

構造化ライトフィールドを用いた高速距離画像計測

松 本 卓 也^{*1} 奥 寛 雅^{*2} 石 川 正 俊^{*1}

High-Speed Real-Time Depth Estimation by Projecting Structured Light Field

Takuya Matsumoto^{*1}, Hiromasa Oku^{*2} and Masatoshi Ishikawa^{*1}

Real-time depth estimation is used in many applications, such as motion capture and human-computer interfaces. However, conventional depth estimation algorithms, including stereo matching or Depth from Defocus, use optimization or pattern matching methods to calculate the depth from the captured images, making it difficult to adapt these methods to high-speed sensing. In this paper, we propose a high-speed and real-time depth estimation method using structured light that varies according to the projected depth. Vertical and horizontal stripe patterns are illuminated, and focus is placed at the near and far planes respectively. The distance from the projector is calculated from the ratio of the amount of the blur due to defocus(bokeh) between the stripes projected on the object surface. This method needs no optimization, and the calculation is straightforward. Therefore, it achieves high-speed depth estimation without parallel processing hardware, such as FPGAs and GPUs. A prototype system of this method calculates the depth map in 0.2 ms for 2,500 sampling points using an off-the-shelf personal computer. Although the depth resolution and accuracy are not on par with conventional methods, the background and the foreground can be separated using the calculated depth information.

Key Words: Range Sensor, Depth from Defocus, High-speed Tracking, Saccade Mirror, Structured Light Field

1. 序 論

最近, Kinect [1] や Leap motion [2] といったデバイスが人気を博しているように、動く三次元シーンの復元が注目されている。その応用は PC やゲームなどの入力インターフェイスといった民間向けのものに加えて、車載カメラやロボットの目として用い物体認識を行ったりと重要な役割を果たす。ロボットの目といった、後に環境に応じたフィードバック制御を備える系に距離計測を適用する際には、その処理速度が求められる。高速な計測が必要となる理由として、1) 露光時間内に対象物体が移動することによって生じるモーションブラーを防ぐこと、2) 計測開始から終了までの時間を短くすることによって、後に続くフィードバック制御をより正確で安定なものとすること、が挙げられる。

現存するリアルタイム三次元計測において代表的な手法の一つは、プロジェクタ・カメラ系を用いて、投影した画像の視差から奥行き距離を復元する能動ステレオ法[3] である。この手法は単一画像からの距離推定が可能であるため、高速環境の計測

にしばしば用いられているが、距離復元の際に対応点探索を行う必要がある。これらのリアルタイム計測手法のスループットは代表的なもので 30~100 [fps] であり、さらなる高速化を行うためには、対応点の探索範囲を制限することで 200 [fps] を達成した立石ら[4] の例や、FPGA を用いた独自回路を構成した Watanabe ら[5] の例のように、並列化の部分を改善するというアプローチが取られてきた。ステレオ法は本来、静止画像に対して、正確さを重視した距離画像計測を行うことを目的として考案された手法であり、特別なハードウェアを用いずに、一般的な計算機上でリアルタイム計測を行う目的には適していない。

ほかにも、撮像の際にピントがあてていないことから生じるボケを用いた Depth from Defocus (DFD) と呼ばれる手法も存在する[12]。特に、一定のパターン光を計測対象に投影し、その模様が撮像時にどうぼけるかという情報から距離を復元する能動 DFD 法では、30 [fps] のリアルタイム計測を可能とした例もある[9]。しかし DFD 法も、ボケの情報を求めるためにフーリエ変換を行う必要があり、少なくない計算量を要するため、これ以上の高速化は見込めない。DFD においては、堀田ら[11] の手法のように、投影系において円形ではない特殊な形状の瞳を用いることで、ボケによって画像の高周波成分を失っても情報の復元を効率的に行い、より正確な距離計測を可能とした手法も提案されている。また、ドットパターンのボケ量による大きな変化を距離と対応付けて対象までの距離を復元する手法も

原稿受付 2014 年 12 月 26 日

*¹ 東京大学大学院情報理工学系研究科

*² 群馬大学大学院理工学府

*¹The University of Tokyo

*² Gunma University

■ 本論文は新規性（要素分野）で評価されました。

提案されている[10]。しかし、これらの手法では、ボケの情報と距離との対応付けをマッチングによって得る必要があり、リアルタイム計測には向かない。高速でリアルタイムに距離画像計測を行うためには、ワンショットで、マッチング等の最適化を用いず、低計算量で距離情報を復元できる手法が求められる。

本研究では、投影系によって空間に静的な奥行き情報を埋め込むことで、非常に小さな計算量によって距離を計測するという、高速性を特に重視した新しい距離画像計測手法を提案する。これまで、空間における光線の全情報を記述する関数としてPlenoptic Function [7] やライトフィールド (Light field) [8] と呼ばれる関数が考案されてきたが、それらは空間中の光線情報を取得し保存することを目的としていた。本稿では、投影距離によって画像が変わらるような光をライトフィールドと呼ぶこととし、その投影によって空間に奥行き情報を埋め込むことを考える。提案手法を用いることで、通常の計算機上でも十分に高速な計算速度で距離画像の計測が行えることを示し、それによって可能となる高速なフィードバック制御システムの一例を示す。

