

NAIST-IS-MT0851038

## 修士論文

# 測位精度を考慮したGPSの併用による 動画像からのカメラ位置・姿勢推定

条 秀行

2010年2月4日

奈良先端科学技術大学院大学  
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に  
修士(工学) 授与の要件として提出した修士論文である。

条 秀行

審査委員：

横矢 直和 教授 (主指導教員)

加藤 博一 教授 (副指導教員)

山澤 一誠 准教授 (副指導教員)

佐藤 智和 助教 (副指導教員)

# 測位精度を考慮した GPS の併用による 動画像からのカメラ位置・姿勢推定\*

桑 秀行

## 内容梗概

現実環境を撮影した動画像に対するカメラの位置・姿勢情報は三次元形状復元や自由視点画像生成，拡張現実感をはじめとする多くの分野への応用が可能である．しかし，動画像のみからカメラ位置・姿勢を推定する手法には推定結果に誤差が蓄積するという問題や，スケール情報が得られないという問題がある．このような問題に対し，動画像に加えて GPS を併用することで，広域屋外環境において低い人的コストで絶対位置情報を推定可能な手法が従来から提案されている．しかし，従来手法には，特に GPS の測位精度が低い場合などのように，短時間の観測において平均測位位置が真値から大きく外れる場合においてカメラ位置・姿勢の推定精度が大きく低下するという問題がある．これは，従来手法が GPS の測位誤差として平均値 0 の正規分布を仮定する一方で，実際の短時間における GPS の測位誤差の分布がこのような正規分布で適切に近似できないことに起因している．加えて，多くの手法は GPS の測位精度が GPS 電波の遮蔽や反射，衛星の配置等によって大きく変化することを考慮していない．これらの問題に対して，本論文では，GPS の測位誤差を単一の正規分布でモデル化する代わりに，GPS 受信機の真の位置が GPS の測位位置を中心とする，GPS の測位精度 (RTK-fix, RTK-float 等) に依存した一定の範囲内に存在することを仮定する．この仮定に基づくエネルギー関数を最小化することで，GPS の測位精度に応じて動画像からのカメラの位置・姿勢推定結果を補正する．これにより，GPS の測位精度が低い

\*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT0851038, 2010年2月4日.

場合にカメラ位置・姿勢の推定精度が低下することを防ぐ。実験では，屋外環境下で取得した動画像と GPS 測位値を用いてカメラ位置・姿勢推定を行い，従来手法と比較することで，提案手法の有効性を検証する。

キーワード

カメラ位置・姿勢推定，structure from motion，GPS

# Extrinsic Camera Parameter Estimation from Video Images Considering GPS Positioning Accuracy\*

Hideyuki Kume

## Abstract

Extrinsic camera parameters estimated from video images have been widely used for a number of different fields such as three-dimensional reconstruction, novel view generation and augmented reality. However, most conventional vision-based camera parameter estimation methods have some problems concerning accumulative errors and unknown scale. In order to cope with these problems, GPS coupled with video images are sometimes used to estimate absolute camera position in outdoor environments. However, the accuracy of estimated camera position largely depends on the accuracy of GPS because conventional methods assume that GPS position error is normally distributed. Additionally, most of the methods do not consider the change of GPS positioning accuracy depending on satellite positions and condition of the environment. In order to resolve this problem, instead of the normal distribution assumption for positioning error, this study assumes that true GPS position exists within a certain range whose size depends on the GPS positioning accuracy. Concretely, the proposed method estimates camera parameters by minimizing an energy function that is defined by using the reprojection error and the penalty term for GPS positioning. In

---

\*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT0851038, February 4, 2010.

experiments, the effectiveness of the proposed method has been demonstrated by comparing it with conventional methods.

**Keywords:**

extrinsic camera parameter estimation, structure from motion, GPS

# 目次

|                                  |    |
|----------------------------------|----|
| 1. はじめに                          | 1  |
| 2. カメラ位置・姿勢推定に関する従来研究と本研究の位置付け   | 4  |
| 2.1 センサを用いたカメラ位置・姿勢推定            | 4  |
| 2.2 動画像を用いたカメラ位置・姿勢推定            | 5  |
| 2.2.1 事前知識を用いない手法                | 6  |
| 2.2.2 事前知識を用いる手法                 | 7  |
| 2.3 動画像とセンサの併用によるカメラ位置・姿勢推定      | 9  |
| 2.3.1 事前知識を用いない動画像からの推定手法とセンサの併用 | 9  |
| 2.3.2 事前知識を用いる動画像からの推定手法とセンサの併用  | 10 |
| 2.4 本研究の位置付けと方針                  | 11 |
| 3. GPS の測位精度を考慮したカメラ位置・姿勢推定      | 13 |
| 3.1 カメラ位置・姿勢推定手法の概要              | 13 |
| 3.2 座標系の定義                       | 15 |
| 3.2.1 カメラ位置・姿勢の定義                | 15 |
| 3.2.2 GPS 測地座標系の定義               | 17 |
| 3.3 GPS の測位精度を考慮したエネルギー関数の定義     | 17 |
| 3.3.1 再投影誤差                      | 18 |
| 3.3.2 GPS 測位位置に関するペナルティ          | 19 |
| 3.4 動画像と GPS の併用によるカメラ位置・姿勢推定    | 21 |
| 3.4.1 特徴点の追跡によるカメラ位置・姿勢の推定       | 22 |
| 3.4.2 特徴点の三次元位置の推定と特徴点の追加・削除     | 23 |
| 3.4.3 GPS 測位値を用いた最適化             | 24 |
| 4. 実験                            | 26 |
| 4.1 定点観測による GPS 測位値の誤差分析         | 26 |
| 4.1.1 実験条件                       | 26 |
| 4.1.2 GPS の測位精度に依存した誤差範囲の決定      | 28 |

|       |                                    |    |
|-------|------------------------------------|----|
| 4.2   | 動画像とGPS測位値を用いたカメラ位置・姿勢推定 . . . . . | 30 |
| 4.2.1 | 実験条件 . . . . .                     | 30 |
| 4.2.2 | カメラ位置推定精度の定量的評価 . . . . .          | 33 |
| 5.    | まとめ                                | 37 |
|       | 謝辞                                 | 38 |
|       | 参考文献                               | 39 |
|       | 付録                                 | 44 |
| A.    | GPSから得られる情報の特性                     | 44 |
| A.1   | GPSの種類と測位精度 . . . . .              | 44 |
| A.2   | GPSの測位誤差の要因と特性 . . . . .           | 44 |

## 目 次

|    |                             |    |
|----|-----------------------------|----|
| 1  | カメラ位置・姿勢推定手法の分類             | 4  |
| 2  | Kalkusch らの AR システム [22]    | 8  |
| 3  | 提案手法の処理の流れ                  | 14 |
| 4  | カメラ座標系と世界座標系の関係             | 16 |
| 5  | 再投影誤差                       | 18 |
| 6  | GPS 測位位置に関するペナルティ           | 20 |
| 7  | $n$ の違いによるペナルティ値の変化         | 21 |
| 8  | 最適化の範囲                      | 24 |
| 9  | GPS 測位値の定点観測の様子             | 27 |
| 10 | 定点観測における測位位置の分布 (RTK-fix)   | 29 |
| 11 | 定点観測における測位位置の分布 (RTK-float) | 29 |
| 12 | 入力画像の一部                     | 31 |
| 13 | カメラと GPS 受信機の位置関係           | 32 |
| 14 | 推定されたカメラ位置・姿勢と特徴点の三次元位置     | 34 |
| 15 | 推定された GPS 受信機の位置            | 35 |
| 16 | フレームごとの位置誤差                 | 35 |
| 17 | GPS の種類と測位精度                | 44 |

## 表 目 次

|   |                    |    |
|---|--------------------|----|
| 1 | 従来カメラ位置・姿勢推定手法の特徴  | 11 |
| 2 | カメラ内部パラメータ         | 15 |
| 3 | RTK-GPS 受信機のカタログ仕様 | 27 |
| 4 | 測位精度ごとの誤差の最大値 [mm] | 28 |
| 5 | 回転楕円体のパラメータ [mm]   | 33 |
| 6 | 位置誤差の比較 [mm].      | 36 |

## 1. はじめに

現実環境を撮影した動画像に対するカメラの位置・姿勢情報は三次元形状復元や自由視点画像生成，拡張現実感をはじめとする多くの分野への応用が可能である．近年これらの分野では，広域な屋外環境への手法の適用が試みられ [1-3]，広域な屋外環境を移動しながら撮影された動画像に対するカメラの位置・姿勢を高精度に推定することが求められている．従来から，カメラの位置・姿勢を推定する手法として GPS やジャイロセンサなどのセンサを用いた手法 [4-12]，撮影された動画像を用いた手法 [13-29]，動画像と GPS やジャイロセンサなどのセンサを併用したハイブリッド手法 [1, 30-37] が提案されている．

センサを用いた手法 [4-12] は，GPS などの絶対位置を取得するセンサと加速度計やジャイロセンサなどの相対位置を取得するセンサを組み合わせることで，高速でロバストなカメラ位置・姿勢の推定を実現している．しかし，センサとカメラのキャリブレーション誤差やセンサの計測誤差の影響により，一般に動画像を用いた手法と比較して高精度なカメラ位置・姿勢の推定が難しいという問題がある．

動画像を用いた手法 [13-29] は，撮影された動画像そのものを用いることで高精度なカメラ位置・姿勢の推定を実現している．しかし，動画像のみを用いる手法 [13-20] には推定されるカメラ位置・姿勢に誤差が蓄積するという問題があり，特に広域屋外環境でのカメラ位置・姿勢を必要とするアプリケーションにおいて大きな問題となる．この問題を解決するため，動画像に加えて三次元位置が既知の基準点や CAD モデルなど，環境に関する事前知識を併用することでカメラの絶対位置・姿勢を推定する手法が提案されている [28, 29]．しかし，広域屋外環境での利用においては事前知識の準備に多くの人的コストがかかるという問題がある．

動画像とセンサを併用したハイブリッド手法 [1, 30-37] は，事前知識を用いない動画像からの推定手法とセンサを併用する手法 [1, 30-33] と，事前知識を用いる動画像からの推定手法とセンサを併用する手法 [34-37] に大別できる．特に前者は，主にセンサとして GPS を用いることにより，環境に対する事前知識なしに絶対的なカメラ位置・姿勢を推定することができるため，広域屋外環境でのカメラ位置・

姿勢の推定に適している。これら動画像と GPS を併用する手法は、情報の融合にカルマンフィルタを用いる手法 [1, 30, 31] とバンドル調整 (bundle adjustment) を用いる手法 [32, 33] に大別できる。

カルマンフィルタを用いる手法 [1, 30, 31] は、動画像と GPS の情報を高速に融合できるため、拡張現実感などのリアルタイムアプリケーションに適している。しかし、過去の状態と現在の観測から逐次的に現在の状態を求めるというフィルタの設計から、全体最適化を行い精度を高めることは難しい。一方、バンドル調整を用いる手法 [32, 33] では、動画像からのカメラ位置・姿勢推定において一般的に用いられるバンドル調整の誤差関数に GPS の測位位置に関する誤差の項を追加することで、動画像と GPS の情報を融合している。この手法は、オフライン処理を前提としているため、リアルタイムアプリケーションへの適用は難しいが、誤差関数を最小化することにより動画像全体でカメラ位置・姿勢を最適化することができる。しかし、従来手法では、特に GPS の測位精度が低い場合などのように、短時間の観測において平均測位位置が真値から大きく外れる場合においてカメラ位置・姿勢の推定精度が大きく低下するという問題がある。これは、従来手法が GPS の測位誤差として平均値 0 の正規分布を仮定する一方で、実際の短時間における GPS の測位誤差の分布がこのような正規分布で適切に近似できないことに起因している。

本研究では、三次元形状復元や自由視点画像生成などオフラインで高精度なカメラ位置・姿勢を必要とするアプリケーションへの適用を前提とし、広範囲な環境を移動しながら撮影した動画像に対するカメラの位置・姿勢推定において、比較的高精度な絶対位置情報を推定可能なバンドル調整による GPS と動画像情報の融合手法 [32, 33] を基礎とし、GPS の測位精度が低い場合にカメラ位置・姿勢の推定精度が低下するという問題を解決する。具体的には、GPS の測位誤差を単一の正規分布でモデル化する代わりに、GPS 受信機の真の位置が GPS の測位位置を中心とする、GPS の測位精度 (RTK-fix, RTK-float 等) に依存した一定の範囲内に存在することを仮定し、この仮定に基づくエネルギー関数を最小化することで、GPS の測位精度に応じて動画像からのカメラの位置・姿勢推定結果を補正する。

以下，2章では，カメラ位置・姿勢推定の従来研究と本研究の位置付けについて述べる．3章では，GPSの測位精度を考慮したエネルギー関数の定義とカメラ位置・姿勢推定手法について詳述する．4章では，屋外環境を移動しながら撮影した動画像を用いてカメラの位置・姿勢推定を行い，従来手法と比較することで精度を検証する．最後に5章でまとめと今後の課題について述べる．

