全停留点の直接計算に基づく一般カメラモデルの PnP 問題に対する統一的解法

中野　　学[†a)]　　田治米純二[††]　　野村　俊之[†]

A Unified Solution to the PnP Problem for General Camera Model Based on Direct Computation of All the Stationary Points

Gaku NAKANO[†a)], Junji TAJIME[††], and Toshiyuki NOMURA[†]

あらまし　本論文は，全停留点の直接計算に基づく一般カメラモデルの PnP 問題に対する統一的解法を提案する．提案解法の特徴は，PnP 問題を 3 変数の無制約最適化問題として定式化し，グレブナー基底を用いて全ての停留点を計算することである．目的関数のこう配をゼロとした連立代数方程式は，$n \geq 3$ で互いに独立であり，方程式中の項は三次元座標の分布により変化しない．そのため，提案解法は，$n = 3$ の場合は P3P 問題の複数解，$n \geq 4$ の場合は大域的最適解が得られ，平面にも非平面にも適用可能な統一的解法である．また，提案解法の演算量は n 点の入力に対し $\mathcal{O}(n)$ である．$n = 100$ に対する平均実行時間は，内点法を用いる従来解法と比較して約 14 倍高速であることを実験により示す．

キーワード　PnP 問題，カメラ位置姿勢推定，無制約最適化，グレブナー基底

1. まえがき

PnP 問題（Perspective-n-Points Problem）とは，ワールド座標系における n 点の三次元座標とそれらの点が観測された画像座標から，校正済カメラの位置姿勢を推定する問題である．コンピュータビジョンにおける重要な問題の一つであり，ロボットナビゲーション，AR などで利用されている．PnP 問題は，最少 $n = 3$ で解けることが知られている [1]〜[3]．しかし，$n = 3$ では最大四つの解が存在し，一意に定められない．$n \geq 4$ であれば，カメラと三次元座標が特殊な配置である場合を除き，解が一意に定められるため，様々な解法が提案されている [4]〜[10]．Quan ら [4] は，P3P 解法を $n = 4$，$n = 5$ に拡張している．$n \geq 6$ にも適用可能だが，n 点に対する演算量は $\mathcal{O}(n^5)$ である．Ansar ら [5] は，三次元座標間の距離を表す二次多項式に基づく解法を提案している．Quan らよりも高精度だが，演算量は $\mathcal{O}(n^8)$ である．これらの解法の問題は，n 点に対する演算量が多いことである．そのため，Lepetit ら [6] は，カメラ座標系の基底ベクトルの基点座標を推定する $\mathcal{O}(n)$ の解法を提案している．しかし，係数行列の階数及び三次元座標の分布が平面か非平面かに応じて異なる計算方法が必要である．Lu ら [7] は，Object Space Error（OSE）と呼ぶ誤差関数を定義することで，位置が姿勢を変数とする関数であることと，PnP 問題が姿勢を変数とする制約付き最適化問題（以下，姿勢最適化問題と呼ぶ）であることを示した．そして，位置と姿勢を交互に最適化することで，大域的収束性をもつ反復解法を提案している．ただし，局所最適解に陥ることが多いという欠点がある．それに対し，Schweighofer ら [8] は全ての局所最適解を計算して大域的最適解を得るように Lu らの解法を改良したが，対象は平面に限定されている．更に Schweighofer ら [9] は，姿勢最適化問題の大域的最適解を得るため，二乗和多項式緩和を用いる解法を提案している．Lepetit らよりも高精度だが，反復解法である内点法の演算量が多い．また，Lepetit ら

† 日本電気株式会社情報・メディアプロセッシング研究所，川崎市
Information and Media Processing Labs, NEC, Shimonumabe 1753, Nakahara-ku, Kawasaki-shi, 211–8666 Japan
†† (株) NEC 情報システムズ先端技術ソリューション事業部，川崎市
Advanced Technology Solution Group, NEC Infomatec Systems, Ltd., 6–1 Kitamikata 2-chome, Takatsu-ku, Kawasaki-shi, 213–8511 Japan
a) E-mail: g-nakano@cq.jp.nec.com

と同様に平面と非平面を区別する必要がある．Hmam ら [10] は，三次元座標がカメラ前面に存在する拘束条件を Schweighofer らの解法に加えることで，平面と非平面を区別しない解法を提案している．半正定値緩和を用いることで Schweighofer らよりも約 2 倍高速だが，内点法による演算量の問題は解決していない．このように，従来の P4P 解法は，低演算量のためには条件に応じて異なる計算方法を必要とし，高精度な推定のためには非線形最適化を解く反復解法を必要とする．また，複数解をもつ $n = 3$ と大域的最適解をもつ $n \geq 4$ とを統一的に扱えない．

本論文では，姿勢最適化問題を 3 変数の無制約最適化問題として定式化し，全停留点の直接計算に基づく統一的解法を提案する [11]．目的関数のこう配である連立代数方程式は，$n \geq 3$ で互いに独立であり，方程式中の項は三次元座標の分布により変化しない．そのため，提案解法は，$n = 3$ の場合は P3P 問題の複数解，$n \geq 4$ の場合は大域的最適解が得られ，平面にも非平面にも適用可能な統一的解法である．また，グレブナー基底を利用することで，行列分解により連立代数方程式の解を得る．それゆえ，反復解法を用いる従来解法よりも低演算量である．なお，文献 [11] の実験条件は文献 [10] と異なるため，本論文では文献 [10] と実験条件をそろえて従来解法との比較を行い，更に独自の評価実験を加えることで外れ値算出の原因について考察を行う．

