

マルチタップCMOSイメージセンサを用いた ダイナミック照度差ステレオ法

依田 拓也^{1,a)} 長原 一¹ 谷口 倫一郎¹ 香川 景一郎² 安富 啓太² 川人 祥二²

概要: 照度差ステレオ法では異なる方向から光を照射した3枚以上の輝度画像と、その既知の光源位置から被写体の法線を推定することができる。しかし被写体が運動している場合、撮影した輝度画像間で対応が取れなくなり、推定される法線は正しくない。本研究では、マルチタップCMOSイメージセンサを用いたダイナミック照度差ステレオ法を提案する。マルチタップCMOSイメージセンサでは、フォトダイオードによって光から変換された電荷をピクセル内に蓄積しておくフローティングディフュージョンを複数備えており、それぞれ独立して電荷の蓄積・画像の読み出しができる。露光を分割して繰り返すことで1回の露光時間を短くでき、撮影される複数の画像はほぼ同時に撮影したとみなすことができる。イメージセンサの露光を異なる光源の発光タイミングと同期させると、異なる方向から被写体に光を照射した複数枚の画像をほぼ同時に撮影することができ、照度差ステレオ法によって動的物体に対しても法線を推定することが可能である。本稿では、試作マルチタップCMOSイメージセンサによって動的物体の法線推定が可能であることを示した。また、撮影における設定パラメータとしてマルチタップCMOSイメージセンサの露光時間の変化と露光回数の変化が照度差ステレオ法で推定される法線の精度にどの程度影響があるか比較する実験を行った。実験結果から露光回数が与える影響は無視できる範囲であり、露光時間が短くなるにつれ推定される法線に大きな誤差が現れることが確認できた。

1. はじめに

近年、3次元情報のセンシングは大きな注目を集めている。3次元計測を行うと、対象の物体がどこにあるか、どのくらいの大きさか、どのような形状をしているのかわかる。それらの情報はテレビゲームなどの入力デバイス [1] や製造業における製品の外観検査 [2]、文化財のデジタルアーカイブの作成 [3] などに応用されている。

3次元計測技術には画像処理を用いたものやレーザーによる計測など多くの手法がある。その中でも特にイメージセンサを用いた3次元情報のセンシングは低コスト、小型化が可能であるため多くの手法が提案されている。それらの手法では大まかに受動型のステレオ法 (Passive stereo) と照明を制御する能動型のステレオ法 (Active stereo) に大別できる。受動型のステレオ法にはステレオ法、能動型のステレオ法には Time-of-Flight [4] [5] や照度差ステレオ法 [6] がある。また、3次元情報を取得する技術にも、セ

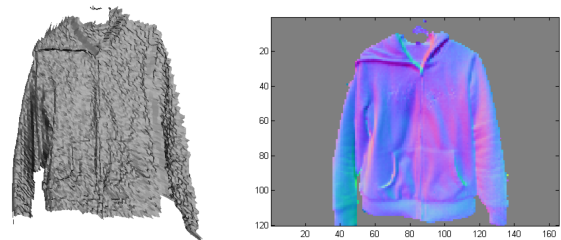


図 1: TOF から作成したメッシュ画像 (左) と照度差ステレオ法による法線 (右) [7]。照度差ステレオ法では被写体の表面を滑らかに推定できている。

ンサと被写体との距離を求める手法と、被写体の法線を推定する手法に分けることができる。上述したステレオ法や Time-of-Flight は被写体との距離を計測する技術であり、照度差ステレオ法は被写体の法線を推定する技術である。被写体との距離を求める方法では微分を行うことにより法線を求めることも可能であるが、センサと被写体の方向に誤差が含まれる場合、何らかの仮定のもとその誤差を修正しなくてはならぬ形状を取得することができない。しかし、照度差ステレオ法では直接法線を推定するためそのような誤差の修正をしなくてすみ、詳細な被写体の法線を推定することができる。

¹ 九州大学 大学院システム情報科学研究院
Graduate School of Information Science and Electrical Engineering,
Kyushu University, Japan

² 静岡大学 電子工学研究所
Research Institute of Electronics, Shizuoka University, Japan

^{a)} yoda@limu.ait.kyushu-u.ac.jp

初期の照度差ステレオ法は Woodham [6] によって提案された。この手法では、法線を推定するために異なる方向から光を当てた画像が少なくとも 3 枚必要である。これらの画像はそれぞれピクセル単位で対応しており、陰影の変化は光源の位置の変化のみからくすることを仮定している。そのため、照度差ステレオ法に用いる画像を撮影する間はカメラや被写体が動いてはならないという条件がある。これらの仮定のもと、照度差ステレオ法ではピクセル単位で被写体の法線の推定を行うことができる。しかし、被写体がカメラに対して動いたり被写体の形状が変化する等の動的なシーンではこれらの仮定が成り立たないため、典型的な照度差ステレオ法では法線を推定することができない。

