

透視射影を考慮した
paraperspective モデルの逐次型因子分解法

藤木 淳, 蔵田 武志
電子技術総合研究所

0. Structure From Motion 問題と因子分解法

SFM 問題とは画像列からカメラ運動と立体形状を復元する問題

⇒ 「カメラモデル」と「座標系の記述」

因子分解法 (Tomasi・Kanade(1992) 他) は立体形状を基準とした座標系

カメラモデルとして透視射影とアフィン射影が考えられる

アフィン射影近似

定式化が容易, 数値計算上安定

モデルとしての誤差が不可避 ⇒ 透視射影に拡張

透視射影の因子分解法では射影的奥行きと呼ばれるパラメータの推定が必要

Strum・Triggs(1996): エピポーラ幾何を用いて推定

⇒ 特徴点の計測誤差に敏感

植芝・富田(1998): 評価関数を用いて推定

⇒ 収束に時間がかかる

Christy・Horaud(1996): Euclid 復元解を用いて推定

⇒ カメラパラメータが既知である必要

1. バッチ処理の因子分解法の特徴と逐次型への拡張

すべての情報を行列の形で表現して同時に扱う

⇒ 全画像を同等に扱うことによる誤差分散

⇒ 比較的良い精度の復元解

求解が線型演算で構成 ⇒ 数値計算上安定

全フレームが同時に必要 ⇒ 実時間処理には逐次型への改良が要求される

Morita・Kanade(1997) の逐次型因子分解法

♡ 形状行列の行空間を近似 ⇒ 初期段階の推定量が信頼できない

藤木他 (1998)

♡ 主成分分析による情報圧縮 ⇒ 全画像の情報をなるべく同等に扱う

提案手法 (1999) (= 藤木他 (1998) + Christy・Horaud(1996))

透視射影による像からアフィン射影による像を推定してアフィン射影モデルの因子分解法の復元精度を高める手法

⇒ 高精度の復元, バッチ処理よりも高速

3. Paraperspective モデルの因子分解法の記述

特徴点の $\left\{ \begin{array}{l} \text{世界座標} \\ \text{画像座標} \end{array} \right\}$ を $\left\{ \begin{array}{l} \mathbf{s}_p^* \\ \mathbf{x}_{fp}^* \end{array} \right\} \implies \mathbf{x}_{fp}^* = M_f \mathbf{s}_p^*$ (特定の特徴点を基準)

$M_f = A_f C_f$ と分解

C_f : カメラの向きを決める行列 (カメラ基底行列)。

$A_f = \frac{1}{\lambda_{f*}} (l \mathbf{I}_2 \mid -\mathbf{x}_{f*})$: Paraperspective カメラを表す行列
 (\mathbf{x}_{f*} : 特定の特徴点の画像座標)

M_f から C_f は求められるので運動 (C_f) の復元は M_f の復元と同等

$\{\mathbf{x}_{fp}^*\}_{f=1,p=1}^{F,P}$ から $\{M_f\}_{f=1}^F$ 及び $\{\mathbf{s}_p^*\}_{p=1}^P$ を復元 \implies SFM 問題を解く。

3. Paraperspective モデルの因子分解法の記述 (2)

特徴点の, 世界座標と画像座標の関係式 $\mathbf{x}_{fp}^* = M_f \mathbf{s}_p^*$ を並べる

$$W_f^* = (\mathbf{x}_{f1}^*, \dots, \mathbf{x}_{fP}^*), \quad S^* = (\mathbf{s}_1^*, \dots, \mathbf{s}_P^*) \implies W_f^* = M_f S^*$$

$$W^* = \begin{pmatrix} W_1^* \\ \vdots \\ W_F^* \end{pmatrix}, \quad M = \begin{pmatrix} M_1 \\ \vdots \\ M_F \end{pmatrix} \implies W^* = M S^*.$$

W^* を $M \times S^*$ に分解 \iff 運動と形状の復元

(W^* : 計測行列, M : 運動行列, S^* : 形状行列)

