

信頼度を考慮した GPS 測位情報の併用による動画像からのカメラ位置・姿勢推定

桑 秀行[†] 穴井 哲治^{††} 佐藤 智和[†]
 武 富貴史[†] 高地 伸夫^{††} 横矢 直和^{†(正会員)}

[†]奈良先端科学技術大学院大学, ^{††}株式会社トプコン研究開発センター

Extrinsic Camera Parameter Estimation from Video Images Considering GPS Positioning Confidence

Hideyuki KUME[†], Tetsuji ANAI^{††}, Tomokazu SATO[†],
 Takafumi TAKETOMI[†], Nobuo KOCHI^{††}, Naokazu YOKOYA^{†(Member)}

[†]Nara Institute of Science and Technology, ^{††}R&D Center, Topcon Corporation

あらまし 本論文では、広域屋外環境における動画像からのカメラ位置・姿勢推定において、測位の信頼度情報を取得可能な GPS を併用することで蓄積誤差を抑制する手法を提案する。従来、動画像からのカメラ位置・姿勢推定で用いられるバンドル調整において、GPS 測位位置を考慮してエネルギー関数を最小化することにより、動画像のみを用いた場合に生じる蓄積誤差を抑制する手法が提案されている。しかし、従来手法は、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 GPS 測位の信頼度を考慮しておらず、特に GPS 測位の信頼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った。また、GPS 受信機から長時間にわたって測位値が得られなかった区間（GPS 測位の欠損区間）後の最適化処理において、蓄積誤差が大きく、初期値が良くない場合には、局所解に陥ることで適切な解が得られないという問題があった。本研究ではこのような問題に対して、エネルギー関数に GPS 測位の信頼度に応じた重みを設定することで、GPS 測位の信頼度を考慮した最適化を行う。また、GPS 測位の欠損区間後の最適化処理では、エネルギー関数を最小化する前に、欠損区間前後の GPS 測位位置を用いて推定されているカメラ位置を補正することで、局所解を回避する。

キーワード：カメラ位置・姿勢推定, structure from motion, GPS

<Summary> This paper proposes a method for estimating extrinsic camera parameters using video images and sparsely acquired GPS data. Although the extended bundle adjustment methods which fuse vision and GPS data have been already proposed and they work finely if GPS data are accurately acquired, existing methods have the following problems: (1) The accuracy of estimated camera position largely depends on the confidence of GPS positioning data because they do not consider the GPS positioning confidence. (2) Solution often converges to local minima when GPS data can not be acquired for long time. In order to solve these problems, we add weighting coefficients depending on the GPS positioning confidence to an energy function for extended bundle adjustment. In order to avoid local minima, estimated camera positions without GPS data are fitted to GPS positions before optimization.

Keywords: extrinsic camera parameter estimation, structure from motion, GPS

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で利用されている。現在までに様々なカメラ位置・姿勢推定手法が提案されているが、なかでも動画像上で特徴点の再投影誤差を最小化することでカメラ位置・姿

勢を推定する手法は、高精度な推定が可能であるという特長から、画像を用いた処理や画像合成を前提とするアプリケーションの前処理として広く用いられている。一方で、近年これらの応用分野では広域な屋外環境への適用が試みられ、広域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。しかし、撮影された動画像のみからカメラの位置・姿勢を推定す

る Structure from Motion 法¹⁾や Visual SLAM 法²⁾、Visual Odometry 法³⁾には、推定されるカメラ位置・姿勢に誤差が蓄積するため、特に広範囲を移動する場合に推定精度が大きく低下するという問題がある。

このような誤差蓄積の問題に対して、動画像に加えて GPS を併用することで未知環境において誤差の蓄積を抑えたカメラ位置・姿勢を推定する手法が提案されている^{4)~11)}。これらの手法は情報の融合法によって、動画像からの推定結果を GPS 測位位置へ当てはめる手法^{4)~6)}、カルマンフィルタを用いる手法^{7)~9)}、バンドル調整を用いる手法^{10),11)}に分類できる。

動画像からの推定結果を GPS 測位位置へ当てはめる手法^{4)~6)}では、動画像から推定される相対的なカメラ位置・姿勢を全体⁴⁾、もしくは部分区間^{5),6)}ごとに回転、並進、スケール変換により GPS 測位位置へ当てはめることで絶対的なカメラ位置・姿勢を推定している。しかし、動画像からの推定による誤差の累積は一般に一樣ではないため、これらの手法を用いて広域を対象とした場合に生じる蓄積誤差を解消することは難しい。

カルマンフィルタを用いる手法^{7)~9)}は、動画像と GPS の情報をフレームごとの逐次的な処理において融合することで、蓄積誤差を抑えたカメラ位置・姿勢を推定する。このアプローチは、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ったフィルタの設計から高速な融合処理が行え、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。反面、入力動画像全体での最適化により精度を高めることは難しい。

バンドル調整を用いる手法^{10),11)}は、動画像からのカメラ位置・姿勢推定で用いられるバンドル調整において、再投影誤差と GPS 測位位置に関する誤差との総和として定義される誤差関数を最小化することにより、動画像と GPS の情報を融合し、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。横地ら¹⁰⁾は、カメラ位置・姿勢を推定することで算出される GPS 受信機の推定位置と GPS の測位位置の間の二乗距離を GPS に関する誤差関数として定義し、これをバンドル調整の誤差関数に加えたエネルギー関数を用いる手法を提案している。Lhuillier¹¹⁾は、横地らのエネルギー関数に対し、再投影誤差がしきい値より大きくならないという不等式制約を加えることで、GPS 測位位置に対する過剰な最適化を抑え、カメラ位置・姿勢と特徴点の三次元位置の間の整合性を保つ手法を提案している。しかし、動画像と GPS の融合という観点において、本質的には横地らの手法¹⁰⁾と同じであり、これらの手法^{10),11)}には共通する次の 2 つの問題がある。

- (i) 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 GPS 測位の信頼度が考慮されていないため、特に GPS 測位の信頼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。

