

ロバスト推定を用いた 一般アフィン射影モデルの因子分解法

蔵田 武志 藤木 淳 坂上 勝彦

✦ 電子技術総合研究所

はじめに

structure-from-motion

- 一般アフィン射影モデルの因子分解法
 - 数値計算的に安定
 - 全画像を均等に扱える
 - 定式化が簡潔

実環境での応用を考える... (video)

- 最初から最小自乗法 (SVD) は使えない
 - ロバスト推定 (LMedS 推定) を利用
- 実時間処理 (毎フレーム結果を出力) するためには...
 - 逐次型因子分解法を拡張

本発表の内容

- 一般アフィン射影 (GAP) モデルの因子分解法の概説
- LMedS 推定を用いた因子分解法
 - LMedS 基準によるアフィン運動行列の復元
 - 外れ値として振舞う特徴点の除去
- 特徴点選択による逐次型ロバスト因子分解法
 - 逐次型因子分解法の概説
 - 逐次型ロバスト因子分解法
- シミュレーションと実画像実験
- むすびと今後の課題

一般アフィン射影 (GAP) モデルの因子分解法

第 1 段階 : アフィン運動・形状の復元

修正計測行列 W' は, 運動行列 M と形状行列 S' に分離可能

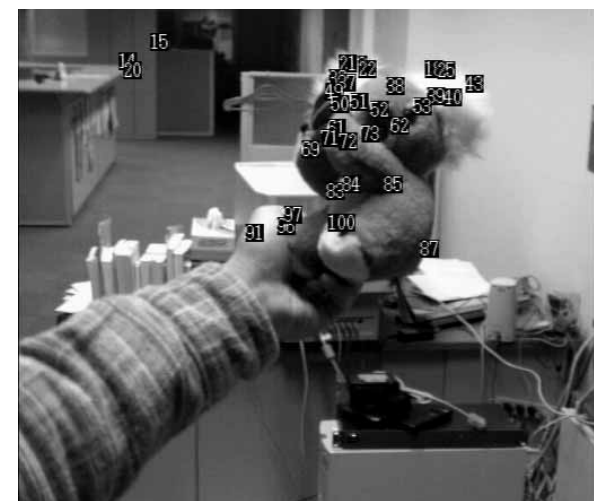
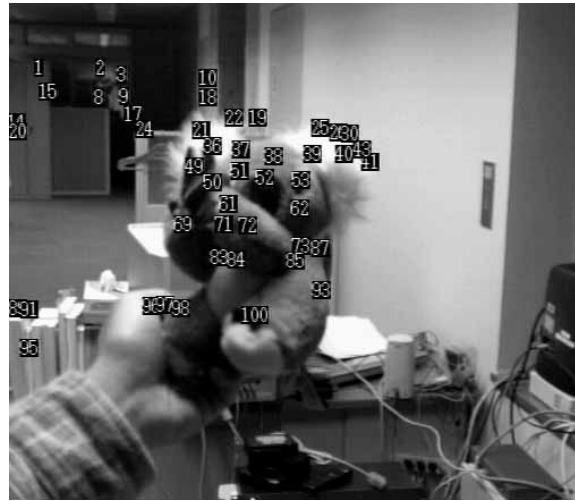
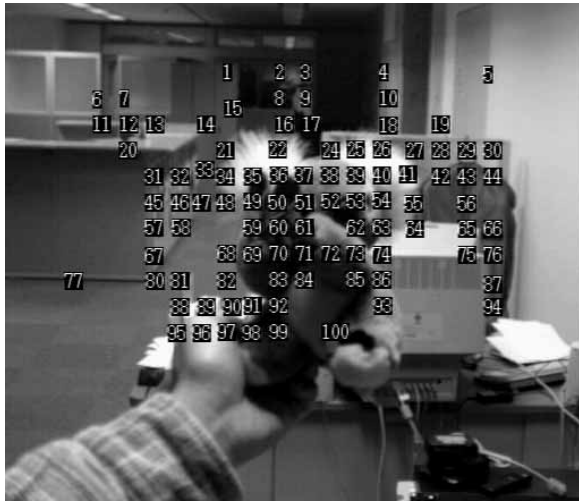
$$W'_{(2F \times P)} = \begin{pmatrix} A_1 C_1 \\ \vdots \\ A_f C_f \\ \vdots \\ A_F C_F \end{pmatrix} S' = \begin{pmatrix} M_1 \\ \vdots \\ M_f \\ \vdots \\ M_F \end{pmatrix} S' = \begin{matrix} M \\ (2F \times 3) \end{matrix} \begin{matrix} S' \\ (3 \times P) \end{matrix}$$

