

論文 / 著書情報
Article / Book Information

題目(和文)	6軸モーションセンサ装着移動単眼カメラによる3次元距離計測
Title(English)	
著者(和文)	赤松俊弘
Author(English)	Toshihiro Akamatsu
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第9866号, 授与年月日:2015年3月26日, 学位の種別:課程博士, 審査員:廣田 薫,長橋 宏,佐藤 誠,柴田 崇徳,室伏 俊明
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9866号, Conferred date:2015/3/26, Degree Type:Course doctor, Examiner:,,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

平成 26 年度

博士（工学）学位論文

6 軸モーションセンサ装着
移動単眼カメラによる
3 次元距離計測

赤松 俊弘

指導教員 廣田 薫 教授

東京工業大学 大学院総合理工学研究科 知能システム科学専攻

要旨

6 軸モーションセンサ装着移動単眼カメラによる 3 次元距離計測

赤松俊弘

6 軸モーションセンサ装着移動単眼カメラによる静止/移動物体の 3 次元距離計測法を提案し，静止/移動物体混在動画のフレーム間対応点に関する静止/移動識別実験を行う．既存の両眼立体視法と比べ，計測精度を保ったまま計算コストをほぼ半減でき，静止/移動対応点を安定して正確に抽出識別できる．単一静止/移動物体の 3 次元距離計測実験では，両眼立体視法と比べ計算コストのほぼ半減を，静止/移動物体が混在する CG 画像を用いた対応点静止/移動識別実験では正答率 0.97，適合率 0.96，再現率 1.00 を，確認している．本論文は，車載カメラや移動ロボットに導入した障害物回避や自動運転システムへの応用の基盤を確立している．

キーワード 3 次元距離計測，6 軸モーションセンサ，移動単眼カメラ，
対応点分類

目次

第 1 章	序論	1
第 2 章	カメラモデルと基礎行列を用いたカメラ校正法	3
2.1	カメラモデル	3
2.2	透視投影	5
2.3	座標変換	8
2.4	正射影カメラ	11
2.5	物理的なカメラモデル	12
2.6	カメラキャリブレーション	16
2.7	ステレオカメラのキャリブレーション法	20
2.8	本章のまとめ	29
第 3 章	6 軸モーショセンサ装着移動単眼カメラによる静止物体の 3 次元 距離計測法	30
3.1	基礎行列を用いた 3 次元距離計測手法および先行研究	30
3.2	6 軸モーショセンサ装着移動単眼カメラによる 基礎行列推定法	31
3.3	6 軸モーショセンサによるカメラ移動・回転角度推定実験.	38
3.4	6 軸モーショセンサ装着移動単眼カメラによる 静止物体 3 次元距離計測実験.	44
3.5	本章のまとめ	51

第 4 章	6 軸モーションセンサ装着移動単眼カメラによる移動物体の 3 次元距離計測法	53
4.1	6 軸モーションセンサ装着移動単眼カメラによる 移動物体 3 次元距離計測.	54
4.2	6 軸モーションセンサ装着移動単眼カメラによる 移動物体 3 次元距離計測実験	56
4.3	本章のまとめ	60
第 5 章	6 軸モーションセンサ装着移動単眼カメラによる 3 次元距離計測のための対応点分類手法	61
5.1	静止物体と移動物体が混在するシーンに おける 3 次元距離計測	61
5.2	6 軸モーションセンサ装着移動単眼カメラによる 対応点分類手法.	63
5.3	対応点分類手法検証実験.	65
5.4	本章のまとめ	82
第 6 章	結論	83
	参考文献	87
	業績リスト	93

第 1 章

序論

3次元計測とは対象の形状や観測者から対象までの距離などを非接触で計測することであり，多視点から得られた複数の画像を用いて計測を行う手法が一般的には良く知られている．これらの手法は古くから研究されているが，物体の形状認識や車両の障害物回避や自動運転，ロボットビジョンなどへの応用が考えられ，今現在でも多くの研究課題を多く残している[1-12]．

3次元計測の手法として用いられるものには，複数の視点から対象を捉えた画像を用いてその対象の特徴点を抽出し，画像上に投影された特徴点の視差から3次元計測を行うものが数多く提案されている[13-27]．特徴点の視差から3次元計測を行うためには，カメラの焦点距離や各視点間のカメラの姿勢などの情報が必要である．しかしながら，それらのカメラパラメータはあらかじめ知ることが困難である場合が多い．エピポラ幾何学を用いて特徴点の対応関係を記述し，カメラパラメータを推定する手法が提案されている[3, 5, 7, 8]．特徴点の対応関係を表した方程式を基礎方程式と呼び，基礎方程式から得られる基礎行列を用いてカメラパラメータを推定することをカメラのセルフキャリブレーションと呼ぶ．

カメラのセルフキャリブレーションを行うには基礎方程式を解いて基礎行列を求める必要があるが，3次元計測の精度は基礎行列の精度に依存する．基礎行列を高精度で計算するためには，対応点を精度良く求めることが重要である．対象の特徴点を検出し，その対応付けを行う手法は数多く研究されている

[29-33]が、特徴点の抽出や対応付けに誤差が生じる可能性は避けられない。また、基礎行列の計算には複数の対応点が必要であるため、誤差のある対応点からより正確なものを選別することは難しい。そのため、基礎行列の計算精度を上げる手法が研究されているが[34-41]、いずれの場合でも計算コストの増大や最悪の場合 3 次元計測自体が困難になってしまう。

本論文では基礎行列を求めるための計算コスト増大や安定性といった問題を解決するために、6 軸モーションセンサ装着移動単眼カメラを用いた、移動および静止物体の 3 次元距離計測を行う手法および 3 次元距離計測のための対応点の分類手法を提案する。この手法は 6 軸モーションセンサによりカメラの動作や姿勢を推定することにより、基礎行列の推定に必要となる対応点の数を抑えることによる計算コスト減少、および対応点を分類し誤対応を除去することによる、安定した 3 次元計測を実現する。

本論文は全 6 章で構成される。第 2 章では 3 次元計測に必要なカメラのモデル化およびカメラのキャリブレーション法について述べる。第 3 章では 6 軸モーションセンサにより移動単眼カメラの動きを推測することにより、静止物体の 3 次元計測に必要な基礎行列の計算コストを削減する手法を提案する。第 4 章では移動単眼カメラでの移動物体に対する 3 次元計測を実現する手法を提案する。第 5 章では 3 次元計測のための対応点分類手法を提案し、静止・移動物体が混在するシーンでの安定した 3 次元計測を実現する。

第 2 章

カメラモデルと基礎行列を用いたカメラ校正法

本章では本論文で用いるカメラモデルとキャリブレーション手法，基礎行列を用いた 3 次元計測手法について述べる．

2.1 カメラモデル

カメラで撮影された画像内の一点と 3 次元空間中の光線の関係は投影・射影関係によって決定される．カメラモデルとはこの投影・射影関係を記述するものである．カメラモデルは正射影モデル，透視投影モデル，ピンホールカメラモデルなど複数のモデルがあるが，本論文ではこのうちピンホールカメラモデルを用いている．ピンホールカメラのモデルを図 2.1 に示す．ピンホールカメラとはスクリーンに小さな穴を開けた非常に単純なカメラである．3 次元空間内の物体に反射した光線はこの穴を通り，撮像素子が置かれる面で結像する．撮像素子が置かれる面は画像面，光線が通過する穴は光学中心，画像面と光学中心との間の距離は焦点距離とそれぞれ呼ばれる．ピンホールカメラの特徴としては，ピントを合わせる必要がない，投影の幾何学的な性質が保存されている、といったものが挙げられる．また，ピンホールカメラでは画像面には倒立した像が結像される．倒立していない像を得る方法としては，スクリーンに対して対称の位置に仮の画像面を配置する方法がある．扱い易さの観点から，画

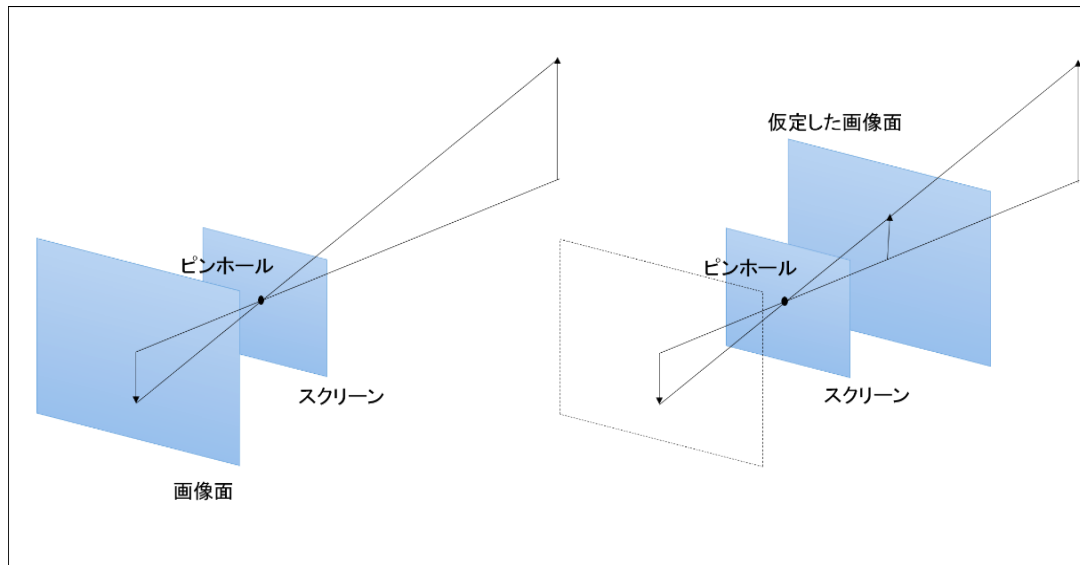


図 2.1：ピンホールカメラ

像面がこのようにスクリーンの前にあると仮定したモデルが一般的であり，本論文で扱うピンホールカメラモデルも同様のものである．

2.2 透視投影

ピンホールカメラモデルについてさらに詳細に記述する．図 2.1 に示したモデルにおいて，スクリーン上にある光学中心を C とし，3次元空間中に C を原点とする座標系 (C, X, Y, Z) を考える．座標軸は画像面上方に X 軸，右方に Y 軸を取るものとする．また，光学中心を通り画像面に直交する方向に Z 軸を取る．この Z 軸をカメラの光軸と呼ぶ．画像面と光軸の交点を光軸点 c とする．光学中心と画像中心の間の距離を焦点距離 f とする．本論文で用いるカメラ座標系は以上のように定義される．

光軸点 c を原点とする画像面上の座標系 (c, u, v) を定義すると，カメラ座標系における3次元空間内の点 $[X \ Y \ Z]^T$ と画像面上の座標系における2次元平面上の点 $[u \ v]^T$ との関係は

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} f \frac{X}{Z} \\ f \frac{Y}{Z} \end{bmatrix} \quad (2.1)$$

となる．このような変換を透視投影と呼ぶ．この関係を，斉次座標を用いて書き直すと

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.2)$$

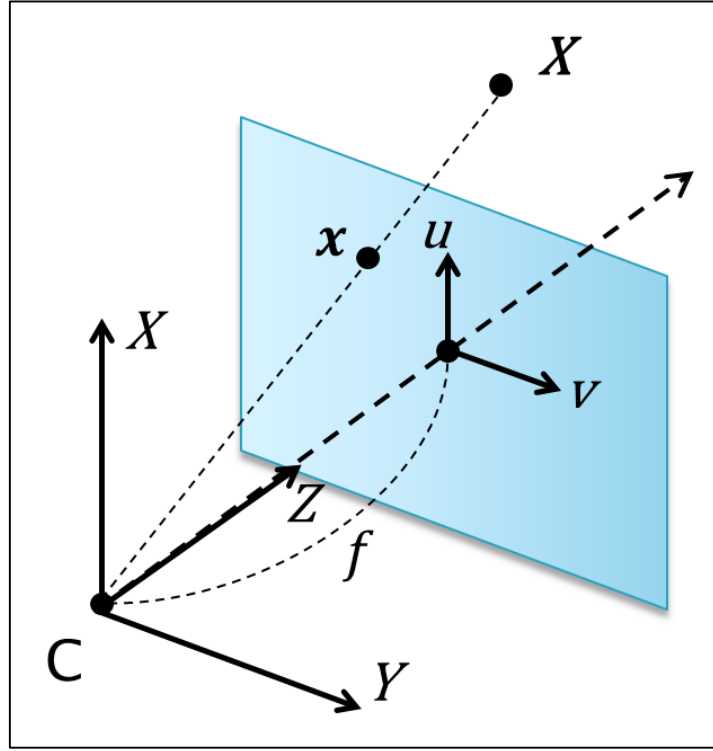


図 2.2：透視投影図

となる．ここで， $\mathbf{x} = [u \ v \ 1]^T$ ， $\mathbf{X} = [X \ Y \ Z \ 1]^T$ ， f を含む行列を $\mathbf{P}$ とするとこの関係は

$$\mathbf{x} = \mathbf{P}\mathbf{X} \quad (2.3)$$

と記述される．これはカメラの投影を射影空間 $\mathbb{P}^3$ から $\mathbb{P}^2$ への線型写像として表現している．この $\mathbf{P}$ を透視投影行列と呼ぶ．ここで， $\mathbf{P}$ を

$$\mathbf{P} = \begin{bmatrix} -\mathbf{P}_1^T & - \\ -\mathbf{P}_2^T & - \\ -\mathbf{P}_3^T & - \end{bmatrix} \quad (2.4)$$

とすると， $\mathbf{P}_n^T$ ， $(n = 1, 2, 3)$ は射影空間中のある平面を表すことになる．すなわち， $\mathbf{P}_1^T$ は $Y-Z$ 平面上の点を表すベクトルであり， $\mathbf{P}_2^T$ は $X-Z$ 平面上の点を表すベクトルである．これらの平面の交差点はカメラの光学中心である．したがって，カメラの光学中心に関して，

$$P \begin{bmatrix} c \\ 1 \end{bmatrix} = \mathbf{0} \quad (2.5)$$

の式が成り立つ.

2.3 座標変換

2.3.1 デジタルデータにおける座標変換

画像をコンピュータ上で扱う場合、カメラから得られた画像はデジタルデータとして画素の集合で表現される。このとき画素は格子状に並んでいるが、この格子のサイズは撮影するカメラのハードウェアやデータのサンプリングレートにより変化する。したがって、デジタルデータ上では画素のサイズに対応した新たな座標系を定義する必要がある。いま、新たな座標系において画素のスケールを $[\kappa_u \ \kappa_v]$ ，光軸点の座標を $[u_0 \ v_0]^T$ とすると、対応する座標変換は

$$\begin{bmatrix} \kappa_u & 0 & u_0 \\ 0 & \kappa_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.6)$$

のように記述できる。

したがって、新たな透視投影行列は

$$\mathbf{P} = \begin{bmatrix} f\kappa_u & 0 & u_0 & 0 \\ 0 & f\kappa_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.7)$$

のようになる。この行列はカメラに固有のパラメータである焦点距離 f や画素の大きさ $[\kappa_u \ \kappa_v]$ ，光軸点 $[u_0 \ v_0]^T$ を含んでいる。したがって、透視投影をより一般的に表現するためにカメラのパラメータの分離を行う必要がある。カメラパラメータを分離すると、透視投影行列は

$$\mathbf{P} = \begin{bmatrix} f\kappa_u & 0 & u_0 \\ 0 & f\kappa_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (2.8)$$

のように表現できる。このカメラは単位カメラと呼ばれるものである。

2.3.2 世界座標を用いた座標変換

3次元空間内の点を表現する場合、これまで説明してきたカメラ座標系に加え、カメラ座標系とは別の独立した座標系を導入したほうが説明がしやすいことがある。このような座標系を世界座標系と呼ぶ。

3次元空間内のある点がそれぞれ、世界座標系で3次元座標ベクトル $[X_w Y_w Z_w]^T$ ，カメラ座標系で3次元座標ベクトル $[X Y Z]^T$ を持つと仮定する。このとき、世界座標系における斉次座標ベクトルを $\mathbf{X}_w$ ，カメラ座標系における斉次座標ベクトルを $\mathbf{X}$ とし、カメラ座標系への座標変換を行う回転行列を $\mathbf{R}$ ，並進ベクトルを $\mathbf{t}$ とすると、互いの座標間の関係は

$$\mathbf{X} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \mathbf{X}_w \quad (2.9)$$

のように記述される。

次に世界座標系と画像座標系とに成り立つ関係を記述する。カメラ座標と画像座標の関係を表す式(2.3)に、単位カメラによる透視投影を表す式(2.8)および、カメラ座標と世界座標の関係を表す式(2.9)を代入することにより、世界座標とカメラ座標の関係を表す式、

$$\mathbf{x} = \begin{bmatrix} f\kappa_u & 0 & u_0 \\ 0 & f\kappa_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_3^T & 1 \end{bmatrix} \mathbf{X}_w \quad (2.10)$$

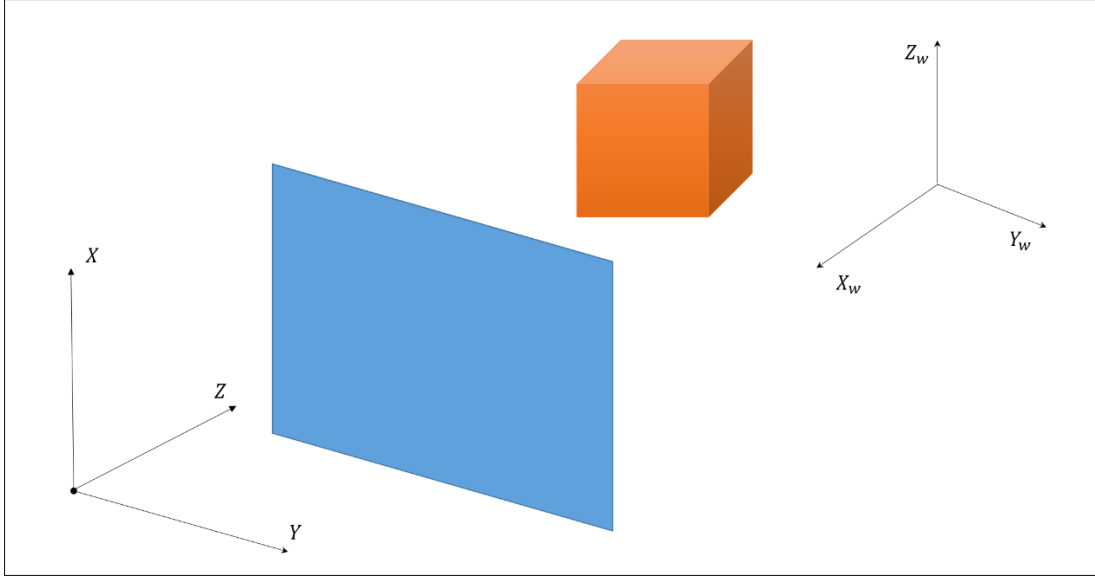


図 2.3：世界座標

が得られる．この式をわかりやすく整理すると

$$\mathbf{x} = \begin{bmatrix} f\kappa_u & 0 & u_0 \\ 0 & f\kappa_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} [\mathbf{R} \quad \mathbf{t}] \mathbf{X}_w \quad (2.11)$$

$$= \mathbf{K} [\mathbf{R} \quad \mathbf{t}] \mathbf{X}_w \quad (2.12)$$

となる．ここで， $\mathbf{K}$ はカメラの内部パラメータ行列と呼ばれ，画面上の座標変換を表す行列であり、カメラに固有のパラメータである．また， $\mathbf{R}$ および $\mathbf{t}$ をカメラの外部パラメータと呼びカメラの姿勢を表している．

2.4 正射影カメラ

次に正射影カメラについて記述する．正射影は透視投影において，カメラから物体間の距離と比較して焦点距離 f が非常に大きくなったときの状態の投影である．式(2.1)のカメラ座標系において， Z 軸と定義されている光軸に沿って f だけ並進した場合を考える．このとき，式(2.11)から透視投影行列は

$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & f \end{bmatrix} \quad (2.13)$$

ように表現される．透視投影行列は定義にスケールを含んでいるため，これは

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix}$$

としても等しくなる．ここで， $f \rightarrow \infty$ とした場合，透視投影行列は

$$\mathbf{P}_{orth} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.14)$$

のように表現される．このとき，画像座標は $u = X$ ， $v = Y$ となる．これが正射影と呼ばれる透視投影である．なお， $\mathbf{P}_{orth}$ の第三行から明らかなように，正射影は無限遠平面である $\boldsymbol{\pi}_{\infty} = [0 \ 0 \ 0 \ 1]^T$ が保存されている．これは正射影が平行性を保存する性質があることを示している．

2.5 物理的なカメラモデル

これまで説明してきたカメラモデルは、理想的な環境を仮定した上で成り立つものである。現実のカメラを考慮すると、それぞれのモデルにはカメラのハードウェア構成上のずれや使用するレンズの収差などによる誤差が存在する。本論文で提案する3次元距離計測法で用いているカメラにおいてもこうした誤差は存在するが、その影響はごく少ないと考えられるため理想的なモデルとしている。ここでは現実的な影響を考慮したカメラモデルについて参考として記述する。

2.5.1 剪断変形を含めたカメラモデル

これまで説明してきたカメラモデルは、画像面と光軸は直行しているという事実を前提として成り立つモデルである。しかしながら、現実のカメラを考えた場合、画像面と光軸が正確に直交していない場合も考えられる。この場合、画像格子が歪んでしまうため、格子が正確な正方形であるという前提が崩れる。この変形は剪断変形と呼ばれ、画像座標系が直交座標系ではなくなってしまうということである。このような剪断変形を考慮すると、内部パラメータ行列は

$$K = \begin{bmatrix} f\kappa_u & f\kappa_u \cos \theta & u_0 \\ 0 & \frac{f\kappa_v}{\sin \theta} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.15)$$

のように表現される。ここで θ は画像座標系の座標軸のなす角である。この内部パラメータ行列は θ に関して非線形であることは明らかである。すなわち、剪断変形を考慮した場合のカメラモデルは非線形になる。

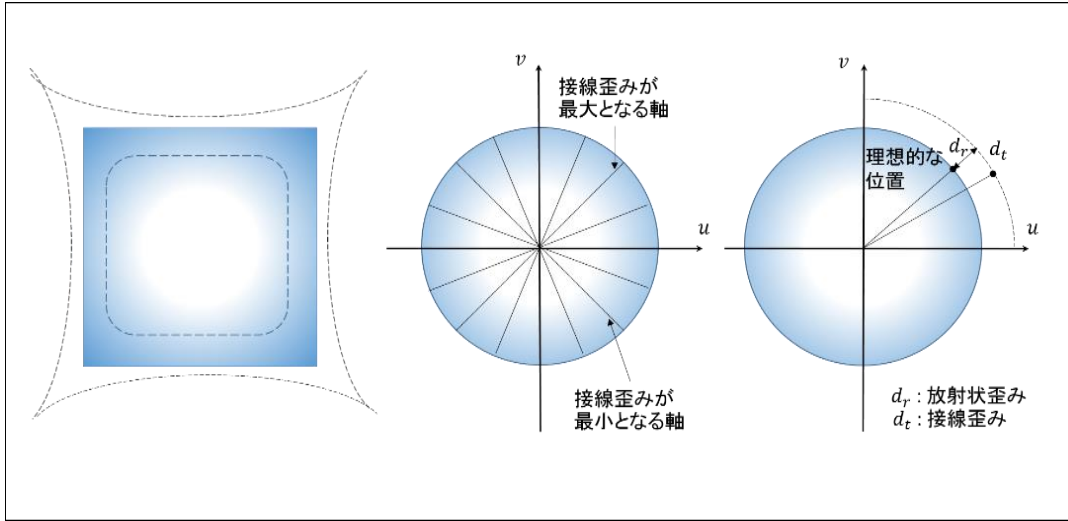


図 2.4：レンズの歪み

2.5.2 レンズ収差を含めたカメラモデル

実際のカメラでは剪断変形に加えて、レンズ収差による歪みも考慮する必要がある。ここでは、そうした歪みの中でカメラモデルに最も大きな影響を与えるとされる、Weng らの提唱したモデルによる放射状歪み、および接線歪み [42] について説明する。