- (ii) 最小化されるエネルギー関数には局所解が存在するため、GPS 受信機から長時間にわたって測位値が得られなかった区間 (GPS 測位の欠損区間) 後の最適化処理において、蓄積誤差が大きく、初期値が良くない場合には、GPS 測位の情報を推定結果全体に反映することが難しい。

本研究では、前者の問題について、従来提案されているエネルギー関数に、GPS 測位の信頼度に応じた重みを設定することで、GPS 測位の信頼度を考慮した最適化を行う¹²⁾。また、後者の問題について、長時間にわたる GPS 測位の欠損区間後の最適化処理では、エネルギー関数を最小化する前に、欠損区間前後の GPS 測位位置を用いて推定されているカメラ位置を補正することで、局所解を回避する。なお、本論文は文献 12) の内容を発展・改良したものである。ただし、GPS 測位位置を用いた初期値補正については、文献 12) に含まれておらず、本論文独自の内容である。

なお、本研究では、広域屋外を対象とした三次元形状復元や自由視点画像生成、拡張現実感などのアプリケーションにおける前処理として、オフラインでカメラ位置・姿勢を推定することを目的とする。

2. GPS 測位の信頼度を考慮したカメラ位置・姿勢推定

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 6 自由度のカメラ位置・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得られるものとする。

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¹⁰⁾を基礎とし、新たに GPS 測位の信頼度に応じた重みを設定したエネルギー関数を用いる。また、バンドル調整の実施前に動画像からのカメラ位置・姿勢の推定結果を GPS 測位値を用いて補正する。提案手法の処理の流れを図 1 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¹³⁾。ここで、現フレーム (第 f フレーム) において、GPS 測位が得られた場合 (GPS 測位時刻に対応する動画像のフレームの集合 G に f が含まれる場合)、現在推定されているカメラ位置と GPS 測位位置から最適化のための初期値を補正 (C) した後に、GPS の測位値を用いた狭区間最適化 (D) を行う。ただし処理 (C) は、得られた GPS 測位が欠損区間後の初めての GPS 測位であった場合 (欠損区間後の初めての GPS 測位に対応するフレームの集合 $G_{\text{recovered}}$ に f が含まれる場合) のみ実施する。これらの処理 (A)~(D) を初期フレームから最終フレームまで繰り返すことで、蓄積誤差を抑えたカメラの位置・姿

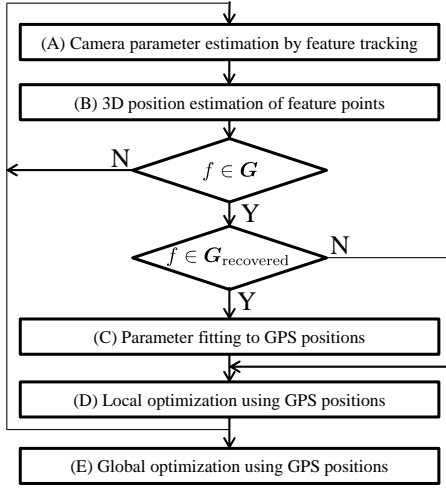


図 1 提案手法の処理の流れ． f はフレーム番号， G は GPS 測位時刻に対応する動画のフレームの集合， $G_{\text{recovered}}$ は欠損区間後の初めての GPS 測位に対応するフレームの集合を表す．

Fig. 1 Flow diagram of the proposed method. f is a frame number, G is a set of frames in which GPS data are observed, $G_{\text{recovered}}$ is a set of frames in which GPS positioning is recovered after GPS-denied section.

勢を逐次的に得た後，最後に動画全体での最適化処理を行う (E)．以下では，まず，本研究で新たに提案する GPS 測位の信頼度を考慮したエネルギー関数について述べる．次に，処理 (C)～(E) について詳述する．

2.1 GPS 測位の信頼度を考慮したエネルギー関数の定義

本研究では，従来のバンドル調整による動画と GPS の融合手法^{10),11)}と同様に，特徴点の再投影誤差と GPS 測位位置に関するエネルギーの重み付き和によって定義されるエネルギー関数を最小化することで，動画と GPS という性質の異なる 2 つの入力に関する誤差を同時に最小化するカメラ位置・姿勢および特徴点の三次元位置を推定する．ただし，提案手法では，従来手法^{10),11)}で用いられているエネルギーに，新たに GPS 測位の信頼度に応じた重みを設定することで GPS 測位の信頼度を考慮する．具体的には，エネルギー関数 E を，特徴点の再投影誤差に関するエネルギー Φ と GPS 測位位置に関するエネルギー Ψ を用いて次のように定義する．

$$E(\mathbf{T}, \mathbf{R}, \mathbf{P}) = \Phi(\mathbf{T}, \mathbf{R}, \mathbf{P}) + \omega \Psi(\mathbf{T}, \mathbf{R}) \quad (1)$$

ただし， \mathbf{T}, \mathbf{R} はそれぞれ第 i フレームにおけるカメラの位置・姿勢 t_i, r_i を全てのフレームについて並べたベクトルを表す．また， \mathbf{P} は特徴点 j の三次元位置 p_j を全ての特徴点について並べたベクトルを表す． ω は GPS 測位位置に関するエネルギーの重み係数である．

式 (1) を用いたエネルギー最小化の枠組みは，従来手法^{10),11)}と提案手法で共通する (従来手法¹¹⁾では，再投影誤差に関するエネルギーの定義は厳密には異なるが，本質的には従来手法¹⁰⁾と同様である) が，提案手法では，GPS 測位位

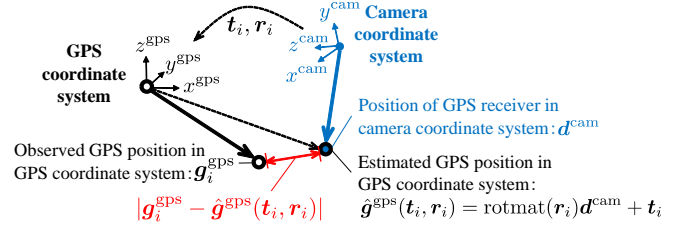


図 2 GPS 測位位置に関するエネルギー
Fig. 2 Energy term with respect to GPS positioning

置に関するエネルギー Ψ の定義が従来手法と異なる．以下，まず，再投影誤差に関するエネルギー Φ について述べ，次に各手法における GPS 測位位置に関するエネルギー Ψ の定義について詳述する．