## 2. カメラ位置・姿勢推定に関する従来研究と本研究の位置付け

本章では，カメラ位置・姿勢推定の従来研究を，図1に示すように，センサを用いた手法 [4-12]，動画像を用いた手法 [13-29]，動画像とセンサを併用したハイブリッド手法 [1,30-37] に分類し，それぞれの手法の特徴と広域屋外環境への手法適用の可能性について述べる．次に，本研究の位置付けと方針について述べる．

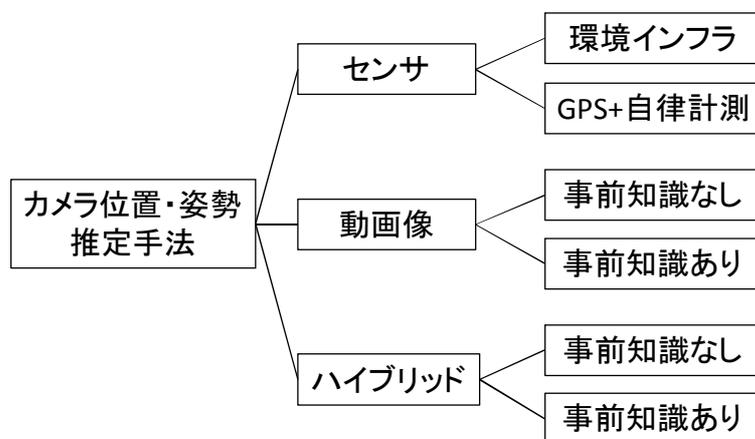


図1 カメラ位置・姿勢推定手法の分類

### 2.1 センサを用いたカメラ位置・姿勢推定

センサを用いたカメラ位置・姿勢推定手法には，環境内に配置された絶対位置・姿勢の計測が可能なインフラを用いる手法 [4-6] と，絶対位置を取得可能なGPSと相対位置・姿勢を取得可能なセンサを組み合わせる手法 [7-12] がある．相対位置・姿勢を取得可能なセンサとしては加速度計やジャイロセンサが用いられる．

インフラを用いる手法 [4-6] は，環境中に設置された超音波センサや赤外線ビーコン，RFID，磁気センサなどを用いてカメラの絶対位置・姿勢を推定する．Newmanら [4] は，超音波センサを建物内に多数設置することにより，広範囲な屋内環境におけるカメラ位置・姿勢の推定を実現している．ただし，この手法で

はセンサの設置に多くの人的コストがかかるという問題がある．これに対して，Tenmoku ら [6] は赤外線ビーコンや RFID といったインフラに加えて，歩数計を用いてインフラ間における相対的な運動を推定することにより，より疎な配置のインフラでカメラ位置・姿勢推定を実現している．しかし，歩数計による位置・姿勢の計測は誤差が蓄積するため，広範囲な環境でカメラの位置・姿勢を推定するためには依然として大量のセンサを設置することが必要となり，設置・維持に関するコストが問題となる．

GPS を用いる手法 [7–12] は，GPS から得られる絶対位置と加速度計やジャイロセンサなどから得られる相対位置・姿勢情報を用いることにより，屋外環境において特別なインフラを設置することなくカメラの絶対位置・姿勢を推定することができる．このような手法の最も基礎的な方法として，GPS とジャイロセンサから得られる値を直接，カメラ位置・姿勢とする手法 [7–9] が用いられている．しかし，一般的な GPS は 1 秒間隔でしか位置情報を取得できないため，動画像のすべてのフレームに対する位置情報を算出することは難しいという問題がある．これに対して，GPS に加えて慣性航法装置やオドメトリ，歩行動作に基づくデットレコニングといった相対位置・姿勢を計測する手法を併用することにより，GPS の位置情報が取得できないフレームにおいてもカメラの位置・姿勢を推定する手法が提案されている [10–12]．しかし，GPS の測位精度は GPS 電波の遮蔽や反射，衛星の配置等によって大きく変化するため，常に高精度の測位を行うことは難しく，GPS の測位精度が低い場合にはカメラ位置の推定精度が低下するという問題がある．

これらセンサのみを用いる手法は，リアルタイムにカメラ位置・姿勢を推定できるが，三次元形状復元や拡張現実感など画像を用いたアプリケーションでの利用を考えた場合，センサとカメラのキャリブレーション誤差や計測誤差に起因して，画素単位での位置合わせが難しいという問題がある．

## 2.2 動画像を用いたカメラ位置・姿勢推定

動画像を用いたカメラ位置・姿勢推定手法は，画像中の特徴点やエッジを追跡，または事前知識と対応付けることにより，カメラの位置・姿勢を推定する．この手

法では撮影された動画像そのものを用いることにより，カメラ位置・姿勢の推定精度を高めやすいという利点がある．以下では，動画像を用いたカメラ位置・姿勢推定手法を，事前知識を用いない手法 [13–20] と，事前知識を用いる手法 [21–29] に大別し，それぞれの手法について詳述する．

### 2.2.1 事前知識を用いない手法

事前知識を用いず，動画像のみからカメラ位置・姿勢を推定する手法は Structure from Motion (SfM) や Simultaneous Localization And Mapping (SLAM) と呼ばれ，入力画像中で特徴点を追跡することにより，カメラ位置・姿勢と特徴点の三次元位置を同時に推定する．これらの代表的な手法に，カメラモデルを線形近似し，線形解法によってカメラ位置・姿勢と特徴点の三次元位置を同時に推定する因子分解法 [13] がある．この手法は，全てのデータが揃ってからの全体処理を前提としており，拡張現実感などのリアルタイムアプリケーションに適用することは難しいが，高速かつ数値計算上安定に解を得ることができるため，オフラインでカメラ位置・姿勢を必要とする三次元形状復元などの分野で従来広く用いられてきた．しかし，カメラモデルを線形近似しているため，カメラ位置・姿勢の推定精度に限界がある．また，動画像の全てのフレームにおいて撮影されている特徴点を用いることを前提としているため，特徴点が連続して観測できない動画像に対するカメラ位置・姿勢の推定は困難である．これらの問題に対して，カメラモデルを透視投影モデルに拡張する手法 [14, 15] や追跡に失敗した特徴点の投影位置を射影復元を用いて補間する手法 [16] が提案されている．しかし，因子分解法に基づく手法において，多くの特徴点が追加・削除されるような動画像を扱うことは原理的に難しい．

一方，カメラモデルとして透視投影を仮定し，特徴点の三次元位置を画像上に投影した座標と，画像上で検出された特徴点の座標の距離の二乗和 (再投影誤差) を最小化することによりカメラ位置・姿勢を推定する手法が提案されている [17–19] ．このような手法はバンドル調整 (bundle adjustment) と呼ばれ，特徴点の出現や消失にも対応可能であり，透視投影モデルを仮定しているため精度の高い推定ができる．従来，バンドル調整はオフラインで高精度なカメラ位置・姿勢を必要と

する三次元形状復元などにおいて、最終的な処理として動画像全体を対象とした最適化処理に利用されてきた [17,18]。これに対し、近年、拡張現実感の分野において、フレーム毎の逐次的な処理でバンドル調整を狭区間を対象として実行することにより、精度の高いカメラ位置・姿勢推定をリアルタイムで実現する手法が提案されている [19]。また、ロボティクスの分野においては、バンドル調整と同様に再投影誤差を指標とし、拡張カルマンフィルタを用いてリアルタイムにカメラ位置・姿勢を推定する手法 [20] が提案されている。

これら動画像のみからカメラ位置・姿勢を推定する手法は、カメラ以外のセンサや事前知識を必要としないため、手軽にカメラ位置・姿勢を推定できるという利点があるが、推定されるカメラ位置・姿勢に誤差が蓄積するという問題があり、特に広域な屋外環境下でのカメラ位置・姿勢を必要とするアプリケーションにおいて大きな問題となる。

### 2.2.2 事前知識を用いる手法

事前知識を用いる手法は、入力画像中の特徴点やエッジを環境に関する事前知識と対応付けることにより、カメラの絶対位置・姿勢を推定する。事前知識としてはマーカ [21–23] や三次元 CAD モデル [24,25]、自然特徴点ランドマーク [26,27] などが用いられる。

マーカを用いる手法 [21–23] は、ARToolkit [21] などのように画像中で検出・識別することが比較的容易なマーカを用い、マーカ上の点の三次元座標と画像上の二次元座標を対応付けることで、マーカを用いて定義される座標系に対するカメラの位置・姿勢を推定する。この手法は、手軽にリアルタイムでのカメラ位置・姿勢推定を実現することができるため、広く用いられてきた。Kalkuschら [22] は、図2のように屋内において多数のマーカを設置することにより、カメラ位置・姿勢推定を行っている。しかし、マーカを用いる手法は環境内に多数のマーカを設置する必要があり、景観を損ねるという問題がある。これに対して中里ら [23] は再帰性反射材を用いた不可視マーカを用いることで、この問題を解決しているが、これを屋外で利用することは難しい。

三次元 CAD モデルを用いる手法 [24, 25] や、自然特徴点ランドマークを用い



図 2 Kalkusch らの AR システム [22]

る手法 [26,27] は、事前に作成された環境の三次元 CAD モデルや三次元位置が既知の特徴点を用いることにより、環境に手を加えることなくカメラ位置・姿勢を推定することができる。Drummond ら [24] は三次元 CAD モデルを入力画像中に投影し、画像中のエッジとの距離を最小化することでカメラ位置・姿勢を推定している。武富ら [27] は、全方位型マルチカメラを用いて撮影した動画像に対して SfM を適用することにより作成された自然特徴点ランドマークを用いて、カメラ位置・姿勢の推定を行っている。これら環境に関する事前知識を用いる手法は、三次元 CAD モデルや自然特徴点ランドマークといった事前知識を用いることで、比較的広域な環境でリアルタイムにカメラ位置・姿勢の推定を行なえるが、事前知識の作成に多くの人的コストがかかるという問題がある。

このような事前知識の準備にかかる人的コストを低減可能なアプローチとして、環境中の一部の領域に対する三次元 CAD モデルや自然特徴点ランドマークと SfM や SLAM といった事前知識を用いない推定手法を併用する手法 [28,29] が提案されている。これらの手法では SfM や SLAM における蓄積誤差の問題を事前知識を用いることで解決している。また、三次元 CAD モデルや自然特徴点ラ

ランドマークが写っていないフレームに対しても SfM や SLAM を用いることでカメラ位置・姿勢を推定できる．従って，環境中の一部に対してのみ事前知識を準備すればよく，準備にかかる人的コストを削減できる．ただし，長時間モデルやランドマークが写らないと誤差が蓄積するという問題が生じる．

## 2.3 動画像とセンサの併用によるカメラ位置・姿勢推定

動画像とセンサの併用によるカメラ位置・姿勢推定手法 [1, 30–37] は，センサを用いた推定手法と動画像を用いた推定手法を併用することにより，それぞれの短所を補い，高精度なカメラ位置・姿勢の推定を実現する．動画像とセンサの併用によるカメラ位置・姿勢推定手法は，事前知識を用いない動画像からの推定手法とセンサを併用する手法 [1, 30–33] と，事前知識を用いる動画像からの推定手法とセンサを併用する手法 [34–37] に大別できる．

### 2.3.1 事前知識を用いない動画像からの推定手法とセンサの併用

事前知識を用いない動画像からの推定手法とセンサを併用する手法 [1, 30–33] は，2.2.1 項で述べた SfM や SLAM における蓄積誤差の問題を，センサから得られる絶対位置を用いることで解決している．一般に広域な屋外環境での利用においては，絶対位置を取得可能なセンサとして GPS が用いられる．これら動画像と GPS を併用する手法は，情報の融合にカルマンフィルタを用いる手法 [1, 30, 31] とバンドル調整を用いる手法 [32, 33] に大別できる．

カルマンフィルタを用いる手法 [1, 30, 31] は動画像と GPS の情報を高速に融合することができるため，拡張現実感などのリアルタイムアプリケーションに適している．Agrawal ら [30] は，動画像から推定されるカメラ位置と GPS の測位位置をカルマンフィルタを用いて融合している．Pollefeys ら [1] は，拡張カルマンフィルタを用いた動画像からの推定手法を基礎とし，GPS，慣性航法装置からの観測をカルマンフィルタに組み込むことにより，より密な情報の融合を実現している．これらカルマンフィルタに基づく手法は，過去の状態と現在の観測から現