2. 原理

2.1 ライトフィールドの投影

本手法では特に画像のボケを用いることで、奥行き方向に構造を持った光を照射することを考える。空間中のある面で合焦するよう投影された画像は、実際の投影面が焦平面と一致していれば焦点のあった鮮明な画像となるが、それらが異なっている場合はボケを生じる。通常、計測においてボケは情報量の落ちるノイズとみなされるが、ボケの量は焦平面からの距離に依存して変化するという性質を利用し、Depth from Defocus と呼ばれる距離計測が可能となる。しかし、静的な1枚の画像を投影するだけでは、投影面が焦平面の前後どちらにあるかがボケの情報からでは推定できない。

そこで本手法では、二つの異なる画像を、投影装置から異なる2距離に合焦させることで奥行き方向に向きを持った光線場を投影することを考えた。計測対象物体が二つの焦平面の間にいると仮定すれば、物体上に投影された二つの画像を分離し、それぞれのボケ量を抽出することで空間中での位置を推定することが可能となる。

具体的にどのような画像を投影するかは、画像の分離やボケ抽出のしやすさ、および投影系の複雑さを考慮することで様々な選ぶことができるが、本手法では最も単純な縦方向の縞と横方向の縞を用いることにした。これらの縦横縞の組を合わせて投影パターンとする。Fig. 1に、投影系に近い側に縦縞が、遠い側に横縞が結像するイメージを表す。

縦横縞の比を用いることには、前後の識別のほかにも利点がある。画像におけるボケの量は、そのコントラストを計算することで求めることができるが、一つの縞のみのコントラストを計算するだけでは、照射対象の持つ光の反射率によってそのコントラストの値が変化し、ボケの情報のみを取り出すことができない。すなわち、反射率の高い物体はコントラストが高く、反射率の低い物体はコントラストが低く見積もられてしまうのである。本手法では、同じ面に投影した二つの縞について、それぞれコントラスト量を計算してその比を取ることで、対象物体

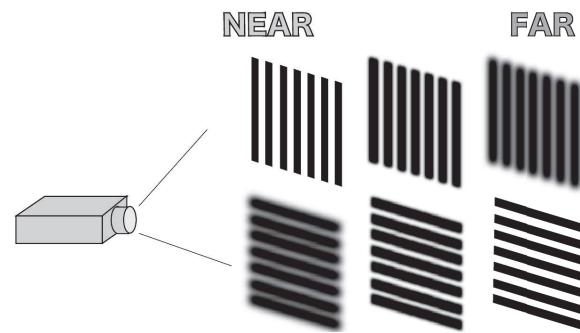


Fig. 1 Projected vertical/horizontal stripes with blurring (bokeh)

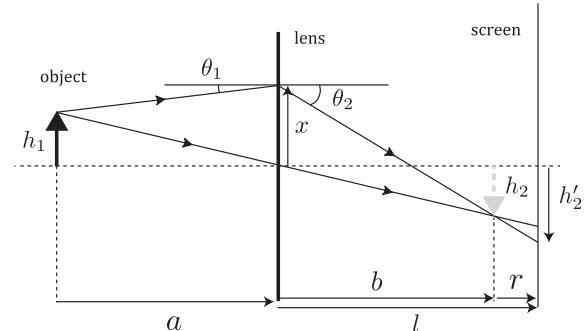


Fig. 2 Model of projected lightfield and variable setting

の持つ性質によるコントラストの変化を低減することができる。

本手法は、距離の情報を投影系のみによって埋め込み、撮像およびその後のプロセスを極力減らすことで高速化を図るアプローチである。また、主に投影系によって計測を行うことで、高速性に加えて、実験システム構築の簡便性もたらす。カメラとプロジェクタ、または2台のカメラ間の空間的関係から距離を復元する従来の手法とは異なり、両者を厳密に位置合わせ（キャリブレーション）しておく必要はない。そのため、ある程度は自由な位置から撮像しても計測を行うことができる。今後、照明系をさらに発展させることによって、より視点の拘束のない計測が可能になると考えられる。

2.2 実験系のモデル化

手法の性能を解析し、パラメータを適切に設定するため、パターンの投影とそのボケをモデル化する。

2.2.1 ボケの大きさ

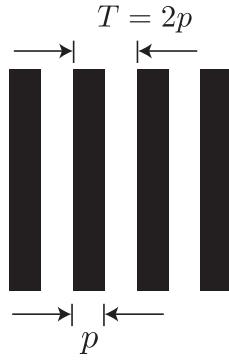
ボケの大きさを考察するため、Fig. 2のように投影モデルを定める。光学系から距離 a 、高さ h_1 の位置から出た光線が、光軸から距離 x の位置で入射角 θ_1 にて光学系に入射し、 θ_2 の方向に射出された後、距離 b の位置で高さ h_2 に合焦し、距離 l で h_2 の高さに照射されるとする。変位はすべて右向きと上向きを正とし、角度は光軸から反時計回りを正とする。

各変数の関係式は以下のようになる。ただし近軸領域を仮定して $\sin \theta = \tan \theta = \theta$ とする。

$$x = h_1 + a\theta_1, \quad h_2 = x + b\theta_2, \quad h'_2 = x + l\theta_2,$$

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f}, \quad \theta_2 = \theta_1 - \frac{x}{f}$$

これより、 h'_2 を x の関数として表すと、

Fig. 3 Vertical stripes with width p

$$\begin{aligned} h'_2(x) &= \frac{x}{b}(b-l) - \frac{l}{a} h_1 \\ &= -\frac{xr}{b} + h'_2 \end{aligned} \quad (1)$$

となり、ボケの大きさ d は、光学系の開口径を D とすると

$$\begin{aligned} d &= |h'_2(x_{\max}) - h'_2(x_{\min})| \\ &= \frac{D}{b}|r| \end{aligned} \quad (2)$$