2.1.1 再投影誤差に関するエネルギー

再投影誤差は，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．本研究では従来手法¹⁰⁾と同様に，再投影誤差に関するエネルギー Φ を次のように定義する．

$$\Phi(\mathbf{T}, \mathbf{R}, \mathbf{P}) = \frac{1}{\sum_{i \in \mathbf{F}} |S_i|} \sum_{i \in \mathbf{F}} \sum_{j \in S_i} \Phi_{ij}(t_i, r_i, p_j) \quad (2)$$

$$\Phi_{ij}(t_i, r_i, p_j) = \mu_j |q_{ij} - \text{proj}(t_i, r_i, p_j)|^2 \quad (3)$$

ただし， \mathbf{F} は動画のフレームの集合， S_i は第 i フレームの画像上で検出される特徴点の集合を表す． q_{ij} は第 i フレームにおける特徴点 j の画像上での検出座標， $\text{proj}(t_i, r_i, p_j)$ は特徴点 j の三次元位置 p_j を第 i フレームにおけるカメラ位置・姿勢 t_i, r_i を用いて画像上へ投影した座標である．また， μ_j は逐次処理で得られた特徴点 j の信頼度である¹³⁾．

2.1.2 GPS 測位位置に関するエネルギー

従来手法¹⁰⁾では，GPS の測位位置に関するエネルギーとして，図 2 に示すように，GPS 測位位置と GPS 受信機の推定位置の間の距離を用い，次のように GPS 測位位置に関するエネルギーを定義している．

$$\Psi_{\text{Yok}}(\mathbf{T}, \mathbf{R}) = \frac{1}{|\mathbf{G}|} \sum_{i \in \mathbf{G}} |g_i^{\text{gps}} - \hat{g}^{\text{gps}}(t_i, r_i)|^2 \quad (4)$$

$$\hat{g}^{\text{gps}}(t_i, r_i) = \text{rotmat}(r_i)d^{\text{cam}} + t_i \quad (5)$$

ただし， \mathbf{G} は GPS 測位時刻に対応する動画のフレームの集合を表す． g_i^{gps} は第 i フレームの GPS 測地座標系における GPS の測位位置， $\hat{g}^{\text{gps}}(t_i, r_i)$ はカメラ座標系における GPS 受信機の位置 d^{cam} を第 i フレームにおけるカメラ位置・姿勢 t_i, r_i を用いて GPS 測地座標系に変換することで算出される，GPS 測地座標系における GPS 受信機の推定位置を表す． $\text{rotmat}(r_i)$ は第 i フレームにおけるカメラ座標系から GPS 測地座標系への回転行列を表す．

従来手法¹¹⁾では，GPS として単独測位の GPS を想定しており，この場合には $|d^{\text{cam}}|$ に対して GPS 測位誤差が大きく

なるため、式 (4) で用いられている GPS 受信機の推定位置 \hat{g}^{GPS} をカメラ位置で代用し、 $\hat{g}_{\text{Lhu}}^{\text{GPS}}(t_i) = t_i$ と定義している。

提案手法では、従来手法では用いられて来なかった GPS 測位の信頼度に応じた重み τ_i を新たに設定し、次のように GPS 測位位置に関するエネルギーを定義する。

$$\Psi(\mathbf{T}, \mathbf{R}) = \frac{1}{|\mathbf{G}|} \sum_{i \in \mathbf{G}} \left(\frac{|g_i^{\text{GPS}} - \hat{g}^{\text{GPS}}(t_i, \mathbf{r}_i)|}{\tau_i} \right)^2 \quad (6)$$

ここで、GPS の精度に影響を与えるものとして、天空上の衛星配置に基づく精度低下率 (DOP), RTK-GPS における解の種類 (RTK-fix, RTK-float) などが考えられる¹⁴⁾が、次章で述べる実験では、解の種類に応じて変動する解の信頼度を実験的に求め、 τ_i を決定する。

2.2 GPS 測位値を用いた最適化

2.2.1 最適化の範囲

図 1 の処理 (A) (B) は、動画像上の情報のみを考慮するため、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生じる。従って、処理 (A) (B) を繰り返すだけでは、逐次処理の終了後に行う全体最適化 (E) において、初期値が真値から大きく離れるため、局所解に陥る可能性が高まる。そこで従来手法¹⁰⁾と同様に、逐次処理において範囲を限定した最適化を行う狭区間最適化処理 (D) と、逐次処理の終了後、動画像全体を対象とした最適化を行う全体最適化処理 (E) を実施する。処理 (D) では、逐次処理において、第 $(f-l)$ フレームから現フレーム (第 f フレーム) を用いて式 (1) で定義したエネルギー関数 E を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。処理 (E) では、処理 (A) ~ (D) 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、動画像全体を用いてエネルギー関数 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。

2.2.2 GPS 測位位置を用いた初期値補正

エネルギー関数 E には局所解が存在し、長時間にわたる GPS 測位の欠損区間においては、処理 (D) による狭区間最適化が働かず、逐次処理における蓄積誤差が大きくなる (図 3(a))。この結果、最適化に用いる GPS 推定位置の初期値が実際の GPS 測位位置から大きく外れることとなり、GPS 測位の情報を推定結果全体に反映することが難しくなる。具体的には、GPS 測位の欠損区間後に取得した GPS 測位を用いた最適化処理において、取得した GPS 測位に対応するフレームのカメラ位置およびその周辺のカメラ位置、特徴点の三次元位置のみが GPS 測位位置に引き寄せられ、局所解に陥ることで、大半の特徴点の三次元位置およびカメラの位置・姿勢が初期値付近から移動せず、蓄積誤差が解消されない (図 3(b))。本研究では、得られた GPS 測位が欠損区間後の初めての GPS 測位であった場合 (欠損区間後の初めての GPS 測位に対応するフレームの集合 $G_{\text{recovered}}$ に f が含まれる