理想的なレンズに対して放射状歪みのあるレンズは像の位置にずれが生じる。このうち、内側へのずれである負の変位は樽型歪み、外側へのずれである正の変位は糸巻き型歪みと呼ばれる。放射状歪みを δ_r とすると、その歪みは

$$\delta_r = k_1 r^3 + k_2 r^5 + k_3 r^7 + \dots \quad (2.16)$$

で表わされる。ここで、 r は光軸点からの距離、 $k_1, k_2, k_3, \dots$ は歪み係数を表す。さらに、この式(2.16)の第一項に着目する。 u 方向と v 方向に対するそれぞれの位置のずれを δ_{ur} および δ_{vr} とすると、

$$\begin{aligned} \delta_{ur} &= k_1 u(u^2 + v^2) \\ \delta_{vr} &= k_1 v(u^2 + v^2) \end{aligned} \quad (2.17)$$

のように表すことができる．式(2.17)における u および v は理想的なレンズモデルにおける像の位置である．放射状歪みは，光軸点を中心として $\pi/4$ 傾いたときに最も大きくなる．

つぎに，接線歪みについて説明する．接線歪みは主に偏心歪みと薄プリズム歪みにより引き起こされる．これらは接線歪みだけではなく，放射状歪みを引き起こす原因にもなる．偏心歪みは，レンズの光軸と光学系の中心がずれている場合など，光学系の誤差が原因となる．偏心歪みは放射状の歪み成分を δ_{rd} ，接線の歪み成分を δ_{td} とするとそれぞれ

$$\begin{aligned}\delta_{rd} &= 3(j_1 r^2 + j_2 r^4 + \dots) \sin(\phi - \phi_0) \\ \delta_{td} &= (j_1 r^2 + j_2 r^4 + \dots) \cos(\phi - \phi_0)\end{aligned}\tag{2.18}$$

のように表される．ここで， ϕ_0 は図 2.4 の接線歪みが最も大きくなる方向における角度である．ここで $\cos \phi = u/r$ ， $\sin \phi = v/r$ ， $p_1 = j_1 \sin \phi_0$ ， $p_2 = j_1 \cos \phi_0$ とし， r に関する 4 次以上の項を省略することにより， u 方向と v 方向の歪み成分 δ_{ud} ， δ_{vd} はそれぞれ

$$\begin{aligned}\delta_{ud} &= p_1(3u^2 + v^2) + 2p_2 uv \\ \delta_{vd} &= 2p_1 uv + p_1(u^2 + 3v^2)\end{aligned}\tag{2.19}$$

のように表される．

薄プリズム歪みはレンズ加工の精度や光学系の配置のずれから生じ，

$$\begin{aligned}\delta_{rp} &= (i_1 r^2 + i_2 r^4 + \dots) \sin(\phi - \phi_1) \\ \delta_{tp} &= (i_1 r^2 + i_2 r^4 + \dots) \cos(\phi - \phi_1)\end{aligned}\tag{2.20}$$

のように表される．ここで， ϕ_1 は図 2.4 の接線歪みが最も大きくなる方向における角度である． $s_1 = i_1 \sin \phi_0$ ， $s_2 = i_1 \cos \phi_0$ とし，偏心歪みの場合と同様に r に関する 4 次以上の項を省略することにより， u 方向と v 方向の歪み成分 δ_{up} ， δ_{vp} はそれぞれ

$$\begin{aligned}\delta_{up} &= s_1(u^2 + v^2) \\ \delta_{vp} &= s_2(u^2 + v^2)\end{aligned}\tag{2.21}$$

のように表される．

これらのレンズ収差を考慮した場合，像が投影される位置のずれは偏心歪みおよび薄プリズム歪みを組み合わせた歪みにより生じる．したがって，それらを合わせた最終的な u 方向と v 方向の歪み成分は

$$\begin{aligned}\delta_u &= s_1(u^2 + v^2) + p_1(3u^2 + v^2) + 2p_2uv + k_1u(u^2 + v^2) \\ \delta_v &= s_2(u^2 + v^2) + p_1(u^2 + 3v^2) + 2p_2uv + k_1v(u^2 + v^2)\end{aligned}\tag{2.22}$$

のように表される．この中で，放射状歪みを表す k_1 を係数とする項が歪みに最も大きく影響する．したがって，一般的にレンズによる歪みを考慮する場合には，この要素のみを適用することが多い[16]．また，レンズ収差による歪みを考慮したカメラモデルも剪断変形を考慮した場合と同様に非線形である．

2.6 カメラキャリブレーション

カメラの持つ透視投影行列を推定し、カメラの内部パラメータおよび外部パラメータを決定することをカメラキャリブレーションという。ここではカメラキャリブレーションの手法として、Zhang[13]の提案した3次元空間中の点の座標を用いたキャリブレーション法および、Triggs[14]の提案した3次元平面上の点対応を用いるキャリブレーション法について紹介する。

2.6.1 既知の物体構造を用いたカメラキャリブレーション法

カメラキャリブレーション法の一つは、既知である3次元構造物をもつ対象を用いる手法である。これは、カメラにより観測された対象の画像座標と、あらかじめ計測した対象の3次元座標の関係を式として表し、それを解くことによってカメラキャリブレーションを行う手法である。しかしながら、3次元物体をキャリブレーションのための対象として用いる場合には、その構造を正確に計測しておく必要がある。したがって、カメラキャリブレーションを容易にするためには比較的簡易な対象を用いる必要がある。そこでZhangはカメラキャリブレーションを行うための対象として、図2.5に示すようなチェスボード様の平面キャリブレーションパターンを用いる手法を提案している。以下にZhangのカメラキャリブレーション法を説明する。

まず、平面キャリブレーションパターン上に $X-Y$ 座標系を設定する。次に、その平面上の任意の点 $\mathbf{X}_p$ 、($p = 1, \dots, n$)を m 箇所の視点から観測する。このとき、 $Z_p = 0$ である。したがって、 i 番目の視点の内部パラメータ行列を $\mathbf{K}_i$ 、外部パラメータである回転行列を $\mathbf{R}_i$ 、並進ベクトルを $\mathbf{t}_i$ 、($i = 1, \dots, m$)

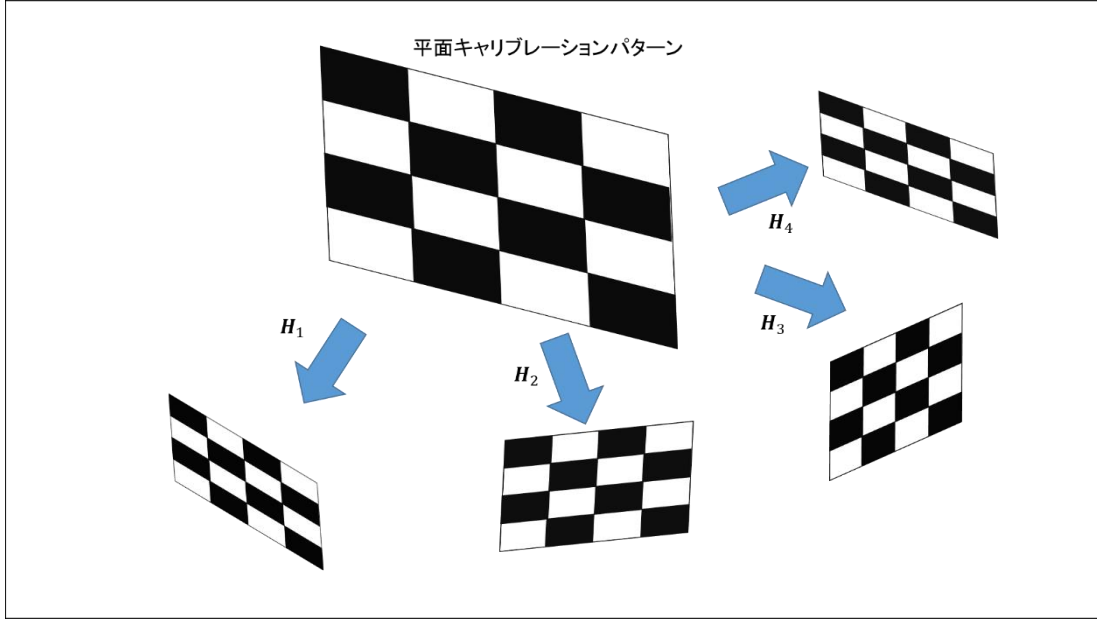


図 2.5 : Zhang のカメラキャリブレーション法

と置くと，平面上の座標 $[u_{ip} \ v_{ip}]^T$ とその 3 次元空間中の座標である

$[X_p \ Y_p \ Z_p]^T$ との関係は斉次座標を用いて

$$\begin{aligned} \begin{bmatrix} u_{ip} \\ v_{ip} \\ 1 \end{bmatrix} &= K_i [R_i \mid t_i] \begin{bmatrix} X_p \\ Y_p \\ 0 \\ 1 \end{bmatrix} \\ &= [h_{i,1} \ h_{i,2} \ h_{i,3}] \begin{bmatrix} X_p \\ Y_p \\ 1 \end{bmatrix} \end{aligned} \quad (2.23)$$

のように表すことができる．ただし， $Z_p = 0$ より回転行列の第 3 列を省略している．ここで，式(2.23)における第 2 式の射影変換を H_i とすると， 3×3 の射影変換である $K_i^{-1}H_i$ の 3 列はそれぞれ回転行列 R_i の第 1 列および第 2 列と並進ベクトル t_i の定数倍に等しくなる．したがって， $\omega_i \cong K_i^{-1}K_i^{-T}$ とすると

$$h_{i,1}^T \omega_i h_{i,2} = 0 \quad (2.24)$$

$$h_{i,1}^T \omega_i h_{i,1} - h_{i,2}^T \omega_i h_{i,2} = 0 \quad (2.25)$$

が成り立つ．ここで， $\mathbf{K}_i$ の各要素が各視点間で変わらないと仮定すると， $\boldsymbol{\omega}_i$ も不変であるので， $\boldsymbol{\omega}_i = \boldsymbol{\omega}$ となる．したがって，この $\boldsymbol{\omega}$ における 6 つの未知数は 3 つの異なる視点における $\mathbf{H}_i$ を用いて推定することが可能である．その後，推定した $\boldsymbol{\omega}$ を用いて $\boldsymbol{\omega} = \mathbf{K}_i^{-1} \mathbf{K}_i^{-T}$ と分解することにより，内部パラメータ行列 $\mathbf{K}_i$ を計算することができる．

2.6.2 特徴点の平面拘束条件を利用したカメラキャリブレーション法

本節では Triggs のカメラキャリブレーション法について説明する．Zhang の手法は既知の平面キャリブレーションパターンにおける特徴点の画像面上の座標と 3 次元座標の関係から平面射影変換 $\mathbf{H}_i$ を求め，カメラキャリブレーションを行う手法であるが，Triggs の手法では平面キャリブレーションパターンを用いて異なる視点間における特徴点の画像面上の点対応から平面射影変換を推定し，カメラキャリブレーションを行う手法である．このように，3 次元空間中の既知の点を全く利用しないカメラキャリブレーションをセルフキャリブレーションと呼ぶ．

いま， m 箇所の異なる視点における n 点の平面キャリブレーションパターン上の特徴点 $\mathbf{X}_p$ ，($p = 1, \dots, n$) に対し，画像 i と画像 j における $\mathbf{X}_p$ の像をそれぞれ $\mathbf{x}_{ip}$ ， $\mathbf{x}_{jp}$ とすると， $\mathbf{x}_{ip}$ と $\mathbf{x}_{jp}$ の間の関係は平面射影変換 $\mathbf{H}_{ij}$ を用いて

$$\mathbf{x}_{ip} = \mathbf{H}_{ij} \mathbf{x}_{jp} \quad (2.26)$$

のように表現できる．式(2.26)は $\mathbf{H}_{ij}$ に関して線型であるから，特徴点 $\mathbf{X}_p$ の 3 次元座標を用いずに特徴点の点対応のみから平面射影変換 $\mathbf{H}_{ij}$ を推定することが可能である．ここで， $\mathbf{H}_{ij}$ は

$$\mathbf{H}_{ij} = \mathbf{H}_i \mathbf{H}_j^{-1} \quad (2.27)$$

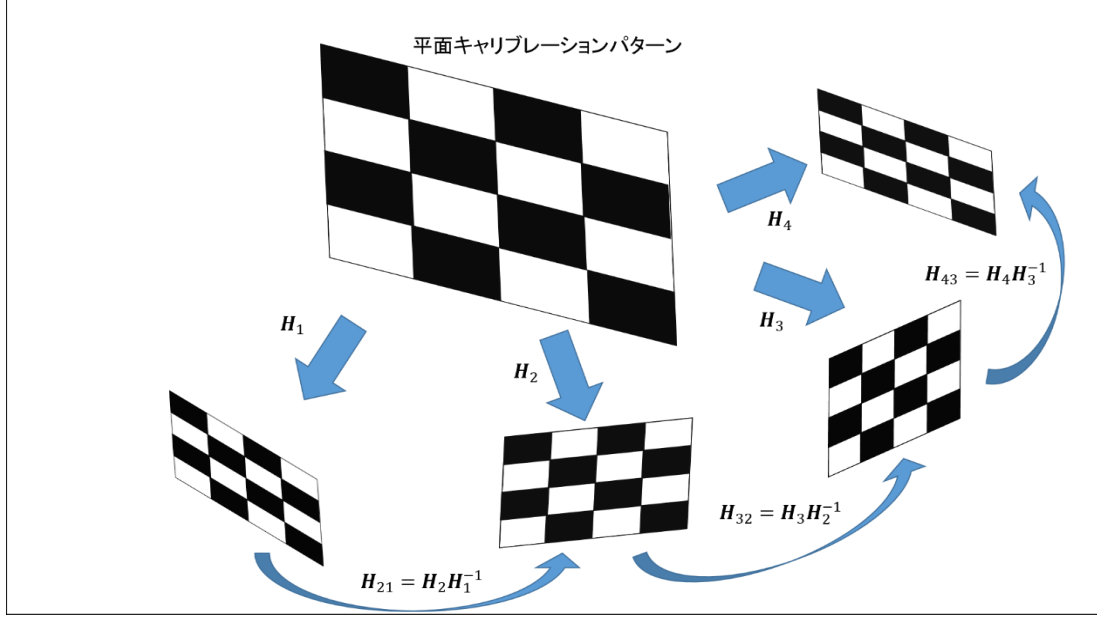


図 2.6 : Triggs のカメラ校正法

のように $\mathbf{H}_i$ および $\mathbf{H}_j$ に分解できる(図 2.6). すなわち, $\mathbf{H}_i$ および $\mathbf{H}_j$ はそれぞれ平面キャリブレーションパターンと画像 i および画像 j 間の平面射影変換を表す. ここで, 式(2.27)の両辺の定数倍を λ_{ij} とすると

$$\begin{bmatrix} \lambda_{11}\mathbf{H}_{11} & \lambda_{12}\mathbf{H}_{12} & \cdots & \lambda_{1m}\mathbf{H}_{1m} \\ \lambda_{21}\mathbf{H}_{21} & \lambda_{22}\mathbf{H}_{22} & \cdots & \lambda_{2m}\mathbf{H}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m1}\mathbf{H}_{m1} & \lambda_{m2}\mathbf{H}_{m2} & \cdots & \lambda_{mm}\mathbf{H}_{mm} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_m \end{bmatrix} [\mathbf{H}_1^{-1} \ \mathbf{H}_2^{-1} \ \cdots \ \mathbf{H}_m^{-1}] \quad (2.28)$$

となる. ここで, λ_{ij} は

$$\lambda_{ij} = \frac{\text{Trace}\left((\mathbf{H}_{i1}\mathbf{H}_{1j})\mathbf{H}_{ij}^T\right)}{\text{Trace}(\mathbf{H}_{ij}\mathbf{H}_{ij}^T)} \quad (2.29)$$

となる. $\text{Trace}(\cdot)$ は行列の対角和を表す. したがって, 各画像間の特徴点对応を用いて平面射影変換 $\mathbf{H}_{ij}$ を推定し, 式(2.28)を解くことで $\mathbf{H}_i$ を計算することができる. カメラキャリブレーションは, 各視点における平面射影変換 $\mathbf{H}_i$, ($i = 1, \dots, m$)を計算することにより, それらを用いて Zhang の方法と同様に行うことができる.

2.7 ステレオカメラのキャリブレーション法

ここではエピポラ幾何学に基づくステレオカメラのキャリブレーション法について説明する．まず，基礎行列についてエピポラ幾何学を用いて説明し，ステレオカメラを用いた画像間の対応点と基礎行列との関係について記述する．次に，ステレオカメラの基礎行列を推定することにより，ステレオカメラの内部パラメータおよび外部パラメータを計算可能であること，すなわちカメラキャリブレーションが可能であることを説明する．

2.7.1 ステレオカメラにおける対応点とエピポラ幾何学

2 視点間のエピポラ幾何学とは，ステレオカメラにより 3 次元空間中の任意の点がそれぞれの画像面上に写像されるとき，それらの点の対応関係を表すものである．ある 3 次元空間中の特徴点 $\mathbf{X}_p = [X_p \ Y_p \ Z_p \ 1]^T$ がそれぞれ，射影行列 $\mathbf{P}_i, \ \mathbf{P}_j$ を介して画像 i および画像 j の画像面上にそれぞれ $\mathbf{x}_{ip} = [u_i \ v_i \ 1]^T$ および $\mathbf{x}_{jp} = [u_j \ v_j \ 1]^T$ と写像されるものとする．画像 i および画像 j に対するカメラの光学中心をそれぞれ $\mathbf{c}_i$ および $\mathbf{c}_j$ と置くと，図 2.7 に示すように $\mathbf{x}_{ip}, \ \mathbf{x}_{jp}, \ \mathbf{c}_i, \ \mathbf{c}_j$ は特徴点 $\mathbf{X}_p$ と同一平面上にある．したがって， $\mathbf{x}_{ip}$ と $\mathbf{c}_i$ を結ぶ直線と $\mathbf{x}_{jp}$ と $\mathbf{c}_j$ を結ぶ直線は $\mathbf{X}_p$ において 1 点で交わることが分かる．次に，2 画像間の対応点である $\mathbf{x}_{ip}$ と $\mathbf{x}_{jp}$ の間に成り立つ関係を考える．対応点である $\mathbf{x}_{ip}$ および $\mathbf{x}_{jp}$ と 3 次元空間中の特徴点 $\mathbf{X}_p$ の間の関係はそれぞれ

$$\begin{aligned}\mathbf{x}_{ip} &= \mathbf{P}_i \mathbf{X}_p \\ \mathbf{x}_{jp} &= \mathbf{P}_j \mathbf{X}_p\end{aligned}$$

のように記述することができる． $\mathbf{x}_{ip}$ は点 $\mathbf{X}_p$ を投影した点であるので，この 2 点は同一直線上にある．この 2 点の関係は，直線上の 2 点を用いて

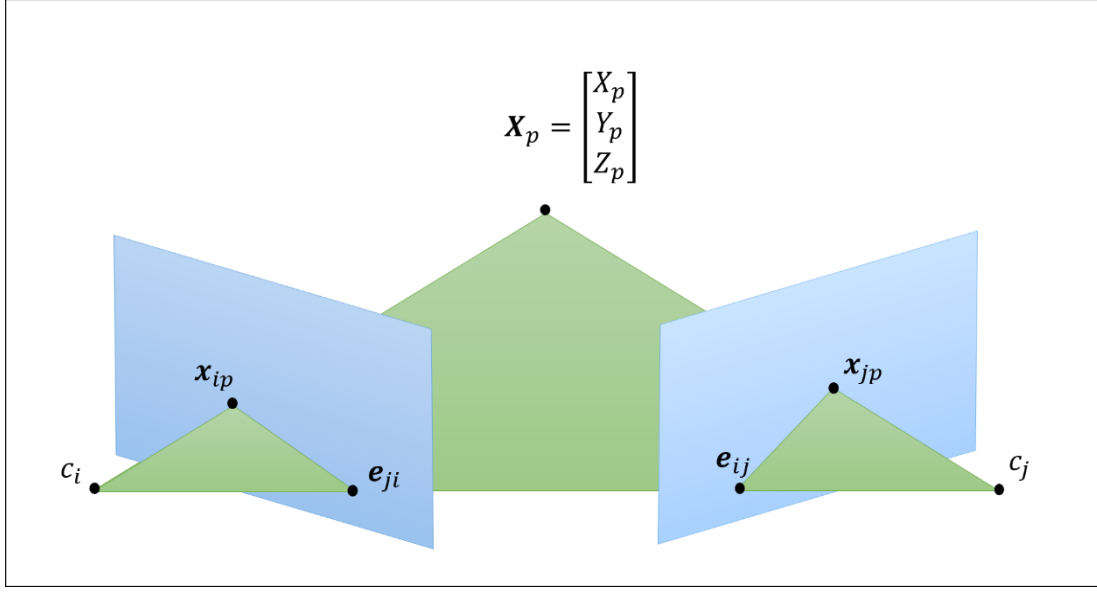


図 2.7：ステレオカメラと点対応

$$\mathbf{X}_{p,\lambda} = \mathbf{P}_i^+ \mathbf{x}_{ip} + \lambda \begin{bmatrix} c_i \\ 1 \end{bmatrix} \quad (2.30)$$

のように表すことができる．ここで， $\mathbf{P}_i^+$ は $\mathbf{P}_i$ の擬似逆行列を表す． $\mathbf{P}_i \mathbf{P}_i^+ = \mathbf{I}_{3 \times 3}$ および式(2.5)を用いると， $\mathbf{P}_i \mathbf{X}_{p,\lambda} = \mathbf{x}_{ip}$ となることが分かる．ここで， $\mathbf{P}_j$ を用いて $\mathbf{X}_{p,\lambda}$ を投影することを考える．その投影像である $\mathbf{l}_{jp}$ は式(2.30)で選んだ2点を結ぶ直線となるため

$$\begin{aligned} \mathbf{l}_{jp} &= \left(\mathbf{P}_j \begin{bmatrix} c_i \\ 1 \end{bmatrix} \right) \times \left(\mathbf{P}_j \mathbf{P}_i^+ \mathbf{x}_{ip} \right) \\ &= [\mathbf{e}_{ij}]_{\times} \mathbf{P}_j \mathbf{P}_i^+ \mathbf{x}_{ip} \end{aligned} \quad (2.31)$$

