

# 点像分布関数設計と深度マップ正則化に基づく スナップショット空間超解像 ToF センシング

河内 穂高<sup>1,a)</sup> 中村 友哉<sup>1,b)</sup> 榎原 靖<sup>1,c)</sup> 八木 康史<sup>1,d)</sup>

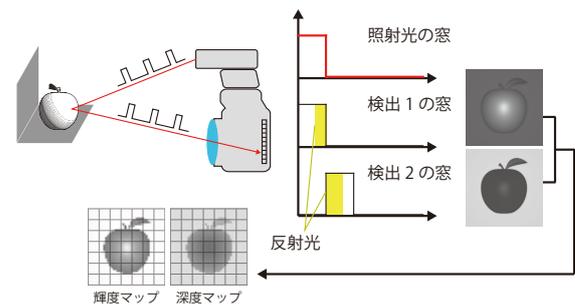
**概要：**光飛行時間 (ToF) センサは深度マップ計測デバイスとして普及しているが、通常のイメージセンサと比較して空間解像度が低い。この問題を解決するために、本研究では点像分布関数設計と深度マップ正則化に基づく圧縮センシング型スナップショット空間超解像法を提案する。Bicubic 補間と提案手法の比較数値実験を行い、正則化項の設計が適合するシーンに対して深度マップの実効分解能及び精度が改善されることを確認した。

**キーワード：**圧縮センシング, ToF センシング, 深度マップ

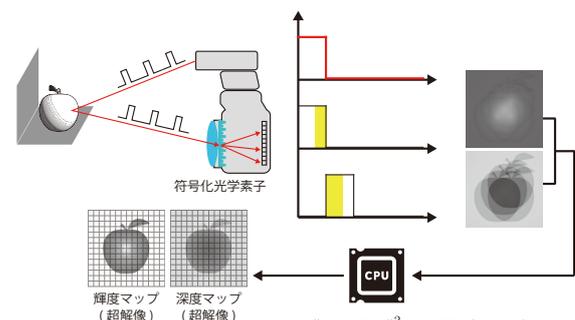
## 1. はじめに

近年、深度マップの利用先はロボティクスから自動運転まで多岐にわたり、深層学習などの発展も受けて、深度マップを入力とするシステムの普及が進んでいる。より高解像度な深度マップを用いれば従来の深度マップを用いたシステムの精度向上が見込める。また従来の解像度では実現できなかった視覚システムの構築も可能になる。これまで様々な深度マップの取得方法が研究されてきており、中でも光飛行時間を位相差により計測する位相差 ToF 法は安価かつ小型なデバイスの設計が可能であるため急速な普及を遂げている [1]。しかし位相差 ToF センサは、一般的なイメージセンサに比べて検出系の複雑化により解像度が制限されるため、高解像度化をハードウェア設計によって実現するのは困難である。そのため、センサ画素数を増やすことなく高解像度化することが必要となる。

センサデバイスによる物理的解像度制約を超えるために、本研究ではデジタル超解像の適用を検討する。モデル駆動で実現可能な超解像手法として代表的なものは、複数回撮影を前提としたマルチフレーム超解像手法である。その有効性については過去に十分実証されており、実用化もされているが、センシングにおいては時間分解能が犠牲



(a) 従来の ToF 法



(b) 提案手法

図 1 研究概要

<sup>1</sup> 大阪大学  
Osaka University

a) kawachi@am.sanken.osaka-u.ac.jp

b) nakamura@am.sanken.osaka-u.ac.jp

c) makihara@am.sanken.osaka-u.ac.jp

d) yagi@am.sanken.osaka-u.ac.jp

となる。

一方近年、符号化撮像と圧縮センシングを用いたスナップショット超解像法が研究されている [2]。超解像の実現のためには、イメージセンサのナイキスト周波数を越える

帯域の情報, 言い換えればサブ画素情報を計測及び再構成することが重要である. マルチフレーム超解像法は複数回撮影によってサブ画素情報を直接計測するのに対し, 圧縮センシング超解像法は符号化光学系を介した撮像によりサブ画素情報を符号化計測する. 符号化計測画像そのものはイメージセンサ解像度であるため, 計測画像から超解像画像を再構成するためには不良設定逆問題を解く必要があるが, 計測対象物体をスパースモデリングすることにより解の唯一性を保証した逆解析を実現できる. この枠組は圧縮センシング [3] と呼ばれ, 画像センシングにおける応用も研究されている [4].