と表せる。

2.2.2 像倍率

一般に、画像の投影においてもカメラによる撮像と同様に、合焦の距離はレンズの公式によって決まる。

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f} \quad (3)$$

投影像の拡大率は投影距離 b とパターンとレンズ間の距離 a の比によって決まるが、十分遠方に投影する場合は $a \simeq f$ とみなせるため、倍率 β は $\beta = |b/a| \simeq |b/f|$ と表される。

Fig. 3 に示すような幅 p 、周期 $T = 2p$ の縞を光学系を通して投影すると、結像面では幅 $w = (b/f)p$ の縞として投影される。以後投影面での周期を $T' = 2w$ と表す。

2.2.3 点像分布関数

一般に光学系による結像は、点像分布関数 (Point Spread Function) によって表現される。点像分布関数は、点光源が光学系によってどのように拡散するかを表し、光学系のインパルス応答と考えることができる。円形瞳を持った収差のない光学系を仮定した場合、焦平面での点像分布関数は光軸に関して対称で、動径方向には一次のベッセル関数 $J_1(x)$ を用いて表されることが知られている [14]。さらに、投影面が焦平面と一致しないことによりこの画像がボケることを考え、ボケによる点像分布関数は、焦平面からの距離によって決まるボケ幅 d を 3 倍とした二次元ガウス関数

$$g_d(\mathbf{x}) = \exp\left(-\frac{\|\mathbf{x}\|^2}{2(d/3)^2}\right) \quad (4)$$

を用いて表すこととする。

2.3 距離情報の復元

投影されたパターンから距離情報を復元する手法を説明する。

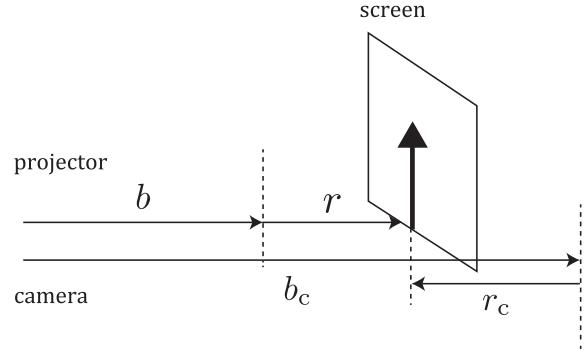


Fig. 4 Projector-camera model

2.3.1 コントラストの計算

カメラ画像上の各計測点 $P(x, y)$ に対して、 $P(x, y)$ を中心にして縦横方向にそれぞれ N ピクセルを切り出して一次元配列 $l_x[n]$, $l_y[n]$ とし、その配列から距離情報を推定する。距離情報は縞模様のボケ量に対応するが、ボケがローパスフィルタとして作用することによって縞模様は縞と同じ周期を持つ正弦波に近づき、その周期に対応する空間周波数成分のコントラスト値を求めてボケの量を推定する。縞の空間周波数を与えるフーリエ係数の次数を k として、コントラスト量 Θ を以下の式で定義する。

$$\Theta = \left| \frac{1}{\sum l[n]} \sum_{n=1}^N l[n] \cdot e^{-jkn} \right|. \quad (5)$$

ただし j は虚数単位、 N は系列 $l[n]$ の長さであり、周辺光量による影響を取り除くため、系列の和で正規化を行っている。 $l_x[n]$ と $l_y[n]$ からそれぞれ Θ を求め、その比から距離の推定を行う。

縦方向と横方向のコントラスト Θ_v および Θ_h は、二つの縞が観測できる区間内では、距離の増加に伴って一様に増加または減少するため、その比 Θ_v/Θ_h と距離を一対一に対応づけることが可能で、あらかじめ比と距離のキャリブレーションを行うことで、計測されたコントラスト量から距離情報を復元することができる。

2.3.2 ボケによるコントラスト量の変化

投影位置によって変化するボケ量によって、コントラストがどのように変化するかを考察するため、**Fig. 4** のような投影-撮像系モデルを仮定する。投影系、撮像系からそれぞれの焦平面までの距離を b , b_c とし、焦平面から投影面までの符号付き距離をそれぞれ r , r_c とする。

一次元のボケ関数を

$$g_d(x) = \exp\left(-\frac{x^2}{2(d/3)^2}\right) \quad (6)$$

とすると、 $b(x)$ のフーリエ変換 $G(\nu)$ は以下の式で得られる。

$$G_d(\nu) = \exp\left(-\frac{\nu^2}{2(3/d)^2}\right) \quad (7)$$

式 (5) によって定義されるコントラスト値は、系列 $l[n]$ をフーリエ変換した信号の、空間周波数 $k = 2\pi/T'$ での値の絶対値に対応するが、ボケ関数がその空間周波数に与える影響は、

投影位置によってボケ関数および投影縞の像倍率が変化することを考慮し、両者が撮像系から同じ距離にあるとして計算する必要がある。ボケ幅 $d = D|r|/b$ は投影面での幅であることに注意すると、周期 $T = 2p$ の縞を距離 $b+r$ の面に投影したとき、投影面での周期は $\hat{T}' = 2p(b+r)/f$ となる。