のように表される．ここで， $[\mathbf{e}_{ij}]_{\times}$ は $\mathbf{e}_{ij}$ の歪み対称行列であり， $\mathbf{e}_{ij}$ との外積を表す．なお， $\mathbf{e}_{ij}$ はエピポールと呼ばれ，画像 i におけるカメラの光学中心 c_i を画像 j の画像面上に投影した像である． $\mathbf{x}_{ip}$ および $\mathbf{x}_{jp}$ は特徴点 $\mathbf{X}_p$ それぞ

れの画像面に対する投影であるので、 $\mathbf{l}_{jp}$ と $\mathbf{x}_{ip}$ は共線となる。したがって、

$\mathbf{l}_{jp}$ と $\mathbf{x}_{ip}$ の間には

$$\begin{aligned} \mathbf{x}_{jp}^T \mathbf{l}_{jp} &= 0 \\ \Leftrightarrow \mathbf{x}_{jp}^T [\mathbf{e}_{ij}]_{\times} \mathbf{P}_j \mathbf{P}_i^+ \mathbf{x}_{ip} &= 0 \\ \Leftrightarrow \mathbf{x}_{jp}^T \mathbf{F}_{ji} \mathbf{x}_{ip} &= 0 \end{aligned} \quad (2.32)$$

の関係が成り立つ。ここで、 $\mathbf{F}_{ji}$ は $\mathbf{x}_{ip}$ および $\mathbf{x}_{jp}$ 間の線形変換であり、

$$\mathbf{F}_{ji} = \mathbf{x}_{jp}^T [\mathbf{e}_{ij}]_{\times} \mathbf{P}_j \mathbf{P}_i^+ \quad (2.33)$$

である。この線形変換行列 $\mathbf{F}_{ji}$ を基礎行列と呼び、直線 $\mathbf{l}_{jp}$ を $\mathbf{x}_{ip}$ に関するエピポーララインと呼ぶ。式(2.32)はエピポーラ方程式と呼ばれ、対応点間の拘束条件を表している。また、ここでは $\mathbf{x}_{ip}$ と $\mathbf{P}_j$ に関する投影関係を記述することにより $\mathbf{F}_{ji}$ を計算したが、は $\mathbf{x}_{jp}$ と $\mathbf{P}_i$ に関する投影関係を記述することによっても同様の拘束条件である $\mathbf{x}_{ip}^T \mathbf{F}_{ij} \mathbf{x}_{jp} = 0$ を計算可能である。2つの基礎行列 $\mathbf{F}_{ij}$ と $\mathbf{F}_{ji}$ の間の関係を記述すると

$$\mathbf{F}_{ij} = \mathbf{F}_{ji}^T \quad (2.34)$$

となる。また、 $\mathbf{e}_{ij}^T \mathbf{F}_{ji} = \mathbf{0}^T$ の定義を用いることにより、基礎行列 $\mathbf{F}_{ij}$ 、 $\mathbf{F}_{ji}$ とエピポール $\mathbf{e}_{ij}$ 、 $\mathbf{e}_{ji}$ の間に成立する関係を記述すると

$$\mathbf{e}_{ij}^T \mathbf{F}_{ji} = \mathbf{0}^T \quad (2.35)$$

$$\mathbf{F}_{ji} \mathbf{e}_{ij} = \mathbf{0}^T \quad (2.36)$$

$$\mathbf{e}_{ij}^T \mathbf{F}_{ij} = \mathbf{0}^T \quad (2.37)$$

$$\mathbf{F}_{ji} \mathbf{e}_{ji} = \mathbf{0}^T \quad (2.38)$$

となる。ここから、基礎行列には定数倍の不定性があり、正則でないことが分かる。したがって、その自由度は7である。

2.7.2 基礎行列と射影復元

3次元空間中の特徴点 $\mathbf{X}_p$ をそれぞれの画像面に投影した像は $\mathbf{x}_{ip}$, $\mathbf{x}_{jp}$ であり, それぞれの射影行列を $\mathbf{P}_i$, $\mathbf{P}_j$ とすると, これらの間には

$$\begin{bmatrix} \mathbf{P}_i & \mathbf{x}_{ip} & \mathbf{0} \\ \mathbf{P}_j & \mathbf{0} & \mathbf{x}_{jp} \end{bmatrix} \begin{bmatrix} \mathbf{X}_p \\ -\lambda_{ip} \\ -\lambda_{jp} \end{bmatrix} = \mathbf{0} \quad (2.39)$$

の関係が成り立つ. ここで, λ_{ip} , λ_{jp} は零ではないスケールを表す. この式(2.39)に対する行列式を展開すると

$$\begin{aligned} 0 = & u_{ip}u_{jp}|\mathbf{p}_{i,2}\mathbf{p}_{i,3}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| + v_{ip}u_{jp}|\mathbf{p}_{i,3}\mathbf{p}_{i,1}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| \\ & + u_{ip}v_{jp}|\mathbf{p}_{i,2}\mathbf{p}_{i,3}\mathbf{p}_{j,3}\mathbf{p}_{j,1}| + v_{ip}v_{jp}|\mathbf{p}_{i,3}\mathbf{p}_{i,1}\mathbf{p}_{j,3}\mathbf{p}_{j,1}| \\ & + u_{jp}|\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| + v_{jp}|\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{j,3}\mathbf{p}_{j,1}| \\ & + u_{ip}|\mathbf{p}_{i,2}\mathbf{p}_{i,3}\mathbf{p}_{j,1}\mathbf{p}_{j,2}| + v_{ip}|\mathbf{p}_{i,3}\mathbf{p}_{i,1}\mathbf{p}_{j,1}\mathbf{p}_{j,2}| + |\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{j,1}\mathbf{p}_{j,2}| \end{aligned} \quad (2.40)$$

が得られる. ただし, $\mathbf{p}_{i,k}$, $\mathbf{p}_{j,k}$, ($k = 1, 2, 3$) はそれぞれ $\mathbf{P}_i$ と $\mathbf{P}_j$ の第 k 行の縦ベクトルを表す. 式(2.32)と式(2.40)は互いに同等性を持つことを利用すると, 基礎行列 $\mathbf{F}_{ji}$ は射影行列 $\mathbf{P}_i$ と $\mathbf{P}_j$ を用いて

$$\mathbf{F}_{ji} = \begin{bmatrix} |\mathbf{p}_{i,2}\mathbf{p}_{i,3}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| & |\mathbf{p}_{i,3}\mathbf{p}_{i,1}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| & |\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| \\ |\mathbf{p}_{i,2}\mathbf{p}_{i,3}\mathbf{p}_{j,3}\mathbf{p}_{j,1}| & |\mathbf{p}_{i,3}\mathbf{p}_{i,1}\mathbf{p}_{j,3}\mathbf{p}_{j,1}| & |\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{j,3}\mathbf{p}_{j,1}| \\ |\mathbf{p}_{i,2}\mathbf{p}_{i,3}\mathbf{p}_{j,1}\mathbf{p}_{j,2}| & |\mathbf{p}_{i,3}\mathbf{p}_{i,1}\mathbf{p}_{j,1}\mathbf{p}_{j,2}| & |\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{j,1}\mathbf{p}_{j,2}| \end{bmatrix} \quad (2.41)$$

と表現される. また, 式(2.49)と式(2.5)を用いることにより, 射影行列 $\mathbf{P}_i$ と $\mathbf{P}_j$ により画像 i と画像 j の光学中心を表現することができる. すなわち, これらの光学中心をそれぞれ $\mathbf{P}_i$, $\mathbf{P}_j$ を用いて画像面に投影したものがエピポール $\mathbf{e}_{ij}$, $\mathbf{e}_{ji}$ であるので, その関係を記述すると

$$\mathbf{e}_{ij} = \begin{bmatrix} |\mathbf{p}_{i,1}\mathbf{p}_{j,1}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| \\ |\mathbf{p}_{i,2}\mathbf{p}_{j,1}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| \\ |\mathbf{p}_{i,3}\mathbf{p}_{j,1}\mathbf{p}_{j,2}\mathbf{p}_{j,3}| \end{bmatrix} \quad (2.42)$$

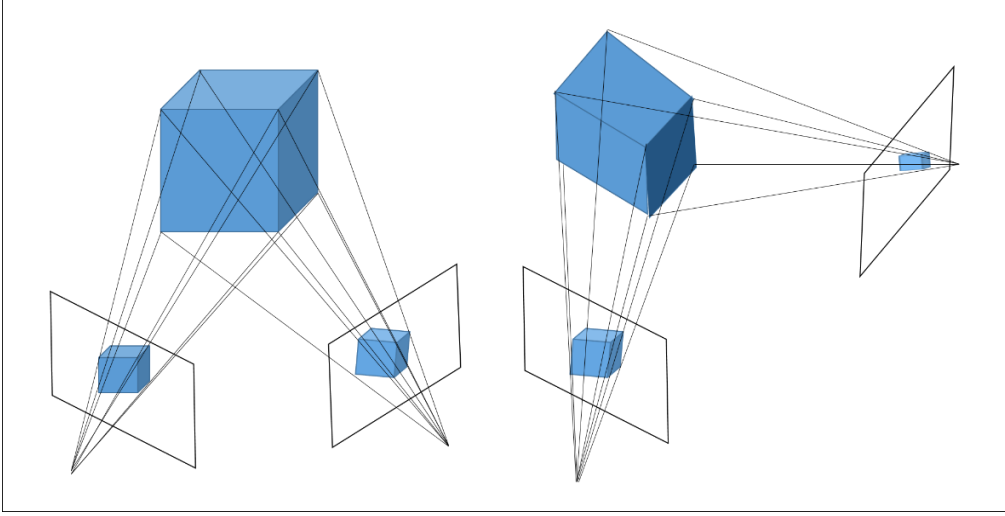


図 2.8：射影復元

$$\mathbf{e}_{ij} = \begin{bmatrix} \mathbf{p}_{j,1}\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{i,3} \\ \mathbf{p}_{j,2}\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{i,3} \\ \mathbf{p}_{j,3}\mathbf{p}_{i,1}\mathbf{p}_{i,2}\mathbf{p}_{i,3} \end{bmatrix} \quad (2.43)$$

のように表現される．これらの式(2.41)，式(2.42)，式(2.43)はそれぞれ，基礎行列 $\mathbf{F}_{ji}$ ，エピポール $\mathbf{e}_{ij}$ ， $\mathbf{e}_{ji}$ に対してグラスマン–ケイリー代数[43]を適用した表現と同じものである．

4×4 の射影変換行列 $\mathbf{H}$ により，射影行列 $\mathbf{P}_i$ ， $\mathbf{P}_j$ および特徴点 $\mathbf{X}_p$ を変換したものをそれぞれ $\mathbf{P}'_i = \mathbf{P}_i\mathbf{H}$ ， $\mathbf{P}'_j = \mathbf{P}_j\mathbf{H}$ ， $\mathbf{X}'_p = \mathbf{H}^{-1}\mathbf{X}_p$ とする．このとき，それぞれを投影した像である $\mathbf{x}_{ip}$ ， $\mathbf{x}_{jp}$ は不変である．これは図 2.8 に示すように，カメラ位置を変換するのに伴いその対象を変形することと同等であるといえる．このとき，式(2.41)を用いると， $\mathbf{P}'_i$ ， $\mathbf{P}'_j$ における基礎行列 $\mathbf{F}'_{ji}$ は $\mathbf{F}_{ji}$ ， $\mathbf{H}$ および $\mathbf{F}_{ij}$ を用いて

$$\mathbf{F}'_{ji} = \mathbf{F}_{ji}|\mathbf{H}^T| = \mathbf{F}_{ji} \quad (2.44)$$

のように記述される．射影変換 $\mathbf{H}$ は基礎行列に対して影響を与えないので，式(2.44)から 2 つの射影行列対は射影座標系とは完全に独立したものであるということがいえる．したがって，基礎行列の推定により，ステレオカメラをキャ

リブレーションすることが可能であるといえる．このような3次元構造復元を射影復元と呼ぶ．なお，基礎行列が既に分かっているステレオカメラのことを弱キャリブレーション済みであるという．基礎行列 F_{ji} に対するステレオカメラの射影行列 P_i ， P_j はそれぞれ

$$P_i = [I_{3 \times 3} \mid \mathbf{0}] \quad (2.45)$$

$$P_j = \left[[e_{ij}]_{\times} F_{ji} + e_{ij} v^T \mid \lambda e_{ij} \right] \quad (2.46)$$

のように表される[30]．ここで， v は 3×1 のベクトルであり， λ は任意のスカラーを表す．

2.7.3 基本行列とセルフキャリブレーション法

射影行列 P_i ， P_j の基礎行列 F_{ji} は式(2.33)からも分かるように，それぞれのカメラの内部パラメータと外部パラメータから構成される．そこで，逆に2つのカメラの内部パラメータと外部パラメータを用いて基礎行列を記述することを考える．世界座標系を P_i の光学中心およびその向きを基準として設定することにより， P_i ， P_j は

$$P_i = K_i [I_{3 \times 3} \mid \mathbf{0}]$$

$$P_j = K_j [R_j \mid t_j]$$

と表すことができる．ここで， K_i ， K_j はそれぞれカメラの内部パラメータ行列， R_j および t_j はカメラ j の外部パラメータを表す．このようなカメラの対における基礎行列を記述すると

$$\begin{aligned} F_{ji} &= [e_{ij}]_{\times} K_j [R_j \mid t_j] [K_i [I_{3 \times 3} \mid \mathbf{0}]]^+ \\ &= [e_{ij}]_{\times} K_j [R_j \mid t_j] \begin{bmatrix} I_{3 \times 3} \\ \mathbf{0}^T \end{bmatrix} K_i^{-1} \end{aligned}$$

$$= [\mathbf{e}_{ij}]_{\times} \mathbf{K}_j \mathbf{R}_j \mathbf{K}_i^{-1} \quad (2.47)$$

のように表現される．さらに，

$$\mathbf{F}_{ji} = \mathbf{K}_j^{-1} [\mathbf{K}_j^{-1} \mathbf{e}_{ij}]_{\times} \mathbf{R}_j \mathbf{K}_i^{-1} \quad (2.48)$$

$$= \mathbf{K}_j^{-1} [\mathbf{t}_j]_{\times} \mathbf{R}_j \mathbf{K}_i^{-1} \quad (2.49)$$

となる．ただし，ここで任意の正則行列 $\mathbf{M}$ とベクトル $\mathbf{v}$ に対して， $[\mathbf{v}]_{\times} \mathbf{M} = \mathbf{M}^{-T} [\mathbf{M}^{-T} \mathbf{v}]_{\times}$ が成り立つことに加え， $\mathbf{e}_{ij} \cong \mathbf{K}_j \mathbf{t}_j$ であること用いている．これは $\mathbf{t}_j$ がカメラ j を基準としたときのカメラ i の光学中心であるためである．ここで，

$$\mathbf{E}_{ji} = [\mathbf{t}_j]_{\times} \mathbf{R}_j \quad (2.50)$$

$$= \mathbf{K}_j^T \mathbf{F}_{ji} \mathbf{K}_i \quad (2.51)$$

で表される行列 $\mathbf{E}_{ji}$ を基本行列と呼ぶ．基本行列の方程式(2.51)を基礎行列の方程式(2.32)と比較すると同一の構造を持つことがわかる，したがって，基本行列を用いることによっても点対応から基本行列を推定すること可能であるといえる．ただし，点対応を用いて計算された基本行列は定数倍の不定性を持つため，これによって復元された構造と元の構造との関係は相似になる．このような構造復元をユークリッド復元と呼ぶ．

次に，クルッパ方程式を導出する方法について説明する．クルッパ方程式とは，基礎行列を用いてカメラの内部パラメータに対して制約を課すための式である．このクルッパ方程式を解くことにより，座標が既知である点を用いることなくカメラキャリブレーションを行うことが可能である．式(2.47)から回転行列 $\mathbf{R}_j$ を消去すると

$$\begin{aligned} \mathbf{F}_{ji} \mathbf{K}_i &= [\mathbf{e}_{ij}]_{\times} \mathbf{K}_j \mathbf{R}_j \\ \Leftrightarrow \mathbf{F}_{ji} \mathbf{K}_i \mathbf{K}_i^T \mathbf{F}_{ji}^T &= [\mathbf{e}_{ij}]_{\times} \mathbf{K}_j \mathbf{R}_j \mathbf{R}_j^T \mathbf{K}_j^T [\mathbf{e}_{ij}]_{\times} \end{aligned}$$

$$\Leftrightarrow \mathbf{F}_{ji} \boldsymbol{\omega}_i^* \mathbf{F}_{ji}^T = [\mathbf{e}_{ij}]_{\times} \boldsymbol{\omega}_j^* [\mathbf{e}_{ij}]_{\times} \quad (2.52)$$

となる．ただし， $\boldsymbol{\omega}_i^* = \mathbf{K}_i \mathbf{K}_i^*$ ， $\boldsymbol{\omega}_j^* = \mathbf{K}_j \mathbf{K}_j^*$ である．この式(2.52)をクルッパ方程式と呼ぶ．カメラの内部パラメータが変化しないと仮定すると，単純なセルフキャリブレーションはこのクルッパ方程式を解くことで行うことができる．この仮定を用いると $\boldsymbol{\omega}_i^* = \boldsymbol{\omega}_j^* = \boldsymbol{\omega}^*$ となるため，クルッパ方程式は

$$\mathbf{F}_{ji} \boldsymbol{\omega}^* \mathbf{F}_{ji}^T = [\mathbf{e}_{ij}]_{\times} \boldsymbol{\omega}^* [\mathbf{e}_{ij}]_{\times} \quad (2.53)$$

と表すことができる．式(2.53)の両辺は $\boldsymbol{\omega}^*$ に関して線形である．したがって，クルッパ方程式より推定した $\boldsymbol{\omega}^*$ は $\boldsymbol{\omega}^* = \mathbf{K} \mathbf{K}^T$ と分解できることからカメラの内部パラメータ行列を計算することができる．

式(2.53)に関して，基礎行列の特異値分解を利用することにより簡易な表現で記述することができる．まず，基礎行列を

$$\mathbf{F}_{ji} = \mathbf{A} \boldsymbol{\Sigma} \mathbf{B}^T = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3] \begin{bmatrix} s_1 & & \\ & s_2 & \\ & & 0 \end{bmatrix} [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \mathbf{b}_3]^T \quad (2.54)$$

のように特異値分解する．ここで， $\mathbf{F}_{ji}^T \mathbf{e}_{ij} = 0$ であることから， $\mathbf{e}_{ij} \cong \mathbf{a}_3$ となる．この式(2.54)を式(2.53)に代入することで

$$\begin{aligned} \mathbf{A} \boldsymbol{\Sigma} \mathbf{B}^T \boldsymbol{\omega}^* \mathbf{B} \boldsymbol{\Sigma} \mathbf{A}^T &= [\mathbf{a}_3]_{\times} \boldsymbol{\omega}^* [\mathbf{a}_3]_{\times} \\ \Leftrightarrow \boldsymbol{\Sigma} \mathbf{B}^T \boldsymbol{\omega}^* \mathbf{B} \boldsymbol{\Sigma} &= \mathbf{A}^T [\mathbf{a}_3]_{\times} \boldsymbol{\omega}^* [\mathbf{a}_3]_{\times} \mathbf{A} \\ \Leftrightarrow \begin{bmatrix} s_1 & & \\ & s_2 & \\ & & 0 \end{bmatrix} \mathbf{B}^T \boldsymbol{\omega}^* \mathbf{B} \begin{bmatrix} s_1 & & \\ & s_2 & \\ & & 0 \end{bmatrix} &= [\mathbf{a}_2 \quad -\mathbf{a}_1 \quad \mathbf{0}]^T \boldsymbol{\omega}^* [\mathbf{a}_2 \quad -\mathbf{a}_1 \quad \mathbf{0}] \\ \Leftrightarrow \begin{bmatrix} s_1^2 \mathbf{b}_1^T \boldsymbol{\omega}^* \mathbf{b}_1 & s_1 s_2 \mathbf{b}_1^T \boldsymbol{\omega}^* \mathbf{b}_2 & 0 \\ s_1 s_2 \mathbf{b}_1^T \boldsymbol{\omega}^* \mathbf{b}_2 & s_2^2 \mathbf{b}_2^T \boldsymbol{\omega}^* \mathbf{b}_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} &= \begin{bmatrix} \mathbf{a}_2^T \boldsymbol{\omega}^* \mathbf{a}_2 & -\mathbf{a}_2^T \boldsymbol{\omega}^* \mathbf{a}_1 & 0 \\ -\mathbf{a}_1^T \boldsymbol{\omega}^* \mathbf{a}_2 & \mathbf{a}_1^T \boldsymbol{\omega}^* \mathbf{a}_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

となる．ここで，最後の式の両辺の要素を比較することでの方程式

$$\frac{\mathbf{a}_2^T \boldsymbol{\omega}^* \mathbf{a}_2}{s_1^2 \mathbf{b}_1^T \boldsymbol{\omega}^* \mathbf{b}_1} = -\frac{\mathbf{a}_1^T \boldsymbol{\omega}^* \mathbf{a}_1}{s_1 s_2 \mathbf{b}_1^T \boldsymbol{\omega}^* \mathbf{b}_2} = \frac{\mathbf{a}_1^T \boldsymbol{\omega}^* \mathbf{a}_1}{s_2^2 \mathbf{b}_2^T \boldsymbol{\omega}^* \mathbf{b}_2} \quad (2.55)$$

を得る[30]. この方程式のうち，独立なものは 2 つであるため，3 視点以上の画像の基礎行列を用いることにより $\boldsymbol{\omega}^*$ を推定することができる.

