

波線パターン投影による高速表面形状計測

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

1. はじめに

物体の三次元形状の計測は、多様な分野での応用が期待され研究が盛んに行われてきた分野の一つである。実際、その用途から接触式のものや非接触のものまで様々なものが製品化され実用されている。近年では、特に動物体の形状を計測できる手法に注目されている。その筆頭として、Microsoft 社の Kinect[1]があげられる。Kinect は、人体の形状を実時間で計測・解析することで、ユーザが身体動作を用いて直感的な操作が可能なゲームインタフェースである。ゲームインタフェースとしての用途以外にも、安価で使用可能であることや SDK が用意されていることなどから、ロボットや拡張現実の分野など、様々な分野で利用しようとする研究が進められている。しかし、Kinect をはじめとする現在利用可能な動物体の三次元形状を計測するセンサは、静的なシーンを計測するセンサほど高精度・高密度な計測ができず、さらに多岐にわたる応用のために精度・密度の向上が望まれている。精度・密度が向上できれば検査用途にも利用可能となる。

三次元形状を計測する方法は、非接触なものに限っても、Time-of-Flight などのレーザスキャン方式や、ステレオカメラなどによる受動ステレオ法、プロジェクタとカメラを用いた能動ステレオ法など、多数存在している。動物体の形状を計測するには、1 回の観測で全体の計測を行う必要がある。そのため、受動ステレオ法や、単一の画像を投影する空間符号化方式の能動ステレオ法が、動物体の計測に適しているとされ盛んに研究されている[1]-[4]。能動ステレオ法は受動ステレオ法に比べ、対象のテクスチャの影響が少なく、テクスチャのないような物体も形状計測が可能という利点があり、本手法では能動ステレオ法を用いる。

能動ステレオ法とは、プロジェクタからパターン光を投影し、パターン光が投影された物体をカメラで撮影することで、プロジェクタとカメラ間の対応をもとに形状を計測する手法である。カメラで撮影された画像中の投影パターンがプロジェクタから投影されたパターンのどの部分かを特定し、三角測量により物体の

形状を計測する。そのため、対応点をいかにして密に精度よく獲得するかが計測の精度に直結する。

我々は単色の波線グリッドパターンを用いることで高精度・高密度な形状計測を実現する。本手法は直線ではなく波線を用い、縦線と横線が交わる位相を変化させることで、交点付近の形状に特異性をもたせ、対応点探索に用いる。さらに、グリッドパターンを用いることで、その接続性を利用することで対応点探索の安定性の向上を図る。また、波線自体をぼかしておくことで、局所領域における画素単位のマッチングによって高密度の計測を可能にする。本手法では、高速カメラを用いることで、非常に高速に動作する物体の計測も可能である。また、投影するパターンが単色であるため、マスクパターンを通すだけでパターンの投影が可能である。そのため、非常に小型な計測を装置の作成も可能であり、内視鏡に応用することで、人体内部の形状すら計測可能である。レーザー光を用いることで環境光の影響を削減し、屋外での計測も可能となると考えられる。

本稿では、能動ステレオ法において単色の波線グリッドパターンを用いることで、物体の形状を高精度・高密度に計測する手法について解説する。さらに、本手法を用いた形状計測システムの例を紹介する。

2. 能動ステレオ法

能動ステレオ法による形状計測では、プロジェクタから投影するパターンと、カメラで撮影された画像で対応をとり、対応のとれた点を用いて三角測量することでそのカメラからの距離を推定する。そのため、プロジェクタ・カメラ間の対応をいかにして取得するかが、能動ステレオ法の肝となる。その対応の取り方によって、能動ステレオ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、その変化を観測することで対応付けの情報を得る。時系列の情報をを用いることでピクセル単位での対応情報の埋め込みが可能であるが、刻一刻と状態の変化する動物体の形状計測には不向きであると言える。

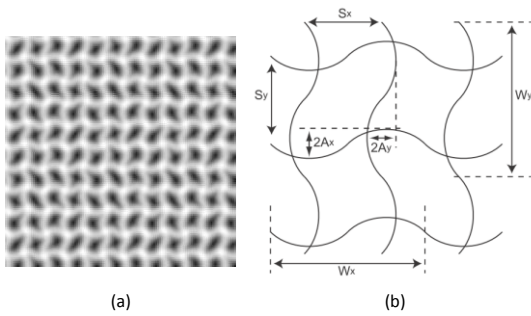


図1 (a)波線グリッドの例, (b)パラメータ.
(S_x, S_y は隣接する波線の間隔, W_x, W_y は波線の波長, A_x, A_y は波線の振幅)

