Vol.2010-CVIM-173 No.4
2010/9/5

平面を利用した直接法による カメラキャリブレーション

杉 本 茂 樹[†1]　奥 富 正 敏[†1]

本稿では，レンズ歪みを含むカメラパラメータのキャリブレーションを，直接法による画像レジストレーションに基づいて行う手法を提案する．本手法では，任意のテクスチャを持つ平面を撮影した複数の画像を利用して，これらの画像が全て一致するように，最適化手法を用いてカメラパラメータを推定する．直接法では，全ての画素が推定に寄与できるため，画像中の特徴点を対応を利用する従来法よりも高精度な推定が期待できる．また，チェスボード等の特別なパターンを必要としないだけでなく，特徴点の入力も不要であることから，校正物体を利用する手法よりも手順を簡略化することが可能と考えられる．合成画像および実画像実験を用いて本手法の有効性を検証した結果を報告する．

A Direct Method for Camera Calibration from Planar Scene

Shigeki Sugimoto[†1] and Masatoshi Okutomi[†1]

In this paper we present a direct method for camera self-calibration from planar scenes. We estimate the camera parameters including lens distortion coefficients that minimizes a SSD (sum of squared differences) function composed with the pixel values of multiple images viewing an arbitrary textured plane. The contribution of every pixel to the measurements can improve the precision of the estimated parameters compared with existing feature-based self-calibration techniques. In addition, this method enables us to offer an easy calibration interfaces which doesn't require specific calibration patterns on the calibration target, such as chessboard patterns, nor user input of corner positions on the target. We report the validity of the method through the experiments using synthetic images and real images.

†1 東京工業大学大学院理工学研究科機械制御システム専攻.
Department of Mechanical and Control Engineering, Graduate School of Science and Engineering, Tokyo Institute of Technology.

1. はじめに

カメラの焦点距離やレンズ歪みなど，カメラ固有の幾何学的特性を求めるカメラキャリブレーションは，ステレオ法などの 3 次元計測には不可欠な技術である．キャリブレーション手法には，3 次元情報が既知の校正用物体を利用する方法と，特定の校正用物体を利用しないセルフキャリブレーションがある．

前者の場合，キャリブレーションのための特別なターゲットを用意する必要があるものの，比較的容易に作成可能な校正用平面を利用した Zhang の手法[1] と，それを実装・改良したツールやライブラリ[2),3)] の普及により，近年では，技術者であれば比較的手軽にキャリブレーションを行うことができるようになっている．一方，後者の場合，特別なターゲットは不要なことや，カメラごとに異なるパラメータを求めることが可能などの利点がある[4)] ものの，基本的にシーンの 3 次元情報の復元を伴うため，問題の自由度が大きくなる．このため，多数の画像を利用した structure from motion において，バンドル調整パラメータに含める方法[5)] が代表的である．いずれの方法においても，最近の研究のトレンドは，全方位カメラなどの特殊カメラのためのモデル記述とそのキャリブレーション[6)–8)] へと移行している．

一方で，これまでのキャリブレーション手法は，全て画像間の特徴点を利用した手法であることに気がつく．すなわち，これらの方法では，画像中の疎な部分のみを用いて，画像全体の特性を記述するモデルパラメータを推定している．これに対し，直接法[9)] を利用したキャリブレーション方法が考えられる．直接法は，画像間の輝度値の最小化に基づいた方法であるため，画像中の全画素が計測に寄与できる．これにより，特徴点対応に基づく手法と比較して精度の向上が期待できる他，より自由度の高いカメラモデルにも拡張できる可能性がある．また，画像間のレジストレーションを行う直接法に基づく方法は基本的にセルフキャリブレーションとなり，これにより，特徴点の入力が必要な校正用物体を利用した既存手法よりもキャリブレーションの手間を簡略化することが可能と考えられる．

そこで，本稿では，直接法に基づくキャリブレーション手法を提案する．ここでは，任意のテクスチャを持つ平面を撮影した複数の画像を用い，直接法による画像レジストレーションに基づいて，標準的なカメラモデルパラメータの推定を行う．より具体的には，レンズ歪みを含むカメラパラメータ用いて画像間の座標変換を定式化し，画像輝度値の SSD (Sum

© 2010 Information Processing Society of Japan

Vol.2010-CVIM-173 No.4
2010/9/5

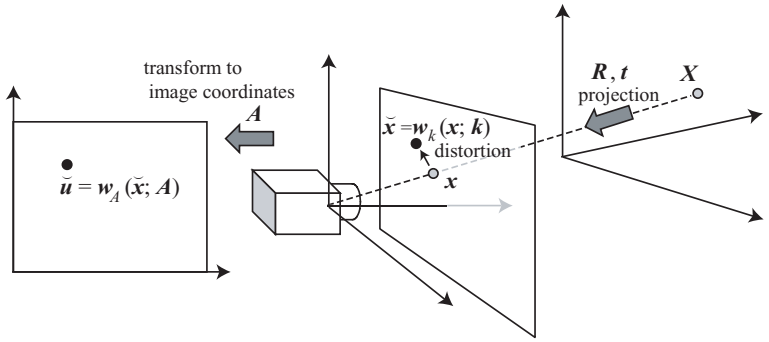


図 1 Camera model

of Squared Differences) で表されたコスト関数を最小化することによってパラメータを推定する．定式化では，画像間の座標変換にレンズ歪みの逆変換を含んでしまうと最適化が極めて難しくなるため，その逆変換を含まないような工夫を行っている．さらに，基準画像と他の画像とのペアにおいて，基準画像を他方の画像へ合わせる backward warp コストを用いることで，収束性能の向上を図っている．合成画像と実画像を用いた実験を通じ，本手法の有効性を検証する．

2. 撮像モデルおよび画像間の座標変換

本手法では，標準的なカメラモデル[1),10)] のパラメータをキャリブレーション対象とする．本節では，このカメラモデルを概説するとともに，そのモデルに基づいて，レンズ歪みをもつカメラで平面を撮影した際に生じる画像間の座標変換を記述する．