在の状態を逐次的に推定する．このため，入力全体で最適化を行い精度を高めることは難しい．

一方，バンドル調整を用いる手法 [32,33] は，動画像からのカメラ位置・姿勢推定において用いられるバンドル調整の誤差関数に，GPS の測位位置に関する誤差の項を追加することで，動画像と GPS の情報を融合する．横地ら [32] は，GPS の測位誤差が単一の正規分布で近似できることを前提に，定義した誤差関数を最小化することで，カメラ位置・姿勢を推定している．この手法は，オフライン処理を前提としているため，リアルタイムアプリケーションへの適用は難しいが，誤差関数を最小化することにより動画像全体でカメラ位置・姿勢を最適化することができる．しかし，実際には，GPS の測位精度は GPS 電波の遮蔽や反射，衛星の配置等によって大きく変化し，常に高精度の測位を行うことは困難である．このため，GPS の測位精度が低い場合にカメラ位置・姿勢の推定精度が大きく低下するという問題がある．これに対して，Anai ら [33] は GPS 受信機から得られる精度情報に応じて正規分布の分散を変化させることにより，測位精度の変化を考慮する手法を提案している．しかし，GPS の測位精度が低い場合など，短時間の観測において平均測位位置が真値から大きく外れる場合において，推定精度が低下するという問題は依然として解消されていない．これは，GPS の測位誤差として平均値 0 の正規分布を仮定する一方で，実際の短時間における GPS の測位誤差の分布がこのような正規分布で適切に近似できないことに起因している．

### 2.3.2 事前知識を用いる動画像からの推定手法とセンサの併用

事前知識を用いる動画像からの推定手法とセンサを併用する手法 [34–37] では，動画像による絶対位置・姿勢の推定と加速度計やジャイロセンサなどのセンサによる相対位置・姿勢の計測を組み合わせることにより，カメラ位置・姿勢推定のロバスト性の向上を実現している．Jiang ら [34] や Reitmayr ら [35] はジャイロセンサによる姿勢の計測結果と環境の三次元モデルを用いた動画像からの推定を拡張カルマンフィルタを用いて融合することにより，カメラの回転に対してロバストに位置・姿勢の推定を行う手法を提案している．Klein ら [36] や松田ら [37] は画像中の特徴を環境の三次元モデルや自然特徴ランドマークと対応付ける際に

ジャイロセンサから得られる角速度情報からモーションブレンダーを考慮することにより、モーションブレンダーが生じた画像に対するカメラ位置・姿勢推定のロバスト性を向上させている。これら事前知識を用いる手法は、2.2.2項で述べた動画像と事前知識を用いる手法と同様に、リアルタイムでカメラ位置・姿勢を推定することができるが、広域屋外環境での利用においては事前知識の準備に多くの人的コストがかかるという問題がある。

## 2.4 本研究の位置付けと方針

前節までで述べたように、従来、カメラの位置・姿勢を推定する手法は数多く提案されている。環境インフラや環境に関する事前知識を用いる手法は、蓄積誤差のないカメラ位置・姿勢を推定できるが、インフラの設置や事前知識の準備にかかる人的コストが高く、広範囲な環境への適用は難しい。従って、広範囲な環境におけるカメラ位置・姿勢の推定には、特別なインフラの設置や事前知識を用いない手法が適している。このような手法として (1)GPS と自律計測を用いる、(2) 動画像のみを用いる、(3) 動画像と GPS を用いる、という 3 つのアプローチが存在する。これらの手法の特徴を表 1 にまとめる。

事前知識を用いない動画像からの推定手法は、センサなど動画像以外を必要とせず、動画像のみからカメラの位置・姿勢を推定できるが、推定結果に誤差が蓄積するという問題がある。これに対して、GPS と自律計測を併用する手法や動画像と GPS を併用する手法では、蓄積誤差のないカメラ位置・姿勢を推定することができる。ただし、センサのみを用いる手法は、リアルタイムにカメラ位置・

表 1 従来のカメラ位置・姿勢推定手法の特徴

| カメラ位置・姿勢の推定手法   | 推定精度     | 蓄積誤差          | ロバスト性 | 全体最適化 | リアルタイム性 | 適用分野※       |
|-----------------|----------|---------------|-------|-------|---------|-------------|
| センサ(GPS+自律計測)   | ×:低い     | ○:なし          | ○:高い  | ×:難しい | ○:あり    | AR RN       |
| 動画像             | 全体処理     | ○:高い          | ×:あり  | ×:低い  | ○:可能    | ×:なし 3DM FV |
|                 | 逐次処理     | ○:高い          | ×:あり  | ×:低い  | ×:難しい   | ○:あり AR RN  |
| ハイブリッド(動画像+GPS) | カルマンフィルタ | △:GPSの測位精度に依存 | ○:なし  | ○:高い  | ×:難しい   | ○:あり RN 3DM |
|                 | バンドル調整   | △:GPSの測位精度に依存 | ○:なし  | ○:高い  | ○:可能    | ×:なし 3DM FV |

※ AR…拡張現実感 RN…ロボットナビゲーション 3DM…三次元形状復元 FV…自由視点画像生成

姿勢を推定できる反面，センサとカメラのキャリブレーション誤差や計測誤差に起因して，カメラ位置・姿勢の高精度な推定が難しいという問題がある．これらの理由により，広範囲な環境を移動しながら撮影した動画像に対するカメラの位置・姿勢推定においては，事前知識を用いない動画像からの推定手法と GPS を併用する手法が適している．しかし，従来手法には GPS の測位精度が低い場合にカメラ位置・姿勢の推定精度が低下するという問題があった．

そこで，本研究では，三次元形状復元や自由視点画像生成などオフラインで高精度なカメラ位置・姿勢を必要とするアプリケーションへの適用を前提とし，事前知識を用いない動画像からの推定手法と GPS を併用する従来手法の中でも，全体最適化を行うことにより高精度な絶対位置情報を推定可能なバンドル調整による GPS と動画像情報の融合手法 [32,33] を基礎とし，GPS の測位精度が低い場合にカメラ位置・姿勢の推定精度が低下するという問題を解決する．提案手法では，GPS の測位誤差を単一の正規分布でモデル化する代わりに，GPS 受信機の真の位置が GPS の測位位置を中心とする，GPS の測位精度 (RTK-fix, RTK-float 等) に依存した一定の範囲内に存在することを仮定し，この仮定に基づくエネルギー関数を最小化することで，GPS の測位精度に応じて動画像からのカメラの位置・姿勢推定結果を補正する．

### 3. GPS の測位精度を考慮したカメラ位置・姿勢推定

#### 3.1 カメラ位置・姿勢推定手法の概要

本研究では、カメラと GPS 受信機の相対位置関係を固定して撮影した動画像と GPS の測位値から、GPS 測地座標系におけるカメラの位置・姿勢を推定する。ただし、カメラと GPS 受信機の相対位置関係は既知であり、動画像と GPS の測位値は同期しているものとする。

提案手法はバンドル調整による GPS と動画像情報の融合手法 [32,33] を基礎とし、これらの手法と同様に、特徴点の再投影誤差と GPS の測位位置に関するエネルギーを同時に最小化することにより、動画像からのカメラ位置・姿勢の推定結果を補正する。提案手法の処理の流れを図 3 に示す。まず、フレーム毎の処理として、特徴点の追跡によるカメラ位置・姿勢推定 (A)、特徴点の三次元位置の推定と特徴点の追加・削除 (B) を行う。ここで、一定フレーム間隔  $k$  毎に、GPS の測位値を用いた狭区間最適化 (C) を行い、カメラの位置・姿勢と特徴点の三次元位置を最適化する。これにより、逐次的に蓄積誤差を抑えたカメラの位置・姿勢を得る。これらの処理 (A) ~ (C) を初期フレームから最終フレームまで繰り返すことで、逐次的にカメラ位置・姿勢および特徴点の三次元位置の初期値を得た上で、最後に動画像全体での最適化処理を行う (D)。以下では、まず、本研究で用いる座標系を定義し、新たに提案する GPS の測位精度を考慮したエネルギー関数について述べる。次に、処理 (A) ~ (D) について詳述する。

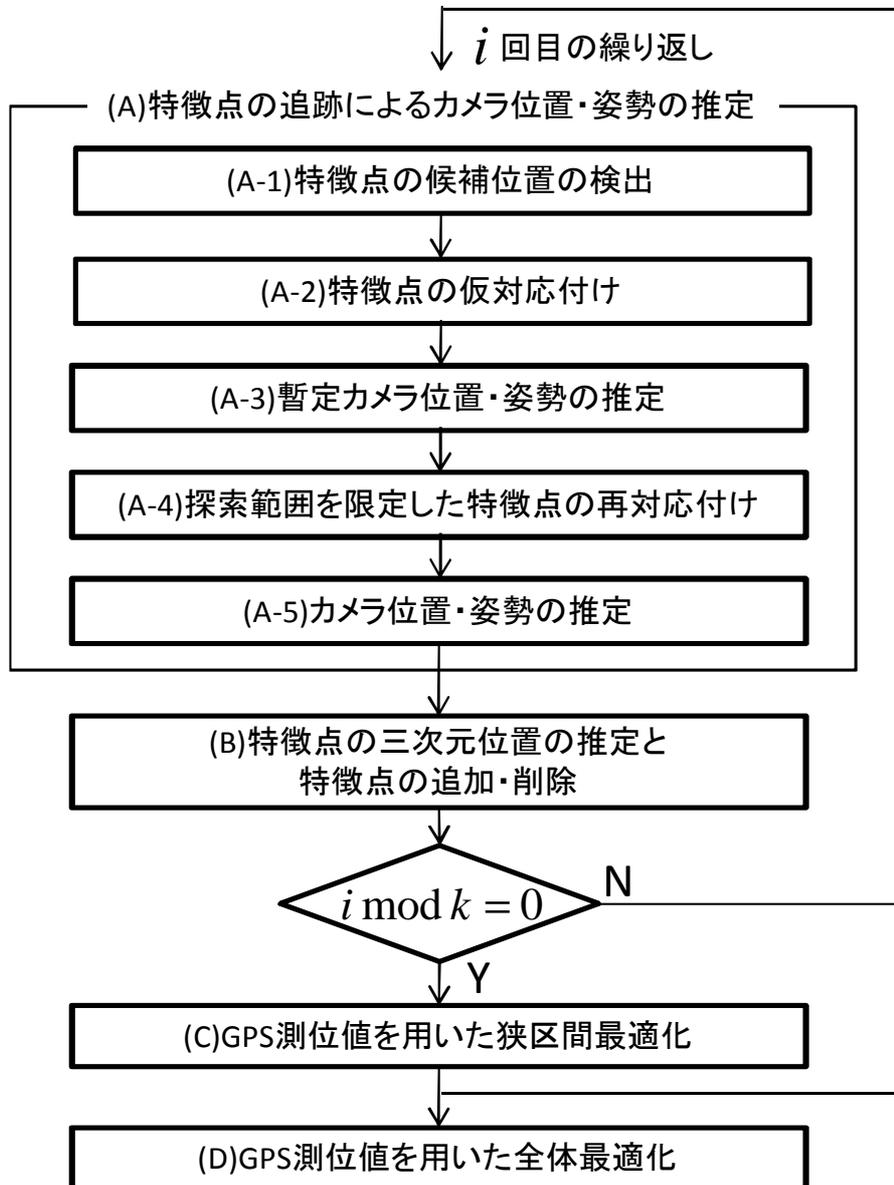


図 3 提案手法の処理の流れ

## 3.2 座標系の定義

### 3.2.1 カメラ位置・姿勢の定義

本論文で用いるカメラ内部パラメータを表 2 に示す．これらのカメラ内部パラメータは，同一の動画像中では変化しないものとし，Tsai の手法 [38] などによって，あらかじめ計測しておくものとする．図 4 に示すように，ここでは第  $i$  フレームにおけるカメラの位置・姿勢を，世界座標系からカメラ座標系への変換行列  $M_i$  で表す．すなわち，本論文ではカメラの姿勢パラメータとして，カメラ座標系の 3 軸に対する回転角  $(r_1, r_2, r_3)$  と，カメラの位置パラメータとして，カメラ座標系における平行移動成分  $(t_1, t_2, t_3)$  を用い，変換行列  $M$  を以下のように定義する．

$$M = \begin{pmatrix} c_1 c_3 + s_1 s_2 s_3 & s_1 c_2 & -c_1 s_3 + s_1 s_2 c_3 & t_1 \\ -s_1 c_3 + c_1 s_2 s_3 & c_1 c_2 & s_1 s_3 + c_1 s_2 c_3 & t_2 \\ c_2 s_3 & -s_2 & c_2 c_3 & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

ただし，

$$\begin{aligned} s_1 &= \sin r_1, & s_2 &= \sin r_2, & s_3 &= \sin r_3 \\ c_1 &= \cos r_1, & c_2 &= \cos r_2, & c_3 &= \cos r_3 \end{aligned} \quad (2)$$

ここで，世界座標系における特徴点  $j$  の三次元位置を  $p_j = (X_j, Y_j, Z_j)^T$ ，第  $i$  フレームのカメラ座標系における特徴点  $j$  の三次元位置を  $c_{ij} = (X'_{ij}, Y'_{ij}, Z'_{ij})^T$

