

単色波線グリッドパターンを用いたプロジェクタ・カメラ能動ステレオ

糟谷 望[†] 阪下 和弘[‡] 佐川 立昌[†] 古川 亮^{††} 川崎 洋^{‡‡}

[†] 独立行政法人 産業技術総合研究所 〒305-8569 茨城県つくば市梅園 1-1-1 つくば中央第2

[‡] 大阪大学 産業科学研究科 〒567-0047 大阪府茨木市美穂ヶ丘 8-1

^{††} 広島市立大学 情報科学研究科 〒731-3194 広島県広島市安佐南区大塚東 3-4-1

^{‡‡} 鹿児島大学 理工学研究科 〒890-0065 鹿児島県鹿児島市郡元 1-21-40

E-mail: [†] {nozomu.kasuya, ryusuke.sagawa}@aist.go.jp, [‡] sakashita@am.sanken.osaka-u.ac.jp

^{††} ryo-f@hiroshima-cu.ac.jp, ^{‡‡} kawasaki@ibe.kagoshima-u.ac.jp

あらまし 本稿では、プロジェクタとカメラを用いた能動ステレオによる形状計測手法を提案する。提案手法では、プロジェクタから静的なパターンを投影し単一画像から復元を行うため、高フレームレートでの動物体の形状計測が可能である。単色パターンによる高精度・高密度なワンショット形状計測の実現を目指し、単色波線グリッドパターンを用いることを提案する。単色のパターンを用いることで、利用できるプロジェクタが増え、応用範囲の拡大が望める。波線グリッドパターンを用いることで交点に空間的な特異性をもたせ、ステレオ法による形状計測の精度向上を図る。また、画像マッチングを利用して画素単位の補間を行うことで高密度な計測を実現する。実験により提案する単色波線グリッドパターンを用いて形状計測が実現できることを確認する。

キーワード 三次元形状復元, ワンショット, 能動ステレオ, 波線グリッド

Projector-Camera Active Stereo using Single-colored Wave Grid Pattern

Nozomu KASUYA[†] Kazuhiro SAKASHITA[‡] Ryusuke SAGAWA[†]

Ryo FURUKAWA^{††} and Hiroshi KAWASAKI^{‡‡}

[†] AIST 1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568 Japan

[‡] Osaka University 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047 Japan

^{††} Hiroshima City University 3-4-1 Ozuka-Higashi, Asa-Minami-ku, Hiroshima, 731-3194 Japan

^{‡‡} Kagoshima University 1-21-40 Korimoto, Kagoshima-city, Kagoshima, 890-0065 Japan

E-mail: [†] {nozomu.kasuya, ryusuke.sagawa}@aist.go.jp, [‡] sakashita@am.sanken.osaka-u.ac.jp

^{††} ryo-f@hiroshima-cu.ac.jp, ^{‡‡} kawasaki@ibe.kagoshima-u.ac.jp

Abstract This paper proposes a method to reconstruct the shape of moving objects by projector-camera active stereo. We reconstruct the shape from single image where a static pattern is cast by a projector, so this method is able to reconstruct the fast moving objects by capturing at a high frame rate. Aiming to reconstruct with high precision and dense, we propose using “single-colored wave grid pattern”. A single-colored pattern enables to use various projectors. To reconstruct dense shapes, we apply pixel-wise interpolations using image matching. We applied the proposed method to real moving object reconstruction to confirm its effectiveness.

Keyword 3D reconstruction, One-shot, Active stereo, Wave grid

1. はじめに

近年、動物体を含む動的シーンの三次元形状計測が注目され、盛んに研究されている。しかし、現在利用可能な動物体スキャナは、静的シーンを計測する三次元スキャナに比べ、精度・密度が低いといった問題がある。動物体スキャナの精度・密度向上が実現できれば、医療や流体解析など様々な分野での応用が可能になると考えられる。

物体の形状計測を行う手法としては、カメラのみを

用いる受動ステレオ法や、プロジェクタとカメラを用いた能動ステレオ法、レーザ光の反射光が戻るまでの時間によって深度を求める Time-of-Flight 方式など様々な手法が存在する。中でも、プロジェクタ・カメラ系によって構造化光を投影する能動ステレオ法は動物体の形状データを得るのに適しているとして熱心に開発・研究されている[1]-[4]。