2.1 カメラモデル

本手法で利用するカメラモデルを図 1 に示す．同図では，空間中の点 $\boldsymbol{X} = (X, Y, Z)^T$ が正規化画像上の点 $\boldsymbol{x} = (x, y)^T$ へ投影され，次いで点 $\boldsymbol{x}$ がレンズ歪みによって点 $\breve{\boldsymbol{x}} = (\breve{x}, \breve{y})^T$ に移動し，最後に正規化画像上の点 $\breve{\boldsymbol{x}}$ が画像上の点 $\breve{\boldsymbol{u}} = (\breve{u}, \breve{v})^T$ （画素単位）に変換される様子を示している．ここでは，レンズ歪が付与された座標を ˘ 付きで表している．また，以下では，点の同次座標を ˜ 付きで表し，同値関係（定数倍の不定性を許して等しい関係）を ∼ で表すものとする．

3 次元座標 $\boldsymbol{X}$ から正規化画像座標 $\boldsymbol{x}$ への投影は次式で表される．

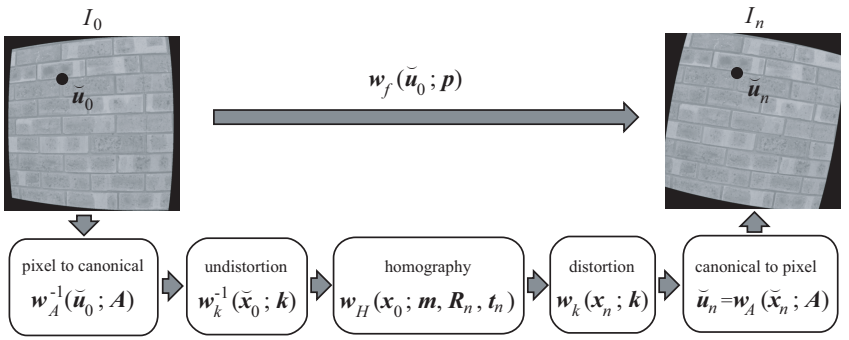


図 2 Coordinate transform

$$\tilde{\boldsymbol{u}} \sim [\boldsymbol{R}(\boldsymbol{\omega})|\boldsymbol{t}]\tilde{\boldsymbol{x}}_w \tag{1}$$

ただし，$\boldsymbol{R}(\boldsymbol{\omega})$ は，3 次元ベクトル $\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3)^T$ によってパラメータ化された回転行列を示し，Rodrigues の回転公式[11)] によって表される．また，$\boldsymbol{t} = (t_1, t_2, t_3)^T$ は並進ベクトルを示す．

レンズ歪みによる正規化画像上の $\boldsymbol{x}$ から $\breve{\boldsymbol{x}}$ への変換は次のように表される[10)]．

$$\breve{\boldsymbol{x}} = \begin{cases} x(1 + k_1 r^2 + k_2 r^4) + 2k_3 xy + k_4(r^2 + x^2) \\ y(1 + k_1 r^2 + k_2 r^4) + k_3(r^2 + x^2) + 2k_4 xy \end{cases}, \quad \text{where} \quad r = \sqrt{x^2 + y^2} \tag{2}$$

ただし，k_1, k_2 は半径方向歪み係数，k_3, k_4 は接線方向歪み係数を表す．以下本稿では，$\boldsymbol{k} = (k_1, k_2, k_3, k_4)^T$ とし，式 (2) の変換を，$\breve{\boldsymbol{x}} = \boldsymbol{w}_k(\boldsymbol{x}; \boldsymbol{k})$ のようにベクトル値関数で表記する．

正規化画像座標から画像座標への変換は次式で表される．

$$\breve{\tilde{\boldsymbol{u}}} = \boldsymbol{A}\breve{\tilde{\boldsymbol{x}}} \tag{3}$$

ただし，$\boldsymbol{A}$ は 3×3 の上三角内部パラメータ行列である．以下本稿では，式 (3) の変換を $\breve{\boldsymbol{u}} = \boldsymbol{w}_A(\breve{\boldsymbol{x}}; \boldsymbol{A})$ のようにベクトル値関数で表記する．

© 2010 Information Processing Society of Japan

Vol.2010-CVIM-173 No.4
2010/9/5

2.2 画像間の座標変換

静止した平面を異なる視点位置から撮影した $N+1$ 枚の画像を考え[*1]，基準画像を I_0，その他の画像を $I_n, (n=1,2,\cdots,N)$ とする．また，基準画像上の座標を $\breve{\boldsymbol{u}}_0$，他方の画像上の座標を $\breve{\boldsymbol{u}}_n$ とする．$\boldsymbol{p}$ をすべての座標変換パラメータを含むベクトルを表すものとすると，$\breve{\boldsymbol{u}}_0$ から $\breve{\boldsymbol{u}}_n$ への座標変換 $\boldsymbol{w}_f(\breve{\boldsymbol{u}}_0;\boldsymbol{p})$ は，次式のように表される (図 2)．

$$\breve{\boldsymbol{u}}_n = \boldsymbol{w}_f(\breve{\boldsymbol{u}}_0;\boldsymbol{p}) = \boldsymbol{w}_A(\breve{\boldsymbol{x}}_n;\boldsymbol{A}) \circ \boldsymbol{w}_k(\boldsymbol{x}_n;\boldsymbol{k}) \circ \boldsymbol{w}_H(\boldsymbol{x}_0;\boldsymbol{m},\boldsymbol{\omega}_n,\boldsymbol{t}_n) \circ \boldsymbol{w}_k^{-1}(\breve{\boldsymbol{x}}_0;\boldsymbol{k}) \circ \boldsymbol{w}_A^{-1}(\breve{\boldsymbol{u}}_0;\boldsymbol{A}) \quad (4)$$

ただし，$\circ$ は変換の合成を示しており，$\boldsymbol{w}_A^{-1}, \boldsymbol{w}_k^{-1}$ は，それぞれ $\boldsymbol{w}_A, \boldsymbol{w}_k$ の逆変換を示す．また，$\boldsymbol{w}_H(\boldsymbol{x}_0;\boldsymbol{m},\boldsymbol{\omega}_n,\boldsymbol{t}_n)$ は平面射影変換を示すベクトル値関数であり，その平面射影変換は次のように表される[12]．

$$\tilde{\boldsymbol{x}}_n \sim \boldsymbol{H}\tilde{\boldsymbol{x}}_0 \quad \text{where} \quad \boldsymbol{H} = \boldsymbol{R}(\boldsymbol{\omega}_n) + \boldsymbol{t}_n\boldsymbol{m}^T \quad (5)$$

ここで，$\boldsymbol{m} = (m_1, m_2, m_3)^T$ は，基準カメラ座標系において平面を一意に定める平面パラメータベクトルである．本稿では，基準画像座標の他方への変換 $\boldsymbol{w}_f$ を forward warp と呼ぶ．