2.8 本章のまとめ

本章ではピンホールカメラを射影幾何学に基づいてモデル化し、さらにそのキャリブレーション法について記述した。まず、ピンホールカメラモデルに関して、斉次座標系でピンホールカメラの投影が空間中のユークリッド変換、射影変換、アフィン変換の合成で記述できることを示した。次に、カメラの内部パラメータおよび外部パラメータを求めるためのカメラキャリブレーション法の既存手法について説明した。最後に、ステレオカメラのキャリブレーション法がエピポーラ幾何学を用いて基礎行列を求めることにより可能であることを示した。

第 3 章

6 軸モーションセンサ装着移動単眼カメラ による静止物体の 3 次元距離計測法

3.1 基礎行列を用いた 3 次元距離計測法および先行研究

2 次元画像から 3 次元位置を計測する手法は対象までの距離推定や形状計測などに用いられ，ロボットビジョンや動画像解析の分野へも応用されている．3 次元形状計測の既存研究としては，カメラの動きに制限を加えたもの[1]やカメラに加え対象までの距離を測るセンサを用いる手法[2]，画像の対応点からの基礎行列推定を用いる手法[3-8] などが提案されている．既存研究[1] ではカメラに平行移動のみを許して距離の計測を行っている．また[2] では Time of Flight カメラという，対象までの距離が計測できるセンサを用いたカメラによる形状計測手法を提案している．しかし，これらの手法を一般的な撮影環境へ用いることは難しい．それに対し，対応点から基礎行列を推定する手法[3-8] では，カメラのみを用い，カメラの動きにも制限を課さない手法ではあるが，推定に多数の対応点が必要となり，計算コストもかかることに加え，実寸での計測にはカメラの移動距離もしくは対応点の空間的位置が既知である必要がある．

3.2 6軸モーションセンサ装着移動単眼カメラによる

基礎行列推定法

本研究では、6軸モーションセンサを取り付けたカメラを用いて任意の視点から撮影した2枚の画像に対し、センサの計測データからカメラの外部パラメータおよびカメラの焦点距離を推定し、対象の距離を計測する手法を提案する。また、推定した距離から3次元位置を再構成する。図3.1に提案手法のアウトラインを示す。提案手法では、カメラの外部パラメータ、すなわち並進ベクトルと回転行列をセンサの測定データから推定することで、カメラの動きに制限を課さない距離計測が可能である。また、理論上少なくとも1組の対応点からの距離計測が可能であり、計算コストを抑えることができる。さらに、センサによりカメラの移動距離が推定できるので実寸での計測が可能である。本研究で6軸モーションセンサを用いたのは、近年モバイル機器などに6軸モーションセンサが搭載されることが多くなり、一般的に利用できる機会の多いデバイスであると考えためである。提案手法を評価するために、まず6軸モーションセンサから得られるデータにより、外部パラメータが精度よく推定できることを示す。次に6軸モーションセンサを取り付けたウェブカメラから得られる実画像と外部パラメータを実際に入力に用いて距離計測を行い、計測さ

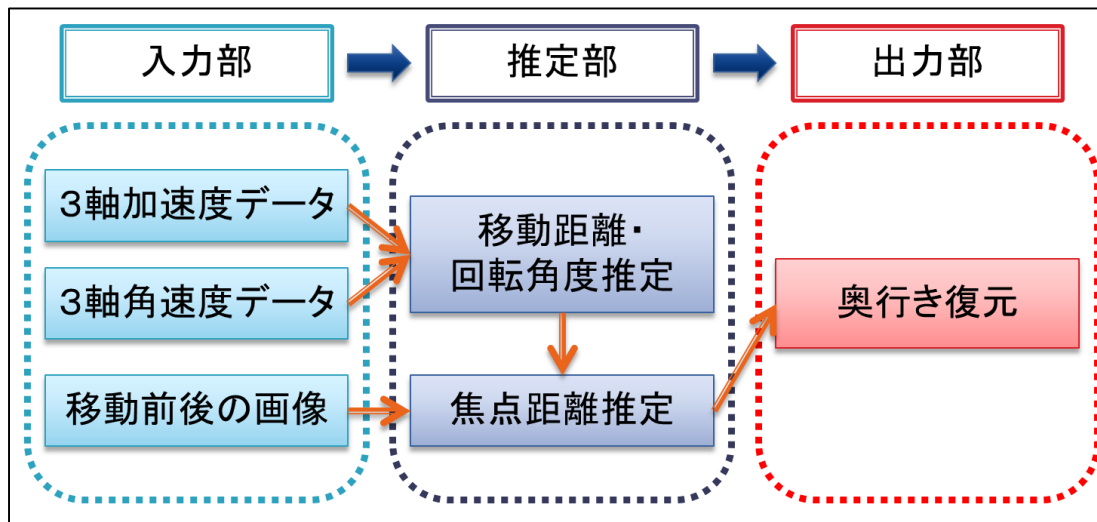


図 3.1：提案手法のアウトライン

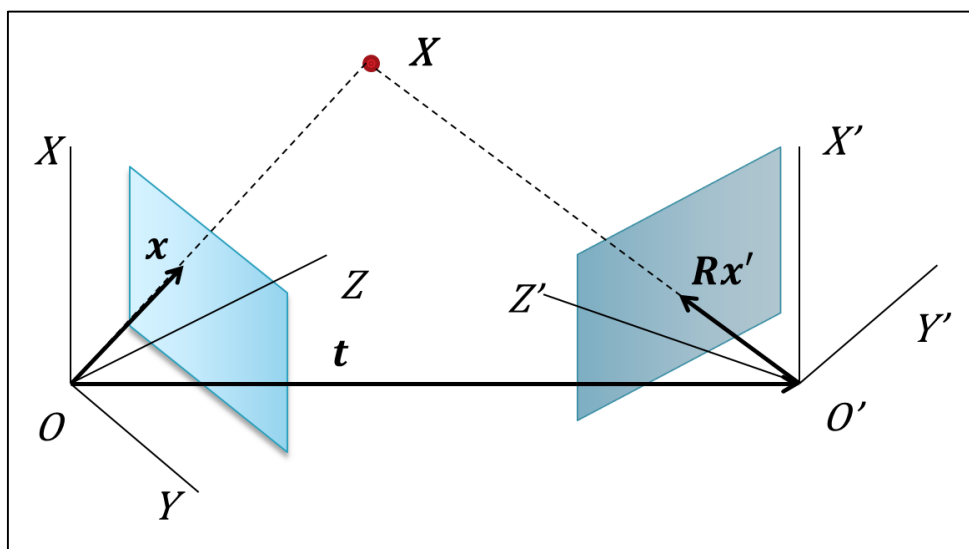


図 3.2：デバイスと特徴点の関係

れた3次元位置を比較することにより提案手法の有効性を確認する。また、計測結果に対しゲージ不変量である長さの比と角度を比較することにより精度を評価する。

3.2.3 6軸モーションセンサによる外部パラメータ推定

カメラの外部パラメータはカメラ間の相対的な位置と姿勢を表すものである。したがって、第 1 画像のカメラ座標系に対する第 2 画像のカメラ位置を表すものが並進ベクトル、カメラの姿勢を表すものが回転行列であり、それぞれ t および R で表す(図 3.2)。

3.2.3.1 並進ベクトルの推定

センサにより観測される加速度から並進ベクトルを求める際、単純な 2 階積分による推定手法ではセンサのバイアスや機械的ノイズのため大きな誤差が生じてしまう。したがって、並進ベクトルの計算にはその誤差を低減するために、軌道を多項式によって近似し、その係数を推定する手法[44]を用いる。

3.2.3.2 回転行列の推定

角速度センサにより観測される角速度は x , y および z 軸の各軸に対する角速度であるが、実際の回転はある 1 つの軸に対して行われる。したがって、ここではその任意軸に対する回転を考える。単位ベクトル $U=(U_x;U_y;U_z)^T$ を回転軸、回転角度を θ とする回転行列 R は

$$R = I + \sin \theta [U]_{\times} + (1 - \cos \theta) [U]_{\times}^2 \quad (3.1)$$

で表される。ここで I は単位行列を表す。角速度センサによって観測される x , y および z 軸に対する角速度をそれぞれ ω_x , ω_y および ω_z とする。ベクトル量としての角速度の方向は、右ねじを角速度が示す回転方向に回転させたときに進む方向すなわち回転軸に沿った方向となる。任意軸に対する回転の例を図 3.3 に示す。各軸で観測される角速度は実際の角速度を各軸方向に分解したものである。実際の角速度とその方向は各角速度を合成したものおよびその方向となる。したがって、回転軸 U は

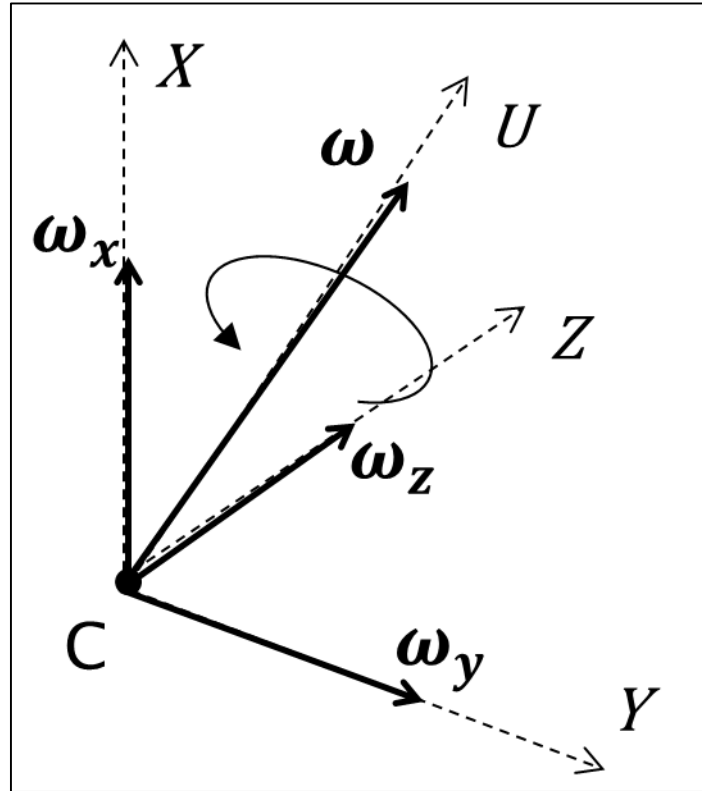


図 3.3 : 3 軸の角速度と任意軸回転

$$U = \frac{1}{\sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2}} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (3.2)$$

となる．また角速度は

$$\omega = \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (3.3)$$

回転角度 は

$$\theta = \omega \Delta t \quad (3.4)$$

となる．各サンプリング時間での回転行列を R_m とすると，回転行列は各時刻での回転行列の総乗で表されるので

$$\mathbf{R} = \prod_{m=1}^M \mathbf{R}_m \quad (3.5)$$

となる．

3.2.4 焦点距離の計算

次に，得られた並進ベクトル，回転行列を用いてカメラの焦点距離 f を計算する．3次元空間中の点 X が第1画像において $\mathbf{x} = [u \ v \ 1]^T$ ，第2画像において $\mathbf{x}' = [u' \ v' \ 1]^T$ と観測されたとする．また，計算した並進ベクトルを

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (3.6)$$

回転行列を

$$\mathbf{R} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \quad (3.7)$$

とする．また，カメラの移動前後で焦点距離は変化しないとすると，内部パラメータは $\mathbf{K} = \mathbf{K}'$ であるので，基礎行列 $\mathbf{F}$ は

$$\mathbf{F} = \begin{bmatrix} \frac{R_{31}t_y - R_{21}t_z}{f^2} & \frac{R_{32}t_y - R_{22}t_z}{f^2} & \frac{R_{33}t_y - R_{23}t_z}{f} \\ \frac{R_{11}t_z - R_{31}t_x}{f^2} & \frac{R_{12}t_z - R_{32}t_x}{f^2} & \frac{R_{13}t_z - R_{33}t_x}{f} \\ \frac{R_{21}t_z - R_{11}t_y}{f^2} & \frac{R_{22}t_z - R_{12}t_y}{f^2} & \frac{R_{23}t_z - R_{13}t_y}{f} \end{bmatrix} \quad (3.8)$$

のように表される．式(3.19)を式(10)に代入し， f について解くと

$$f = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (3.9)$$

$$a = R_{23}t_x - R_{13}t_y \quad (3.10)$$

$$b = u(R_{21}t_x - R_{11}t_y) + v(R_{22}t_x - R_{12}t_y) \\ + u'(R_{33}t_y - R_{23}t_z) + v'(R_{13}t_z - R_{23}t_x) \quad (3.11)$$

$$c = u\{u'(R_{31}t_y - R_{21}t_z) + v'(R_{11}t_z - R_{31}t_x)\} \\ + v\{u'(R_{32}t_y - R_{22}t_z) + v'(R_{12}t_z - R_{32}t_x)\} \quad (3.12)$$

となる．式(3.9)により求まる f は鏡像解の関係にある．これはシーンがカメラの前方にあっても後方にあっても数学的には同じ透視変換の式となることによる．したがって，鏡像解を除去するため焦点距離が実数かつ次の 2.2.5 において計測した距離が正となるものを真の焦点距離とする．

2.2.5 距離の計測

図 4 にカメラの位置と距離の関係を示す．第 1, 第 2 画像のカメラ座標系の原点から光軸に平行に測った距離の距離をそれぞれ Z, Z' とおく．同様に，それぞれの座標系の原点から画像上の点へのベクトルを $\mathbf{x}, \mathbf{x}'$ とおくと

$$\bar{\mathbf{x}} = \begin{bmatrix} u/f \\ v/f \\ 1 \end{bmatrix} \quad (3.13)$$

$$\bar{\mathbf{x}}' = \begin{bmatrix} u'/f \\ v'/f \\ 1 \end{bmatrix} \quad (3.14)$$

となる．ここで，第 2 画像のカメラ座標系は第 1 画像のカメラ座標系に相対的に $\mathbf{R}$ だけ回転しているので，第 2 画像上のベクトル $\mathbf{x}'$ は第 1 画像のカメラ座標系に対して $\mathbf{R} \mathbf{x}'$ と表される．したがって，第 1 画像のカメラ座標系における 3 次元空間中の点の座標を考えると

$$Z\bar{\mathbf{x}} = \mathbf{t} + Z'\mathbf{R}\bar{\mathbf{x}}' \quad (3.15)$$

の関係が成り立つ。式(37) に対して、両辺と $\mathbf{R}\mathbf{x}'$ との外積をとり、整理すると

$$Z = (\mathbf{t} \times \mathbf{R}\bar{\mathbf{x}}', \mathbf{n}) \quad (3.16)$$

となる。同様に式(3.27) の両辺と $\mathbf{x}$ との外積をとり、整理すると

$$Z' = (\mathbf{t} \times \bar{\mathbf{x}}, \mathbf{n}) \quad (3.17)$$

となる。ここで $(\cdot; \cdot)$ は内積を表し、 $\mathbf{n}$ は

$$\mathbf{n} = \frac{\bar{\mathbf{x}} \times \mathbf{R}\bar{\mathbf{x}}'}{\|\bar{\mathbf{x}} \times \mathbf{R}\bar{\mathbf{x}}'\|^2} \quad (3.18)$$

とする。ここで3.3 章で得られた焦点距離のうち、距離が正の値をとるものを真の焦点距離とする。また、3次元空間中の点 $\mathbf{X}$ の第1 カメラ座標系における位置は

$$\mathbf{X} = Z\bar{\mathbf{x}} \quad (3.19)$$

となる。

3.3 6 軸モーションセンサによるカメラ移動・回転角度推定実験

本論文に述べた手法を用いて、センサから得られた加速度および角速度データから移動距離および回転角度を計算し、精度を評価する。また得られた結果から多項式近似に用いる最大次数を決定する。移動の程度は手による移動の範囲を考え、移動距離に関しては 30cm 程度までの距離とし、回転角度に関しては 90 度程度までの角度に限定し、数値を変えて 10 回ずつ行う。

3.3.1 移動距離と最大近似次数の比較

多項式の最大次数として 5 次、7 次、9 次、11 次の 4 つを設定してそれぞれの移動距離を推定し、その結果を実際の移動距離と比較する。また多項式近似による結果との比較のために、単純な 2 階積分による移動距離も計算する。

図 3.4 に 2 階積分およびそれぞれ最大次数が 5 次、7 次、9 次、11 次の多項式近似による推定結果を示す。表 3.1 に推定手法ごとの推定した移動距離と実際の移動距離との平均誤差率を示す。誤差率 E_r は推定値を V_e 、実測値を V_m として

$$E_r = \frac{|V_e - V_m|}{V_m} \quad (3.20)$$

で求まる。2 階積分による結果はかなりばらつきが見られ、実際の移動距離よりも大きな値となっていることがわかる。2 階積分と比べ、5 次では誤差にそれほど違いは見られないが、次数が上がるにつれ実際の距離との誤差は小さく抑えられている。9 次近似と 11 次近似の結果を比較すると、平均誤差率はともに 0.03 と差はなく、実際の移動距離とほぼ同様の推定結果を出しているといえる。また、図 3.5 から図 3.9 に計測された加速度を 3 次から 11 次までの多項式で近似した例を示す。近似された加速度の変化をしてみると、9 次

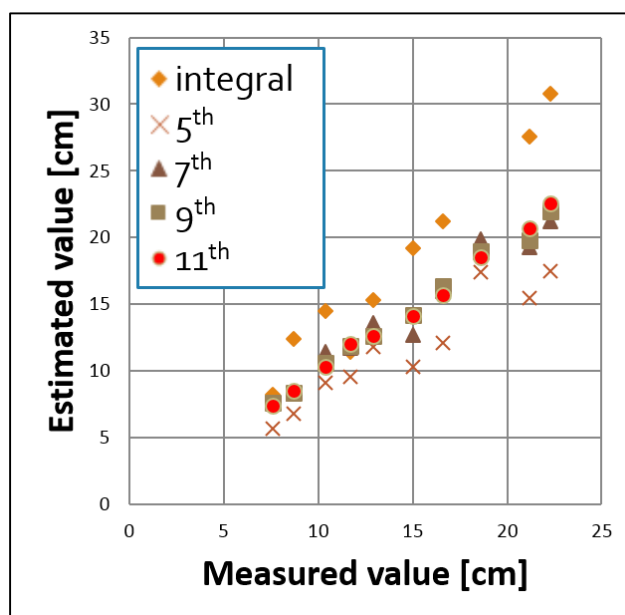


図 3.4：6 軸モーショセンサによる移動距離推定結果

表 3.1：移動距離の推定結果に対する平均誤差率

推定手法	平均誤差率
2 階積分	0.24
5 次近似	0.20
7 次近似	0.06
9 次近似	0.03
11 次近似	0.03

までは滑らかな曲線を描いており，手によるカメラの動きをよく再現できていると言えるが，11 次では曲線の形が崩れオーバーフィッティングが起こっている．したがって，精度と計算コストおよびオーバーフィッティングの問題を考慮し，移動距離の推定に用いる近似の最大次数は 9 次とする．