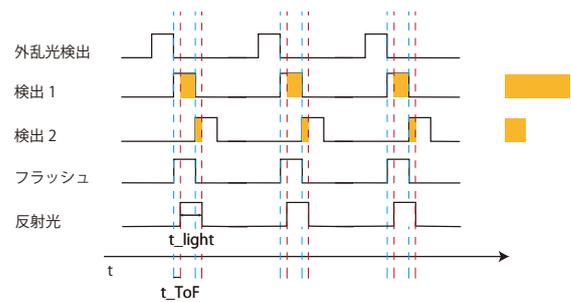
本研究では, この圧縮センシング技術を位相差 ToF イメージングに適用し, スナップショット超解像距離画像センシングを実現する. 図 1 のように, 通常のレンズ光学系では低解像度な観測しかできなかったのに対して, 位相変調素子等を瞳面におくことで, 符号化した画像を観測し, 深度に正則をかけて再構成することで, 高解像度な深度マップを取得することができる.

## 2. 関連研究

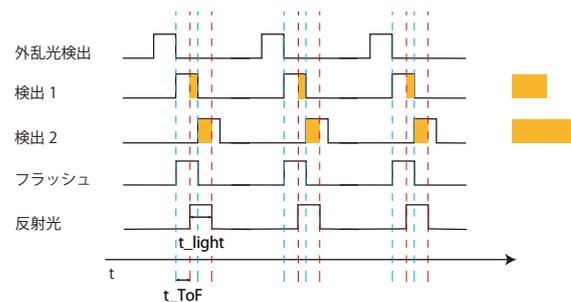
空間画像センシングにおいて, 符号化と圧縮センシングを用いたスナップショット超解像法が提案されている [2]. 当該手法では, 符号化光学系によりサブ画素情報を符号化計測し, その後演算系で超解像画像を再構成する. 画像再構成は一般的にスパース正則化付誤差最小化問題を解くことで実装される. 空間画像センシングでは, 計測対象物体画像と計測符号化画像は一度の線形変換で関係づけられるため, 反復解法における正則化の実装は単に推定画像をスパース化することで実装可能である. 一方で間接 ToF センシングでは, 計測対象が深度マップであり, 計測符号化画像とは非線形な関係となる. そのため, 当該手法の間接 ToF センシングへの適用可能性は自明ではない.

複数回計測に基づく圧縮センシング型超解像 ToF センシングが提案されている [5]. 当該手法では, 物体の結像面に高解像度なマイクロミラーアレイデバイス (DMD) を置き, 複数回ランダムな強度反射率を実装することで, 低解像度なイメージセンサの帯域を超える画像及び深度マップを再構成する. この時, 圧縮センシング法を用いることで, 再構成データの画素数を遙かに下回る撮像回数で超解像情報を再構成できる. 本研究では, 開口面に符号化開口を置き, 超解像情報を分散多重計測することでスナップショットでの超解像センシングの実現を目指す.

近年, 機械学習を用いた圧縮センシング超解像技術も広く研究されている [6]. データ駆動型の超解像法は, 高速な画像再構成が可能である一方で, 精度のシーン依存性や説明可能性が問題となる. 本研究では情報計測への応用を念頭に, モデル駆動型の超解像法を検討する.