3. 直接法によるカメラキャリブレーション

本手法の目的は，任意のテクスチャを持つ平面を撮影した複数の画像から，$\boldsymbol{A}, \boldsymbol{k}, \boldsymbol{m}$ および $\boldsymbol{\omega}_n, \boldsymbol{t}_n, (n=1,2,\cdots,N)$ を推定することにある．本節では，直接法を利用してこれらのパラメータを推定する方法について述べる．

3.1 Forward Warp Cost

式 (4) の座標変換を用いることにより，最小化すべきコスト関数が次式のように定義できる．

$$C_f = \sum_{n=1}^{N} \sum_{\breve{\boldsymbol{u}}_0 \in \mathrm{ROI}_0} \left(I_0[\breve{\boldsymbol{u}}_0] - I_n[\boldsymbol{w}_f(\breve{\boldsymbol{u}}_0;\boldsymbol{p})]\right)^2 \quad (6)$$

ただし，ROI_0 は，基準画像上に定めた処理領域 (ROI: region of interest) を示す．

しかし，式 (6) を最小化することは極めて困難である．これは，$\boldsymbol{w}_f$ にはレンズ歪みの逆変換 $\boldsymbol{w}_k^{-1}$ が含まれることに起因する．式 (6) の最小化のためには，基本的に $\boldsymbol{w}_f$ の偏微分

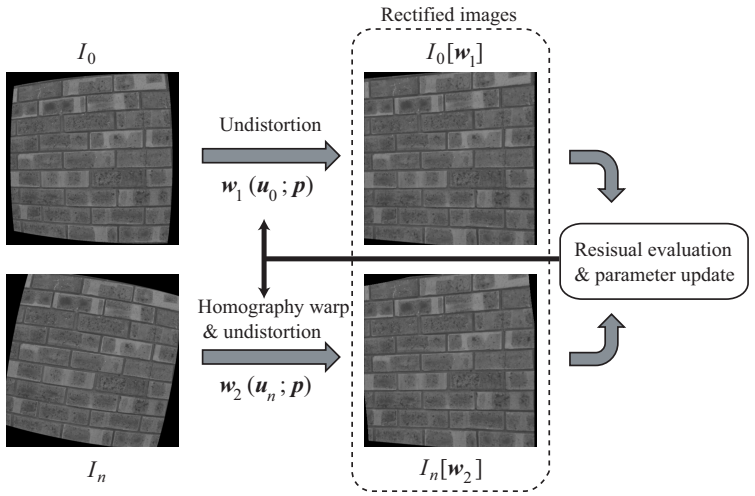


図 3 Parameter update by fowrad warp cost

計算が必要となるが，$\boldsymbol{w}_k^{-1}$ を解析的に計算できないため，$\boldsymbol{w}_f$ を偏微分することは難しい．

そこで，次のような等価なコスト関数を考える．

$$C'_f = \sum_{n=1}^{N} \sum_{\breve{\boldsymbol{u}}_0 \in \mathrm{ROI}_0} \left(I_0[\boldsymbol{w}_1(\boldsymbol{u}_0;\boldsymbol{p})] - I_n[\boldsymbol{w}_2(\boldsymbol{u}_0;\boldsymbol{p})]\right)^2, \quad (7)$$

where

$$\boldsymbol{w}_1(\boldsymbol{u}_0;\boldsymbol{p}) = \boldsymbol{w}_A(\breve{\boldsymbol{x}}_0;\boldsymbol{A}) \circ \boldsymbol{w}_k(\boldsymbol{x}_0;\boldsymbol{k}) \circ \boldsymbol{w}_A^{-1}(\boldsymbol{u}_0;\boldsymbol{A}), \quad (8)$$

$$\boldsymbol{w}_2(\boldsymbol{u}_0;\boldsymbol{p}) = \boldsymbol{w}_A(\breve{\boldsymbol{x}}_n;\boldsymbol{A}) \circ \boldsymbol{w}_k(\boldsymbol{x}_n;\boldsymbol{k}) \circ \boldsymbol{w}_H(\boldsymbol{x}_0;\boldsymbol{m},\boldsymbol{\omega}_n,\boldsymbol{t}_n) \circ \boldsymbol{w}_A^{-1}(\boldsymbol{u}_0;\boldsymbol{A}). \quad (9)$$

$\boldsymbol{w}_1, \boldsymbol{w}_2$ は，式 (4) において，$\boldsymbol{w}_H \circ \boldsymbol{w}_k^{-1}$ の間に単位変換 $\boldsymbol{e} = \boldsymbol{w}_A^{-1} \circ \boldsymbol{w}_A$ を挿入することによって得られる．すなわち，

$$\boldsymbol{w}_f(\breve{\boldsymbol{u}}_0;\boldsymbol{p}) = \boldsymbol{w}_A(\breve{\boldsymbol{x}}_n;\boldsymbol{A}) \circ \boldsymbol{w}_k(\boldsymbol{x}_n;\boldsymbol{k}) \circ \boldsymbol{w}_H(\boldsymbol{x}_0;\boldsymbol{m},\boldsymbol{\omega}_n,\boldsymbol{t}_n) \circ \boldsymbol{w}_A^{-1}(\boldsymbol{u}_0;\boldsymbol{A}) \circ \boldsymbol{w}_A(\boldsymbol{x}_0;\boldsymbol{A}) \circ \boldsymbol{w}_k^{-1}(\breve{\boldsymbol{x}}_0;\boldsymbol{k}) \circ \boldsymbol{w}_A^{-1}(\breve{\boldsymbol{u}}_0;\boldsymbol{A}) \quad (10)$$

$$= \boldsymbol{w}_2(\boldsymbol{u}_0;\boldsymbol{p}) \circ \boldsymbol{w}_1^{-1}(\boldsymbol{u}_0;\boldsymbol{p}) \quad (11)$$

と書き，$\boldsymbol{w}_1^{-1}$ の逆変換を基準画像に作用させることにより，式 (7) が得られる．このようにして，$\boldsymbol{w}_k^{-1}$ を含まないコスト関数を定義できる．

図 3 は，式 (7) の 2 つの座標変換関数 $\boldsymbol{w}_1, \boldsymbol{w}_2$ の役割を示している．式 (8) は，キャリブ

