

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

横地 裕次[†], 池田 聖[†]
佐藤 智和[†], 横矢 直和[†]

本論文では、屋外環境下で広範囲にカメラを移動し取得した動画と GPS による位置情報からカメラの外部パラメータを推定する手法を提案する。提案手法は、動画中の特徴点の追跡により外部パラメータを推定する手法を基に、GPS による位置情報を用いた外部パラメータの最適化により、累積的な推定誤差を軽減するものである。さらに、GPS の位置情報を用いて推定された暫定的な外部パラメータを特徴点の追跡処理にフィードバックすることにより、特徴点の追跡自体についても精度向上を図る。シミュレーションによる精度評価実験および屋外実環境下で取得した動画と GPS による位置情報を用いた実験により、環境中の三次元位置が既知な点を用いて累積誤差を軽減する手法と同程度の精度で外部パラメータが推定可能であることを確認した。

Estimation of Extrinsic Camera Parameters Based on Feature Tracking and GPS Data

YUJI YOKOCHI,[†] SEI IKEDA,[†] TOMOKAZU SATO[†]
and NAOKAZU YOKOYA[†]

This paper describes a novel method for estimating extrinsic camera parameters using an image sequence and sparse position data acquired by GPS. This method is based on tracking a number of image features and optimizing extrinsic parameters using position data from GPS. The position data are actually used to prevent the algorithm from incorrectly tracking image features. Moreover, the position data are also used for optimizing extrinsic parameters over the whole image sequence. The proposed method allows us to estimate extrinsic parameters accurately with little accumulated estimation errors even from a long image sequence. The scale of estimated extrinsic parameters is automatically determined by position data acquired by GPS. The validity of the proposed method is demonstrated through experiments of estimating extrinsic parameters for both synthetic and real outdoor scenes.

1. はじめに

現実環境を撮影した動画からのカメラ外部パラメータの推定は、複合現実感、三次元形状復元、新視点画像生成など様々な分野における基礎技術として利用されている。これらの応用分野では、近年、広範囲な環境への手法の適用が試みられ^{1),2)}、広範囲に移動するカメラの外部パラメータを精度良く推定することが求められている。しかし、動画のみから外部パラメータを推定する手法^{3),4)}では、外部パラメータの推定結果に蓄積誤差が発生するため、広範囲に移動する

場合に著しく推定精度が低下する問題がある。また、推定される外部パラメータは、フレーム間での相対的な位置・姿勢関係を表すものであり、スケールについては不定である。そのため、複数回に分けて撮影した動画の外部パラメータの統合や、閉曲線上の経路で撮影した動画の外部パラメータ推定のように、同一シーンが連続しないシーケンス上に現れる場合、推定されるパラメータ間の不整合が問題となる。

これらの問題は、従来から、動画と現実環境に関する事前知識の使用^{5)~8)}や、動画を問わずに複数のセンサから得られる値を直接外部パラメータとすること⁹⁾により解決されてきた。環境の事前知識を用いる手法では、三次元的な位置関係が既知のランドマーク群^{5),6)}や、三次元 CAD モデル^{7),8)}が用いられている。文献^{5),6)}の手法では、三次元位置が既知のランドマークと三次元位置が未知の画像特徴点（以後、特徴点）をフレーム毎に追跡し、その再投影誤差を最小化

[†] 奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute
of Science and Technology
現在、株式会社 本田技術研究所 栃木研究所
Presently with Tochigi R&D Center, Honda R&D Co.,
Ltd.

することで外部パラメータを推定する．文献^{7),8)}の手法では，現実環境の三次元 CAD モデルの輪郭と入力画像上のエッジを対応付けることで外部パラメータを推定する．これらの手法では，ランドマークや三次元 CAD モデルの現実環境での三次元位置情報を用いることで蓄積誤差の問題を解決している．しかし，ランドマークや三次元 CAD モデルに含まれる三次元位置情報を屋外で広範囲に取得することは困難であり，また，基準とする三次元情報の獲得のためには，人手による作業が必要であるという問題がある．

一方，RTK-GPS，磁力計，2軸角度センサーの値から算出される値を，直接動画のカメラ外部パラメータとする手法⁹⁾では，RTK-GPSを用いることによって，誤差の蓄積しない数 cm 精度の位置情報（以後，GPS 測位値）を獲得できる．しかし，センサの組み合わせによってカメラの外部パラメータを推定する手法では，カメラとセンサ類の位置関係が必要であるが，これを正確に計測することは難しい．また，一般的な GPS では，位置情報の取得レートが動画より低い 1Hz であり，カメラの運動に含まれる高周波成分を復元することは難しい．

そこで本論文では，これらの問題を解決するために，屋外環境下を広範囲に移動するカメラで取得した動画と GPS 測位値からカメラ外部パラメータを推定する手法を提案する．提案手法では，動画中の特徴点の追跡と GPS 測位値を用いた最適化に基づきカメラ外部パラメータを推定する．特徴点の追跡では，ランドマーク・特徴点追跡による手法⁶⁾と同様に，特徴点の仮対応づけから推定される暫定的な外部パラメータを用いて特徴点の画像上での探索範囲を限定し，誤追跡を防ぐ．GPS 測位値を用いた最適化では，動画から推定される外部パラメータに基づく特徴点の再投影誤差及び外部パラメータから推定される GPS 受信機の位置の推定誤差を同時に最小化することで誤差の蓄積を防ぐ．本手法では，現実環境に関する事前知識や，GPS 受信機とカメラの間の正確なキャリブレーションを必要とせず，GPS の測位座標系における外部パラメータを推定できる特長がある．そのため，乗り物などで屋外を広範囲に移動しながら撮影した動画に CG を重畳表示する，景観シミュレーションや映画の特撮などへの利用法が考えられる．

ただし，本手法では，動画全体にわたりカメラに対する GPS 受信機の位置は固定されているものとし，容易に計測できる両センサ間の距離については既知，カメラ座標系における GPS 受信機の方位については推定するものとする．また，本論文で扱う全てのカ

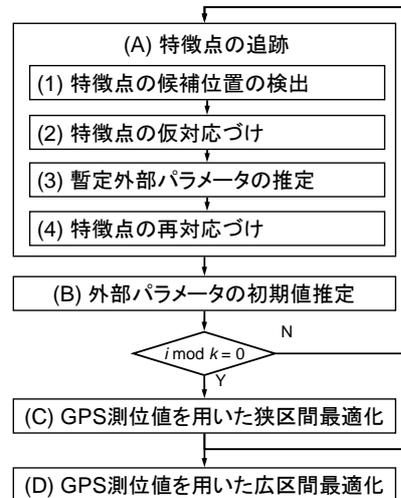


図 1 提案手法の処理手順

Fig. 1 Procedure of the proposed method.