(a) 近距離を撮像した場合



(b) 遠距離を撮像した場合

図 2 撮像距離とタイミングチャート

## 3. 提案手法

本研究では, 間接 ToF 法の中でもパルス駆動形間接 ToF 法 [7] を研究対象とする. パルス駆動形間接 ToF 法で遠距離及び近距離をそれぞれ撮像した場合のタイミングチャートを図 2 に示す. この方式では各画素ごとに, 外乱光のみを受光するセンサ, 近距離で反射した光を多く受光するセンサ (検出 1), 遠距離で反射した光を多く受光するセンサ (検出 2) の 3 種類のセンサで構成される. (a) は近距離の点を画素で観測した場合で, 検出 1 回目により多くの反射光を受光している. 対して (b) では, 検出 2 回目により多くの光を受光している.  $t_{\text{ToF}}$  は光飛行時間を表しておりパルス幅の  $t_{\text{light}}$  との比が検出 1 回目と 2 回目の和と検出 2 回目の比と一致している. また外乱光検出フェーズでは, 照射光の反射光は検出されないため, 外乱光の各画素ごとの大きさを計測することができる. 以上よりこの 3 回の検出から全画素の距離情報を統合することで, 単一シーンの深度マップ  $\mathbf{d} \in \mathbb{R}^{N_{\text{low}} \times 1}$  が得られる. ただし  $N_{\text{low}}, N_{\text{high}} \in \mathbb{R}$  はそれぞれ低解像度画像と高解像度画像のサイズを表す. 各時間窓画像から深度マップを算出する過程は次式で定式化される. ただし, 本研究では特に断りのない限り二次元画像はラスタスキャンを行い一元化したものとする.

$$\begin{aligned} \mathbf{r} &= \mathbf{Q}_2 \odot (\mathbf{Q}_1 + \mathbf{Q}_2), \\ \mathbf{d} &= \frac{c}{2} t_{\text{light}} \mathbf{r}. \end{aligned} \quad (1)$$

ここで,  $\mathbf{Q}_1, \mathbf{Q}_2 \in \mathbb{R}^{N_{\text{low}} \times 1}$  は各時間窓で検出された計測

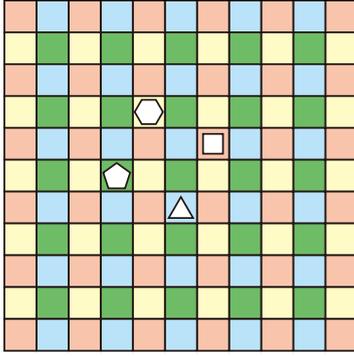


図 3 適切な点像分布関数例

輝度画像,  $\mathbf{r} \in \mathbb{R}^{N_{\text{low}} \times 1}$  はそれらの輝度比,  $t_{\text{light}} \in \mathbb{R}$  は光照射時間,  $c$  は空気中の光速,  $\circ$  はアダマール除算を表す.

圧縮センシングを用いた超解像を行うためには対象物体のサブ画素空間情報を可逆符号化した観測が必要となる. そこで本研究では, 開口面で符号化光学素子を置くことにより, 点像分布関数, すなわち撮像系のインパルス応答を積極設計することで, この光学的符号化を実装する. 点像分布関数は, 全てのサブ画素が均等に観測されるように設計する. 問題設定として, 縦横それぞれ 2 倍の超解像を考える. また低解像度へのダウンサンプリングは 4 画素のうちの 1 つのみの値を採用することにする. これはレンズ光学系の場合にセンサ同士の間サブ画素情報が消失する状態を想定している. 画像は高解像度のとき 4 倍の画素数によって構成されることと, それぞれの位置関係を考慮すると図 3 のように色分けすることが可能となる. 低解像度レンズ光学系ではこの 4 色のうちの 1 色のみしか観測することができない. 4 点のスポットを持つインパルス応答を考えると, 図 3 のように 4 点の配置が 4 色均等に含まれる配置ならば, 各画素を均等に符号化することが可能となる. よって適切に符号化を行うことのできる点像分布関数とは, 4 色のスポットの値のそれぞれの合算値が全て等しい分布と考えることができる. 本研究で用いる点像分布関数はこの条件を満たすように設計することとする.

この符号化により低解像度な観測であってもサブ画素情報を計測画像内に残すことができる. 符号化開口を介した輝度イメージングは, シフト不変性に基づき点像分布関数のアナログ空間での畳み込み演算によって表現される.