表 2 カメラ内部パラメータ

|  |                      |
|--|----------------------|
| 焦点距離 (mm)                                    | F                    |
| CCD のサイズ (mm)                                | $ccd_x \times ccd_y$ |
| 画像のサイズ (画素)                                  | $scr_x \times scr_y$ |
| 投影中心座標 (画素)                                  | $(C_x, C_y)$         |
| スケール係数                                       | $s_x$                |
| レンズ歪み係数 ( $\text{mm}^{-2}, \text{mm}^{-4}$ ) | $\kappa_1, \kappa_2$ |

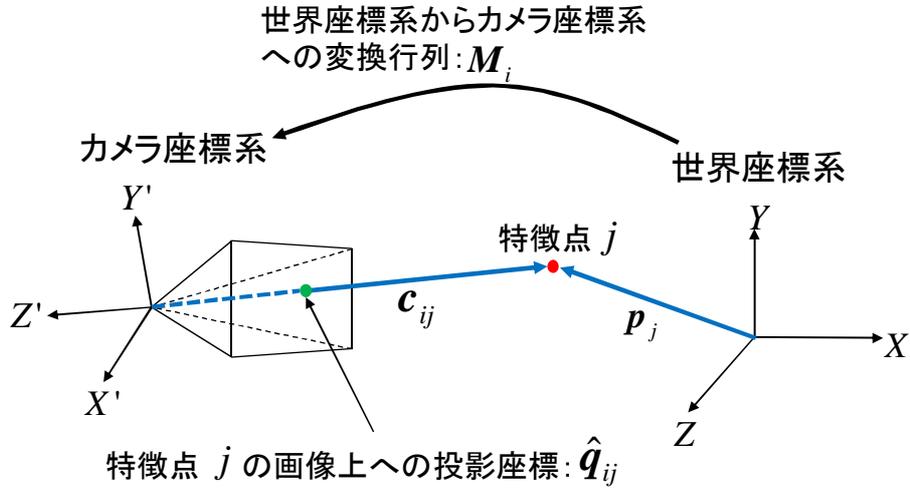


図 4 カメラ座標系と世界座標系の関係

とすれば，世界座標系とカメラ座標系の変換は以下のように記述される．

$$\begin{pmatrix} X'_{ij} \\ Y'_{ij} \\ Z'_{ij} \\ 1 \end{pmatrix} = M_i \begin{pmatrix} X_j \\ Y_j \\ Z_j \\ 1 \end{pmatrix} \quad (3)$$

レンズ歪みを考慮しなければ，特徴点  $j$  の三次元座標  $c_{ij}$  の画像上への投影座標  $q_{ij} = (u, v)$  は，以下のように定義される．

$$q_{ij} = (u, v) = \left( -\frac{scr_x}{ccd_x} \frac{X'_{ij}}{Z'_{ij}} F s_x + C_x, -\frac{scr_y}{ccd_y} \frac{Y'_{ij}}{Z'_{ij}} F + C_y \right) \quad (4)$$

実際には，レンズ歪みの影響があるため，入力画像上において  $(u, v)$  は  $(\hat{u}, \hat{v})$  に投影される．本手法では，以下の式を用いて，特徴点の入力画像上での座標  $(\hat{u}, \hat{v})$  をレンズ歪みを取り除いた座標  $(u, v)$  に変換して用いる [38]．

$$(u, v) = ((1 + \kappa_1 \rho^2 + \kappa_2 \rho^4)(\hat{u} - C_x) + C_x, (1 + \kappa_1 \rho^2 + \kappa_2 \rho^4)(\hat{v} - C_y) + C_y) \quad (5)$$

$$\rho = \sqrt{\left( \frac{ccd_x}{scr_x s_x} (\hat{u} - C_x) \right)^2 + \left( \frac{ccd_y}{scr_y} (\hat{v} - C_y) \right)^2} \quad (6)$$

### 3.2.2 GPS 測地座標系の定義

GPS から得られる位置情報は緯度，経度，高度である．しかし，これらを座標系の各軸とした場合，基底ベクトルが直交せず，式 (3) のような線形変換で座標変換を扱えないという問題が生じる．また，緯度によって経度 1 秒の長さが大きく異なるため，スケール等の取り扱いも煩雑となる．そこで，本論文では，対象となるシーン中における任意の点を世界座標系の原点とする局所座標系に，緯度，経度，高度情報を変換して扱い，これを GPS 測地座標系と呼ぶ．GPS 測地座標系における，第  $i$  フレームの GPS 測位位置  $g_i$  は，GPS から得られる緯度，経度，高度を  $(B_i, L_i, H_i)$ ，GPS 測地座標系原点の緯度，経度，高度を  $(B_o, L_o, H_o)$  とすると，以下の式により算出される．

$$g_i = \begin{pmatrix} \sin B_i \cos L_i & \sin B_i \sin L_i & -\cos B_i \\ -\sin L_i & \cos L_i & 0 \\ \cos B_i \cos L_i & \cos B_i \sin L_i & \sin B_i \end{pmatrix} (\mathbf{x}_i - \mathbf{x}_o) \quad (7)$$

$$\mathbf{x}_k = \begin{pmatrix} (R_k + H_k) \cos B_k \cos L_k \\ (R_k + H_k) \cos B_k \sin L_k \\ (R_k(1 - \epsilon^2) + H_k) \sin B_k \end{pmatrix} \quad (8)$$

$$R_k = \frac{a}{\sqrt{1 - \epsilon^2 \sin^2 B_k}} \quad (9)$$

$$\epsilon^2 = f(2 - f) \quad (10)$$

ただし， $a$  は地球の長軸半径 (6, 378, 137m)， $f$  は地球の偏平率 (1/298.257223563) である．

### 3.3 GPS の測位精度を考慮したエネルギー関数の定義

本研究では，動画像を対象とした三次元復元の評価尺度として広く用いられている再投影誤差と，GPS に関して本研究で新たに定義するエネルギー項の和を最小化することで，カメラ位置・姿勢および特徴点の三次元位置を推定する．具体的には，エネルギー関数  $E$  を， $i$  フレームにおける再投影誤差  $\Phi_i$  と GPS 測位

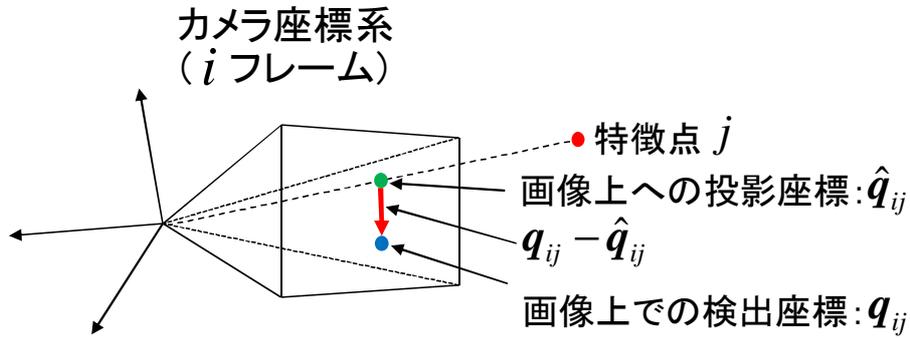


図 5 再投影誤差

位置に関するペナルティ項  $\Psi_i$  を用いて以下のように定義する .

$$E = \sum_{i \in F} \Phi_i + \omega \sum_{i \in F_g} \Psi_i \quad (11)$$

ただし ,  $\omega$  は重み係数 ,  $F$  は動画像のフレームの集合 ,  $F_g$  は GPS 測位時刻に対応する動画像のフレームの集合を表す . 以下 , 再投影誤差に関するエネルギー  $\Phi_i$  と GPS 測位位置に関するペナルティ項  $\Psi_i$  について詳述する .

### 3.3.1 再投影誤差

再投影誤差は , 図 5 に示すように , 特徴点の三次元位置を画像上に投影した座標と , 画像上で検出された特徴点の座標の距離の二乗和であり , 動画像を対象としたバンドル調整において一般的に用いられている . 本研究では , 再投影誤差に関するエネルギー  $\Phi_i$  を以下のように定義する .

$$\Phi_i = \frac{1}{|P_i|} \sum_{j \in P_i} \mu_j (q_{ij} - \hat{q}_{ij})^2 \quad (12)$$

ただし ,  $P_i$  は第  $i$  フレームの画像上で検出される特徴点の集合である . また ,  $q_{ij}$  は , 3.4.1 項で述べる手法により検出された , 第  $i$  フレームにおける特徴点  $j$  の画像上での座標 ,  $\hat{q}_{ij}$  は 3.2.1 項で定義した , 特徴点  $j$  の三次元位置を画像上へ投影した座標を表す .  $\mu_j$  は逐次処理で得られた特徴点  $j$  の信頼度であり , 特徴点  $j$

が第  $(w - m)$  フレームから第  $w$  フレームまで追跡された場合において，以下の式で算出される [28] .

$$\mu_j = \frac{1}{2\sigma_j^2} = \frac{m+1}{2} \left( \sum_{i=w-m}^w (\mathbf{q}_{ij} - \hat{\mathbf{q}}_{ij})^2 \right)^{-1} \quad (13)$$

ただし， $\sigma_j^2$  は特徴点  $j$  の再投影誤差の分散を表す .

### 3.3.2 GPS 測位位置に関するペナルティ

従来手法 [32,33] は，GPS の測位誤差が平均値 0 の正規分布で近似できることを前提に，以下のような GPS 測位位置に関するペナルティ項を定義している .

$$\hat{\Psi}_i = (\mathbf{M}_i \mathbf{g}_i - \mathbf{d})^2 \quad [32] \quad (14)$$

$$\hat{\Psi}_i = \frac{1}{\tau^2} (\mathbf{M}_i \mathbf{g}_i - \mathbf{d})^2 \quad [33] \quad (15)$$

ただし， $\mathbf{M}_i$  は 3.2.1 項で定義した，第  $i$  フレームにおける GPS 測地座標系からカメラ座標系への変換行列， $\mathbf{d}$  はカメラ座標系における GPS 受信機の位置， $\mathbf{g}_i$  は第  $i$  フレームの GPS 測地座標系における GPS の測位位置を表す . また， $\tau$  は GPS 受信機から得られる GPS 測位位置の RMS を表す . 前述したように，これらのペナルティ項を用いた場合において，GPS の平均測位位置が真値から大きく外れる誤差が生じた場合に，推定精度が低下するという問題がある .

本研究では，GPS の測位誤差を正規分布でモデル化する代わりに，GPS 受信機の真の位置が GPS の測位位置を中心とする，GPS の測位精度 (RTK-fix, RTK-float 等) に依存した一定の範囲内に存在するという仮定を用いる . 一般に GPS の測位誤差は水平方向と高度方向で異なることから，本論文では，GPS の水平方向の誤差と高度方向の誤差を独立に扱う . すなわち，図 6 に示すように，一定の範囲を円柱であるとし，GPS 測位位置に関するエネルギー  $\Psi_i$  を以下のように定義

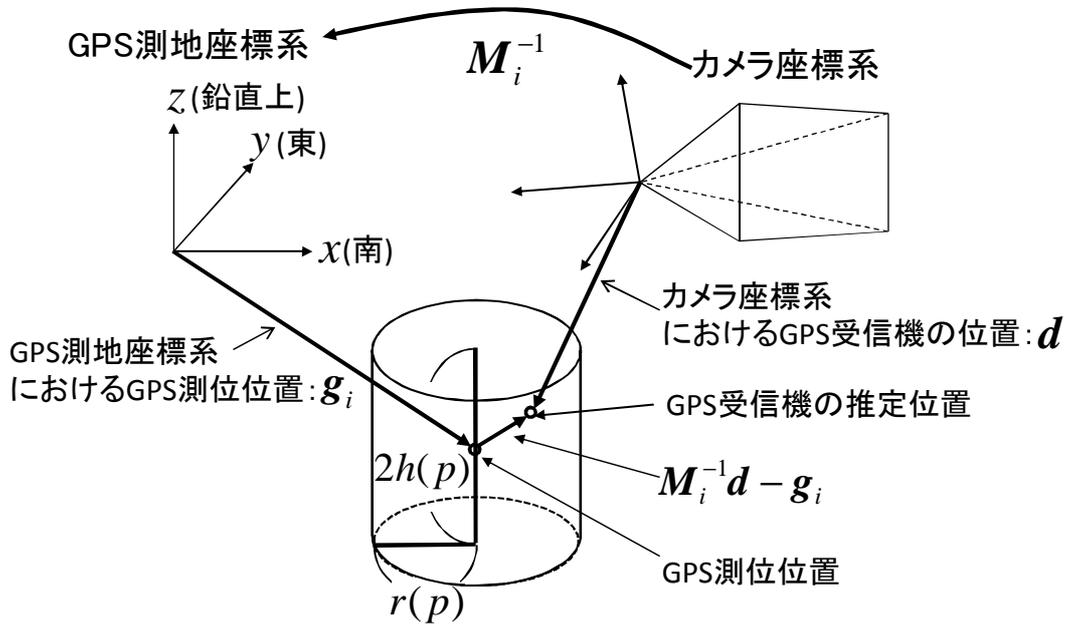


図 6 GPS 測位位置に関するペナルティ

する .

$$\Psi_i = \left( \frac{1}{r(p)} \sqrt{x_i^2 + y_i^2} \right)^{2n} + \left( \frac{1}{h(p)} z_i \right)^{2n} \quad (16)$$

$$\begin{pmatrix} x_i \\ y_i \\ z_i \\ 1 \end{pmatrix} = M_i^{-1} d - g_i \quad (17)$$

ただし,  $M_i^{-1}$  は第  $i$  フレームにおけるカメラ座標系から GPS 測地座標系への変換行列,  $r$  は円柱の半径,  $h$  は円柱の高さの  $1/2$ ,  $p$  は GPS の測位精度を表す. また,  $n$  はあらかじめ与える定数であり,  $n$  に大きな自然数を与えることで, GPS 受信機の位置が測位位置を中心とした円柱の中に存在する場合には  $\Psi_i$  が 0 に近い値となり, 円柱の外となる場合には非常に大きな値となる. 例えば,  $e = (\frac{1}{h}z)^{2n}$  のグラフは, 図 7 のようになり,  $n$  を大きく設定するにつれてグラフの傾きが急激に大きくなるのが分かる. ここでは,  $\Psi_i$  を微分可能な形で定義することで, 後述する最適化処理における勾配計算を可能としている.