メラの内部パラメータは既知であり，予めカメラのレンズ歪みや焦点距離などを考慮したキャリブレーション¹⁰⁾が行われているものとする．動画と GPS 測位値の同期についても得られているものとする．

以下，2章では，提案手法の流れ及び各処理について詳述する．3章では提案手法の有効性を確認するためにシミュレーションによる実験及び実環境のデータを用いた実験を示し，最後に4章で本論文をまとめる．

2. GPS 測位値を用いた外部パラメータ推定

提案手法では，計算コストの問題から従来手法による一括処理が難しいような広範囲にカメラが移動する環境を想定し，動画と GPS 測位値から，特徴点の三次元位置，カメラ座標系における GPS 受信機の方位，カメラの外部パラメータを推定する．本手法は，主に特徴点追跡と外部パラメータの最適化処理からなり，両処理に GPS の位置情報を反映させることで，誤差の蓄積を防止する．

本手法の流れを図 1 に示す．まずフレーム毎の処理として，(A) 特徴点の追跡，(B) カメラ外部パラメータの初期値推定を行い，次に一定フレーム間隔 k で (C) GPS 測位値を用いた最適化を行う．処理 (C) では，蓄積誤差の影響によって後続のフレームで特徴点の対応付けの精度が低下することを防ぐために狭区間での最適化を行う．処理 (D) では，処理 (C) よりも広い区間で多くの GPS 測位値を用いて外部パラメータの最適化を行う．以下では，まず処理 (C)，(D) で共通に用いる GPS 測位値を用いた誤差関数を定義し，その最適化手法について述べる．次に処理 (A) ~ (D)

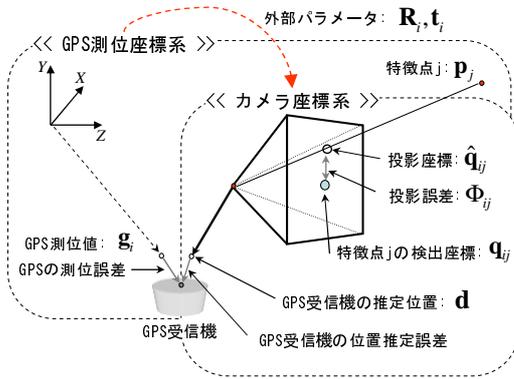


図 2 座標系と誤差の定義

Fig. 2 Coordinate systems of camera and GPS.

についてそれぞれ詳述する。

2.1 GPS 測位値を用いた誤差関数の定義とその最適化

本節では、まず、従来の特徴点のみによるカメラパラメータ推定で一般的に用いられる再投影誤差について説明する。次に、カメラと GPS 受信機の幾何学的関係をモデル化し、GPS 受信機に関する誤差 Ψ を定義する。最後に、GPS 受信機に関する誤差 Ψ と画像の特徴点の再投影誤差 Φ を組み合わせて、新たな誤差関数 E を定義し、その最小化方法について説明する。

特徴点の再投影誤差の定義

再投影誤差は、特徴点追跡に基づく外部パラメータ推定手法で、一般的に用いられる誤差関数であり^{3),4),6)}、再投影誤差の二乗和を最小化することでカメラパラメータを推定する手法は、バンドル調整法と呼ばれている¹¹⁾。特徴点 j の第 i フレーム上での再投影誤差 Φ_{ij} は、推定された三次元位置の画像上への投影座標 \hat{q}_{ij} と画像上で検出される座標 q_{ij} を用いて、以下のように表される。

$$\Phi_{ij} = |q_{ij} - \hat{q}_{ij}| \quad (j \in S_i) \quad (1)$$

ただし、 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。

GPS に関する誤差の定義

GPS の測位誤差には、真の GPS 受信機の位置を中心に発生する期待値 0 の局所的な測位誤差と平均観測位置が真値から外れる大域的な誤差の 2 種類が考えられる。本論文では、前者の局所的な測位誤差のみを扱う。

一般に、GPS 測位値と外部パラメータの推定結果に誤差が含まれない場合、第 i フレームの外部パラメータ R_i, t_i とそれに対応する GPS 測位値 g_i 及び、カメラ座標系での GPS 受信機の位置 d には、以下の関

係が成り立つ。

$$R_i g_i + t_i = d \quad (i \in \mathcal{F}) \quad (2)$$

ただし、 \mathcal{F} は、GPS の計測時刻に対応する動画のフレームの集合を表す。ここで、図 2 に示すように、GPS 測位値 g_i 及び外部パラメータ R_i, t_i に誤差があると、式 (2) は成り立たない。そこで本研究では、外部パラメータ R_i, t_i から推定される GPS 受信機の位置と GPS 測位値との差を GPS 受信機に関する誤差関数 Ψ_i とし、以下のように定義する。

$$\Psi_i = |R_i g_i + t_i - d| \quad (3)$$

特徴点と GPS による誤差の定義とその最適化

ここでは、外部パラメータ推定に用いる最適化の誤差関数 E を、式 (1) で示した画像上の再投影誤差 Φ_{ij} 及び、式 (3) の GPS 位置の誤差 Ψ_i を用いて以下のように定義する。

$$E = \frac{\omega}{|\mathcal{F}|} \sum_{i \in \mathcal{F}} \Psi_i^2 + \frac{1}{\sum_i |S_i|} \sum_i \mu_i \sum_{j \in S_i} w_j \Phi_{ij}^2 \quad (4)$$

