

汎用プロジェクトを用いる高精度 レンジファインダのためのキャリブレーション

Calibration Method for a Highly Precise Range Finder using Versatile Projector

田中文武, 出口光一郎, 岡谷貴之

Fumitake Tanaka, Koichirou Deguchi, Takayuki Okatani

東北大学大学院情報科学研究科

Graduate School of Information Sciences, Tohoku University.

キーワード: 基礎行列 (fundamental matrix), 内部パラメータ (internal parameters),
外部パラメータ (external parameters), 自由度 (degree of freedom)

連絡先: 〒980-8579 仙台市青葉区荒巻字青葉6-6-01 東北大学大学院
情報科学研究科 システム情報科学専攻 出口・岡谷研究室

Tel: (022)795-7017, Fax: (022)795-7017, E-mail: tanaka@fractal.is.tohoku.ac.jp

1. はじめに

三次元形状計測用レンジファインダとは, 物体の立体形状を非接触で取得しコンピュータに取り込む装置である。人の顔照合, 整形や矯正のシミュレーション, 文化財の形状の解析や保存, 部品の検査や設計など, 様々な用途に使われている。

現在, このようなレンジファインダの製品が数多く市販されているが, それらは一般に高価であり, また装置も大掛かりとなっているものが多い。その主な理由は, 通常, レンジファインダの主要構成部品であるカメラとプロジェクトが, 装置内に固定された設計となっていることによる。プロジェクトによる対象表面へのパターンの投影像をカメラで撮影して, ステレオの原理によって三次元情報を取得するため, カメラとプロジェクトの距離 (基線長) はある程度大きくとる必要がある。また, カメラとプロジェクト間の位置姿勢決めの

精度がそのまま計測精度につながるので, 両者を精度良く固定することが必要だということもある。

近年, 映像投影用の汎用データプロジェクトが, 安価に入手できるようになってきている。これらのプロジェクトの画像表示能力は, 年々画素数が増えるなど向上し, しかもコンパクトになってきている。このような汎用データプロジェクトを, 市販されているカメラとPCを組み合わせれば, 非常に安価にレンジファインダを構成することができる。精度に限って言えば, 上述のような専用の製品に引けをとらないものが実現できる。

ただし, このような汎用プロジェクトとカメラの組み合わせによるレンジファインダの構成には問題があって, それは, カメラとプロジェクトの位置姿勢関係を求めるためのキャリブレーションの問題である。上述のような製品と異なりプロジェクトとカメラの位置関係はまったく自由であるの

で、両者の間の位置姿勢関係を求めることは、高い精度の実現には本質的に重要である。

そのためのキャリブレーションを行う一つのやり方は、形状が正確にわかっている物体（たとえば立方体など）を用いて、事前にキャリブレーションを行うことである。しかしながら、この場合、プロジェクタやカメラをセットするたびにキャリブレーションの作業が必要となり、手軽さがおおいに損なわれる。そこで考案されたのが、マルチカメラを使った三次元形状復元の研究から編み出されたセルフキャリブレーションの考え方を、今のプロジェクタ-カメラのキャリブレーションに応用する方法である。この方法では、実際に形状を計測したデータのみを使って、キャリブレーションと形状計測を同時に行うもので、事前にキャリブレーション用の物体を使用する必要がなくなる。この方法には、絶対的なスケールが分からない問題はあるものの、それを除けば非常に便利な方法であり、汎用プロジェクタを使ってレンジファインダを構成する上で鍵となる技術である。

通常、プロジェクタとカメラがそれぞれ1台ずつあれば、このようなレンジファインダが実現できる。本研究では、上述のようにプロジェクタとカメラの位置関係がまったく自由である条件を逆に活かし、セルフキャリブレーションの機能を発展させることを考えた。具体的には、カメラの位置を動かして複数の画像を撮影し、その画像を用いて、レンジファインダとしての機能・性能向上の可能性を追求した（複数画像の撮影は、一台のカメラを動かすのではなく複数台のカメラを利用しても良い）。このように画像を撮影する視点の数を増やすことにより、二種類の利点が期待される。一つは、計測範囲の拡大と精度の向上という量的な性能向上であり、もう一つが、キャリブレーションで決定できるパラメータの自由度が増えるという質的な機能向上である。後者をより詳細に述べると、従来の1視点のみを用いるシステムに比べて、カメラとプロジェクタが持つ多数の光学パラメータのうち、事前に分かっていたいなければならぬものの数を減らせるということである。

