

## 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

佐藤 智和<sup>†a)</sup> 池田 聖<sup>†</sup> 横矢 直和<sup>†</sup>

Determining Extrinsic Camera Parameters of an Omni-Directional Multi-Camera System from Multiple Video Streams

Tomokazu SATO<sup>†a)</sup>, Sei IKEDA<sup>†</sup>, and Naokazu YOKOYA<sup>†</sup>

あらまし 複数のカメラユニットから構成される全方位型マルチカメラシステムを用い、移動を伴って取得した全方位動画像をもとに、カメラブロックの位置・姿勢パラメータであるカメラ外部パラメータを推定する手法を提案する。本論文ではまず、三次元位置が既知の基準マーカと全方位画像から、PnP (Perspective n-Point) 問題を解くことにより、カメラブロックの外部パラメータを統合的に推定する手法について述べる。次に、動画像中のキーフレームにおいて基準マーカを画像上で指定し、三次元位置が未知の特徴点 (自然特徴点) と同時に自動追跡することで、基準マーカが観測できないカメラ位置においてもカメラブロックの外部パラメータを推定する手法について述べる。実験では、屋内・屋外の実環境を撮影した全方位動画像を用いてカメラ外部パラメータを推定し、定量的評価を行う。

キーワード 全方位型マルチカメラシステム, PnP 問題, Structure from motion, 特徴点追跡

### 1. ま え が き

近年、全方位を一度に撮影可能な全方位型カメラシステムが多数開発され、テレプレゼンスや遠隔ロボット操縦、ビデオサーベイランスなどの幅広い分野に应用されはじめている [1] ~ [6]。これら全方位カメラの中でも、複数のカメラユニットを放射状に配置することで一つのカメラブロックを構成し、一度に全方位を撮影する全方位型マルチカメラシステム (Omni-directional Multi-camera System, 以降 OMS) [4] ~ [6] は、カメラシステムの校正が難しい反面、高解像度であることや、視線方向によって解像度があまり変化しないという特長から、注目され始めている。しかし、各カメラユニットの投影中心の不一致を考慮した統合的なカメラ外部パラメータの推定手法が従来提案されていないために、OMS の利用はテレプレゼンスなどのカメラの位置・姿勢パラメータを必要としない分野に限定

されてきた。そこで本論文では、移動を伴って撮影された全方位動画像中のキーフレームで絶対指標 (基準マーカ) を与えることで、高精度に絶対位置・姿勢を推定する手法について述べる。本手法では、推定されるカメラパラメータと画像情報を屋外三次元モデリングに応用することを想定しており、町並みにおける交差点など、キーとなる地点 (キーフレーム) において絶対指標が設定・計測できることを前提としている。以下では、従来手法の概要と問題点について述べる。

まず、単眼カメラにおけるカメラ外部パラメータの推定手法について述べる。単眼カメラで撮影される 1 枚の画像において、複数の基準マーカの三次元座標とその画像上の二次元座標の組から、カメラ外部パラメータを推定する問題は、PnP 問題 (Perspective n-Point problem) と呼ばれ、現在までに多くの研究がなされている [7] ~ [9]。一般に、PnP 問題は 6 点以上の基準マーカを画像上で観測できれば、最小二乗法によって解くことができる [10]。しかし、三次元位置が既知の基準マーカが常に画像フレーム内に撮影されている必要があるため、広い範囲の移動を伴って撮影された動画像の各フレームに対してこれらの手法を適

<sup>†</sup> 奈良先端科学技術大学院大学, 生駒市

Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma-shi, 630-0192 Japan

a) E-mail: tomoka-s@is.naist.jp

用し、カメラの外部パラメータを推定するためには、多数の基準マーカを広く配置することが必要となる。また、OMS への応用を考えた場合、複数カメラの投影中心の不一致を考慮した PnP 問題の解法が従来提案されていないために、そのままでは手法を適用できないという問題がある。

一方で、単眼カメラで撮影される動画像において、画像上の三次元位置が未知の特徴点（以降、自然特徴点）を追跡することで、カメラの外部パラメータを推定する Structure from motion と呼ばれる手法が研究されている [11] ~ [13]。また、これらを全方位カメラに適用する手法 [14] ~ [16] が複数提案されている。これらの手法では、三次元位置が既知の基準マーカを必要としないが、自然特徴点の誤追跡や量子化誤差の影響によるカメラパラメータの推定誤差が累積するために、多くの画像を扱うことが困難であるという問題がある。また、これらの手法はカメラが一点透視投影であることを仮定しているため、複数の投影中心をもつ OMS への適用は困難である。

そこで本研究では、OMS を構成する複数のカメラユニットで撮影される動画像間にわたって基準マーカと自然特徴点を同時に追跡することで、広範囲の移動を伴う長時間の全方位動画像においてもカメラ外部パラメータを安定に推定する手法を提案する。本手法では、図 1 に示すように、まず移動を伴って撮影した全方位動画像中のキーフレームにおいて、基準マーカを画像上で指定する (A)。次に、基準マーカと自然特徴点を複数の動画像に渡って自動追跡することで、カメラ外部パラメータと自然特徴点の三次元位置をフレームごとに逐次推定する (B)。このとき、OMS のカメラブロックの外部パラメータは、基準マーカと自然特徴点の三次元位置及び画像上の追跡座標を用いて、PnP 問題をフレームごとに解くことで得る。最後に、キーフレームにおいて指定した基準マーカを用いて動画像全体で投影誤差の和を最小化することで、カメラ外部パラメータの累積的な推定誤差を最小化し、精度の高いカメラ外部パラメータを得る (C)。本手法では、ステージ (B) において、カメラユニット間の投影中心の不一致を考慮した PnP 問題の解法が必要となるため、本論文では OMS における PnP 問題についても新たに統合的な解法を提案する。ただし、本研究で扱う OMS は、位置関係が固定された複数のカメラユニットで構成されており、カメラブロックの内部パラメータ（各カメラユニット間の位置・姿勢関係及び各カメ

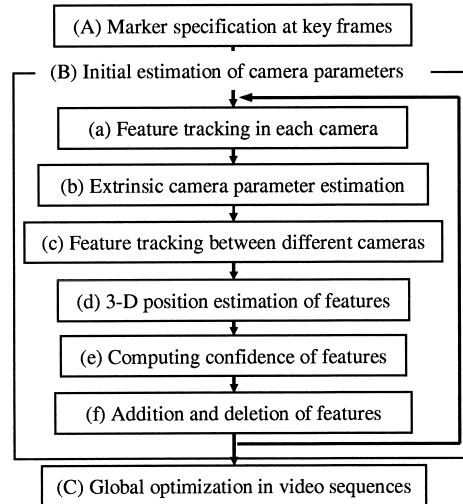


図 1 特徴点追跡によるカメラ外部パラメータ推定の処理手順

Fig. 1 Flow diagram of extrinsic camera parameter estimation by tracking features.

