

高密度 3 次元形状復元のための単色波線パターンを用いた グリッド能動ステレオ

阪下 和弘[†] 佐川 立昌^{††} 古川 亮^{†††} 川崎 洋^{††††} 八木 康史[†]

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

^{††} 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

^{†††} 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 丁目 21-40

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

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

あらまし 本論文では、テクスチャを持つ動物体の再構成手法を提案する。提案手法はプロジェクタ・カメラ系を利用し、投影された静的パターンによって単一画像からの復元を行う。これにより、高フレームレートでの動物体計測が可能である。提案手法の目標は、(1) 単色パターンによるワンショット三次元計測の実現、(2) サブピクセル精度の対応点推定による高精度計測、及び、(3) 環境光源の影響下での計測の実現である。これらの目標を実現するために、以下の手法を提案する。すなわち、(1) 波線パターンのグリッドによる暗黙的な対応情報の利用、(2) プロジェクタの投影パターンとカメラ画像とのグリッドベースステレオ法による一意的な対応推定、(3) (準)画素単位の補間および最適化による密な形状計測、および(4) 単色パターンの利用による物体表面テクスチャに安定な処理と利用可能なプロジェクタの種類の拡大である。実験により、提案手法が上記目標を実現していることを示す。

キーワード 3次元形状復元, ワンショット, 能動ステレオ, 波線グリッド

1. はじめに

最近、動く 3 次元シーンの復元が注目されている。例えば、人体を実時間計測してその動きを解析することで、デバイスの装着が不要なインタフェースを実現するゲーム用製品が大きな成功を納めている [1]。さらに、そのような製品を自律移動するロボットの目として利用する研究も進められており、動物体計測の重要性が強く認識されるようになってきている。このような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 3 次元スキャナほど高精度・高密度な計測ができないが、もう一段の精度および密度の向上が実現できれば、医療応用や流体解析など、応用範囲が格段に広がると考えられる。

動いている対象の形状計測を行う方法は、カメラのみを用いるステレオ手法から Time-of-Flight (TOF) 方式によるレーザスキャンまで多数存在する。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手法は、動いている物体の形状データを得るのに適しているとして最近熱心に関係・研究されている [1]~[4]。

構造化光を投影する手法は、大まかに時間符号化法と空間符号化法に分類される。特に空間符号化法は、単一画像から形状復元する (ワンショットスキャン) 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。このため盛んに研究されているが、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 2 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ひずみ等による復号化誤りなども起きやすい。

効率的に対応情報をパターンに埋め込む方法として、色情報を使うことが考えられる。複数の色を用いて、複数のビット情報を各点に埋め込む方法は広く利用されてきた [3], [5]~[8]。しかし、色情報を利用する場合、目標物体の表面は RGB の各色要素を十分に反射する必要がある、また、市販のプロジェクタでは各色要素間のスペクトル分布が干渉するため、各画素の色の決定には誤差が生じやすい。このため、色を利用しない空間符号化手法として、点線パターンやグリッドパターンなどを利用する方法が提案されているが、これまでのところ曖昧性や密度低下の問題を完全には解決できていない。本論文では、単色の波線によるグリッドパターンを用いることでこれを解決する。さらに、復元アルゴリズムとしてグリッドベースステレオを提案する。これを実現するための具体的な手法を、以下に述べる。

波線によるグリッドパターン：波状のグリッドパターンを利用することで、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る。提案手法では、位置情報がパターンに陽に符号化されるのではなく、パターンの特異性が、パターンマッチングによって、対応点候補の評価に暗黙的に利用される点が特徴である。この特異性は、本手法がステレオに基づく手法のため精度向上に大きく寄与する。

プロジェクタ・カメラ系へのステレオ法の拡張：提案手法は、カメラ対による受動計測手法における、画素単位のステレオ法をプロジェクタ・カメラ系に拡張したものとみなすことができる。本手法では、撮影画像上でグリッドのラインが接続されていれば、正規化を含む大域的最適化が実現される。

画像マッチングを利用した画素単位の補間：グリッドパターンによる形状復元の計測密度は疎であるが、本論文では(準)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り、密な画素についての奥行き推定が可能である。

単色かつ1枚のみの静的パターン：提案システムでは単色のパターンを1枚だけ利用するため、目標物体のテクスチャに対して頑健性が高い。また、これにより、利用可能なプロジェクタの種類も多くなる。例えば、レーザープロジェクタを用いれば、環境光による計測への影響を減らすことが可能となる。

2. 関連研究

能動計測手法としては、TOF スキャナや能動ステレオによるシステムが一般に良く知られている。動物体の能動計測手法も、それぞれによる手法が研究されている。

多くの TOF レーザスキャナでは、計測したい対象に向けて点レーザが照射され、それが検出器に戻るまでの時間が計測される。このため、一度に一点のみが計測され、広い範囲を短時間で計測することには適さない。そこで、動物体などを計測するために、広範囲に対して時間的に変調された光を照射し、二次元センサーの各画素でその変調を観測することで、距離画像の計測を実現する装置がある [9], [10]。しかし、現状では他の光源による外乱に弱く、解像度も通常のカメラと比較して低い。