以下、まず、視点の数を2つとした場合のキャリ

ブレーションの原理を述べる。さらに、その方法の有効性と、視点の数の増加に伴う精度向上の程度を確かめるため、シミュレーション実験を行ったのでその結果を示す。

2. 原理

2.1 ステレオ

レンジファインダによる三次元形状の計測を行う原理として、プロジェクタとカメラによるステレオ法を用いる。カメラが三次元空間中の点 (X, Y, Z) を撮影するとき、その点が写る画像上の座標 (u, v) とは

$$\omega \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{P} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1)$$

と与えられる。ここで、 ω は0でない定数である。また、 \mathbf{P} は透視投影行列（projection matrix, 3×4 行列）と呼ばれ、内部パラメータを表す行列を \mathbf{K} 、外部パラメータを \mathbf{R}, \mathbf{t} とすると

$$\mathbf{P} = \mathbf{K}(\mathbf{R} \mid -\mathbf{R}\mathbf{t}) \quad (2)$$

という関係がある。ここで、 \mathbf{R} は回転行列（ 3×3 行列）で \mathbf{t} は3次元の平行移動ベクトルである。この関係は、プロジェクタの光学系についても、プロジェクタが投影する画像上の点 (u, v) とその三次元空間中の投影点 (X, Y, Z) の間に全く同じ関係が成り立つ。したがって、プロジェクタの内部パラメータを表す行列を \mathbf{K}_p 、カメラの内部パラメータを表す行列を \mathbf{K}_c 、プロジェクタ座標系をワールド座標系とし、プロジェクタ画像とカメラ画像に対する透視投影行列をそれぞれ $\mathbf{P}_p, \mathbf{P}_c$ とすることで、プロジェクタとカメラの透視投影行列を

$$\mathbf{P}_p = \mathbf{K}_p(\mathbf{I} \mid \mathbf{0}) \quad (3)$$

$$\mathbf{P}_c = \mathbf{K}_c(\mathbf{R} \mid -\mathbf{R}\mathbf{t}) \quad (4)$$

とできる。式(3),(4)と式(1)の関係から、三次元空間中の同一点 (X, Y, Z) のプロジェクタの画像座標 (u_p, v_p) と、カメラの画像座標 (u_c, v_c) より (X, Y, Z) を求めることが計測の原理である。Fig.1のように

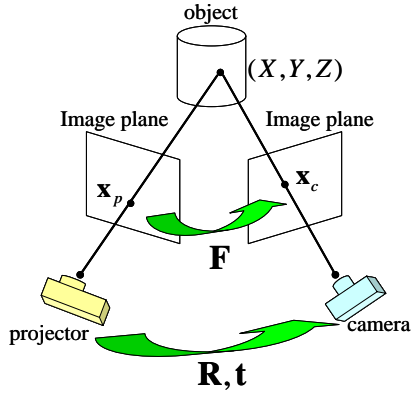


Fig. 1 ステレオによる対応点と基礎行列

プロジェクタから x_p を通る直線と、カメラから x_c を通る直線との交点が (X, Y, Z) となる。 (X, Y, Z) を求めるためには、プロジェクタとカメラの内部パラメータと外部パラメータが必要となる。

プロジェクタとカメラの内部パラメータと外部パラメータを知る方法の1つは、計測前にキャリブレーション用の物体（立方体や平面など）を使って、これらのパラメータを予め求めておくことである。この方法は、キャリブレーション精度が高く、市販のレンジファインダについてもこのような方法を用いている。一方で、計測前にキャリブレーションが必要であり、二度手間となる。特に、ここで与えているような単体のプロジェクタとカメラを組み合わせるレンジファインダを構成する場合には、大きな手間となる。