以上より、ボケ幅 d を持ったボケ関数 $g_d(x)$ による、投影面上での空間周波数 $2\pi/\hat{T}' = \pi f/p(b+r)$ のコントラスト量変化の割合を A とすると、

$$\begin{aligned} A &= G_d \left(\frac{\pi f}{p(b+r)} \right) \\ &= \exp \left\{ -\frac{\pi^2}{18} \left(\frac{fd}{p(b+r)} \right)^2 \right\} \\ &= \exp \left\{ -\frac{\pi^2}{18} \left(\frac{Df}{bp} \frac{r}{b+r} \right)^2 \right\} \quad (8) \end{aligned}$$

を得る。式 (8) から分かるように、幅 p の縞を投影した際のコントラスト値の変化は、投影系の開口径 D 、焦平面までの距離 b 、焦平面と投影面の距離 r によって決まる。

これはカメラによる撮像においても同様で、撮像系のパラメータ D, b, r によって、撮像時のボケによるコントラスト変化を表すことができる。

また、パターンの投影・撮像の際には、光学系の性能によりコントラストが低下する。厳密には撮像光学系のMTF（変調伝達関数）を求めて計算する必要があるが、収差のない円形瞳からなる光学系を仮定した場合、コントラスト低下は以下の式で与えられる [13]。

$$H(\mu) = 1 - \frac{\mu}{\pi} \sqrt{1 - \left(\frac{\mu}{2} \right)^2} - \frac{2}{\pi} \sin^{-1} \left(\frac{\mu}{2} \right). \quad (9)$$

ただし、 μ は規格化空間周波数で、 $\mu = D/(\lambda f)$ 、 $0 \leq \mu \leq 2$ である。

以上より、一つの縞模様を投影した際、その投影距離によって変化するコントラスト量 Θ の理論値を見積もることができる。

3. 評価実験

提案手法の持つ性能を評価し、本手法を用いて実時間距離計測実験を行った。

3.1 実験装置

パターンの遮蔽と取得画像の回転・スケール変換を不要とするため、ビームスプリッタを用いて投影系と撮像系の光学中心を一致させることで、Fig. 5 に示すような同軸系を構成した。これにより、得られる画像上の縞のピクセル周期は常に一定で、式 (6) において縞の空間周波数を与える添字 k は一意に求まる。

投影系における照明は 250 [W] メタルハライドランプ、投影レンズは焦点距離 20 [mm]、絞り開放 f/2.8 のものを用い、撮像素子は XIMEA 社の xiQ MQ003CG-CM、解像度は 480 × 640 ピクセルであり、フレームレートは最高で 500 [fps] である。撮像レンズは焦点距離 50 [mm] のものを用いた。画像処理は Intel Core i7-3630QM 2.4GHz を備えた PC 上で、OpenMP にて 8 コアによる並列分散処理を行った。

投影パターンは、光学チャートとして用いられるロンキー

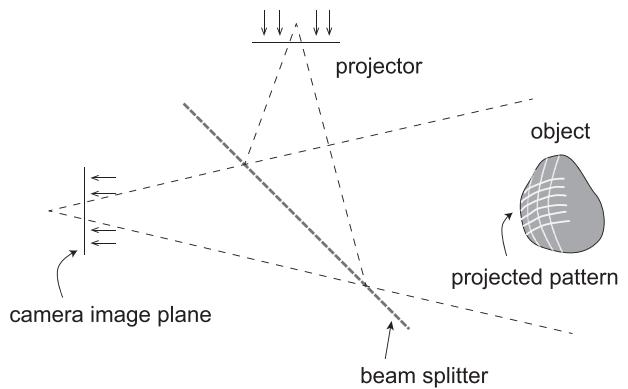


Fig. 5 Coaxial setting of projector and camera

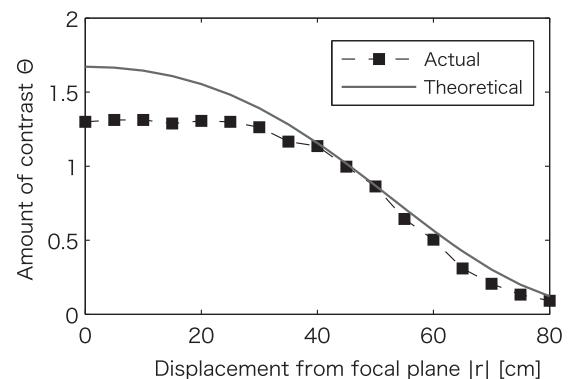


Fig. 6 Relation between depth $|r|$ and amount of contrast Θ

ルーリングを用いた。ロンキールーリングは Fig. 3 に示すような 1 方向の縞が等間隔で刻まれているガラス板である。今回は $T = 0.05$ [mm] のものを 2 枚用いた。

2 枚のロンキールーリングを 0.25 [mm] の間隔をあけて配置し、投影先で各縞が別の位置に焦平面を形成するようにした。

3.2 距離に対するコントラスト量の変化

2.3.2 節に示した、距離に対するコントラスト量の変化を実測し、理論値と比較した結果を Fig. 6 に示す。投影系と撮像系の焦点位置を等しく $b = 2,000$ [mm] とし、撮像系の f/8.0 としたとき、焦点位置からの変位 r を負の方向へ変化させた際のコントラスト応答を示している。なお光の波長 $\lambda = 500$ [nm] とした。図より、焦平面付近で $|r|$ が小さい時を除き、理論値と実測値が一致していることが分かる。 $|r|$ が小さいときに実測値が横ばいになるのは、撮像系の開口径が小さく被写界深度が広くなるため、ボケ量を撮像素子が識別できていないためだと考えられる。2 枚の縞のコントラストの比を用いる場合には、一方が焦平面付近であっても、もう一方のコントラストから距離の変化を識別することができるため、問題にはならないと考えられる。