輝度画像の圧縮センシング [3] や, 複数回撮影に基づく既存の圧縮 ToF センシング [5] では, 各時間窓で検出された符号化計測輝度画像から再構成される各時間窓に対応する物体輝度画像  $\mathbf{f}_1, \mathbf{f}_2 \in \mathbb{R}^{N_{\text{high}} \times 1}$  に対して全変動を用いて正則化をかけており, 式 (2) に示すような最適化問題を各時間窓に対応する物体輝度画像に対して解くことになる.

$$\mathbf{f}_i^* = \arg \min_{\mathbf{f}_i} \text{TV}(\mathbf{f}_i) \quad (2)$$

$$\text{subject to } \mathbf{g}_i = \Phi \mathbf{f}_i.$$

ここで,  $\mathbf{g}_i \in \mathbb{R}^{N_{\text{low}} \times 1}$  は各時間窓に対応する低解像符号化観測輝度画像,  $\Phi \in \mathbb{R}^{N_{\text{low}} \times N_{\text{high}}}$  はインパルス応答の畳み込みを行う観測行列,  $\text{TV}(\cdot)$  は二次元全変動の評価関数とする. 二次元全変動の評価関数は一般にエッジなどの情報を残したまま平滑化を行うための評価関数で式 (3) のように表現される [8]. ただし  $\nabla_x \mathbf{f}_i$  及び  $\nabla_y \mathbf{f}_i$  は画像  $\mathbf{f}$  を横方向と縦方向にそれぞれ微分した結果の  $i$  番目の要素を表す.

$$\text{TV}(\mathbf{f}) = \sum_i \sqrt{\nabla_x f_i^2 + \nabla_y f_i^2}. \quad (3)$$

輝度画像に対する空間的な滑らかさを用いた正則化は, 多くの応用での有用性が過去研究により示されている. 一方で, ToF センシングの場合は最終的な計測の興味が深度マップであるため, 必ずしも輝度画像に対して正則化を行う必然性はない. むしろ, 多くの状況では輝度画像よりも深度マップの方がよりスパースに表現できることが直感的には予想され, 正則化の対象としてより適していると考えられる. そこで本研究では深度マップに対して全変動による正則化をかける. 深度マップは観測時の符号化のための線形な畳み込み演算だけでなく式 (1) の非線形変換を経て得られるため, 式 (2) の最適化問題をそのまま適用することはできない. そこで本研究では, 距離情報を含む各時間窓に対応する物体輝度画像は距離情報である位相差比率画像  $\mathbf{R} \in \mathbb{R}^{N_{\text{high}} \times 1}$  とシーンの輝度画像  $\mathbf{P} \in \mathbb{R}^{N_{\text{high}} \times 1}$  の 2 つの独立変数に従属する変数であると考え, 各時間窓に対応する物体輝度画像を式 (4) のように表す.

$$\mathbf{f}_1 = \mathbf{P} \circ (\mathbf{1} - \mathbf{R}), \quad (4)$$

$$\mathbf{f}_2 = \mathbf{P} \circ \mathbf{R}.$$

ここで  $\circ$  はアダマール積,  $\mathbf{1} \in \mathbb{R}^{N_{\text{high}} \times 1}$  は全要素が 1 のベクトルを表す. これを用いて最適化問題を式 (5) のように定義する.

$$\mathbf{R}^*, \mathbf{P}^* = \arg \min_{\mathbf{R}, \mathbf{P}} \text{TV}(\mathbf{R}) \quad (5)$$