計量拘束

$$C_f \text{ が直交行列} \iff M_f M_f^T = A_f C_f C_f^T A_f^T = A_f A_f^T$$

因子分解法とは

「Affine 復元」, 「Euclid 復元」の順で運動と形状を復元 (W^* を分解)

3. Paraperspective モデルの因子分解法の記述 (3)

(1) Affine 復元

W^* を $\hat{M}_{(2F \times 3)}$ と $\hat{S}^*_{(3 \times P)}$ の積に一時的に分解 (ex. SVD)

\implies この時点で Affine 復元は完了している

\implies (C_f の直交性を考慮し) Euclid 復元解に修正

(2) Euclid 復元

Euclid 計量に修正する正則行列 A の決定 ($M = \hat{M}A$, $S^* = A^{-1}\hat{S}^*$)

A の拘束条件 $\implies \hat{M}_f Q \hat{M}_f^T = A_f^T A_f$ ($Q = AA^T$)
(ここで $\hat{M} = (\hat{M}_1^T, \dots, \hat{M}_F^T)^T$).

$\implies Q$ を計算

$\implies A = LU_3$ $U_3 \in O(3)$ ($Q = LL^T$: Cholesky 分解)

U_3 : 世界座標の取り方の自由度

$\implies \det U_3$ の正負に対応する互いに鏡映な 2 つの解

4 Christy・Horaud(1996) の手法 (を因子分解法の枠組で記述)

透視射影の像から paraperspective 射影の像を反復推定する手法

特徴点の $\left\{ \begin{array}{l} \text{paraperspective 射影} \\ \text{透視射影} \end{array} \right\}$ による画像座標を $\left\{ \begin{array}{l} \mathbf{x}_{fp}^{*\text{para}} \\ \mathbf{x}_{fp}^{*\text{per}} \end{array} \right\}$ 。
(特定の特徴点を基準)

$$\implies \mathbf{x}_{fp}^{*\text{para}} = \mu_{fp}^* \mathbf{x}_{fp}^{*\text{per}}, \quad \mu_{fp}^* = \frac{\lambda_{fp}}{\lambda_{f*}}$$

(λ_{fp} : 射影的奥行き, λ_{f*} : 特定の特徴点 * の奥行き)

μ_{fp}^* は特徴点 * の奥行きを 1 とした相対射影的奥行き

射影的奥行きを考慮した計測行列 W^* :

$$W^* = \begin{pmatrix} \mu_{11}^* \mathbf{x}_{11}^* & \cdots & \mu_{1P}^* \mathbf{x}_{1P}^* \\ \vdots & \ddots & \vdots \\ \mu_{F1}^* \mathbf{x}_{F1}^* & \cdots & \mu_{FP}^* \mathbf{x}_{FP}^* \end{pmatrix} \quad (1)$$

観測された \mathbf{x}_{fp}^* は透視射影の像とみなせるから μ_{fp}^* を反復推定し、paraperspective 射影の像を求める。

4.1 C-H 手法のアルゴリズム

- (0) $\mu_{fp}^* = 1$ とおき, 因子分解法により Euclid 復元。このとき一方の復元解を $M^{(+)}, S^{*(+)}$, その鏡映解を $M^{(-)}, S^{*(-)}$ とし,

$$\mu_{fp}^{*(\pm)} \leftarrow \frac{\lambda_{fp}^{*(\pm)}}{\lambda_{f*}^{(\pm)}}$$