2. 一般カメラモデルに対する PnP 問題

2.1 一般カメラモデル

一般カメラモデルとは，ステレオカメラや多焦点カメラを扱うための，透視投影モデルの拡張である [9], [12]．一般カメラモデルの例を図 1 に示す．一般カメラモデルにおける三次元座標の射影は式 (1) で表される．

$$\lambda_i \mathbf{v}_i = \mathbf{R}\mathbf{X}_i + \mathbf{t} - \mathbf{c}_i \tag{1}$$

ここで，$\mathbf{R}$ はカメラ姿勢を表す 3×3 の回転行列，$\mathbf{t}$ はカメラ位置を表す 3×1 のベクトル，$\mathbf{X}_i$ は i 番目（$1 \leq i \leq n$）の三次元座標を表す 3×1 のベクトル，$\mathbf{v}_i$ は第 3 成分を 1 に斉次化した画像座標を表す 3×1 のベクトル，$\mathbf{c}_i$ はカメラ座標系における $\mathbf{v}_i$ の投影中心を表す 3×1 のベクトル，λ_i は $\mathbf{X}_i$ への奥行を表すスカラである．PnP 問題においては，$\mathbf{X}_i$，$\mathbf{v}_i$，$\mathbf{c}_i$ が既知で，$\mathbf{R}$，$\mathbf{t}$，λ_i が未知である．式 (1) において，$\mathbf{c}_i = \mathbf{0}$ とすると，透視投影モデルとなる．

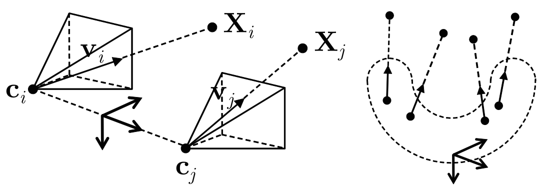

図 1 一般カメラモデルの例
Fig. 1 Examples of General Camera Model.

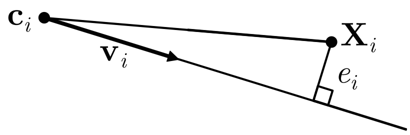

図 2 Object Space Error
Fig. 2 Object Space Error.

2.2 Object Space Error

Object Space Error（OSE）は，文献 [7] で透視投影モデルに対する PnP 問題の誤差関数として定義され，文献 [9] で一般カメラモデルに拡張された．文献 [9] では，一般カメラモデルへの拡張に関する記述が省略されているため，本節では改めてその導出を行い，OSE による PnP 問題の定式化を説明する．

$\mathbf{X}_i$ と $\mathbf{v}_i$ が与えられたとき，式 (1) の二乗誤差は式 (2) で表される．

$$\|\lambda_i \mathbf{v}_i - (\mathbf{R}\mathbf{X}_i + \mathbf{t} - \mathbf{c}_i)\|^2 \tag{2}$$

式 (2) を λ_i で微分しゼロとおき，式 (2) を最小化する λ_i を求める．それを式 (2) に代入して λ_i を消去すると，式 (3) を得る．

$$e_i = \|\mathbf{V}_i(\mathbf{R}\mathbf{X}_i + \mathbf{t} - \mathbf{c}_i)\|^2 \tag{3}$$

ここで，$\mathbf{V}_i = \mathbf{I} - \dfrac{\mathbf{v}_i \mathbf{v}_i^T}{\mathbf{v}_i^T \mathbf{v}_i}$ である．

式 (3) が一般カメラモデルに対する OSE である．図 2 に示すように，OSE は，カメラ座標系において $\mathbf{X}_i$ を $\mathbf{v}_i$ へ射影したベクトルの L2 ノルムである．$\mathbf{R}$ に関する拘束条件を式 (3) に加えると，OSE による PnP 問題は式 (4) で表される．

$$\begin{aligned} &\min_{\mathbf{R},\,\mathbf{t}} \quad e = \sum_{i=1}^{n} \|\mathbf{V}_i(\mathbf{R}\mathbf{X}_i + \mathbf{t} - \mathbf{c}_i)\|^2 \\ &\text{s.t.} \quad \det(\mathbf{R}) = 1,\, \mathbf{R}\mathbf{R}^T = \mathbf{I} \end{aligned} \tag{4}$$

2.3 姿勢最適化問題の定式化

本節では，PnP 問題を姿勢最適化問題として定式化

する [9]．まず，式 (4) を最小化する $\mathbf{t}_{opt}$ を求めるため，式 (4) を $\mathbf{t}$ で微分すると式 (5) を得る．

$$\frac{\partial e}{\partial \mathbf{t}} = 2\sum_{i=1}^{n} \mathbf{V}_i(\mathbf{R}\mathbf{X}_i + \mathbf{t} - \mathbf{c}_i) = \mathbf{0}_{3\times 1} \tag{5}$$