能動ステレオ法は、大まかに時間符号化法と空間符号化法に分類される。空間符号化法は単一の画像から

形状を復元する手法のため、動物体を高フレームレートで計測するのに適している。しかし、投影パターン全体の中から一意に特定可能な対応点情報を二次元パターンに直接埋め込むにはある程度大きな領域が必要なため、復元密度が低くなりがちであり、また表面形状変化によるパターンの歪み等による復号化誤りなども起きやすい。

本研究では、単色の波線グリッドパターンを用いることでこれを解決する。波状のグリッドパターンを用いることで、グリッドの交点における空間的なパターンの特異性を向上させ、画素単位の補間および画像マッチングに基づく最適化によって密な形状計測を実現する。単色で形状計測が可能であるため、ビデオプロジェクタはもちろん、単色光にマスクをかけるだけで実現できるため、超小型化やレーザー光による屋外計測も可能であり、利用用途が格段に広がると考えられる。

2. 関連研究

プロジェクタからパターン光を投影し、カメラで撮影される構造化光を用いて形状データを得る能動ステレオ手法は、プロジェクタで投影したパターンとカメラで撮影された画像との対応点を密にとる必要がある。これを解決する方法は、時間的符号化方式と空間的符号化方式に大別される[5]。

時間的符号化方式では、複数のパターンを順に投影し、その変化に各画素位置の情報が符号化される。このため、複数のパターンを投影する必要があり、様々な高速化のためのアプローチが取られているが、本質的に動物体計測には適さないと言える。

一方で、空間符号化方式は、投影パターンを切り替える必要がないため、カメラの撮影速度での形状計測が可能であり、動物体計測に適している。しかし、一枚のパターンに画素位置の情報を埋め込む必要があり、情報を付与するのに大きな領域が必要となるなど、計測密度や安定性が低下しがちである。これを軽減する手法として、複数の色を用いる方法[3][5]-[7]や、特異な点線を利用する方法[8]、二次元的なパターンに情報を埋め込む手法などがある[1][9]。しかし、精度・解像度および安定性のすべてにおいて十分な性能を満たしている手法は未だに存在していない。

提案手法は単色のグリッドパターンを用いた空間符号化方式の形状計測法である。グリッドパターンを用いる手法には、対応点の曖昧性と線の接続判定の誤りによる誤復元の問題が存在する[2]-[4]。この問題を解決するために複数の色を用いたグリッドパターンの利用が提案されているが、これらの手法は対応物体のテクスチャなどに影響されるため、計測が安定しない。提案手法では、単色の波線グリッドパターンを用いることで、これらの問題を解決する。

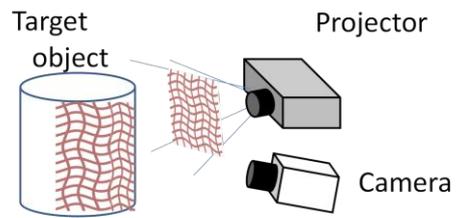


図 1 単色波線グリッドパターンによる形状計測システム

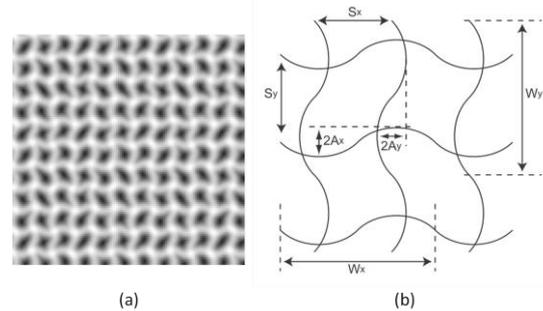


図 2 (a)波線グリッドの例, (b)パラメータ。
(S_x, S_y は隣接する波線の間隔, W_x, W_y は波線の波長, W_x, W_y は波線の振幅)

3. 単色波線グリッドパターンを用いた能動ステレオ

提案する能動ステレオ手法は、図 1 に示すように、プロジェクタから図 2(a)に示す静的なパターンを計測対象に投影し、カメラでそれを撮影し解析することで三次元形状計測を実現する。