により, それぞれの解に対して $\mu_{fp}^{*(\pm)}$ を更新 (以下, 複号同順)。

- (1) $\mu_{fp}^{*(\pm)}$ を用いて計測行列 $W^{(\pm)}$ を更新し, 因子分解法による Euclid 復元。
- (2) それぞれの $W^{*(\pm)}$ から得られる 2 組の復元解のうち $M^{(\pm)}, S^{*(\pm)}$ と同じ向きをみたす組を用いて $M^{(\pm)}, S^{*(\pm)}$ を更新。
- (3) $\mu_{fp}^{*(\pm)}$ を更新 ($\mu_{fp}^{*(\pm)}$ の変化が閾値 ϵ 以下でなければ (1) に戻る)。
- (4) 推定されて $M^{(\pm)}, S^{(\pm)}$ から透視射影による画像座標 \hat{x}_{fp}^* を再計算して, 実際に観測された画像座標とのずれを小さくする方を復元解として選ぶ。

4 透視射影を考慮した逐次型因子分解法

4.1 逐次型因子分解法には何が必要か

復元結果は世界座標系の取り方の自由度を持つ

- ⇒ 世界座標系によって、同一復元解の表現が異なる
- ⇒ 推定結果の更新には世界座標の重ね合わせが必要

時間と共に画像枚数が増える

- ⇒ 実時間処理には画像情報の圧縮が不可欠

特定の誤差の多い画像の影響を受けないために、すべての画像のもつ情報をまんべんなく圧縮する必要がある。

Euclid 復元を行なうので、計量拘束も考慮した圧縮が必要

- ⇒ 単に計測行列を圧縮 (主成分分析) するだけでは駄目
- ⇒ 運動行列を圧縮して計量拘束を抽出

隠れ点の処理や新しい特徴点の追加についても考慮する必要がある。

4.1 世界座標の重ね合わせ

$S_{\text{ob}}^{(G)}$ を $S_{\text{ref}}^{(G)}$ に重ねる直交行列を直接最小 2 乗推定

$\implies S_{\text{ref}}^{(G)} S_{\text{ob}}^{(G)\text{T}}$ の SVD を UDV^{T} とすると, 求める直交行列は UV^{T}

4.2 隠れ点の処理

観測された点の世界座標だけを更新 \implies 隠れた点は更新しない。

4.3 新しい特徴点の追加

画像集合 $\mathcal{F} = \{f_1, \dots, f_J\}$ 全体において観測された (射影的奥行きを考慮した) 新しい特徴点の画像座標 $\mathbf{X}_{\langle \mathcal{F} \rangle, \text{new}}^* = (\mu_{f_1, \text{new}}^* \mathbf{x}_{f_1, \text{new}}^{*\text{T}}, \dots, \mu_{f_J, \text{new}}^* \mathbf{x}_{f_J, \text{new}}^{*\text{T}})^{\text{T}}$

$W_{\langle \mathcal{F} \rangle}^*, M_{\langle \mathcal{F} \rangle}, S_{\langle \mathcal{F} \rangle}^*$: \mathcal{F} における計測行列, 運動行列及び形状行列 (既知)

$$\begin{cases} W_{\langle \mathcal{F} \rangle}^* = M_{\langle \mathcal{F} \rangle} S_{\langle \mathcal{F} \rangle}^* \\ \mathbf{X}_{\langle \mathcal{F} \rangle, \text{new}}^* = M_{\langle \mathcal{F} \rangle} \mathbf{s}_{\text{new}} \end{cases} \iff \begin{cases} S_{\langle \mathcal{F} \rangle}^* = \mathcal{T}_{\langle \mathcal{F} \rangle} W_{\langle \mathcal{F} \rangle}^* \\ \mathbf{s}_{\text{new}} = \mathcal{T}_{\langle \mathcal{F} \rangle} \mathbf{X}_{\langle \mathcal{F} \rangle, \text{new}}^* \end{cases}$$

$M_{\langle \mathcal{F} \rangle}$ の一般逆行列 $\mathcal{T}_{\langle \mathcal{F} \rangle}$ は画像座標から世界座標への変換行列

$$(0) \mu_{f_j, \text{new}}^* = 1, \quad (1) \mathbf{s}_{\text{new}}^* \leftarrow \mathcal{T}_{\langle \mathcal{F} \rangle} \mathbf{X}_{\langle \mathcal{F} \rangle, \text{new}}^*, \quad (2) \mu_{f_j, \text{new}}^* \leftarrow \frac{\lambda_{f_j, \text{new}}}{\lambda_{f_j^*}}$$

4.4 運動情報の抽出

運動行列はカメラ運動だけでなく観測している空間の計量拘束を与える

⇒ Euclid 復元を行なうには, 計量拘束を考慮した運動情報の圧縮

空間の計量は, 互いに線型独立な 3 つの基底間の計量で決定される。

⇒ どのような基底を選ぶか? そのときの計量は?