そのため大きく分けて 2 つの動的シーンに対して法線を推定する照度差ステレオ法が提案されている。1 つはマルチスペクトル照明を用いた手法であり、もう 1 つはハイスピードカメラを用いた手法である。Christensen ら [8] や Hernández ら [9] がマルチスペクトル照明を用いた照度差ステレオ法を提案している。この手法では赤色、青色、緑色等の周波数の異なる色照明を異なる方向から被写体に当て、イメージセンサで得られたカラー画像を RGB の 3 チャンネルに分割し、それぞれの画像を照度差ステレオ法に適用した。この手法では色照明の照射は 1 度ですむため、複数の画像間で被写体の位置や形状が変化するなどしてピクセル単位で対応が取れなくなるといった問題が発生しない。しかし、この手法では色照明を用いるため、被写体のアルベドが異なる箇所があると法線の推定を行うことができないという問題がある。

Vlasic ら [10] は 240fps の、Malzbender ら [11] は 500fps のハイスピードカメラをそれぞれ用いて撮影を行い、照度差ステレオ法を適用した。ハイスピードカメラでは高速に撮影を行うことが可能であり、それらの画像間での被写体の位置や形状の変化は殆ど無いとみなすことができるため、動的シーンに対しても法線を推定することができる。しかし、一般的にハイスピードカメラは高価であり、シャッタースピードを短くするために露光時間が短くなってしまふ。そのため、SN 比が低くなり、推定される法線にノイズが含まれてしまふ。そのためにより強い照明を用いたり、感度がより良いカメラが必要となってしまう。

本稿では、マルチタップ CMOS イメージセンサを用いて動的シーンに対しても法線を推定できる照度差ステレオ法を提案する。マルチタップ CMOS イメージセンサでは、露光後にフォトダイオードで生成された電荷をピクセル毎に分割して異なるフローティングディフュージョン (FD; Floating Diffusion) に蓄積することで、露光を分割し複数枚の画像をほぼ同時に取得する。我々は光源をマルチタップイメージセンサの露光と同期させることによって、異なる方向から被写体に光を当てた画像を同時に撮影した。この画像を照度差ステレオ法に適用することによって、動的

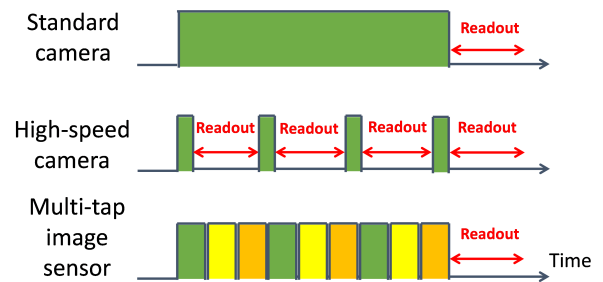


図 2: 通常のカメラとハイスピードカメラ, マルチタップ CMOS イメージセンサの露光のタイミング図の比較. 通常のカメラとハイスピードカメラはそれぞれ 1 ピクセルあたり 1 つの FD を備える. 通常のカメラが 1 枚の画像を撮影する間, ハイスピードカメラは複数枚の画像を撮影することができる. しかし, 露光時間は短くなってしまふという欠点がある. マルチタップイメージセンサでは短い露光を繰り返すことで SN 比が良い画像を複数枚撮影することができる.

シーンに対しても法線の推定を行うことができる。

2. マルチタップ CMOS イメージセンサ

イメージセンサは携帯電話やカメラなど多くの場面で使われている。そのイメージセンサは CCD (Charge-Coupled Device) イメージセンサと CMOS (Complementary Metal Oxide Semiconductor) イメージセンサに大別できる。その中でも CCD イメージセンサは感度が良くデジタルカメラなど多くの場面で使われてきた。しかし、CCD は一般的に構造が複雑であり、消費電力が高いという欠点がある。それに対し CMOS イメージセンサは安価であり消費電力も少ないため、携帯電話などの小型デバイスに利用されている。また、CMOS イメージセンサでは周辺回路の集積化が可能であるため、高速なイメージセンサとしても利用されるようになっている。通常 CMOS イメージセンサは 1 つのフォトダイオードと 1 つの FD をそれぞれのピクセルごとに持つ。フォトダイオードでは光電効果によって光を電荷に変換し、生成された電荷は FD に蓄積される。蓄積された電荷は FD から読み出され、それらの FD から読み出された電荷が 1 枚の画像の 1 ピクセルに対応する。

露光時間と読み出し時間のタイミング図の比較を図 2 に示す。上段のタイミング図は通常イメージセンサが、1 枚の画像を取得するタイミング図を表している。中段ではハイスピードカメラが 4 枚の画像を高速に撮影するときのタイミング図を示している。一般的にハイスピードカメラでは、通常イメージセンサが 1 枚の画像を取る間に複数枚の画像の撮影が可能である。しかし、ハイスピードカメラでは露光時間が短くなるため、取得される画像の SN 比は悪くなる。

本稿では Han ら [5] が提案したマルチタップ CMOS イ

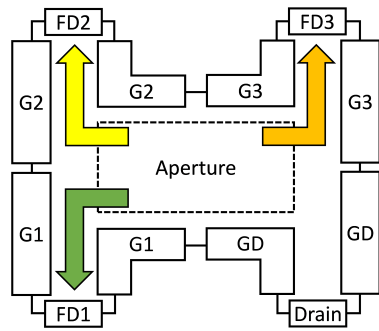


図 3: マルチタップ CMOS イメージセンサの 1 ピクセルあたりの構造の模式図。緑色と黄色と橙色の矢印はそれぞれアパチャーで生成された電荷の流れを表す。これらの色は図 2 に示した露光のタイミング図の最下段の色と対応する。