ラユニットの内部パラメータ) は校正済みで既知とする。また、基準マーカは、全方位動画像中の各キーフレームにおいてそれぞれ 6 点以上が観測可能であるものとする。

以下、2. ではまず、ステージ (B) で利用する、OMS を構成するカメラユニット間の投影中心の不一致を考慮した PnP 問題の統合的な解法について述べる。3. では、基準マーカと自然特徴点を複数の動画像上で同時に追跡し逐次 PnP 問題を解くことで、カメラ外部パラメータを推定する手法について述べる。4. では、市販の OMS を用いた実験と定量的評価実験について報告する。最後に、5. でまとめと今後の課題について述べる。

## 2. OMS における PnP 問題の解決

本章では、複数の基準マーカの三次元位置とその画像上の座標から、PnP 問題を統合的に解くことによって、OMS のカメラブロックのカメラ外部パラメータを推定する手法を提案する。本手法は、複数のカメラユニットに撮像される基準マーカを統合的に扱うため、カメラユニット全体で合計 6 点以上の基準マーカが観測できることを想定している。以下では、まず、OMS のカメラ外部パラメータと基準マーカの投影誤差について定義を行う。次に、定義した投影誤差を最小化することで、OMS のカメラ外部パラメータを推定する

手法について述べる．

### 2.1 カメラ外部パラメータと投影誤差の定義

まず、本論文で扱う座標系について定義する．図 2 に示すように、一般に OMS はカメラブロックの特定の位置に固定された複数のカメラユニットから構成されている．本論文では、OMS のカメラブロックの位置・姿勢を表すカメラブロック座標系を設定し、世界座標系からカメラブロック座標系への変換行列  $M$  を、位置  $(t_1, t_2, t_3)$ ・姿勢  $(r_1, r_2, r_3)$  の 6 パラメータを用いて以下のように定義する．

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

$$= \begin{pmatrix} \mathbf{R}(r_1, r_2, r_3) & (t_1, t_2, t_3)^T \\ \mathbf{0} & 1 \end{pmatrix} \quad (2)$$

ただし、 $\mathbf{R}$  は  $3 \times 3$  の回転行列である．また、カメラブロック座標系から  $k$  番目のカメラユニットのローカルカメラ座標系への変換を、同じく  $4 \times 4$  の行列  $T_k$  で表す．このとき、世界座標系から  $k$  番目のカメラユニットのローカルカメラ座標系への変換行列を  $L_k$  とすれば、 $L_k = T_k M$  である．

次に、基準マーカの投影誤差について定義する．以降、基準マーカの画像上での検出座標は画像ひずみ補正後のものとし、記述の簡単化のため、カメラ内部パラメータである各カメラユニットの焦点距離及びアスペクト比を 1、投影中心の座標を  $(0, 0)$  とする．このとき、基準マーカ  $p$  の 3 次元座標  $S_p = (x_p, y_p, z_p, 1)^T$  と、その  $k$  番目のカメラユニットの画像上への投影座標  $(u_p, v_p)$  には、以下の関係が成り立つ．

$$(au_p, av_p, a, 1)^T = L_k S_p = T_k M S_p \quad (3)$$

ただし、 $a$  は媒介変数である．一般に、式 (3) によって

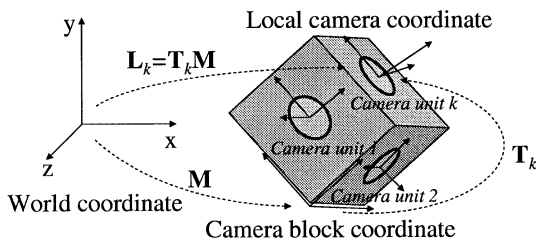


図 2 OMS における座標系の関係  
Fig. 2 Coordinate systems of an OMS.

求まる投影座標  $(u_p, v_p)$  は、基準マーカの検出誤差や量子化誤差の影響によって、画像上での基準マーカの検出座標  $(\hat{u}_p, \hat{v}_p)$  とは一致しない．一般的な単眼カメラにおいては、この  $(\hat{u}_p, \hat{v}_p)$  と  $(u_p, v_p)$  の距離の二乗和を評価尺度とし、最小化することでカメラパラメータを推定する手法が広く用いられている [13], [17] ~ [19]．本研究においても、これらの手法と同様に、各カメラユニットの画像上への基準マーカの投影誤差の二乗和を評価尺度とし、誤差関数  $E$  を以下のように定義する．

$$E = \sum_{k=1}^n \sum_{p \in F_k} \{ (u_p - \hat{u}_p)^2 + (v_p - \hat{v}_p)^2 \} \quad (4)$$

ただし  $F_k$  は、 $k$  番目のカメラユニットの画像上において観測可能な基準マーカの集合を、 $n$  は OMS を構成するカメラユニットの数を表す．

### 2.2 投影誤差の最小化によるカメラ外部パラメータの推定

前節において誤差関数  $E$  として定義した基準マーカの各画像面への投影誤差の和を最小化することで、自由度 6 のカメラ外部パラメータ  $M$  を推定する．この問題は、非線形最小化問題となるため、最適な解を得るためには良い初期値が必要となる．

本研究ではまず、線形手法によって 12 自由度の  $M$  を初期値として算出し、6 自由度への補正を行った後に、最急降下法によって  $E$  を最小化する  $M$  を算出する．以下では、それぞれの手順について詳述する．

#### 2.2.1 線形演算による初期値の算出

ここでは、 $n$  台のカメラに写る合計  $j$  個 ( $j \geq 6$ ) の基準マーカを統合的に扱いながら線形に PnP 問題を解くための手法について述べる．まず、世界座標系から  $k$  番目のカメラユニットのローカルカメラ座標系への変換行列  $L_k$  を三つの行ベクトル  $(l_{x_k}, l_{y_k}, l_{z_k})$  を用いて以下のように再定義する．

$$L_k = T_k M = \begin{pmatrix} l_{x_k} \\ l_{y_k} \\ l_{z_k} \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5)$$