この共分散行列を良く表現する基底は? そのときの計量は?

$$M \longrightarrow F\Lambda E \text{ (SVD)} : \begin{cases} \Lambda: M \text{ の主成分,} \\ E: \text{主成分ベクトル,} \\ F: \text{観測という演算} \implies \text{本質的でない.} \end{cases}$$

E は運動空間を表現する正規直交基底の最小 2 乗推定量

Λ は基底方向の分散 \implies 基底方向の情報の「確度」。

$M = \Lambda E$ によって, 運動情報 (計量拘束) の抽出 (圧縮) の実現

運動空間の計量拘束は $MM^T = \Lambda^2$

行列のサイズは常に $3 \times 3 \implies$ 計算時間の短縮

これを主成分運動行列と呼ぶ。

4.4 運動情報の抽出 (2)

$W^* = MS$ によって, 圧縮された計測行列が求まる

これを主成分計測行列と呼ぶ。

(注) 単に計測行列を主成分分析しただけでは計量拘束が求まらない

運動情報, 計測行列は以下のように圧縮される

$$\begin{cases} W^* = MS^* \\ M_f M_f^T = A_f A_f^T \end{cases} \implies \begin{cases} W^* = \mathcal{M} S^* \\ \mathcal{M} \mathcal{M}^T = \Lambda^2 \end{cases}$$

つまり, W^*, M, A_f から $\{W^*, \mathcal{M}, \Lambda\}$ を求めることが運動情報の抽出である

4.5 アルゴリズム

(0) $k(\geq 3)$ 枚の画像から C-H 手法を用いて Euclid 的復元

$$\Rightarrow \left\{ \begin{array}{l} \text{第 } k \text{ 主成分運動行列 } \mathcal{M}_{[k]} \\ \text{第 } k \text{ 主成分計測行列 } \mathcal{W}_{[k]} \\ \text{計量拘束 } \mathcal{M}_{[k]} \mathcal{M}_{[k]}^T = \Lambda_{[k]}^2 \end{array} \right\} \text{ を求める。}$$

ここで得られた解を世界座標の基準とする

(1) 第 $f(> k + 1)$ 逐次計測行列の構成 (観測された特徴点のみから構成)

$$\begin{pmatrix} \mathcal{W}_{[f-1]}^* \\ \mathcal{W}_f^* \end{pmatrix} \quad (\text{各画像座標は, 着目した特徴点からの相対座標})$$

(2) C-H 手法を用いて Euclid 的復元し, 復元結果を基準世界座標系で表現。

$$\begin{pmatrix} \mathcal{W}_{[f-1]}^* \\ \mathcal{W}_f^* \end{pmatrix} = \begin{pmatrix} \mathcal{M}_{[f-1]} \\ M_f \end{pmatrix} S^* \text{ により, 第 } f \text{ 運動行列はの下の } 2 \text{ 行。}$$

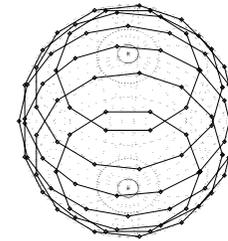
(3) 隠れ点の処理及び新しい特徴点の追加 \Rightarrow 第 f 画像を用いた復元は完了

$$(4) \text{ 運動情報の抽出 } \Rightarrow \left\{ \begin{array}{l} \text{第 } f \text{ 主成分運動行列 } \mathcal{M}_{[f]} \\ \text{第 } f \text{ 主成分計測行列 } \mathcal{W}_{[f]} \\ \text{計量拘束 } \mathcal{M}_{[f]} \mathcal{M}_{[f]}^T = \Lambda_{[f]}^2 \end{array} \right\} \text{ を求める。}$$