メージセンサを用いる。このイメージセンサでは、1 ピクセルあたり 4 つの FD を備えている。図 3 に、そのマルチタップイメージセンサの 1 ピクセルの構成の模式図を示す。このマルチタップイメージセンサでは 4 つのゲート (G1, G2, G3, GD) と 3 つの FD (FD1, FD2, FD3) とドレインを備えている。ドレインは余分な電荷を捨てるために使われる。図 3 に緑色の矢印で示したとおり、G1 が High になった時に、Aperture で生成された電荷は FD1 に蓄積される。同様に、G2 が High になった時には黄色の矢印で示したとおり、Aperture で生成された電荷が FD2 に蓄積され、G3 が High になった時には橙色の矢印で示したとおり、電荷は FD3 に蓄積される。これらのゲートが順に High になるサイクルを繰り返すことで図 2 の下段に示したような露光を行うことができる。図 3 の矢印の色は図 2 の下段のタイミング図の露光の色と対応している。FD に蓄積された電荷をそれぞれ読みだすことで図 2 の下段の露光に対応した画像を取得することができる。

このようなマルチタップイメージセンサは Kinect V2 [1] やその他 [4] などの ToF(Time-of-Flight) イメージセンサに用いられている。ToF カメラでは光源から照射された光は被写体までの距離に応じてカメラで観測されるまでの時間が異なり、各タップが On になるタイミングをずらすことで被写体までの距離に応じて輝度値が異なる画像を取得できる。これら画像間で異なる輝度値の比から被写体までの距離を求めることができる。提案する手法では、このマルチタップイメージセンサを用いることで照度差ステレオ法によって動的シーンの法線の推定を行う。

3. マルチタップイメージセンサを用いたダイナミック照度差ステレオ法

我々は動的シーンに対して照度差ステレオ法で法線を推定することができるカメラと照明のシステムを提案する。提案するシステムでは、Han ら [5] が提案したマルチタップ CMOS イメージセンサを用いる。このシステムでは 3

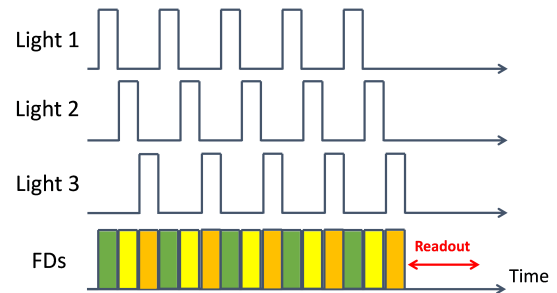


図 4: 光源とタップの同期タイミング図。それぞれの光源の色はマルチタップイメージセンサの各ピクセルの FD に対応する。各ピクセルのそれぞれの FD では対応した光源からの光を変換した電荷のみを蓄積する。

表 1: プロトタイプカメラの性能。

Number of pixels	413(H) × 240(V) (Total pixels)
Pixel size	16.8 μ m × 16.8 μ m
Exposure time	0.2ms × 40iteration = 8ms for each image.
Readout time	13.2ms
Frame rate	20fps
Focal length of lens	12.5mm

つの制御可能な光源を用いており、それらの光源はイメージセンサの露光と同期している。図 4 に示したとおり、それぞれの光源の On と Off はイメージセンサの露光と同期しており、それぞれの光源は露光と同様に何度も On と Off が繰り返される。FD1 から読み出される画像には light1 からの光のみが含まれ、同様に FD2 から読み出される画像、FD3 から読み出される画像にも light2, light3 からの光のみが含まれる。それぞれの光源が On になる時間と露光時間は 1 μ s まで短くすることができ、異なる光源から光を当てた動的シーンの画像を取得するのに十分な早さである。

以上の構想に基づき、プロトタイプのカメラシステムを作成した。作成したシステムの写真を図 5 に示す。本システムで用いた文献 [5] のマルチタップ CMOS イメージセンサの性能を表 1 に示す。各光源は Arduino Uno によって制御されており、Arduino Uno ではマルチタップイメージセンサのゲート信号と同期して光源の On と Off を切り替える。Arduino Uno を用いているため、光源の On と Off は 1 μ s の速度で行うことができる。

照度差ステレオ法では光源の位置の変化によってのみ画像の輝度値が変化することを仮定している。しかし、提案するシステムではイメージセンサのタップの感度が異なるため、同じ光源を用いて撮影した画像でもそれぞれの FD から読み出される画像の輝度値は異なる。そのため、画像の輝度値の変化は光源の変化のみならず、撮影された画像はそのままでは照度差ステレオ法に用いることはできない。よって同じシーンを撮影した画像の輝度値の合計を用いて、以下に示す数式で撮影された画像の正規化を行う。