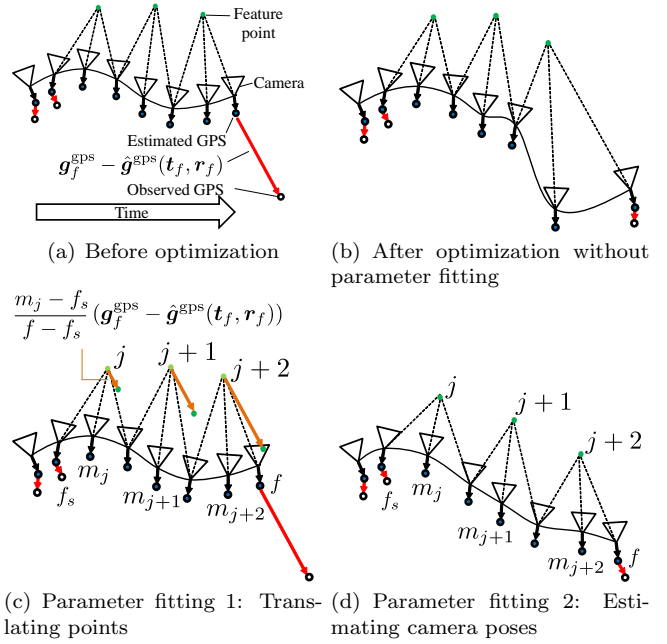


図 3 GPS 測位位置を用いた初期値補正
Fig. 3 Parameter fitting to GPS positions

場合), エネルギー関数を最小化する前に、推定されているカメラ位置・姿勢、特徴点の三次元位置を GPS 測位位置を用いて補正することで、最適化処理前に動画から得られるカメラの運動情報と GPS 測位値の整合性を高め、局所解を回避する。ここで、本補正処理は欠損区間後に得られた GPS 測位の信頼度が著しく低い場合を除き実施する。これは、一般に GPS 測位の欠損区間後の初めての GPS 測位は信頼度が低いことが多く、場合によっては信頼度の高い GPS 測位が得られないまま次の欠損区間に入ることがあるためである。また、欠損区間後に信頼度の高い GPS 測位が得られた場合には、後に行われる GPS 測位の信頼度を考慮した最適化処理により、信頼度の低い GPS 測位の影響は抑えられる。

ここで、初期値の補正に区間に対する回転、並進、スケール変換による GPS 測位位置への単純な当てはめ^{4)~6)}を用いると、欠損区間とそれ以前の GPS 測位を用いて推定が行われている区間間で不整合が生じることで、欠損区間の初期フレーム付近で再投影誤差が著しく増加し、良い初期値が得られない。そこで、欠損区間の最終フレームから欠損区間の初期フレームにかけて、補正によるカメラ位置・姿勢の変化を連続的に小さくする。具体的には、以下の手順で特徴点の三次元位置とカメラ位置・姿勢を補正することで、特定のフレームにおける再投影誤差の大幅な増加を抑えつつ、GPS に関する誤差を低下させる。

- (i) 欠損区間内で観測された特徴点 j の三次元位置 p_j を次のように更新する (図 3(c))。

$$p_j \leftarrow p_j + k_j (g_f^{\text{GPS}} - \hat{g}^{\text{GPS}}(t_f, \mathbf{r}_f)) \quad (7)$$

$$k_j = \begin{cases} \frac{m_j - f_s}{f - f_s} & ; m_j - f_s > 0 \\ 0 & ; \text{otherwise} \end{cases} \quad (8)$$

ただし, f_s は欠損区間の初期フレーム, m_j は特徴点 j が観測されているフレーム番号の中間値を表す.

- (ii) 欠損区間内のフレームについて, 手順 (i) で移動させた特徴点を用いて PnP 問題を解くことでカメラ位置・姿勢を算出する (図 3(d)).

なお, 本補正を行う場合には狭区間最適化に用いる区間を GPS 測位が得られなかった区間全体 (第 f_s フレームから現フレーム (第 f フレーム)) とする.

3. 実験

本節では, 提案手法における GPS 測位の信頼度に応じたエネルギー項に対する重みの効果と, 最適化のための初期値補正の効果を検証するために, 信頼度の低い GPS 測位が多く含まれるデータ (実験データ 1), 長時間にわたる GPS 測位の欠損区間を含むデータ (実験データ 2), および対応する動画像を用いてカメラ位置・姿勢の推定を行い, 従来手法¹⁰⁾との位置推定精度の定量的な比較を行う. 以下では, まず, 2つの実験に共通する実験条件, 提案手法で用いる GPS 測位の信頼度に応じた重み τ_i の決定方法について述べた後, それぞれのデータを用いた実験について詳述する.

3.1 実験条件

本実験では, 車両に取り付けたビデオカメラ (Point Grey Research Grasshopper2) を用いて, 約 0.9km の路上区間を走行しながら撮影した 194 秒間の動画像 (解像度 1600×1200 画素, プロGRESSIVE 撮影, 2756 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, カメラ位置・姿勢の推定を行った. 今回の実験では, 上空が開けた環境において計測を行ったため, 約 91% の GPS 測位値が信頼度の高い RTK-fix として取得されたが, 後述する実験ではこれらを真値として扱う. 入力とする GPS 測位値 (実験データ 1, 2) 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, 仮想的に複数の GPS 衛星が遮蔽されていたとみなし, それらの GPS 衛星からの信号を用いずに測位を行うことで作成した. 各データの詳細はそれぞれ 3.3 節, 3.4 節で述べる. ここで用いられる GPS 衛星からの電波は, 実際に取得されたものであり, 建物による反射や樹木による減衰など現実環境における様々な測位精度の低下要因を含んでいる. 位置推定精度は, 真値と GPS の推定位置 $\hat{g}^{\text{GPS}}(t_i, r_i)$ を比較することで算出した.

図 4 に入力画像の一部を示す. 同図に示すように, 本実験は人工物と自然物を共に含む, 一般的な市街地環境下で実施し, 本実験における車両の運動は, 右左折や坂道の登り下り, 旋回による傾きなどの一般的な 6 自由度の運動を含んでいる. また, 本実験では, 実験の施行ごとの逐次処理における特徴点追跡精度の違いを排除するために, あらかじめ佐藤らの手法¹³⁾を適用することで得た, 誤対応を自動で排除した特徴点



図 4 入力画像の例
Fig. 4 Examples of input images

の追跡結果を比較するすべての手法に共通して用いた.