一方、能動ステレオによる計測手法では、多くの場合、点あるいはラインレーザを照射し、これを走査することで対象を計測する。このため、計測に時間がかかることから動物体の計測には向いていなかった。ビデオプロジェクタなどの面状光源を利用することで計測時間が短縮できるが、このためには点や線光源の場合と違って対応点問題を解く必要がある。その解決方法としては、大きく時間符号化方式と空間符号化方式がある [5]。

時間符号化方式では、複数のパターンを投影し、パターンの各点の時間変化の中に情報が符号化される。このため、動物体計測には適さないと言えるが、この欠点を補うためにいくつかの方法が提案されている。例えば、パターンを高速に切り替える方法 [11]、位相パターンを用いて、必要な枚数を減らす方法 [12]、あるいは、DMD パターンを使う方法 [13] 等である。また、通常の能動ステレオとは少し異なるアプローチとして、カメラを2台以上用いて、時間変化するパターンを照射する、時空間ステレオ法 [14] 等も提案されている。現在、動き推定を組み合わせる等して 100fps 程度での計測に成功した例も

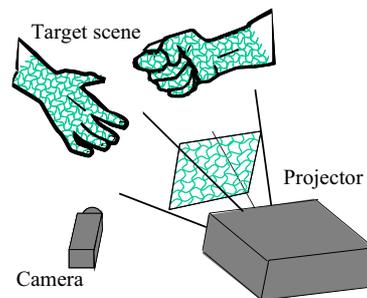


図 1 スキャンシステム：複数の波線が投影され、交点が検出され復元に利用される。

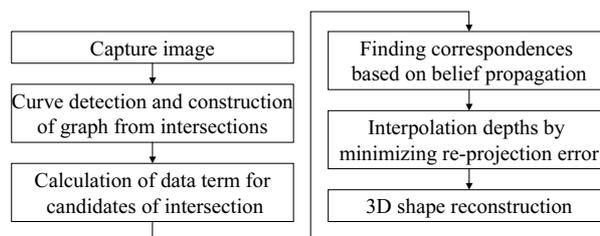


図 2 アルゴリズムの概要

紹介されているが、複数フレームの情報を必要とするため、高速な物体の計測には本質的に適さない。

一方で、空間符号化方式では、静的パターンを用いて、一枚のみの入力画像から形状を復元するため、動物体計測に向いている。しかし、パターンの空間的な領域に情報を埋め込む必要性から解像度が低く、かつ、対象物体表面のテクスチャや形状によりパターンが歪むため、不安定になりやすい。このため、これを軽減する方法が多数提案されている。例えば、同じ組み合わせの無いように複数の色の帯を用いる方法 [15], [16]、特異な点線を利用する手法 [17], [18]、二次元的なパターンに情報を埋め込む手法 [1], [19] 等がある。しかし、精度・解像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る手法は未だ存在していない。

提案手法はグリッドパターンの接続性を利用した空間符号化による手法である。この方法は、対応点の曖昧性と、検出された線の接続判定の誤りによる誤復元の問題があることが知られている [2] ~ [4]。従来法では、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。しかし、これらの手法では、対象物体表面の反射率やテクスチャに影響されるため、計測が安定しない。提案手法では、単色のグリッドを用いることにより、グリッドパターンと多色パターンの両方の問題を同時に解決することができる。

3. 手法の概要

提案手法のシステムは、プロジェクタとカメラの一台ずつからなる (図 1)。プロジェクタは図 3(b) に示す静的パターンを投影する。このパターンは、単色で、正弦波形状の縦横の曲線をグリッド状に配置したものである (詳細については 4. 章で説明する)。パターンは静的であ

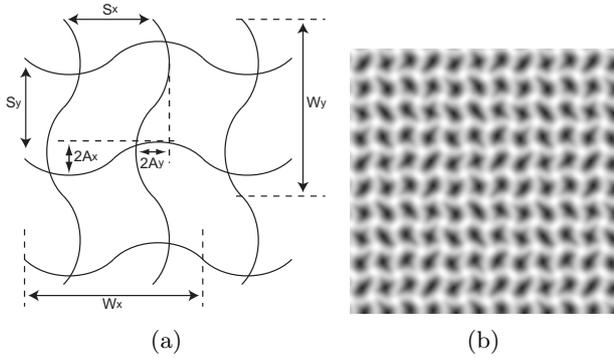


図 3 (a) 波線グリッドのパラメータ. S_x, S_y は隣接する波線の間隔, W_x, W_y は波線の波長, A_x, A_y は波の振幅である. (それぞれ縦, 横線に対応する). (b) 波線グリッドの例.

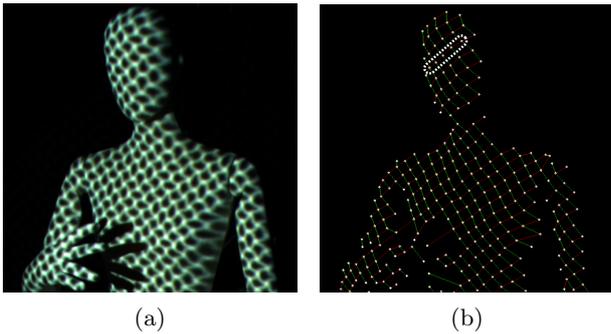


図 4 縦横の波線検出: (a) 入力画像, (b) 検出されたグリッド. 破線の円内部の点は, 同一の横線 ID を持つ.