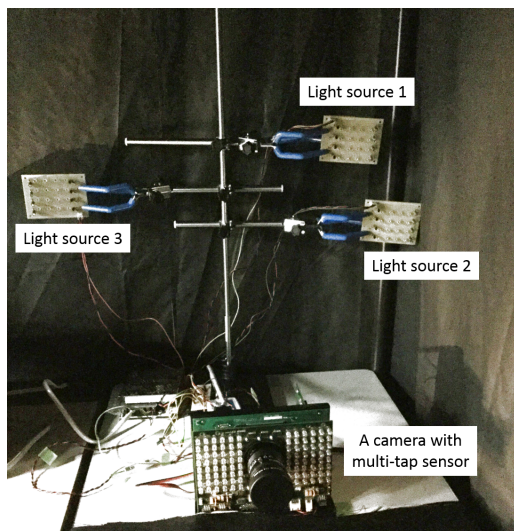


図 5: 光源とマルチタップ CMOS イメージセンサからなる本稿で作成したプロトタイプシステムの写真。光源はそれぞれ 16 個の LED からなり、オブジェクトとそれぞれの光源との距離が等しくなるように光源を配置した。カメラ横の LED は使用していない。

$$I_n^i(x, y) = I_c^i(x, y) \frac{\sum_{x,y} I_b^1(x, y)}{\sum_{x,y} I_b^i(x, y)}, \quad (1)$$

ここで、 $I_c^i(x, y)$ と $I_n^i(x, y)$ は撮影された画像と、それを正規化した画像を表している。 I_b^i は同じシーンを撮影した画像を表しており、この画像の輝度値の差はマルチタップイメージセンサの各タップの感度の差を表している。このタップの感度の差のみを表している画像はすべての撮影された画像の正規化に用いることができる。 i は各 FD に対応する通し番号を表している。

上記のようにして撮影された画像を、照度差ステレオ法 [6] に適用する。照度差ステレオ法では 3 枚の画像と既知の光源から被写体表面の法線を推定する。もし被写体がランバート拡散反射面であると仮定できる場合、以下の式によって画像の輝度値を計算することができる。

$$\begin{bmatrix} I_n^1 \\ I_n^2 \\ I_n^3 \end{bmatrix} = N \begin{bmatrix} L^1 \\ L^2 \\ L^3 \end{bmatrix}, \quad (2)$$

ここで、 N は被写体表面の法線ベクトルの行列を表しており、 L^i は各 FD と同期している光源の位置を表す。事前にキャリブレーションによって L^i を求めておくことで、以下に示す数式のように撮影した画像から被写体の法線を推定することができる。式 (3) で L^\dagger は擬似逆行列を表す。

$$N = \begin{bmatrix} I_n^1 \\ I_n^2 \\ I_n^3 \end{bmatrix} \begin{bmatrix} L^1 \\ L^2 \\ L^3 \end{bmatrix}^\dagger. \quad (3)$$

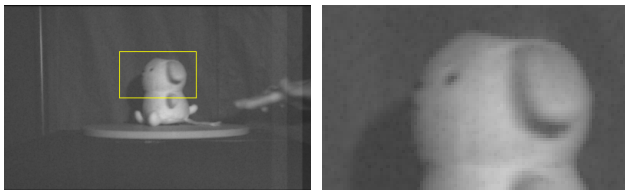
4. 実験

この章では我々の提案手法で動的シーンに対して法線を推定することができることを示す。我々が用いたマルチタップ CMOS イメージセンサは露光時間を自由に設定することができる。そのイメージセンサを用いて、動的シーンに対して通常のイメージセンサ (FD を 1 つしか備えていないイメージセンサ) の設定と各 FD で露光を繰り返すマルチタップイメージセンサの設定で撮影を行い比較した。外光の影響をなくすために撮影は暗室の中で行った。被写体にはターンテーブルの上に置いた人形を用い、そのターンテーブルを手動で回転させて撮影を行った。図 6 に通常のイメージセンサの設定で撮影を行った画像と、それらから推定した法線マップを示す。図 7 に提案するマルチタップイメージセンサで露光を繰り返して撮影を行った時の画像と、それらから推定した法線マップを示す。それぞれの図で、黄色の四角形で囲まれた範囲を拡大して右側に示している。図 6 では、通常のイメージセンサの設定で撮影した画像の対応が取れてなく、被写体の位置が動いてしまっている。そのため、図 6g に示すように推定された法線にはエラーが含まれてしまっている。それに対し図 7 では、撮影された 3 枚の画像で被写体の位置は変化なく、推定された法線に被写体が運動していることによって生じるエラーはない。

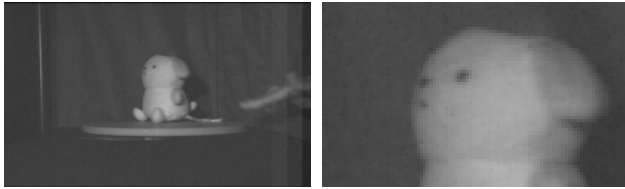
また、ハイスピードカメラのような短い露光時間では撮影される画像の輝度値が小さくなり、推定される法線にノイズが含まれる。外光の影響をなくすために暗室の中で実験を行い、被写体は静止させたまま長い露光時間と短い露光時間を設定して撮影を行った。長い露光時間では 1 枚あたり 10ms の露光を行い、短い露光では 1 枚あたり 4ms の露光を行った。図 8 と図 9 にその実験の結果を示す。図 8 と図 9 ではそれぞれ比較しやすくするために画像の輝度値を調整してある。図 8 では図 9 に比べ画像が暗く、そのためノイズが生じている。