3.3 対象物体の反射率がコントラスト量に与える影響

本手法では、投影された縦縞と横縞のコントラスト量の比を取ることで、対象物体の反射率によるコントラストの変化の影響を低減させ、対象の材質に依らない計測を可能とする。反射率の異なる三つの素材として、

- (i) 白色ボール紙



Fig. 7 Left:Green pencase, Right:Hand

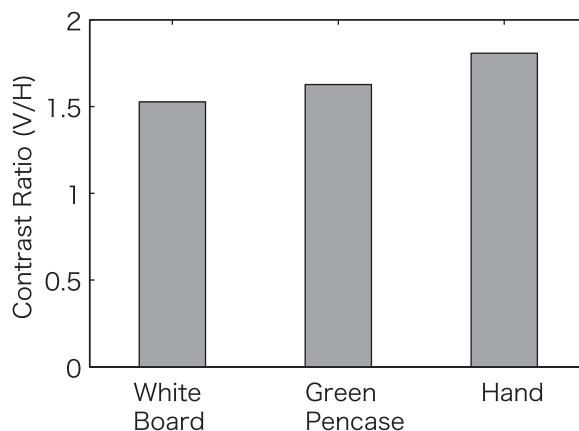
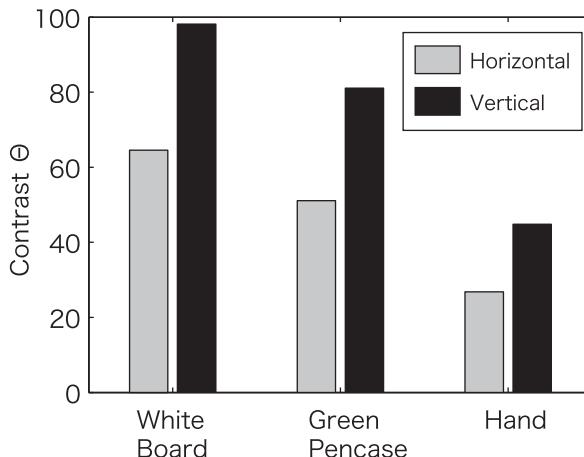


Fig. 8 Amount of contrast(upper) and its ratio(lower) vs. the material of objects

(ii) 緑色布地ペンケース (Fig. 7 左)

(iii) 手 (Fig. 7 右)

を用意し、パターンを投影しコントラストを算出した。なお反射率は (i) から (iii) の順に小さくなる。

本手法で用いた投影パターンを照射し、同じ位置に配置した各対象上の 1 点でコントラスト量を計算し、時間平均した値を Fig. 8 上に示す。いずれの物体上でも縦縞のほうがコントラストが強く出ているが、反射率の低い物体ほどその値は小さくなり、ボケではなく光の吸収によってコントラストが低下していることが分かる。それに対して縦縞と横縞のコントラストの比を取ることによって、Fig. 8 下に示すように反射率の影響を低

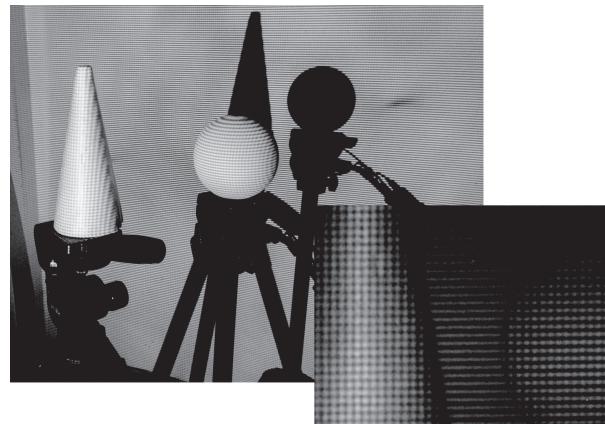


Fig. 9 Left: Experiment overview, Right: Image from camera

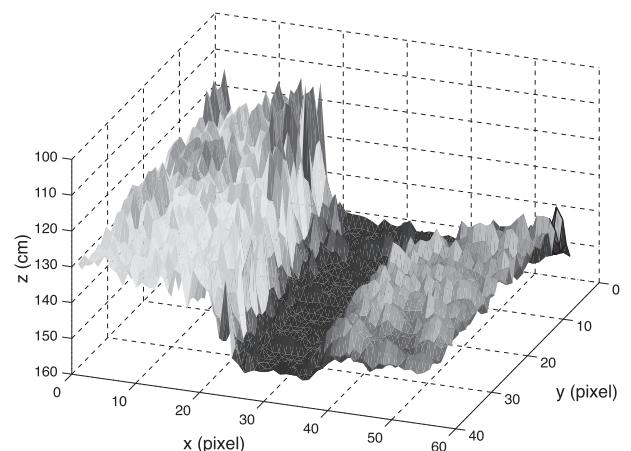


Fig. 10 Result of depth map reconstruction

減することができる事が分かる。

Fig. 8 では反射率の低い物体ほど縦縞と横縞の比が大きく出ているが、これは、手のような物体上では表面下散乱等の影響で、縦縞と横縞が独立に反射せず、例えば縦縞を構成する縞の明線が乱反射を起こし、横縞の暗線のコントラストを下げてしまうなどの影響が考えられる。