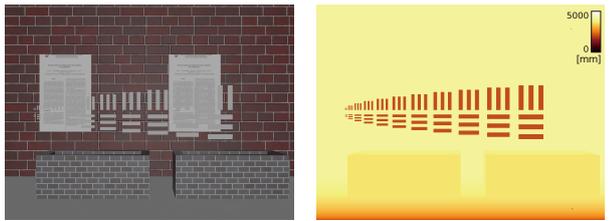
$$\text{subject to } \mathbf{g}_1 = \Phi(\mathbf{P} \circ (\mathbf{1} - \mathbf{R}))$$

$$\mathbf{g}_2 = \Phi(\mathbf{P} \circ \mathbf{R}).$$

本研究では, 最適化問題 (5) を, 式 (6) のようなデータ忠実項と正則化項の和で表現されるコスト関数の最小化問題に変形し, これを連続最適化と捉え反復アルゴリズムにより解く. 今回は, 正則化項に二次元全変動 [8] を選択する.

$$\begin{aligned} \mathbf{R}^*, \mathbf{P}^* = \arg \min_{\mathbf{R}, \mathbf{P}} & \tau \text{TV}(\mathbf{R}) \\ & + \|\mathbf{g}_1 - \Phi(\mathbf{P} \circ (\mathbf{1} - \mathbf{R}))\|_2^2 \\ & + \|\mathbf{g}_2 - \Phi(\mathbf{P} \circ \mathbf{R})\|_2^2. \end{aligned} \quad (6)$$

ここで,  $\tau \in \mathbb{R}$  は正則化項の係数を表す.



(a) 解像度チャートの RGB 画像 (b) 解像度チャートの深度マップ



(c) 深度の変動が大きい対象の RGB 画像 (d) 深度の変動が大きい対象の深度マップ

図 4 シミュレーション対象

#### 4. 実験

本研究では、提案手法の妥当性を数値シミュレーションにより確認した。シミュレーションを行う上でいくつかの条件を仮定した。一般に位相差 ToF 法で物体を撮影すると、外乱光の影響を受ける。これはあらかじめ外乱光の値を計測しておくことで除去が可能だが、外乱光は常に変化するもので厳密には外乱光検出で検出した外乱光と各時間窓に対応する物体輝度画像に含まれる外乱光とは、差が存在する。今回のシミュレーションでは、この部分を理想化して外乱光に変化がないものとした。また本シミュレーションでは、撮影の対象として CG レンダリングソフトウェアを用いて 3D モデルを作成し、得られる深度画像と RGB 画像から擬似的な各時間窓に対応する物体輝度画像を計算した。対象としては、解像度チャートが深度マップに現れるように作成したシーンと提案手法の苦手とする深度の変動が大きいシーンを利用した。図 4 に対象のシーンを示す。

シーンを低解像度空間で観測し、縦横  $2 \times 2$  倍の超解像を目的とした。この際に対象を符号化して観測するために、サブ画素を均等に観測するように設計した図 5 に示すような点像分布関数を用いた。図 6 および 7 に高解像度な時間画像の真値と実際に観測される低解像度な符号化画像を示す。

輝度画像の撮像過程は、十分高解像度な空間での点像分布関数の畳み込み演算、イメージセンサの標準化によるダウンサンプリングの実装、設定深度に基づく時間窓の適用、及びノイズ印加によりシミュレーションした。計測画像の画素数は  $240 \times 320$ 、再構成画像の画素数は  $480 \times 640$  と設定した。計測ノイズとして、40dB の加法性白色ガウスノイズを計測画像に印加した。再構成処理は、式 (6) の最

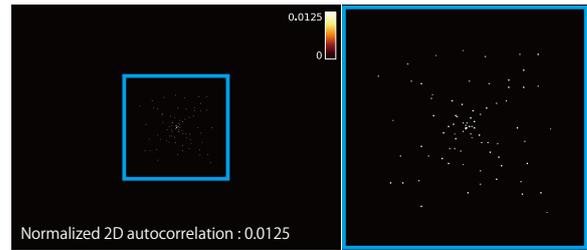
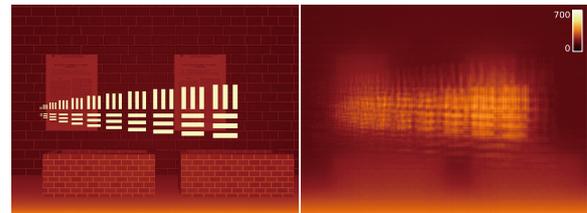
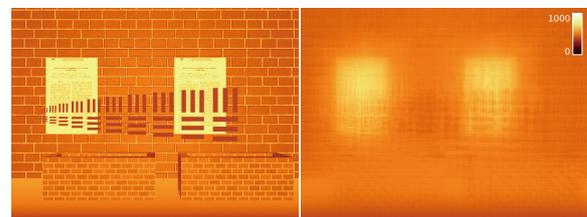