投影するパターンは、単色の正弦波形状の縦横線を格子状に配置したものであり、本稿ではこれを波線グリッドと呼ぶ。パターンが静的であるため、非常に高い撮影速度での計測も可能である。

波線グリッドを用いた形状計測手法では、まず撮影した画像から Sagawa ら[3]が提案した手法を用いて線検出を行う。検出された縦横の波線から交点を算出し、交点をノードとしたグラフを作成する。各ノードに対応するエピポーラ線の位置をプロジェクタパターン上で計算し、エピポーラ線上でパターンの交点が存在する場合はそれを対応候補点とする。接続情報と形状特徴から各点における最適な対応点の組み合わせを求め、カメラ中の画素と対応するプロジェクタの交点座標とで三角測量することで三次元位置を計測する。このままでは計測は疎であるため、各画素における深度を、補間とパターンの観測画像の画素単位のマッチングを用いて求めることで、密な三次元形状を得る。

4. 単色波線グリッドによる暗黙的符号化

本稿では、プロジェクタ画像面から投影するパターンをプロジェクタ画像、カメラ画像面で撮影される画像をカメラ画像と呼ぶ。

ステレオ法で三次元形状を計測するためには、プロジェクト画像とカメラ画像の対応を一意に取る必要がある。提案手法で用いる波線グリッドは、画像処理によって一意に対応が決まるパターンではなく、対応の優先順位に関する情報を与えるパターンである。波線は単純なパターンで、画像中での線として検出しやすく、輝度値のピークを算出することにより、サブピクセル精度で位置を獲得できる。

波線グリッドでは、交点の配置が波線の間隔と波長で決定される。縦波線の間隔と横波線の波長の整数倍でないかぎり、交点位置の位相にずれが生じ、交点の形状を対応付けの特徴として用いることができる。この交点の形状は、隣接する波線の間隔と波長によって決まるサイクルで同一形状が現れる。しかし、各サイクル内において識別可能なパターンであり、ステレオマッチングでは、対応点候補はエッジ線上の点に限られているため、対応の候補となる交点を少数に絞り込める。そこで、プロジェクト画像の交点とエッジ線が適当な距離以内に位置している場合、その交点を対応点候補の1つとして選択し、5章に示す最適化で対応点を決定する。

5. 波線グリッド上でのステレオ法

対応点候補の中で最適な対応点を見つけるために、局所的なマッチングと正則化を利用したエネルギー最小化による最適化を行う。

グリッドは、カメラ画像上での波線グリッドの交点（以下、これを格子点と呼ぶ）からなるノード $p \in V$ と、格子点の接続を表すエッジ $(p, q) \in U$ からなる。ただし、 p, q は格子点であり、 V は格子点の集合、 U はグリッドグラフのエッジ集合である。格子点 p の対応点候補を $t_p \in T_p$ （ここで T_p は格子点 p の候補点集合）とすると、対応の集合をパラメータとして、ステレオマッチングのエネルギーを以下のように定義する。

$$E(T) = \sum_{p \in V} D_p(t_p) + \sum_{(p, q) \in U} W_{pq}(t_p, t_q) \quad (1)$$

ここで、 $D_p(t_p)$ は p の対応点を t_p に割り当てる場合のデータ項、 $W_{pq}(t_p, t_q)$ は t_p と t_q を隣り合う格子点に割り当てるための正規化項であり、以下の式で定義される。

$$W_{pq}(t_p, t_q) = \begin{cases} 0 & t_p \text{ と } t_q \text{ が同一波線上} \\ \lambda & \text{それ以外の場合} \end{cases} \quad (2)$$

データ項はカメラ画像とプロジェクト画像と格子点周りの SSD によって計算する。ただし、検出された格子点位置には誤差が存在し、カメラで観測された画像は計測対象物体の表面形状によって歪むことがあるため、対象物体における格子点周辺領域を格子点の接平面で近似したパッチを考え、マッチングコストを計算する。