なお, カメラと GPS 受信機は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^{cam} およびカメラの内部パラメータはあらかじめキャリブレーション済みで既知とした. また, 動画像と GPS の測位値は手動で開始時刻を同期させ, 狭区間最適化の使用フレームを $l = 200$ とした. 文献 10) に示されているシミュレーション実験の結果より, 式 (1) で用いられる GPS 測位位置に関するエネルギー項の重みを $\omega = 10^{-9}$ に設定した. これらのパラメータは, 比較する手法で共通して用いた. エネルギー関数の最小化には, E のヤコビ行列が疎になることから, 疎な行列を考慮した Levenberg-Marquardt 法 (sparseLM)¹⁵⁾を用いた.

3.2 GPS 測位の信頼度に応じた重み τ_i の決定

本実験では, GPS として RTK-GPS を用いるが, この場合には解の種類 (RTK-fix, RTK-float) 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, これを用い, 式 (6) で用いられる GPS 測位の信頼度に応じた重み τ_i を次のように定義する.

$$\tau_i = \begin{cases} 1 & ; i \in G_{\text{fix}} \\ \tau_{\text{float}} & ; \text{otherwise} \end{cases} \quad (9)$$

ただし, G_{fix} は信頼度の高い RTK-fix 解の GPS 測位値に対応するフレームの集合である. また, 信頼度の低い RTK-float 解の GPS 測位値に対する重み τ_{float} は, 事前に GPS 測位の定点観測を行うことによって取得した信頼度の高い GPS 測位位置の RMS rms_{fix} と信頼度の低い GPS 測位位置の RMS rms_{float} から $\tau_{\text{float}} = rms_{\text{float}}/rms_{\text{fix}} = 107.4$ と決定した.

3.3 信頼度が低い GPS 測位値を多く含むデータに対する実験 (実験 1)

本実験では, 2.1.2 で述べた信頼度に応じた重みの有効性を確認するために, 以下の複数の手法によりカメラ位置・姿勢推定を行い, 位置推定精度を比較する.

手法 A: GPS 測位の信頼度を考慮しない従来手法¹⁰⁾

手法 B: 手法 A において信頼度の高い RTK-fix の測位値のみを用いる手法

手法 C: GPS 測位の信頼度を考慮した提案手法

手法 D: GPS 測位値を用いず, 動画像のみから推定を行う手法

表 1 位置誤差の比較 (実験 1) [m]
Table 1 Comparison of position errors (Experiment 1) [m]

Method	Average	Std. dev.	Max
RTK-fix	0.003	0.003	0.010
RTK-float	1.844	1.093	6.641
A	1.553	1.182	6.326
B	4.298	3.944	14.041
C	1.217	1.070	4.073
D	37.189	13.033	69.388

表 2 狭区間最適化の使用フレーム数 l と平均位置誤差の関係 (実験 1) [m]
Table 2 Relationship between l and average position errors (Experiment 1) [m]

l	50	100	200	400	800
Avg. error	13.431	1.284	1.217	1.232	1.219

ただし、手法 A, B, C は全て著者による実装であり、違いは GPS に関するエネルギー関数の定義のみである。また、手法 D は GPS 測地座標系において位置誤差を比較するため、推定終了後に真値に対して平行移動、回転、スケール変化による当てはめを行う。

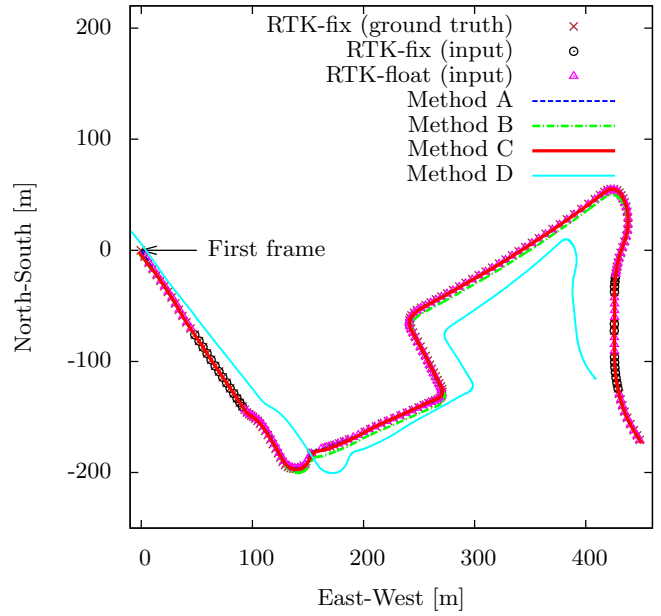
ここでは、入力として、一般市街環境における建物や樹木などによる GPS 衛星の遮蔽の結果生じる GPS 測位の信頼度の低下を想定し、スポット的に得られる信頼度の高い RTK-fix の測位値と大部分 (全 GPS 測位の約 86%) を占める信頼度の低い RTK-float の測位値からなるデータを用いる。