もう1つの方法として、自己校正（セルフキャリブレーション）がある。本手法では、この方法を用いる。

2.2 基礎行列

Fig.1のように、物体の表面上にプロジェクタからあるパターンを投影し、そのパターンをカメラで撮ったとき、プロジェクタ画像上の点 $x_p = (u_p, v_p, 1)^T$ とカメラ画像上での対応する点 $x_c = (u_c, v_c, 1)^T$ との間には

$$x_p^T F x_c = 0 \quad (5)$$

という関係がある。ここで、 F は基礎行列と呼ばれる 3×3 の行列である。式(5)よりスケール倍の不

定性がある（7点以上の対応点が与えられれば F を線形計算で求めることができる）。

また、式(5)は互いの対応点がある直線（エピ極線）に乗っていることを意味しており、それぞれの画像上ですべてのエピ極線はそれぞれのエピ極 e_p, e_c を通る。エピ極は基礎行列の固有値0に対する固有ベクトルとして、

$$F e_p = 0, F^T e_c = 0 \quad (6)$$

で与えられる。よって基礎行列 F は、上記したように式(5)よりスケール倍の不定性があることと、ランクが2であることから自由度は7である²⁾。つまり、この基礎行列からパラメータを推定する場合、推定することのできるパラメータの数は最大で7個ということになる。

2.3 1プロジェクタ, 1カメラシステム

1つのプロジェクタと1つのカメラを用いてレンジファインダを構成させることを考える。プロジェクタ、カメラともに未校正である。プロジェクタやカメラの内部パラメータ行列は、

$$K = \begin{pmatrix} f & sf & u_0 \\ 0 & \alpha f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7)$$

となる。ここで、 f は焦点距離、 s はスキュー、 α はアスペクト比、 (u_0, v_0) は画像中心の座標である。

また、外部パラメータの R は3自由度があり、 t は3次元のベクトルであるが、その大きさは決まらず方向のみを表すため、自由度は2となる。よって外部パラメータの推定には5つのパラメータが必要となる。よって、基礎行列から推定することのできるパラメータの数は最大で7個であるため、内部パラメータは最大で2つのパラメータしか計算できない。そのうち、スキュー、アスペクト比、そして画像中心はプロジェクタやカメラに固有であるのに対して、焦点距離は使用するたびに調整し変化させることが多いため、プロジェクタとカメラの焦点距離 f_p, f_c を基礎行列から推定するパラメータとする。それ以外は一般に不変なので事前のキャリブレーションで決定しておく。

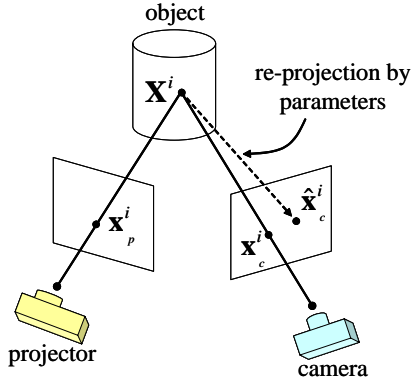


Fig. 2 パラメータより再投影したカメラの対応点

2.3.1 パラメータの推定

2.3節で述べた7つのパラメータをLevenberg-Marquardt法を用いて推定する．その評価関数は

$$\sum_i^n \left(\|u_c^i - \hat{u}_c^i\|^2 + \|v_c^i - \hat{v}_c^i\|^2 \right) \rightarrow \min \quad (8)$$

とする．ここで，対応点の位置を $x_c^i = (u_c^i, v_c^i, 1)$ ，物体表面上の i 番目の点 $X^i = (X^i, Y^i, Z^i)$ からパラメータよりカメラ画像に再投影した i 番目の対応点の位置を $\hat{x}_c^i = (\hat{u}_c^i, \hat{v}_c^i, 1)$ ，対応点の総数を n とする．この評価関数の意味するところは，Fig.2のように x_c^i と \hat{x}_c^i との距離の二乗和が最小になるようにパラメータを決定するということである．ここで，1.節で述べたように X^i も推定する各点であるので， Z^i もパラメータとする (Z^i が求めれば式(1), (3)より X^i, Y^i も求めることができる)．