るため, 撮影に同期は必要ない. そのため, 非常に高いFPSでの計測が可能である.

提案手法による復元アルゴリズムを図2に示す. まず撮影した画像から線検出を行う. これはICCV09に佐川らによって提案された手法[3]を用いており, Belief propagation(BP)による最適化により, 単色のグリッド状の線を, 縦と横に安定して別々に検出することができる. 検出された縦横の線から交点を算出し, 交点をノードとしたグラフを作る. 次に各ノードに対応するエピソード線の位置をプロジェクタパターン上で計算し, そのライン上にグリッドの交点がある場合, これを対応候補とする. この時, 通常は複数の対応候補が見つかるので, BPを利用して各点における対応候補の最適な組み合わせを求める. このままでは復元結果は疎であるので, 最後に, 各画素での深さを, 補間とパターンと観測画像の画素単位のマッチングを利用して求めることで, 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

プロジェクタとカメラは幾何的には同じ枠組みで考えることができる. 本論文では, プロジェクタ画像面から投影するパターンはプロジェクタ画像, カメラ画像面で撮像される画像をカメラ画像と呼ぶことにする.

空間符号化によって, カメラ画像とプロジェクタ画像間で一意な対応付けを得るために, 従来法では大きな窓

サイズを持つ複雑なパターンが必要であった. さらに, 精度の向上には広い基線長が望ましいが, 観測されるパターンは大きく歪むため, 現実的には復号化が困難になる. したがって, 検出と復号化が簡単な, 単純でありながら一意性の高いパターンが望ましい.

本論文では, 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をもちいる. 具体的には, 縦, 横の波線からなるグリッドパターンを用いる. 波線パターンは単純なパターンであるため, 画像中での曲線として検出しやすく, 輝度値のピークを計算することにより, その位置をサブピクセル精度で得ることができる.

波線として周期的な正弦波パターンを用いるが, 波線グリッドは対応点検出に有効な情報を持つ. 提案手法では縦, 横の波線の交点を特徴点として用いる. 交点の配置は, 波線の間隔と波長で決定される. 一定の間隔と波長を持つ波線を用いるが, 下記に述べるように, 縦波線の間隔が横波線の波長の整数倍でない場合, 交点位置の位相がずれが生じる. そのため, 局所的なパターンが周囲の交点と相違することになり, 対応付けのための特徴量として用いることができる.

交点周囲の局所パターンは, 投影パターン全体の中で一意性を持つものではない. 図3(a)において, S_x, S_y は隣接する波線の間隔, W_x, W_y は波長とすると, $N_x = \text{lcm}(S_x, W_x)/S_x, N_y = \text{lcm}(S_y, W_y)/S_y$ を用いて, 同一パターンが, 横, 縦軸に沿ってそれぞれ N_x, N_y 本の波線ごとに起こる. ここで $\text{lcm}(a, b)$ は a と b の最小公倍数であり, 添字 x, y はそれぞれ横, 縦軸に沿った値を表すものとする. しかしながら, 局所パターンは, 各サイクルにおいて識別可能なパターンである. 図3(b)は, $S_x = 10, S_y = 11, W_x = W_y = 14, A_x = A_y = 1$ (単位は画素) からなるパターンの例である. この例では, 1周期は縦線7本, 横線14本となる. よって, $98 (= 7 \times 14)$ 種類の交点が1周期で構成される矩形の中に存在する.

ステレオマッチングでは, 対応点候補はエピソード線上の点に限られている. あるプロジェクタ画像の交点とエピソード線が適当な距離以内に位置している場合, そのプロジェクタ画像の交点是对応点候補の1つとして選ばれる. 候補の数は, カメラ画像の交点位置に依存する. 対応点候補は, プロジェクタ画像において疎に分布するため, 画素単位で候補点を探索する通常のスレオ視と比べて, 対応候補の数は劇的に少なくなる.

5. グリッド上でのプロジェクタ・カメラ系ステレオ法

対応点の最適な組み合わせを見つけるために, 局所的なマッチングと正則化を利用した新しい最適化手法を提案する. 提案手法では, 波状パターンのグリッドを利用し, 図4(b)に示すような線検出を行う. カメラ画像上での波状グリッドにおける縦横線の交点を以後格子点と呼

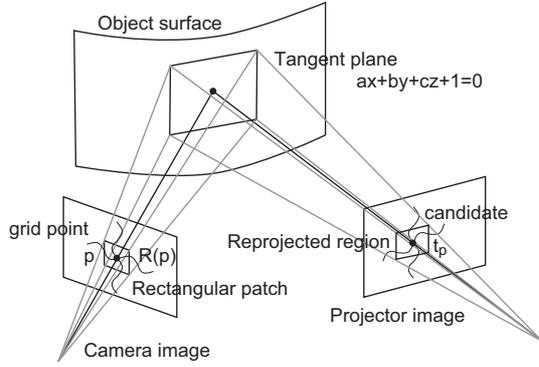


図5 格子点周囲の四角パッチ p がプロジェクタ画像面に再投影される． t_p は p の対応候補である．

ぶ．もし複数の格子点がグリッド線で接続されていれば，それらはプロジェクタ画像上で同一の波線の上にある．このことは，対応点を決定する際の正規化に利用される．格子点の接続性は線検出によって得る．ただし，線検出による接続性の判定は誤る場合がある．特に，形状の不連続な境界付近では，このような誤りが多くなる．このため，我々は，各格子点に対応点を割り当てるために，グリッド上に定義したエネルギーの最小化を利用する．