式 (5) より，$\mathbf{t}_{opt}$ は式 (6) で表される．

$$\begin{aligned}\mathbf{t}_{opt} &= -\left(\sum_{i=1}^{n}\mathbf{V}_i\right)^{-1}\sum_{i=1}^{n}\mathbf{V}_i(\mathbf{R}\mathbf{X}_i - \mathbf{c}_i)\\ &= -\left(\sum_{i=1}^{n}\mathbf{V}_i\right)^{-1}\sum_{i=1}^{n}\mathbf{V}_i[\mathbf{A}_i| - \mathbf{c}_i]\begin{bmatrix}\mathbf{r}\\ 1\end{bmatrix}\\ &= \mathbf{B}\begin{bmatrix}\mathbf{r}\\ 1\end{bmatrix}\end{aligned} \tag{6}$$

ここで，$\mathbf{A}_i = \begin{bmatrix}\mathbf{X}_i^T & \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3}\\ \mathbf{0}_{1\times 3} & \mathbf{X}_i^T & \mathbf{0}_{1\times 3}\\ \mathbf{0}_{1\times 3} & \mathbf{0}_{1\times 3} & \mathbf{X}_i^T\end{bmatrix}$，$\mathbf{R}$ の第 j 行を $\mathbf{r}_j$ として $\mathbf{r} = [\mathbf{r}_1^T,\ \mathbf{r}_2^T,\ \mathbf{r}_3^T]^T$ である．また，$\mathbf{B} = -\left(\sum_{i=1}^{n}\mathbf{V}_i\right)^{-1}\sum_{i=1}^{n}\mathbf{V}_i[\mathbf{A}_i| - \mathbf{c}_i]$ である．

次に，$\mathbf{t}_{opt}$ を式 (4) に代入すると，式 (7) を得る．

$$\begin{aligned}&\min_{\mathbf{R}} \quad \mathbf{r}^T\mathbf{M}_1\mathbf{r} + 2\mathbf{m}_2^T\mathbf{r} + m_3\\ &\text{s.t.} \quad \det(\mathbf{R}) = 1, \mathbf{R}\mathbf{R}^T = \mathbf{I}\end{aligned} \tag{7}$$

ここで，

$$\begin{aligned}&\begin{bmatrix}\mathbf{M}_1 & \mathbf{m}_2\\ \mathbf{m}_2^T & m_3\end{bmatrix}\\ &= \sum_{i=1}^{n}([\mathbf{A}_i| - \mathbf{c}_i]+\mathbf{B})^T\mathbf{V}_i([\mathbf{A}_i| - \mathbf{c}_i]+\mathbf{B})\end{aligned} \tag{8}$$

である．

以上で，式 (7) に示すように，PnP 問題が姿勢最適化問題として定式化された．

3. 提 案 解 法

提案解法は，まず，正規化されていない四元数を用いて姿勢最適化問題を無制約化する．次に，無制約姿勢最適化問題のこう配をゼロとした連立代数方程式を解いて全ての停留点を得る．連立代数方程式の解法にはグレブナー基底を用いる．最後に，停留点から大域的最適解を選択し，$\mathbf{R}$ と $\mathbf{t}$ を得る．

3.1 姿勢最適化問題の無制約化

正規化されていない四元数 $\mathbf{q} = [1,\ a,\ b,\ c]^T$ を用いると，$\mathbf{R}$ は式 (9) で表される．

$$\mathbf{R} = \frac{1}{\|\mathbf{q}\|^2}\hat{\mathbf{R}} \tag{9}$$

ここで，

$$\hat{\mathbf{R}} = \begin{bmatrix}1+a^2-b^2-c^2 & 2(ab-c) & 2(ac+b)\\ 2(ab+c) & 1-a^2+b^2-c^2 & 2(bc-a)\\ 2(ac-b) & 2(bc+a) & 1-a^2-b^2+c^2\end{bmatrix} \tag{10}$$

である．

次に，式 (9) を式 (1) に代入し，両辺を $\|\mathbf{q}\|^2$ 倍すると式 (11) を得る．

$$\|\mathbf{q}\|^2\lambda_i\mathbf{v}_i = \hat{\mathbf{R}}\mathbf{X}_i + \|\mathbf{q}\|^2(\mathbf{t} - \mathbf{c}_i) \tag{11}$$

式 (11) を用いて，**2. 3** と同様に最適化問題を導出すると，式 (12) を得る．

$$\min_{a,b,c} \quad \hat{\mathbf{r}}^T\mathbf{M}_1\hat{\mathbf{r}} + 2\|\mathbf{q}\|^2\mathbf{m}_2^T\hat{\mathbf{r}} + \|\mathbf{q}\|^4 m_3 \tag{12}$$

ここで，$\hat{\mathbf{R}}$ の第 j 行を $\hat{\mathbf{r}}_j$ として $\hat{\mathbf{r}} = [\hat{\mathbf{r}}_1^T,\ \hat{\mathbf{r}}_2^T,\ \hat{\mathbf{r}}_3^T]^T$ である．