5. 考察

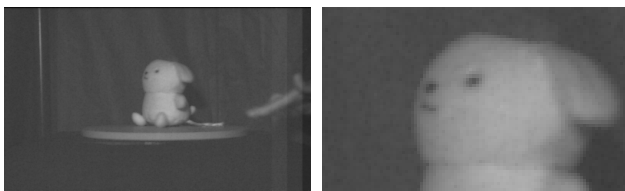
この章では撮影可能な被写体の運動速度やカメラの露光時間、露光の繰り返し回数などの最適なパラメータについて考察する。動的シーンに対して被写体表面の法線を推定するためには露光時間をできるだけ短くすればよい。しかし、光源のスイッチングの応答速度や撮影した画像の輝度値など露光時間を短くすることにも限界がある。また、高速に光源のスイッチングを行うためにはシステムも高価になりやすい。そのため、ここではそれらのパラメータのトレードオフについて述べる。



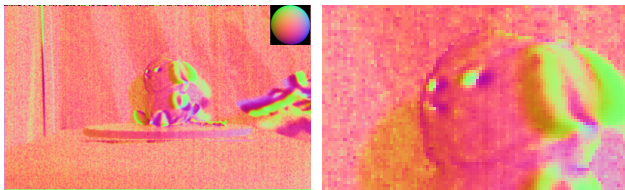
(a) 撮影した画像 1. (b) 図 6a の拡大画像.



(c) 撮影した画像 2. (d) 図 6c の拡大画像.



(e) 撮影した画像 3. (f) 図 6e の拡大画像.



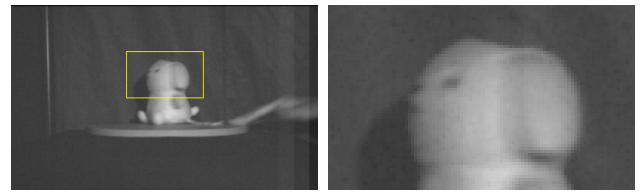
(g) Normal map. (h) 図 6g の拡大画像.

図 6: 動的シーンに対して通常のイメージセンサの露光の設定で撮影した画像と、それらから推定した法線マップ。右側にそれぞれの画像を拡大した画像を示している。拡大している範囲は図 6a に黄色の枠線で囲んである。それぞれの画像で被写体の位置が異なっており、そのため推定される法線にエラーが生じている。

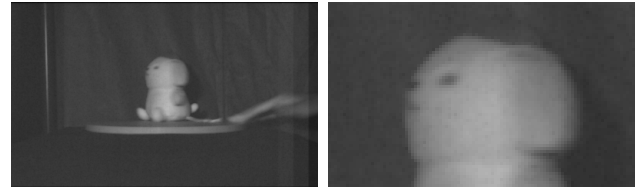
5.1 露光時間の変化によるノイズの比較

図 8 と図 9 ですでに露光時間が短すぎると推定される法線にノイズが多く含まれてしまうことについて述べた。そのため、ここでは露光時間と被写体の法線の推定のトレードオフについて調べた。比較実験では、露光時間を 20ms と十分長くして撮影した画像から推定された法線をリファレンスとし、露光時間が短くなった時に法線はリファレンスからどの程度異なってくるかを Root Mean Squared Error(RMSE) によって評価した。式 (4)、式 (5) で、 w と h はそれぞれ撮影される画像の横と縦のサイズを表す。また、 N_r と N_c はそれぞれリファレンスの法線ベクトルと撮影した画像から推定した法線ベクトルを表している。

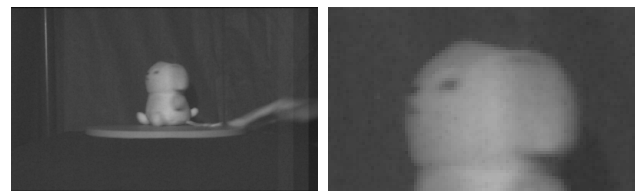
$$\theta(x, y) = \arccos \frac{N_r(x, y) \cdot N_c(x, y)}{|N_r(x, y)| |N_c(x, y)|}. \quad (4)$$



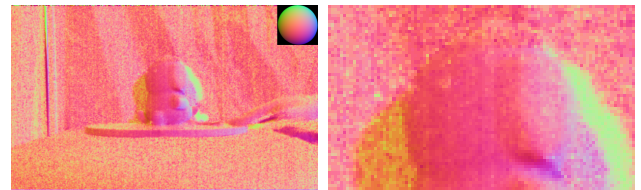
(a) 撮影した画像 1. (b) 図 7a の拡大画像.



(c) 撮影した画像 2. (d) 図 7c の拡大画像.



(e) 撮影した画像 3. (f) 図 7e の拡大画像.



(g) Normal map. (h) 図 7g の拡大画像.