3.3.1 位置推定精度の定量的な比較

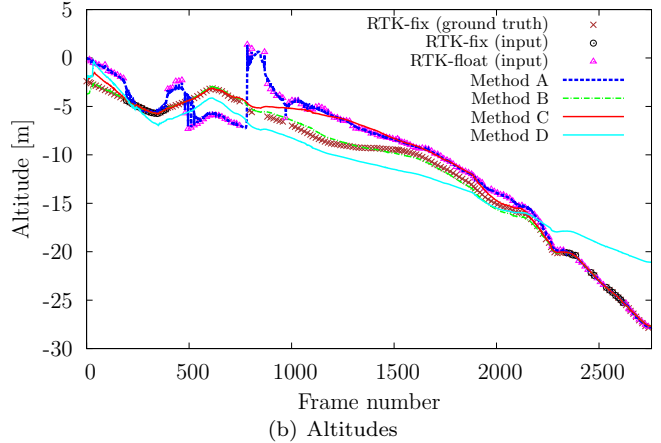
各手法により得られた GPS 推定位置を図 5 に示す。真値との比較による、GPS 推定位置の誤差および入力として用いた GPS 測位位置の誤差について、フレームごとの位置誤差を図 6 に、位置誤差の平均、標準偏差、最大値を表 1 に示す。図 6 より、GPS 測位の信頼度を用いない手法 A は、信頼度の低い RTK-float の測位値に大きく影響されていることがわかる。また、手法 B は、GPS 測位値を用いていない 500 ~ 2,000 フレーム付近の誤差が大きく、信頼度の高い GPS 測位値のみを用いるだけでは蓄積誤差を十分に解消できないことがわかる。これらに対して提案手法 (手法 C) は、蓄積誤差を抑えつつ、RTK-float による誤差も抑制しており、表 1 より、手法 A, B と比較して最も精度良く GPS 受信機の位置を推定できていることがわかる。これは、GPS 測位の信頼度に応じた重みの効果であると考えられる。また、図 5 より、動画のみを用いる手法 D は蓄積誤差の影響が大きく、蓄積誤差は GPS 測位位置に対する平行移動、回転、スケール変化による当てはめだけでは解消されないことがわかる。

3.3.2 パラメータに関する考察

提案手法 (手法 C) において、狭区間最適化の使用フレーム数 l を 50, 100, 200, 400, 800 と変化させカメラ位置・姿勢推定を行った場合の平均位置誤差を表 2 に示す。表 2 から、 $l = 50$ の場合には誤差が大きく、それ以外の場合には大きな差がないことがわかる。このことから、狭区間最適化の使用



(a) Horizontal 2D positions



(b) Altitudes

図 5 推定された GPS 受信機の位置 (実験 1)
Fig. 5 Estimated GPS positions (Experiment 1)

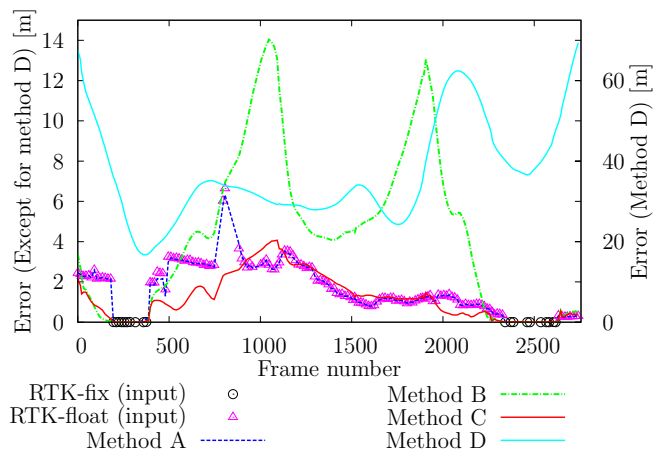


図 6 フレームごとの位置誤差 (実験 1)

Fig. 6 Position errors in each frame (Experiment 1)

フレーム数が短すぎる場合には、逐次処理において誤差を十分に解消できないが、それ以外の場合には精度に大きく影響しないことが分かる。

手法 C において、信頼度の低い GPS 測位の重み τ_{float} を

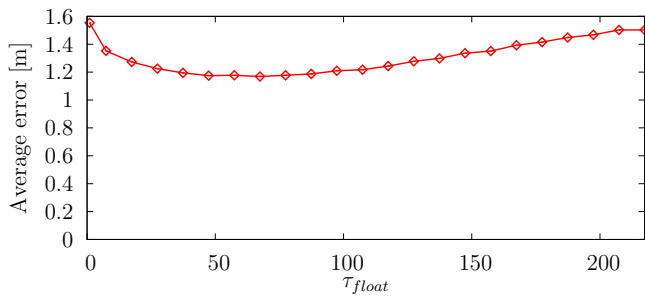


図 7 重み τ_{float} と平均位置誤差の関係 (実験 1)

Fig. 7 Relationship between weight τ_{float} and average position errors (Experiment 1)

実験で用いた $\tau_{float} = 107.4$ を中心として変化させた場合の平均位置誤差を図 7 に示す. $\tau_{float} = 1$ の場合は, RTK-fix と RTK-float を区別しない従来手法¹⁰⁾に対応する. 図 7 から, 定点観測で得られた τ_{float} は最適なものではなかったが, τ_{float} を変化させても精度に大きな影響はなく, 対象とするシーンでの定点観測を行い決定した τ_{float} を用いることの有効性が確認できる.

3.4 長時間にわたる GPS 測位の欠損区間を含むデータに対する実験 (実験 2)

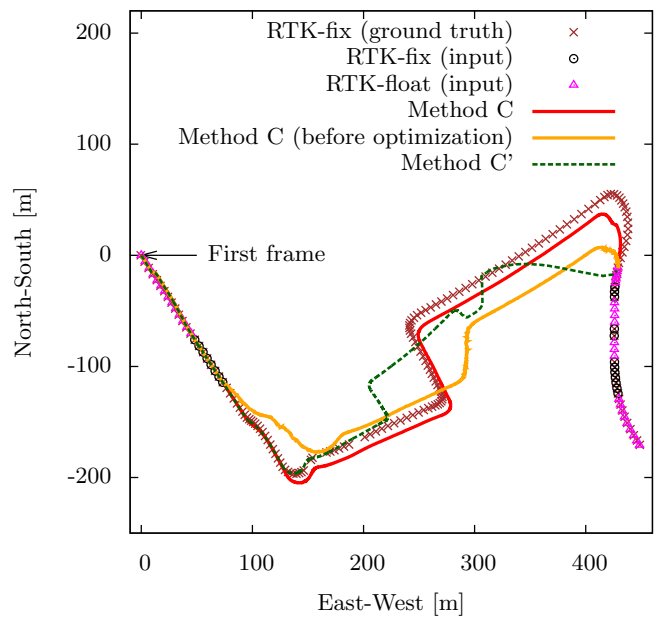
本実験では, 2.2.2 で述べた初期値補正手法の有効性を確認するため, 以下の二手法によりカメラ位置・姿勢推定を行い, 位置推定精度を比較する.