上式により、式 (3) は以下のように変形できる．

$$\begin{aligned} l_{x_k} S_p - \hat{u}_p l_{z_k} S_p &= 0, \\ l_{y_k} S_p - \hat{v}_p l_{z_k} S_p &= 0 \end{aligned} \quad (6)$$

このとき、行列  $M$  の各成分をパラメータベクトル

$\mathbf{m} = (m_{11}, \dots, m_{14}, m_{21}, \dots, m_{24}, m_{31}, \dots, m_{34})^T$  で表せば、式 (6) は合計  $j$  点の基準マーカについて、以下の形にまとめることができる。

$$\mathbf{A}\mathbf{m} - \mathbf{s} = \mathbf{0} \quad (7)$$

ただし、 $k_i$  を  $i$  番目の基準マーカが撮影されたカメラユニットの番号と定義すれば、 $\mathbf{A}$ ,  $\mathbf{s}$  は以下のように表される。

$$\begin{pmatrix} s1(k_i) & s2(k_i) & s3(k_i) & s4(k_i) \\ s5(k_i) & s6(k_i) & s7(k_i) & s8(k_i) \end{pmatrix} = \begin{pmatrix} 1 & 0 & -\hat{u}_i & 0 \\ 0 & 1 & -\hat{v}_i & 0 \end{pmatrix} \mathbf{T}_{k_i} \quad (i = 0, \dots, j) \quad (8)$$

$$\mathbf{A} = \begin{pmatrix} s1(k_1)\mathbf{S}_1^T & s2(k_1)\mathbf{S}_1^T & s3(k_1)\mathbf{S}_1^T \\ \vdots & \vdots & \vdots \\ s1(k_j)\mathbf{S}_j^T & s2(k_j)\mathbf{S}_j^T & s3(k_j)\mathbf{S}_j^T \\ s5(k_1)\mathbf{S}_1^T & s6(k_1)\mathbf{S}_1^T & s7(k_1)\mathbf{S}_1^T \\ \vdots & \vdots & \vdots \\ s5(k_j)\mathbf{S}_j^T & s6(k_j)\mathbf{S}_j^T & s7(k_j)\mathbf{S}_j^T \end{pmatrix} \quad (9)$$

$$\mathbf{s} = (-s4(k_1), \dots, -s4(k_j), -s8(k_1), \dots, -s8(k_j))^T \quad (10)$$

式 (7) において、 $\mathbf{m}$  以外のパラメータはすべて既知である。

ここで、式 (6) より、 $|\mathbf{A}\mathbf{m} - \mathbf{s}|^2$  は、各特徴点の画像上の投影誤差をそれぞれ  $(\mathbf{z}_k \mathbf{S}_p)^2$  倍した和、すなわち、画像面上での投影誤差の和ではなく三次元空間中の誤差の和を意味している。本論文では、 $|\mathbf{A}\mathbf{m} - \mathbf{s}|^2$  を最小化する  $\mathbf{m}$  を線形に求め、カメラ外部パラメータの初期値とする。一般に、 $|\mathbf{A}\mathbf{m} - \mathbf{s}|^2$  を最小化するパラメータベクトル  $\mathbf{m}$  は、基準マーカの数  $j$  が 6 点以上あれば、一般化逆行列により以下のように求めることができる。

$$\mathbf{m} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{s} \quad (11)$$

ただし、基準マーカの三次元座標  $\mathbf{S}_p$  の分布範囲が、カメラ間の距離に対して非常に大きい場合、 $\mathbf{s}$  のスケールが  $\mathbf{A}$  に対して非常に小さくなるために、式 (11) は不安定となり、 $\mathbf{m}$  が誤差に敏感となってしまう。そこで、 $\mathbf{m}$  の各要素を用いて、新たなパラメータベクトル  $\mathbf{m}' : m'_{ij} = m_{ij}/m_{34}$  を定義し、 $\mathbf{s}$  を  $\mathbf{0}$  で近似する。これらを用い、 $\mathbf{A}\mathbf{m}' = \mathbf{0}$  を解くことで、スケールを除く 11 自由度をもつカメラパラメータの初期値  $\mathbf{M}$  を安定に求めることができる。次に、求めた  $\mathbf{M}$  の回転行列成分を正規化することでスケールを決定する。

## 2.2.2 カメラ外部パラメータの自由度の補正

前述の手法によって得られた 12 パラメータの  $\mathbf{M}$  は、正規直交基底を満たしておらずユークリッド復元ではないために、これを用いた三次元復元結果にはひずみが生じる。そこで、これを 3 自由度の位置パラメータ  $(t_1, t_2, t_3)$  と 3 自由度の姿勢パラメータ  $(r_1, r_2, r_3)$  からなる 6 自由度のカメラパラメータに補正する。 $\mathbf{M}$  の定義である式 (1), (2) から、OMS の位置成分  $(t_1, t_2, t_3)$  は容易に  $(m_{14}, m_{24}, m_{34})$  として決定可能である。また、姿勢パラメータ  $(r_1, r_2, r_3)$  は回転行列に対応する残りの 9 パラメータ  $\hat{\mathbf{R}}$  から、特異値分解を用いる手法 [20] によって以下のように最適な回転成分を計算する。

$$\mathbf{R}(r_1, r_2, r_3) = \mathbf{U} \text{diag}(1, 1, \det(\mathbf{UV}^T)) \mathbf{V}^T \quad (12)$$

ここで  $\mathbf{U}$  と  $\mathbf{V}$  はそれぞれ、 $\hat{\mathbf{R}}$  の左特異ベクトル行列、右特異ベクトル行列である。また、 $\text{diag}(\mathbf{x})$  は  $\mathbf{x}$  を対角成分にもつ対角行列を表す。

## 2.2.3 投影誤差の非線形最小化

6 自由度に補正済みのカメラ外部パラメータ  $(r_1, r_2, r_3, t_1, t_2, t_3)$  を初期値として、式 (4) で定義した誤差関数  $E$  を最急降下法によって最小化する。本手法では、各パラメータを以下のように更新し、収束するまで繰り返すことで、最適なパラメータの推定を行う。

$$r_i \leftarrow r_i - l_{r_i} \frac{\delta E}{\delta r_i}, t_i \leftarrow t_i - l_{t_i} \frac{\delta E}{\delta t_i} \quad (i = 1, 2, 3) \quad (13)$$