5.1 エネルギー最小化のためのデータ項

まず，我々はすべての対応候補についてマッチングコストを計算し，これをエネルギー最小化におけるデータ項として利用する．コストは，撮影された画像と，プロジェクタ画像（パターン画像）との SSD として計算する．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対象物体の形状によって歪むので，四角形の領域に対する単純な SSD はデータ項には適さない．そこで，我々は，対象物体における格子点において，その周辺の領域を格子点の接平面で近似したパッチを考える．これにより，高精度なマッチングコストを計算し，対応点をサブピクセル精度で計算する．

図5に示すように，対象表面における格子点の周辺形状が局所的に平面であると仮定すると，その平面は $ax + by + cz + 1 = 0$ と表される．ただし， a, b, c は平面パラメータである．このパラメータを画像のひずみを考慮して SSD 最小化により求める．そのアルゴリズムを以下に示す：

- (1) カメラ画像における格子点周囲の四角パッチを3次元接平面に投影し，プロジェクタ画像に再投影する．
- (2) 再投影されたパッチとプロジェクタ画像との輝度値 SSD を計算する．
- (3) a, b, c を変数として SSD を最小化する．
- (4) 上記のステップを数回繰り返す．

a, b, c の初期値は，接平面が，視差によって計算される格子点の三次元位置を含み，カメラの画像面に対して平行になるように設定され，SSD は以下の式で表される．

$$SSD_{a,b,c}(p) = \sum_{p' \in R(p)} (I_c(p') - I_p(H_{a,b,c}(p')))^2. \quad (1)$$

ただし， $R(p)$ は p 周辺の四角パッチであり， $H_{a,b,c}(p')$ は p' をプロジェクタ画像面に再投影した場合の変形を表す． $I_c(\cdot)$ と $I_p(\cdot)$ はそれぞれカメラ画像とプロジェクタ画像における輝度値を表す．

5.2 グリッド上でのエネルギー最小化による対応推定

ここで，グリッドは，格子点から成るノード $p \in V$ と，格子点の接続を表すエッジ $(p, q) \in U$ から成る．ただし p, q は格子点であり， V は格子点の集合であり， U はグリッドグラフのエッジ集合である．格子点 p は，対応点の候補 $t_p \in T_p$ を持つ．ここで T_p は格子点 p の対応点候補の集合である．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：

$$E(T) = \sum_{p \in V} D_p(t_p) + \sum_{(p,q) \in U} W_{pq}(t_p, t_q). \quad (2)$$

ただし $T = \{t_p | p \in V\}$ であり， $D_p(t_p)$ は p の対応点を t_p に割り当てる場合のデータ項である． $W_{pq}(t_p, t_q)$ 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る．

上記データ項は，前節で述べた方法で計算された SSD の値を取る．正規化項は以下の式で定義される．

$$W_{pq}(t_p, t_q) = \begin{cases} 0 & t_p \text{ と } t_q \text{ が同一波線上} \\ \lambda & \text{それ以外の場合} \end{cases} \quad (3)$$

ただし λ はユーザが決める定数である．定義されたエネルギーは，BP 法 [20] によって最小化される．

エネルギー最小化を利用する利点として，上記の隣接による正規化項が「ソフトな条件」となる点が挙げられる．実際のデータでは，線検出の誤りなどにより，誤った接続が生成される場合があるので，上記の点は重要である．先行研究 [3] では，誤った接続を，3次元再構成の前の線検出段階で排除する必要があるが，提案手法では，誤った接続の除去と3次元再構成は同時に実行されるため，より高密度，高精度な復元が可能である．

6. 格子点間の補間による密な形状生成

グリッドベースドステレオ法により，疎な格子点における対応が得られる．次のステップとして，全ての画素の情報を利用して密な対応を求める．このためにまず，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さを求める．次に，これらの標本画素の深さを変数として，カメラ画像とプロジェクタ画像の輝度の差を最小化する．既に Sagawa らも先行研究 [8] で高密度な再構成を提案しているが，これは検出されたグリッド線間の補間に基づくものであった．これに対し，提案手法では，(準)画素それぞれに独立した深さ推定が，photo-consistency に基づく最適化によって実現される．

続く節では，より詳細に最適化の手順を述べる．

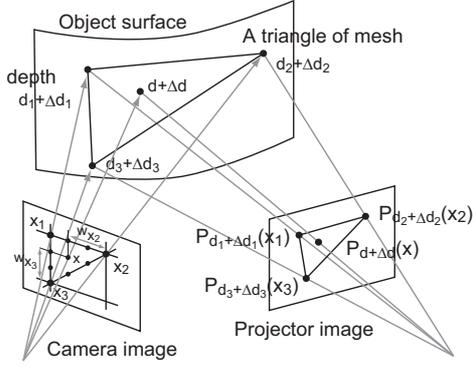


図 6 x_1, x_2, x_3 は再標準化された標本画素である．全画素の深さは標本画素の補間によって求められる．各画素は深さ情報を利用してプロジェクタ画像面に再投影される．再投影は、 $P_{D+\Delta D}(x)$ で表される．各画素の再投影には $D + \Delta D$ の部分ベクトルが利用される．