以上より，姿勢最適化問題は，式 (12) の無制約非線形最適化問題に帰着する．

無制約非線形最適化問題を解くには，停留点を求め，停留点が大域的最適解かどうかを調べればよい．停留点は，目的関数のこう配をゼロとした連立代数方程式の実数解である．変数ベクトルを $\mathbf{x} = [a, b, c]^T$，3×20 の係数行列を $\mathbf{N}$，20×1 の項ベクトルを $\mathbf{z} = [a^3, a^2b, ab^2, b^3, a^2c, abc, b^2c, ac^2, bc^2, c^3, a^2, ab, b^2, ac, bc, c^2, a, b, c, 1]^T$ とすると，解くべき連立代数方程式は，式 (12) を $\mathbf{x}$ で微分してゼロとした式 (13) で表される．

$$\begin{aligned}\mathbf{N}\mathbf{z} &= \frac{d\hat{\mathbf{r}}}{d\mathbf{x}}^T\mathbf{M}_1\hat{\mathbf{r}} + 2\mathbf{m}_2^T\hat{\mathbf{r}}\mathbf{x} + \|\mathbf{q}\|^2\frac{d\hat{\mathbf{r}}}{d\mathbf{x}}^T\mathbf{m}_2\\ &\quad + 2\|\mathbf{q}\|^2 m_3\mathbf{x}\\ &= \mathbf{0}\end{aligned} \tag{13}$$

ここで，$\dfrac{d\hat{\mathbf{r}}}{d\mathbf{x}} = \left[\dfrac{\partial\hat{\mathbf{r}}}{\partial a},\ \dfrac{\partial\hat{\mathbf{r}}}{\partial b},\ \dfrac{\partial\hat{\mathbf{r}}}{\partial c}\right]$ である．

式 (13) は 3 個の三元三次式であり，次のような性質をもっている．まず，$\mathbf{z}$ の各項は，三次元座標の分布やカメラモデルに依存しない．つまり，平面でも非平面でも $\mathbf{z}$ は変化しないし，$\mathbf{c}_i = \mathbf{0}$ の透視投影モデルでも $\mathbf{c}_i \neq \mathbf{0}$ の一般カメラモデルでも $\mathbf{z}$ は変化しない．変化するのは $\mathbf{N}$ のみである．次に，$n \geq 3$ の場合，式

(8) より $\mathrm{rank}(\mathbf{M}_1) \geq 3$ となり，式 (13) は互いに独立である．以上より，$n \geq 3$ であれば，三次元座標の分布とカメラモデルに関係なく式 (13) は解ける．

3.2 グレブナー基底による停留点計算

式 (13) の解の個数は，ベズーの定理より $3^3 = 27$ 個である．このような複雑な多変数連立代数方程式の解法として，u 終結式法，隠密変数法，固有値分解法などが知られている [13]．u 終結式法は，$u_0 + u_1 a + u_2 b + u_3 c = 0$（$u_i$ は独立変数）なる新たな方程式を加え，u_i についての終結式を因数分解して解を求める手法である．巨大な行列式を数式処理で解く必要がある上に，浮動小数点による因数分解は非常に不安定なため，実用的ではない．隠密変数法は，任意の 1 変数を定数とみなして終結式を計算する手法である．例えば，式 (13) において a を定数とすると，b と c を変数とする 3 個の二元三次式となり，u 終結式法よりも行列式のサイズを小さくできる．また，因数分解は不要である．しかし，変数ごとに解が得られ，それらの組合せについては何も情報は与えない．1 変数の解のみが必要なときは有効な手法だが，本論文のように全ての解の組合せが必要な問題には不適である．固有値分解法は，任意の項に対する倍写像行列を計算し，その固有値と固有ベクトルを解とする手法である．全ての解の組合せが得られる上に，広く研究されている固有値分解に帰着するため数値計算と相性がよく，本論文に適している．

倍写像行列の計算には終結式に基づく手法とグレブナー基底に基づく手法があり，得られる倍写像行列は同一である [13]．本論文では，Kukelova ら [14] によるグレブナー基底に基づく手法を利用する．グレブナー基底とは，多変数の連立代数方程式を，等価でかつ解きやすい形に変換した多項式である [15]．「等価」とは，グレブナー基底の解が元の連立代数方程式の解と一致することである．Kukelova らの手法は，グレブナー基底に現れる項を事前に求め，係数行列 $\mathbf{N}$ から倍写像行列を計算する．試行ごとのグレブナー基底計算を省略できるため，高速に解が得られる．Kukelova らのウェブサイト(注1)で MATLAB コードが公開されており，同サイトでは Kukelova らの手法を様々なコンピュータビジョンの問題に適用した文献が数多く紹介されている．このように，高速性と効果が実証されていることが，Kukelova らの手法を選択する理由である．

Kukelova らの手法を式 (13) に適用すると，89×116 行列のガウス・ジョルダン消去法と 27×27 倍写像行列の固有値分解に帰着する．**4.** の実験では，項を適切に並べ換えることでガウス・ジョルダン消去法を LU 分解に置き換える手法 [16] を用いて，更に計算量を削減している．

3.3 大域的最適解の選択

以下（1）～（4）の処理を行い，**3.2** で得られた 27 個の解から $n = 3$ の複数解若しくは $n \geq 4$ の大域的最適解を選択する．

（1） 虚数解はあり得ないため，実数解を抽出．
（2） 局所最適性を満たすため，式 (12) のヘッセ行列が正定値である解を抽出．
（3） 式 (13) は，カメラが三次元座標の方向を向くことを保証しないため，カメラ前面に三次元座標が存在する解を抽出．
（4） （1）～（3）を満たす解の中で，
  i. $n = 3$ の場合，全ての解を選択．
  ii. $n \geq 4$ の場合，式 (12) を最小化する大域的最適解を選択．