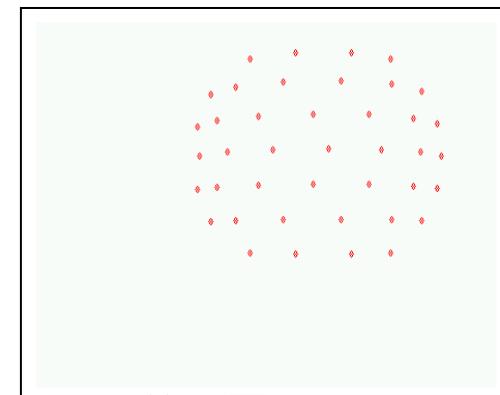
4. 実験

4.1 データ生成

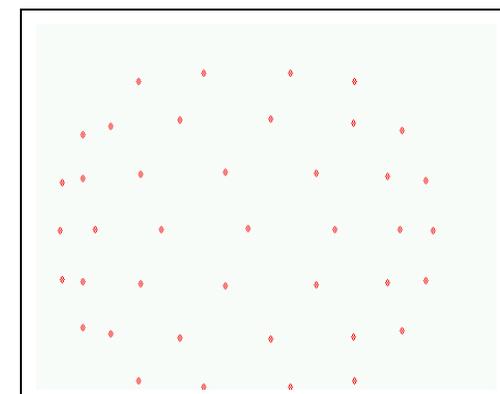
- 焦点距離は 84mm(35mm カメラ相当)
- $\{s_p\}_{p=1}^{92}$ は一辺が 150mm の球面上に生成
- 121 フレーム分の画像座標は透視射影にて生成
- 画像座標は sub-pixel 単位で標本化
- 第 f カメラ座標基底行列
 $C_f = R_y(\pi f/60)$ 。
- 奥行きは球の中心が 1200mm から 800mm となるよう単調に変化
- 初期推定量は, 初めの 10 画像から推定
- 連続する 10 画像で観測された特徴点を追加