6.1 全画素における深さの計算

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、その画素の深さ d_x は

$$d_x = \frac{-1}{a_x u + b_x v + c_x}, \quad (4)$$

で計算される．ここで、パラメータ a_x, b_x, c_x は画素ごとに計算される．各画素における a_x は以下のように補間される：

$$a_x = \frac{\sum_p G(|p-x|) a_p}{\sum_p G(|p-x|)}. \quad (5)$$

ただし、 p は格子点、 $G(\cdot)$ はガウス関数、 $|p-x|$ は p と x の距離である． b_x と c_x は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

最適化において、全ての画素の深さを、独立変数として扱い、全画素の深さを推定することは可能である（画素単位の深さ推定）．しかし、本論文では、画素を縦横について3画素ごとに再標準化し、これらの標本画素について深さを推定する（準画素単位の深さ推定）．これにより、得られるメッシュの解像度は、より適切なものになる．なぜなら、通常、投影するパターンの解像度は画像解像度より低いので、単純に全ての画素について深さ推定を行うと、かえって精度低下を招くことがある．これを解決するためには、受動ステレオ法で行われるように、ある程度の範囲を持つマッチング窓を利用する等の方法が考えられるが、計算コストが増大する．これに対し、提案手法では以下の手法により精度を犠牲にせず、点と変数の数を減らし、効率的な計算を実現する．

まず、図 6 に示すように、標本画素を頂点として、三角メッシュを作成する．標本画素以外の画素における深さは線形補間される．繰り返し計算による最適化のために、深さを、 d_x の微小変位 Δd_x を変数として近似する．図 6 における画素 x の深さは以下のように表される．

$$d_x + \Delta d_x = [1 - w_{x_2} - w_{x_3}, w_{x_2}, w_{x_3}] \begin{bmatrix} d_{x_1} + \Delta d_{x_1} \\ d_{x_2} + \Delta d_{x_2} \\ d_{x_3} + \Delta d_{x_3} \end{bmatrix}.$$

ただし、 w_{x_2} と w_{x_3} は線形補間の重みである．

ここで、 $D + \Delta D$ を、 $d_x + \Delta d_x$ を全ての標本画素について集めたベクトルとする．プロジェクタ画像（パターン）に対する再投影誤差は、標本画素ばかりでなく、全ての画素について以下の式で求められる．

$$E(\Delta D) = \sum_x (I_c(x) - I_p(P_{D+\Delta D}(x)))^2 + \gamma \sum_{x,x'} (\Delta d_x - \Delta d_{x'})^2. \quad (6)$$

ここで $P_{D+\Delta D}(x)$ はプロジェクタ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' は隣接する頂点である． γ は利用者が決める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される．再投影と最小化を、解の収束まで交互に繰り返すことで、 D を決定する．

7. 実験

7.1 シミュレーションデータによる評価

我々は、提案手法をシミュレーションデータによって評価し、色を2色用いる手法 [8] と比較した、結果を図 7 に示す．実験では、Stanford 大学の形状データ集 [21] の bunny データを目標形状として利用した．目標形状には、上列 (a) ではテクスチャが無く、中列 (f) では格子模様がマッピングされている．提案手法への入力画像を、レイトレーシングソフト [22] で生成したものを (b)(g) に示す．図 7(c),(h) は、頭部における、提案手法のグリッド検出結果を示す．頭部、耳、胴体の境界において、グリッドの接続性が誤って検出された所があるが、これらはステレオマッチングの処理において正しく切断されている．図 7(d)(i) は2色を利用する手法 [8] の入力画像であり、2色の線を3本用いて8周期の局所的ID情報が符号化されている．この手法は、目標物体にテクスチャが無い場合には (e) に示すように成功するが、(j) に示すようにテクスチャがある場合には、パターンの色情報が乱され、ID情報の復号化と3次元復元に失敗した．

上記実験の定量的な評価を行うために、(b),(d),(g) の対応の誤差を計算した．カメラ画像の各画素に対応するプロジェクタ画像の座標は既知であるので、復元結果において推定された対応点と正解の対応点との誤差を画像面上での距離で算出した．図 7(k), (l), (m) に (b), (d), (g) の誤差をそれぞれ画像化したものであり、明るい画素は誤差が大きいことを示す．同図 (n) に示された表に、画素を単位とした二乗誤差の平方根 (RMSE) を示す．RMSE の値は、復元された全ての対応点について計算されたものと1画素を超える外れ値を除いた対応点に