それに対し、空間符号化方式は単一のパターンで形状計測を行うため、動物体の形状計測に適している。しかし、単一のパターンに対応点情報を埋め込む必要があり、計測密度が低下しがちである。この問題を解決するため、複数の色を利用した方法が広く用いられてきた[3][5]-[7]が、色情報を利用する場合には、プロジェクタの各色スペクトルの干渉や計測物体の反射特性の影響から、カメラ画像中の色の決定に誤差が生じ、安定した形状計測にはいたっていない。他にも、特異な点線を利用する手法[8]や、二次元的なパターンに情報を埋め込む手法[1][9]などが提案されている。しかし、精度・密度および安定性の全てにおいて十分な性能を満たしていない。

提案手法は、単色のグリッドパターンを用いることで、多色パターンの問題を解決し、グリッドパターンの接続性を利用することで安定性が向上する。

3. 単色波線グリッドを用いた能動ステレオ

提案手法では、プロジェクタから図1(a)に示すような波線グリッドを計測対象物体に投影する。それをカメラで撮影し、撮影されたカメラ画像と投影パターンの対応点を取得し、三角測量することで形状の計測を行う。投影パターンは正弦波形状の縦横の波線を格子状に配置したものである。本稿ではこれを波線グリッドと呼ぶ。プロジェクタから投影されるパターンはこの1パターンのみで、静的なものであるため、カメラとプロジェクタの同期も必要なく、高速度カメラを使うことで高フレームレートでの形状計測も可能である。また、単色のパターンであるため、マスクパターンの後ろから単色光を照射するだけの単純な機構でもパターン光の投影を実現できる。

提案手法では、事前準備としてプロジェクタ・カメラの内部パラメータおよびプロジェクタ・カメラ間の外部パラメータをキャリブレーションにより求めておく。そして、物体にパターン光を投影し、カメラで撮影した画像から線形状を検出し、交点をノードとした

交点グラフを作成する。各ノードにおけるパターン光における対応候補は、プロジェクタ・カメラの幾何関係から限定できる。その候補点の中からもっともらしいものを交点グラフの Belief Propagation (BP) [10]による最適化によって求める。ここまでで疎な対応関係が得られる。最後に、カメラ画像中のノード間の局所領域において画素単位の画像マッチングを行うことで密な対応を取得し、密な三次元形状を獲得する。

4. 波線グリッドによる暗黙的符号化

本稿では、プロジェクタが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、画像間の対応を一意に取得する必要がある。空間符号化方式の能動ステレオ法では、一般的にパターン光に符号化された情報をカメラ画像中から抽出することで対応点を得る。提案手法で用いる波線グリッドでは、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り、グリッドの接続性を利用してもっともらしいものを選択することで最適な対応を取得する。

波線グリッドは図1に示すように、縦横それぞれ一定の波長 W_x, W_y 、振幅 A_x, A_y の正弦波を S_x, S_y の間隔でグリッド状に並べることで生成する。プロジェクタ画像中の交点位置、および交点周りの形状は、これらをパラメータとして決定される。縦横の波線の間隔と波長が互いに整数倍でない場合、交点位置の位相にずれが生じ、交点周りの局所的パターンが周囲の交点とは異なる形状を持つことになる。これを対応付けのための特徴として用いる。この交点の形状は、波線の間隔と波長の最小公倍数の間隔で同一形状が現れることになる。しかし、プロジェクタとカメラの配置から対応点はエピポーラ線と呼ばれる直線上にしか存在しないため、キャリブレーション誤差を含めてもその直線から数ピクセル以内の交点に対応点を限定できる。

5. 波線検出による交点グラフの作成

カメラ画像からの線検出には Sagawa ら[3]が提案した BP を用いた手法を用いる。この手法では、縦横それぞれ別々に BP を用いたピクセル分類を行うことで画像中の線を安定してサブピクセル精度で検出できる。検出された縦横の線から交点を算出し、交点をノード、波線をエッジとした交点グラフを作成する。ノード間エッジで結ばれていれば、それらのノードはプロジェクタ画像上で同一の波線の上であり、隣り合っていることを示す。ただし、カメラ画像における線検出では、特に不連続な境界付近で誤判定を含むことが多くなり、必ずしも信頼出来るとは言えない。そこで、エネルギー最小化法を用いて尤もらしい対応を求めてい