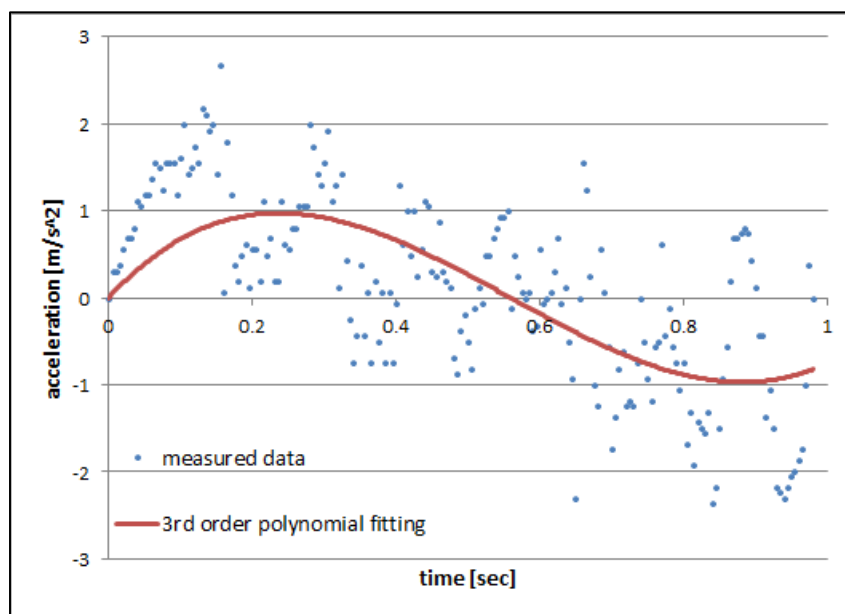


図 3.5：3次近似のときの加速度

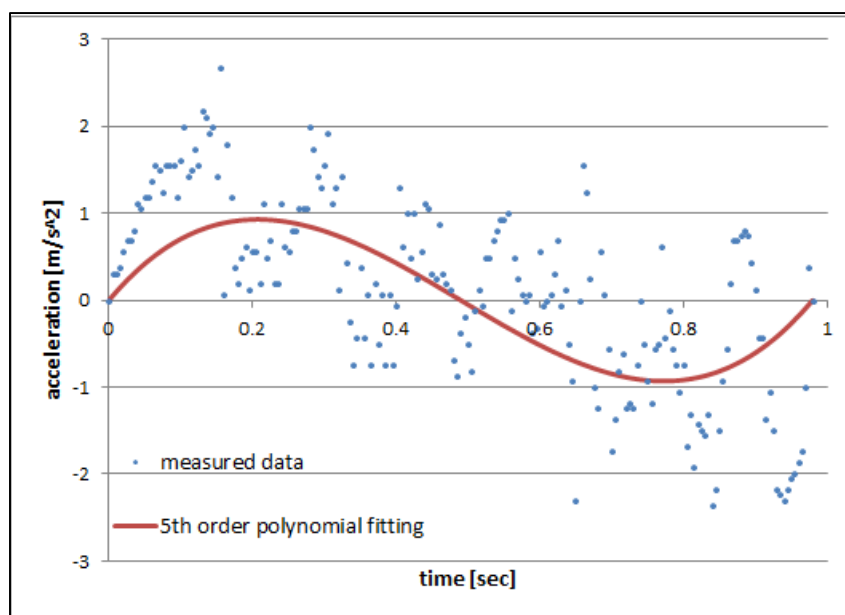


図 3.6：5次近似のときの加速度

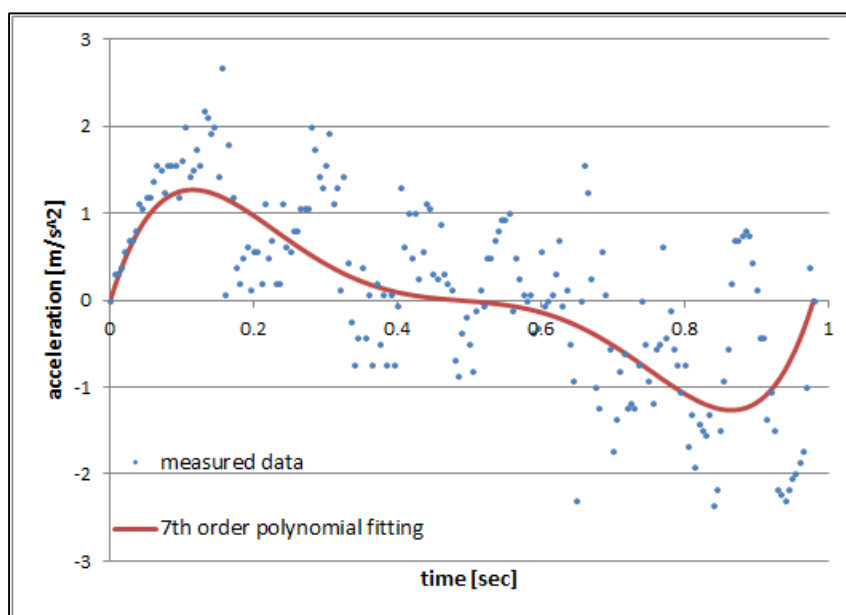


図 3.7：7次近似のときの加速度

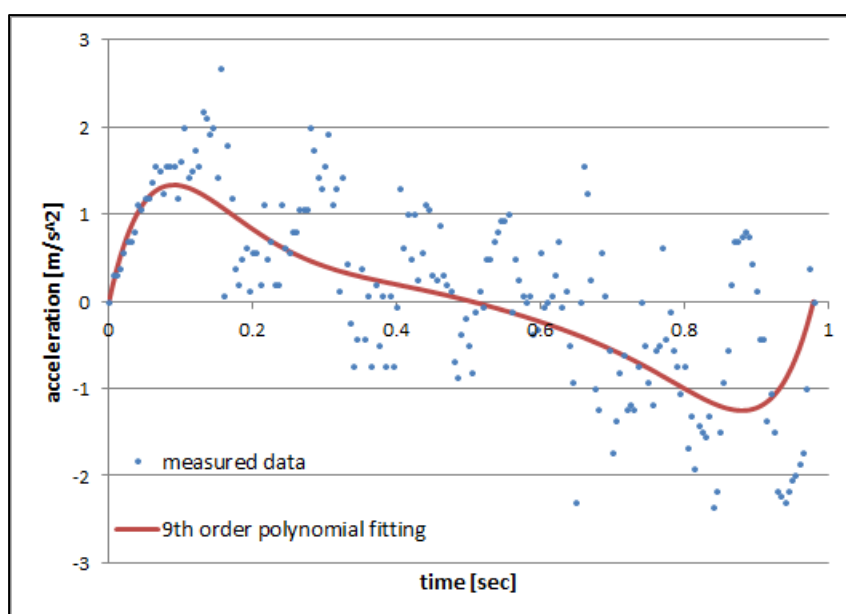


図 3.8：9次近似のときの加速度

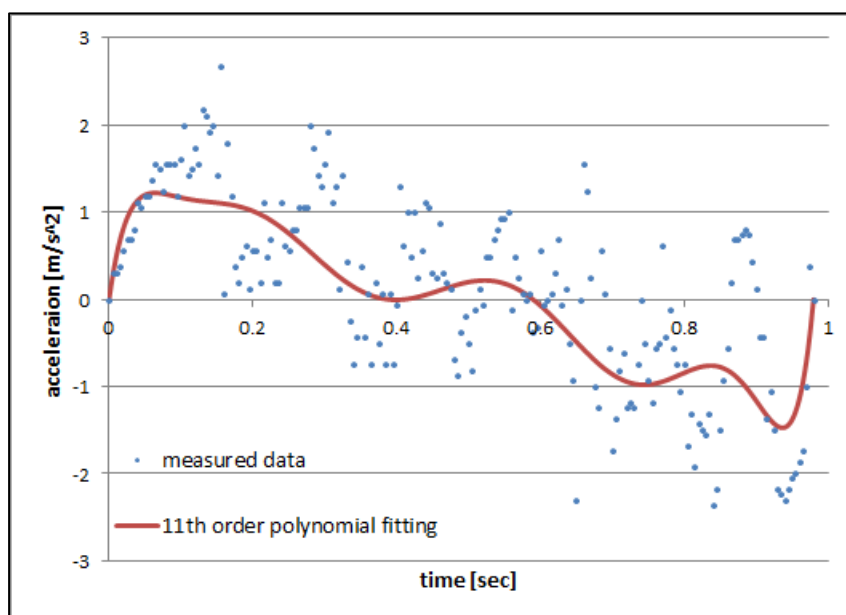


図 3.9：11 次近似のときの加速度

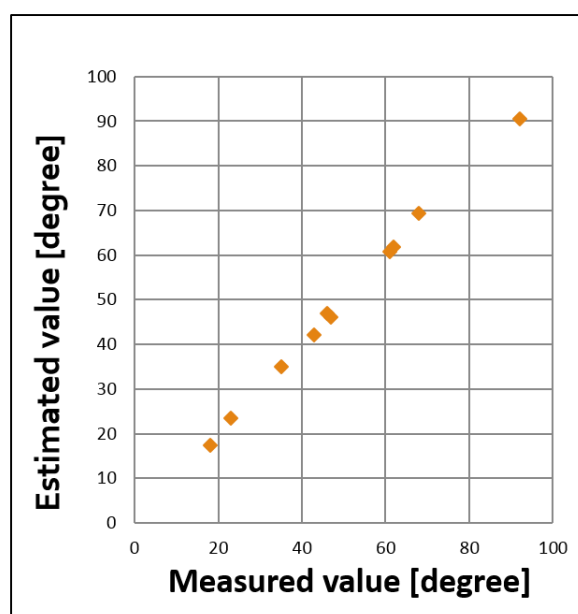


図 3.10：回転角度推定結果

3.3.2 回転角度の比較

回転角度についても実際の角度との比較を行い、その精度を評価する。図3.10 に実際の回転角度と推定した回転角度を示す。推定結果と実際の角度を比較してみるとほぼ同様の値が得られ、平均誤差率も0.02 とほとんど誤差はみられない。したがって、回転角度についてもセンサから得られた計測データから実際の回転角度をよく表すことができているといえる。

3.4 6 軸モーションセンサ装着移動単眼カメラによる

静止物体 3 次元距離計測実験

3.4.1 実験環境

実験にはウェブカメラとして Logitech 社製の QcamConnect, 6 軸モーションセンサとして MicroStone 社製の MVPRF8-AC を用いる。また、カメラの解像度は 640×480 , センサの測定周期は 5msec とする。計算機環境は IntelCore2 Duo CPU 2.54GHz, 2GB RAM, Windows 7, C++言語と OpenCV2.3 のライブラリを用いて実装する。

3.4.2 距離計測の精度評価

実際に撮影した画像から手対応点を抽出し、推定した外部パラメータを用いてそれぞれの点に対し焦点距離を算出し、距離を計測する。実験には、任意の視点から撮影した 2 組の画像を用いる。図 3.11, 図 3.12 および図 3.13, 図 3.14 に実験に用いた入力画像の組と抽出した対応点を示す。抽出した対応点の数は図 3.12 で 33 個, 図 3.14 で 20 個ずつである。焦点距離は抽出した対応点ごとに計算し、その平均をカメラの焦点距離とする。図 3.15, 図 3.16 は原画像をカメラ座標系に対して垂直な方向から眺めたものに対して、計測した対応点の 3 次元位置を重ね合わせたものである。これらから、計測した 3 次元位置は全体的に実際の対象の距離や形状をよく再現できていることがわかる。しかし、カメラの手前にある点に対して、奥にある点の方がずれが多少大きくなっている。手前の点と奥の点の計測精度の差を、ゲージ不変量である長さの比と角度を用いて比較する。具体的には、図 3.12 で抽出した点を用い



図 3.11：入力画像（缶）



図 3.12：抽出した特徴点（缶）



図 3.13：入力画像（人形）



図 3.14：抽出した特徴点（人形）



図 3.15：対応点の3次元計測結果（缶）

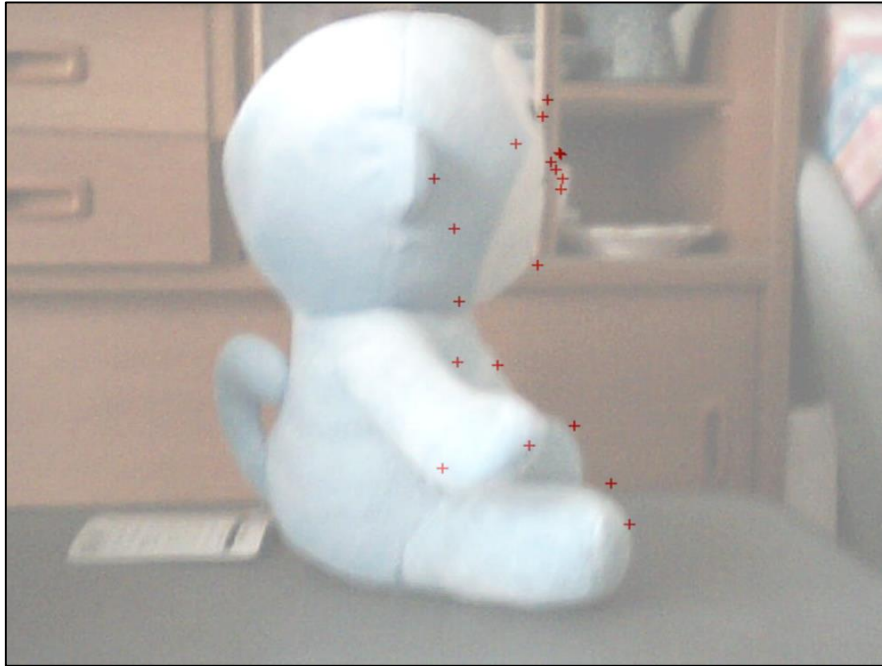


図 3.16：対応点の3次元計測結果（人形）

て2辺を作り，実測値と計測した3次元位置による計算値に対して辺の長さの比となす角を比較する．図 3.17 が手前，図 3.18 が奥の比較に用いた点と2辺である．表 3.2 と表 3.3 にそれぞれの結果を示す．手前の点に対しては，2辺の比および角度は実測値とほぼ等しいことが確認できる．しかし，奥の点に対しては2辺の比は実測値とほぼ変わらないものの，角度に多少の誤差がある．



図 3.17：手前の点に関するゲージ不変量



図 3.18：奥の点に関するゲージ不変量

表 3.2：図 11 に対するゲージ不変量の比較

	計算値	実測値	計算値と実測値の比
2 辺の比	3.23	3.02	1.07
角度[degree]	89.6	90.0	0.096

表 3.3：図 13 に対するゲージ不変量の比較

	計算値	実測値	計算値と実測値の比
2 辺の比	1.09	1.10	0.99
角度[degree]	96.1	90.0	1.07

3.4.3 距離計測比較実験

つぎに、従来法であるステレオカメラ手法を比較対象とし、提案手法との3次元計測結果の比較を行う。実験対象は同様に缶、および人形とする。比較実験の結果を表 3.4 に示す。実験対象を缶としたとき、計算時間は提案手法で 0.26 秒、ステレオカメラ手法で 0.83 秒である。また平均奥行き誤差は提案手法で 1.6 mm、ステレオカメラ手法で 0.9 mm である。この結果から、計算時間では提案手法のほうが約 3.2 倍速いものの、計測精度では約 1.8 倍の誤差が見られる。つぎに実験対象を人形とした場合、計算時間は提案手法で 0.30 秒、ステレオカメラ手法で 0.97 秒、平均奥行き誤差は提案手法で 2.1 mm、ステレオカメラ手法で 1.3 mm である。こちらの結果でも、計算時間では提案手法のほうが約 3.2 倍速いものの、計測精度では約 1.6 倍の誤差が見られる。どちらの結果からも提案手法は従来法と比べ、計算速度は速いものの精度の点では劣っているように思われる。しかしながら、実際の誤差は従来法と比較して 1 mm 以内であり、カメラから対象までの距離の 1 %以下に抑えられている。すなわち、その精度はほぼ同程度だと考えられる。したがって、提案手法は従

来法であるステレオカメラ手法と比較して、ほぼ同等の精度を保ったまま高速での3次元距離計測が可能であるといえる。

表 3.4：提案手法とステレオカメラ手法との比較実験結果

対象 手法	缶		人形	
	提案手法	ステレオカメラ 手法	提案手法	ステレオカメラ 手法
計算時間 [sec]	0.26	0.83	0.30	0.97
平均奥行き 誤差[mm]	1.6	0.9	2.1	1.3

3.5 本章のまとめ

6 軸モーションセンサを用いてカメラの外部パラメータを計算することで、任意視点画像から対象の距離を計測する手法を提案した。外部パラメータをセンサの計測データから直接推定することで、自由な環境での撮影が可能である。また、理論上 1 組の対応点を用いるのみで距離の計測が可能であり、計算時間も少なく抑えることができる。さらにカメラの移動距離が分かるため、実寸での計測が可能である。提案手法では、6 軸モーションセンサから得られた 3 軸加速度および角速度データからカメラの外部パラメータを推定する。並進ベクトルの推定には多項式近似を用い、回転行列はサンプリング時間ごとの任意軸回転を用いて

推定する。得られた外部パラメータおよび、対応点と基礎行列の関係から焦点距離を計算する。焦点距離については鏡像の関係から 2 つの値が得られるが、距離を計測する時点で正の値をとるものを真の焦点距離とする。最終的に、得られた距離を用いて対応点の 3 次元位置を計測する。

6 軸モーションセンサにより計算された外部パラメータの精度を検証するために、実際に計測した移動距離および回転角度との比較を行った。移動の程度は手による移動の範囲を考え、移動距離に関しては 30cm 程度までの距離とし、回転角度に関しては 90 度程度までの角度に限定し、数値を変えて 10 回ずつ行った。センサのサンプリング周期は 5 ミリ秒に設定した。実験結果から、並進ベクトルについては近似の次数が高くなるほど精度が高くなることを確認した。さらに、並進ベクトルの推定の精度は 9 次と 11 次で平均誤差率 0.03 とほぼ変わらないため、精度と計算コストおよびオーバーフィッティングを考慮し、9 次近似で十分な精度の推定が可能であることを示した。また、回転角度についても高い精度での推定が可能であることを確認した。また、提

案手法の有効性を評価するために、任意視点から撮影した2組の画像を用いて距離計測実験を行い、カメラに対して手前および奥の計測点からつくった2辺を用いて、ゲージ不変量である辺の比およびなす角を実測値と比較した。計測点と実測値との辺の比は手前で1.07、奥で0.99、角度比については手前で0.996、奥で1.07となり、ほぼ正確な距離計測ができることを確認した。また、従来法手あるステレオカメラ手法と比較した結果、計算速度で約3.2倍の向上が見られたが、精度では約1.7倍の低下が確認された。しかしながら、実際の誤差はカメラから対象までの距離の1%以下であり、結果としてほぼ同等の精度を保っていることが確認できた。

第4章

6軸モーションセンサ装着移動単眼カメラ による移動物体の3次元距離計測法

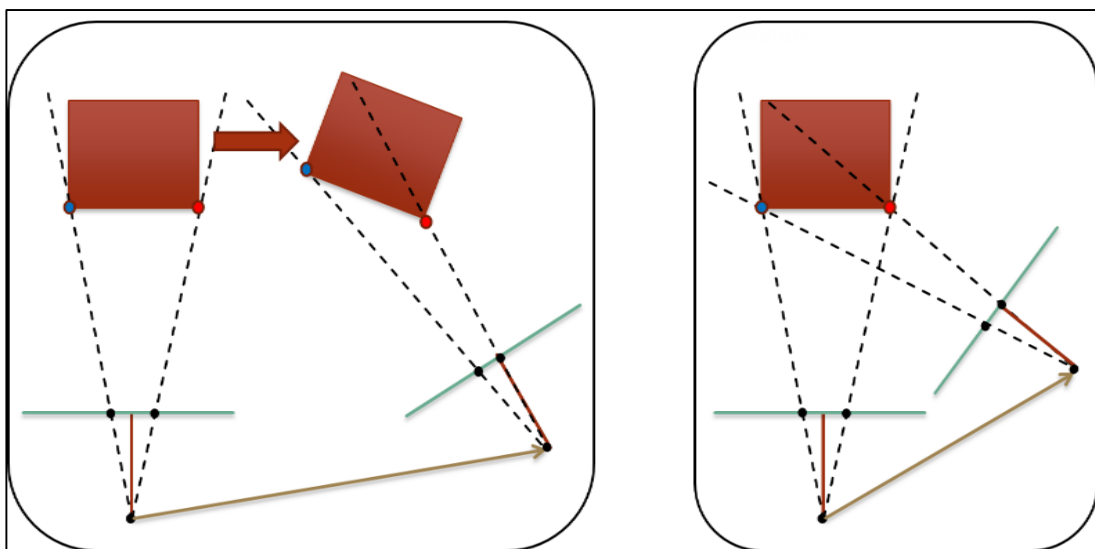
カメラで撮影した2枚の画像から基礎行列を用いて対象の3次元計測を行う場合、2枚の画像の対応点の視差を用いて計測を行うため、対象かカメラのどちらかが静止していなければならない[45, 46]。したがって、移動物体を対象とする3次元計測は固定単眼カメラによる手法か、ステレオカメラを用いて2枚の画像を同時に撮影する手法を用いる。固定単眼カメラを用いる手法ではカメラから対応点までの距離に加え対象の移動方向や移動距離を測定可能であるが、適用できる環境が限定されてしまう。一方ステレオカメラを用いる手法はカメラが静止・移動にかかわらず計測可能であるが、固定単眼カメラと比較して計算コストが高く、カメラも移動する状況では対象の移動方向や移動距離は計測できない[47-64]。本章では、計算コストと適用環境の観点から移動単眼カメラによる移動物体の3次元計測を6軸モーションセンサを用いて実現する手法を提案し、その計算コストと計測精度をステレオカメラを用いた手法と比較し評価する。

4.1 6軸モーションセンサ装着移動単眼カメラによる

移動物体3次元距離計測

4.1.1 仮基礎行列の推定

基礎方程式を用いた3次元計測手法では、対象とカメラのいずれかは静止しているという前提のもとに基礎行列が推定される。したがって、対象とカメラの両方が移動する場合、基礎方程式から推定される基礎行列は対象が静止しカメラのみが移動していると仮定した状態のものである。図4.1は対象とカメラの両方が移動する場合の実際の位置関係と推定された基礎行列が表す対象とカメラの仮の位置関係を示したものである。カメラと対象の両方が移動する場合とカメラのみが移動する場合とでは、カメラと対象の間の相対的な位置関係は同じである。したがって、計算された仮のカメラ間の位置を、6軸モーションセンサから得られるカメラの実際の動作によって修正することにより、移動後の対象の位置も推定が可能となる。



(a) 実際の対象とカメラの位置関係

(b) 推定された対象とカメラの位

4.2 6 軸モーションセンサ装着移動単眼カメラによる

移動物体の 3 次元距離計測実験

4.2.1 実験環境

本実験に使用するウェブカメラは logicool 社製の Qcam Connent (図 4.2), 6 軸モーションセンサは Micro Stone 社製の MVP-RF8-AC (図 4.3)である. カメラの解像度は 640×480 , 6 軸モーションセンサのサンプリング時間は 5 msec とする. 計算機環境は Intel Corei7 CPU 3.40GHz, 8GB RAM, OS は Windows 7 である. コーディングには C++を用い, ライブラリとして OpenCV 2.3 を利用している.

4.2.2 移動物体の対応点 3 次元計測実験

提案手法をもちいた移動物体の 3 次元計測精度を検証するために, 移動物体から抽出した対応点とデバイス間の距離を推定し, それぞれの 3 次元位置を計測, 実際の位置と比較する. また, 既存のステレオカメラ手法による推定結果との比較も行う. 提案手法で用いる入力画像 1 組および抽出した対応点を図 4.4 にステレオカメラ手法で用いる入力画像 2 組および抽出した対応点を図 4.5 および図 4.6 に示す. このとき提案手法とステレオカメラ手法とで結果を比較するため, 提案手法の入力画像はステレオカメラ手法で用いる入力画像の右側の画像を用いる. 対応点の抽出には特徴点記述子として Oriented-BRIEF (ORB)[33], 検出器として Features from Accelerated Segment Test (FAST)[32]を用い, Random sample consensus (RANSAC)を用いて誤対応を除去している. 提案手法の入力画像からは 68 組, ステレオカメラ手法の入力



図 4.2：提案手法に用いるカメラ



図 4.3：提案手法に用いる 6 軸モーショセンサ

画像からは 74 組の対応点を抽出している．提案手法により計測された対応点とデバイスとの位置関係を図 4.7 および 4.8 に，ステレオカメラ手法により計測された対応点とカメラとの位置関係を図 4.9 および 4.10 に示す．画像内の黒点が計測された対応点の位置，赤と青の三角形がデバイスおよびカメラの位置である．計測された対応点の位置と実際の位置とをランダムに 10 点ずつ抜き出し比較した結果，提案手法による再投影誤差は 0.28 pixel，ステレオカメラ手法による平均誤差率は 0.09 pixel である．

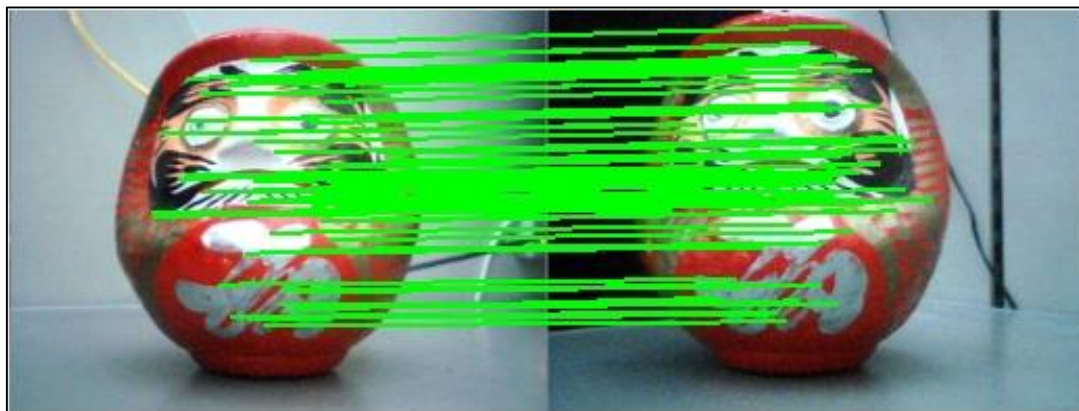


図 4.4：提案手法の入力画像

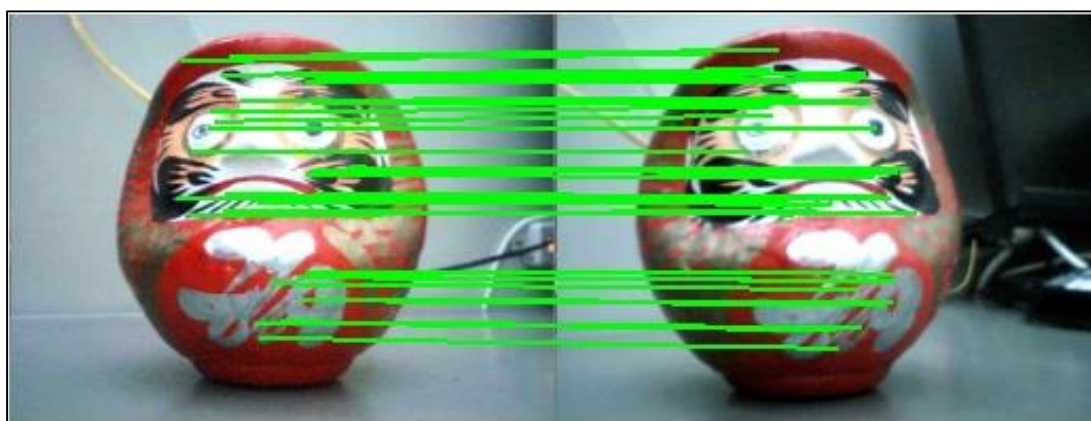


図 4.5：ステレオカメラ手法の移動前の入力画像

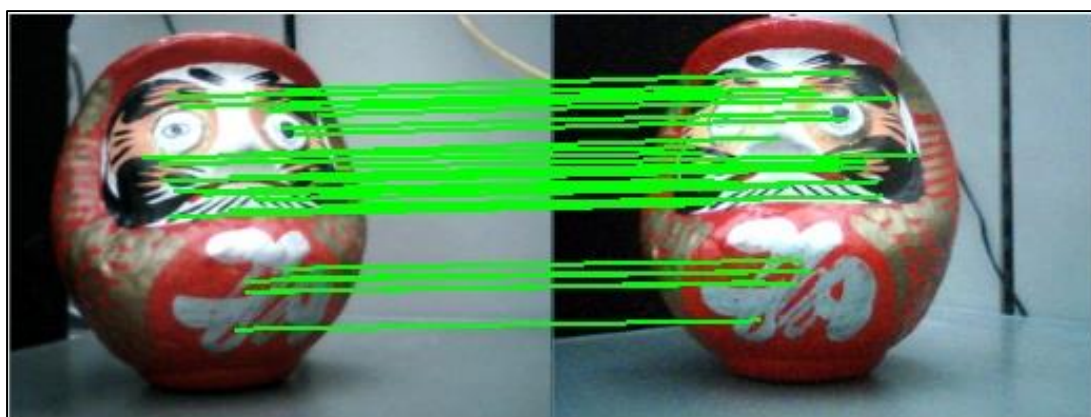


図 4.6：ステレオカメラ手法の移動後の入力画像

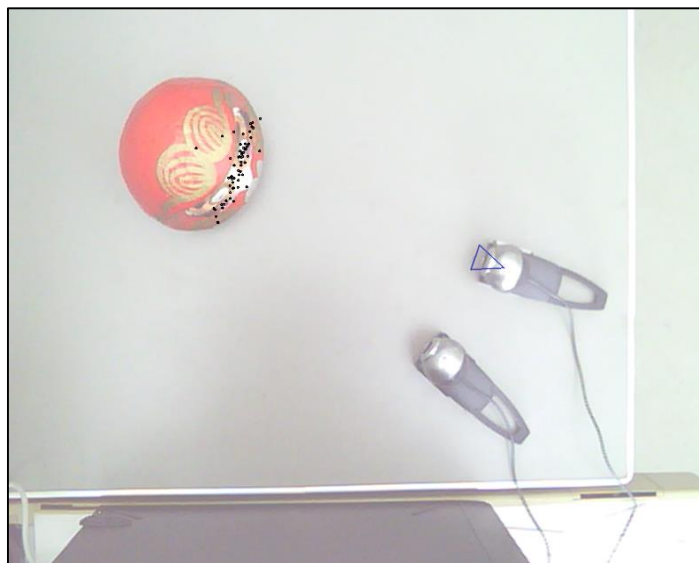


図 4.7：提案手法の移動前の対応点復元結果



図 4.8：提案手法の移動後の対応点復元結果

この結果から提案手法は既存のステレオカメラ手法と比較して、精度の観点からは多少劣るものの、実際の誤差は1 m に対し 2.8 cm という非常に小さな値であり、自動車やモバイルロボットの障害物回避システムというような実用には耐えうるものと判断する。続いて、それぞれの計測時間を計測し比較して結