ただし  $(l_{r_1}, l_{r_2}, l_{r_3}, l_{t_1}, l_{t_2}, l_{t_3})$  は導関数のスケール項で、これらは各パラメータ更新の時点で  $E$  が最小となるように動的に決定する。既に述べたように、2.2.1 で得られる線形解は、各基準マーカの投影誤差の重み付き和を最小化したものであるため、この繰返しに用いる初期値は、真値に近い値であると考えられる。このため、最急降下法によって、誤差関数  $E$  の大域最小解を得る最適な OMS のカメラ外部パラメータ  $\mathbf{M}$  を算出することができる。

## 3. 特徴点追跡による OMS のカメラ外部パラメータの推定

本章では、OMS を構成する複数のカメラユニットで撮影された動画像系列にわたって特徴点を追跡し、前章で述べた PnP 問題の解法を利用することで、移動する OMS のカメラ外部パラメータを推定する手法について述べる。本研究では、Structure from motion の分野で一般的に用いられる自然特徴点に加えて、三

次元位置が既知の基準マーカをキーフレームで指定することで、世界座標系に対するカメラブロックの絶対位置・姿勢を推定する。また、複数動画像系列全体での最適化により、逐次処理で発生する累積的な推定誤差を最小化する。

本手法では、図 1 に示したように、まず初期フレーム、中間フレーム、最終フレームのようなキーフレームにおいて基準マーカを画像上で指定し (A)、指定された基準マーカと自然特徴点を動画像上で同時に自動追跡することで、カメラ外部パラメータと自然特徴点の三次元位置の初期値を算出する (B)。次に、キーフレームで指定した基準マーカを用い、特徴点の投影誤差を動画像全体で最小化することによって、最適なカメラ外部パラメータを推定する (C)。

以下では、まず、前章と同様に OMS の動画像における特徴点の投影誤差を定義する。次に、定義した誤差関数を最小化することで、動画像から OMS のカメラ外部パラメータと自然特徴点の三次元位置を推定する手法について述べる。

### 3.1 全方位動画像における投影誤差の定義

前章において式 (4) で定義した基準マーカの投影誤差  $E$  を動画像に拡張する。本手法では、動画像の第  $f$  フレーム ( $f = 1, \dots, q$ ) における誤差関数を以下のように定義する。

$$E_f = \sum_{k=1}^n \sum_{p \in \mathbf{F}_{kf}} W_p \{ (u_{fp} - \hat{u}_{fp})^2 + (v_{fp} - \hat{v}_{fp})^2 \} \quad (14)$$

ただし、 $W_p$  は特徴点  $p$  の信頼度による重みであり、特徴点  $p$  の投影誤差の分散の逆数として算出する [19]。また、 $\mathbf{F}_{kf}$  は第  $f$  フレームにおいて  $k$  番目のカメラユニットで観測できる特徴点の集合である。

本研究では、全方位動画像全体での誤差関数  $E_{total}$  を  $E_f$  を用いて以下のように表す。

$$E_{total} = \sum_{f=1}^q E_f \quad (15)$$

一方、第  $f_s$  フレームから第  $f_e$  フレームまで追跡される特徴点  $p$  の投影誤差の和は以下のように表される。

$$EF_p(f_s, f_e) = \sum_{f=f_s}^{f_e} \{ (u_{fp} - \hat{u}_{fp})^2 + (v_{fp} - \hat{v}_{fp})^2 \} \quad (16)$$

## 3.2 全方位動画像からのカメラ外部パラメータの推定

本節では、基準マーカと自然特徴点を自動追跡し、前章で述べた PnP 問題を各フレームで解くことでカメラブロックの外部パラメータを逐次的に推定する手法、誤差関数  $E_{total}$  を動画像全体で最小化することでカメラ外部パラメータを最適化する手法、について順に述べる。本節で述べる手法は、我々が従来提案した、単眼動画像からのカメラ外部パラメータ推定手法 [19] を全方位動画像に対して拡張したものである。

### 3.2.1 特徴点追跡によるカメラ外部パラメータの逐次推定

以下では、基準マーカと自然特徴点を自動追跡し、2. で述べた PnP 問題を各フレームで解くことで、OMS のカメラ外部パラメータの初期値を逐次推定する手法について述べる。本手法では、フレームごとに新たな自然特徴点を自動で検出し、三次元位置が既知の特徴点と未知の特徴点を同時に追跡することで、三次元位置が未知の特徴点の三次元位置を推定・更新する。三次元位置が推定された自然特徴点は、基準マーカと同様、PnP 問題を解くために利用できるため、基準マーカが不可視となるフレームにおいてもカメラ外部パラメータの推定を継続できる。

以下では、第  $f$  フレームにおける逐次処理の各ステップ (a) ~ (f) について処理の順に詳述する。

(a) 各カメラ画像内での特徴点追跡： 基準マーカは、一般的なテンプレートマッチングの手法を用いて、キーフレームから一定フレームの間、自動で追跡する。自然特徴点は、Harris オペレータ [21] により追跡の容易な画像特徴を検出して特徴点の候補位置とし、ロバスト推定 [22] によって誤追跡を検出しながら自動で追跡する。本ステップにおける特徴点追跡は、それぞれのカラユニットの画像内でのみ行う。

(b) カメラ外部パラメータの推定： 2. で述べた PnP 問題の解法を用いることで、ステップ (a) で追跡された特徴点の画像上の座標と三次元座標から、カメラブロックのカメラ外部パラメータ  $M_f$  を推定する。ここでは、式 (14) で定義した誤差関数  $E_f$  を用いることで、各特徴点の信頼度を考慮したカメラパラメータの算出を行う。

(c) カメラ画像間での特徴点追跡： ステップ (a) の追跡処理において、画像からのフレームアウトによって、追跡が失敗した自然特徴点について、他のカメラ画像への追跡処理を行う。ここでは、図 3 に示すよ

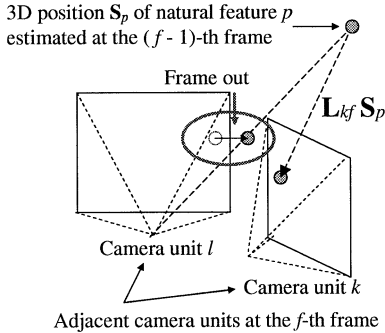


図 3 異なるカメラ画像間にわたる特徴点の追跡  
Fig. 3 Feature tracking between images of adjacent camera units.