ただし、 w_j は逐次処理で得た特徴点 j の信頼度であり、再投影誤差 Φ_{ij} の分散の逆数とする。 μ_i は各フレームに対する重み係数を表す。 ω は Ψ_i 及び Φ_{ij} に対する重み係数であり、予め、特徴点の検出位置精度および GPS の測位精度を考慮し、後述するシミュレーションにより、真値に対する誤差が最小になるような ω を算出しておく。一般に、シミュレーションと実利用時の環境では、GPS の測位回数 $|\mathcal{F}|$ と特徴点数 $\sum_i |S_i|$ が異なるので、本研究では、これらの値に依らず重み係数 ω を一定値に決めるために、 $|\mathcal{F}|$ 及び $\sum_i |S_i|$ により右辺の各項を正規化する。実利用時において撮影対象までの奥行きや使用する GPS の測位精度がこのシミュレーションと大きく異なる場合は、重み係数 ω を求め直す必要がある。

式 (4) の誤差関数 E を用いた最適化では、外部パラメータ R_i, t_i 、特徴点の三次元位置 p_j 、及びカメラ座標系における GPS 受信機の方位をパラメータとして、誤差関数 E を勾配法によって最小化する。ただし、この誤差関数 E には、局所解が多数存在するため、単純な勾配法では、大域最小解を得ることが難しい。そこで、最適化の各フレームの重み μ_i を変化させることで、局所解を回避する。具体的には、GPS 測位値があるフレームの重みを大きく設定することおよび、すべてのフレームの重みを等価に設定することを、最適化処理が収束するステップごとに交互に繰り返す。誤差関数 E の重み μ_i を変化させても大域最小解を与えるパラメータに大きな差はなく、局所解の位置のみが大きく変化するため、局所解を回避できると考えられる。

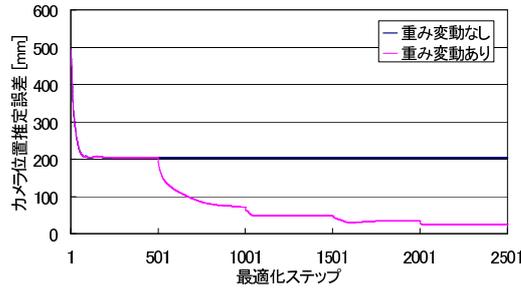


図 3 重み変動の効果
Fig. 3 Effect of changing weight.

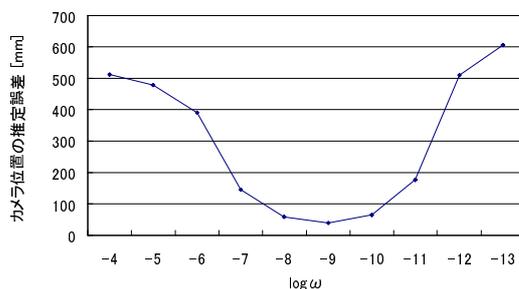


図 4 重み係数 ω の決定
Fig. 4 Decision of weighting factor ω .

この局所解回避策は経験的に得たものであり、効果については予めシミュレーションによる予備実験により確認している。図 3 は、真値に誤差を加えることで人為的に作成した初期値を提案手法により最適化した際のカメラ位置の推定誤差の変化を表す。このシミュレーションは、後述する 3.1 節で示すシミュレーションと同様の条件で行い、重み μ_i は、500 ステップ毎に変化させている。同図から、重み変動を行わない場合は、カメラ位置の平均推定誤差が 200mm 程度のときに局所解に陥っているが、重み変動を行うことにより、局所解が回避されていることが分かる。

また、重み係数 ω を決定するために、予め重み係数 ω が変化したときのカメラ位置の推定誤差の変化を予備実験により調べておく。予備実験では、3.1 節で示すシミュレーションと同程度のスケールの環境を使用し、カメラの移動経路については、異なるものを設定した。また、特徴点の検出精度や GPS の測位精度は、環境に大きく依存しないと考え、後述する 3.1 節のシミュレーションと同じとする。我々が行った予備実験では、図 4 に示すように、 $\omega = 10^{-9}$ のときに最も真値に近いカメラ位置が得られた。この予備実験で使用した仮想環境と類似する環境で提案手法を用いる場合は、重み係数を $\omega = 10^{-9}$ に設定する。

2.2 各処理の詳細

本節では、先に定義した特徴点と GPS に関する誤差関数 E を用いたカメラパラメータの推定手順について、図 1 に示したフローチャートの各ステップ (A) ~ (D) を順に説明する。本手法において、ステップ (A)、(B) は毎フレーム、(C) は一定フレーム毎に実行され、(D) は最終的な最適化処理として 1 回のみ実行される。

(A) 特徴点の追跡

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の対応関係を求める。特徴点には、回転・拡大縮小等の画像の変形に対して頑健に同一の位置が検出されるエッジの角や交点を用いる。さらに、特徴点の仮対応づけの結果から統計的手法によって誤対応結果を排除し、それにより求まる暫定的な外部パラメータを用いて、特徴点の対応関係を修正する。これにより、特徴点の周辺類似パターンとの誤対応を防止する。

具体的な処理手順としては、まず、(1) Harris オペレータ¹²⁾ で求められる特徴量が極大値となる位置を特徴点の候補位置とする。次に、(2) 前フレーム上の各特徴点の近傍領域を現フレーム上でパターン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。(3) 特徴点の仮の追跡位置から、LMeds 基準を用いたロバスト推定¹³⁾ によって、暫定外部パラメータを推定する。最後に、(4) 暫定外部パラメータを用いて、前フレームにおけるすべての特徴点の三次元位置から、現フレーム上での出現位置を予測し、ステップ (2) よりも小さな探索窓を用いて対応づけを行い、追跡位置を修正する。

(B) 外部パラメータの初期値推定

処理 (A) で求めた現フレーム (第 i フレーム) 上の特徴点に関する再投影誤差の二乗和を最小化することで特徴点の三次元位置、及び現フレームの外部パラメータである回転 R_i および並進 t_i を推定する。再投影誤差の二乗和は、以下のように表される。

$$\sum_{h=1}^i \mu_h \sum_j w_j \Phi_{hj}^2 \quad (5)$$

(C) GPS 測位値を用いた狭区間最適化

処理 (C) では、処理 (A)、処理 (B) で得られるカメラ外部パラメータを初期値として、前節で定義した E を最小化することで、GPS 測位値を反映させた外部パラメータを算出する。さらに、この処理で算出されたパラメータを処理 (A) にフィードバックすることで、特徴点対応付けの精度向上を図る。

具体的には、図 5 に示すように、処理 (A) 及び (B) で処理の対象となる第 i フレームに対して、第

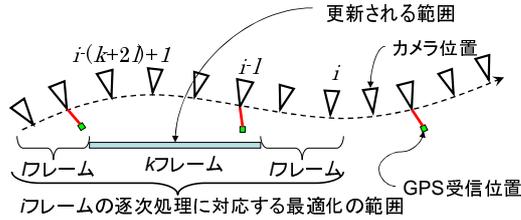


図5 最適化の範囲
Fig. 5 Range of optimization.