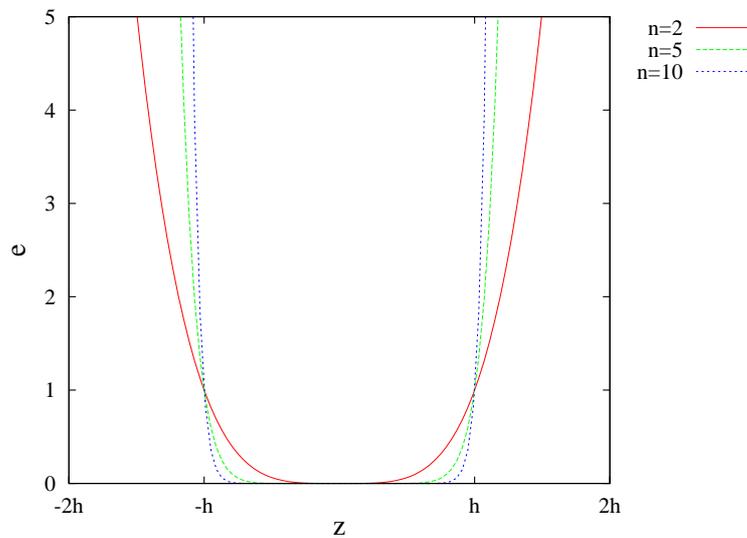


図 7  $n$  の違いによるペナルティ値の変化

なお，付録 A に示すように，GPS の精度に影響を与えるものとして，GPS の種類 (単独測位，DGPS，RTK-GPS)，RTK-GPS における解の種類 (RTK-fix，RTK-float)，DOP などが考えられる．後述する実験では，GPS として RTK-GPS を用いるが，この場合には解の種類に応じて測位精度が大きく異なることが知られているため，これを測位精度  $p$  として用いる．また，GPS の測位誤差はさまざまな要因によるため，理論的に大きさを見積もることは難しい．従って，それぞれの測位精度  $p$  における円柱のパラメータ  $r(p), h(p)$  は，カメラ位置・姿勢を推定する環境において GPS の定点観測を行うことで実験的に決定する必要がある．ただし，本研究では，GPS の測位値に含まれる，少数の誤差が非常に大きい外れ値は，GPS からの情報と動画像からの情報の整合性を検証することにより除去できるものとする．

### 3.4 動画像と GPS の併用によるカメラ位置・姿勢推定

本節では，図 3 に示した提案手法の処理の流れの各処理 (A) ~ (D) について詳述する．3.1 節で述べたように，処理 (A),(B) はフレーム毎，処理 (C) は一定フレーム毎，処理 (D) は最終的な最適化処理として 1 回のみ実行される．

### 3.4.1 特徴点の追跡によるカメラ位置・姿勢の推定

逐次処理 (A) においては, 特徴点を入力画像上で検出し, 前フレームと現フレームで対応関係を求めることにより, カメラの位置・姿勢を推定する. 本研究では特徴点として, 回転・拡大などの変形に対してロバストに同一の点を検出可能な輝度エッジのコーナーを用いる. また, 特徴点の仮の追跡を行い, ロバスト推定を用いて暫定的なカメラ位置・姿勢を推定することで, 追跡すべき特徴点の探索範囲を限定し, 周辺類似パターンとの誤対応を防止する [28]. 以下に具体的な処理の流れを示す.

- (A-1) Harris オペレータ [39] の出力が極大となる点を特徴点の追跡位置の候補とする.
- (A-2) 前フレームにおける各特徴点の近傍領域を現フレーム上で SSD(Sum of Squared Differences) を評価尺度としたテンプレートマッチングによって探索し, 現フレームにおける仮の追跡位置を決定する.
- (A-3) 仮の追跡位置から, LMedS 基準を用いたロバスト推定 [40] によって誤対応を排除し, 暫定的なカメラの位置・姿勢を得る.
- (A-4) 暫定的なカメラの位置・姿勢と特徴点の三次元位置から, 現フレームにおける特徴点の出現位置を予測し, 周辺の小さな領域に対して対応付けを行い, 追跡位置を確定する.
- (A-5) 追跡された特徴点の画像上の位置と三次元位置から, カメラの位置・姿勢を推定する.

(A-3), (A-5) におけるカメラ位置・姿勢は, 追跡された特徴点の画像上の座標と特徴点の三次元位置から, 以下の式を最小化することで推定する.

$$E_i = \sum_{j \in P_i} \mu_j (\mathbf{q}_{ij} - \hat{\mathbf{q}}_{ij})^2 \quad (18)$$

$E_i$  は, GPS の測位値を陽に用いることなく算出されるが, 3.4.3 項で述べる GPS 測位値を用いた最適化処理が, 一定フレーム毎に実行されることで, カメラ位置・

姿勢と特徴点の三次元位置が GPS 測地座標系上に配置され，最終的には GPS 測地座標系におけるカメラの位置・姿勢を求めることができる．

### 3.4.2 特徴点の三次元位置の推定と特徴点の追加・削除

現フレームまでのカメラ位置・姿勢と特徴点の画像上での検出座標を用いて，以下の式を最小化することにより，特徴点の三次元位置を推定し，更新する（処理 B）．

$$E_j = \sum_{i \in S_j} (q_{ij} - \hat{q}_{ij})^2 \quad (19)$$

ただし， $S_j$  は特徴点  $j$  が検出されたフレームの集合を表す．

また，本研究では，広域な屋外環境を対象とした動画像を扱うため，カメラの動きによって特徴点の出現・消失が起こる．そこでフレーム毎に自動で特徴点の追加・削除を行なう [28]．特徴点の追加は，追跡された特徴点の候補から，以下の全てを満たすものを選択することで実現する．

- 信頼度が一定値以上
- 各画像間のマッチング誤差が一定値以下
- Harris オペレータによる特徴量が一定値以上
- 推定される三次元位置と各フレームのカメラの投影中心を結ぶ直線群の成す最大角が一定値以上

また，以下のいずれかの条件を満たす特徴点を削除する．

- 信頼度が一定値以下
- 各画像間のマッチング誤差が一定値以上
- 再投影誤差が一定値以上

なお，本処理では処理 (A) と同様に，GPS 測位値を陽に使用していないが，GPS 測位値を用いた最適化によって，最終的には GPS 測地座標系における特徴点の三次元位置が得られる．

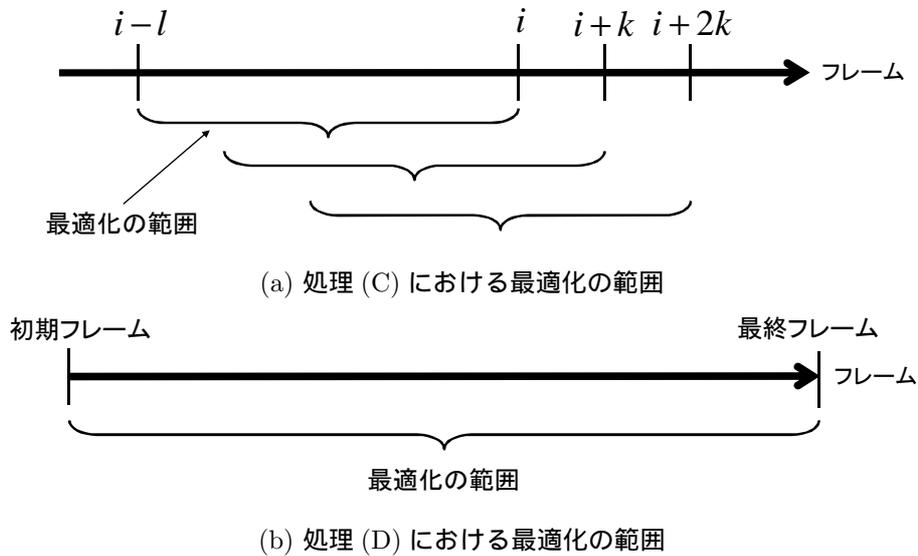


図 8 最適化の範囲

### 3.4.3 GPS 測位値を用いた最適化

GPS 測位値を用いた最適化には、図 8 に示すように、逐次処理において範囲を限定して行う狭区間最適化 (C) と、逐次処理の終了後、動画像全体を対象として行う全体最適化 (D) がある。

処理 (A), (B) は、動画像上の情報のみを考慮した推定を行うため、得られたカメラの位置・姿勢と特徴点の三次元位置には蓄積誤差が存在する。従って、処理 (A), (B) を繰り返すだけでは、以下の問題が生じる。

1. 蓄積誤差の影響により、3.4.1 項で述べた特徴点の追跡に失敗する。
2. 逐次処理の終了後に行う全体最適化 (D) において、初期値が真値から離れるため、大域最適解を求めることが難しくなる。

そこで処理 (C) では、逐次処理において、図 8(a) に示すように、第  $(i-l)$  フレームから現フレーム (第  $i$  フレーム) を用いて、式 (11) で定義した誤差関数  $E$  を勾配法によって最小化することにより、GPS の測位値を用いてカメラの位置・姿勢と特徴点の三次元位置を最適化する。具体的には、第  $i$  フレームにおけるカメラの姿勢  $(r_{1i}, r_{2i}, r_{3i})$ 、位置  $(t_{1i}, t_{2i}, t_{3i})$  を以下の式を逐次的に繰り返すことで更新

し、 $E$  を最小化する。

$$\begin{aligned} r_{1i} &\leftarrow r_{1i} - l_{r_{1i}} \frac{\delta E}{\delta r_{1i}}, & r_{2i} &\leftarrow r_{2i} - l_{r_{2i}} \frac{\delta E}{\delta r_{2i}}, & r_{3i} &\leftarrow r_{3i} - l_{r_{3i}} \frac{\delta E}{\delta r_{3i}}, \\ t_{1i} &\leftarrow t_{1i} - l_{t_{1i}} \frac{\delta E}{\delta t_{1i}}, & t_{2i} &\leftarrow t_{2i} - l_{t_{2i}} \frac{\delta E}{\delta t_{2i}}, & t_{3i} &\leftarrow t_{3i} - l_{t_{3i}} \frac{\delta E}{\delta t_{3i}} \end{aligned} \quad (20)$$

次に、特徴点  $j$  の三次元位置  $(X_j, Y_j, Z_j)$  を以下の式を逐次的に繰り返すことで更新し、 $E$  を最小化する。

$$X_j \leftarrow X_j - l_{X_j} \frac{\delta E}{\delta X_j}, \quad Y_j \leftarrow Y_j - l_{Y_j} \frac{\delta E}{\delta Y_j}, \quad Z_j \leftarrow Z_j - l_{Z_j} \frac{\delta E}{\delta Z_j} \quad (21)$$

ただし、矢印は更新を意味する。また、勾配法のスケール係数  $(l_{r_{1i}}, l_{r_{2i}}, l_{r_{3i}}, l_{t_{1i}}, l_{t_{2i}}, l_{t_{3i}})$ 、 $(l_{X_j}, l_{Y_j}, l_{Z_j})$  は、各パラメータの更新時点で  $E$  が最小となるように直線探索を行い決定する。これらカメラ位置・姿勢の最適化と特徴点の三次元位置の最適化を  $E$  が収束するまで交互に繰り返す。

処理 (D) では、処理 (A) ~ (C) で得られたカメラ位置・姿勢、特徴点の三次元位置、特徴点の信頼度を初期値として、図 8(b) に示すように、動画像全体を用いて誤差関数  $E$  を最小化することにより、カメラ位置・姿勢の推定精度の向上を図る。