2.3.2 パラメータの初期値

2.3.1節の方法は非線形の数値的最小化であるため，パラメータの初期値が必要である．これらは以下のようにして求める．

まず，プロジェクタとカメラの焦点距離 f_p, f_c の値を決める．次に，基礎行列 f を計算し，その焦点距離を用いて基礎行列から外部パラメータの初期値を決める．具体的には，基礎行列はプロジェクタとカメラの内部パラメータ行列 K_p, K_c から

$$F = (K_p^{-1})^T E K_c^{-1} \quad (9)$$

という関係がある．ここで， E は基本行列 (essential matrix) と呼ばれる 3×3 の行列であり， $t = (t_1, t_2, t_3)^T$

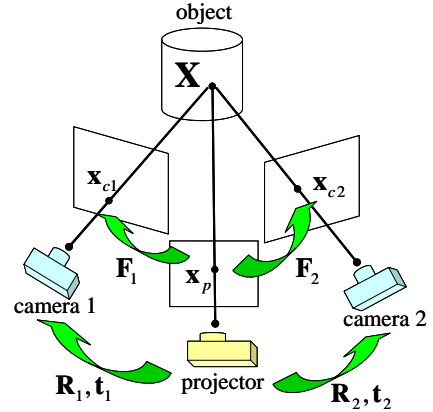


Fig. 3 1プロジェクタ,2カメラシステム

に対して

$$T = \begin{pmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{pmatrix} \quad (10)$$

となる行列を用いて，

$$E = TR \quad (11)$$

という関係がある．この関係から E を分解することにより外部パラメータ R, t を決定することができる³⁾，これを初期値とする．

最後に，上記したように決定したパラメータの初期値を用いて，プロジェクタとカメラのステレオによって各点の奥行き Z^i の初期値を決定する．

2.4 1プロジェクタ,2カメラシステム

1プロジェクタ,2カメラシステムは，カメラの位置を変え2視点から物体を計測する方法である．移動前のカメラをカメラ1，移動後のカメラをカメラ2とすると，プロジェクタとカメラ1，プロジェクタとカメラ2との間でそれぞれ基礎行列を求めることができる．よって，基礎行列の自由度を14とし，推定することのできるパラメータの数を14個にすることができる．

2.4.1 1プロジェクタ,2カメラシステムのパラメータ

1プロジェクタ,2カメラ - システムにおいて基礎行列から推定するパラメータは，まずプロジェク

タとカメラ1との間の外部パラメータ R_1, t_1 ,そしてプロジェクタとカメラ2との間の外部パラメータ R_2, t_2 が必要である.また,プロジェクタ,カメラ1,カメラ2のそれぞれの焦点距離 f_p, f_{c1}, f_{c2} は使用するたびに調整し変化させることが多いため,推定するパラメータとする.2カメラの場合,データが増えたことで推定できる内部パラメータの数を増やすことができる.そこで,プロジェクタの垂直方向の画像中心 v_{0p} はプロジェクタによる個体差がかなりあり,プロジェクタ自体の機能によって v_{0p} が変化する場合があるため, v_{0p} をパラメータとする.

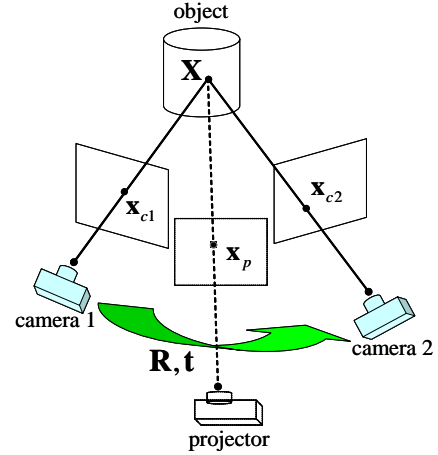


Fig. 4 2カメラシステム

2.4.2 パラメータの推定

パラメータの推定は,2.3.1節と同様に行うが,最適化を行うための評価関数は

$$\sum_i^{n_1} \left(\|u_{c1}^i - \hat{u}_{c1}^i\|^2 + \|v_{c1}^i - \hat{v}_{c1}^i\|^2 \right) + \sum_j^{n_2} \left(\|u_{c2}^j - \hat{u}_{c2}^j\|^2 + \|v_{c2}^j - \hat{v}_{c2}^j\|^2 \right) \rightarrow \min \quad (12)$$