図 4.9：ステレオカメラ手法の移動前の対応点復元結果

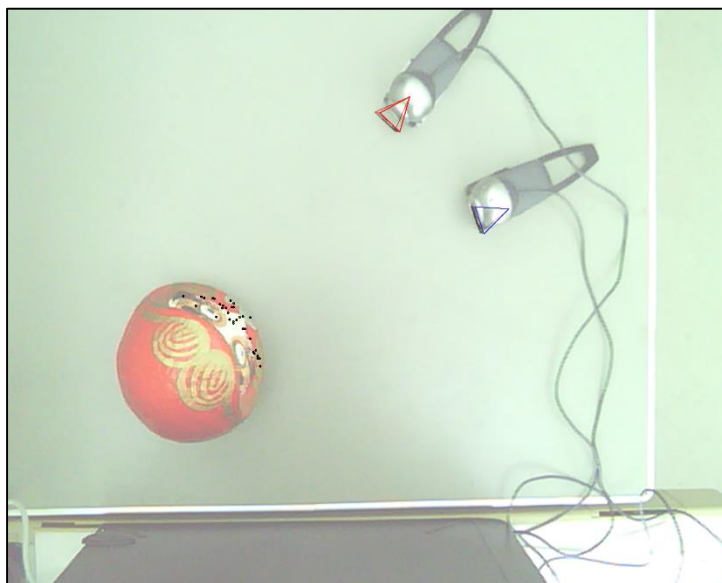


図 4.10：ステレオカメラ手法の移動後の対応点復元結果

果，提案手法の計測時間は 2.67 秒，ステレオカメラ手法での計測時間は 4.68 秒であり，提案手法はステレオカメラ手法と比較し 1.75 倍のコスト削減を確認している．以上の結果から，提案手法はステレオカメラ手法と比較し，実用に必要な精度を保ったまま約 2 倍のコスト削減を実現している．

4.3 本章のまとめ

6軸モーションセンサ装着移動単眼カメラを用いて移動物体の3次元計測を行う手法を提案した。計測された対応点の位置と実際の位置とをランダムに10点ずつ抜き出し比較した結果、提案手法による再投影誤差は0.28 pixel、ステレオカメラ手法による平均誤差率は0.09 pixelであった。この結果から提案手法は既存のステレオカメラ手法と比較して、精度の観点からは多少劣るものの、実際の誤差は1 m に対し2.8 cm という非常に小さな値であり、自動車やモバイルロボットの障害物回避システムというような実用には耐えうるものと判断される。それぞれの計測時間を計測し比較して結果、提案手法の計測時間は2.68 秒、ステレオカメラ手法での計測時間は4.68 秒であり、提案手法はステレオカメラ手法と比較し1.75 倍のコスト削減が確認された。以上の結果から、提案手法はステレオカメラ手法と比較し、実用に必要な精度を保ったまま約2 倍のコスト削減を実現しているといえる。

第 5 章

6 軸モーションセンサ装着

移動単眼カメラによる

3 次元距離計測のための対応点分類手法

5.1 静止物体と移動物体が混在するシーンにおける 3 次元計測

静止ステレオカメラを用いて単一（静止あるいは移動）対象物までの距離や形状を計測するには、対象物を撮影した左右 2 枚の画像から抽出した対応点を用い、基礎方程式を解いて 3×3 の基礎行列 $\mathbf{F}$ を求めるのが一般的である。基礎行列を求めるためには対応点が最低 7 組必要であることに加え、障害物回避用の車載カメラや移動ロボットの観測画像などの実応用を想定すると、フレームごとに 2 枚ずつの画像が必要であるため計算コストもかかる。移動単眼カメラを用いる場合、フレームごとに必要な画像は 1 枚に抑えられるが、撮影シーン内に静止対象物と移動対象物が混在している状況では、基礎行列を求めるのに必要な静止対象物の対応点 7 組を選別するのは難しい。また、どちらの場合も対応点を抽出する時点での誤対応や対応点座標の誤差の問題、さらに、対応関係をより正確に修正するためのバンドル調整コストの問題がある。

そこで本論文では、6 軸モーションセンサ（3 軸加速度センサと 3 軸角速度センサを組み合わせたもの）を単眼移動カメラに装着したデバイスを用いて、デバイスの回転角度と移動距離情報を利用して対応点を静止・移動点へ分類する

手法を提案し、それらを基礎行列の同定に利用する 3 次元計測手法の基礎を確立する.

提案手法では、静止・移動混在の複数対象物が撮影された単眼カメラ 2 フレームの対応点を、2 フレーム間で静止・移動に分類することが出来る. また、誤対応の除去および対応点座標のより正確な組を選別することによるバンドル調整時のコスト削減も可能である.

静止・移動対象物が混在するシーンを CG で作成し、計算機上で Python 2.7 を用いて提案手法による対応点の分類を行う.

5.2 6軸モーションセンサ装着移動単眼カメラによる

対応点分類手法

5.2.1 対応点の分類条件

3次元計測には基礎行列が必要であるが，移動単眼カメラが静止対象物と移動対象物が混在しているシーンを観測している場合には，基礎行列の推定に必要な7組以上の静止対象物の対応点を選別するのが難しい．そこで基礎行列を求めるために，6軸モーションセンサを移動単眼カメラに装着したデバイスを用いて各対応点を静止・移動点に分類する手法を提案する．

基礎行列はカメラの内部パラメータおよび外部パラメータに分解でき，

$$\mathbf{F} = \mathbf{K}'^T \mathbf{R}[\mathbf{t}]_{\times} \mathbf{K}, \quad (5.1)$$

と表される．ここで $\mathbf{K}$ および $\mathbf{K}'$ はカメラの内部パラメータ， $\mathbf{R}$ はカメラの回転行列， $\mathbf{t}$ は並進ベクトルを表し， $[\cdot]_{\times}$ は歪対称行列の演算子である．いまカメラの内部パラメータの内，アスペクト比を1，歪み係数を0，画像中心を $(0,0)$ ，カメラの移動前後で焦点距離 f が変わらないと仮定すると，

$$\mathbf{K} = \mathbf{K}' = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (5.2)$$

となり，これを式(1)に代入すると

$$\mathbf{p}'^T \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{R}[\mathbf{t}]_{\times} \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{p} = 0, \quad (5.3)$$

となる．ここで $\mathbf{R}$ および $\mathbf{t}$ を6軸モーションセンサにより推定した結果を用いると，式(5.3)は f についての二次方程式に帰着できる．すなわち式(5.3)が異符号の解をもつことが静止対象物の対応点の候補となるために必要な条件で

ある．ここで，式 (5.3) の左辺を展開し，各項の係数を A ， B ， C と置く．さらに全体を $Crt(f)$ と置くと，

$$Crt(f) = Af^2 + Bf + C = 0, \quad (5.4)$$

となる．このとき

$$\mathbf{R}[\mathbf{t}]_{\times} = \mathbf{E} = \begin{pmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{pmatrix}, \quad (5.5)$$

$\mathbf{p} = (x, y, 1)^T$ ， $\mathbf{p}' = (x', y', 1)^T$ と置くと，式 (5.5) の係数はそれぞれ

$$A = xx'E_{11} + xy'E_{21} + yx'E_{12} + yy'E_{22}, \quad (5.6)$$

$$B = xE_{31} + yE_{32} + x'E_{13} + y'E_{23}, \quad (5.7)$$

$$C = E_{33} \quad (5.8)$$

となる．したがって，静止対象物の対応点候補となる条件は A と C が異符号であること，すなわち

$$AC < 0 \quad (5.9)$$

である．

5.2.2 誤対応の除去と座標誤差の評価

静止対象物の対応点でなくても，誤対応や移動対象物がデバイスと平行に移動するなどの特殊な条件化で式 (5.9) を満たしてしまう場合がある．これらを排除するために候補から静止対象物の対応点を選別する手法を提案する．候補のうち静止対象物の対応点同士は同一の基礎行列で表現されるため式 (5.4) の解である f も同一である．したがって，式 (5.4) の軸を用い評価関数 Evl を

$$Evl = -\frac{B}{2A}, \quad (5.10)$$

と定義すると，静止対象物以外の対応点の Evl の値は外れ値となり，除去できる．また Evl は対応点の正確さを評価する指標としても利用でき， Evl の頻度の

高い範囲の対応点がより正確なものとみなせるので、それらを基礎行列の推定に用いることで推定精度を高めコストを削減することができる。

5.3 対応点分類手法検証実験

5.3.1 実験環境

提案手法の妥当性を検証するために、シミュレーションによる対応点分類実験を行う。入力画像は Blender[64]により作成した CG 画像を用いる。CG 画像はデバイスの移動前および移動後の 2 枚である。これら 2 枚の入力画像より対応点を抽出する。対応点は ORB および FAST を用いて抽出する。これらを提案手法により静止対象物と移動対象物に分類し、その分類精度を確認する。実験に用いる計算機は Windows 7 Professional で 64 bit OS, CPU は Intel® Core™ i5-3210M CPU 2.50 GHz, 実装メモリ (RAM) は 8.00 GB である。システムは Python 2.7 を用いてコーディングしている。

5.3.2 対応点分類シミュレーション実験

実験に使用するデバイスの移動前の画像を図 5.1, 移動後の画像を図 5.2 に示す。入力画像の物体の内、立方体が静止対象物、車とヘリコプタが移動対象物である。車は床上を平行移動、四角錐は空中を平行および回転移動している。各対象物とデバイスの移動方向および位置関係を図 5.3 に示す。デバイスに対し上方を x 軸, 右方を y 軸, 前方を z 軸とすると, デバイスの並進ベクトル $\mathbf{t}$ は $t_x = -0.50$, $t_y = -1.00$, $t_z = 1.00$, 回転角度は x 軸回りに $\pi/12$ rad である。

入力画像より ORB および FAST を用いて抽出した対応点を図 5.4 に示す。抽出された対応点は全 39 組であり, 内 24 組が静止対象物である立方体に属する対応点である。

次に対応点分類の精度を検証する。提案手法により静止対象物の対応点であると判定された組は全 39 組中 25 組, そのうち真に静止対象物に属する組は 24 組であり, 残りの 1 組は誤対応の組である (表 5.1)。

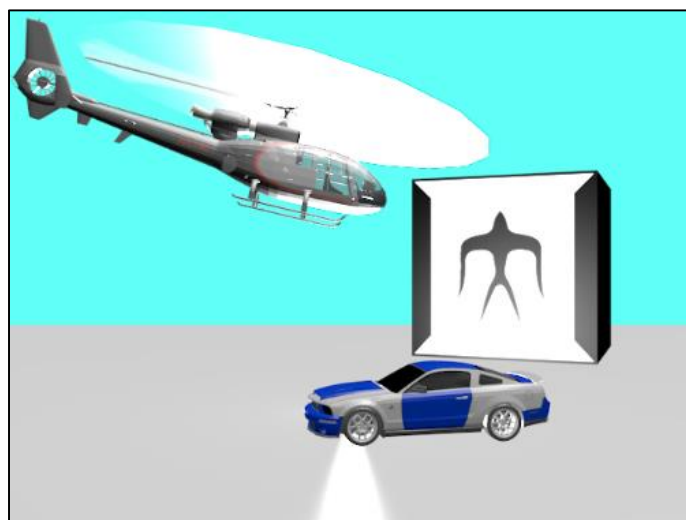


図 5.1：シミュレーション実験におけるデバイス移動前の入力画像

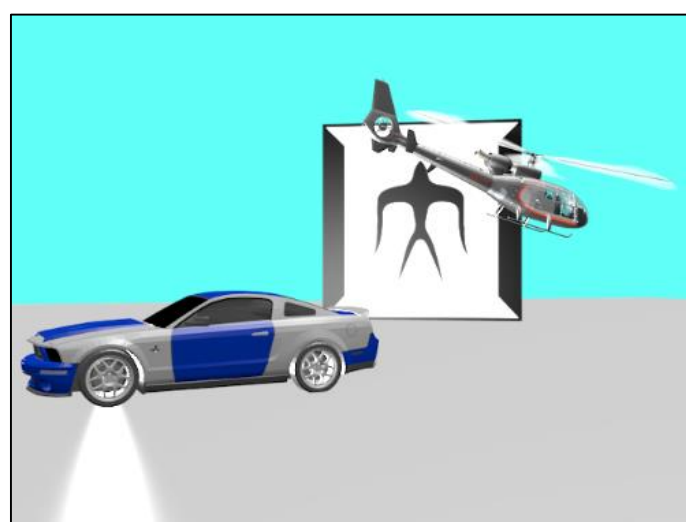


図 5.2：シミュレーション実験におけるデバイス移動後の入力画像

表 5.1：シミュレーション実験における対応点分類結果

True Positive	24
False Positive	1
False Negative	0
True Negative	14
Total	39

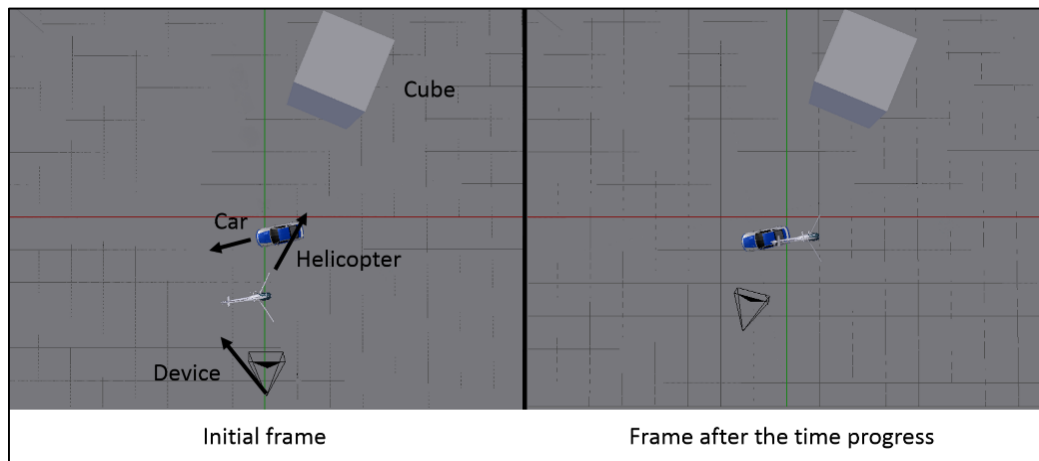


図 5.3：デバイスおよび各対象物の移動前後の位置関係に対する俯瞰図

図 5.5 に静止対象物に属すると判定された対応点を示す．この結果を指標にまとめると，正答率は 0.97，適合率は 0.96，再現率は 1.00 となり，F 値は 0.98 となる（表 5.2）．

最後に誤対応を除去が可能であるかを検証する．図 5.6 は誤対応を含め静止対象物の対応点と判定された 25 組の E_{vl} の値のヒストグラムである．誤対応の 1 組のみが外れ値となっていることが確認でき，さらに -0.01 から -0.008 の範囲の頻度が高くなっていることから，この範囲の対応点の精度がより高いことがわかる．図 5.7 は E_{vl} を用いて外れ値を除去した対応点の図である．除去には

表 5.2：シミュレーション実験における対応点分類結果に対する指標

Accuracy	0.97
Precision	0.96
Recall	1.00
F-measure	0.98

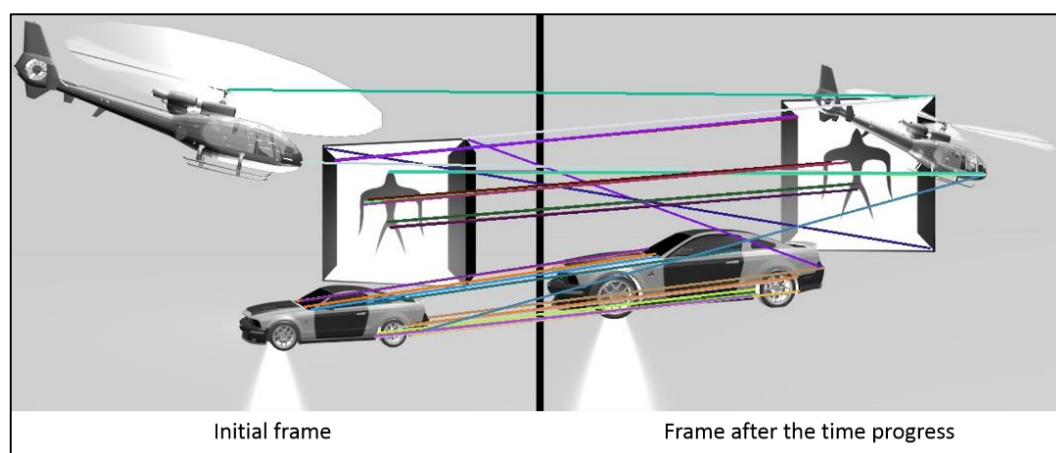


図 5.4：シミュレーション実験において抽出された対応点

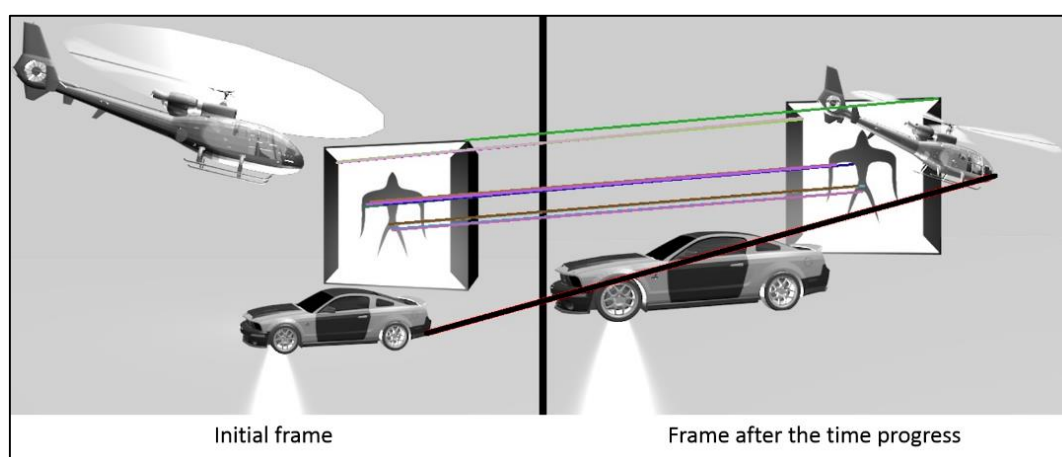


図 5.5：シミュレーション実験において
静止物体の対応点と分類された対応点

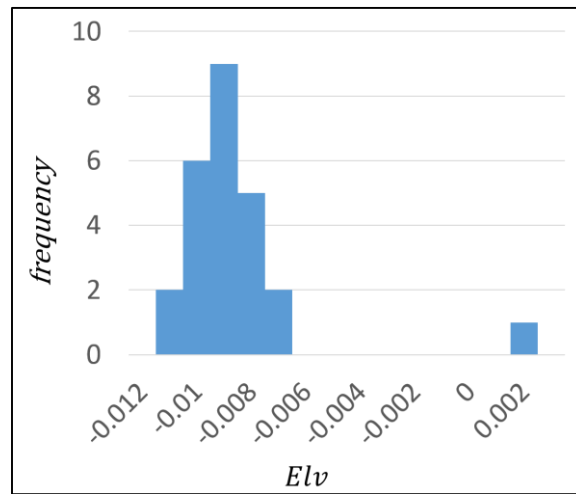


図 5.6：シミュレーション実験における
静止物体に属する対応点の各候補に対する Evl の値のヒストグラム

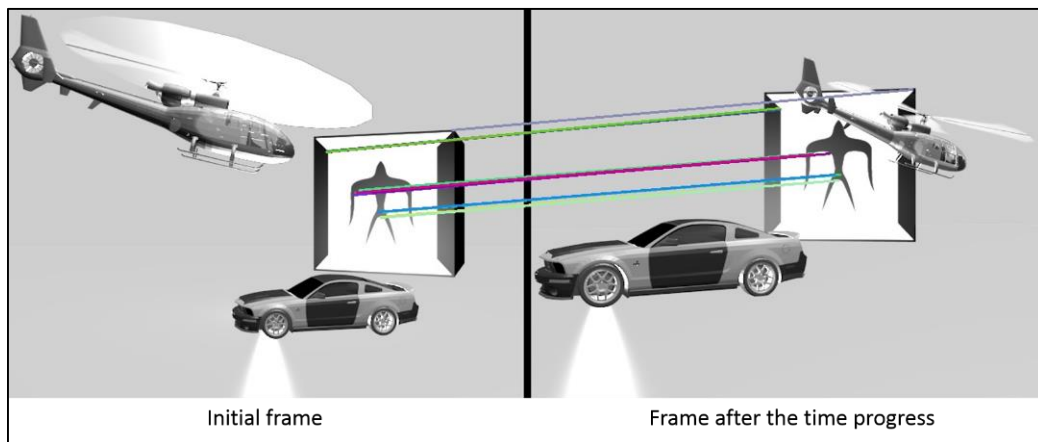


図 5.7：シミュレーション実験において
最終的に静止物体に分類された対応点

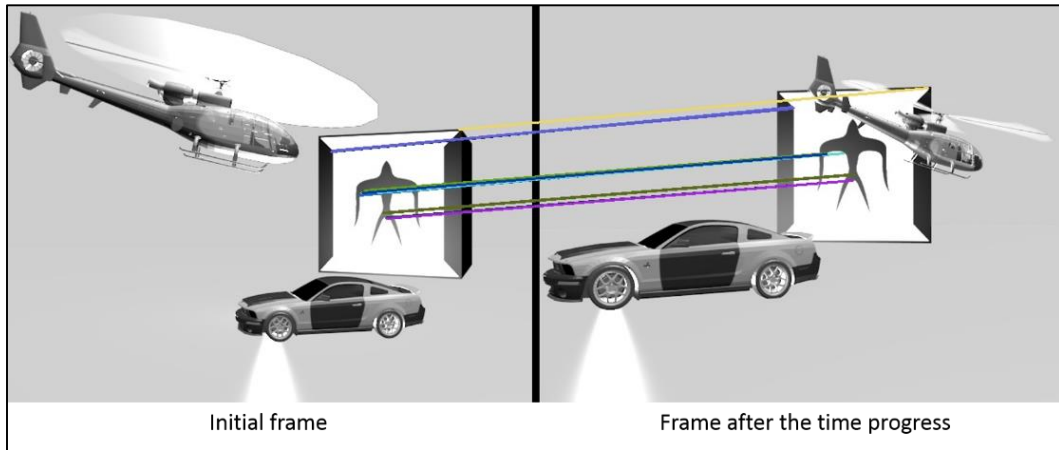


図 5.8 : $\sigma^2 = 0.1$ のときの静止物体に分類された対応点

Interquartile Range (IQR) を用いた外れ値検定を用いている．図 5.5 と比較すると車とヘリコプタの誤対応が除去されており，結果として，正確に静止物体の対応点のみを抽出できていることが確認できる．