うに、ステップ (b) で得られたカメラ外部パラメータ  $M_f$  を用い、世界座標系から各ローカルカメラ座標系への変換行列  $L_{kf}(=T_k M_f)$  によって、前フレームで推定済みの自然特徴点の三次元位置  $S_p$  をすべてのカメラユニットの画像上に投影する。ここで、各画像への投影位置の周辺に Harris オペレータにより算出される特徴点の候補が存在する場合には、その点を追跡位置とする。候補位置が存在しない場合には、その特徴点を削除する。

(d) 自然特徴点の三次元位置の推定：第  $f$  フレームまで追跡されているすべての自然特徴点について、式 (16) に示した  $EF_p(fs(p), f)$  を最小化することで自然特徴点の三次元位置  $S_p$  を毎フレームで推定・更新する。ただし、 $fs(p)$  は自然特徴点  $p$  が最初に検出されたフレーム番号である。

(e) 自然特徴点の信頼度の算出：自然特徴点の画像上の追跡誤差をガウス分布で近似することにより、特徴点  $p$  の信頼度  $W_p$  を投影誤差の分散の逆数として定義し算出する [19]。この信頼度は、毎フレーム更新する。

(f) 自然特徴点の追加と削除：特徴点の信頼度やテンプレート間の誤差などの複数の評価尺度を用いて、自然特徴点の追加・削除を自動的に行う [19]。上述のステップ (a) ~ (f) を、初期フレームから最終フレームまで繰り返すことで、全方位動画像から OMS のカメラ外部パラメータと自然特徴点の三次元位置の初期値を推定することができる。

### 3.2.2 複数動画像系列全体での最適化

式 (15) に示した誤差関数  $E_{total}$  を、最急降下法によって最小化することで、逐次推定で得られたカメラ外部パラメータ  $M_f(f=1, \dots, q)$  と自然特徴点の三

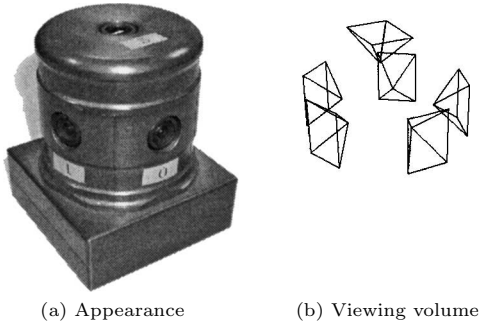
次元位置  $S_p$  を最適化する。ここでは、2.2.3 において述べた最適化の手法をすべての動画像のフレームに適用することで最適化を行う。ただし、ここでは、自然特徴点の三次元位置  $S_p = (x_p, y_p, z_p, 1)^T$  もパラメータであり、これらは導関数  $(\frac{\delta E_{total}}{\delta x_p}, \frac{\delta E_{total}}{\delta y_p}, \frac{\delta E_{total}}{\delta z_p})$  を用いることで、カメラの外部パラメータと同様の最適化処理を行う。自然特徴点の信頼度  $W_p$  としては、逐次処理で得られた値を用いる。また、キーフレームで指定した基準マーカ  $p$  の信頼度  $W_p$  としては、自然特徴点の信頼度の平均に対して十分に大きな値を設定する。これにより、キーフレームにおける基準マーカの三次元位置及び画像上の座標を真値とみなした動画像全体での最適化処理が可能となる。

このような誤差関数  $E_{total}$  が収束するまで最適化を行うことで、累積的な誤差を最小化するカメラ外部パラメータと自然特徴点の三次元位置が推定される。この最適化は非線形最小化問題となるが、初期値が真値に近ければ解は大域最小解に収束しやすくなり、また高速に求めることができる。

## 4. 実験

提案手法を用いて OMS のカメラ外部パラメータを推定し、定量的評価を行う。本章ではまず、2. で述べた PnP 問題の解法を用い、計算機シミュレーションによって定量的にカメラ外部パラメータの推定精度を評価する。次に、実環境において、実際に移動を伴って撮影した全方位動画像から 3. で述べた手法でカメラ外部パラメータを推定し、真値と比較することで提案手法の有効性を確認する。

本実験では、計算機シミュレーション・実環境での実験ともに、PointGreyResearch 社製の全方位型マルチカメラシステム Ladybug [23] を用いた。図 4 に示すように、Ladybug には水平方向に五つ、上向きに一つの合計六つのカメラユニットが放射状に配置されており、各カメラユニットはそれぞれ  $768 \times 1024$  画素の解像度の画像を秒間 15 枚撮影可能である。また Ladybug によって、全天球の 75% の視野を一度に撮影できる。本実験において、各カメラユニットの内部パラメータ（焦点距離、アスペクト比、投影中心、ひずみ係数）及びカメラ間の位置・姿勢の関係は、マーカボードとレーザ距離計測機器を用いた手法 [6] によって、あらかじめ測定した。この測定結果に基づく各カメラの視錐台を図 4(b) に示す。このキャリブレーション結果から、水平方向のカメラユニット間の距離



(a) Appearance (b) Viewing volume

図 4 全方位型マルチカメラシステム “Ladybug”

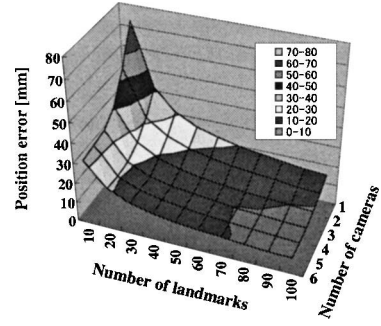
Fig. 4 Omni-directional multi-camera system “Ladybug”.

は  $40 \pm 2$  mm, また垂直方向の距離は  $46 \pm 4$  mm であった.