(1)式を Belief Propagation[10]によって最小化するこ

とで最適な対応点を得る。これにより、格子点の誤った接続と三次元再構成を同時に実行され、より高精度な復元が可能となる。

6. 格子点間の補間による密な形状生成

上述したグリッドベースのステレオ法により、疎な格子点における対応が得られる。次に、カメラ画像の全画素の情報を利用して密な対応を求める。まず、高密度に再標本化された標本画素について、グリッド点の補間により深度を求める。次に、これらの標本画素の深度を変数として、カメラ画像とプロジェクト画像の輝度差を最小化する。これによって準画素それぞれに独立した深度推定が photo-consistency に基づく最適化によって実現される。

最適化において、すべての画素の深度を独立変数として扱い、画素単位の深度推定を行うことは可能である。しかし、通常は投影するパターンの解像度がカメラの解像度より低いため、単純に全画素について深度推定を行うと、かえって精度低下を招くことがある。そこで、本稿では、画素を縦横3画素ごとに再標本化し、これらの標本画素について深度を推定する、準画素単位の深度推定を行う。

6.1. 各画素における深度計算

カメラ中心から画素 x へ向かう方向ベクトルを $(u, v, 1)$ と表すと、その画素の深度 d_x は、画素ごとに計算されるパラメータ a_x, b_x, c_x を用いて、

$$d_x = \frac{-1}{a_x u + b_x v + c_x} \quad (3)$$

で計算される。各画素における a_x は、周囲の格子点 p におけるパラメータを用いて以下のように補間される：

$$a_x = \frac{\sum_p G(|p-x|) a_p}{\sum_p G(|p-x|)} \quad (4)$$

ただし、 $G(\cdot)$ はガウス関数、 $|p-x|$ は p と x の距離である。また、 b_x, c_x も a_x と同様な加重平均で求められる。

6.2. リサンプルによる準画素単位の奥行き推定

繰り返し計算による最適化のために、深度を d_x の微小変位 Δd_x を変数として近似する。

ここで、 $D + \Delta D$ を、 $dx + \Delta d_x$ を全ての標本画素について集めたベクトルとする。プロジェクト画像に対する再投影誤差は、全画素について以下の式で求められる。

$$E(\Delta D) = \sum_x \left(I_c(x) - I_p(P_{D+\Delta D}(x)) \right)^2 + \gamma \sum_{x, x'} (\Delta d_x - \Delta d_{x'})^2 \quad (5)$$

ここで $P_{D+\Delta D}(x)$ はプロジェクト画像への再投影位置を表す。各画素の再投影には、 $D + \Delta D$ の一部が利用される。 x と x' は隣接する頂点、 γ は正規化パラメータである。パラメータ ΔD は上記誤差を最小化するように決定