[*1] 実際のキャリブレーションでは，静止したカメラを用いて様々な姿勢の平面を撮影することを想定しているが，カメラが動き平面が静止したケースを想定しても一般性は失わない．

© 2010 Information Processing Society of Japan

レーション済みの画像の歪みを除去する際に広く利用されている座標変換であり，パラメータ $\boldsymbol{k}$ が正しければ，$I_0[\boldsymbol{w}_1]$ は基準画像の歪みが完全に除去された結果となる．一方，式 (9) は，式 (8) に平面射影変換が加わったものである．同図に示すように，式 (7) を最小化することは，現在までに推定されたパラメータを用いて画像をレクティフィケーションし，その残差をパラメータ更新に反映させるアプローチを意味する．

一方で，式 (6) と式 (7) とでは，座標変換関数の入力座標が異なっていることに注意が必要である．式 (6) における座標変換 $\boldsymbol{w}_f$ の入力は観測された基準画像上の座標 $\breve{\boldsymbol{u}}_0$ であったのに対し，式 (7) で利用される $\boldsymbol{w}_1, \boldsymbol{w}_2$ の入力は歪みのない基準画像座標 $\boldsymbol{u}_0$ である．$\boldsymbol{u}_0$ の値は未知であるため，レクティフィケーションされた画像を想定し，その画像座標を入力として利用する．

ただし，この場合，観測した基準画像中の平面テクスチャをすべてを有効に活用できないことがある．例えば，図 3 で示しているように，I_0 で観測した平面の一部（右上と右下）が $I_0[\boldsymbol{w}_1]$ では欠けてしまうことがあり，このようなケースでは，欠けた部分のテクスチャは推定に寄与しない．さらに，基準画像のサンプリングが疎になることで，$I_0[\boldsymbol{w}_1]$ にエイリアシングなどのアーティファクトが生じる可能性もある．このような問題を避けるためには，レクティフィケーション画像を大きめにすることや，その解像度を高めにするなどの対処が有効と考えられる．その他，$\boldsymbol{u}_0$ は基準画像上に定めた ROI_0 とは一致しないため，$\boldsymbol{w}_1$ が ROI_0 内に入っているかのチェックが必要である．

3.2 Backwarp Warp Cost

一般的な 2 枚の画像のレジストレーションでは，どちらか一方をテンプレートとし，他方の画像をテンプレートに合わせる．このような forward warp を用いたコストの最小化に対し，基準画像を他方の画像に合わせるための backward warp を用いたコスト関数を合わせて利用することで，レジストレーションの安定性や精度の改善が可能である．このような方法は，パラメータ数が多く容易に局所解に陥りやすい非剛体レジストレーションにおいて利用されている[13)–15)]．

本問題においても，パラメータ数が比較的多いことや，カメラの回転と並進に内在する曖昧性[16)]，および画角（焦点距離）と半径方向歪みとの間の曖昧性[6)] などによる局所解が考えられる．また，両方の画像をサンプリングすることによって生じるぼけがレジストレーションの真値まわりでの収束性能を劣化させる可能性があり，これらのことを考慮すると，backward warp を利用する方法が有効と考えられる．

一方で，本手法では複数画像を利用しているため，基準画像と n 番目画像のペア以外でも画像ペアを組むことが可能である．しかし，最適化における計算コストが膨大になるため，ここでは基準画像との画像ペアのみを利用する．

Backward warp を用いたコストは次式で表される．

$$C'_b = \sum_{n=1}^{N} \sum_{\breve{\boldsymbol{u}}_n \in \mathrm{ROI_n}} \left(I_n[\boldsymbol{w}_3(\boldsymbol{u}_n; \boldsymbol{p})] - I_0[\boldsymbol{w}_4(\boldsymbol{u}_n; \boldsymbol{p})] \right)^2, \tag{12}$$

where

$$\boldsymbol{w}_3(\boldsymbol{u}_n; \boldsymbol{p}) = \boldsymbol{w}_A(\breve{\boldsymbol{x}}_n; \boldsymbol{A}) \circ \boldsymbol{w}_k(\boldsymbol{x}_n; \boldsymbol{k}) \circ \boldsymbol{w}_A^{-1}(\boldsymbol{u}_n; \boldsymbol{A}), \tag{13}$$

$$\boldsymbol{w}_4(\boldsymbol{u}_n; \boldsymbol{p}) = \boldsymbol{w}_A(\breve{\boldsymbol{x}}_0; \boldsymbol{A}) \circ \boldsymbol{w}_k(\boldsymbol{x}_0; \boldsymbol{k}) \circ \boldsymbol{w}_H^{-1}(\boldsymbol{x}_n; \boldsymbol{m}, \boldsymbol{\omega}_n, \boldsymbol{t}_n) \circ \boldsymbol{w}_A^{-1}(\boldsymbol{u}_n; \boldsymbol{A}). \tag{14}$$

ただし，$\mathrm{ROI_n}$ は n 番目の画像に定めた処理領域を示す．また，$\boldsymbol{w}_3$ は，$\boldsymbol{w}_1$ と全く同じベクトル値関数であり，入力座標のみが異なる．さらに，$\boldsymbol{w}_4$ は，$\boldsymbol{w}_2$ における $\boldsymbol{w}_H$ がその逆変換になったものであり，$\boldsymbol{w}_H^{-1}$ は，Sherman-Morrison の公式[17)] を用いて次式のように書ける．

$$\tilde{\boldsymbol{x}}_0 \sim \boldsymbol{H}^{-1} \tilde{\boldsymbol{x}}_n \quad \text{where} \quad \boldsymbol{H}^{-1} = \boldsymbol{R}^T(\boldsymbol{\omega}_n) \left(\frac{\boldsymbol{I} - \boldsymbol{t}_n \boldsymbol{m}^T \boldsymbol{R}^T(\boldsymbol{\omega}_n)}{1 + \boldsymbol{m}^T \boldsymbol{R}^T(\boldsymbol{\omega}_n) \boldsymbol{t}_n} \right) \tag{15}$$