とする.ここで, $(u_{c1}^i, v_{c1}^i), (\hat{u}_{c1}^i, \hat{v}_{c1}^i)$ はそれぞれカメラ1の画像の*i*番目の対応点と,物体表面上の*i*番目の点 X^i とパラメータからカメラ1の画像に再投影した対応点の座標であり, $(u_{c2}^j, v_{c2}^j), (\hat{u}_{c2}^j, \hat{v}_{c2}^j)$ はカメラ2の画像についてのものである.また, n_1, n_2 はそれぞれカメラ1と2の対応点の総数である.

また,パラメータの初期値についても2.3.2節と同様に行うが,物体表面上の点 X がカメラ1,カメラ2の両方で撮られているとき, X の成分 Z はプロジェクタとカメラ1とのステレオによって求めた Z^i と,プロジェクタとカメラ2とのステレオによって求めた Z^j との平均とする.

2.5 2カメラシステム

2カメラを使う場合,カメラのみのステレオによって三次元形状の計測を行うことができる.その場合には対応点の抽出が大きな問題になるが,プロジェクタの存在はその対応点抽出の助けとなっているだけと見ることもできる.しかし,同時に

プロジェクタの存在が精度の向上をもたらすことも予想できる.そこで,比較のために,以下の実験では2カメラのみのステレオによる形状の計測を行った.

2カメラシステムにおける推定するパラメータは,カメラ1とカメラ2の焦点距離 f_{c1}, f_{c2} と外部パラメータ R, t より7個となる.評価関数は

$$\sum_i^n \left(\|u_{c2}^i - \hat{u}_{c2}^i\|^2 + \|v_{c2}^i - \hat{v}_{c2}^i\|^2 \right) \rightarrow \min \quad (13)$$

とする.ここで, (u_{c2}^i, v_{c2}^i) は物体表面上の*i*番目の位置にある点 X^i から2.3節と同様にして求めたカメラ2の画像へ再投影した対応点の位置である.

各パラメータの初期値についても,2.3.2節と同様にして決める.

3. シミュレーション実験

3.1 実験方法

まず,プロジェクタとカメラの内部パラメータを設定し,プロジェクタとカメラ1,2のそれぞれについての外部パラメータを設定する.三次元空間中の仮想的な立方体の辺上の44点の対応点を,プロジェクタ,カメラ1,カメラ2についてそれぞれ式(1)より求める.その際,カメラ1とカメラ2の対応点には標準偏差が0.1~0.5のノイズを加える.プロジェクタの対応点には,投影するパターンを自身で作成するためノイズを加えない.その対応

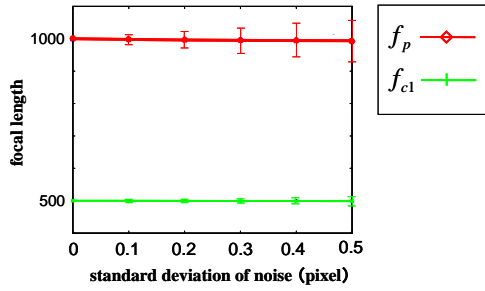


Fig. 5 1プロジェクタ,1カメラ (カメラ1) システムの焦点距離の平均と標準偏差 (真値: $f_p = 1000, f_{c1} = 500$)