$i - (k + 2l) + 1$ フレームから第 i フレームの特徴点追跡結果および GPS 測位値を用いて狭区間での誤差関数 E の最小化を行い、第 $i - (k + l) + 1$ フレームから第 $i - l$ フレームの k フレーム分の外部パラメータのみを更新する。これにより、更新する区間に対して、前後のフレームにおける特徴点の追跡結果及び、GPS 測位値を最適化に反映させる。さらに、一度に複数のフレーム (k フレーム) を更新し、 k フレーム間隔で本処理 (C) を行うことで計算量の増大を軽減する。ただし、多数の追跡済みの特徴点が、推定結果を更新する区間 k フレーム内に写るように、 l は数十フレームに設定する。また、特徴点追跡が外部パラメータの蓄積誤差により破綻する前に GPS 測位値を反映させる必要があるため、 k は数フレームに設定する。

(D) GPS 測位値を用いた広範囲最適化

処理 (D) では、処理 (A) ~ (C) を繰り返すことで得られるカメラ外部パラメータを特徴点の三次元位置を初期値として、処理 (C) よりも広い区間で、外部パラメータの最適化を行う。これにより、特徴点追跡と GPS 測位の両情報に含まれる局所的な誤差の平均化を図る。ただし、本研究では、処理 (C) において間欠的に絶対指標を用いた最適化処理を行うため、時間的に十分離れた 2 フレームに関する外部パラメータや特徴点の三次元位置などの情報は、互いに影響しないと考えられる。そこで処理 (C) での最適化範囲 $k + 2l$ よりも十分に長い複数の $k' + 2l'$ フレームのシーケンスに動画を分割し、各シーケンスを独立に最適化する。ただし、処理 (C) と同様に、最適化区間内に外部パラメータを更新する区間 k' とその前後に更新しない区間 l' を設け、区間 l' を十分大きくとることで、各シーケンスの最適化における独立性を確保する。ただし、十分な数の GPS 測位値を用いて測位誤差を平均化するために、 l' は数百フレームに設定する。また、分割されたシーケンス間の独立性を確保するために、 k' についても数百フレームに設定する。

3. 実験

本章では、まず、式 (4) で定義した誤差関数 E の最小化によって最適化処理 (D) で推定される外部パラメータの精度を定量的に評価するために、実環境から作成したデータを用いたシミュレーション実験の結果を示す。次に、提案手法における特徴点検出位置及び GPS 測位値の誤差モデルや特徴点对応付けの精度を含め、提案手法の有効性を確認するために、実環境で取得した動画と GPS 測位値から外部パラメータを推定する。また、推定された外部パラメータを用いて動画上に仮想物体を重畳表示するマッチムーブを行うことで、外部パラメータが正しく推定されていることを確認する。最後に、動画全体の一括最適化によって推定される外部パラメータと、動画を複数のシーケンスへ分割して最適化し、最後に統合して得られる外部パラメータの比較を行うことで、分割による最適化の有効性を確認する。

3.1 シミュレーションによる精度評価

シミュレーションによる評価では、提案手法の処理 (D) によって推定される最終的な外部パラメータの推定精度を定量的に評価する。また、従来手法であるランドマーク・特徴点追跡による手法⁶⁾と精度を比較することにより、提案手法の有効性を確認する。

本シミュレーションで用いる仮想環境を図 6 に示す。ここで、図中の曲線はカメラの移動経路を、錘台は 5 フレーム毎のカメラの姿勢を表し、点群は仮想環境中に配置した特徴点の三次元座標を表す。また、グリッドの一边は実環境での 20m に相当する。仮想環境中の特徴点には、事前に屋外環境を撮影した全方位動画を入力として、佐藤らの手法¹⁴⁾を用いて三次元復元を行い、これによって推定された特徴点の三次元位置を本シミュレーションの真値として用いた。本実験では、このカメラの移動経路とは異なる経路で仮想カメラを同図に示すように a 地点から b 地点へ移動させ、仮想的に 990 フレームの特徴点検出画像を取得した。ただし、仮想カメラの内部パラメータには、次節で述べる実環境で用いたカメラと同一のものを用いた。本シミュレーションにおける入力では、1 フレームあたり平均 200 個の特徴点が画像上で観測された。また、カメラ座標系における GPS 受信機の位置を (600, 600, 600) (単位 mm) と設定し、GPS 測位値には、上記 GPS 受信機の位置に誤差を付加し、これを 1Hz に相当するフレーム間隔として 15 フレーム毎に与えた。特徴点の画像上での検出座標 q_{ij} および、GPS 測位値 g_s に付加した誤差を表 1 に示す。ま

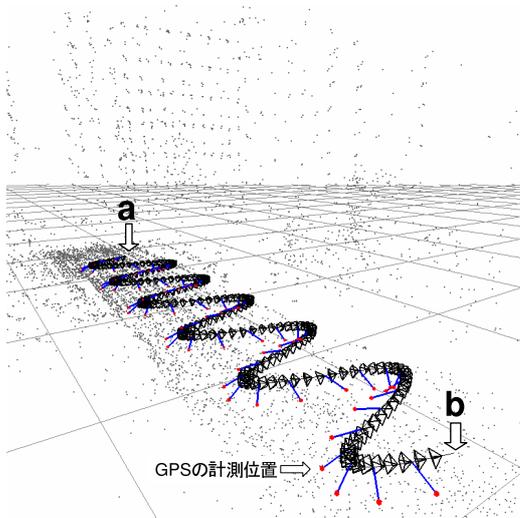


図 6 シミュレーション環境
Fig. 6 Environment for simulation.