3.3 最 適 化

本手法で最小化するコストは，上述した 2 つのコストを加算した次式のようなものである．

$$C(\boldsymbol{p}) = C'_f(\boldsymbol{p}) + C'_b(\boldsymbol{p}) \tag{16}$$

本手法では，Gauss-Newton 法を利用した最適化を利用する．パラメータの推定値を $\bar{\boldsymbol{p}}$ とし，更新量を $\Delta\boldsymbol{p}$ とする．式 (16) に Gauss-Newton 法を適用すると，次式が得られる．

$$\Delta\boldsymbol{p} = -\boldsymbol{Q}^{-1}\boldsymbol{b} \tag{17}$$

where

$$\boldsymbol{Q} = \sum_{n=1}^{N} \left(\sum_{\breve{\boldsymbol{u}}_0 \in \mathrm{ROI_0}} (\boldsymbol{J}_1 - \boldsymbol{J}_2)^T (\boldsymbol{J}_1 - \boldsymbol{J}_2) + \sum_{\breve{\boldsymbol{u}}_n \in \mathrm{ROI_n}} (\boldsymbol{J}_3 - \boldsymbol{J}_4)^T (\boldsymbol{J}_3 - \boldsymbol{J}_4) \right) \tag{18}$$

$$\boldsymbol{b} = \sum_{n=1}^{N} \left(\sum_{\breve{\boldsymbol{u}}_0 \in \mathrm{ROI_0}} (\boldsymbol{J}_1 - \boldsymbol{J}_2)^T (I_0[\boldsymbol{w}_1(\boldsymbol{u}_0; \bar{\boldsymbol{p}})] - I_n[\boldsymbol{w}_2(\boldsymbol{u}_0; \bar{\boldsymbol{p}})]) + \sum_{\breve{\boldsymbol{u}}_n \in \mathrm{ROI_n}} (\boldsymbol{J}_3 - \boldsymbol{J}_4)^T (I_n[\boldsymbol{w}_3(\boldsymbol{u}_n; \bar{\boldsymbol{p}})] - I_0[\boldsymbol{w}_4(\boldsymbol{u}_n; \bar{\boldsymbol{p}})]) \right) \tag{19}$$

© 2010 Information Processing Society of Japan

Vol.2010-CVIM-173 No.4
2010/9/5

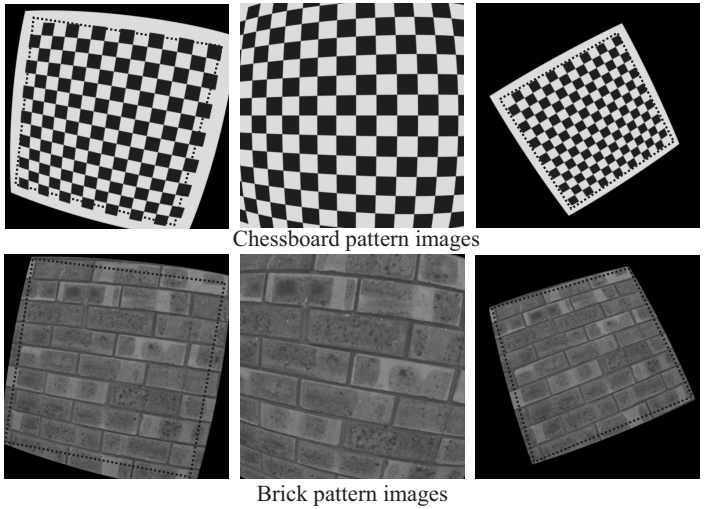


図 4 Examples of synthetic images

ただし，$\boldsymbol{J}_1 = \frac{\partial I_0}{\partial \boldsymbol{w}_1}\frac{\partial \boldsymbol{w}_1}{\partial \boldsymbol{p}}$, $\boldsymbol{J}_2 = \frac{\partial I_n}{\partial \boldsymbol{w}_2}\frac{\partial \boldsymbol{w}_2}{\partial \boldsymbol{p}}$, $\boldsymbol{J}_3 = \frac{\partial I_n}{\partial \boldsymbol{w}_3}\frac{\partial \boldsymbol{w}_3}{\partial \boldsymbol{p}}$, $\boldsymbol{J}_4 = \frac{\partial I_0}{\partial \boldsymbol{w}_4}\frac{\partial \boldsymbol{w}_4}{\partial \boldsymbol{p}}$ である．

実装上の注意点としては，ヘッセ行列 $\boldsymbol{Q}$ の条件数を改善するために画像座標には正規化なこと，および，正規化画像座標上で原点から遠く離れた点において，レンズ歪みを適用した結果が折り返さないように，折り返しを検出することが挙げられる．

4. 実験結果

本手法の有効性を示すために，合成画像および実画像を用いた比較実験を行った．推定対象のパラメータは，$\boldsymbol{A}, \boldsymbol{k}, \boldsymbol{m}, \boldsymbol{\omega}_n, \boldsymbol{t}_n, (n = 1, 2, \cdots, N)$ であるが，$\boldsymbol{A}$ については，対角成分 2 個 (a_{11}, a_{22}) と画像中心 (a_{13}, a_{23}) の計 4 個の要素を推定対象とする．また，カメラ並進 $\boldsymbol{t}$ と平面姿勢パラメータ $\boldsymbol{m}$ の間にはスケール不定性が存在するため，$\boldsymbol{m}$ の第 3 成分は 1 に固定している．本手法の実装は C++言語で行い，Windows-XP の PC(Intel Core2Quad 2.8GHz) 上で実行させた．

4.1 合成画像実験

チェスボードパターンと煉瓦パターンの平面テクスチャ用いて合成画像を生成した．画像は 512×512[pix] サイズのモノクロ画像であり，共通パラメータとして，内部パラメータを

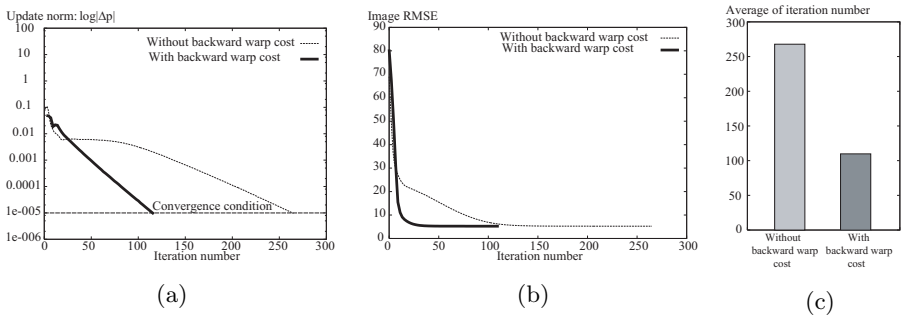


図 5 Difference of convergence behavior with/without backward warp cost: (a) Update norm, (b) Image RMSE, (c) Mean iteration number over 30 trials.