図 5 設計した点像分布関数 (スポット数:80)

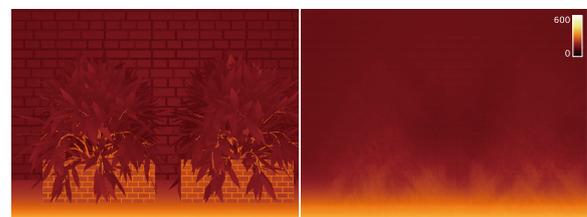


(a) 前半の時間窓に対する画像

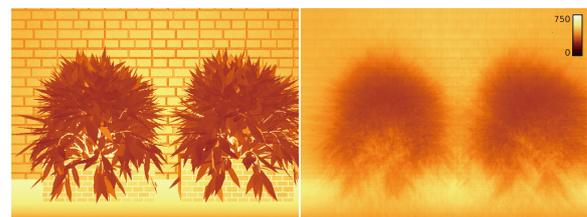


(b) 後半の時間窓に対する画像

図 6 解像度チャートの各時間画像の高解像度画像 (左, 擬似的真値) と観測される低解像符号化画像 (右)



(a) 前半の時間窓に対する画像



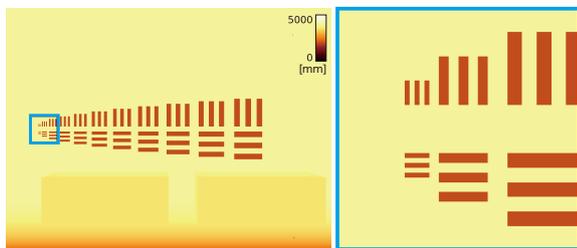
(b) 後半の時間窓に対する画像

図 7 深度の変動の大きい対象の各時間画像の高解像度画像 (左, 擬似的真値) と観測される低解像符号化画像 (右)

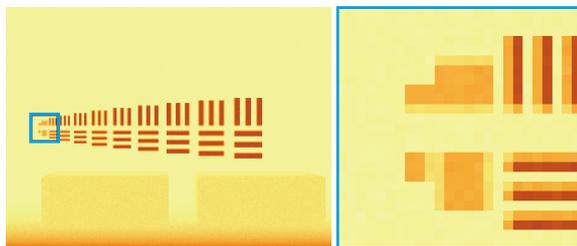
小化問題を反復アルゴリズムにより解くことで実装した。アルゴリズムとして TwIST [9] を用いた。また、精度向上のため、適応的な学習係数や運動量を考慮した解法である Adam [10] を併用した。再構成における初期解としては、低解像度にダウンサンプリングした対象を Bicubic 補間によって高解像度化した深度マップを利用した。再構成時における二次元全変動の係数  $\tau$  は  $1e-5$  を初期値として徐々

に小さくなるように動的に設定した。

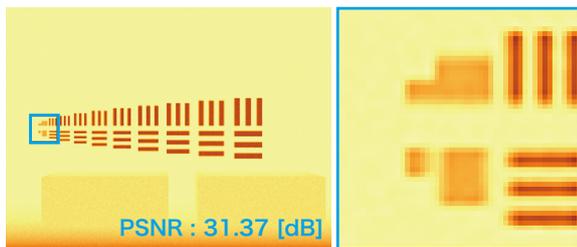
図8および図9に深度マップの真値、ノイズの印加された低解像度深度マップから Bicubic 補間により得た高解像度深度マップ、及び提案手法による再構成で得た高解像度深度マップを示す。定量的な評価としては、ピーク信号対ノイズ比 (PSNR) 比較で 0.83dB 低下している。これは、手前の花壇の部分を見ると補間アルゴリズムに対してノイズによる変動が大きいことから分かるように、スナップショット符号化計測では、ノイズが増幅されて再構成されることが原因である。一方で図8の拡大画像に注目すると、Bicubic 補間による高解像度化の場合にはエッジが鈍っている。またサブ画素情報が欠損しているため、3本の線が分離不可能となっている部分がある。一方で提案手法ではエッジが真値に近い状態で残っている上に、補間では分離不可能な情報まで分離できている。