第 2 段階 : 計量拘束

第 1 段階の分離は線形変換の自由度を持つため, カメラ座標基底の正規直交性から得られる計量条件を用いて, ユークリッドな運動と形状を復元する

$$\hat{M}_f Q \hat{M}_f^T = A_f A_f^T, \quad Q = \begin{matrix} A A^T \\ (3 \times 3) \end{matrix}$$

LMedS 推定を用いた因子分解法



A: [frame 1: 100 points] B: [frame 33: 48 points] C: [frame 99: 35 points]

入力時系列画像 (99 フレーム) と特徴点追跡結果

実環境での特徴点追跡結果には...

他物体に属するもの, 追跡処理の失敗による誤対応などが含まれる



特徴点 (計測行列の列) を LMedS 推定により選択

基本戦略

- ランダムに選ばれる 4 点の特徴点からアフィン運動を復元
(ランク落ちしている場合は選びなおす)
- 復元された運動を用いて全特徴点のアフィン形状を復元
- 復元された形状の逆投影像と画像上の特徴点座標との自乗誤差を時系列で累積
(LMedS 推定で用いる評価値)



LMedS 基準による準最適なアフィン運動行列を選択

試行回数

特徴点 4 点すべてが外れ値でない確率を 99.9% とするためには、

40% の外れ値 — 50 回以上

50% の外れ値 — 72 回以上

本発表の実験... 100 回の試行

外れ値として振舞う特徴点の除去

- 選択されたアフィン運動に従わない特徴点を除去
(外れ値以外の特徴点の観測誤差が正規分布に従うと仮定)
- 残りの点全てを用いて運動と形状を復元

逐次型ロバスト因子分解法

実時間で結果を出力したい（新しいフレームが入力されたらすぐに結果を出力）



逐次型因子分解法をロバスト型に拡張

逐次型因子分解法の概説

基本戦略

- 計算量をフレーム数に関わらず一定にする
⇒ 運動情報（計量条件）と計測行列を PCA で圧縮
- 各フレームで世界座標系を固定する
⇒ 2 時点の世界座標を結ぶ直交変換を形状行列から最小自乗法により推定

アルゴリズム

[1] アフィン構造の復元：
$$W'_{[f]} = \begin{pmatrix} \mathcal{W}'_{[f-1]} \\ W'_f \end{pmatrix} \implies \hat{M}_{[f]} \hat{S}'_{[f]}, \quad \hat{M}_{[f]} = \begin{pmatrix} \hat{\mathcal{M}}_{[f]} \\ \hat{M}_f \end{pmatrix}$$

[2] 計量を復元し，ユークリッドな運動と形状 $(\bar{M}_{[f]}, \bar{S}'_{[f]})$ を求める：

圧縮された運動情報の計量条件：
$$\hat{\mathcal{M}}_{[f]} Q_{[f]} \hat{\mathcal{M}}_{[f]}^T = \Lambda_{[f-1]}^2$$