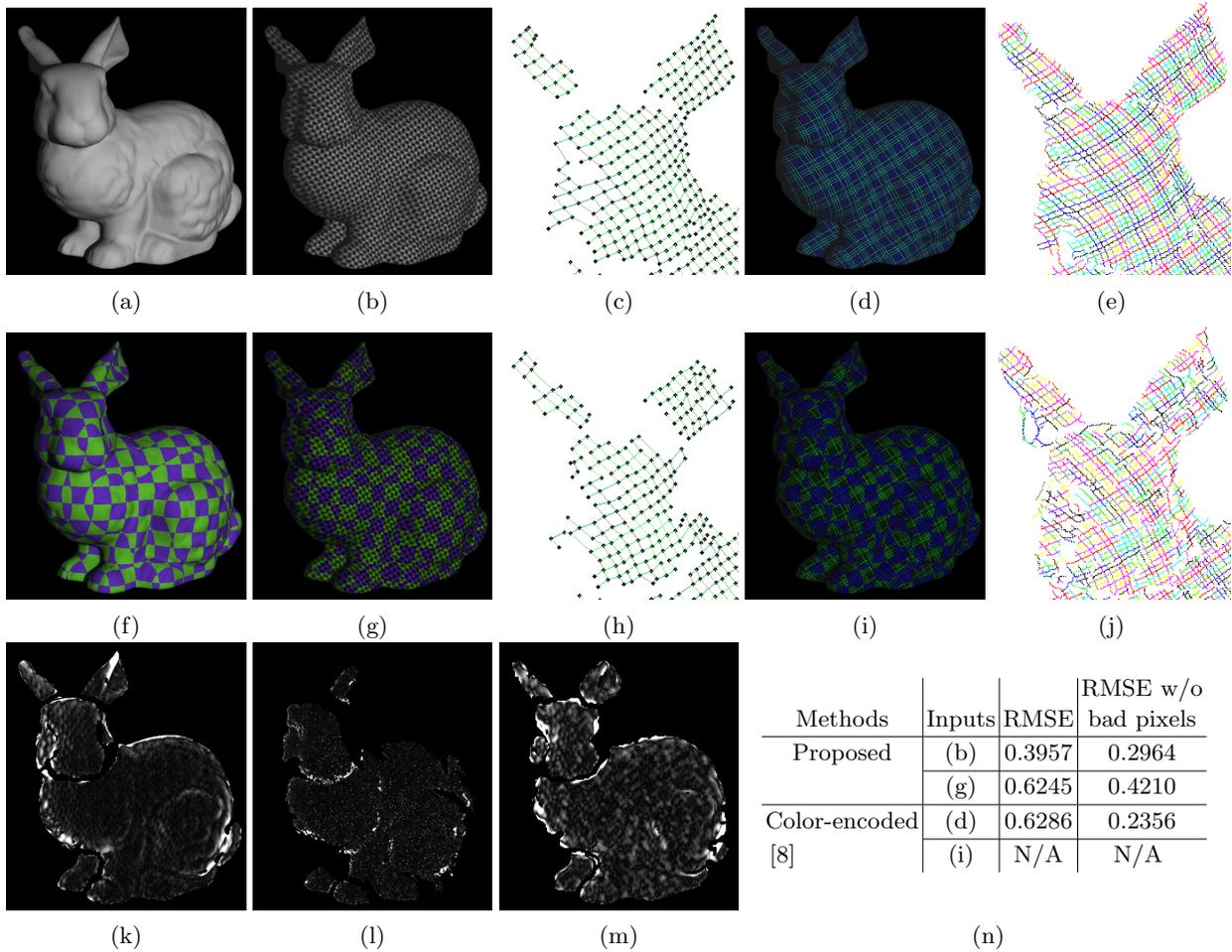


図7 提案手法と色符号化による方法 [8] の比較 .

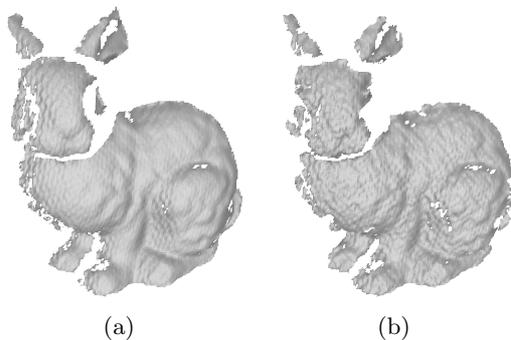


図8 グリッドベースステレオで再構成されたメッシュと補間によって高精度化されたメッシュ. 図7(b)は (a) の入力画像であり, 図7(g)は (b) の入力画像である .

ついて計算されたものである . この表から , テクスチャが無い場合 , 全ての画素についての RMSE では , 提案手法が色を利用する手法 [8] よりも良いのに対し , 外れ値を除いた RMSE では色を利用する手法 [8] が提案手法よりも良かったことがわかる . この理由は以下のように考えられる . 色を利用する手法では , 周期的に現れる線パターン of 局所 ID (位相) から対応を求めるので , 局所 ID 情報が正確に計算されれば精度が高い . しかし , 局所 ID の復号に失敗すると , 大きな誤差が生じる . このような誤差は図7(l)のごま塩状の雑音として観測される . また , 格子模様がマッピングされている場合 , 色を利用する手

法では復元そのものが失敗したのに対し , 提案手法では誤差が少し大きくなったものの形状復元に成功した . したがって , 提案手法は , 特にテクスチャがある場合 , 色を利用した方法よりも頑健で精度が高いと言える .

図8に , 提案手法によって復元されたポリゴンメッシュを示す . 入力画像は図7(b),(g)である . 図8の形状は補間による密な復元の結果である . 本実験の条件ではカメラ , プロジェクタ間のベースラインが大きく , 例えば100画素程度の視差が存在するが , ステレオ復元では探索範囲を指定することなく正しく対応が得られた . さらに , 補間と最適化により密な対応点を得ることができた .