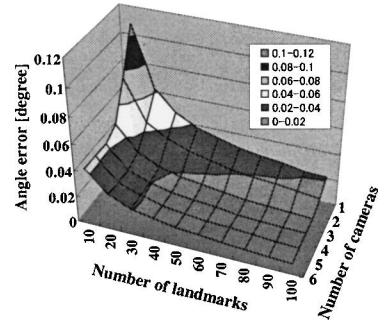
#### 4.1 計算機シミュレーションによる PnP 解法の定量的評価実験

本シミュレーションでは, 2. で述べた PnP 問題の解法によるカメラ外部パラメータの推定精度について, 定量的評価を行う. 本実験では, 仮想空間内で, 実物と同じパラメータをもつ仮想 Ladybug を固定し, その周囲半径 5 m から 50 m の間に, 基準マーカをランダムに配置した. これらの基準マーカは, 仮想 Ladybug の各カメラ画像上に投影された後に, 検出誤差としてガウスノイズが付加され, 更に画素単位に量子化される. 本実験では, ガウスノイズとして, 標準偏差が 1.0 画素のものを用いた.

このような仮想環境内で, 仮想 Ladybug において利用するカメラユニットの数, 及び基準マーカの数を変化させながら, 2. で述べた手法によってカメラ外部パラメータを推定し, カメラ位置と姿勢に関して, 真値との誤差を測定した. 図 5 (a), (b) に測定されたカメラ位置と姿勢の推定誤差を示す. 同図から, (a), (b) とともに, カメラの数と基準マーカ数の増加に伴って, 誤差が単調減少していることが確認できる. 一般的に, PnP 問題においては, 基準マーカの分布範囲のスケール変化は位置の推定誤差のスケール変化としてのみ現れることから, 本実験で想定した 5 ~ 50 m 以外の基準マーカの分布スケールに対しても同様に, カメラ数と基準マーカ数の増加に伴って誤差が単調減少すると考えられる. 特に, 基準マーカ数が一定でもカメラの数が増加すれば推定精度が向上することから, 単眼カメラに比べて OMS を用いることが PnP 問題の解決において推定精度の面で有利であることが分かる.



(a) Position error



(b) Angle error

図 5 カメラ位置・姿勢の推定誤差 (計算機シミュレーション)

Fig. 5 Errors in camera block position and angle (simulation).

#### 4.2 実環境における全方位動画画像からのカメラ外部パラメータの推定実験

3. で述べた特徴点追跡によるカメラ外部パラメータ推定手法の有効性を確認するため, Ladybug を用いて屋内・屋外の二つの異なる実環境を全方位動画として撮影し, 実験を行った. 両実験ともに, 実環境中の自然特徴点の一部をレーザー距離計測器トータルステーション (Leica TCR1105XR) を用いて三次元計測し, 基準マーカとした. また, 本実験では以下の複数のパターンの組合せをキーフレームとして設定し, 基準マーカの画像上の位置は, キーフレームでのみ手動で与えた.

(パターン 1) 初期・最終フレーム

(パターン 2) 初期・中間・最終フレーム

(パターン 3) 初期・最終フレームを含めた 1/4 区間おきのフレーム

##### 4.2.1 屋内環境における実験

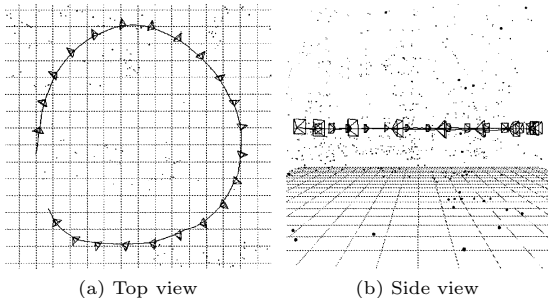
本実験においては, 入力として図 6 に示す屋内環境をそれぞれ 450 フレームからなる六つの動画像系列として歩きながら撮影した. まず, 3. で提案した手法に



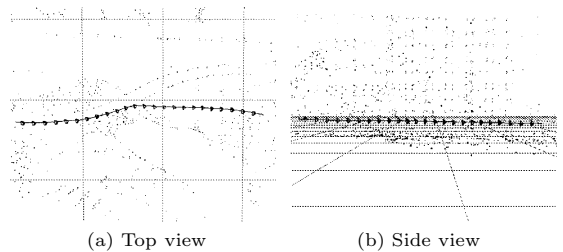
図 6 入力に用いた全方位動画の 1 フレーム (屋内環境)  
 Fig. 6 Sampled frames of input image sequences obtained by six cameras (indoor scene).



図 8 入力に用いた全方位動画の 1 フレーム (屋外環境)  
 Fig. 8 Sampled frames of input image sequences obtained by six cameras (outdoor scene).



(a) Top view (b) Side view  
 図 7 カメラ外部パラメータと自然特徴点の三次元位置の推定結果 (屋内環境)  
 Fig. 7 Result of extrinsic camera parameter estimation (indoor scene).



(a) Top view (b) Side view  
 図 9 カメラ外部パラメータと自然特徴点の三次元位置の推定結果 (屋外環境)  
 Fig. 9 Result of extrinsic camera parameter estimation (outdoor scene).

より、カメラ外部パラメータを推定した。本実験において、自然特徴点は各フレームの画像上で六つのカメラに対して平均 440 点が自動で追跡された。また、逐次処理における自然特徴点の平均投影誤差は 2.1 画素であった。次に、パターン 1~3 のキーフレームの画像上で基準マーカを指定し、それぞれのパターンにおいて、カメラ外部パラメータと自然特徴点の三次元位置を動画全体で最適化した。図 7 に、パターン 1 をキーフレームとした場合の、Ladybug のカメラユニット 1 に関するカメラ外部パラメータの推定結果と自然特徴点の三次元位置の推定結果を示す。同図において、曲線はカメラユニットの移動の軌跡を、錘台は 20 フレームごとのカメラユニットの姿勢を表す。また、黒点は推定された自然特徴点の三次元位置である。本実

験で推定されたカメラパスの長さは 29 m であり、どのパターンのキーフレームを用いた場合にも、パスの不連続は見られなかった。

#### 4.2.2 屋外環境における実験

図 8 に示す複数の建物を含む大学キャンパスにおいて、屋内実験と同様に歩きながら撮影を行った。本実験で入力として用いた六つの動画像系列は、それぞれ 500 フレームからなる。逐次処理による特徴点追跡の結果、各フレームの画像上で六つのカメラに対して平均 530 点の自然特徴点が自動で追跡された。また、自然特徴点の平均投影誤差は 1.6 画素であった。図 9 に、パターン 1 における Ladybug のカメラユニット 1 に関するカメラ外部パラメータの推定結果と、自然特徴点の三次元位置の推定結果を示す。屋内実験と同様に、本実験で推定されたカメラパスの長さは 29 m であり、どのパターンでキーフレームを設定した場合にも、パ