手法 C: 提案手法

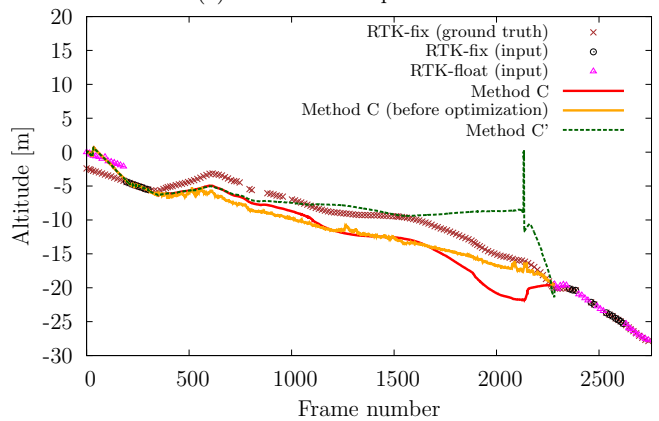
手法 C': 提案手法において初期値補正を行わない場合

ここでは, 入力として, 一般市街環境における建物や樹木などによる GPS 衛星の遮蔽の結果生じる GPS 測位の欠損を想定し, このような想定下では十分に長時間であると考えられる GPS 測位の欠損区間 (139 秒間, 315 ~ 2,271 フレーム, 全区間の 72%) を含むデータを用いる. また本実験では GPS として RTK-GPS を用いるが, この場合には信頼度の低い RTK-float 解であっても欠損区間における蓄積誤差に対して GPS 測位位置の誤差が十分に小さいと想定されるため, 初期値補正は GPS 測位が RTK-fix 解, RTK-float 解のいずれの場合にも実施する.

手法 C, C' を用い, GPS 測位の欠損区間後に得られた最初の GPS 測位を用いて推定を行った結果を図 8 に示す. 手法 C については 2.2.2 で述べた補正処理直後の位置と最適化後の結果を共に示している. また, 手法 C, C' により全体最適化後に得られた GPS 推定位置と真値との比較による位置誤差および入力として用いた GPS 測位位置の誤差について, フレームごとの位置誤差を図 9 に, 位置誤差の平均, 標準偏差, 最大値を表 3 に示す. 図 8(a), (b) は軸のスケールが大幅に異なることに注意が必要であるが, 同図より, 最適化のための初期値補正を行わない手法 C' は, GPS 測位に対応するフレームのカメラ位置およびその周辺のカメラ位置のみが GPS 測位位置に引き寄せられ, 欠損区間のフレームにおける蓄積誤差は解消されていないことがわかる. これに対



(a) Horizontal 2D positions



(b) Altitudes

図 8 欠損区間後の推定された GPS 受信機的位置 (実験 2)

Fig. 8 Estimated GPS positions after GPS-denied section (Experiment 2)

して, 手法 C では, 欠損区間後の初めての GPS 測位は信頼度の低い RTK-float 解であったが, 蓄積誤差に対して GPS 測位の誤差は小さく, 初期値補正により真値に近い推定結果が得られている. また, その後, 欠損区間に対する最適化を行うことで, 推定結果全体においてより真値に近い経路上に推定結果が得られている.

図 9 より, 手法 C' では GPS 測位の欠損区間および最終フレーム付近において誤差が大きくなっている. 前者については, 上述したように, GPS 測位の情報を GPS 欠損区間に反映できなかったためであると考えられる. また, 後者については, GPS 測位の欠損区間後の最適化によって, GPS 測位に対応するフレームのカメラ位置およびその周辺のカメラ位置のみが GPS 測位位置に引き寄せられたため, 動画像からの推定において, 特徴点の三次元位置とカメラ位置・姿勢の間の整合性が取れなくなり, 動画像からの推定の誤差が大きくなったためであると考えられる. これに対して, 手法 C では, 動画像の初期フレームと最終フレーム付近に存在する

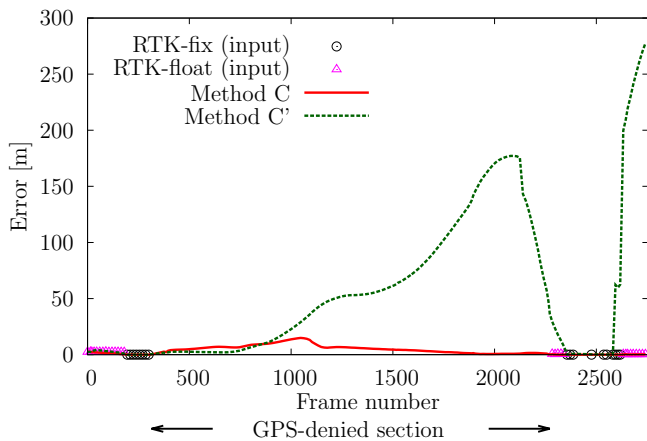


図9 フレームごとの位置誤差 (実験2)

Fig. 9 Position errors in each frame (Experiment 2)

表3 位置誤差の比較 (実験2) [m]
Table 3 Comparison of position errors (Experiment 2) [m]

Method	Average	Std. dev.	Max
RTK-fix	0.006	0.009	0.038
RTK-float	1.315	0.989	2.594
C	3.593	3.617	14.935
C'	60.299	67.829	277.530

GPS 測位値を用いて、動画全体で蓄積誤差を最小化している。ただし、今回の実験では欠損区間が長時間にわたるため、提案手法による結果が、前節の実験における信頼度の高いGPS測位のみを用いる手法Bと同程度の精度となっており、このような長い欠損でも精度の高いカメラ位置・姿勢推定を実現しようとする場合には、今後、他のセンサの併用も検討する必要がある。