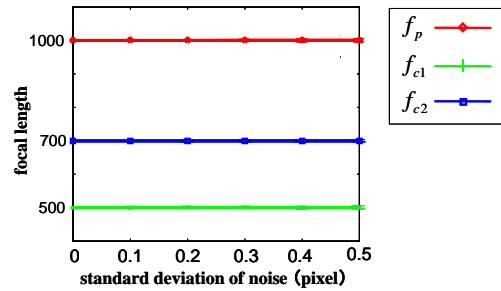


Fig. 7 1プロジェクタ,2カメラシステムの焦点距離の平均と標準偏差 (真値: $f_p = 1000, f_{c1} = 500, f_{c2} = 700$)

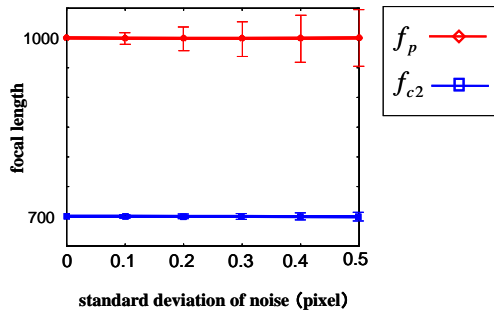


Fig. 6 1プロジェクタ,1カメラ (カメラ2) システムの焦点距離の平均と標準偏差 (真値: $f_p = 1000, f_{c2} = 700$)

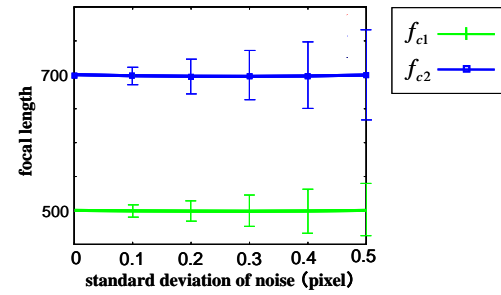


Fig. 8 2カメラシステムの焦点距離の平均と標準偏差 (真値: $f_{c1} = 500, f_{c2} = 700$)

点から, 1プロジェクタ,1カメラシステム, 1プロジェクタ,2カメラ - システム, そして2カメラ - システムのキャリブレーションを行い, 立方体の形状を計測する. これを各標準偏差のノイズに対して100回行い, 各手法を比較する.

3.2 実験結果

Fig.5に1プロジェクタ,1カメラ (カメラ1) システム, Fig.6に1プロジェクタ,1カメラ (カメラ2) システム, Fig.7に1プロジェクタ,2カメラシステム, Fig.8に2カメラ - システムのそれぞれの各標準偏差のノイズにおける焦点距離の平均と標準偏差のグラフを示す. 折れ線グラフは各標準偏差における焦点距離の平均を表し, 各標準偏差のノイズにおける縦線が焦点距離の標準偏差を表している.

Fig.5~8をみると, プロジェクタ, カメラ1, カメラ2の焦点距離の平均は各標準偏差のノイズに対して真値とほぼ等しくなっている. 1プロジェクタ, 1カメラシステムと2カメラ - システムにおいてノイ

ズの標準偏差が大きくなるにつれて, 焦点距離の標準偏差もはっきりと大きくなっている. つまり, ノイズの標準偏差を大きくしていくと, 計測毎の推定されたパラメータのばらつきが大きくなっているということが言える. しかし, 1プロジェクタ, 2カメラ - システムでは, ノイズの標準偏差を変化させてもあまり焦点距離の標準偏差が変化せず, 計測毎のばらつきがあまり見られなかった. つまり, 1プロジェクタ, 2カメラシステムにおいてプロジェクタとカメラの焦点距離ともに安定して推定できていると言える.

Fig.9に1プロジェクタ, 2カメラシステムにおける垂直方向の画像中心 v_{0p} の平均と標準偏差を示す. Fig.9をみると, v_{0p} の平均はノイズの標準偏差を大きくしていくと, 多少変化していくもののあまり大きな変化ではないのが分かる. また, v_{0p} の標準偏差もノイズの標準偏差を大きくしていくと, 多少変化していくもののあまり大きな変化ではない.