スの不連続は見られなかった。

#### 4. 2. 3 カメラ外部パラメータの定量的評価

推定されたカメラ外部パラメータを真値と比較することで定量的な評価を行った。一般に動画画像のカメラ外部パラメータの真値を測定することは困難であるため、本実験では、基準マーカをすべてのフレームの画像上で手動で指定し、2. で提案した手法を用いてすべてのフレームでカメラ外部パラメータを算出することで、真値を作成した。表 1 に真値の作成に用いた基準マーカの条件を示す。作成された真値は、フレームごとに独立に算出されるため、累積的な推定誤差を含ま

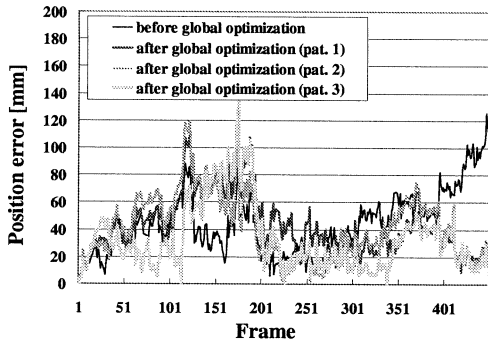
ないが、特徴点位置の指定の誤差及び量子化誤差に起因するカメラ位置・姿勢の推定誤差が発生する。ここではまず、表 1 の条件に基づき、4. 1 で述べた計算機シミュレーションと同様のシミュレーションを行うことで、真値算出における推定誤差の平均値を算出した。ただし、基準マーカを手動で指定する際の誤差として、1.0 画素の標準偏差をもつガウスノイズ及び画素単位での量子化誤差を想定した。計算機シミュレーションの結果から、本実験で作成したカメラ外部パラメータの真値には、少なくとも屋内環境での実験に関して各フレームで平均 5.4 mm, 0.014°, 屋外環境で平均 15 mm, 0.014° のカメラ位置・姿勢の推定誤差が含まれていることが分かった。

次に、屋内環境、屋外環境において推定されたカメラ外部パラメータと作成された真値を比較し、カメラ位置・姿勢の推定誤差を定量的に評価した。図 10、図 11 に、それぞれ、屋内環境、屋外環境における各パターンでのカメラ位置・姿勢の推定誤差を示す。また、キーフレームの与え方の各パターンにおける推定

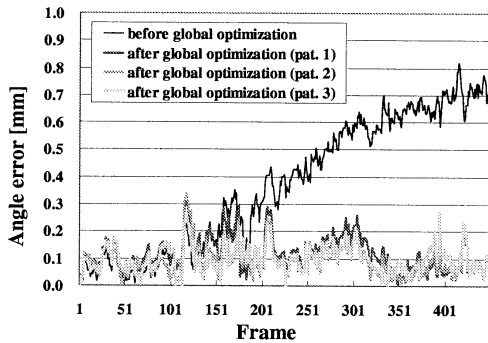
表 1 カメラ外部パラメータの真値の作成に用いた基準マーカの条件

Table 1 Conditions of feature landmarks for generating ground truth of extrinsic camera parameters.

	indoor	outdoor
average number of feature landmarks	53	51
average distribution range of feature landmarks (m)	5 ~ 22	10 ~ 61



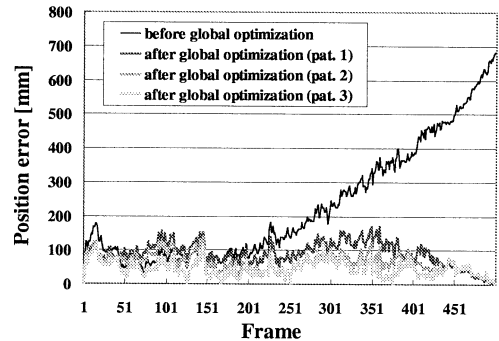
(a) Position error



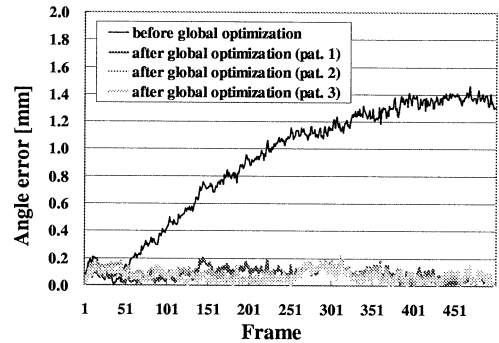
(b) Angle error

図 10 カメラ位置・姿勢の推定誤差 (屋内環境)

Fig. 10 Errors in estimated camera path and posture (indoor scene).



(a) Position error



(b) Angle error

図 11 カメラ位置・姿勢の推定誤差 (屋外環境)

Fig. 11 Errors in estimated camera path and posture (outdoor scene).

表 2 カメラ位置・姿勢の推定誤差

Table 2 Estimation errors in position and angle.

(a) Position errors (average/standard deviation) [mm]

	indoor	outdoor
before global optimization	44/23.6	230/173.9
after global optimization (pat. 1)	43/22.2	90/34.1
after global optimization (pat. 2)	41/23.5	66/26.9
after global optimization (pat. 3)	33/24.9	47/22.6

(b) Angle errors (average/standard deviation) [degree]

	indoor	outdoor
before global optimization	0.37/0.25	0.92/0.44
after global optimization (pat. 1)	0.12/0.06	0.09/0.04
after global optimization (pat. 2)	0.09/0.06	0.08/0.04
after global optimization (pat. 3)	0.09/0.05	0.07/0.04

誤差の平均値と標準偏差を表 2 に示す。これらの結果から、どのパターンにおいても、全体最適化によって累積的なカメラパラメータの推定誤差が低減していることが確認できる。また、キーフレームを増やすことで、推定精度の向上を図ることが可能であることを確認できる。先に述べた真値算出時の推定誤差を考慮すれば、パターン 3 におけるカメラ位置の推定誤差の平均値は、屋内環境で 27~39 mm、屋外環境で 32~62 mm の範囲であると考えられる。本実験で利用した OMS:Ladybug において、カメラ位置に関して 40 mm 程度の位置の推定誤差の影響は、OMS から 20 m 以上離れば画像上への投影位置で 1 画素以内に収まる [6] ため、両実験環境において推定されたカメラ外部パラメータは、画素単位での推定精度を必要とするような画像解析による三次元モデル復元等への応用が可能なレベルであるといえる。