選択された解 $[a, b, c]$ を式 (9) に代入して $\mathbf{R}$ が，$\mathbf{R}$ を式 (6) に代入して $\mathbf{t}$ が得られる．

4. 実　　験

シミュレーション実験により，提案解法を評価した．まず，透視投影モデルにおいて P3P 問題の解が得られるか調べた．次に，一般カメラモデルにおける位置姿勢推定誤差と演算量を評価した．比較する各解法について，以下の略称を用いる．

UPnP	提案解法（Unified PnP solution）
SDR+CG	Hmam らによる内点法の解を初期値として共役こう配法を適用する解法 [10]
EPnP+GN	Lepetit らによるガウスニュートン最適化付き EPnP 法 [6]
LHM	Lu らによる反復解法 [7]
MLHM	Schweighofer らによる平面を対象とする改良 LHM 法 [8]
FB	Fischer らによる P3P 解法 [2]

プログラムは，全て MATLAB R2011a で実装及び実行した．SDR+CG は，Hmam より提供して頂いた

(注1)：http://cmp.felk.cvut.cz/minimal/automatic_generator.php

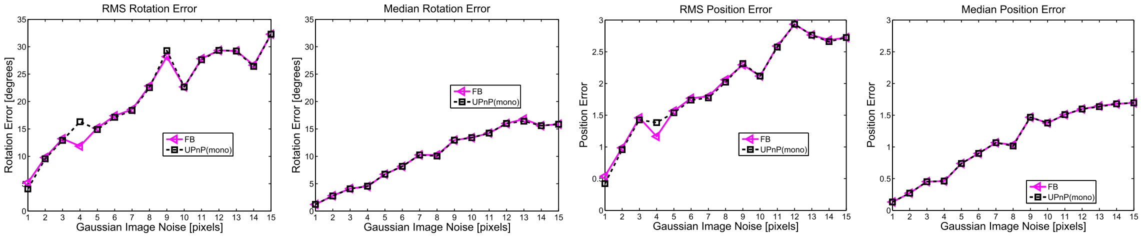

図 3　$n=3$，透視投影モデルにおける位置姿勢推定の RMS 誤差と誤差の中央値
Fig. 3　RMS and Median estimation errors for 3 uniformly distributed points under perspective camera model.

透視投影モデルに対応するコードを，筆者らが一般カメラモデルに拡張した．内点法のソルバーは SeDuMi 1.3(注2) を用いた．従来解法の中では SDR+CG が最も高精度であるため，その他の手法は一般カメラモデルに拡張していない．また，FB は筆者らが実装した．実験に使用した PC の構成は，Windows XP SP3, Core2 Duo E8500（3.16 GHz），メモリ 2 GByte である．

実験条件や誤差評価方法は基本的に文献 [10] に従い，不記載のパラメータは独自に設定した．カメラの内部パラメータ $\mathbf{K}$ は，光学中心 [320, 240]，焦点距離 800 とした．画像サイズは 640×480 である．n 点の三次元座標群 $\mathbf{X}_i$ を生成する空間の大きさは各実験の節で説明する．$\mathbf{X}_i$ をカメラに射影した座標 $\mathbf{u}_i$ に標準偏差 σ のガウスノイズを加え，$\mathbf{v}_i=\mathbf{K}^{-1}(\mathbf{u}_i+N(0,\sigma^2))$ とした．また，一般カメラモデルとして，透視投影モデルのカメラを間隔 d 離して横に二つ並べたステレオカメラを設定した．ここで，左カメラの投影中心をカメラ座標系の原点とした．これは，式 (1) において，$\mathbf{X}_i$ が左カメラで観測されたときは $\mathbf{c}_i=\mathbf{0}$，右カメラで観測されたときは $\mathbf{c}_i=[-d,0,0]^T$ であることを意味する．各実験においては $d=0.5$ とした．実験結果の凡例において，各解法の後に続く stereo は一般カメラモデルを表し，mono または不記載は透視投影モデルを表す．

真値 $\mathbf{R}_{true}$，$\mathbf{t}_{true}$ に対する推定値 $\mathbf{R}_{est}$，$\mathbf{t}_{est}$ の誤差を式 (14) で評価する．

$$\begin{aligned} E_t &= \|\mathbf{p}_{true}-\mathbf{p}_{est}\| \\ &= \|\mathbf{R}_{est}^T\mathbf{t}_{est}-\mathbf{R}_{true}^T\mathbf{t}_{true}\| \\ E_R &= 2\cos^{-1}(q_0) \end{aligned} \tag{14}$$

ここで，$q_0=0.5\sqrt{1+\mathrm{trace}(\mathbf{R}_{est}\mathbf{R}_{true}^T)}$ である．また，実験結果を示す際，E_R はラジアンから度数に変換している．評価結果は，RMS（Root Mean Square）誤差と誤差の中央値をプロットする．中央値を示すのは，外れ値（局所最適解や全く値が異なる解）を算出するか確認するためである．つまり，RMS 誤差と中央値の差が少なければ少ないほど，外れ値を算出することが少ない安定した解法といえる．