くことで、信頼度の低い接続の除去と、カメラ画像中の交点とプロジェクタ画像中の交点の最適な対応の取得を同時に行う。

6. 交点グラフのエネルギー最小化による

最適な対応点の決定

交点グラフの各ノードの対応点は、プロジェクタとカメラの幾何関係から対応候補点を絞り込むことができる。キャリブレーション誤差や、画像歪みなどの影響を考慮して、各ノードに対応するエピポーラ線をプロジェクタ画像に引いた際に、プロジェクタ画像上の交点とその線が十分に近い場合に対応点候補の一つとして選択する。ここで選択された対応点候補の中から最適なものをBPによって決定する。

BPのマッチングパラメータは交点周りの形状をカメラ画像とプロジェクタ画像で比較したデータ項と、ノードの接続性による正規化項の和として以下の式のように定義する。

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

ここで、 $p \in V$ は交点グラフのノード、 $(p,q) \in U$ はエッジ、 t_p はノード p の対応点候補の一つである。 $D_p(t_p)$ が p の対応点を t_p に割り当てる場合のデータ項、 $W_{pq}(t_p, t_q)$ がエッジ (p,q) に対する正規化項である。

データ項 $D_p(t_p)$ は、上述した交点周りの形状特徴の類似度を用いることで、似ているものほど小さくように設定する。具体的には、カメラ画像とプロジェクタ画像の交点周りのSSD (Sum of Squared Difference) によって計算する。ただし、カメラ画像中の交点形状は物体表面の形状の影響を受けて歪むことがあるため、対象における交点周辺領域を、交点の接平面で近似したパッチでコストを計算する。この時、接平面の推定は周辺の交点位置によって行う。

正規化項 $W_{pq}(t_p, t_q)$ は、エッジで結ばれたノードが同一線上にあり、隣り合っていることを示していることから、ノード p と q に隣り合う交点が割り当てられた際に最小となるよう設定する。そこで、 t_p と t_q が同一波線上にあれば0、それ以外の場合にはユーザが決めた定数 λ をとる。 λ を大きくすると、カメラ画像中の線検出結果によるノードの接続を大きく信頼し、逆に小さい場合には接続性の依存度が下がりエッジを切断することが多くなる。

(1)式をBPによって最小化することで、エッジで結ばれているノードでエネルギーが伝播され、大域的な最適解を取得できる。

7. 画素単位の最適化による密な形状計測

上述した交点グラフの最適化により、波線グリッドの交点における対応が得られ、疎な形状を計測できる。次に、画素単位の補間と画像マッチングに基づく最適

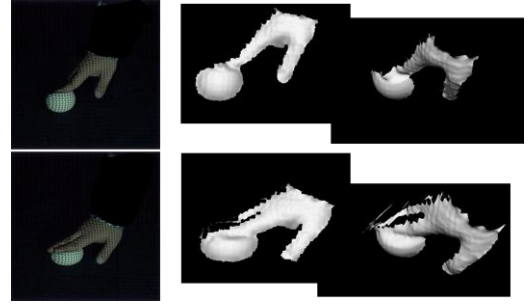


図2 動物体の形状計測の入力(左)と結果(右)

化により、密な形状推定を行う。

投影する波線グリッドは図1(a)に示すように、1ピクセルの線ではなく、ガウス関数をかけておくことで交点以外の画素にも情報をもたせ、局所的な画素値のマッチングを行うことで密な対応を取得できる。

まず、各画素において、周囲の交点の情報をもとに、接平面のパラメータをガウス補間することで算出する。補間によって求めた接平面のパラメータを a_x, b_x, c_x とすると、カメラ画像中の画素 $x = (u, v)$ における深度 d_x は以下の式で計算される。

$$d_x = \frac{-1}{a_x u + b_x v + c_x} \quad (2)$$

そして、このように求めた深度を変数として繰り返し計算による画像マッチングによる最適化を行い、より適切な深度を計算する。繰り返し計算のため、深度 d_x の微小変位 Δd_x を変数として近似する。最適化では、カメラ画像中の画素 x における輝度値 $I_c(x)$ と、画素 x をプロジェクタ画像に再投影した画素 $P(x)$ の輝度値 $I_p(P(x))$ の差を最小化する。ただし、隣接画素との深度差が大きくなり、形状に細かな凹凸ができないよう正規化項を導入し、以下の式を用いる。

$$E(\Delta D) = \sum_x \left(I_c(x) - I_p(P_{D+\Delta D}(x)) \right)^2 + \gamma \sum_{x, x'} (\Delta d_x - \Delta d_{x'})^2 \quad (3)$$