生成された特徴点



第 1 画像

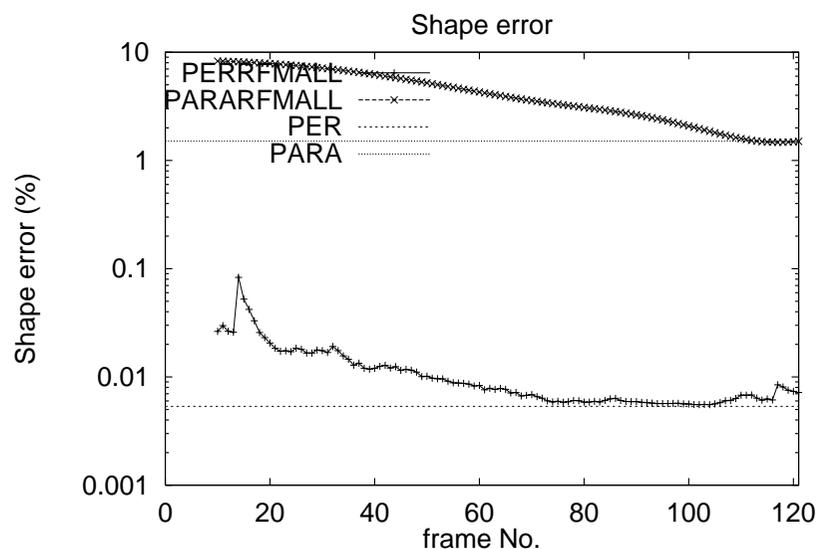


第 121 画像

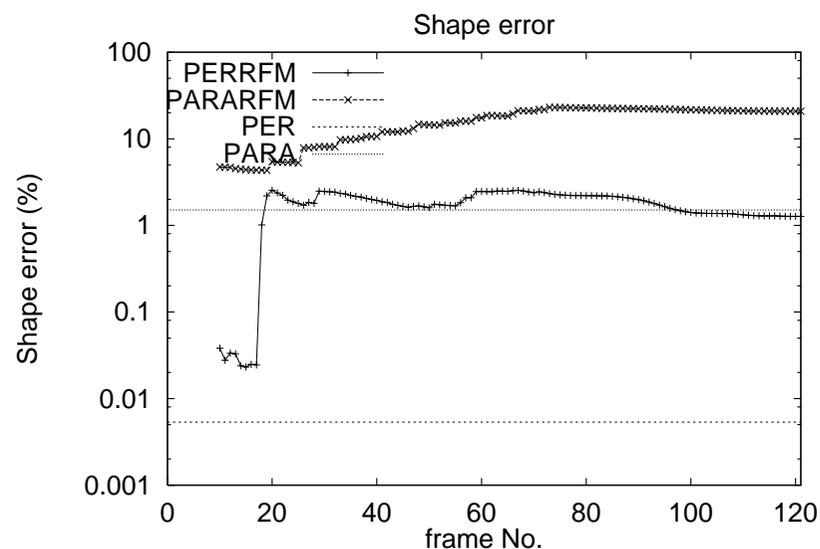
5.3 提案手法の性能評価

形状復元

$$e_{\text{shape}}(f) = \frac{\|S^{(G)\text{true}} - S^{(G)\text{estimated}}\|}{\|S^{(G)\text{true}}\|} \times 100(\%).$$



隠れなし



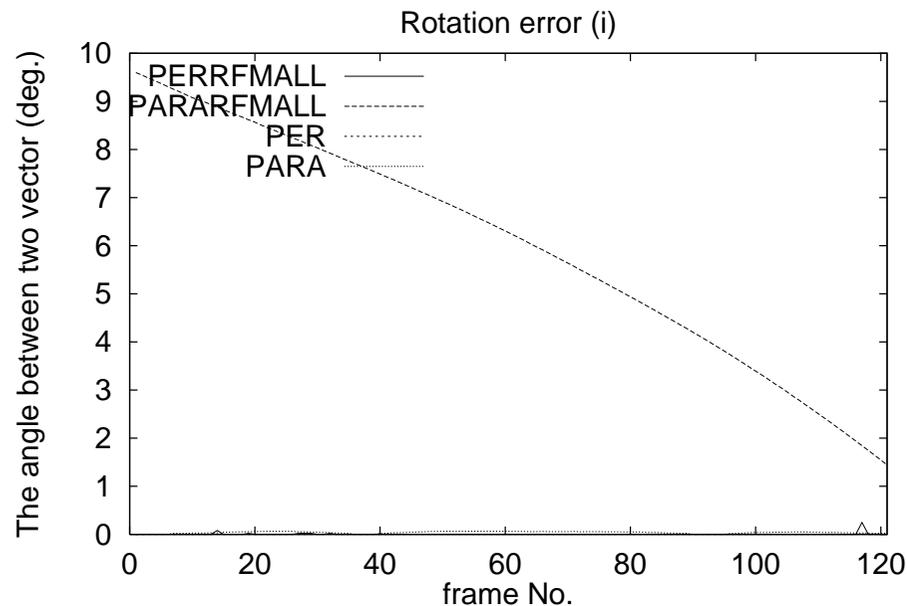
隠れあり

- 提案手法による形状復元精度は非常に高精度。
- 隠れや点の追加を行なっても、良い精度の復元。
- 2 pixel 程度のガウシアンノイズに対し誤差は 0.3% 程度のみ増加。