4. まとめ

本論文では、バンドル調整によるGPS測位情報と動画からのカメラ位置・姿勢推定において、GPS測位の信頼度が低い場合に推定精度が低下するという問題に対して、従来手法で提案されているエネルギー関数にGPS測位の信頼度に応じた重みを設定することで問題を解決する手法を提案した。また、初期値によってGPS測位の情報を推定結果全体に反映することが難しい場合においても、エネルギー関数を最小化する前に、カメラ位置とGPS測位位置を用いて初期値を補正することで、局所解を回避する手法を提案した。実験では、屋外環境で取得した動画とGPSの測位値を用いてカメラ位置・姿勢を推定し、従来手法との推定精度の定量的な比較を行った。実験の結果、提案手法は従来手法より精度良くGPS受信機の位置を推定できていることから、カメラ位置も高精度に推定できていると考えられる。

本研究の課題として、エネルギー関数で用いられる重みの決定法があげられる。最適な重みは環境によって変化すると考えられるため、実験で用いた重みを他の環境に適用した場合に精度が低下する可能性がある。これに対し、今後、環境に

応じた重みの自動決定法の開発を行う。また、長時間にわたるGPS測位の欠損区間における蓄積誤差は依然として大きい。これを解消するために他のセンサの併用を検討する。

謝 辞

本研究の一部は文部科学省特別経費「アンビエント環境知能研究創出事業」、日本学術振興会科学研究費補助金基盤研究(A) No. 23240024による。

参考文献

- 1) M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, R. Koch: "Visual Modeling with a Hand-held Camera", *Int'l J. of Computer Vision*, Vol. 59, No. 3, pp. 207-232 (2004).
- 2) G. Klein, D. Murray: "Parallel Tracking and Mapping for Small AR Workspaces", *Proc. of Int'l. Symp. on Mixed and Augmented Reality*, pp. 225-234 (2007).
- 3) D. Nistér, O. Naroditsky, J. Bergen: "Visual Odometry", *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 652-659 (2004).
- 4) 仲野潤一, 西村邦裕, 谷川智洋, 廣瀬通孝: "Visual SLAM と GPS を用いた位置情報付き画像による都市空回記録の研究", *電子情報通信学会論文誌 (D)*, Vol. J93-D, No. 6, pp. 901-908 (2010).
- 5) Y. Bok, Y. Jeong, D. G. Choi, I. S. Kweon: "Capturing Village-level Heritages with a Hand-held Camera-laser Fusion Sensor", *Int'l J. of Computer Vision*, Vol. 94, No. 1, pp. 36-53 (2011).
- 6) L. Wei, C. Cappelletti, Y. Ruichek, F. Zann: "GPS and Stereovision-based Visual Odometry: Application to Urban Scene Mapping and Intelligent Vehicle Localization", *Int'l J. of Vehicular Technology*, Vol. 2011, 17 pages (2011).
- 7) M. Pollefeys, D. Nistér, J. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S. Kim, P. Merrell, C. Salmi, S. Sinha, B. Talton, L. Wang, Q. Yang, H. Stewénius, R. Yang, G. Welch, H. Towles: "Detailed Real-time Urban 3D Reconstruction from Video", *Int'l. J. of Computer Vision*, Vol. 78, No. 2-3, pp. 143-167 (2008).
- 8) M. Agrawal, K. Konolige: "Real-time Localization in Outdoor Environments using Stereo Vision and Inexpensive GPS", *Proc. of Int'l Conf. on Pattern Recognition*, pp. 1063-1068 (2006).
- 9) D. Schleicher, L. M. Bergasa, M. Ocana, R. Barea, E. Lopez: "Real-time Hierarchical GPS Aided Visual SLAM on Urban Environments", *Proc. of Int'l Conf. on Robotics and Automation*, pp. 4381-4386 (2009).
- 10) 横地裕次, 池田聖, 佐藤智和, 横矢直和: "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).
- 11) M. Lhuillier: "Incremental Fusion of Structure-from-motion and GPS Using Constrained Bundle Adjustments", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 34, No. 12, pp. 2489-2495 (2012).
- 12) T. Anai, N. Fukaya, T. Sato, N. Yokoya, N. Kochi: "Exterior Orientation Method for Video Image Sequences with Considering RTK-GPS Accuracy", *Proc. of Int'l Conf. on Optical 3-D Measurement Techniques*, pp. 231-240 (2009).
- 13) 佐藤智和, 神原誠之, 横矢直和, 竹村治雄: "マーカと自然特徴点の追跡による動画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).

- 14) 臼井澄夫: “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).
- 15) M. I. Lourakis: “Sparse Non-linear Least Squares Optimization for Geometric Vision”, Proc. of European Conf. on Computer Vision, pp. 43-56 (2010).

(2013年7月4日 受付)

(2013年10月15日 再受付)



武 富 貴 史

2006年佐世保高専・専攻科・電気電子工学専攻卒。2011年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。現在、同大情報科学研究科助教。博士(工学)。



桑 秀 行

2008年同志社大学工学部知識工学科中退。2010年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。2011~2012年カーネギーメロン大学滞在。現在、奈良先端科学技術大学院大学情報科学研究科博士後期課程に在学中。



高 地 伸 夫

1979年東京都立大学工学部卒業。同年(株)トプコン入社。測量, 眼科医用, 半導体画像処理の研究開発に従事。現在三次元画像計測, 点群処理に関する研究開発に携わる。2013年北海道大学大学院博士後期課程修了。研究開発センター研究主幹。博士(情報科学)。



穴 井 哲 治

1997年東京電機大学理工学研究科建設工学専攻修士課程修了。2000年同大学博士後期課程満期退学。同年同大学建設環境工学科嘱託助手。2005年株式会社トプコン入社。現在に至る。



横 矢 直 和 (正会員)

1974年大阪大学基礎工学部情報工学科卒。1979年同大学院博士後期課程修了。同年電子技術総合技術研究所入所。1986~87年マッギル大・知能機械研究センター客員教授。1992年奈良先端科学技術大学院大学・情報科学センター教授。現在、同大理事・副学長, 情報科学研究科教授。工学博士。



佐 藤 智 和

1999年大阪府立大工学部情報工学科卒。2003年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。同年同大情報科学研究科助教。2010~11年チェコ工科大客員研究員。現在、奈良先端科学技術大学院大学情報科学研究科准教授。コンピュータビジョンの研究に従事。博士(工学)。