た, 外部パラメータ R_i , t_i の初期値は, 表 2 に示す誤差を真値に付加して作成した. 2.1 節で定義した式 (4) 誤差関数 E 中の重み係数 ω は, 予めシミュレーションにより得た値 10^{-9} とした. また, 式 (4) 中の各フレームの重み係数 μ_i については, GPS 測位値が得られていないフレームで 1.0 とし, GPS 測位値が得られているフレームでは, 2.0 と 1.0 を交互に用いた. 逐次処理中に行う最適化処理 (C) においては, 更新対象フレーム $k = 5$, 重複フレーム $l = 22$ フレームと設定した. 処理 (D) では, 動画全体を一括で最適化した. 初期フレームにおける外部パラメータの初期値は, 手で大まかな姿勢を与え, 位置に関しては GPS 測位値を用いた. また, 提案手法の精度評価のために, ランドマーク・特徴点追跡による手法⁶⁾ によって上述と同様の条件で外部パラメータを推定した. ただし, ランドマークは特徴点の信頼度 w_{ij} を十分大きくしたものと与え, 全ての画像上での位置を与えた. また, ランドマークを与えるフレームの数や設定方法により外部パラメータの推定精度が変化する可能

表 1 観測データに付加する誤差
Table 1 Errors in observational data.

特徴点の検出座標	標準偏差 0.6 画素のガウスノイズ +量子化誤差
GPS の計測誤差	標準偏差 30mm のガウスノイズ

表 2 外部パラメータの初期値に付加する誤差
Table 2 Errors in initial extrinsic parameters.

カメラ位置	標準偏差 500mm のガウスノイズ
カメラ光軸の向き	標準偏差 0.020rad のガウスノイズ

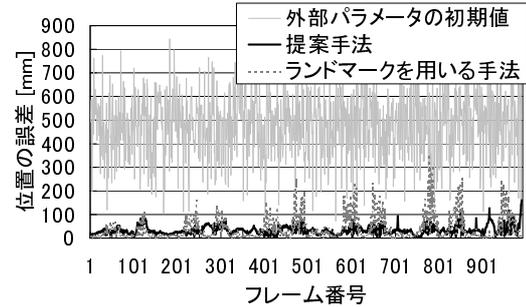


図 7 カメラ位置の推定誤差
Fig. 7 Errors of estimated camera position.

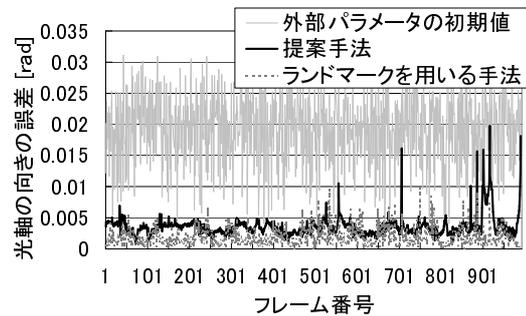


図 8 カメラ光軸方向の推定誤差
Fig. 8 Errors of estimated camera direction.

性があるため, 全てのフレームで 15 個以上のランドマークが観測されるように設定した.

以上の条件で, 特徴点の三次元位置, 外部パラメータ, カメラ座標系における GPS 受信機の方角を最適化した. 実験の結果, 外部パラメータは, カメラ位置に関して平均誤差 32.9mm, 標準偏差 18.0mm, カメラ光軸の向きに関して平均誤差 0.0036rad, 標準偏差 0.0016rad で推定された.

図 7 および図 8 に初期値として与えた真値からの誤差, 提案手法における誤差, ランドマーク・特徴点追跡による手法における誤差を示す. ただし, 図 7 は, 外部パラメータの位置に関する誤差, 図 8 については, 光軸の向きに関する誤差を示す. 実験の結果, ランドマークを用いる手法では, カメラ位置に関しては平均誤差 39.8mm, カメラ光軸の向きに関しては平均誤差 0.0019rad の精度で推定された. これに対して提案手法では, カメラ位置に関しては平均誤差 32.9mm, カメラ光軸の向きに関しては平均誤差 0.0036rad の精度で推定された.

提案手法では, 人手による計測が必要な事前知識を用いることなく, カメラの位置に関してランドマーク・特徴点追跡による手法と同程度の精度が得られているが, カメラの光軸に関しては提案手法で誤差が 2 倍程

度に大きくなった。これは、ランドマークを絶対指標とする手法と GPS を指標とする手法で、カメラの位置・姿勢の拘束のされ方が異なるためであると考えられる。本実験環境のようにランドマークが主に遠方に分布する場合、画像上でのランドマークの位置がカメラの並進の変化よりも回転の変化に敏感に反応することから、ランドマークを用いる場合は、姿勢に関する拘束が強く、結果として推定精度が高くなると考えられる。これに対して、GPS を用いる場合には、GPS 受信機の観測位置とカメラの投影中心間の距離が強く拘束されるため、カメラ位置に関する推定精度が従来手法よりも若干高まった反面、カメラ姿勢に関する精度は劣っていると考えられる。

3.2 実環境データを用いた外部パラメータの推定

本節では、提案手法における特徴点検出位置及び GPS 測位値の誤差モデルや特徴点对応付けの精度を含めて、提案手法の有効性を確認するために、実環境で取得した動画と GPS 測位値から外部パラメータを推定する。また、推定された外部パラメータを用いて動画上に仮想物体を重畳表示するマッチムーブを行うことで、外部パラメータが正しく推定されていることを確認する。

実験では、ワイドレンズ (Sony VCL-HG0758) を取り付けられたビデオカメラ (Sony DSR-PD-150) と GPS 受信機 (Nikon LogPakII, 測位精度 $\pm 3.0\text{cm}$) を図 9 に示すように車に固定して取り付けた。このような車両を用いて、約 1.1km の距離を平均時速 16.5km の速度で走行し、得られた 3600 枚の画像 (解像度 720×480 画素, 14.985fps, プロGRESSIVE撮影) 及び、RTK 法により 1 秒間隔で計測された GPS 測位値を入力として用いた。ただし、動画と GPS 測位値は手動で開始時刻を同期させ、カメラの内部パラメータは予め Tsai の手法¹⁰⁾ によりキャリブレーションし、その値を用いた。また、カメラと GPS 受信機間の距離は、手動で計測した 1020mm を既知な情報として設定した。初期フレームにおける外部パラメータは、前節と同様に、姿勢に関しては大まかな初期値を与え、位置に関しては GPS 測位値を用いた。更新対象フレーム k, k' および重複フレーム l, l' は、前節と同様に設定した。