$(a_{11}, a_{22}, a_{13}, a_{23}) = (512, 512, 255.5, 255.5)$，レンズ歪みパラメータを $\boldsymbol{k} = (-0.25, 0, 0, 0)$ とした．

合成画像生成における $\boldsymbol{m}$ および $\boldsymbol{\omega}_n, \boldsymbol{t}_n$ の設定方法は以下のとおりである．テクスチャパターンを 1 m 四方の平面板とみなして物体座標系 $\boldsymbol{X}_p$ を設定し，$\boldsymbol{X}_p$ から i 番目のカメラ座標系 $\boldsymbol{X}_i, (i = 0, 1, \cdots, N)$ への座標変換 $\boldsymbol{X}_i = \boldsymbol{R}(\boldsymbol{\omega}_i)\boldsymbol{X}_p + \boldsymbol{t}_i$ を考える．そして，$\boldsymbol{\omega}_i$ の各要素の範囲を $[-20, 20]$(degree)，$t_1^{(i)}, t_2^{(i)}$ の範囲を $[-0.1, 0.1]$(meter)，$t_3^{(i)}$ の範囲を $[0.7, 1.5]$ (meter) としてランダムにパラメータを生成する．$\boldsymbol{\omega}_0, \boldsymbol{t}_0$ から平面パラメータ $\boldsymbol{m}$ が算出でき，$\boldsymbol{\omega}_n, \boldsymbol{t}_n, (n = 1, 2, \cdots, N)$ は $i = 0$ と $i = n$ の相対姿勢から算出できる．

一方で，合成画像の生成には，レンズ歪み変換 $\boldsymbol{w}_k$ の逆変換が不可欠である．ここでは，10 万分の 1 画素以下の精度が得られるように画素ごとに最適化[18] をすることで逆変換マップを生成した．また，合成画像生成の際にエイリアシングの発生を避けるため，4 倍サイズの画像を生成してから平滑化した後にリサンプリングした．さらに，すべての画像には標準偏差 2.0(graylevel) の正規ノイズを加えた．

このようにして，それぞれのテクスチャについて画像 100 枚を生成した．生成された合成画像のサンプルを図 4 に示す．生成した画像には手動で ROI を設定しており，画像中の破線によってこれらを示している．ただし，平面領域すべてが画像に入っているとみなせる場合には，画像全体を ROI とした．後述する実験では，100 枚の画像から 10 枚 $(N = 9)$ をランダムに抽出して実験に利用する．

4.1.1 Backward Warp Cost の効果

まず，本手法における backward warp コストの効果を検証した．図 5 は，backward warp

© 2010 Information Processing Society of Japan

コストの有無による収束の違いを示している．同図 (a) と (b) は，あるチェスボードパターン画像 10 枚を利用した場合において，パラメータ更新ベクトルのノルム $|\Delta \boldsymbol{p}|$ について，最適化の繰返し演算ごとの変化と，コスト値の変化の様子をそれぞれ示している．また，同図 (c) では，100 枚のチェスボードパターン画像からランダムに抽出した 10 枚の画像を 30 セット利用して，収束条件に達するまでの平均繰返し演算回数を評価したものである．

ここでは，本手法の真値周辺での最適化特性を確認することを目的としており，最適化の初期値は，$\boldsymbol{k} = \boldsymbol{0}$ とするのみで，他のパラメータは真値をそのまま入力している．

合成画像実験では，収束条件として $|\Delta \boldsymbol{p}| < 10^{-5}$ を利用しており，同図 (a) を見ると，forward warp コストのみを利用した場合と比べ，backward warp コストを追加的に利用することにより，早く収束条件に達していることがわかる．また，同図 (b) においても，backward warp コストを追加的に利用することにより，コスト値が早く小さくなっている．さらに，同図 (c) により，平均的な繰返し回数は，backward warp コストを利用することにより半分以下になっていることが示されている．

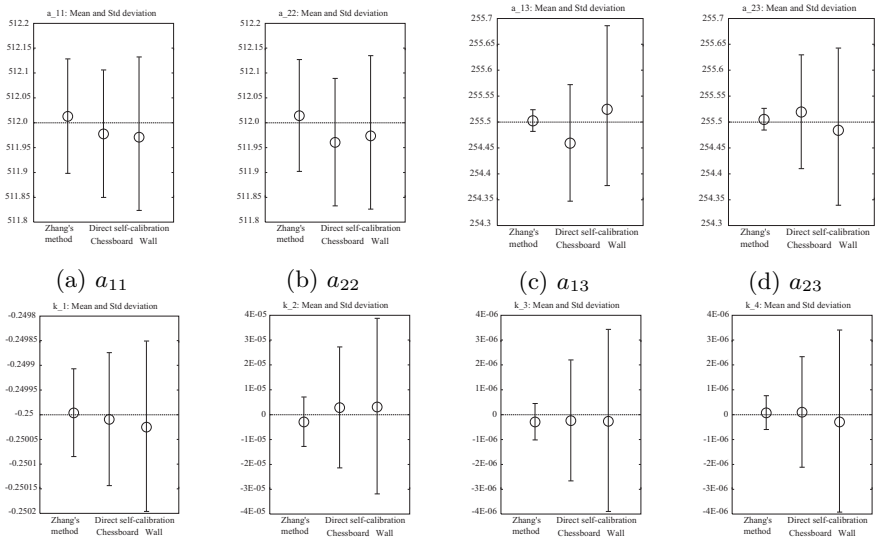
4.1.2 推定精度

次に，本手法の推定精度を検証した．ここでは，チェスボード画像 10 枚に対して，Matlab Toolbox を使って Zhang の手法を適用したケースと，チェスボード画像 10 枚，および，煉瓦テクスチャ画像 10 枚それぞれに本手法を適用したケースを比較した．いずれのケースでも，30 セットの画像について推定を行い，各パラメータについてその平均値と標準偏差を評価した．Zhang の手法と本手法をチェスボード画像に適用したケースでは，同じ 30 セットの画像を利用した．

図 6 は，$\boldsymbol{A}$ と $\boldsymbol{k}$ の各パラメータについて，その平均値と標準偏差を示している．各グラフにおいて，一番左が Zhang の手法，中央がチェスボード画像に本手法を適用した結果，一番右が煉瓦テクスチャ画像に本手法を適用した結果を示している．また，中央の破線は真値を示している．

いずれのパラメータにおいても，本手法の標準偏差は Zhang の手法よりも大きい．特に，画像中心 a_{13}, a_{23} のパラメータの値が相対的に悪い．これは，本手法にカメラ回転と並進に関する弱い曖昧性[16),19)] があり，これによって各カメラの光学中心の位置が精度よく定まっておらず，その影響が画像中心の精度に現れているものと推察される．しかし，セルフキャリブレーションである本手法が Zhang の手法に匹敵する精度を実現しているといえる．

本手法において，チェスボード画像を利用した場合と煉瓦テクスチャを用いたケースでは，相対的にチェスボードの方が高い精度が得られている．これは，一般的な画像よりも，



(a) a_{11} (b) a_{22} (c) a_{13} (d) a_{23}

(e) k_1 (f) k_2 (g) k_3 (h) k_4

図 6 Accuary of parameters estimated.