## 4. 実験

本章では、カメラと GPS を移動させながら取得した動画像と GPS の測位値を用いてカメラ位置・姿勢の推定を行い、従来手法との定量的な推定精度の比較を行うことで、提案手法の有効性を示す。ここではまず、GPS を固定して長時間観測を行うことで、3.3.2 項で述べた提案手法で用いる測位精度ごとの位置計測の誤差範囲を決定し、次に決定したパラメータを用いた推定結果により評価を行う。

### 4.1 定点観測による GPS 測位値の誤差分析

本実験では、GPS 測位値の定点観測を行うことで、式 (16) で定義した GPS の測位位置に関するペナルティで用いる、測位精度ごとの円柱のパラメータ  $r(p)$ ,  $h(p)$  を決定する。なお、3.3.2 項で述べたように、GPS の測位精度  $p$  としては GPS の解の種類 (RTK-fix, RTK-float) を用いる。

#### 4.1.1 実験条件

RTK-GPS 受信機 (TOPCON GR-3) を図 9 に示すような環境に固定し、約 5 時間の連続測位を行った。表 3 に利用した RTK-GPS 受信機のカタログ仕様を示す。測位は、上空が開けており、後述する実験においてカメラ位置・姿勢推定を行う環境と概ね同じ環境で行った。GPS の測位周期は 1Hz であり、インターネット回線を通して補正データを受信することで、1 台の RTK-GPS 受信機で RTK 測位が可能な仮想基準点方式で測位を行った。また、本実験では、GPS 測位の誤差を算出するため、得られた全ての RTK-fix データの平均位置を、定点観測を行った場所の真値とすることで、各測位値に対する誤差を算出した。



図 9 GPS 測位値の定点観測の様子

表 3 RTK-GPS 受信機のカatalog仕様

| 機種名          | GR-3  |
|--------------|---|
| 使用搬送波        | GPS L1/L2/L2C (2 周波)<br>GLONASS L1/L2 (2 周波)                    |
| 水平測位精度 (RTK) | $\pm(10 \text{ mm} + 1.0 \text{ ppm} \times D)$ m.s.e<br>D:測定距離 |
| 垂直測位精度 (RTK) | $\pm(15 \text{ mm} + 1.0 \text{ ppm} \times D)$ m.s.e<br>D:測定距離 |

#### 4.1.2 GPS の測位精度に依存した誤差範囲の決定

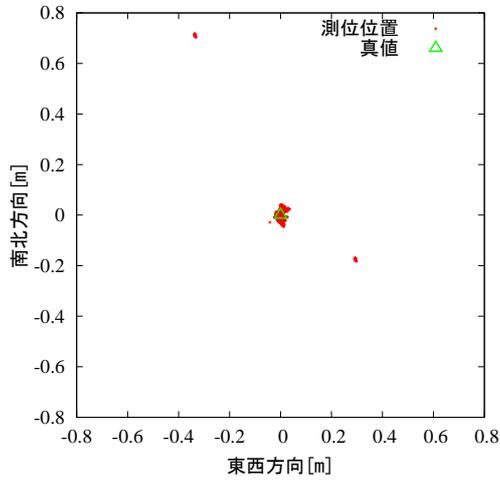
定点観測によって得られた GPS の測位位置の分布を，測位精度ごとに図 10，11 に示す．また，表 4(a) に，測位精度ごとの真値からの誤差の最大値を示す．これらの結果から，測位精度によって誤差の大きさが大きく異なることおよび，水平方向と高度方向で測位位置のばらつき具合が異なることが分かる．表 4(b) に，測位精度ごとに，水平方向，高度方向それぞれについて誤差が大きい 5% の測位値を除去した結果を示す．外れ値を除去することにより，誤差の最大値が大幅に小さくなっていることおよび，図 10，11 に示した測位位置の分布から，GPS の測位値には誤差が極端に大きい外れ値が含まれていることが確認できる．以降の実験では，GPS 測位の外れ値は動画像からの情報と GPS からの情報の整合性を検証することにより除去できるものとし，表 4(b) に示した，各測位精度における水平方向の誤差の最大値と高度方向の誤差の最大値を，それぞれ円柱のパラメータ  $r(p)$ ， $h(p)$  とする．

ただし，以降の実験では，カメラと GPS 受信機は移動しているため， $r(p)$ ， $h(p)$  の決定にはこの移動量を考慮する必要がある．すなわち，本研究で用いる式 (11) のエネルギー関数は，1 フレーム以下の精度での画像と GPS の同期を扱えないため，画像の取得と GPS 測位の時刻には，0.5 フレーム分のずれが生じる可能性がある．そこで，実際の円柱のサイズとしては，表 4(b) に示した水平方向，高度方向それぞれの値に，0.5 フレームの間に人が動く距離として 37mm と 5mm を足した値を用いる．<sup>1</sup>

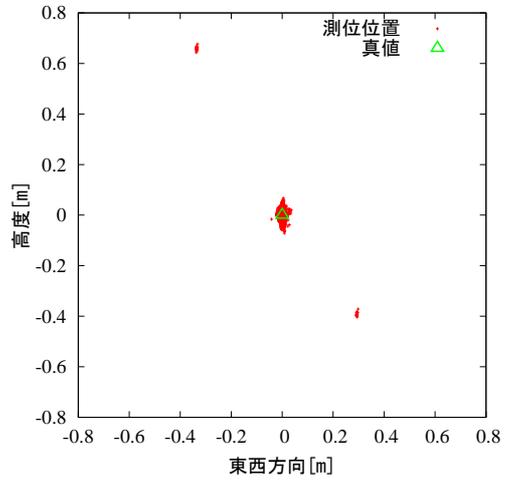
表 4 測位精度ごとの誤差の最大値 [mm]

| 測位精度      | (a) 全データ |       | (b) 外れ値除去後 |      |
|-----------|----------|-------|------------|------|
|           | 水平方向     | 高度方向  | 水平方向       | 高度方向 |
| RTK-fix   | 793      | 677   | 29         | 41   |
| RTK-float | 12792    | 20424 | 3778       | 9504 |

<sup>1</sup>後述する実験では，人が手に持ったカメラで歩行中に撮影された 15fps の動画像を用いる．人の歩行速度は時速 4km であると言われているため，0.5 フレーム間の移動量は 37mm となる．また，上下運動は 15cm 程度であると言われているため，0.5 フレーム間の移動量は 5mm となる．

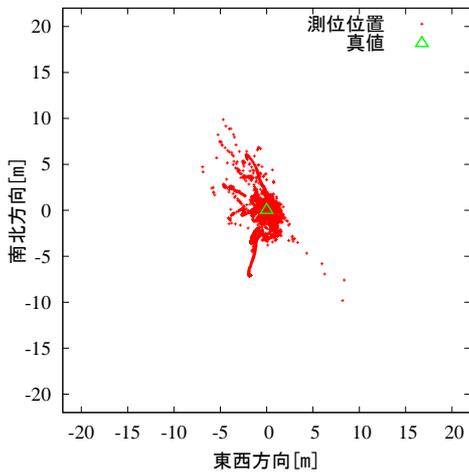


(a) 東西-南北平面

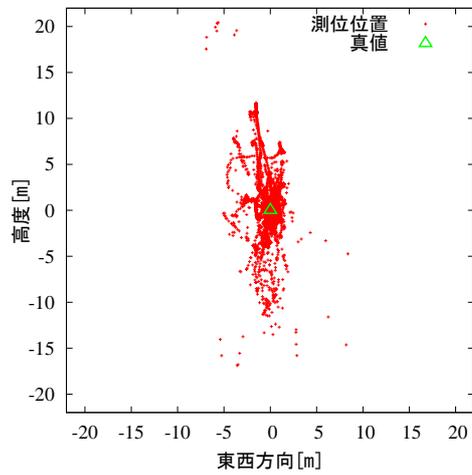


(b) 東西-高度平面

図 10 定点観測における測位位置の分布 (RTK-fix)



(a) 東西-南北平面



(b) 東西-高度平面

図 11 定点観測における測位位置の分布 (RTK-float)

## 4.2 動画像と GPS 測位値を用いたカメラ位置・姿勢推定

本実験では、動画像と GPS の測位値を用いてカメラ位置・姿勢の推定を行い、従来のバンドル調整による GPS と動画像情報の融合手法 [32,33] との定量的な推定精度の比較により、提案手法の有効性を確認する。

### 4.2.1 実験条件

本実験では、ワイドレンズ (Sony VCL-HG0758) を取り付けたビデオカメラ (Sony DSR-PD-150) を用いて約 70m の距離を移動しながら、手持ちで撮影した動画像 (解像度  $720 \times 480$  画素, プログレッシブ撮影, 15 fps, 1110 フレーム) と RTK-GPS 受信機 (TOPCON GR-3) から得られた 1 秒間隔で計測された GPS 測位値を入力として用いてカメラ位置・姿勢の推定を行った。入力画像の一部を図 12 に示す。ただし、GPS 測位値は全て RTK-fix として取得され、これを本実験における真値として用いる。入力は、得られた GPS 測位値から後処理ソフトウェア (TOPCON Tools) を用いて、仮想的に GPS 衛星をマスクすることにより、RTK-float のデータを作成し、これを用いた。

カメラと GPS 受信機は図 13 に示すように、相対位置関係を固定し、カメラ座標系における GPS 受信機の位置  $d$  は、トータルステーションで計測した基準点の三次元位置を画像上で指定することによって得たカメラ位置・姿勢と、トータルステーションで計測した GPS 受信機の位置から算出した。カメラの内部パラメータはあらかじめ Tsai の手法 [38] によって求め、動画像と GPS の測位値は手動で開始時刻を同期させた。また、式 (11) で用いられる GPS 測位位置に関するペナルティの重みを  $\omega = 10^{-8}$  とし、狭区間最適化を行うフレーム間隔を  $k = 15$ 、使用フレームを  $l = 500$  に設定した。

本実験では、提案手法の有効性を確認するため、以下の手法とカメラ位置の推定精度を比較する。

手法 A: GPS 測位値を用いず、動画像のみから推定を行う手法。

手法 B: GPS の誤差として正規分布を仮定し、GPS の測位精度を考慮しない従来手法 [32]。



(a) 第 1 フレーム



(b) 第 150 フレーム



(c) 第 300 フレーム



(d) 第 450 フレーム



(e) 第 600 フレーム



(f) 第 750 フレーム



(g) 第 900 フレーム



(h) 第 1110 フレーム

図 12 入力画像の一部



図 13 カメラと GPS 受信機の位置関係

手法 C: GPS の誤差として正規分布を仮定し，正規分布の分散を GPS の測位精度を考慮して変化させる従来手法 [33] .

手法 D: 提案手法において，GPS の誤差範囲を回転楕円体とした場合 .

ただし，ここでは手法 A によって推定されるカメラ位置・姿勢を GPS 測地座標系で評価するため，第 1 フレームから第 30 フレームまでのカメラ位置・姿勢を画像上で三次元位置が既知の基準点を指定することによって与えた . また，これらの基準点は手法 B, C, D, および提案手法において特徴点の追跡処理の初期化のためにも使用した .