7.2 実データによる実験

実データによる実験では , 1600×1200 画素のカメラと 1024×768 画素のプロジェクタを使用した . 画像系列は30FPSで撮影され , Intel Core i7 2.93GHz と NVIDIA GeForce 580GTX を備えた PC を利用した . 5. 章および6. 章で述べたアルゴリズムは CUDA で実装された . 線検出は , CPU 上にシングルスレッドで実装された .

まず , 提案された波状パターンの効果を示すために , 波状パターンと直線パターンを比較した . 提案された波状パターン (図3(b)) で復元された3次元再構成結果を図9に示す . 図9(a)に入力画像を , (b)に , 5. 章のプロジェクタ・カメラ系ステレオ法による復元結果を示す .

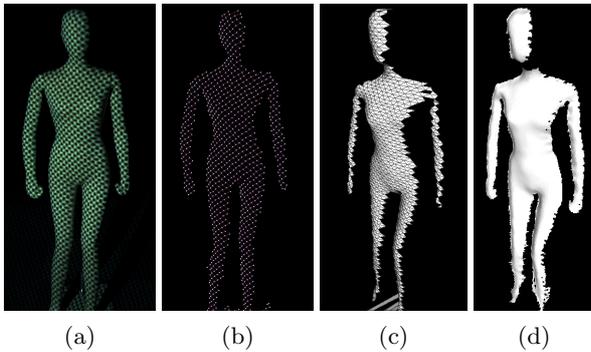


図9 提案された波パターンによる3次元再構成：(a) 入力画像，(b) ステレオマッチング後の検出グリッド (c) グリッドの再構成結果，(d) 補間後の再構成結果．

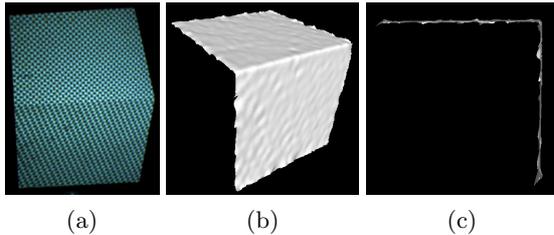


図10 立方体形状の再構成結果：(a) 入力画像，(b) 3次元再構成結果，(c) 他方向から見た形状．

物体の不連続部分（マネキンの頭部と首の境界）におけるグリッド線は、ステレオ処理の過程において正しく切断された。図9(c)に提案手法による3次元復元の結果を示す。グリッド点の数は943であり、各グリッド点における対応点候補数の平均値は41であった。ステレオ処理にかかった時間は0.22秒であった。探索範囲は画像全体に設定したにも関わらず、グリッドパターンが画素数と比べて疎であることにより、計算コストは小さかった。さらに、6.章の補間手法により、図9(d)に示すような密な形状が生成された。三次元モデルの頂点数は25,938であった。最適化における繰り返し回数は5であり、補間処理にかかった時間は0.59秒であった。線検出を含めた総計算時間は4.87秒であった。

次に、図10に示す立方体形状の計測により、提案手法の精度を評価した。立方体の大きさは一辺0.2mであり、カメラからの距離は1.0mであった。復元された立方体の各面に平面を当てはめ、各平面からの誤差をRMSEで算出した。2個の面のRMSEの平均値は0.36mmであり、面との角度は88.8度（正解は90.0度）であった。この誤差は、実用的に十分小さいと考えられる。

単色の静的パターンの重要な利点として、パターンを照射するデバイス選択の幅が広がることが挙げられる。そこで、単一波長の光を投影するレーザプロジェクタを利用した復元実験を行った。投影される光のエネルギーが狭帯域に集中しているため、適当なバンドパスフィルタを用いることで、環境光の影響下においても、投影パターンを観測できる。図11に、実験結果を示す。(a)は実験環境であり、目標が外部光源から強く照らされていることがわかる。しかし、(b)に示すように、バンドパス

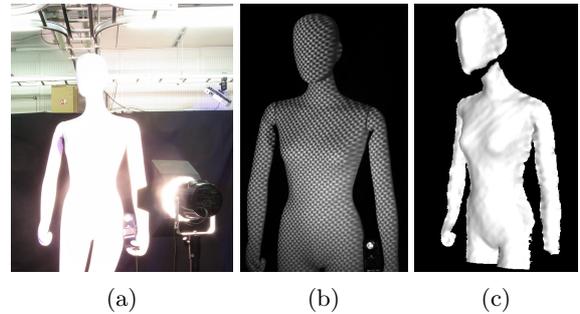


図11 環境光の影響下における再構成結果．

フィルタにより、照射されたパターンが明確にされ、(c)に示すように、3次元復元を正しく行うことができた。

最後に、運動する物体の計測結果を示す。図12と図13は、手の開閉の運動と、パンチ動作中の人体の動きを計測したものである。提案手法はワンショット復元であるため、対象物体が高速に動いていても、フレームごとに独立して3次元復元を行うことができる。