3.4 距離計測実験

提案手法を用いて 1 枚の画像から距離計測を行った結果を Fig. 9, Fig. 10 に示す。撮像系から 1,200 [mm] (手前側) の位置に直円すい、1,400 [mm] (奥側) の位置に球が配置してあり、Fig. 9 左のように投影光を照射した部分を、同軸に配置したカメラで撮影し、Fig. 9 右のような画像を得る。カメラの計測領域は、2,000 [mm] 先で約 200×150 [mm] となり比較的狭いが、これは投影縞の細さを識別できるよう撮像系の倍率を設定する必要があるためで、より太い縞を投影し、撮像系では広角のレンズを用いることで画角を広げることができる。

得られた画像から縦横 10 ピクセル間隔で計 2,500 点をサンプリングし、提案手法によってコントラスト比を計算すると Fig. 10 のようになる。なお、切り出した一次元配列の長さ N は 32 ピクセルとし、この範囲におよそ 2 周期の縞が投影される。そのため、縞を表すフーリエ係数の次数は $k = 2$ と設定した。

Table 1 Relation between exposure time and standard deviation (f/4.0)

Exposure time(ms)	Contrast(avg.)	SD/avg. (%)
20	162	3.50%
15	171	3.49%
10	191	4.04%
5	100	11.8%

Table 2 Relation between exposure time and standard deviation (f/8.0)

Exposure time(ms)	Contrast(avg.)	SD/avg. (%)
20	95	9.86%
15	95	11.5%
10	184	15.3%
5	201	22.1%

Fig. 10 から分かるように、球や円すいの丸みといった細かい奥行き形状は判別できないものの、計測範囲内での前後関係は識別できていることが分かる。

また 2,500 点のサンプリングに対して 8 コア並列処理を用いることで 0.2 [ms] での計算速度を達成した。これはフレームレートに換算すると 5,000 [fps] であり、撮像のフレームレートと比べても非常に高速に距離計測を行うことが可能である。

3.5 計測精度

本手法の持つ距離計測の精度について実験を行った。本手法では投影されたパターンのコントラストから距離情報を復元しているため、光量が落ちるほど撮像素子のノイズに対する SN 比が低下する。そこで画像上の 1 点に対して約 2,000 フレームの間、縞 1 枚のコントラスト量を計算し、その平均値に対する標準偏差の割合を、撮像系の絞りが f/4.0, f/8.0 の 2 とおりの場合について計測した。結果を **Table 1**, **Table 2** に示す

表から分かるとおり、露光時間が短くなるに従ってコントラスト量の標準偏差は大きくなっています。500 [fps] に相当する 5 [ms] の露光においては、標準偏差は f/4.0 で 11.8%, f/8.0 で 22.1% まで増大する。そのため、フレームレートと計測の精度にはトレードオフの関係がある。

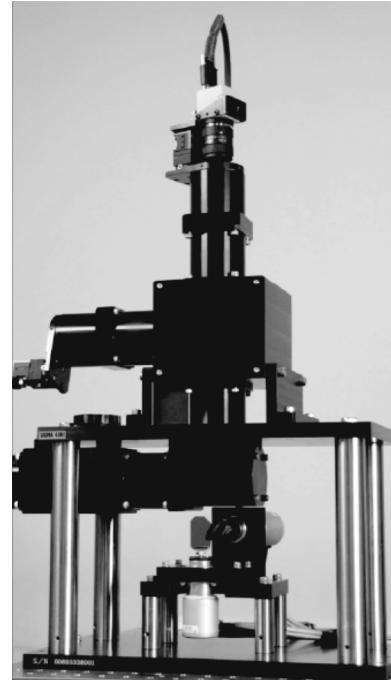
3.6 計算量

本手法では、精度は従来手法よりも低いが、FPGA 等の並列計算ハードウェアを用いることなく、2,500 点に対して 0.2 [ms] という非常に高速な計算速度を達成した。この高速性を達成する要因として、フーリエ変換を用いないことが挙げられる。Nayer ら [9] の手法に代表される通常の DFD による距離復元では、画像のフーリエ変換に対して演算を行い、ボケのパラメータ量を抽出している。一般にサイズ $S \times S$ の画像のフーリエ変換の計算量は、計算量 $\mathcal{O}(S \log S)$ の高速フーリエ変換を、各行、各列ごとに行うため、

$$2 \times S \times \mathcal{O}(S \log S) = \mathcal{O}(S^2 \log S) \quad (10)$$

である。例えば、画像サイズ 512×512 [pixel] とした場合、計算量はおよそ $\mathcal{O}(2^{20})$ となる。フーリエ変換を含む画像処理ではこの部分の計算量がボトルネックになり、大規模な並列計算の必要性を有する。

対して本手法では、計測点 1 点あたりの計算量は画像サイズ

**Fig. 11** High-speed gaze controller by rotation mirrors (Saccade Mirror)

によらず、定数である探索ウインドウ幅 N のみによって決まる。これを m 点に対して行ったとしても、全体の計算量は $\mathcal{O}(mN)$ で行える。

これは、計測システム設備等の制限から、精度ないしは解像度を落としたとしても高速な計測が行えることを示しており、例えば計算能力の低くない汎用 PC やスマートフォン上でのリアルタイム距離画像計測が行える可能性を見込めるところとなる。

3.7 距離計測を用いたフィードバック制御

本計測手法の持つ高速性を実証するために、高速・低レイテンシのフィードバック制御が必要とされるシステムにおいて距離計測を行った。

3.7.1 高速光軸制御ユニット

奥村らは、**Fig. 11** に示すサッカードミラー (Saccade Mirror) と呼ばれるデバイスを開発し、ms オーダーで光軸方向を制御し、高速飛翔物体を 1 [ms] の応答速度で追尾する「1 [ms] オートパン・チルト」というシステムを開発した [6]。このシステムは、高速カメラによって得られた画像における対象物体の位置を 1 [ms] ごとに計算し、物体の変位の差分に合わせてパン・チルトの 2 軸に配置されたガルバノミラーを回転させることによって視線方向を更新し物体を追尾している。