チェスボード画像の方がコントラストが高くテクスチャが明瞭であるため，レジストレーションの際にノイズの影響を受けにくいからと考えられる．

なお，本手法における計算時間は，10 枚の画像を利用した場合，Intel Core2Quad 2.8GHz の PC において，1 回の繰返し計算あたり約 15 秒であった．

4.2 実画像実験

Pointgray 社の Bumblebee2 (640×480 pix) を用いて実画像を撮影した[⋆1]．撮影対称はチェスボードとポスターの 2 種類であり，それぞれの種類について 30 枚を取得した．撮影した画像のサンプルを図 7 に示す．本実験では，これらの画像からランダムに 10 枚を抽出して推定に利用する．ここでは，平面領域が画像全体に入るように撮影しており，ROI は設定していない．また，実画像はノイズの影響を受けていると考え，本手法を適用する際には，5×5 の正規フィルタによって平滑化した．

⋆1 Bumblebee2 はステレオカメラであるが，右画像のみ利用している．

© 2010 Information Processing Society of Japan

Vol.2010-CVIM-173 No.4
2010/9/5

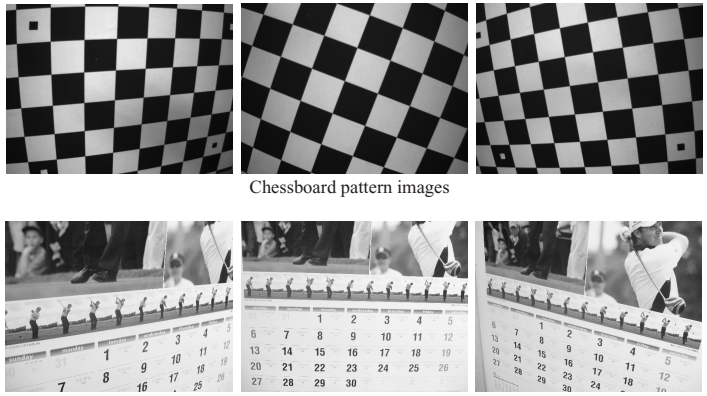
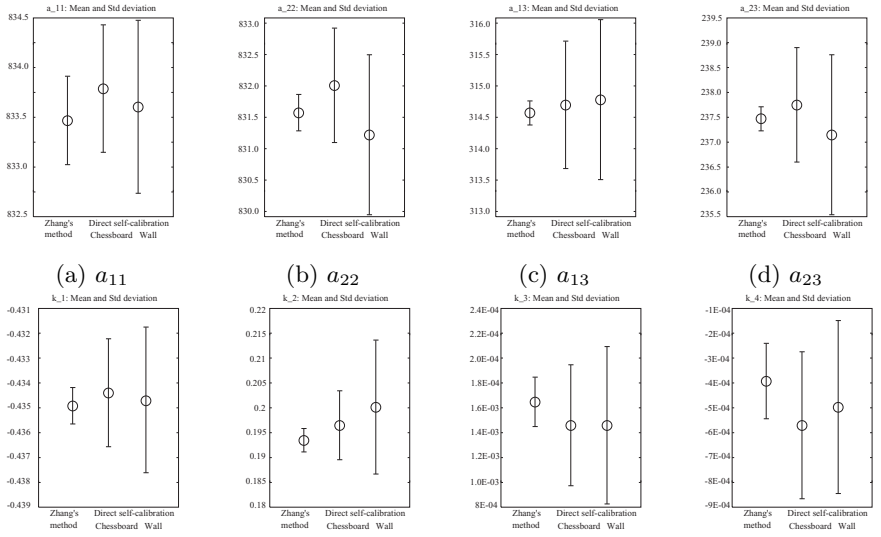


図 7 Examples of real images

実画像を用いた推定では，本手法の初期値に対する安定性を高めることと，適切な初期値を定める方法が必要となる．前者については，低い解像度の画像から推定を行う階層的推定手法[20] を利用する．本実験では 3 段の推定を行っており，各段階での収束条件は，画像サイズの小さいほうから $|\Delta \boldsymbol{p}| < 10^{-2}, |\Delta \boldsymbol{p}| < 10^{-3}, |\Delta \boldsymbol{p}| < 10^{-4}$ とした．一方，初期値については，レンズ歪みは大きくないと仮定し，次のような方法で初期値を定めた．

まず，画像を撮影することで得られるおおよその画角と画像サイズを用い，内部パラメータ $\boldsymbol{A}$ の対角成分と画像中心の初期値を定める．次に，基準画像と他の画像との間の特徴点対応から平面射影変換行列を 9 個求める．既存の直接法[21] を用いて精度を上げた後に，これらから内部パラメータの影響を取り除き，さらに，特異値分解を利用して各行列を $\boldsymbol{R}, \boldsymbol{t}, \boldsymbol{m}$ に分解する[22]．このとき得られた各 $\boldsymbol{R}$ はそのまま初期値として利用する．そして，各 $\boldsymbol{m}$ のノルムが一致するように $\boldsymbol{t}$ のスケールを合わせることで，$\boldsymbol{t}$ の初期値を得る．このようにすると，N 個の $\boldsymbol{m}$ が生成されるが，画像間のオーバラップが最も大きい画像間の $\boldsymbol{m}$ を選択することで，平面パラメータの初期値を得る．ポスター画像に対しては，SIFT[23] を適用して画像間の対応を求めた．繰返しパターンのチェスボード画像には SIFT を適用できないので，対応を手動入力とした．本実験ではこの初期値設定方法で問題は発生しなかったが，より精度の高い初期値を必要とする場合は，平面シーンから内部パラエータ行列を推定するための既存手法[24),25)] を利用する方法が利用できる．

実画像実験においても，合成画像のときの精度評価と同様に，Matlab Toolbox を使って



(a) a_{11} (b) a_{22} (c) a_{13} (d) a_{23}

(e) k_1 (f) k_2 (g) k_3 (h) k_4

図 8 Mean and standard deviation of parameters estimated.