## 5. む す び

全方位型マルチカメラシステム (OMS) によって得られる全方位動画画像から、カメラブロックの外部パラメータを推定する手法を提案し、定量的な評価実験を行った。本論文では、まず、OMS における PnP 問題とその解法について述べた。次に、キーフレームで基準マーカを指定し、自然特徴点を自動追跡することで、全方位動画画像のカメラ外部パラメータを精度良く推定する手法を提案した。

実験では、まず計算機シミュレーションによって、提案した PnP 問題の解法によるカメラ外部パラメータの推定精度について検証し、OMS を用いたカメラパラメータの推定が、単眼カメラに比べて有利であることを示した。次に、屋内・屋外の実環境において精度良くカメラ外部パラメータを推定できることを定量

的に示した。今後、提案手法で推定したカメラ外部パラメータを用い、マルチベースラインステレオ法 [24] を適用することで、屋外環境の三次元モデル化を行う予定である。

## 文 献

- [1] K. Miyamoto, "Fish eye lens," J. Opt. Soc. Am., vol.54, no.2, pp.1060-1061, 1964.
- [2] K. Yamazawa, Y. Yagi, and M. Yachida, "Omnidirectional imaging with hyperboloidal projection," Proc. Int. Conf. on Intelligent Robots and Systems, vol.2, pp.1029-1034, 1993.
- [3] S.K. Nayar, "Catadioptric omnidirectional cameras," Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, pp.482-488, 1997.
- [4] J. Shimamura, H. Takemura, N. Yokoya, and K. Yamazawa, "Construction and presentation of a virtual environment using panoramic stereo images of a real scene and computer graphics models," Proc. 15th IAPR Int. Conf. on Pattern Recognition, vol.IV, pp.463-467, 2000.
- [5] H. Tanahashi, K. Yamamoto, C. Wang, and Y. Niwa, "Development of a stereo omni-directional imaging system (SOS)," Proc. IEEE Int. Conf. on Industrial Electronics, Control and Instrumentation, pp.289-294, 2000.
- [6] S. Ikeda, T. Sato, and N. Yokoya, "High-resolution panoramic movie generation from video streams acquired by an omnidirectional multi-camera system," Proc. IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent System, pp.155-160, 2003.
- [7] R. Horand, B. Conio, and O. Le Boulleux, "An analytic solution for the perspective 4-point problem," Comput. Vis. Graph. Image Process., vol.47, pp.33-44, 1989.
- [8] J.S.C. Yuan, "A general photogrammetric method for determining object position and orientation," IEEE Trans. Robot. Autom., vol.5, no.2, pp.129-142, 1989.
- [9] R. Krishnan and H.J. Sommer, "Monocular pose of a rigid body using point landmarks," Comput. Vis. Image Understand., vol.55, pp.307-316, 1992.
- [10] R. Klette, K. Schluns, and A. Koschan, eds., Computer Vision: Three-dimensional Data from Image, Springer, 1998.
- [11] P. Beardsley, A. Zisserman, and D. Murray, "Sequential updating of projective and affine structure from motion," Int. J. Comput. Vis., vol.23, no.3, pp.235-259, 1997.
- [12] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," Int. J. Comput. Vis., vol.9, no.2, pp.137-154, 1992.
- [13] M. Pollefeys, R. Koch, M. Vergauwen, A.A. Deknuydt, and L.J.V. Gool, "Three-dimensional scene reconstruction from images," Proc. SPIE,

- vol.3958, pp.215–226, 2000.
- [14] J. Gluckman and S. Nayer, “Ego-motion and omnidirectional cameras,” Proc. 6th Int. Conf. on Computer Vision, pp.999–1005, 1998.
- [15] M. Etoh, T. Aoki, and K. Hata, “Estimation of structure and motion parameters for a roaming robot that scans the space,” Proc. 7th Int. Conf. on Computer Vision, vol.I, pp.579–584, 1999.
- [16] C.J. Taylor, “VideoPlus,” Proc. IEEE Workshop on Omnidirectional Vision, pp.3–10, 2000.
- [17] R. Szeliski and S.B. Kang, “Recovering 3D shape and motion from image streams using non-linear least squares,” J. Vis. Commun. Image Represent., vol.6, no.1, pp.10–28, 1994.
- [18] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon, “Bundle adjustment a modern synthesis,” Proc. Int. Workshop on Vision Algorithms, pp.298–372, 1999.
- [19] T. Sato, M. Kanbara, N. Yokoya, and H. Takemura, “Dense 3-D reconstruction of an outdoor scene by hundreds-baseline stereo using a hand-held video camera,” Int. J. Comput. Vis., vol.47, no.1-3, pp.119–129, 2002.
- [20] K. Kanatani, Statistical Optimization for Geometric Computation: Theory and Practice, Elsevier Science, 1998.
- [21] C. Harris and M. Stephens, “A combined corner and edge detector,” Proc. Alvey Vision Conf., pp.147–151, 1988.
- [22] M.A. Fischler and R.C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” Commun. ACM, vol.24, no.6, pp.381–395, 1981.
- [23] Point Grey Research Inc., “Ladybug,” <http://www.ptgrey.com/products/ladybug/index.html>
- [24] M. Okutomi and T. Kanade, “A multiple-baseline stereo,” IEEE Trans. Pattern Anal. Mach. Intell., vol.15, no.4, pp.353–363, 1993.

(平成 16 年 6 月 3 日受付, 9 月 14 日再受付)



佐藤 智和 (正員)

1999 阪府大・工・情報工卒。2003 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 本会学術奨励賞受賞。情報処理学会, IEEE 各会員。



池田 聖 (学生員)

2001 広島大・理・物理卒。2003 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了。現在, 同大博士後期課程に在学中。IEEE 会員。



横矢 直和 (正員)

1974 阪大・基礎工・情報工卒。1979 同大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。1986~1987 マツギル大・知能機械研究センター客員教授。1992 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 情報処理学会論文賞受賞。工博。情報処理学会, 日本バーチャルリアリティ学会, 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。