図 7: 動的シーンに対して提案手法であるマルチタップイメージセンサの露光の設定で撮影した画像と、それらから推定した法線マップ。右側にそれぞれの画像を拡大した画像を示している。拡大している範囲は図 7a に黄色の枠線で囲んである。それぞれの画像で被写体の位置が対応しており、推定された法線に、被写体の位置の変化によるエラーは見られない。

$$RMSE = \sqrt{\frac{1}{wh} \sum_{x=0}^w \sum_{y=0}^h \theta(x, y)^2}. \quad (5)$$

我々は暗室の中で露光時間を変化させながらそれぞれ 100 回ずつ法線の推定を行った、ここでは露光時間が推定される法線に与えるノイズについて評価するため、静的シーンを撮影した。図 10 にそれぞれの設定での RMSE を示す。図 10 に示したとおり、露光時間を 8ms よりも短くした時に RMSE は 0.3 より大きくなってしまふ。よってそれぞれの画像の露光時間は 8ms 以上にしたほうがノイズが少なくてよい。

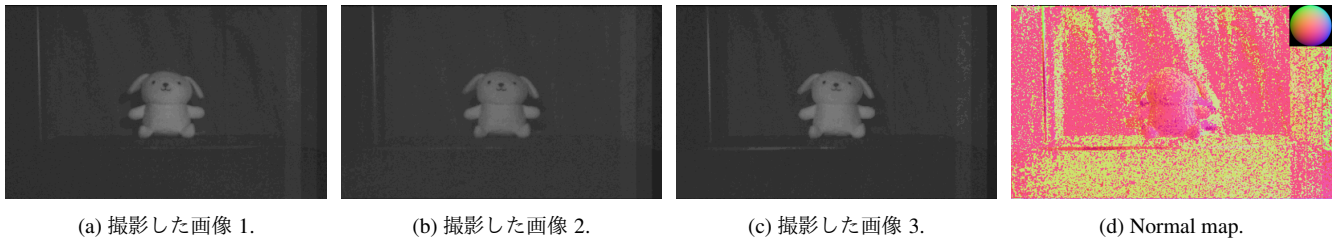


図 8: 短い露光時間で撮影した画像と、それらから推定した法線マップ。それぞれの画像 1 枚あたりの露光時間は 4ms である。画像は見やすくするために輝度値を 8 倍して表示している。露光時間が短いと SN 比が悪くなり、推定される法線に多くのノイズが生じている。

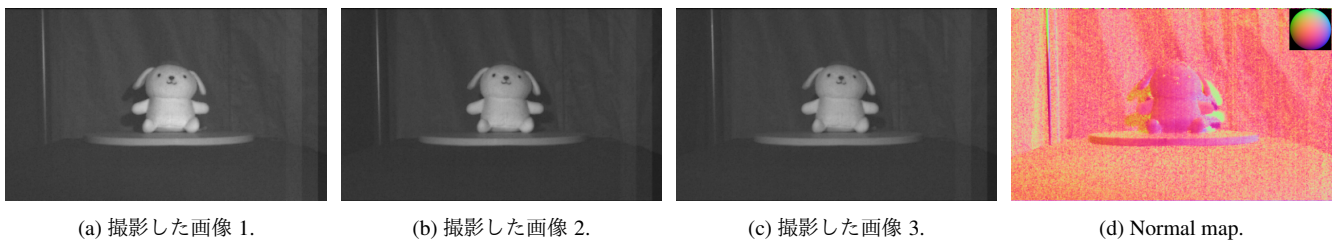


図 9: 長い露光時間で撮影した画像と、それらから推定した法線マップ。それぞれの画像 1 枚あたりの露光時間は 10ms である。画像は見やすくするために輝度値を 1.5 倍して表示している。露光時間が長いと SN 比による法線のノイズは生じていない。

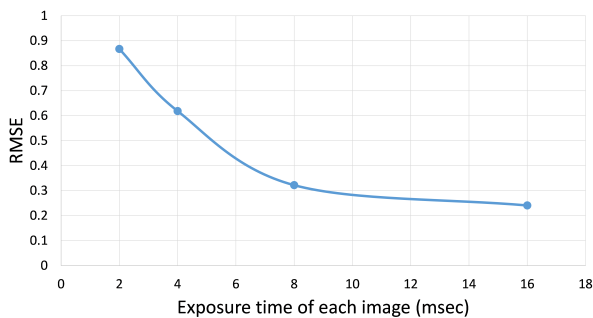


図 10: リファレンスの法線と、撮影した画像から推定された法線との間でどの程度角度に差が出るのかを RMSE によって評価した。静的なシーンを暗室の中で、それぞれの露光時間の設定で 100 枚撮影した。動的シーンに対して法線を推定するためには露光時間は短くなっていないとされないが、露光時間が短くなると推定される法線に多くの誤差が含まれてしまう。特に露光時間が 8ms よりも短いと RMSE は 0.3 を超えてしまうため、各画像で露光時間は 8ms 以上取った方がよい。

5.2 被写体の運動速度とカメラのシャッタースピードとの関係

前節では露光時間と、画像の低い輝度値が推定される法線に与える誤差について述べた。ここで、マルチタップイメージセンサで撮影できる最大の被写体速度について述べなければならない。マルチタップイメージセンサで撮影する時間と、その時に被写体の動きが画像に現れない最大の