4.1　P3P 問題に対する推定誤差

透視投影モデル，$n=3$ として UPnP を評価し，P3P 問題が解けるか実験した．複数解が得られた場合，真値との誤差が最も小さい解を $\mathbf{R}_{est}$，$\mathbf{t}_{est}$ とした．点群は，原点を中心とする $2\times2\times2$ の立方体内に，一様分布に従いランダムに 3 点生成した．カメラは原点から 6 離れた位置にランダムに置いた．また，ノイズの標準偏差は $1\leq\sigma\leq15$，試行回数は各 200 回である．実験結果を図 3 に示す．横軸は σ，縦軸は誤差を表す．P3P 問題は最少点数の PnP 問題のため，FB 以上の高精度化はできない．図 3 より，FB と UPnP の誤差分布はほとんど同じである．これにより，UPnP により従来解法と同等精度で P3P 問題が解けることが示された．

4.2　非平面におけるノイズに対する推定誤差

非平面における UPnP の位置姿勢推定誤差を評価した．点群は，原点を中心とする $4\times4\times4$ の立方体内に，一様分布に従いランダムに 6 点生成した．カメラは原点から 10 離れた位置にランダムに置いた．ノイズの標準偏差，試行回数は **4.1** と同様である．実験結果を図 4 に示す．UPnP は，EPnP よりも高精度であるが SDR+CG には劣っている．特に，高ノイズレベルにおいて外れ値の影響が顕著である．外れ値の原因については **5.** で考察する．従来解法の中で，LHM はほとんど局所最適解に収束してしまい，文献 [10] の結果よりも悪い．しかし，LHM は大域的最適解への収束は何ら保証していないため，不自然な結果ではない．

(注2)：SeDuMi 1.3: http://sedumi.ie.lehigh.edu/

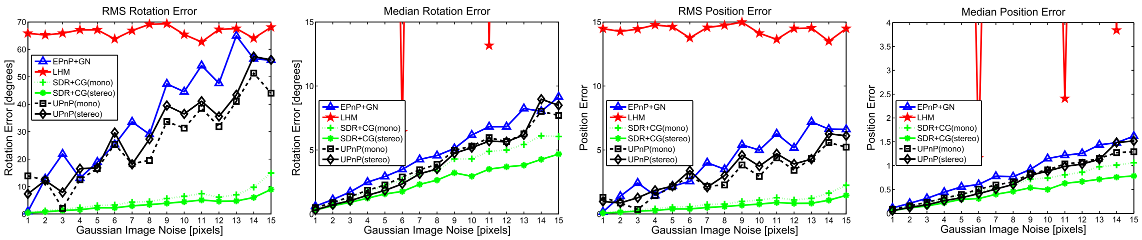

図 4　$n=6$，非平面におけるノイズに対する位置姿勢推定の RMS 誤差と誤差の中央値
Fig. 4　RMS and Median estimation errors for 6 uniformly distributed non-planar points.

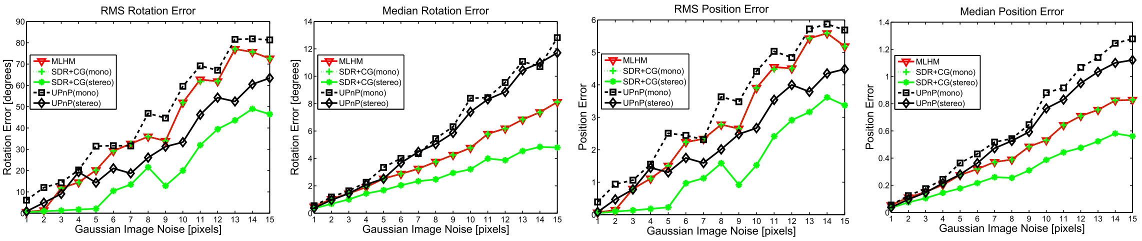

図 5　$n=10$，平面におけるノイズに対する位置姿勢推定の RMS 誤差と誤差の中央値
Fig. 5　RMS and Median estimation errors for 10 uniformly distributed coplanar points.

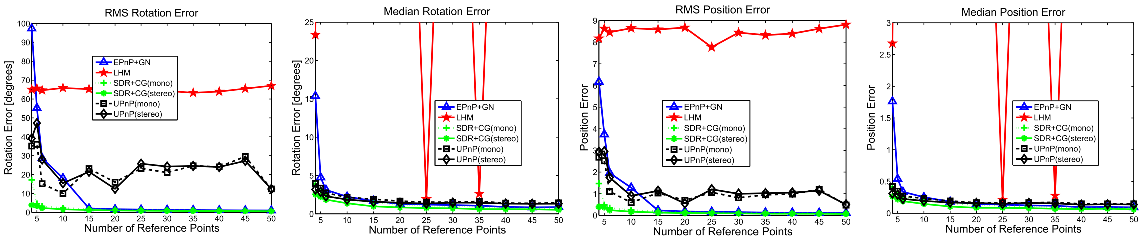

図 6　非平面における点数の変化に対する推定誤差
Fig. 6　RMS and Median estimation errors for each number of non-planar points.