Zhang の手法を適用したケースと，チェスボード画像，およびポスター画像それぞれに本手法を適用した．10 枚の画像 30 セットについて推定を行い，各パラメータについてその平均値と標準偏差を評価した．図 8 に結果を示す．

同図の各パラメータの標準偏差は，その大きさが合成画像のときよりもはるかに大きくなっているいるものの，基本的な傾向は合成画像のときと同じである．ここでは，本手法の収束条件を合成画像のときよりも緩和しているが，約 80 回の繰返し演算を行うことで収束条件に達しており，テクスチャの違いによる回数の相違はほとんどなかった．

5. ま と め

本稿では，直接法によるセルフキャリブレーション手法を提案した．ここでは，平面を撮影した複数の画像を直接法によってレジストレーションすることで，レンズ歪みを含むカメラパラメータを推定した．レンズ歪みの逆変換を利用しない座標変換の定式化を行い，forward warp コストと backward warp コストを両方利用して最適化を行った．実験結果

© 2010 Information Processing Society of Japan

では，セルフキャリブレーションであっても，校正物体を利用した場合に近い推定精度が得られた．

今後は，既存のセルフキャリブレーションとの比較を行う予定である．また，現時点では，本手法の最適化の収束が遅く，計算コストもかなり大きいため，これらについては改善方法を検討してゆく予定である．

参 考 文 献

1) Zhang, Z.: Flexible Camera Calibration by Viewing A Plane From Unknown Orientation, *IEEE International Conference on Computer Vision*, Vol.1, pp.666–673 (1999).
2) Bouguet, J.-Y.: Camera Calibration Toolbox for Matlab, http://www.vision.caltech.edu/bouguetj/.
3) Bradski, G. and Kaehler, A.: *Learning OpenCV*, Oreilly & Associates Inc (2008).
4) Hemayed, E.E.: A Survey of Camera Self-Calibration, *IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp.351–357 (2003).
5) Lourakis, M. I.A. and Argyros, A.A.: SBA: A Software Package for Generic Sparse Bundle Adjustmen, *ACM Transactions on Mathematical Software*, Vol.36, No.2, pp. 1–30 (2009).
6) Hartley, R. and Kang, S.B.: Parameter-Free Radial Distortion Correction with Center of Distortion Estimation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.29, No.8, pp.1309–1321 (2007).
7) Tardif, J.P. and Sturm, P.: Self-Calibration of a General Radially Symmetric Distortion Model, *European Conference on Computer Vision*, Vol.4 (2006).
8) Tardif, J.P., Sturm, P. and Roy, S.: Plane-based Self-calibration of Radial Distortion, *IEEE International Conference on Computer Vision*, pp.1–8 (2007).
9) Irani, M. and Anandan, P.: About direct methods, *International Workshop on Vision Algorithms : Theory and Practice*, Vol.1883, Springer-Verlag, pp.267–277 (1999).
10) Weng, J., Cohen, P. and Herniou, M.: Camera Calibration With Distortion Models and Accuracy Evaluation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.14, No.10, pp.965–980 (1992).
11) Brockett, R.W.: Robotic Manipulators and the Product of Exponentials Formula, *International Symposium on Mathematical Theory of Networks and Systems*, pp. 120–127 (1983).
12) Hartley, R. and Zisserman, A.: *Multiple View Geometry in Computer Vision (Second Edition)*, Cambridge University Press (2003).
13) Cachier, P. and Rey, D.: Symmetrization of the Non-Rigid Registration Problem using Invertion-Invariant Energies: Application to Multiple Sclerosis, *International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 472–481 (2000).
14) Rogelj, P. and Kovacic, S.: Symmetric Image Registration, *SPIE Medical Imaging 2003: Image Processing*, Vol.5023 (2003).
15) Christensen, G.E. and Johnson, H.J.: Consistent Image Registration, *IEEE Transactions on Medical Imaging*, Vol.20, No.7, pp.568–582 (2001).
16) Daniilidis, K. and Nagel, H.: The Coupling of Rotation and Translation in Motion Estimation of Planar Surfaces, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.188–193 (1993).
17) Press, W.H., Flannery, B.P., Teukolsky, S.A. and Vetterling, W.T.: Numerical Recipes in C (日本語版)，技術評論社, 第 9 刷 (1998).
18) DeVilliers, J.P., Leuschnerb, F.W. and Geldenhuysb, R.: Centi-pixel Accurate Real-time Inverse Distortion Correction, *SPIE International Symposium on Optomechatronic Technologies*, Vol.7266, pp.11–8 (2008).
19) Ke, Q. and Kanade, T.: Transforming Camera Geometry To A Virtual Downward-looking Camera: Robust Ego-motion Estimation and Ground-Layer Detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol.1, pp.390–397 (2003).
20) Bergen, J.R., Anandan, P., Hanna, K.J. and Hingorani, R.: Hierarchical Model-based Motion Estimation, *European Conference on Computer Vision*, pp.237–252 (1992).
21) Baker, S. and Matthews, I.: Lucas-Kanade 20 Years On: A Unifying Framework, *International Journal of Comoputer Vision*, Vol.56, No.3, pp.221–255 (2004).
22) Faugeras, O. and Lustman, F.: Motion And Structure From Motion in A Piecewise Planar Environment, *International Journal of Pattern Recognition and Artificial Intelligence*, Vol.2, No.3, pp.485–503 (1988).
23) Lowe, D.: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol.2, No.60, pp.91–110 (2004).
24) Malis, E. and Cipolla, R.: Multi-view Constraints between Collineations: Application to Self-Calibration from Unknown Planar Structures., *European Conference on Computer Vision*, pp.610–624 (2000).
25) Malis, E. and Cipolla, R.: Camera Self-Calibration from Unknown Planar Structures Enforcing the Multiview Constraints between Collineations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24, No.9, pp.1268–1272 (2002).

© 2010 Information Processing Society of Japan