5.3.3 6軸モーションセンサのノイズに対する分類手法の頑健性の検証

実際の 6 軸モーションセンサの性能を考慮すると，その計測データにはノイズが含まれていると考えなければならない．したがって，提案する分類手法が 6 軸モーションセンサのノイズに対しどの程度の頑健性を有するかを検証する．6 軸モーションセンサのノイズは，デバイスの移動ベクトルおよび回転角度に対し平均 0.0，分散 σ^2 正規分布に従うノイズを付加することによりシミュレーションする．付加するノイズの分散およびそのときに得られた移動ベクトルと回転角度，さらに各誤差を表 5.3 に示す．図 5.1 および図 5.2 の入力画像に対し，各ノイズを付加した 6 軸モーションセンサのデータを用いて対応点分類を行った結果を図 5.8, 5.9, 5.10, 5.11, 5.12 に示す．また，そのときの各対応点の E_{vl} の値をヒストグラムに表したものを図 5.13, 5.14, 5.15, 5.16, 5.17 に示す．

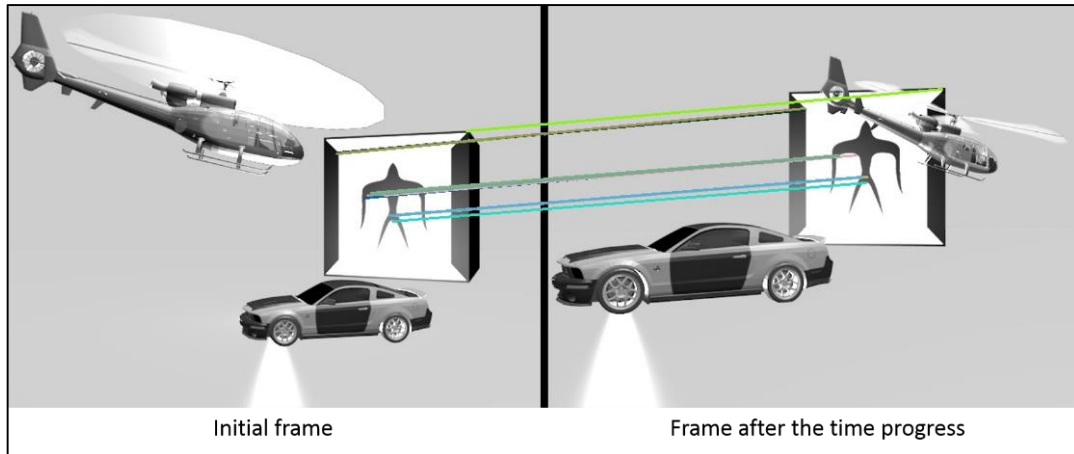


図 5.9 : $\sigma^2 = 0.2$ のときの静止物体に分類された対応点

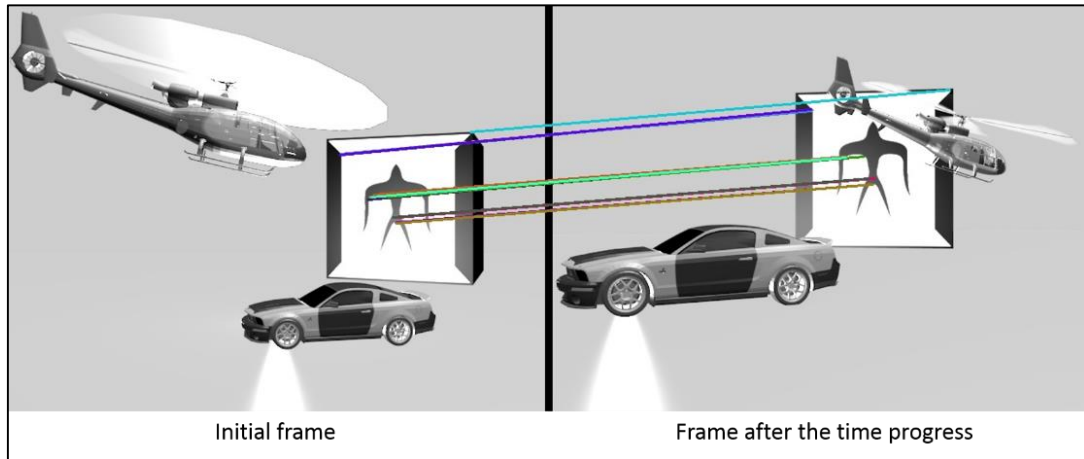


図 5.10 : $\sigma^2 = 0.3$ のときの静止物体に分類された対応点

Fig. 8, 9, 10 および Fig. 13, 14, 15 の結果から，付加する正規分布の分散 σ^2 が 0.3 の場合までは正確に立方体の対応点のみを抽出できていることが確認できるが，Fig. 11, 12 および Fig. 16, 17 から分散 σ^2 が 0.4 以上になると E_{vl} の値にばらつきが見られ，結果として誤対応を除去しきれていないことが確認できる．したがって，6 軸モーションセンサのノイズに対しては 20% 程度の誤差までは問題なく対応点の分類ができ，実際の計測環境においても 6 軸モーションセンサを用いた対応点分類手法が有効であることを示している．

表 5.3：移動ベクトルおよび回転角度に付加する正規分布の標準偏差と

付加後の値および誤差率

σ^2	0.1	0.2	0.3	0.4	0.5
t	$\begin{pmatrix} -0.46 \\ -0.95 \\ 1.10 \end{pmatrix}$	$\begin{pmatrix} -0.42 \\ -1.21 \\ 1.11 \end{pmatrix}$	$\begin{pmatrix} -0.71 \\ -0.91 \\ 0.85 \end{pmatrix}$	$\begin{pmatrix} -0.76 \\ -0.82 \\ 1.33 \end{pmatrix}$	$\begin{pmatrix} -0.72 \\ -0.69 \\ 1.27 \end{pmatrix}$
Average error of t	0.08	0.16	0.22	0.32	0.34
θ	0.29	0.22	0.28	0.21	0.37
Error of θ	0.11	0.16	0.07	0.20	0.41

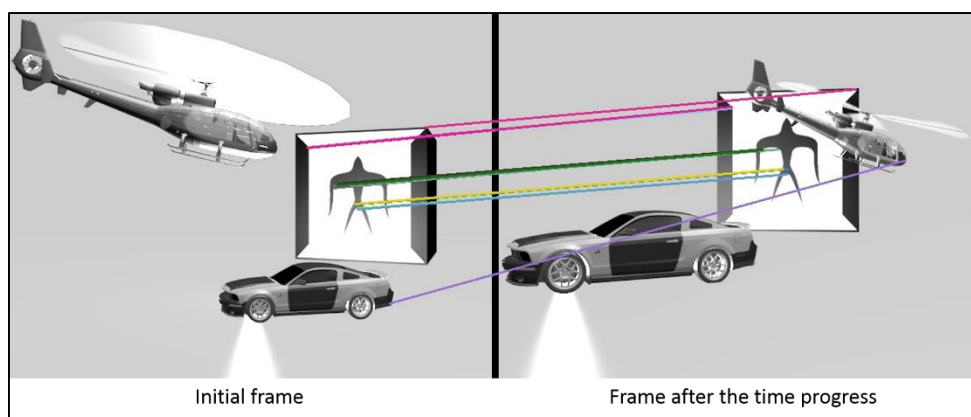


図 5.11： $\sigma^2 = 0.4$ のときの静止物体に分類された対応点

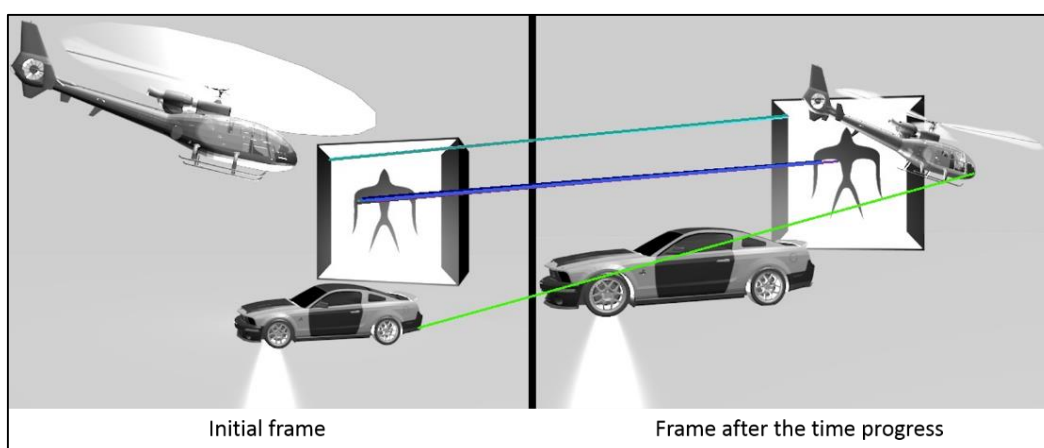


図 5.12： $\sigma^2 = 0.5$ のときの静止物体に分類された対応点

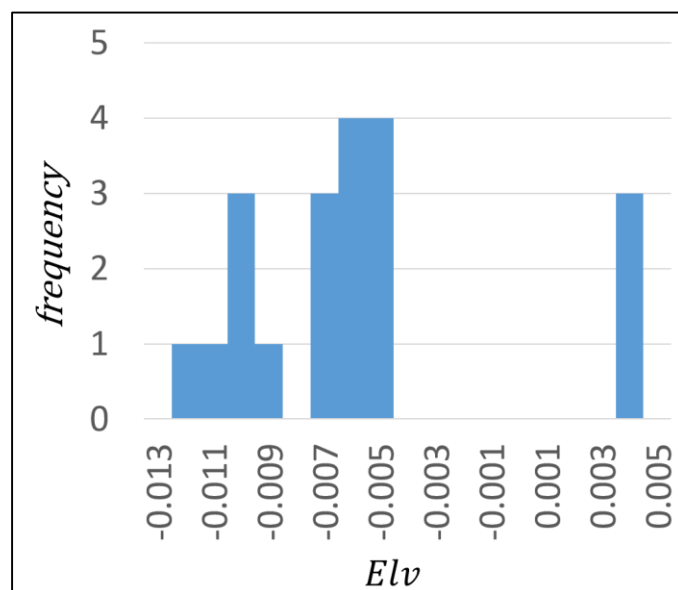


図 5.13 : $\sigma^2 = 0.1$ のときの Elv のヒストグラム

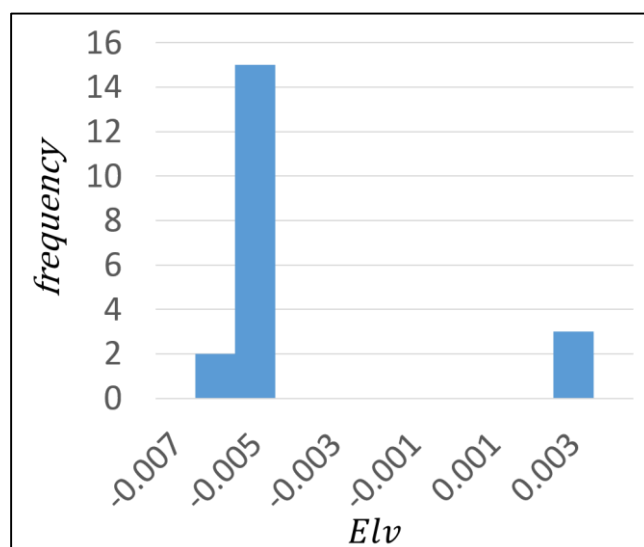


図 5.14 : $\sigma^2 = 0.2$ のときの Elv のヒストグラム

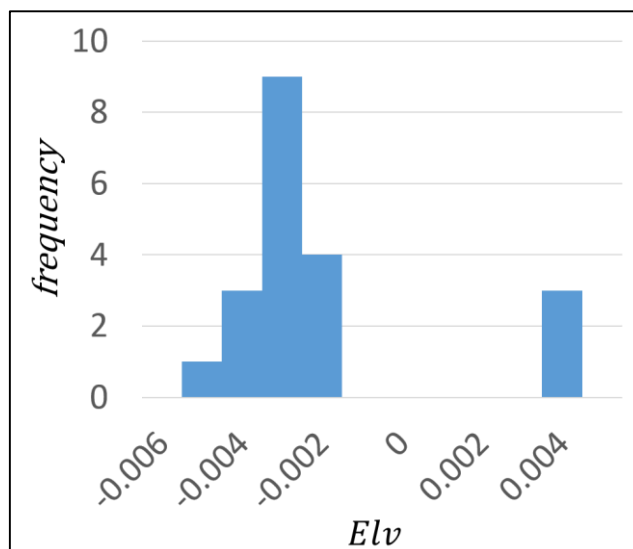


図 5.15 : $\sigma^2 = 0.3$ のときの Elv のヒストグラム

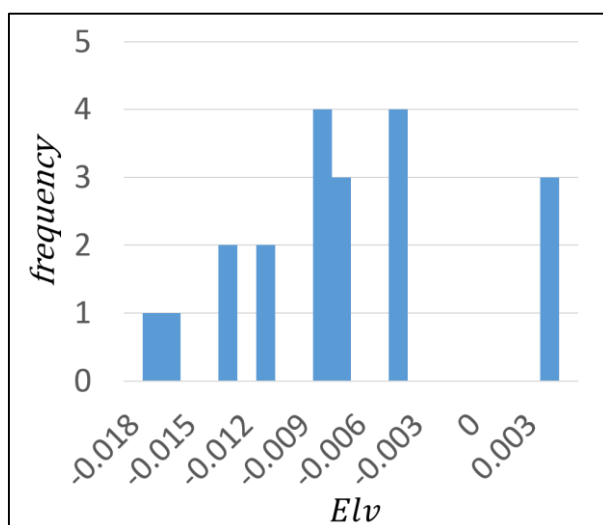


図 5.16 : $\sigma^2 = 0.4$ のときの Elv のヒストグラム

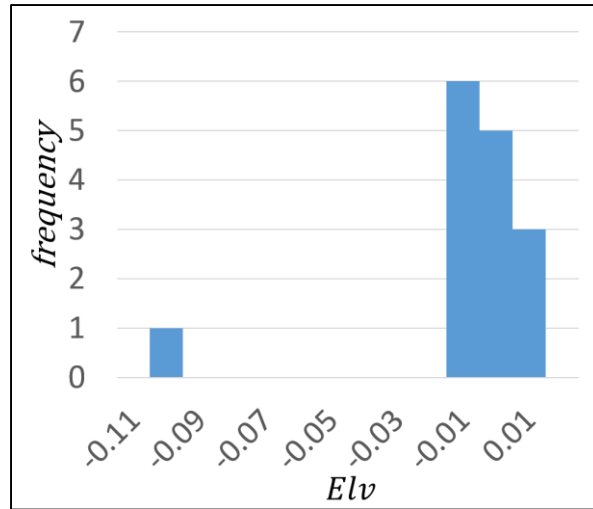


図 5.17 : $\sigma^2 = 0.5$ のときの Elv のヒストグラム

5.3.4 実画像を用いた対応点分類実験

提案する対応点分類手法が実環境においても有効であることを検証するために、実画像を用いた対応点分類実験を行う。実験に用いる入力画像を図 5.18 および図 5.19 に示す。デバイスの移動前が図 5.18，移動後が図 5.19 である。これらの入力画像において、移動物体は車とヘリコプタ，それ以外が静止物体である。対応点の抽出および対応付けには ORB および FAST を用いる。図 5.20 は抽出された対応点を表したものである。抽出された対応点は全部で 96 組であり，そのうち静止物体に属する対応点は 62 組である。これらの対応点に対し提案する分類手法を適用した結果を表 5.4 に示す。また，図 5.21 に提案手法により静止物体に属すると分類された対応点を示す。ただし，このうち太い赤線で示された対応点は，誤って静止物体に属すると分類された対応点である。この分類結果において，66 組の対応点が静止物体に属すると分類されているが，このうち実際に静止物体に属する対応点は 5 組であり，残りの 8 組は誤って分類された対応点である。さらに，4 組の静止物体に属する対応点が誤って静止物体に属さない対応点として分類されている。この結果を指標としてま



図 5.18：実画像実験におけるデバイス移動前の入力画像



図 5.19：実画像実験におけるデバイス移動後の入力画像

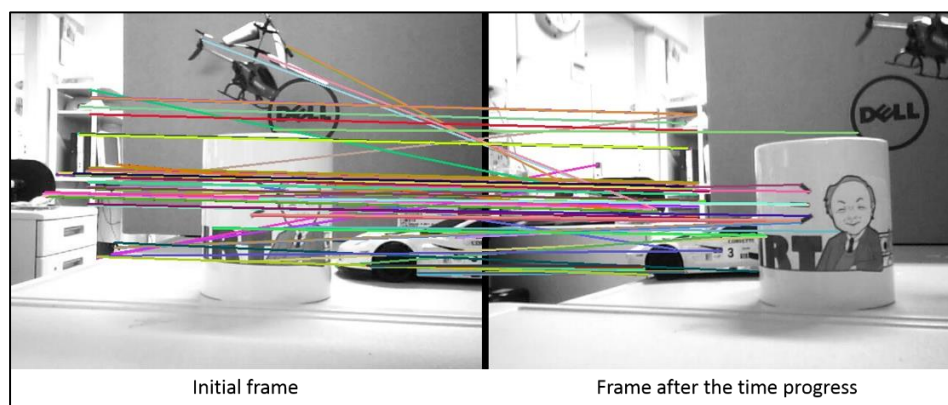


図 5.20：実画像実験において ORB および FAST により入力画像より
抽出された対応点

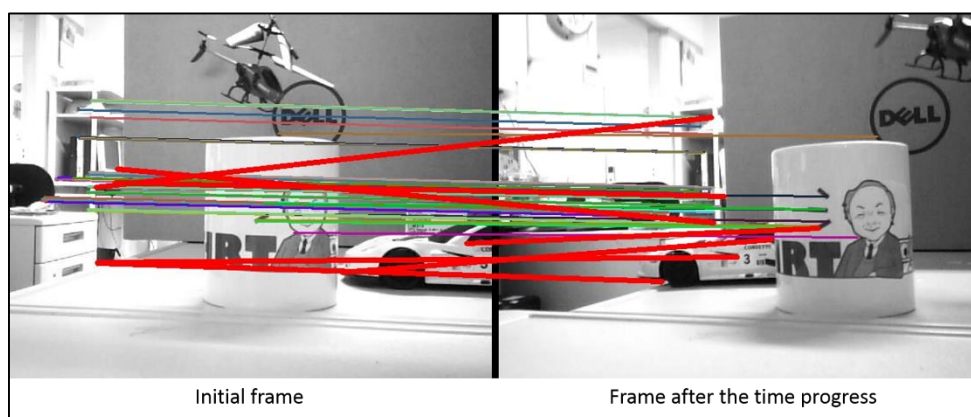


図 5.21：実画像実験において静止物体に属すると分類された対応点

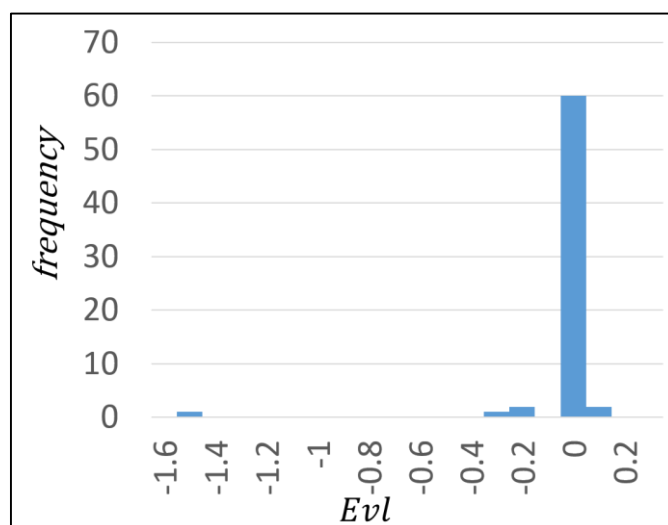


図 5.22：実画像実験における静止物体に属する対応点の各候補に対する
 Evl の値のヒストグラム

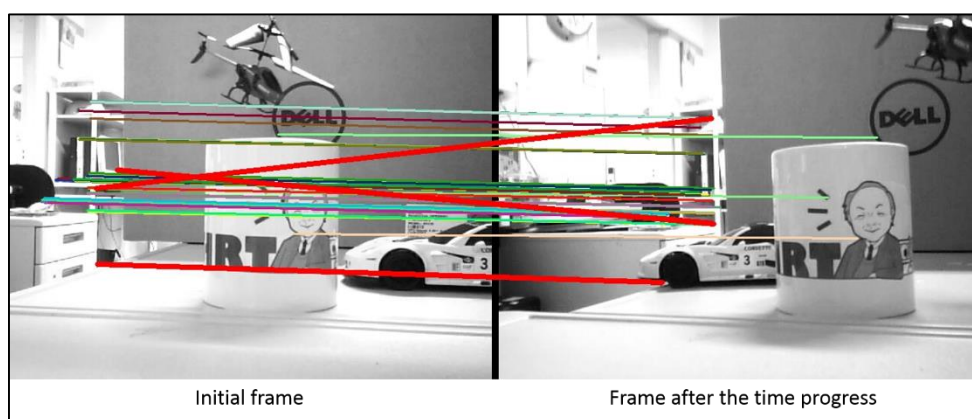


図 5.23：実画像実験における外れ値を除去した後の対応点

とめたものを表 5.5 に示す．次に，静止物体に属すると分類された各対応点に対し Evl の値を計算する．その結果をヒストグラムにまとめたものを図 5.22 に示す．また，外れ値を除去した後の対応点を図 5.23 に示す．この結果から，外れ値として 11 組の対応点が除去されているが，そのうち静止物体に属さない対応点は 4 組であり，7 組の静止物体に属する対応点を除去してしまっている(表 5.6)．これらの最終的な分類結果を指標として表したものが表 5.7 である．ここから，正答率，適合率，再現率および F 値はそれぞれ 0.84, 0.93, 0.87 および 0.82 である．この結果から，提案手法を用いずに抽出されたすべての対応点からランダムに 1 点を選んだ場合，それが静止物体に属している確率はおおよそ 68% であるが，提案手法をにより静止物体に属すると分類された対応点から 1 点を選んだ場合に，それが実際に静止物体に属している確率はおおよそ 93% である．したがって，提案手法を用いることにより実環境においても対応点の分類は有効であり，3 次元距離計測においてもその安定性を向上させることができると考える．