奥村らの手法では対象物体の認識に際して、HSV 空間での色認識を行い二値化することで、高速な物体検出を実現している。しかしこの手法では、背景が対象物体と同じ色をしている場合に物体の認識が行えないという欠点があった。一般的に動物体追跡は、手前側にある物体を対象に行われるため、物体までの距離を用いて物体検出を行うことは有用であるが、数 ms という低レイテンシで距離計測を行うことのできる手法はこれまでに存在しなかった。今回はこのシステムの物体認識に提案手法

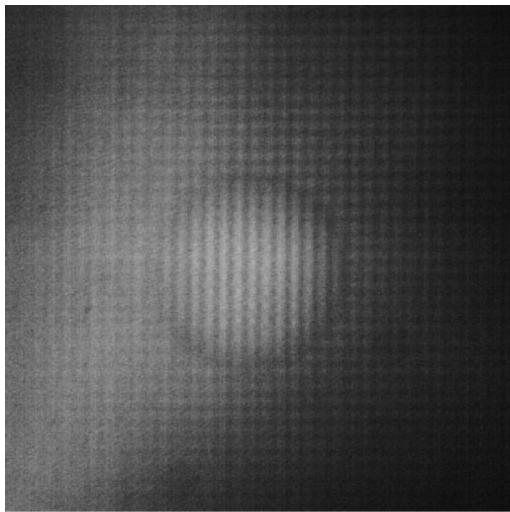


Fig. 12 Image from high-speed camera (2 [ms] exposure)

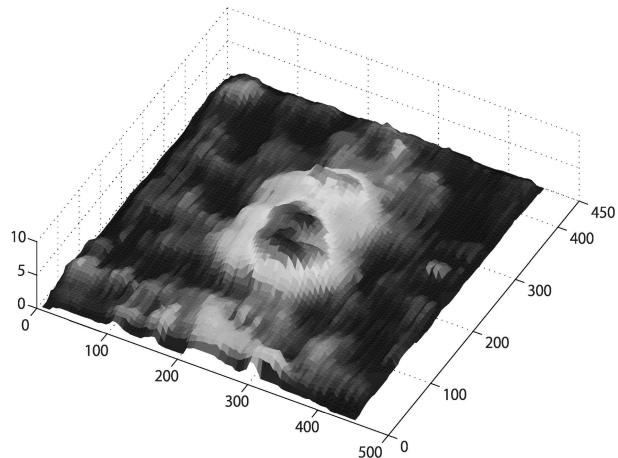


Fig. 13 Reconstruction of contrast ratio from camera image

を組み込み、動物体追跡を行った。

3.7.2 実験結果

本実験では、3.1節で用いた照明を照射した上で、Phototron社の高速カメラ IDP-Express R2000 によって 500 [fps] にて撮像を行う。サッカードミラーはハーフミラーを通して高速カメラと同軸に照明を投影することが可能である。投影レンズには焦点距離 35 [mm]、絞り f/1.8 のものを用い、撮像には焦点距離 25 [mm]、絞り f/2.8 のレンズを用いた。

本手法を組み込んだ、サッカードミラーによる高速動物体トラッキングは、以下の手順による。

- (1) 高速カメラにより、Fig. 12 に示すような 512×512 ピクセルの Bayer 画像を毎秒 500 フレーム取得する。
- (2) Bayer 画像をグレイスケールに変換する。
- (3) グレイスケール画像の中央 $1/4$ 部分を切り出し、8 ピクセルごとに提案手法で距離を推定する (Fig. 13)。
- (4) 一定の閾値を定め、距離画像を二値化する。
- (5) 二値化された画像の重心を対象物体の位置として更新し、PID 制御によってミラー角度を更新する。

投影系から約 130 [cm] 離れた位置で白色ボール紙を背景とし

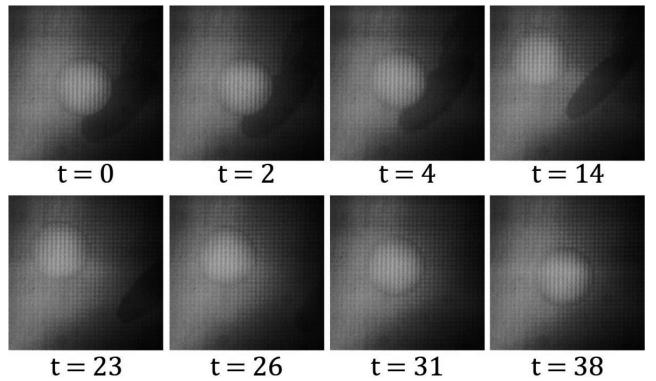


Fig. 14 Result of tracking of white ball

て、白色発泡スチロール製の球を打ち返した様子を Fig. 14 に示す。従来の色認識をベースにした画像処理では、このような同色の背景を有する環境でのトラッキングは非常に困難であったが、本手法で距離情報を基にトラッキングを行うことで、高速環境においても対象の追尾に成功している。

ボールがラケットに接触した瞬間を 0 フレームめとして、14 フレーム目には跳ね返ったボールが画像中心から離れるも、その後ミラーの駆動によって 38 フレームめには画像中心に復帰していることが分かる。