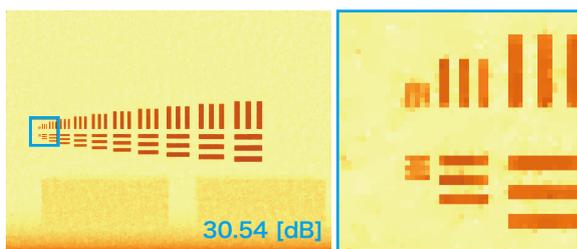
(a) 真値



(b) 低解像度観測深度マップ



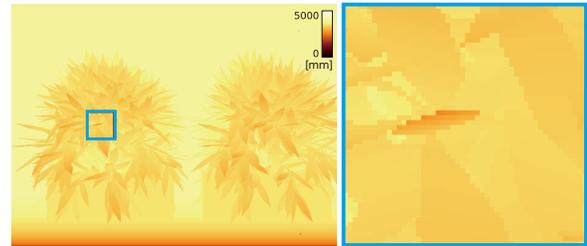
(c) Bicubic 補間



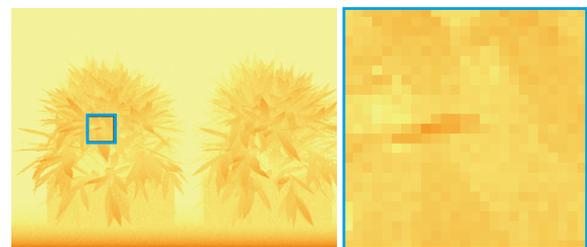
(d) 提案手法

図8 解像度チャートの各再構成結果

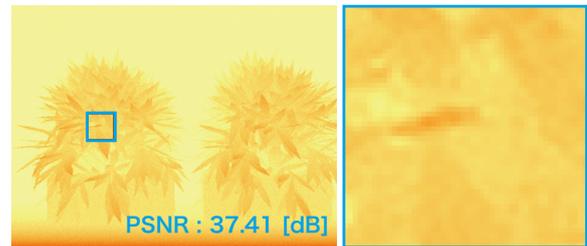
一方で図9に示すような深度の変動が大きい対象に対しては、定量的な評価では、PSNR 比較で 6.69dB 低下した。これは二次元全変動による正則化では対象シーンをスパース表現できないためだと考えられる。このような対象に対して高い精度の再構成結果を得るには、二次元全変動以外の正則化を用いる必要がある。



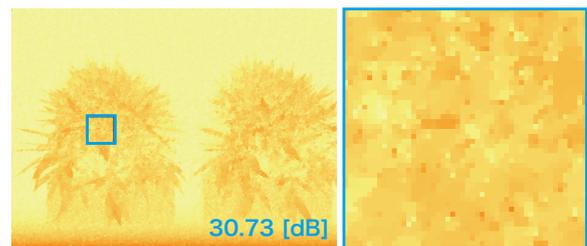
(a) 真値



(b) 低解像度観測深度マップ



(c) Bicubic 補間



(d) 提案手法

図9 深度の空間変動が大きい対象の各再構成結果

## 5. おわりに

本研究では、点像分布関数を設計し、深度画像の二次元全変動によるスパース性を利用した圧縮センシングを用いることでシングルショットでのデジタル超解像 ToF イメージングを提案した。この手法を用いることにより、PSNR 比較では二次元全変動によるスパース表現が有効な対象に対して補間アルゴリズムと比較して 0.83dB の低下