まず、逐次処理において GPS 測位値を用いた最適化を行うことにより、安定した特徴点追跡と、誤差の蓄積しない外部パラメータ推定が実現されることを確認するため、逐次処理において GPS 測位値を用いて最適化する場合と、動画のみから推定する場合を比較した。ただし、蓄積誤差の影響を確認するために、

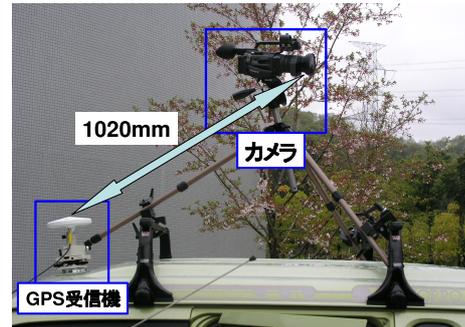


図 9 ビデオカメラと GPS 受信機の位置関係
Fig. 9 Positions of camera and GPS receiver.

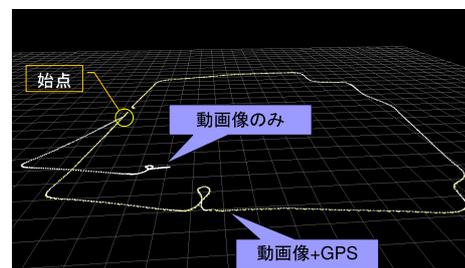
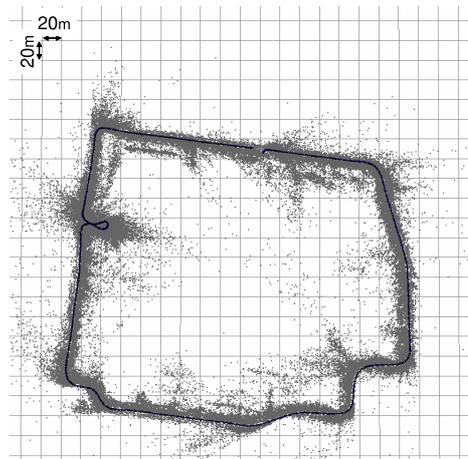


図 10 逐次処理で得られた外部パラメータ
Fig. 10 Extrinsic parameters estimated in initial estimation.

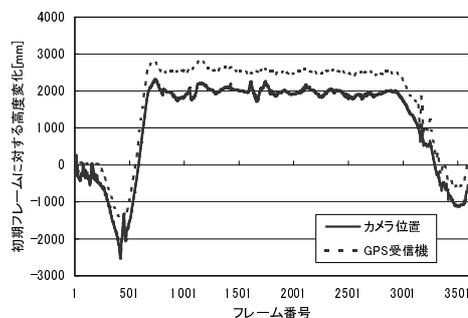
双方共に第 1 フレームと第 15 フレームの位置と姿勢を与え、ランドマークは用いなかった。動画のみから外部パラメータを推定した結果と、GPS 測位値を用いた最適化を行う逐次処理で推定された外部パラメータを図 10 に示す。

実験の結果、動画のみから外部パラメータを推定した場合は、図 10 に示すように第 1409 フレームで追跡される特徴点が減少し、以降の外部パラメータは推定されなかった。これに対し、GPS 測位値を用いた提案手法の逐次処理では、初期フレームから最終フレームまで途切れることなく、1 フレームあたり平均約 300 個の特徴点が追跡された。このことから、GPS 測位値を利用することで、利用しない場合よりも安定に一定数以上の特徴点が対応付けられていることが分かる。

次に、逐次処理で得られた外部パラメータを初期値として動画全体で最適化した。全体最適化によって推定された特徴点の三次元位置と外部パラメータを図 11 に示す。図中 (a) の曲線は推定されたカメラの移動経路を表し、点群は外部パラメータ推定に用いられた特徴点を表す。同図から分かるように、カメラパスは滑らかに推定されている。また、同図 (b) は、推定されたカメラ位置及び GPS 測位値の高度変化を表



(a) 水平移動軌跡



(b) 高度変化

図 11 カメラ外部パラメータの推定結果

Fig. 11 Result of estimated extrinsic parameters.

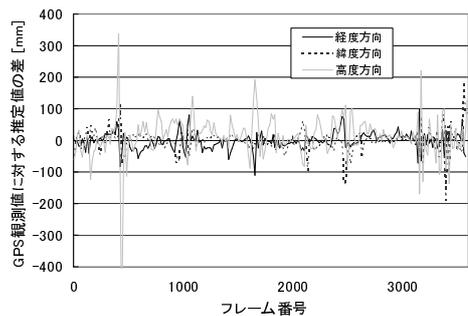


図 12 GPS 観測位置と推定位置の関係

Fig. 12 Errors of of estimated GPS positions.

す。同図から、カメラ位置が GPS 受信機の高度変化に沿うように推定されていることが分かる。図 12 に GPS 観測位置と推定位置の関係を示す。各成分の平均絶対値は、緯度方向が 18.5mm、経度方向が 21.2mm、高度方向が 43.1mm であった。高度方向の誤差が水平方向の 2 倍程度であることは、一般的な GPS 受信機

の高度方向の測位精度が水平方向よりも低いことと矛盾がない。

推定された外部パラメータを用いて仮想物体を動画画像に合成したマッチムーブの結果を図 13 に示す。仮想物体は、GPS による計測位置および GPS 受信機設置時の道路面からの高さから計算される道路面上に配置した。図 13 に示すシーンでは、仮想物体が現実環境中でほぼ一定の位置に合成されていることから、外部パラメータに大きな推定誤差が含まれないことを確認した (http://yokoya.naist.jp/pub/movie/yokochi/match_move.mpg)。しかし、図 14 に示すように、第 995 フレームから第 1030 フレームでは、仮想物体が一定の位置には静止せず、フレームが進むにつれ場所が移動した。これは、RTK-GPS の測位精度が環境中の遮蔽物により低下したことが考えられる。実際に、図 15 に示すように、RTK-GPS から出力される信頼度を表す標準偏差が第 995 フレームから第 1030 フレームにおいて増加している。本研究では、遮蔽物などの影響を受ける場合については考慮していないが、このような大域的な誤差に対応するためには、GPS 測位値の信頼度等のセンサ情報を考慮したに誤差関数を導入する必要がある。