4.3　平面におけるノイズに対する推定誤差

平面における UPnP の位置姿勢推定誤差を評価した．点群は，原点を中心とする 2×2 の正方形内に，一様分布に従いランダムに 10 点生成した．カメラは原点から 6 離れた位置にランダムに置いた．ノイズの標準偏差，試行回数は **4.1** と同様である．文献 [10] では，カメラに対する平面の傾きは 45 度と固定している．しかし，このような固定条件は実応用において一般的でなく，様々な角度に対するロバスト性が検証できない．そのため，本実験では傾きは固定しない．実験結果を図 5 に示す．UPnP は，非平面だけでなく平面にも適用できることを示している．しかし **4.2** と同様に，UPnP は従来解法に比べ，高ノイズレベルにおいてやや低精度である．これも外れ値の影響と考えられる．

4.4　非平面における点数の変化に対する推定誤差

非平面における点数の変化に対する位置姿勢推定誤差を評価した．点群は，原点を中心とする $2\times2\times2$ の立方体内に，一様分布に従いランダムに 4 点から 50 点生成した．カメラは原点から 6 離れた位置にランダムに置いた．ノイズの標準偏差は $\sigma=5$ に固定し，試行回数は **4.1** と同様である．実験結果を図 6 に示す．横軸は n，縦軸は誤差を表す．LHM を除く従来解法は点数の増加とともに精度が増す一方で，UPnP は精度向上の度合が低い．これについての考察は **5.** で行う．

4.5　演　算　量

非平面，ノイズの標準偏差 $\sigma=2$ に対して演算量を評価した．点数は $4\le n\le1000$，試行回数は各 100 回である．平均実行時間を図 7 に示す．図 7 において，左図は $4\le n\le1000$ の全範囲を示し，右図は $4\le n\le30$ の拡大図である．横軸は n，縦軸は実行

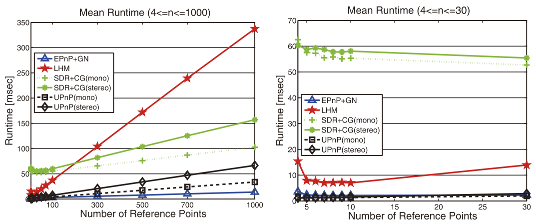

図 7　平均実行時間の比較
Fig. 7　Mean runtime comparisons.

時間（単位：ミリ秒）を表す．

UPnP は非常に高速であり，$4 \leq n \leq 30$ の範囲内では従来解法の中で最速な EPnP+GN に匹敵，若しくは上回る．SDR+CG（mono）と比べると，$n = 100$ における平均実行時間は，UPnP（mono）が 8.36 ミリ秒，SDR+CG（mono）が 110.1 ミリ秒であり，約 14 倍高速である．最も高速なのは，$n = 4$ における約 100 倍である．従来解法はいずれも反復解法であるため，$n = 4$ と $n = 5$ では一時的に局所最適解に近づき，演算量が増加する [9]．しかし，UPnP は全ての局所最適解から大域的最適解を選択するため，そのようなことは起きない．UPnP の演算量は，完全に $\mathcal{O}(n)$ である．また，一般カメラモデルを用いた場合，演算量の増加は SDR+CG よりも少ない．

5. 考　　察

本章では，提案解法が外れ値を算出する原因について考察する．外れ値の原因として，(1) 数値計算誤差，(2) 真値のノルム $\|\mathbf{q}\|$ が大きいため式 (13) が解けない，(3) 式 (13) がノイズの影響を受ける，の 3 種類が考えられる．まず，原因 (1) を検証する．SDR+CG では真値近くに収束したが，UPnP では外れ値となった入力に対し，倍々精度浮動小数点にて再度実験を行った．その結果，倍々精度浮動小数点でも UPnP は同じ外れ値を算出した．そのため，数値計算誤差が原因である可能性は低い．次に，原因 (2) を検証する．UPnP は，四元数の一つを定数にしているため，無限大と無限小を用いないと回転行列を完全に表現できない．つまり，真値のノルム $\|\mathbf{q}\|$ が大きい場合は数値計算が不安定になる可能性がある．しかし，原因 (1) の考察より，ノルムの大きさが原因である可能性は低い．最後に，原因 (3) について検証する．非平面の点数 $4 \leq n \leq 50$，ノイズの標準偏差 $1 \leq \sigma \leq 15$ の間で変化させ，各 200 回試行したうちの外れ値の割合を図 8 に示す．経験的に $E_R \geq 20$ [degree] を外れ値と

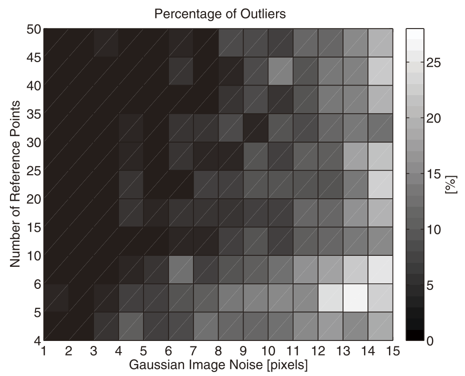

図 8　外れ値の割合
Fig. 8　Percentage of outlier.