を確認したが、従来の超解像なしの ToF イメージングで取得した深度マップを、補間アルゴリズムによって高解像度化した場合には、サブ画素情報の消失の影響で分離不可能な画素が分離可能であることを確認した。一方で二次元全変動によるスパース表現が有効ではない対象に対しては補間アルゴリズムと比較して 6.69dB の精度低下を確認した。分解能の点でも優位性を確認することはできなかった。

以上のことから本研究の提案手法は正則化に二次元全変動を用いる場合、二次元全変動によるスパース表現が有効な場合には補間と比較して有用性が見込めるが、二次元全変動によるスパース表現が有効ではない場合、別の正則化を用いなければ補間と比較して有用性は見込めない。しかし、物理的な設計を考慮すれば、本研究で提案する手法は、光学系の瞳面に設計した点像分布関数を表現する位相変調素子等を置くだけで実現が可能であるため [11]、仮に超解像の困難な対象に対しても素子を着脱することにより対応が可能である。

今後の課題は、符号化イメージングのノイズに対する頑健性の向上させるため、観測結果に対して符号化された画像に特化させたデノイズ処理を追加することで、ノイズのある環境下でも再構成精度の低下しないモデルを構築することが挙げられる。

## 謝辞

本研究は JST 創発的研究支援事業 (JPMJFR206K) の助成を受け実施された。

## 参考文献

- [1] Kawahito, S., Halin, I. A., Ushinaga, T., Sawada, T., Homma, M. and Maeda, Y.: A CMOS time-of-flight range image sensor with gates-on-field-oxide structure, *IEEE Sensors Journal*, Vol. 7, No. 12, pp. 1578–1586 (2007).
- [2] Monakhova, K., Yanny, K., Aggarwal, N. and Waller, L.: Spectral DiffuserCam: Lensless snapshot hyperspectral imaging with a spectral filter array, *Optica*, Vol. 7, No. 10, pp. 1298–1307 (2020).
- [3] Donoho, D.: Compressed sensing, *IEEE Transactions on Information Theory*, Vol. 52, No. 4, pp. 1289–1306 (online), DOI: 10.1109/TIT.2006.871582 (2006).
- [4] Stern, A.: *Optical Compressive Imaging*, CRC Press (2017).
- [5] Li, F., Chen, H., Pediredla, A., Yeh, C., He, K., Veeraraghavan, A. and Cossairt, O.: CS-ToF: High-resolution compressive time-of-flight imaging, *Optics express*, Vol. 25, No. 25, pp. 31096–31110 (2017).
- [6] Sitzmann, V., Diamond, S., Peng, Y., Dun, X., Boyd, S., Heidrich, W., Heide, F. and Wetzstein, G.: End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging, *ACM Transactions on Graphics (TOG)*, Vol. 37, No. 4, pp. 1–13 (2018).
- [7] 安富啓太, 川人祥二: Time-of-Flight カメラ, 映像情報メディア学会誌, Vol. 70, No. 11, pp. 880–885 (オンライン), DOI: 10.3169/itej.70.880 (2016).
- [8] Rudin, L. I., Osher, S. and Fatemi, E.: Nonlinear total variation based noise removal algorithms, *Physica D: Nonlinear Phenomena*, Vol. 60, No. 1, pp. 259–268 (online), DOI: [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F) (1992).
- [9] Bioucas-Dias, J. M. and Figueiredo, M. A. T.: A New TwIST: Two-Step Iterative Shrinkage/Thresholding Algorithms for Image Restoration, *IEEE Transactions on Image Processing*, Vol. 16, No. 12, pp. 2992–3004 (online), DOI: 10.1109/TIP.2007.909319 (2007).
- [10] Kingma, D. P. and Ba, J.: Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980* (2014).
- [11] Boominathan, V., Adams, J. K., Robinson, J. T. and Veeraraghavan, A.: Phlatcam: Designed phase-mask based thin lensless camera, *IEEE transactions on pattern analysis and machine intelligence*, Vol. 42, No. 7, pp. 1618–1629 (2020).