最後に、処理 (D) における分割最適化の効果を検証するために、図 16 に示すように、重複フレーム l' を 200 フレーム、更新フレーム k' を 900 フレームとし、動画を 4 つのシーケンスに分割し、各シーケンスのみに最適化後に統合した。次に、動画を一括で最適化して得られる外部パラメータを真値とし、分割最適化によって得られた外部パラメータの誤差を検証した。図 17 に、分割最適化によって得られた外部パラメータの誤差を示す。本実験における平均誤差は、6mm であった。この値は、3.1 節で示した実験において示した、カメラ位置の推定誤差 32.9mm に対して比較的小さく、分割最適化により一括での最適化に近い推定結果が得られていることが分かる。計算時間は、Pentium4 3GHz、メモリ 1GB の PC を用いた場合、一括最適化が約 7 時間であったのに対して、分割最適化では各シーケンスがそれぞれ約 1.5 時間であった。誤差が大きなフレームに関しては、3 番目のシーケンスにおいて推定をした外部パラメータを更新しない区間 l' の大きさが不十分であった可能性がある。実際、図 17 に示すように、区間 l' を 2 倍にすることにより、上述した大きな誤差が減少し、平均 0.6mm となった。以上により、計算機に搭載されるメモリの量の制限など、一括で動画全体を最適化できないような場合においても、最適化の際に外部パラメータを更

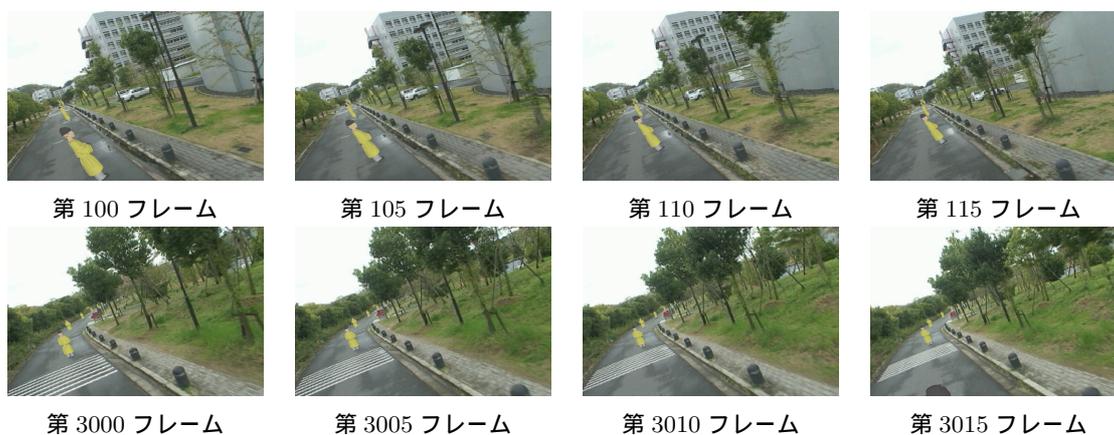


図 13 推定された外部パラメータを用いたマッチムーブの結果

Fig. 13 Result of match move.

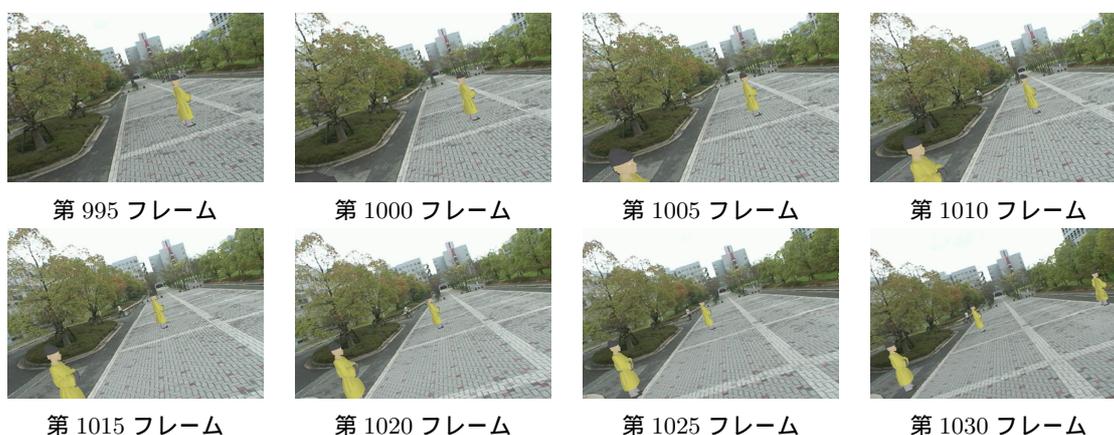


図 14 パラメータ推定の誤差によって仮想物体が静止しないシーン

Fig. 14 Examples of incorrect match move.

新しい区間 l' のフレーム数を十分大きく設定することで、一括で最適化する場合に近い結果が得られることが分かった。

4. ま と め

動画像と GPS 測位値からカメラの外部パラメータを推定する手法を提案した。提案手法では、画像上で定義される特徴点の再投影誤差と GPS 測位値によって定義される位置誤差を同時に最小化することで、外部パラメータを各フレームに対して推定することができる。

シミュレーション実験及び実環境データを用いた実験により、提案手法は、ランドマーク・特徴点追跡による手法に近い精度でカメラパラメータを推定可能であることを確認した。提案手法により、現実環境に関する事前知識の取得に要する人的コストの大幅な削減

が可能である。しかし、GPS の電波をさえぎる遮蔽物に囲まれるような状況下では、外部パラメータの推定精度の低下が見られた。

今後は、上記のような環境中に遮蔽物が多数存在する場合に対応するために、GPS 測位値の信頼度等のセンサ情報を考慮した誤差関数の導入を検討する。また、異なるシーケンスにおいて同一シーンが撮影される場合に再出現する特徴点を対応づけすることで、さらなる精度向上を図る。

参 考 文 献

- 1) Feiner, S., MacIntyre, B., Höllerer, T. and Webster, A.: A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment, *Proc. 1st IEEE Int. Symp. on Wearable Computers*

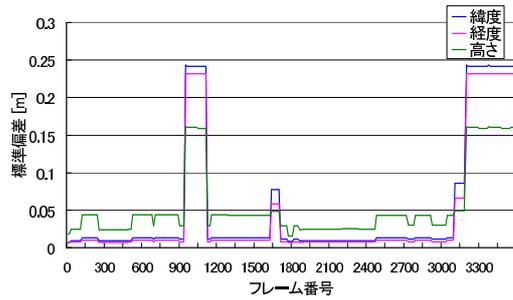


図 15 GPS 測位値に含まれる標準偏差

Fig. 15 Standard deviations in GPS data.