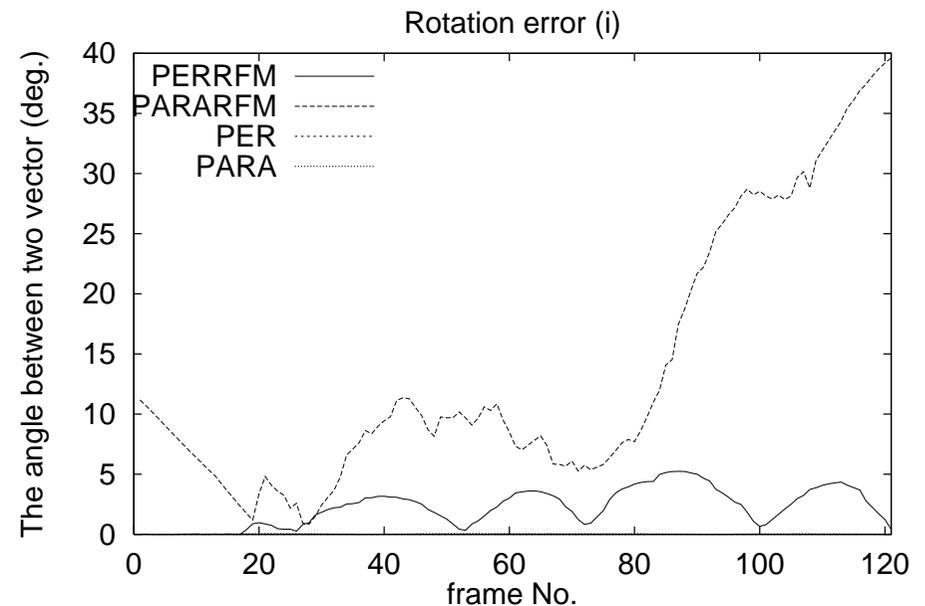
運動復元 (2つのベクトルの成す角)

$$e_i(f) = \cos^{-1} \left| \left| (\mathbf{i}_f^{\text{estimated}})^T \mathbf{i}_f^{\text{true}} \right| \right|. \quad (\text{度}) \quad (2)$$