画像処理の速度は、高速カメラから得られる Bayer 画像のグレイスケール変換、および 1,024 点サンプリングの距離計測と重心位置推定で 1 [ms] 以内に収まっている。露光時間を 2 [ms] とし 500 [fps] で駆動しているのは、光量を稼ぎ SN 比を高めるためであり、照明の強度や撮像素子感度の改善を行えば、より高速なフィードバック制御が可能となる。

4. 結 論

本研究では、投影の距離によって画像が変化するライトフィールドを投影することで、撮像側でのプロセスを減らし、ワンショットかつマッチング等を用いずに奥行き情報を求める手法を提案し、その手法の性能を検証した。さらに評価実験を通して、手法の高速性が見込めるかを確認し、精度においても、マッチングを用いないため従来の距離計測手法に比べると劣るもの、前後関係を識別するには十分な奥行き分解能を有することを示した上で、2 [ms] の応答速度を持つ動物体トラッキングシステムへと組み込み、実際に前景背景の分離を非常に高速で行えることを確認した。

今後の課題として、手法の更なる解析を行い、性能をより向上させるための検討が挙げられる。本発表ではライトフィールドとしては非常に単純な、二つの縞を投影するのみであった。どのような画像を投影し、距離情報を復元するかはまだ大いに検討の余地が残されている。その点を効果的に改善することで、分解能や計測範囲といった性能は向上するものと考えられる。

参 考 文 献

- [1] Xbox 360 Kinect, <http://www.xbox.com/en-US/kinect>, 2010.
- [2] Leap Motion Controller, <http://www.leapmotion.com/>, 2013.
- [3] H. Kawasaki, R. Furukawa, R. Sagawa and Y. Yagi: "Dynamic scene shape reconstruction using a single structured light pat-

- tern,” IEEE Trans. CVPR 2008, pp.1–8, 2008.
- [4] 立石雅輝, 石山英俊, 梅田和昇: “マルチスボットレーザプロジェクタを用いた 200Hz 小型距離画像センサの構築”, 日本機械学会論文集 C 編, vol.74, no.739, pp.499–505, 2008.
- [5] Y. Watanabe, T. Komura and M. Ishikawa: “955-Fps Real-Time Shape Measurement of a Moving/Deforming Object Using High-Speed Vision for Numerous-Point Analysis,” Proc. of IEEE Int. Conf. International Conference on Robotics and Automation, pp.3192–3197, 2007.
- [6] K. Okumura, H. Oku and M. Ishikawa: “High-Speed Gaze Controller for Millisecond-order Pan/tilt Camera,” Proc. of IEEE Int. Conf. Robotics and Automation, pp.6186–6191, 2011.
- [7] E.H. Adelson and J.R. Bergen: “The Plenoptic Function and the Elements of Early Vision,” Computational Models of Visual Processing, pp.3–20, 1991.
- [8] M. Levoy and P. Hanrahan: “Light Field Rendering,” Proc. SIGGRAPH’96, pp.31–42, 1996.
- [9] S.K. Nayar, M. Watanabe and M. Noguchi: “Real-time focus range sensor,” IEEE Trans. PAMI, vol.18, no.12, pp.1186–1198, 1996.
- [10] F. Moreno-Noguer, P.N. Belhumeur and S.K. Nayar: “Active refocusing of images and videos,” ACM Trans. Graph., vol.26, no.3, 2007.
- [11] 堀田祐樹, 小野智司, 川崎洋, 木村誠, 高根靖雄: “符号化開口を用いたプロジェクタカメラシステムによるぼけを利用した三次元計測手法”, 電子情報通信学会論文誌 D, vol.J96-D, no.8, pp.1823–1833, 2013.
- [12] A.P. Pentland: “A New Sense for Depth of Field,” IEEE Trans. PAMI, vol.9, no.4, pp.523–513, 1987.
- [13] 左貝潤一: 光学の基礎. コロナ社, 1997.
- [14] J.W. Goodman: Introduction to Fourier Optics. McGraw-Hill, 1968.



松本卓也 (Takuya Matsumoto)
2013 年東京大学工学部計数工学科卒業。現在東京大学大学院情報理工学系研究科システム情報学専攻修士課程在学中。構造照明の投影による高速三次元計測に関する研究に従事。
(日本ロボット学会学生会員)



石川正俊 (Masatoshi Ishikawa)
1977 年東京大学工学部計数工学科卒業。1979 年東京大学大学院工学系研究科計数工学専攻修士課程修了。同年通産省工業技術院製品科学研究所入所。1989 年東京大学計数工学科助教授。1999 年東京大学現在、東京大学大学院情報理工学系研究科創造情報学専攻教授。2002, 2003 年東京大学総長特任補佐。2004 年度東京大学副学長。2005 年度理事・副学長。超並列・超高速ビジョン, センサフュージョン, メタ・パーセプションなどの研究に従事。工学博士。
(日本ロボット学会正会員)



奥 寛雅 (Hiromasa Oku)
1998 年東京大学理学部物理学科卒業。2003 年東京大学大学院工学系研究科計数工学専攻博士課程修了。2003~2005 年科学技術振興機構研究員。2005 年東京大学大学院情報理工学系研究科システム情報学専攻助手。2007 年同助教。2011 年同講師。2014 年群馬大学大学院理工学府准教授、現在に至る。高速光学デバイスと高速画像処理による映像制御に関する研究に従事。博士 (工学)。
(日本ロボット学会正会員)