8. 結 論

本論文では、単色かつ静的なパターンの照射による、単一画像からの三次元形状復元（ワンショット復元）方法を提案した。先行手法と異なり、提案手法では、波線のグリッドパターンにおける各交点のパターンの違いにより対応情報を暗黙的に表現する。そこで、パターンの規則性を乱すことにより、パターンの特異性が大きくなり、安定した解が得られる。また、本論文ではステレオ法による形状復元を、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方法を提案した。復元の最終段階では、グリッドによる復元を補間して、画素単位の深さ推定が行われる。実験では、既存の2色用いるワンショット計測手法と比較して、より安定した結果が得られること、また、移動物体の計測が効果的に行われることを示した。今後の課題として、複数のシステムを利用した全周囲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] “Xbox 360 Kinect,” 2010. <http://www.xbox.com/en-US/kinect>.
- [2] “Dynamic scene shape reconstruction using a single structured light pattern,” CVPR, pp. 1–8, June 23–28 2008.
- [3] “Dense 3d reconstruction method using a single pattern for fast moving object,” ICCV, 2009.
- [4] “One-shot scanning using de bruijn spaced grids,” The 7th IEEE Conf. 3DIM, 2009.
- [5] “A robust-coded pattern projection for dynamic 3D scene measurement,” year, year, Pattern Recognition, pp. 1055–1065, 1998.

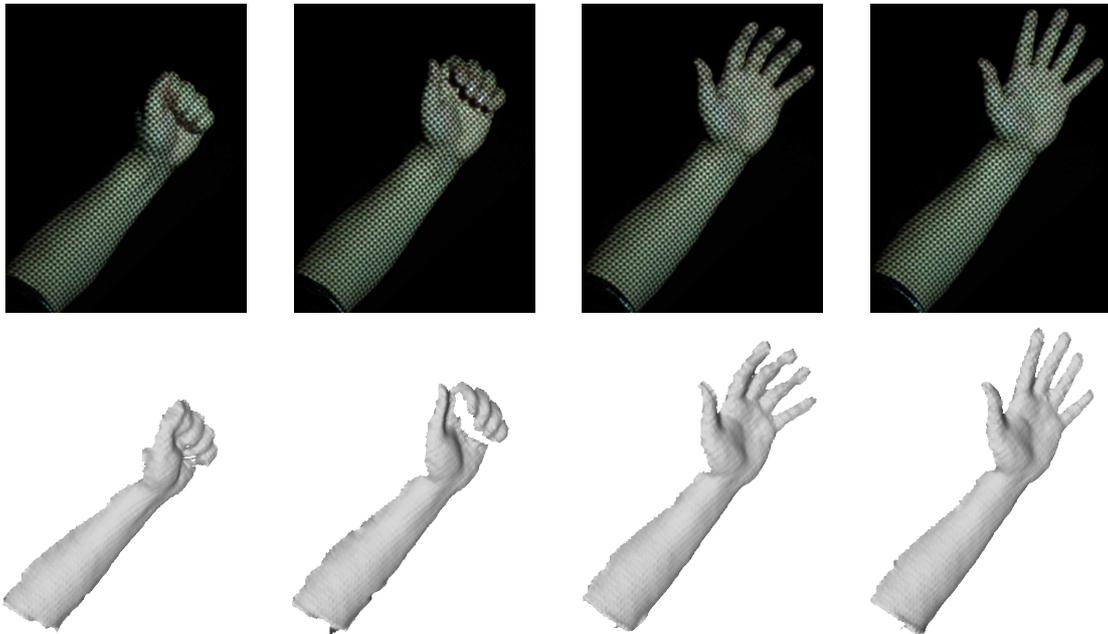


図 12 指を開く動作を再構成した手の三次元形状 .



図 13 パンチ動作を再構成した人体の三次元形状 . 服の詳細形状が計測されている .

- [6] “High-contrast color-stripe pattern for rapid structured-light range imaging,” year, year, ECCV, pp. 95–107, 2004.
- [7] “Rapid shape acquisition using color structured light and multi-pass dynamic programming,” 3DPVT, pp. 24–36, 2002.
- [8] “Dense one-shot 3d reconstruction by detecting continuous regions with parallel line projection,” ICCV, 2011.
- [9] “CanestaVision EP Development Kit,” 2010. <http://www.canesta.com/devkit.htm>.
- [10] “SwissRanger SR-4000,” 2011. <http://www.swissranger.ch/index.php>.
- [11] “Real-time 3D model acquisition,” Proc. SIG-GRAPH, pp. 438–446, 2002.
- [12] “Fast 3D scanning with automatic motion compensation,” CVPR, 2007.
- [13] “Temporal dithering of illumination for fast active vision,” Proc. European Conference on Computer Vision, pp. 830–844, October 2008.
- [14] “Spacetime stereo: Shape recovery for dynamic scenes,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 367–374, June 2003.
- [15] “3-D data acquisition by rainbow range finder,” ICPR, pp. 309–313, 1990.
- [16] “High-resolution, real-time 3D shape acquisition,” Proc. Conference on Computer Vision and Pattern Recognition Workshop, p. 28, 2004.
- [17] “Range sensing by projecting multiple slits with random cuts,” year, SPIE Optics, Illumination, and Image Sensing for Machine Vision IV, pp. 216–224, 1989.
- [18] “United States Patent Application 2009005924,” 2007j.
- [19] “Range image acquisition with a single binary-encoded light pattern,” year, IEEE Trans. on PAMI, pp. 148–164, 1990.
- [20] “Efficient belief propagation for early vision,” year, IJCV, pp. 41–54, 2006.
- [21] “The Stanford 3D Scanning Repository,” <http://www-graphics.stanford.edu/data/3Dscanrep/>, 2012.
- [22] “POV-Ray,” 2004.