速度は以下の数式によって表される。

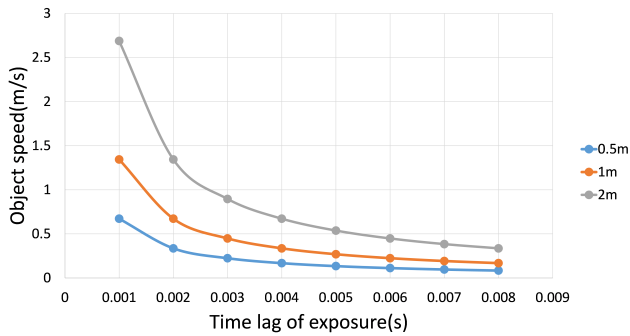
$$v_{max} = \frac{d}{f} \cdot \frac{s}{e}, \quad (6)$$

ここで、 v_{max} は画像中に被写体の速度が現れない最大速度を表し、 f と s はイメージセンサのレンズの焦点距離とイメージセンサのピクセルサイズを表す。本提案手法で用いたマルチタップ CMOS イメージセンサではピクセルサイズは $16.8\mu\text{m}$ である。 d はイメージセンサと被写体との距離を表す。また、 e は図 2 に示したように、出力される各画像間での露光時間のずれを表している。

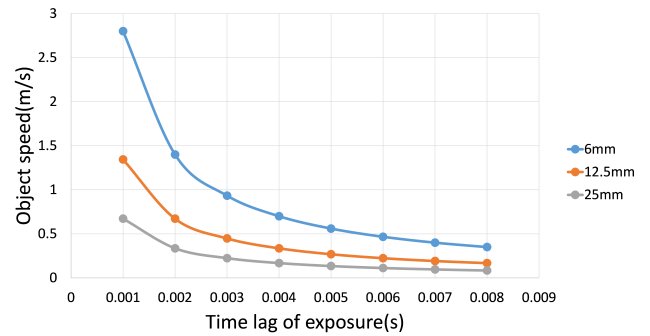
図 11 に式 (6) のグラフを示す。図 11a は被写体の運動速度と露光時間のずれ、そして被写体とイメージセンサとの距離の関係を表している。露光時間のずれが小さければ小さいほど、より速い被写体を撮影することができる。また、同じ露光時間のずれでも、イメージセンサと被写体との距離を離せばより速い被写体の撮影を行うことができる。また、図 11b に示したとおり、イメージセンサのレンズの焦点距離が小さいものを用いることでもより速い被写体の撮影を行うことができる。

5.3 露光回数の変化によるノイズの比較

図 11 が示すように、動的シーンに対して異なる方向から光を当てた画像を同時に取得して法線を推定するためには露光時間のずれを短くすればよい。よって露光を可能な限り分割して小さな露光を何度も繰り返すことで高速に運動する被写体に対しても法線の推定を行うことができる。こ



(a) 露光時間のずれと撮影可能な被写体の最大速度との関係。用いたカメラのレンズは 12.5mm.



(b) 露光時間のずれと撮影可能な被写体の最大速度との関係。カメラと被写体との距離は 1.0m.

図 11: 被写体の運動速度と露光時間のずれ, 及びカメラと被写体との距離とカメラのレンズの焦点距離の関係. 図 11a では, 露光時間のずれが短くなると撮影可能な被写体の最大速度が大きくなる. また, カメラと被写体との距離を長くすることで速い被写体を撮影することが可能である. 図 11b では, カメラと被写体との距離と同様に, 焦点距離の小さなカメラのレンズを用いることでより速い被写体を撮影することが可能である.

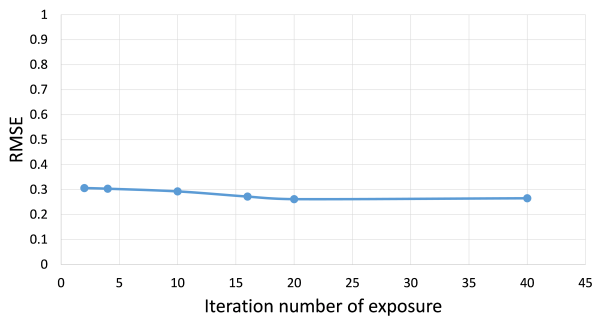


図 12: リファレンスの法線と, 撮影した画像から推定された法線との間でどの程度角度に差が出るのかを RMSE によって評価した. 静的なシーンを暗室の中で, それぞれの露光の分割数で 100 枚撮影した. 結果より, 推定される法線に対して露光の分割数はほとんど影響を与えないことがわかった. そのため, 動的シーンに対して法線を推定するためにはできるだけ露光回数を増やした方がよい.

ここでは, 露光の分割数を増やしていった時に得られる法線にエラーがどのくらい含まれているのかについて述べる. 露光時間の長さによるノイズの影響の比較実験を行った時と同じように, 撮影は暗室内で行い, 静的シーンを撮影した. それぞれの露光の分割数ではそれぞれ 100 枚画像を撮影し, RMSE によってリファレンスからどの程度推定された法線にエラーが含まれるかを調べた. 図 12 にその結果を示す.