手法 D は，3.3.2 節で定義した GPS 測位位置に関するペナルティを決定するための，GPS 測位位置を中心とした一定の範囲を円柱から回転楕円体に変更したものであり，手法 D との比較によって，範囲の形が推定精度に与える影響を検証する . 具体的には，一定の範囲を回転楕円体とした場合の GPS 測位位置に関するペナルティ項  $\Psi_i$  を以下のように定義する .

$$\Psi_i = \left( \frac{1}{\alpha(p)} \sqrt{x_i^2 + \left( \frac{\alpha(p)}{\beta(p)} y_i \right)^2 + z_i^2} \right)^{2n} \quad (22)$$

表 5 回転楕円体のパラメータ [mm]

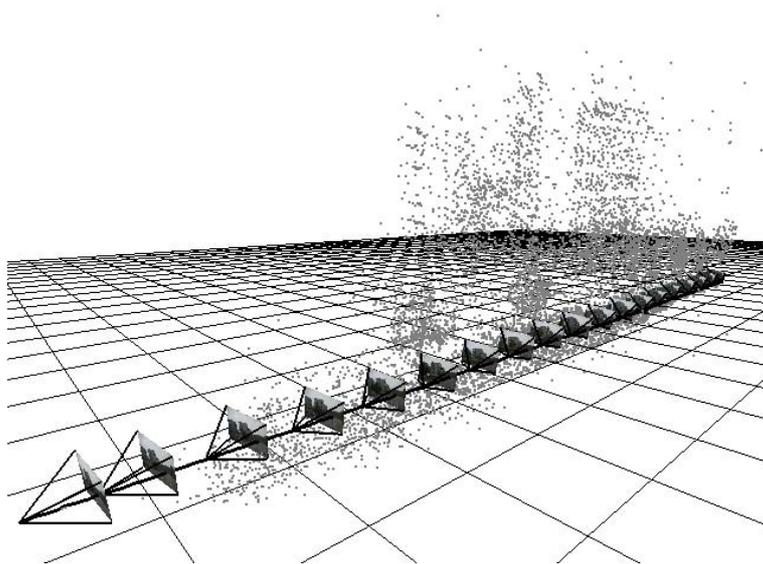
| 測位精度      | $\alpha$ | $\beta$ |
|-----------|----------|---------|
| RTK-fix   | 25       | 55      |
| RTK-float | 4125     | 11366   |

ただし， $\alpha, \beta$  はそれぞれ水平方向，高度方向の回転楕円体の径の  $1/2$  である．円柱の場合と同様に，GPS 受信機の位置が測位位置を中心とした回転楕円体の中に存在する場合には  $\Psi_i$  が 0 に近い値となり，回転楕円体の外となる場合には非常に大きな値となる．また，回転楕円体のパラメータ  $\alpha(p), \beta(p)$  は，4.1.2 節で決定した円柱のパラメータと同様に，GPS 測位値の定点観測の結果から決定する．ただし，回転楕円体の体積が最も小さくなるように，円柱の場合と同数の測位値を外れ値として除去する．決定した回転楕円体のパラメータを表 5 に示す．

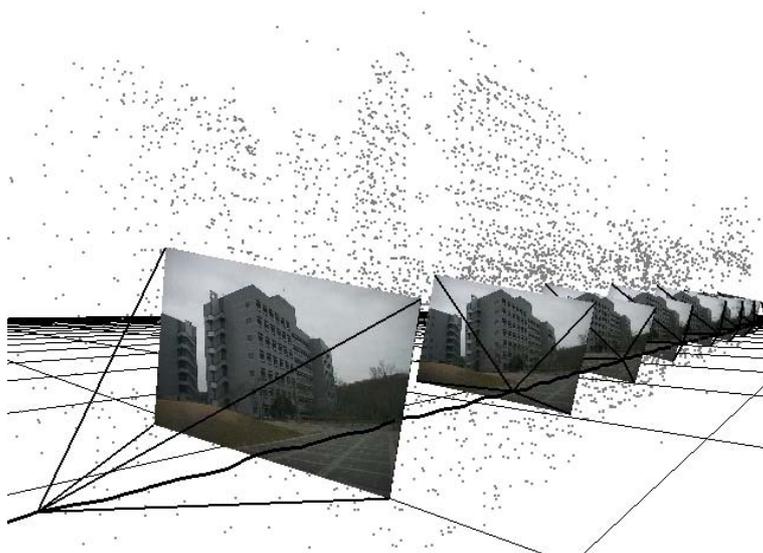
#### 4.2.2 カメラ位置推定精度の定量的評価

提案手法によって推定されたカメラ位置・姿勢と特徴点の三次元位置を図 14 に示す．図中の曲線は推定されたカメラパスを，錐体は 60 フレーム毎のカメラの姿勢を表し，錐体中の画像はそのフレームにおける入力画像を表す．同図から，カメラパスは滑らかに推定されていることが分かる．また，目視によって環境の三次元構造が概ね正しく復元されていることを確認した．

各手法における，水平方向および高度方向の GPS 受信機の位置の推定結果と GPS の測位位置を図 15 に示す．また真値との比較による，フレームごとの位置誤差を図 16 に，位置誤差の平均，標準偏差，最大値を表 6 に示す．図 16 より，動画像のみを用いる手法 A は手法 C, D, および提案手法と比較して，位置誤差の増加率が大きいことから，カメラ位置・姿勢の推定結果が蓄積誤差に大きく影響されていることが分かる．これは，手法 A が動画像のみを用いており，蓄積誤差を抑える外部指標を持たないためである．また，図 15 より，手法 B は測位精度の低い RTK-float の測位位置に大きく影響されていることが分かる．これは，手法 B が GPS の測位精度を考慮せず，GPS の測位誤差は小さいと仮定しているためである．これに対して，手法 C, D, および提案手法は GPS の測位精度を考慮す

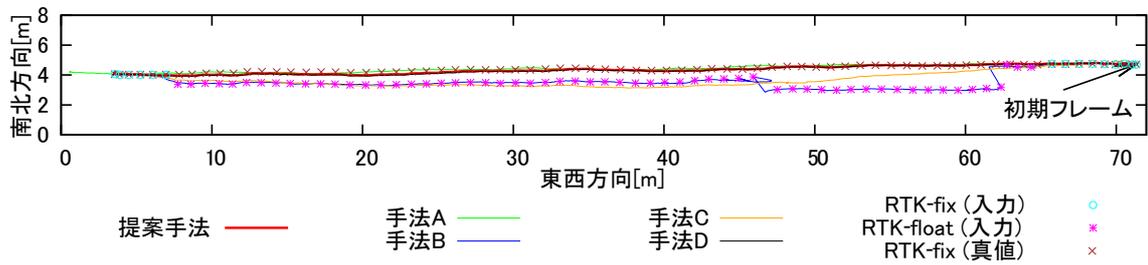


(a) 全体

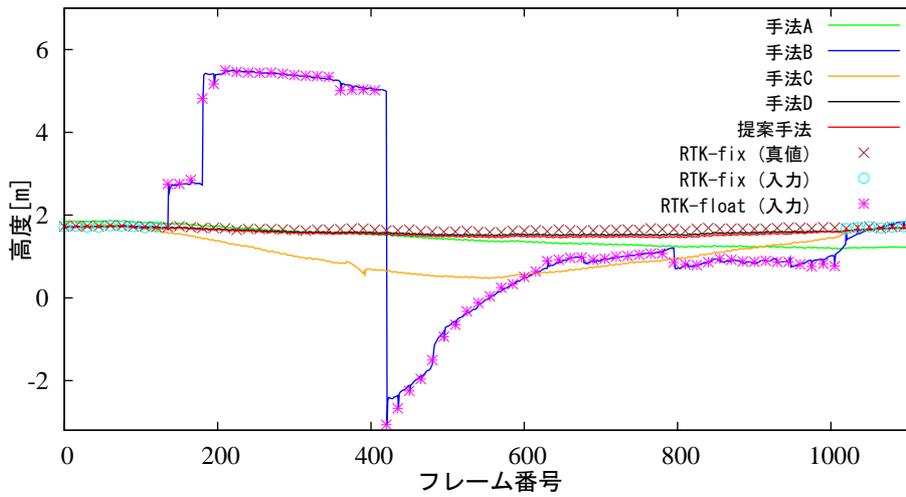


(b) 一部拡大

図 14 推定されたカメラ位置・姿勢と特徴点の三次元位置



(a) 水平方向



(b) 高度方向

図 15 推定された GPS 受信機の位置

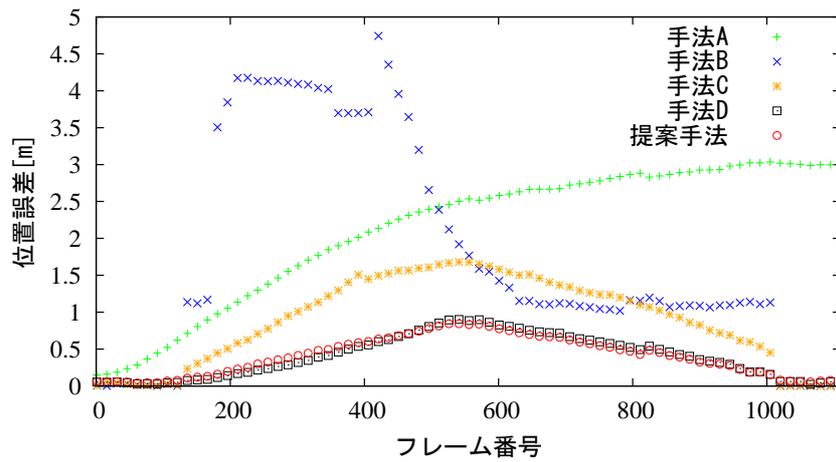


図 16 フレームごとの位置誤差

表 6 位置誤差の比較 [mm].

| 手法   | 平均     | 標準偏差   | 最大     |
|------|--------|--------|--------|
| 手法 A | 2109.6 | 915.8  | 3036.2 |
| 手法 B | 1787.0 | 1497.3 | 4742.5 |
| 手法 C | 894.7  | 582.2  | 1680.7 |
| 手法 D | 405.5  | 286.7  | 901.6  |
| 提案手法 | 406.0  | 261.8  | 844.2  |

ることにより，手法 B より RTK-float による誤差の影響が小さくなっている．これは，GPS の測位精度を考慮することで RTK-float の影響を抑えつつ，動画像の初期フレームと最終フレーム付近に存在する RTK-fix の測位位置を用いて蓄積誤差を最小化しているためである．ただし，手法 C では，手法 D，および提案手法と比較して RTK-float の影響がより大きく，推定されるカメラ位置が RTK-float の測位位置に引き寄せられている．これは，手法 C が GPS の測位誤差として平均値 0 の正規分布を仮定しているのに対し，短時間における RTK-float の平均測位位置が真値から外れているためであると考えられる．手法 D，および提案手法では RTK-float の測位位置に大きく影響されることなく，カメラ位置が推定されている．表 6 から分かるように，本実験では，手法 D，および提案手法は従来手法 C に対して 2 倍程度の精度向上が見られる．また，GPS の測位位置を中心とした一定の範囲を回転楕円体とした手法 D と，円柱とした提案手法を比較して，カメラ位置の推定精度にほぼ差がないことから，本実験においては範囲の形は推定精度に大きく影響しないことが分かる．

## 5. まとめ

本論文では、三次元形状復元や自由視点画像生成などオフラインで高精度なカメラ位置・姿勢を必要とするアプリケーションへの適用を前提とし、GPS と動画像の併用によるカメラ位置・姿勢推定において、GPS 受信機の真の位置が GPS の測位位置を中心とする GPS の測位精度に依存した一定の大きさの円柱内に存在することを仮定し、この仮定に基づくエネルギー関数を最小化することで、GPS の測位精度が低い場合にも精度良くカメラ位置・姿勢の推定が可能な手法を提案した。実験では、屋外環境で取得した動画像と GPS の測位値を用いてカメラ位置・姿勢を推定し、従来手法との定量的な推定精度の比較を行った。実験の結果、GPS の測位精度が低い場合において、提案手法は従来手法より精度良くカメラ位置を推定できることを確認した。このことから、本手法は、屋外環境の三次元形状復元や自由視点画像生成などへ適用できると考えられる。

今後の課題として、GPS 測位に含まれる外れ値の自動除去手法の開発、動画像と GPS の同期の自動化、GPS の解の種類以外の測位精度に影響を与える要因の考慮などがあげられる。外れ値の除去と同期の自動化は、GPS から得られる情報と動画像から得られる情報の整合性を、エネルギー関数の値の大きさを評価尺度として検証することで実現できると考えられる。また、測位精度に影響を与える要因の考慮として、GPS の測位におけるマルチパスなどの環境に依存する誤差を、動画像からの推定によって得られる環境の三次元構造を用いることにより見積もることが考えられる。

## 謝辞

本研究の全過程を通して、懇切なる御指導、御鞭撻を賜りました視覚情報メディア講座 横矢直和 教授に心より感謝致します。また、本研究の遂行にあたり、有益な御助言、御鞭撻を頂いたインタラクティブメディア設計学講座 加藤博一 教授に厚く御礼申し上げます。そして、本研究を進めるにあたり、始終暖かい御指導をして頂いた視覚情報メディア講座 山澤一誠 准教授に深く感謝致します。さらに、本研究を通じて、的確な御助言、御鞭撻を頂いた視覚情報メディア講座 神原誠之 助教に深く御礼申し上げます。また、本研究を行うにあたり、多大なる御助言、御鞭撻を賜った視覚情報メディア講座 佐藤智和 助教に心より感謝致します。特に、佐藤智和 助教には本研究のテーマ設定から本論文の執筆、その他の発表論文の添削、発表練習に至るまで細やかな御指導を頂きました。さらに、本研究の遂行にあたり、貴重な御助言や御指摘を頂きました視覚情報メディア講座 武富貴史氏に深く感謝いたします。そして、本研究を進める上で、御多用の中にもかかわらず貴重な御意見を頂いた株式会社トプコン 穴井哲治氏に深く御礼申し上げます。また、研究室での生活を支えていただいた視覚情報メディア講座事務補佐員 高橋美央 女史に厚く御礼申し上げます。最後に、研究活動だけでなく日々の生活においても大変お世話になった視覚情報メディア講座の諸氏に心より感謝致します。

## 参考文献

- [1] M. Pollefeys, D. Nistér, J. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S. Kim, P. Merrell, et al.: “Detailed real-time urban 3D reconstruction from video,” *Int. J. of Computer Vision*, Vol. 78, No. 2-3, pp. 143–167, 2008.
- [2] S. Knorr, M. Kunter and T. Sikora: “Super-resolution stereo- and multi-view synthesis from monocular video sequences,” *Proc. Int. Conf. on 3-D Digital Imaging and Modeling*, pp. 55–64, 2007.
- [3] A. Stafford, W. Piekarski and B. H. Thomas: “Implementation of god-like interaction techniques for supporting collaboration between outdoor AR and indoor tabletop users,” *Proc. Int. Symp. on Mixed and Augmented Reality*, pp. 165–172, 2006.
- [4] J. Newman, D. Ingram and A. Hopper: “Augmented reality in a wide area sentient environment,” *Proc. Int. Symp. on Augmented Reality*, pp. 77–86, 2001.
- [5] D. Hallaway, T. Höllerer and S. Feiner: “Coarse, inexpensive, infrared tracking for wearable computing,” *Proc. Int. Symp. on Wearable Computers*, pp. 69–78, 2003.
- [6] R. Tenmoku, M. Kanbara and N. Yokoya: “A wearable augmented reality system using positioning infrastructures and a pedometer,” *Proc. Int. Symp. on Wearable Computers*, pp. 110–117, 2003.
- [7] S. Feiner, B. MacIntyre, T. Höllerer and A. Webster: “A touring machine: prototyping 3d mobile augmented reality systems for exploring the urban environment,” *Proc. Int. Symp. on Wearable Computers*, pp. 74–81, 1997.
- [8] R. Azuma, B. Hoff, H. Neely, III and R. Sarfaty: “A motion-stabilized outdoor augmented reality system,” *Proc. IEEE Virtual Reality*, pp. 252–259,

1999.