した．外れ値の出現確率は，点数ではなくノイズレベルに比例している．つまり，ノイズの影響により，式 (13) からは正しい解が得られないことが外れ値の主な原因と考えられる．そのため，**3. 2** で述べたグレブナー基底以外の手法，例えば終結式に基づく倍写像行列計算を用いても外れ値算出は解決できない．これを解決する方法は，SDR+CG と同様に，正規化四元数を用いた最適化問題を解くことである．正規化四元数を用いる場合，拘束条件に対するラグランジュ乗数が加わるため，式 (13) よりも更に複雑な連立代数方程式を解く必要がある．反復解法を用いずにその連立代数方程式が解ければ，SDR+CG と同等精度を高速に実現できる可能性がある．

6. む　す　び

本論文は，一般カメラモデルの PnP 問題に対する統一的解法を提案した．提案解法は，無制約化した姿勢最適化問題の全停留点を直接計算し，大域的最適解を選択する．停留点計算は，目的関数のこう配をゼロとした連立代数方程式を解いて得られる．連立代数方程式は $n \geq 3$ で互いに独立であるため，平面と非平面を区別する必要がない．連立代数方程式の解法としてグレブナー基底を用いることで，高速な演算を実現した．実験により，$n = 3$ の場合は P3P 問題の複数解，$n \geq 4$ の場合は大域的最適解が得られ，平面にも非平面にも適用可能なことを示した．また，$n = 100$ に対する平均実行時間は，内点法を用いる従来解法と比較して約 14 倍高速であることを示した．一方で，提案解法は高ノイズレベルに対する脆弱性があり，その原因が正規化しない四元数表現にあることを実験的に示

した．今後の課題は，反復解法を用いずに正規化四元数による最適化問題を解くことである．

文　　献

[1] J.A. Grunert, "Das pothenotische problem in erweiterter gestalt nebst bber seine anwendungen in der geodasie," Grunerts Archiv fur Mathematik und Physik, pp.238–248, 1841.

[2] 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, June 1981.

[3] R.M. Haralick, C.-N. Lee, K. Ottenberg, and M. Nolle, "Review and analysis of solutions of the three point perspective pose estimation problem," Int. J. Comput. Vis., vol.13, no.3, pp.331–356, Dec. 1994.

[4] L. Quan and Z. Lan, "Linear n-point camera pose determination," IEEE Trans. Pattern Anal. Mach. Intell., vol.21, no.8, pp.774–780, Aug. 1999.

[5] A. Ansar and K. Daniilidis, "Linear pose estimation from points or lines," IEEE Trans. Pattern Anal. Mach. Intell., vol.25, no.5, pp.578–589, May 2003.

[6] V. Lepetit, F. Moreno-noguer, and P. Fua, "EPnP: An accurate $\mathcal{O}(n)$ solution to the PnP problem," Int. J. Comput. Vis., vol.81, no.2, pp.155–166, Feb. 2009.

[7] C.-P. Lu, G.D. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," IEEE Trans. Pattern Anal. Mach. Intell., vol.22, no.6, pp.610–622, June 2000.

[8] G. Schweighofer and A. Pinz, "Robust pose estimation from a planar target," IEEE Trans. Pattern Anal. Mach. Intell., vol.28, no.12, pp.2024–2030, 2006.

[9] G. Schweighofer and A. Pinz, "Globally optimal $\mathcal{O}(n)$ solution to the PnP problem for general camera models," Proc. BMVC '08, Leeds, UK, Sept. 2008.

[10] H. Hmam and J. Kim, "Optimal non-iterative pose estimation via convex relaxation," Int. J. Image Vis. Comput., vol.28, no.11, pp.1515–1523, Nov. 2010.

[11] 中野　学，田治米純一，野村俊之，"一般カメラモデルのPnP問題に対するグレブナー基底を用いた統一的解法，" 画像の認識･理解シンポジウム（MIRU2011），pp.845–851, 2011.

[12] M.D. Grossberg and S.K. Nayar, "A general imaging model and a method for finding its parameters," Proc. ICCV '01, pp.108–115, 2001.

[13] D. Cox, J. Little, and D. O'Shea, Using Algebraic Geometry, Springer-Verlag, New York, 1998. 大杉英史，北村知徳，日比孝之（共訳），グレブナー基底 1，シュプリンガー・フェアラーク東京，2000.

[14] Z. Kukelova, M. Bujnak, and T. Pajdla, "Automatic Generator of Minimal Problem Solvers," Proc. ECCV '08, pp.302–315, Marseille, France, Oct. 2008.

[15] 野呂正之，横山和弘，グレブナー基底の計算　基礎編，東京大学出版会，東京，2008.

[16] M. Byröd, K. Josephson, and K. Åström, "Fast and stable polynomial equation solving and its application to computer vision," Int. J. Comput. Vis., vol.34, no.3, pp.237–255, 2009.

（平成 23 年 10 月 7 日受付，24 年 2 月 20 日再受付）

中野　　学

2006 筑波大・第三学群・工学システム卒．2008 同大大学院システム情報工学研究科博士前期課程了．2008 日本電気（株）入社．以来，コンピュータビジョン，画像処理に関する研究開発に従事．情報処理学会会員．

田治米純二　（正員）

1995 早大・理工・情報学卒．1997 同大大学院修士課程了．1997 日本電気（株）入社．現在は（株）NEC 情報システムズにて，ビデオ符号化及び映像処理に関する研究開発に従事．

野村　俊之　（正員）

1990 名大・工・電気卒．1992 同大大学院修士課程了．1992 日本電気（株）入社．以来，音声符号化，オーディオ符号化及びマルチメディア信号処理に関する研究開発に従事．ISO/IEC JTC1/SC29/WG11（MPEG）においてオーディオ符号化の国際規格化に貢献．IEEE 会員．