図 12 に示したとおり, 分割数を増やしても露光時間を変化させた時の大きな影響は現れていない. 露光回数は, 3 枚の画像を同時に取得する際に 3 枚の画像間にどれくらい時間差が含まれているかに影響するので, より高速に運動する被写体の法線を推定する際には露光回数を大きくして時間差を小さくすればよい.

6. まとめ

照度差ステレオ法は静止物体に対して, 一意に法線を求めることができる. しかし, 運動する被写体では, 高速に被写体を撮影する必要があるため, 既存のシステムでは照度差ステレオ法によって法線を求めることは難しかった. そこで本論文ではマルチタップを備えた CMOS イメージセンサを用いて, 異なる光源から光を当てた画像を同時に取得するシステムについて提案した. 法線を推定する際に, 露光時間と露光回数をどの程度にすればよいのかを露光時間と露光回数を制御できるマルチタップ CMOS イメージセンサを用いて実験を行い, RMSE によって評価した. その結果, 露光時間は 8ms に設定するのが最もノイズの影響が少なく, かつ露光時間が短いことが示された.

今回の実験では法線の推定はオフラインで行っているため, 今後はリアルタイム (15Hz 以上) で法線を推定するシステムを作成していく. また, 今回用いた光源は点光源であり, 照度差ステレオ法の平行光線の仮定に反するため, 画像の端の方では法線を正しく求められていない. 近接点光源を用いた照度差ステレオ法についても, 様々な研究で議論されている. Iwahori ら [12] は, 点光源を使って観測される 4 枚以上の画像のそれぞれの最も輝度が高い点を用いて被写体の法線を求める PSIS(Point Source Illumination Stereo) を提案した. また, 岡部らは近接点光源が未校正照度差ステレオにおける形状復元の不定性を解決することを示した [13]. それらを参考にリアルタイムで画像全体で法線が正しく求められるようシステムを拡張することが必要である.

謝辞 本研究は Microsoft Research CORE 11 program の助成を受けたものです.

参考文献

- [1] Yang, L., Zhang, L., Dong, H., Alelaiwi, A. and Saddik, A.: Evaluating and improving the depth accuracy of Kinect for Windows v2, *IEEE Sensors Journal*, Vol. 15, No. 8, pp. 4275–4285 (2015).
- [2] Landström, A., Thurley, M. J. and Jonsson, H.: Sub-millimeter crack detection in casted steel using color photometric stereo, *Digital Image Computing: Techniques and Applications* (2013).
- [3] Vogiatzis, G., Hernández, C. and Hern, C.: Practical 3D Reconstruction Based on Photometric Stereo, *Computer Vision*, Vol. 285, pp. 313–345 (2010).
- [4] Gokturk, S., Yalcin, H. and Bamji, C.: A time-of-flight depth sensor-system description, issues and solutions, *IEEE Conference on Computer Vision and Pattern Recognition Workshop*, pp. 35–35 (2004).
- [5] Han, S.-M., Takasawa, T., Yasutomi, K., Aoyama, S., Kagawa, K. and Kawahito, S.: A Time-of-Flight Range Image Sensor With Background Canceling Lock-in Pixels Based on Lateral Electric Field Charge Modulation, *IEEE Journal of the Electron Devices Society*, Vol. 3, No. 3, pp. 267–275 (2015).
- [6] Woodham, R. J.: Photometric method for determining surface orientation from multiple images, *Optical Engineering*, Vol. 19, No. 1, pp. 191139–191139 (1980).
- [7] Changpeng Ti, Ruigang Yang, James Davis, Santa Cruz and Zhigeng Pan: Simultaneous Time-of-Flight Sensing and Photometric Stereo with a Single ToF Sensor, *CVPR*, No. 1, pp. 4334–4342 (2015).
- [8] Christensen, P. H. and Shapiro, L. G.: Three-dimensional shape from color photometric stereo, *International Journal of Computer Vision*, Vol. 13, No. 2, pp. 213–227 (1994).
- [9] Hernández, C., Vogiatzis, G., Brostow, G. J., Stenger, B. and Cipolla, R.: Non-rigid Photometric Stereo with Colored Lights, *IEEE International Conference on Computer Vision*, pp. 1–8 (2007).
- [10] Vlastic, D., Peers, P., Baran, I., Debevec, P., Popović, J., Rusinkiewicz, S. and Matusik, W.: Dynamic shape capture using multi-view photometric stereo, *ACM Transactions on Graphics*, Vol. 28, p. 1 (2009).
- [11] Malzbender, T., Wilburn, B., Gelb, D. and Ambrisco, B.: Surface Enhancement Using Real-time Photometric Stereo and Reflectance Transformation., *Rendering Techniques*, Vol. 2006, p. 17 (2006).
- [12] Iwahori, Y., Sugie, H. and Ishii, N.: Reconstructing shape from shading images under point light source illumination, *International Conference on Pattern Recognition*, Vol. 1, pp. 83–87 (1990).
- [13] 岡部孝弘, 佐藤洋一: 近接点光源は未校正照度差ステレオにおける形状復元の不定性を解決するか?, *情処学研報*, Vol. 1106, pp. 2007–157 (2007).