- [9] W. Piekarski, R. Smith and B. H. Thomas: “Designing backpacks for high fidelity mobile outdoor augmented reality,” Proc. Int. Symp. on Mixed and Augmented Reality, pp. 280–281, 2004.
- [10] 神原, 横矢: “RTK-GPS と慣性航法装置を併用したハイブリッドセンサによる屋外型拡張現実感システム”, 画像の認識・理解シンポジウム (MIRU2005) 講演論文集, pp. 933–938, 2005.
- [11] K. Ohno, T. Tsubouchi, B. Shigematsu, S. Maeyama and S. Yuta: “Outdoor navigation of a mobile robot between buildings based on DGPS and odometry data fusion,” Proc. Int. Conf. on Robotics and Automation, pp. 1978–1984, 2003.
- [12] M. Kouroggi, N. Sakata, T. Okuma and T. Kurata: “Indoor/outdoor pedestrian navigation with an embedded GPS/RFID/self-contained sensor system,” Proc. Int. Conf. on Artificial Reality and Telexistence, pp. 1310–1321, 2006.
- [13] C. Tomasi and T. Kanade: “Shape and motion from image streams under orthography: a factorization method,” Int. J. of Computer Vision, Vol. 9, No. 2, pp. 137–154, 1992.
- [14] P. Sturm and B. Triggs: “A factorization based algorithm for multi-image projective structure and motion,” Proc. European Conf. on Computer Vision, pp. 709–720, 1996.
- [15] 植芝, 富田: “奥行パラメータの逐次推定による多視点透視投影画像のための因子分解法”, 電子情報通信学会論文誌 (D), Vol. J81-D2, No. 8, pp. 1718–1726, 1998.
- [16] 上島, 斎藤: “射影幾何学に基づく計測行列補間を用いた因子分解法”, 情報処理学会研究報告 CVIM135-15, pp. 95–102, 2002.

- [17] B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: “Bundle adjustment - a modern synthesis,” Proc. Int. Workshop on Vision Algorithms, pp. 298–372, 1999.
- [18] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops and R. Koch: “Visual modeling with a hand-held camera,” Int. J. of Computer Vision, Vol. 59, No. 3, pp. 207–232, 2004.
- [19] G. Klein and D. Murray: “Parallel tracking and mapping for small AR workspaces,” Proc. Int. Symp. on Mixed and Augmented Reality, pp. 225–234, 2007.
- [20] A. J. Davison, I. D. Reid, N. D. Molton and O. Stasse: “MonoSLAM: Real-time single camera SLAM,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 29, No. 6, pp. 1052–1067, 2007.
- [21] 加藤, M. Billingham, 浅野, 橘啓: “マーカー追跡に基づく拡張現実感システムとそのキャリブレーション”, 日本バーチャルリアリティ学会論文誌, Vol. 4, No. 4, pp. 607–616, 1999.
- [22] M. Kalkusch, T. Lidy, M. Knapp, G. Reitmayr, H. Kaufmann and D. Schmalstieg: “Structured visual markers for indoor pathfinding,” Proc. Int. Augmented Reality Toolkit Workshop, 2002.
- [23] 中里, 神原, 横矢: “不可視マーカーを用いた位置・姿勢推定のための環境構築とユーザ位置・姿勢推定システム”, 日本バーチャルリアリティ学会論文誌, Vol. 13, No. 2, pp. 257–266, 2008.
- [24] T. Drummond and R. Cipolla: “Real-time visual tracking of complex structures,” IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 24, No. 7, pp. 932–946, 2002.
- [25] L. Vacchetti, V. Lepetit and P. Fua: “Combining edge and texture information for real-time accurate 3D camera tracking,” Proc. Int. Symp. on Mixed

and Augmented Reality, 2004.

- [26] E. Royer, M. Lhuillier, M. Dhome and J. M. Lavest: “Monocular vision for mobile robot localization and autonomous navigation,” *Int. J. of Computer Vision*, Vol. 74, No. 3, pp. 237–260, 2007.
- [27] 武富, 佐藤, 横矢: “拡張現実感のための優先度情報を付加した自然特徴点ランドマークデータベースを用いた実時間カメラ位置・姿勢推定”, *電子情報通信学会論文誌 (D)*, Vol. J92-D, No. 8, pp. 1440–1451, 2009.
- [28] 佐藤, 神原, 横矢, 竹村: “マーカと自然特徴点の追跡による動画像からのカメラ移動パラメータの復元”, *電子情報通信学会論文誌 (D-II)*, Vol. J86-D-II, No. 10, pp. 1431–1440, 2003.
- [29] G. Bleser, H. Wuest and D. Stricker: “Online camera pose estimation in partially known and dynamic scenes,” *Proc. Int. Symp. on Mixed and Augmented Reality*, pp. 56–65, 2006.
- [30] M. Agrawal and K. Konolige: “Real-time localization in outdoor environments using stereo vision and inexpensive GPS,” *Proc. Int. Conf. on Pattern Recognition*, Vol. 3, pp. 1063–1068, 2006.
- [31] D. Schleicher, L. M. Bergasa, M. Ocana, R. Barea and E. Lopez: “Real-time hierarchical GPS aided visual SLAM on urban environments,” *Proc. Int. Conf. on Robotics and Automation*, pp. 4381–4386, 2009.
- [32] 横地, 池田, 佐藤, 横矢: “特徴点追跡と GPS 測位に基づくカメラ外部パラメータの推定”, *情報処理学会論文誌: コンピュータビジョンとイメージメディア*, Vol. 47, No. SIG5(CVIM13), pp. 69–79, 2006.
- [33] T. Anai, N. Fukaya, T. Sato, N. Yokoya and N. Kochi: “Exterior orientation method for video image sequences with considering RTK-GPS accuracy,” *Proc. Int. Conf. on Optical 3-D Measurement Techniques*, Vol. I, pp. 231–240, 2009.

- [34] B. Jiang, U. Neumann and S. You: “A robust hybrid tracking system for outdoor augmented reality,” Proc. IEEE Virtual Reality, pp. 3–10, 2004.
- [35] G. Reitmayr and T. Drummond: “Going out: Robust model-based tracking for outdoor augmented reality,” Proc. Int. Symp. on Mixed and Augmented Reality, pp. 109–118, 2006.
- [36] G. Klein and T. Drummond: “Robust visual tracking for non-instrumented augmented reality,” Proc. Int. Symp. on Mixed and Augmented Reality, pp. 113–122, 2003.
- [37] 松田, 池田, 佐藤, 横矢: “自然特徴点ランドマークデータベースと姿勢センサに基づく高速な回転にロバストなカメラ位置・姿勢推定”, 画像の認識・理解シンポジウム (MIRU2007) 講演論文集, pp. 1558–1563, 2007.
- [38] R. Y. Tsai: “An efficient and accurate camera calibration technique for 3D machine vision,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 364–374, 1986.
- [39] C. Harris and M. Stephens: “A combined corner and edge detector,” Proc. Alvey Vision Conf., pp. 147–151, 1988.
- [40] 栗田, 坂上: “ロバスト統計とその画像理解への応用”, 画像の認識・理解シンポジウム (MIRU2000) 講演論文集, pp. 65–70, 2000.
- [41] 安田: “GPS 技術の展望”, 電子情報通信学会論文誌 (B), Vol. J84-B, No. 12, pp. 2082–2091, 2001.
- [42] 臼井: “高精度衛星測位技術の動向と応用”, 電子情報通信学会誌, Vol. 92, No. 9, pp. 768–774, 2009.

## 付録

### A. GPS から得られる情報の特性

本項では，GPS の種類と測位精度，GPS の測位誤差の要因と特性について述べる．

#### A.1 GPS の種類と測位精度

図 17 に GPS の種類と測位精度の関係 [41, 42] を示す．単独測位は 1 台の受信機のみで位置を算出する方式であり，A.2 で述べる誤差を補正することができない．これに対して，DGPS(Differential GPS)，RTK-GPS(Real Time Kinematic GPS) は，正しい位置が既知である基準点を用いることにより，誤差を補正することができる．特に，RTK-GPS は搬送波の位相を観測することにより，数 cm ~ 数 mm レベルの高精度な測位を可能とする．ただし，得られた位相には波数分の不確定が含まれ，これを確定することは整数値バイアスの決定と呼ばれる．整数値バイアスが整数値として求められた解を RTK-fix 解と呼び，実数値として求められた解である RTK-float 解より精度が良いとされる．

#### A.2 GPS の測位誤差の要因と特性

GPS 測位には以下のような誤差が伴う [41, 42] ．

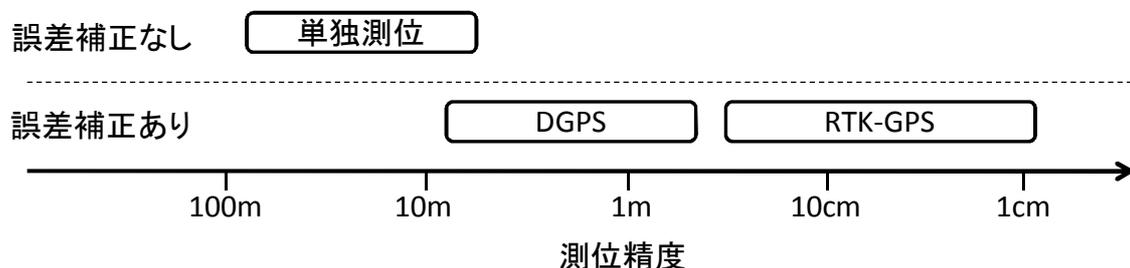


図 17 GPS の種類と測位精度

#### 電離層誤差

搬送波が電離層を通過する際に発生する伝搬遅延により到着時間が長くなり、これに基づいて計算した衛星との距離は実際より長くなる。誤差の量は電離層の活動状況に依存し、地域、季節、時間帯、太陽活動によって大きく異なる。航法データには電離層誤差の補正情報が含まれているが、補正できない誤差が10m以上に達することがある。

#### 対流圏誤差

搬送波が対流圏を通過する際の空気及び水蒸気による遅延であり、大気の状態に依存する。誤差量は1m未満程度である。

#### 衛星軌道誤差，衛星時計誤差

衛星の軌道は地球重力の不均一などにより乱れを生じ、予測との誤差を生ずる。また衛星搭載時計のずれによる誤差も生じる。GPSでは平均1m程度である。

#### マルチパス誤差

建物などほかの物体に反射した搬送波が受信されることで、条件によっては大きな誤差が生じる。都市部では可視衛星数の確保と並ぶ大きな問題である。

#### その他

その他、アンテナ位相中心偏移、受信機処理誤差、雑音などがある。

これらの誤差の内、電離層誤差、対流圏誤差、衛星軌道誤差と衛星時計誤差は距離の近い2地点間では違いが少ない。このため、DGPSやRTK-GPSでは、正しい位置が既知の基準点で得られた観測から、これらの誤差を補正することができる。しかし、基準点との距離が大きくなると、誤差を十分に相殺できず精度が低下する。これらの誤差は短時間では変化しないため、バイアス誤差として現れる。また、マルチパス誤差は測位を行う環境に大きく依存するため、補正することは難しい。

また、GPSによる測位は幾何学的な測量の一種であるため、天空における衛星の幾何学的配置が測位精度に影響する。この精度低下の程度を表す指標として精度劣化係数DOP(Dilution of Precision)がある。