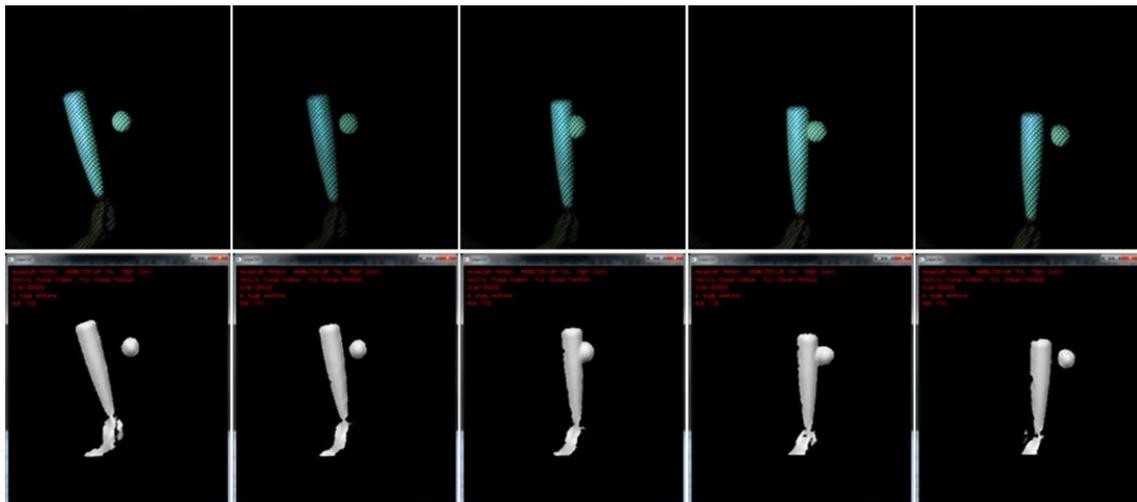


図 3 ボールを打つ動作を再構成したボールとバットの三次元形状
(上段) 撮影画像, (下段) 復元した三次元形状

される。再投影と最小化を解の収束まで繰り返すことで D を決定する。

7. 実験

提案手法を用いることで高速に移動する物体の密な形状計測が可能であることを実験で確認する。実験はハイスピードカメラを用いて、撮影速度 2000fps、解像度 1024x1024 画素で撮影を行う。プロジェクタは 1024x768 画素のものを使用し、波線の間隔が縦 11、横 9 画素、波長 14 画素の波線グリッドパターンを投影する。軟式テニスボールをバットで叩いた際のボールの変化の様子を撮影し、形状を復元する。

図 3 に入力とした撮影画像と復元結果の一部を示す。1 枚目から 5 枚目までは 20msec ほどであり、非常に高速な変化であるボールがバットに当たり潰れている際の三次元形状も計測できているのがわかる。

8. 結論

本稿では、静的なパターンをプロジェクタから投影し、カメラで撮影した単一画像から三次元形状復元をするワンショット形状計測方式において、単色の波線グリッドパターン光を用いることで高精度・高密度に計測する手法を提案した。従来手法のように色を用いることなく、パターンの局所的な形状変化によって特異性を付与した。また、ステレオ法における形状復元を、グリッドの接続性を考慮しながらプロジェクタ・カメラ系に拡張する手法を提案し、画素単位の深度推定を実現した。今後の課題として、複数のシステムを利用した全周計測システムの構築などが考えられる。

謝辞

本研究の一部は、総務省 SCOPE(101710002) および内閣府 NEXT プログラム(LR030) の助成を受けて実施されたものである。

文献

- [1] Microsoft, “Xbox 360 Kinect”, <http://www.xbox.com/en-US/kinect/>
- [2] H. Kawasaki, R. Furukawa, R. Sagawa, Y. Yagi: “Dynamic scene shape reconstruction using a single structured light pattern”, IEEE Conference on Computer Vision and Pattern Recognition (Jun.2008)
- [3] R. Sagawa, Y. Ohta, Y. Yagi, R. Furukawa, N. Asada, H. Kawasaki: “Dense 3d reconstruction method using a single pattern for fast moving object”, IEEE 12th International Conference on Computer Vision, pp1779-1786 (Sep.2009).
- [4] A.O. Ulusoy, F. Calakli, G. Taubin: “One-shot scanning using de bruijn spaced grids”, IEEE 12th International Conference on Computer Vision Workshop (ICCV Workshop), pp1786-1792 (Sep.2009).
- [5] J. Salvi, J. Batlle, E. Mouaddib: “A robust-coded pattern projection for dynamic 3D scene measurement”, Pattern Recognition Letters 19, pp1055-1065 (1998).
- [6] R. Sagawa, H.Kawasaki, R. Furukawa, S. Kiyota: “Dense One-shot 3D Reconstruction by Detecting Continuous Regions with Parallel Line Projection”, IEEE 13th International Conference on Computer Vision, pp.1911-1918 (Nov.2011)
- [7] S. Zhang, P. Huang: “High-Resolution, Real-time 3D Shape Acquisition”, Conference on Computer Vision and Pattern Recognition Workshop (Jun.2004)
- [8] M. Maruyama, S. Abe: “Range sensing by projecting multiple slits with random cuts”, SPIE Vol.1194 Optics, Illumination, and Image Sensing for Machine Vision IV, pp.216-224 (Jun.1993)
- [9] P. Vuytsteke, A. Oosterlinck: “Range image acquisition with a single binary-encoded light pattern”, IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.148-164 (Feb.1990).
- [10] P. F. Felzenszwalb, D. P. Huttenlocher: “Efficient Belief Propagation for Early Vision”, International Journal of Computer Vision, Vol.70, No.1, pp.41-54 (Oct.2006)