表 5.4：実画像実験における対応点分類結果

True Positive	58
False Positive	8
False Negative	4
True Negative	26
Total	96

表 5.5：実画像実験における対応点分類結果の指標

Accuracy	0.86
Precision	0.88
Recall	0.94
F-measure	0.91

表 5.6：実画像実験における外れ値を除去した後の対応点分類結果

True Positive	51
False Positive	4
False Negative	11
True Negative	30
Total	96

表 5.7：実画像実験における外れ値を除去した後の対応点分類結果の指標

Accuracy	0.84
Precision	0.93
Recall	0.82
F-measure	0.87

5.4 本章のまとめ

本章では移動・静止物体が混在するシーンにおいて 3 次元計測を実現するための 6 軸モーションセンサによる対応点分類手法を提案した。移動・静止対象物混在の CG テスト画像を用いた対応点分類実験で、ORB および FAST により抽出された全対応点 39 組中静止対象物に属する対応点は 24 組であるが、提案手法で静止対象物と判定された対応点は誤対応 1 組を含む 25 組である。これにより正答率 0.97, 適合率 0.96, 再現率 1.00 の結果を得ている。さらに対応点ごとに算出した E_{vl} に対して IQR を用いた外れ値検定を適用することにより、誤対応を除去できることを確認している。また、6 軸モーションセンサのノイズに対する提案手法の頑健性の検証を行い、およそ 20% 程度の誤差まで問題なく対応点分類ができていることを確認している。また、実画像を用いた対応点分類実験を行い、抽出された対応点 96 組およびその中の静止物体に属する対応点 62 組に対し、提案手法により静止物体に属すると分類された対応点は 55 組であり、その中で実際に静止物体に属する対応点 51 組を正確に分類できている。これにより、正答率、適合率、再現率および F 値はそれぞれ 0.84, 0.93, 0.87 および 0.82 の結果を得ている。さらに、静止物体の対応点に対する精度は、提案手法を用いない場合と比較して、約 25% 向上している。したがって、提案手法は実環境においても有効であり、3 次元距離計測の安定性を向上させることができると考える。

第 6 章

結論

本論文では、カメラを用いた 3 次元計測における処理コストの削減、適用環境の拡張を目的とし、カメラに 6 軸モーションセンサを装着したデバイスによる静止物体と移動物体の 3 次元計測手法、および 3 次元計測のための対応点分類手法の 3 つの提案を行っている。第 2 章では 6 軸モーションセンサ装着移動単眼カメラを用いた静止物体の 3 次元計測手法を提案しており、その計測精度と計算コストを評価している。第 3 章では 6 軸モーションセンサ装着移動単眼カメラを用いた移動物体の 3 次元計測手法を提案しており、その計測精度と計測コストの評価を行っている。第 4 章では静止物体と移動物体が混在するシーンにおける安定した 3 次元計測を実現するための対応点分類手法を提案しており、その分類精度とノイズへの頑健性を評価している。

第 2 章ではピンホールカメラの投影を射影幾何学に基づいてモデル化し、その校正法について簡単に記述した。前半では外界を $P3$ 、画像面を $P2$ と考え、斉次座標系で線型変換を記述することでピンホールカメラの投影が空間中のユークリッド変換、像面への射影変換、そして像面上のアフィン変換の合成で記述できることを示した。後半ではカメラの結像特性や姿勢の決定方法であるカメラ校正についていくつかの手法を記述した。最後に、ステレオカメラにおける点对応の数学的な記述として、エピポーラ幾何を扱い、これを記述する基礎行列に関していくつかの議論を行った。そして、基礎行列が座標系と独立なステレオカメラの射影計測であることを示した。

第 3 章で提案する 6 軸モーションセンサ装着移動単眼カメラを用いた静止物体の 3 次元計測手法は、デバイスの移動ベクトルおよび回転角度を 6 軸モーションセンサを用いて推定することにより、1 組の対応点のみを用いた基礎行列推定を実現している。6 軸モーションセンサからは 3 軸の加速度と 3 軸の角速度が得られる。移動ベクトルは多項式近似、回転角度は角速度を積分することにより推定している。6 軸モーションセンサに対する予備実験を行い、移動ベクトルの多項式近似による推定誤差およびオーバーフィッティングの影響、回転角度の積分による推定誤差を評価している。移動ベクトル推定では近似多項式の次数が 9 および 11 次ではともに誤差 0.03 と変化がないこと、11 次においてオーバーフィッティングが起きることを確認しており、この結果から 9 次近似が最適であると結論付けられる。回転角度推定では角速度の積分により得られた誤差は 0.02 である。また、提案手法の有効性を評価するために、任意視点から撮影した 2 組の画像を用いて距離計測実験を行い、カメラに対して手前および奥の計測点からつくった 2 辺を用いて、ゲージ不変量である辺の比およびなす角を実測値と比較した。計測点と実測値との辺の比は手前で 1.07、奥で 0.99、角度比については手前で 0.996、奥で 1.07 となり、ほぼ正確な距離計測ができることを確認した。実画像を用いた 3 次元距離計測比較実験より、従来のステレオカメラ手法と比較して 3.2 倍の計算時間の向上を確認した。精度に関しては 1.7 倍低下しているが、実際の誤差はカメラから対象までの距離の 1%以下に抑えられており、従来法との差は 1 mm 以下である。これにより、ほぼ同等の精度を保ったまま高速での 3 次元距離計測が可能であることを確認している。

第 4 章では、6 軸モーションセンサ装着移動単眼カメラによる移動物体の 3 次元計測手法を提案しており、計測された対応点の位置と実際の位置とをラン

ダムに 10 点ずつ抜き出し比較した結果、最投影誤差は提案手法で 0.28 pixel, ステレオカメラ手法で 0.09 pixel である。それぞれの計測時間を計測し比較して結果、提案手法の計測時間はステレオカメラ手法と比較し 1.75 倍のコスト削減を確認している。

第 5 章では移動・静止物体が混在するシーンにおいて 3 次元計測を実現するための 6 軸モーショセンサによる対応点分類手法を提案した。移動・静止対象物混在の CG テスト画像を用いた対応点分類実験で、ORB により抽出された全対応点 39 組中静止対象物に属する対応点は 24 組であるが、提案手法で静止対象物と判定された対応点は誤対応 1 組を含む 25 組である。これにより正答率 0.97, 適合率 0.96, 再現率 1.00 の結果を得ている。さらに対応点ごとに算出した E_{vl} に対して IQR を用いた外れ値検定を適用することにより、誤対応を除去できることを確認している。また、6 軸モーショセンサのノイズに対する提案手法の頑健性の検証を行い、およそ 20% 程度の誤差まで問題なく対応点分類ができていることを確認している。また、実画像を用いた対応点分類実験を行い、提案手法により静止物体に属すると分類された対応点は 55 組であり、その中で実際に静止物体に属する対応点 51 組を正確に分類できている。これにより、正答率、適合率、再現率および F 値はそれぞれ 0.84, 0.93, 0.87 および 0.82 の結果を得ている。さらに、静止物体の対応点に対する精度は、提案手法を用いない場合と比較して、約 25% 向上している。したがって、提案手法は実環境においても有効かつ、3 次元距離計測の安定性を向上させることができると思う。

最後に、6 軸モーショセンサ装着移動単眼カメラは従来法であるステレオカメラ手法よりも計算コストが低く静止・移動物体の混載するシーンにおいても有効であるため、実世界での速い計測が求められる車載カメラによる障害物

回避システムへの適用が考えられる．また，ステレオカメラ手法と比べ，カメラ 1 台と 6 軸モーションセンサによるデバイスは小型化が可能であり，携帯電話などのモバイル機器への搭載も期待できる．また，密な計測を行うことにより対象の形状復元も可能であるため，簡易的な 3 次元スキャナなどへの応用も可能であると考ええる．

参考文献

- [1] Z. Hu and Z. Tan, “Depth recovery and affine reconstruction under camera pure translation”, *Pattern Recognition*, Vol. 40, Issue 10, pp. 2826-2836, 2007.
- [2] S. Fuchs and G. Hirzinger, “Extrinsic and Depth Calibration of To F-cameras”, *Computer Vision and Pattern Recognition*, 2008. CVPR2008. IEEE Conference on, pp. 1-6, 2008.
- [3] R. Hartlry and C. Silpa-Anan, “Reconstruction from two views using approximate calibration”, *Proc. 5th Asian Conf. Comput. Vision*, Vol. 1, pp. 338-343, 2002.
- [4] Mei Han. “Linear and Bilinear Subspace Method for Structure from Motion”. PhD thesis, Carnegie Mellon University, Pittsburgh, PA 15213, February 2002.
- [5] Richard I. Hartley. “Euclidean Reconstruction from Uncalibrated Views”. In *Proceedings of Joint European – US Workshop on Applications of Invariance in Computer Vision*, pp. 237–256, 1993.
- [6] W. Chojnacki, M. J. Brooks, A. van den Hengel and D. Gawley, “A new constrained parameter estimator for computer vision applications”, *Image Vision Comput.*, 22-2 (2004-2), 85–91.
- [7] R. Hartley, “In Defense of the Eight-Point Algorithm”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 6, pp. 580–593, 1997.
- [8] R. Hartley and A. Zisserman, “Multiple View Geometry in Computer Vision Second Edition”, Cambridge University Press. March, 2004.
- [9] Richard I. Hartley and Frederik Schaffalitzky. “Power Factorization: 3D reconstruction with missing or uncertain data”. In *Proceedings of Australia-Japan Advanced Workshop on Computer Vision*, pp. 1–9, 2003.
- [10] Anders Heyden, Rikard Berthilsson, and Gunnar Sparr. “An iterative factorization method for projective structure and motion from image sequences”. *Image Vision Computing*, Vol. 17, No. 13, pp. 981–991, November 1999.
- [11] Anders Heyden and Du Q. Huynh. “Autocalibration via the Absolute Quadric

-
- and Scene Constraints”. In Proceedings of 16th International Conference on Pattern Recognition, Quebec, Canada, Vol. 2, pp. 631–634, August 2002.
- [12] Paul A. Beardsley, Andrew Zisserman, and David W. Murray. “Sequential Updating of Projective and Affine Structure from Motion”. *International Journal of Computer Vision*, Vol. 23, No. 3, pp. 235–259, June 1997.
- [13] Z. Zhang, “A flexible new technique for camera calibration”, *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 22, No. 11, pp. 1330–1334, 2000
- [14] Bill Triggs. “Autocalibration from planar scenes”. In proceedings of European conference on computer vision, Freiburg, Germany, pp. 89-105, 1998
- [15] Olivier D. Faugeras. “Stratification of three-dimensional vision: projective, affine, and metric representations”. *Journal of the Optical Society of America*, Vol. 12, No. 3, pp. 465–484, March 1995.
- [16] Olivier D. Faugeras and Quang-Tuan Luong. “The Geometry of Multiple Images : the laws that govern the formation of multiple images of a scene and some of their applications”. The MIT Press, 2001.
- [17] Y. Leedan and P. Meer, “Heteroscedastic regression in computer vision: Problems with bilinear constraint”, *Int. J. Comput. Vision.*, Vol. 37, No. 2, 127–150, 2006.
- [18] 出口光一郎. “多視点からの透視投影像の因子分解による 3 次元形状計測”. 計測自動制御学会論文集, Vol. 34, No. 10, pp. 1321–1328, October 1998.
- [19] 出口光一郎. “画像と空間—コンピュータビジョンの幾何学—”. 昭晃堂, 1991.
- [20] 出口光一郎. “ロボットビジョンの基礎”. コロナ社, 2000.
- [21] K. Kanatani and N. Ohta, “Comparing optimal three-dimensional reconstruction for finite motion and optical flow”, *J. Elec. Imaging*, Vol, 12, No. 3, 478–488, 2003.
- [22] Y. S. Hung and W. K. Tang. “Projective Reconstruction from Multiple Views with Minimization of 2D Reprojection Error”. *International Journal of Computer Vision*, Vol. 66, No. 3, pp. 305–317, March 2006.
- [23] Hanno Ackermann and Kenichi Kanatani. “Fast Projective Reconstruction: Toward Ultimate Efficiency”. *IPSJ Transactions on Computer Vision and Image*

Media, Vol. 49, No. SIG 6, pp. 68–78, March 2008.

- [24] 山田健人, 金澤靖, 金谷健一, 菅谷保之, “2 画像からの3 次元計測の最新アルゴリズム”, 情報処理学会研究報告, 2009-CVIM-168-15, pp. 1-8, 2009.
- [25] 金谷健一, 三島等, “未校正カメラによる2 画像からの3 次元計測とその信頼性評価”, 情報処理学会研究報告, 2000-CVIM-121-4, pp. 25-32, 2000.
- [26] Z. Zhang, R. Deriche, O. D. Faugeras and Q. Luong, “A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry”, Research Report 2273, 1994.
- [27] O. D. Faugeras. “Stratification of three dimensional vision: projective, affine, and metric representations”, Journal of the Optical Society of America, Vol. 12, No. 3, pp. 465-484, 1995.
- [28] Stan Birchfield. “KLT: An Implementation of the Kanade-Lucas-Tomasi Feature Tracker”. <http://www.ces.clemson.edu/~stb/klt/>.
- [29] D. G. Lowe. “Distinctive image features from scale-invariant keypoints”. International Journal of Computer Vision, Vol. 60, No. 2, pp. 91–110, 2004
- [30] H. Bay, T. Tuytelaars, and L. Van Gool. “Surf: Speeded up robust features”. In European Conference on Computer Vision, May 2006
- [31] E. Rosten and T. Drummond. “Machine learning for highspeed corner detection”. In European Conference on Computer Vision, vol. 1, 2006.
- [32] E. Rosten, R. Porter, and T. Drummond. “Faster and better: A machine learning approach to corner detection”. IEEE Trans. Pattern Analysis and Machine Intelligence, No. 32, pp. 105–119, 2010.
- [33] Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski "ORB: an efficient alternative to SIFT or SURF", Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011.
- [34] Hanno Ackermann and Kenichi Kanatani. “Iterative low complexity factorization for projective reconstruction”. In Proceedings of the 2nd Workshop on Robot Vision, pp. 153–164, 2008.
- [35] Padmanabhan Anandan and Michal Irani. “Factorization with Uncertainty”. International Journal of Computer Vision, Vol. 49, No. 2, pp. 101–116, October

2002.

- [36] Adrien Bartoli. “A Unified Framework for Quasi-Linear Bundle Adjustment”. In Proceedings of IEEE Conference on Pattern Recognition, Quebec City, Canada, pp. 560–563, August 2002.
- [37] Martin Bergtholdt. “Auto-Calibration with Convex Constraints”. PhD thesis, University of Mannheim, Mannheim, Baden-Württemberg, February 2002.
- [38] Aeron M. Buchanan and Andrew W. Fitzgibbon. “Damped Newton Algorithms for Matrix Factorization with Missing Data”. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, Vol. 2, pp. 316–322, June 2005.
- [39] Pei Chen. Heteroscedastic “Low-Rank Matrix Approximation by the Wiberg Algorithm”. IEEE Transactions on Signal Processing, Vol. 56, No. 4, pp. 1429–1439, April 2008.
- [40] Pei Chen. “Optimization Algorithms on Subspaces: Revisiting Missing Data Problem in Low-Rank Matrix”. International Journal of Computer Vision, Vol. 80, No. 1, pp. 125–142, October 2008.
- [41] Pei Chen and David Suter. “Recovering the Missing Components in a Large Noisy Low-Rank Matrix: Application to SFM”. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 8, pp. 1051–1063, August 2004.
- [42] Juyang Weng, Paul Cohen, and Marc Herniou. “Camera Calibration with Distortion Models and Accuracy Evaluation”. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 10, pp. 965–980, 1992.
- [43] Olivier D. Faugeras and Theo Papadopoulos. “Grassmann-Cayley algebra for modeling systems of cameras and the algebraic equations of the manifold of trifocal tensors”. Research Report 3225, INRIA, July 1997.
- [44] T. Kagawa, and Y. Uno, “Gait pattern generation for a power-assist device of paraplegic gait”, Proceedings of The 18th IEEE International Symposium on Robot and Human Interactive Communication, pp. 633–638, 2009.
- [45] Garbriella Csurka and Radu Horaud. “Finding the Collineation Between two Projective Reconstructions”. Research Report 3468, INRIA, August 1998.
- [46] Olivier D. Faugeras. “Three-dimensional computer vision : a geometric

-
- viewpoint”. The MIT Press, 1993.
- [47] Andrew Fitzgibbon, Maurizio Pilu, and Robert B. Fisher. “Direct Least Square Fitting of Ellipses”. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 5, pp. 476–480, May 1999.
- [48] Gene H. Golub and Charles F. Van Loan. “Matrix Computations”. The Johns Hopkins University Press, Third edition, 1996.
- [49] Nicolas Guilbert, Adrien Bartoli, and Anders Heyden. “Affine Approximation for Direct Batch Recovery of Euclidian Structure and Motion from Sparse Data”. International Journal of Computer Vision, Vol. 69, No. 3, pp. 317–333, April 2006.
- [50] N. Chernov and C. Lesort, “Statistical efficiency of curve fitting algorithms”, Comput. Stat. Data Anal., Vol. 47, No. 4, pp. 713–728, 2004.
- [51] W. Chojnacki, M. J. Brooks, A. van den Hengel and D. Gawley, “On the fitting of surfaces to data with covariances”, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 22, No. 11, pp. 1294–1303., 2000
- [52] W. Chojnacki, M. J. Brooks, A. van den Hengel and D. Gawley, “FNS, CFNS and HEIV: A unifying approach”, J. Math. Imaging Vision, Vol. 23, No. 2 , pp. 175–183, 2005.
- [53] 金谷健一, “コンピュータビジョンのためのくりこみ法”, 情報処理学会論文誌, 35-2, pp. 201–209, 1994.
- [54] 金谷健一, “当てはめ問題の最適推定と精度の理論限界”, 情報処理学会論文誌, 36-8, pp. 1865–1873, 1995.
- [55] 金谷健一, “幾何学的補正問題の最適計算と精度の理論限界”, 情報処理学会論文誌, 37-3, pp. 363–370, 1996.
- [56] K. Kanatani, “Geometric Computation for Machine Vision”, Oxford University Press, Oxford, U.K., 1993.
- [57] K. Kanatani, “Statistical Optimization for Geometric Computation: Theory and Practice”, Elsevier Science, Amsterdam, The Netherlands, 1996; Dover, New York, 2005.

-
- [58] 金谷健一, ” くりこみ法その後:波紋と発展”, 情報処理学会研究報告, 2003-CVIM-139-5, pp. 33–40, 2003.
- [59] 金谷 健一, “最尤推定の最適性と KCR 下界”, 情報処理学会研究報告, 2005-CVIM-147-8, pp. 59–64, 2005.
- [60] 金谷 健一, ” くりこみ法の謎を解く”, 情報処理学会研究報告, 2005-CVIM-149-3, pp. 15–22, 2005.
- [61] 金谷健一, “これなら分かる最適化数学-基礎原理から計算手法まで-“, 共立出版, 2005.
- [62] 三島等,金谷健一, “基礎行列の最適計算とその信頼性評価”, 情報処理学会研究報告, 99-CVIM-118-10, pp. 67–74, 1999.
- [63] G. Taubin, “Estimation of planar curves, surfaces, and nonplanar space curves defined by implicit equations with applications to edge and range image segmentation”, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 13, No. 11, pp. 1115–1138, 1991.
- [64] <http://www.blender.org/>

業績リスト

論文

- [J1] **Toshihiro Akamatsu** et al., “*Depth Measurement of Still 3D Object using Camera and 3D Accelerometer*”, International Journal of Computer Vision (投稿済み)
- [J2] **Toshihiro Akamatsu** et al., “*3D Measurement of a Moving Object Using a Moving Camera Attached with a 6-Axis Sensor*”, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 18, No. 5, pp. 736-744, 2014
- [J3] **Toshihiro Akamatsu** et al., “*Still Corresponding Points Extraction Using a Moving Monocular Camera with a Motion Sensor*”, Journal of Advanced Computational Intelligence and Intelligent Informatics Special Issue on ITCA & ISCHIA 2014 (採録済み, 2015 年 3 月出版予定)

査読付き国際会議

- [C1] **Toshihiro Akamatsu** et al., “*3D Measurement of a Moving Object Using 3D Accelerometer Attached to Moving Camera*”, International Symposium on Intelligence Signal Processing, pp.71-76, Funchal, Madeira, Portugal, September 16-18, 2013

-
- [C2] **Toshihiro Akamatsu** et al. , “*3D Measurement of a Moving Object Using a Moving Monocular Camera Attached with a 6-axis sensor*”, Proceedings of Joint Conference of ITCA2014 & ISCIIA2014, pp.110-116, Changsha, China, 15-20 September, 2014

国内会議

- [D1] 赤松俊弘, 董芳艶, 廣田薫 “6 軸センサ付きカメラを用いた三次元計測のための対応点分類手法”, コンピュータショナル・インテリジェンス研究会, PG0008/14/0000-0070, 神奈川, 2014 年 7 月 1 日-2 日

謝辞

本研究を遂行するにあたりご指導を賜りました東京工業大学大学院総合理工学研究科知能システム科学専攻教授廣田薫先生に心から感謝致します。日々熱心なご指導を頂戴しました。あっという間でしたが、廣田研究室で過ごした3年間はかけがえのないものです。3年間お世話になりました。

本研究の審査を引き受けて頂きました東京工業大学大学院総合理工学研究科知能システム科学専攻連携教授柴田崇徳先生，同准教授室伏俊明先生，東京工業大学像情報工学研究所教授長橋宏先生，東京工業大学精密工学研究所佐藤誠教授に深く感謝致します。1年という長期間に渡り審査をして頂きました。この場を借りて厚く御礼申し上げます。

東京工業大学大学院総合理工学研究科知能システム科学専攻助教董芳艶先生，廣田研究室秘書星野治美氏には大変お世話になりました。研究・生活面で様々なアドバイスを頂きました。お二方のご支援がなければ修了できなかったと思います。この場を借りて厚く御礼申し上げます。