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、 γ を利用者が決める正規化パラメータとする。この最小化計算と再投影を、解の収束まで交互に繰り返すことで、画素単位の深度 D を決定する。

8. 動物体の形状計測実験

提案手法を用いて動物体の形状計測が可能であることを実験で確認する。実験には、1024x1024画素のカメラと、1024x768画素のプロジェクタを用いる。投影パターンは、波線の間隔が縦11、横9画素、波長がともに14画素のパターンを用いる。対象は軟式テニスボールを指で凹ませた際の形状の変化を観測したものである。ボールが凹んでいる様子を計測できているのがわかる。

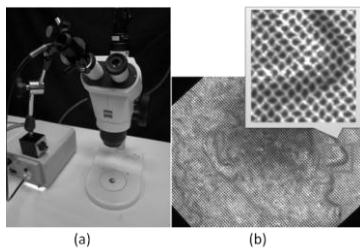


図3 (a)顕微鏡システム,
(b)1セント硬貨に投影された波線グリッド

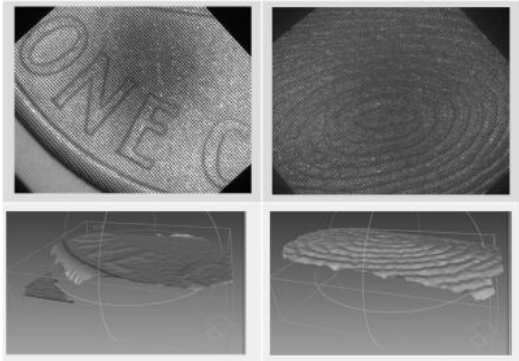


図4 顕微鏡システムによる形状計測結果
(上段)撮影画像,(下段)復元した三次元形状
(左)1セント硬貨,(右)指先(指紋)

9. 顕微鏡への応用

提案手法は、固定した格子パターンの投影により形状計測が可能である。その単純さにより、同様の手法で様々な機器での適応が可能である。その一例として、立体顕微鏡を用いた計測例を示す。構築した顕微鏡システムを図3(a)に示す。双眼の顕微鏡の片側の接眼レンズにマスクパターンを設置し、後ろから光を当てることで計測対象物体にパターン光を投影する。もう一方の接眼レンズから覗くと図3(b)のようにパターンが投影された物体が観測される。投影されたパターンの間隔は約0.05mmである。このシステムを用いることで、0.05~0.10mmといった非常に細かな凹凸をも計測可能である。図4に本システム利用して形状を計測した例として、1セント硬貨と指紋を計測したものを示す。

10. おわりに

本研究では、動物体の形状を高精度・高密度に計測することを目的に、単色の波線グリッドを用いた三次元形状計測手法を提案した。提案手法は単一の静的パターンをプロジェクタから投影するため、動物体の形状計測も可能である。波線グリッドを用いることで、波線の交点周りに局所的特徴を付与し、パターンの暗黙的な符号化を行った。この局所特徴と、グリッドの接続情報を Belief Propagation を用いて最適化すること

で、安定した形状計測を実現し、さらに、交点以外の画素において、補間と画像マッチングによって対応を密に取得することで、画素単位の形状計測を実現した。本稿で示した、顕微鏡を用いたシステムなどの他にも、高速度カメラによる高速に変化する移動・変化する物体の形状計測も可能である。

今後は、今回開発した計測手法について、マルチメディア、医療、スポーツ、材料解析など、従来手法では計測が十分に行われなかった分野への応用を進めるなど、本手法の応用範囲を広げていく予定である。また、カメラ・プロジェクタの台数を増やすことで全周の形状計測へ拡大も目指す。

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 (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 (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 (.2009).
- [5] J. Salvi, J. Battle, 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 (2011)
- [7] S. Zhang, P. Huang: "High-Resolution, Real-time 3D Shape Acquisition", Conference on Computer Vision and Pattern Recognition Workshop (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 (1993)
- [9] P. Vuylsteke, A. Oosterlinck: "Range image acquisition with a single binary-encoded light pattern", IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.148-164 (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 (2006)