$\mathbf{j}_f, \mathbf{k}_f$ についても同様に評価した。



隠れなし

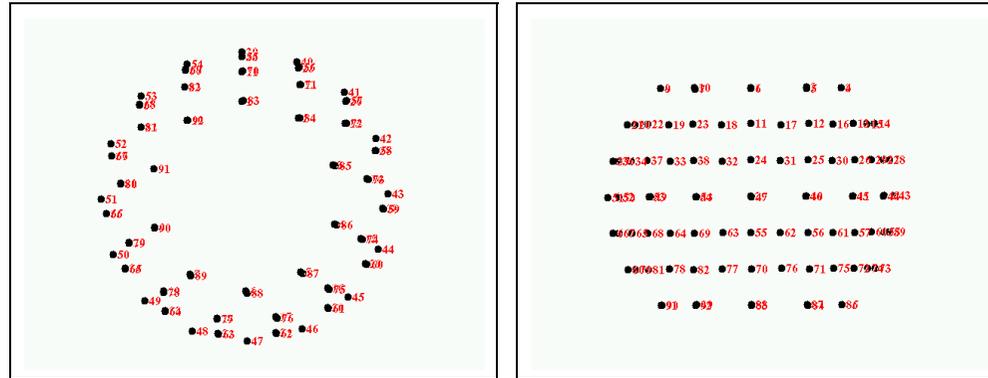


隠れあり

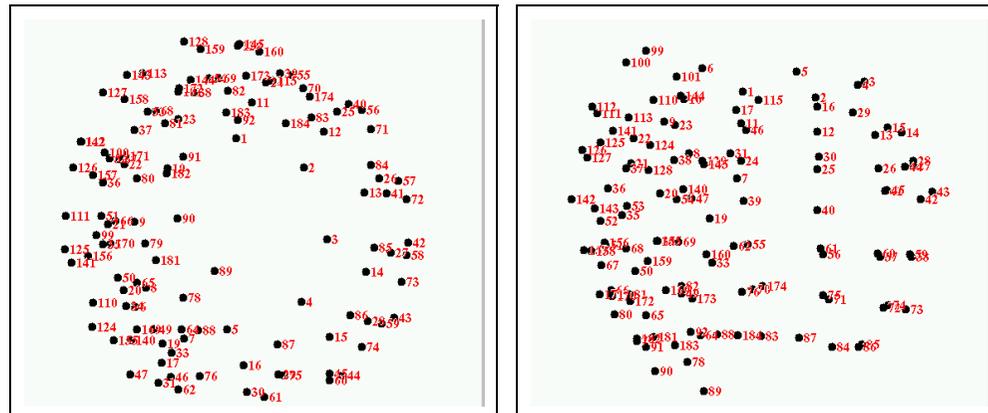
- 提案手法による形状復元精度は非常に高精度。
- 隠れや点の追加を行なっても、良い精度の復元。

復元された形状 (左は極から, 右は赤道からみた形状)

提案手法

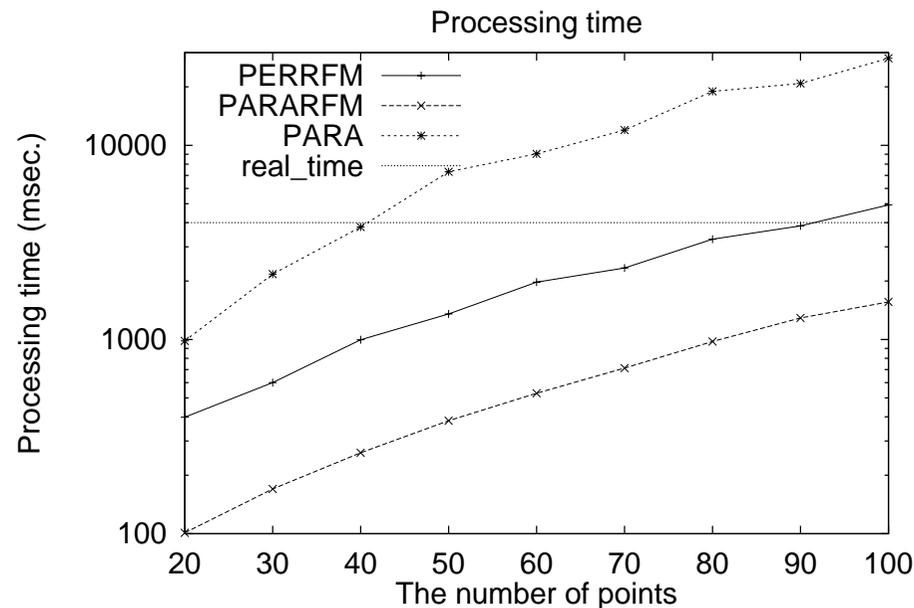


アフィン射影の逐次型因子分解法 (藤木他 (1998)) による復元結果



計算時間 (Pentium II 300MHz PC)

- 特徴点は 10 点から 100 点まで増やし, 画像枚数は 121 枚で固定
- 計算時間には特徴点の選択や追跡にかかった時間は含まれていない
- オリジナルの計算時間は各フレームにおける計算時間の合計



- 提案手法はバッチ処理に比べて計算時間を大幅に短縮。
(90 点程度までなら実時間で処理できる)

6. 結論

以下のような性能を持つ逐次型因子分解法を提案した。

- ♡ 各フレームにおいて、運動と形状を高精度で推定。

透視射影を考慮

- ♡ 高速で安定した手法であり、80 特徴点程度であれば、他の処理との同期のための待ち時間を含めても実時間に対応できる。

主成分運動行列 $\mathcal{M}_{[f]}$ によるサイズの小型化

- ♡ 全画像をなるべく同等に扱い、特定の画像の影響を強く受けにくい。

主成分運動行列 $\mathcal{M}_{[f]}$ による運動情報の抽出