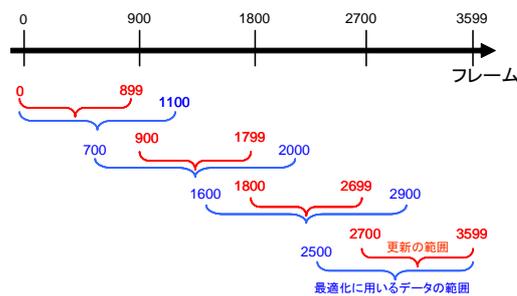


図 16 分割最適化における各シーケンスの分割範囲

Fig. 16 Range of optimization of divided sequence.

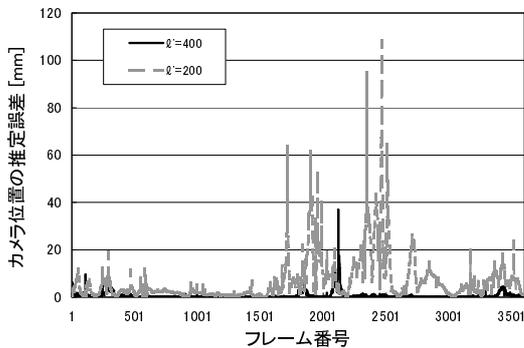


図 17 分割最適化の精度

Fig. 17 Result of optimization of divided sequence.

(ISWC1997), pp. 208–217 (1997).

- 2) 遠藤隆明, 片山昭宏, 田村秀行, 廣瀬通孝: 写実的な広域仮想空間構築のための画像補間手法, 日本バーチャルリアリティ学会論文誌, Vol. 7, No. 2, pp. 185–192 (2002).
- 3) Fitzgibbon, A. W. and Zisserman, A.: Automatic Camera Recovery for Closed or Open Image Sequences, *Proc. 5th European Conference on Computer Vision (ECCV)*, pp. 311 – 326 (1998).
- 4) Pollefeys, M., Koch, R., Vergauwen, M., Deknuydt, B. and Gool, L. V.: Three-

dimensional Scene Reconstruction from Images, *Proc. SPIE*, Vol. 3958, pp. 215–226 (2000).

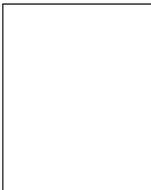
- 5) Davison, A. J.: Real-Time Simultaneous Localisation and Mapping with a Single Camera, *Proc. 9th IEEE Int. Conf. on Computer Vision*, Vol. 2, pp. 1403–1410 (2003).
- 6) 佐藤智和, 神原誠之, 横矢直和, 竹村治雄: マーカと自然特徴点の追跡による動画像からのカメラ移動パラメータの復元, 電子情報通信学会論文誌 (D-II), Vol. J86-D-II, No. 10, pp. 1431–1440 (2003).
- 7) Comport, A.I., Marchand, É. and Chaumette, F.: A Real-time Tracker for Markerless Augmented Reality, *Proc. 2nd ACM/IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR2003)*, pp. 36–45 (2003).
- 8) Vacchetti, L., Lepetit, V. and Fua, P.: Combining edge and texture information for real-time accurate 3D camera tracking, *Proc. 3rd IEEE and ACM Int. Symp. on Mixed and Augmented Reality (ISMAR2004)*, pp. 48–57 (2004).
- 9) Güven, S. and Feiner, S.: Authoring 3D Hypermedia for Wearable Augmented and Virtual Reality, *Proc. 7th IEEE Int. Symp. on Wearable Computers (ISWC2003)*, pp. 118–126 (2003).
- 10) Tsai, R. Y.: An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 364–374 (1986).
- 11) Triggs, B., P. McLauchlan, R. H. and Fitzgibbon, A.: Bundle Adjustment A Modern Synthesis, *Vision Algorithms: Theory and Practice*, pp. 298 – 375 (2000).
- 12) Harris, C. and Stephens, M.: A Combined Corner and Edge Detector, *Proc. Alvey Vision Conf.*, pp. 147–151 (1988).
- 13) 栗田多喜夫, 坂上勝彦: 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, Vol. Vol. I, pp. 65–70 (2000).
- 14) 佐藤智和, 池田聖, 横矢直和: 複数動画像からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定, 電子情報通信学会論文誌 (D-II), Vol. J88-D-II, No. 2, pp. 347–357 (2005).

(平成 17 年 5 月 25 日受付)

(平成 17 年 12 月 21 日採録)

**横地 裕次**

2003 年奈良工業高等専門学校専攻科電子情報工学専攻修了。2005 年奈良先端科学技術大学院大学情報科学研究科修了。現在、株式会社 本田技術研究所 栃木研究所 勤務。

**池田 聖 (学生会員)**

2001 年広島大学理学部物理学科卒。2003 年奈良先端科学技術大学院大学情報科学研究科博士前期課程修了。現在、同大博士後期課程に在学中。電子情報通信学会、情報処理学会、IEEE 各会員。

**佐藤 智和 (正会員)**

1999 年大阪府立大学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。現在、同大情報科学研究科助手。コンピュータビジョンの研究に従事。2001 年電子情報通信学会学術奨励賞受賞。電子情報通信学会、情報処理学会、IEEE 各会員。

**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了。同年電子技術総合研究所入所。以来、画像処理ソフトウェア、画像データベース、コンピュータビジョンの研究に従事。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、同大情報科学研究科教授。1990 年情報処理学会論文賞受賞。2005 年情報処理学会フェロー、電子情報通信学会フェロー。工博。電子情報通信学会、情報処理学会、人工知能学会、日本認知科学会、映像情報メディア学会、IEEE 各会員。