Non-Local Image Dehazing, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1674–1682 (2016).

[3] Cai, B., Xu, X., Jia, K., Qing, C. and Tao, D.: DehazeNet: An End-to-End System for Single Image Haze Removal, *IEEE Transaction on Image Processing*, Vol. 25, No. 11, pp. 5187–5198 (2016).

[4] Chartrand, R. and Yin, W.: Iteratively reweighted algorithms for compressive sensing, *The IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3869–3872 (2008).

[5] Fattal, R.: Dehazing Using Color-Lines, *ACM Transactions on Graphics (TOG)*, Vol. 34, No. 1 (2014).

[6] Fujimura, Y., Iiyama, M., Hashimoto, A. and Minoh, M.: Photometric Stereo in Participating Media Considering Shape-Dependent Forward Scatter, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 7445–7453 (2018).

[7] Gu, J., Nayar, S. K., Belhumeur, P. N. and Ramamoorthi, R.: Compressive Structured Light for Reconstructing Inhomogeneous Participating Media, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 3, pp. 555–567 (2013).

[8] Gupta, M., Nayar, S. K., Hullin, M. B. and Martin, J.: Phasor Imaging: A Generalization of Correlation-Based Time-of-Flight Imaging, *ACM Transaction on Graphics*

(*TOG*), Vol. 34, No. 5 (2015).

[9] He, K., Sun, J. and Tang, X.: Single Image Haze Removal Using Dark Channel Prior, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 33, No. 12, pp. 2341–2353 (2011).

[10] Heide, F., Hullin, M. B., Gregson, J. and Heidrich, W.: Low-budget Transient Imaging using Photonic Mixer Devices, *ACM Transactions on Graphics (TOG)*, Vol. 32, No. 4 (2013).

[11] Heide, F., Xiao, L., Kolb, A., Hullin, M. B. and Heidrich, W.: Imaging in scattering media using correlation image sensors and sparse convolutional coding, *Optics Express*, Vol. 22, No. 21, pp. 26338–26350 (2014).

[12] Holland, P. and Welsch, R. E.: Robust regression using iteratively reweighted least-squares, *Communications in Statistics Theory and Method*, Vol. 6, No. 9, pp. 813–827 (1977).

[13] Li, B., Peng, X., Wang, Z., Xu, J. and Feng, D.: AOD-Net: All-in-One Dehazing Network, *The IEEE International Conference on Computer Vision (ICCV)*, pp. 4770–4778 (2017).

[14] Li, Z., Tan, P., Tang, R. T., Zou, D., Zhou, S. Z. and Cheong, L.: Simultaneous Video Defogging and Stereo Reconstruction, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4988–4997 (2015).

[15] Narasimhan, S. G., Nayar, S. K., Sun, B. and Koppal, S. J.: Structured Light in Scattering Media, *Proceedings of the Tenth IEEE International Conference on Computer Vision*, Vol. I, pp. 420–427 (2005).

[16] Nishino, K., Kratz, L. and Lombardi, S.: Bayesian Defogging, *International Journal of Computer Vision*, Vol. 98, No. 3, pp. 263–278 (2012).

[17] Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X. and Yang, M.: Single Image Dehazing via Multi-scale Convolutional Neural Networks, *European Conference on Computer Vision (ECCV)*, pp. 154–169 (2016).

[18] Satat, G., Tancik, M. and Rasker, R.: Towards Photography Through Realistic Fog, *The IEEE International Conference on Computational Photography (ICCP)*, pp. 1–10 (2018).

[19] Tanaka, K., Mukaigawa, Y., Funatomi, T., Kubo, H., Matsushita, Y. and Yagi, Y.: Material Classification using Frequency- and Depth-Dependent Time-of-Flight Distortion, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 79–88 (2017).

[20] Treibitz, T. and Schechner, Y. Y.: Active Polarization Descattering, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 3, pp. 385–399 (2009).

[21] Tsotsios, C., Angelopoulou, M. E., Kim, T. and Davison, A. J.: Backscatter Compensated Photometric Stereo with 3 Sources, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2259–2266 (2014).

[22] Wipf, D. and Nagarajan, S.: Iterative Reweighted l1 and l2 Methods for Finding Sparse Solutions, *IEEE Journal on Selected Topics in Signal Processing*, Vol. 3, No. 2, pp. 317–329 (2010).

[23] 鈴木大介, 田中賢一郎, 北野和哉, 船富卓哉, 向川康博: 極座標系の原点移動モデルを用いた ToF 計測における霧の影響除去, 情報処理学会研究報告, Vol. 2018-CVIM-210, No. 49 (2018).