Fig.10は各手法による立方体の計測誤差を示している. ここで, 計測誤差とは, 立方体の1辺の長さを100としたときの各边上点の真の位置との距

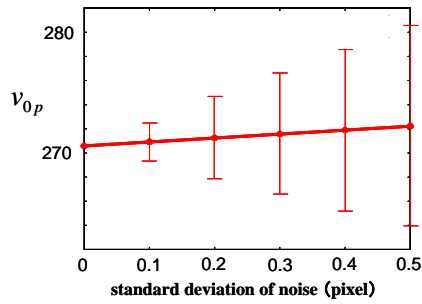


Fig. 9 1プロジェクタ,2カメラ - システムの垂直方向の画像中心の平均と標準偏差 (真値: $v_{0p} = 270$)

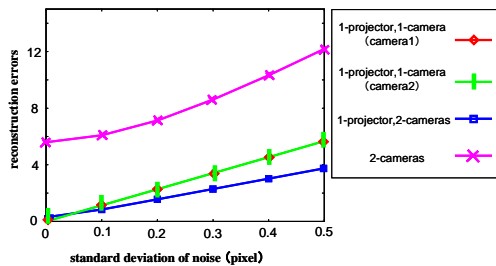


Fig. 10 計測誤差

離の平均である。Fig.10をみると、各手法ともにノイズの標準偏差が大きくなると計測誤差が大きくなっている。そして、各手法間で比較すると1プロジェクタ,2カメラシステムによる計測が最も精度が良く、逆に2カメラ - システムによる計測が最も計測精度が悪いということが言える。

4. 実画像による実験

Fig.11に計測の様子を示す。グレイコード⁴⁾による空間コード化法を用いてプロジェクタ画像とカメラ画像の対応点をとった。そして、その対応点からプロジェクタとカメラのキャリブレーションをし、顔の石膏像の形状を計測した Fig.12に計測した顔の石膏像を示す。

Fig.12より、石膏像の形状を計測することができるのが分かる。

5. まとめ

カメラを移動させることによって視点を増やした1プロジェクタ,2カメラシステムによるキャリブレーション手法を示した。



Fig. 11 計測の様子

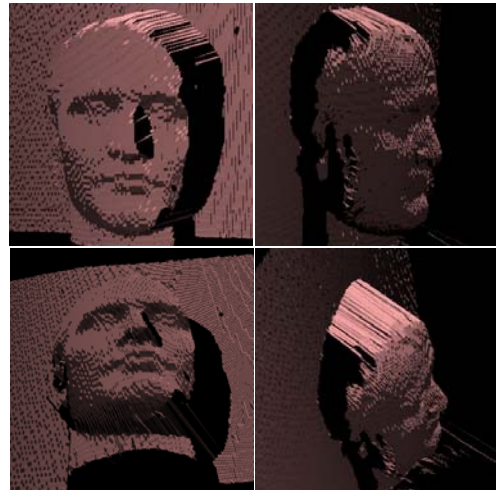


Fig. 12 顔の石膏像の計測結果

1プロジェクタ,2カメラシステムによってプロジェクタの垂直方向の画像中心 v_{0p} をパラメータとして含めてキャリブレーションを行うことができた。さらにカメラを移動させる、もしくはカメラの台数を増やすことなどにより、推定することのできるパラメータの数を多くすることもできるであろう。また、視点を増やすことによって計測領域が広がるだけでなく、計測の精度が向上することを確認することができた。

参考文献

- 1) 出口 光一郎: ロボットビジョンの基礎, コロナ社 (2000)
- 2) 出口 光一郎: 多視点からの透視投影画像の因子分解による3次元形状復元, 計測自動制御学会論文集, 34-10, 1321/1328 (1998)
- 3) 金谷 健一, 松永 力: 基礎行列の分解: 焦点距離の直接表現, 情報処理学会研究報告, 120-7, 49/56 (2000)
- 4) 高辻誠也, 佐川立昌, 越後富夫, 八木康史: グレイコードパターンを利用したレンズ歪みの補正手法, 画像の認識・理解シンポジウム, OS6A-21, (2005)