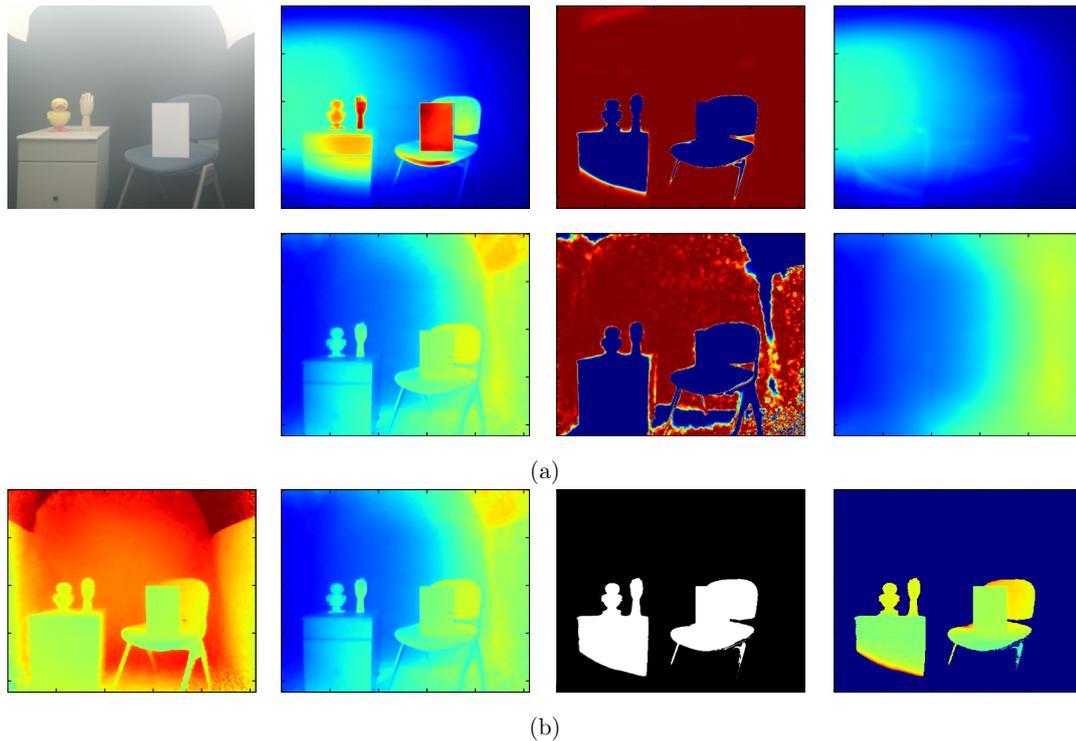


図 9: 霧の濃度が低い場合の実験結果. (a) 散乱成分とマスクの推定. 上段が振幅の推定結果で, 下段が位相の推定結果である. (b) 最終的な物体領域と復元した奥行き. 左から, 霧がない場合の奥行き, 霧がある場合の奥行き, 最終的な物体領域, 復元した奥行きである.

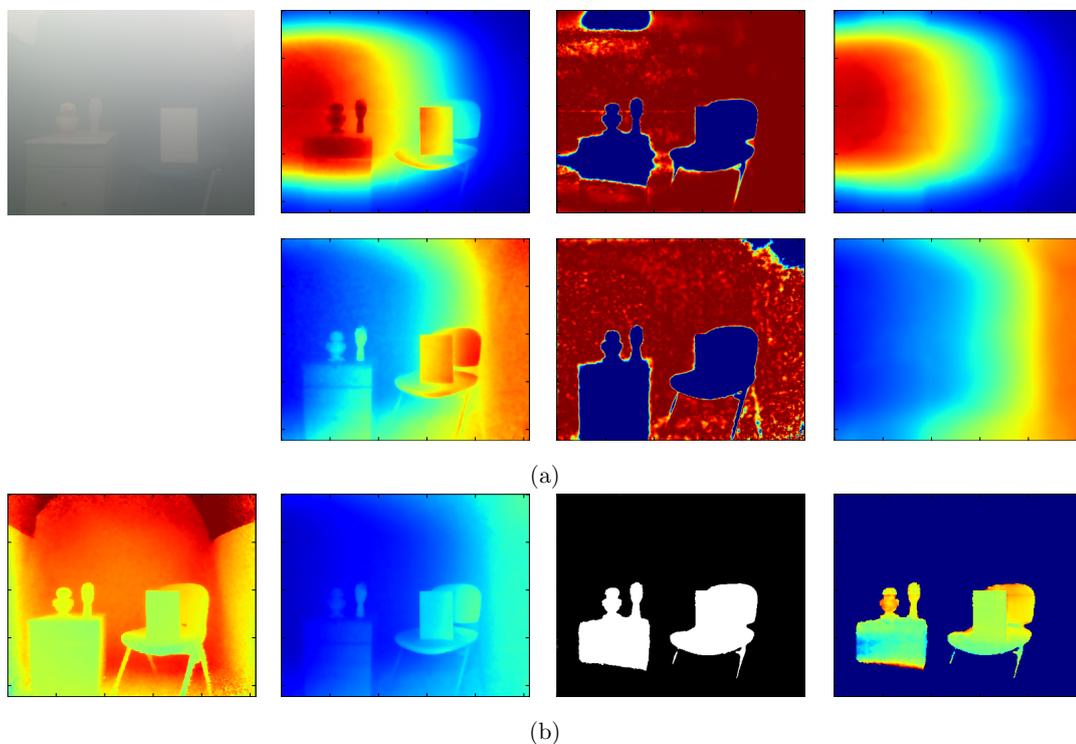


図 10: 霧の濃度が高い場合の実験結果. (a) 散乱成分とマスクの推定. 上段が振幅の推定結果で, 下段が位相の推定結果である. (b) 最終的な物体領域と復元した奥行き. 左から, 霧がない場合の奥行き, 霧がある場合の奥行き, 最終的な物体領域, 復元した奥行きである.