新しいフレームの計量条件：
$$\hat{M}_f Q_{[f]} \hat{M}_f^T = A_f A_f^T$$

[3] 世界座標系を固定 $(\bar{S}'_{[f]}$ を $S'_{[f-1]}$ に変換する直交行列 \mathcal{E}_f を求める)：

$$M_f = \bar{M}_f \mathcal{E}_f^T, \quad S'_{[f]} = \mathcal{E}_f \bar{S}'_{[f]}$$

[4] 計測行列と運動情報 (計量条件) の圧縮

(運動行列を SVD $M_{[f]} \implies F_{[f]} \Lambda_{[f]} E_{[f]}$)：

運動情報 (計量条件) の圧縮：
$$\mathcal{M}_{[f]} = F_{[f]}^T M_{[f]} (= \Lambda_{[f]} E_{[f]})$$

(3×3)

計測行列の圧縮：
$$\mathcal{W}'_{[f]} = \mathcal{M}_{[f]} S'_{[f]} (= F_{[f]}^T W'_{[f]})$$

(3×P)

特徴点選択による逐次型ロバスト因子分解法

アルゴリズム

[1] アフィン構造の復元 :

← LMedS 推定を導入 (外れ値とみなされる特徴点は除去される)

[2] 計量の復元 :

← 計量拘束は特徴点数に依存しない

[3] 世界座標系を固定 ($\bar{S}'_{[f]}$ を $S'_{[f-1]}$ に変換する直交行列 \mathcal{E}_f を求める) :

← 共通する特徴点を利用

[4] 計測行列と運動情報 (計量条件) の圧縮 :

計測行列の圧縮 : $\underset{(3 \times P)}{W}'_{[f]} = \mathcal{M}_{[f]} S'_{[f]}$

← 過去の復元結果を利用し, 誤って外れ値と判断された点を復帰

シミュレーション

設定

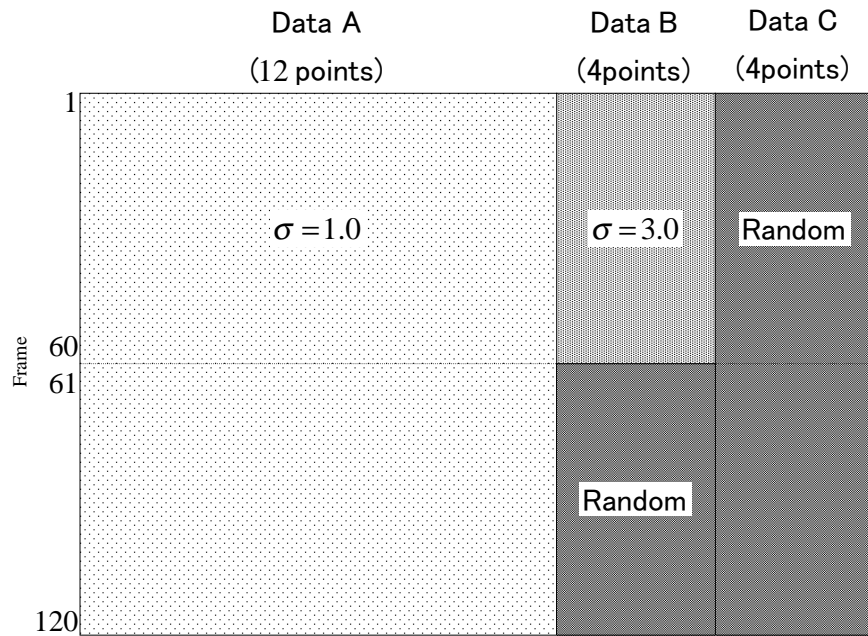
- $\{s_p\}_{p=1}^{20}$ は一辺 200mm の立方体内にランダムに一様分布
- カメラはピンホールカメラを仮定．焦点距離は 35mm カメラ相当で84mm
- 第 $f(= 120)$ カメラ座標基底行列 C_f

$$C_f^T = R_x\left(\frac{4}{3}f\right) R_y(0) R_z\left(\frac{1}{3}f\right) \quad (f = 1, \dots, 60)$$

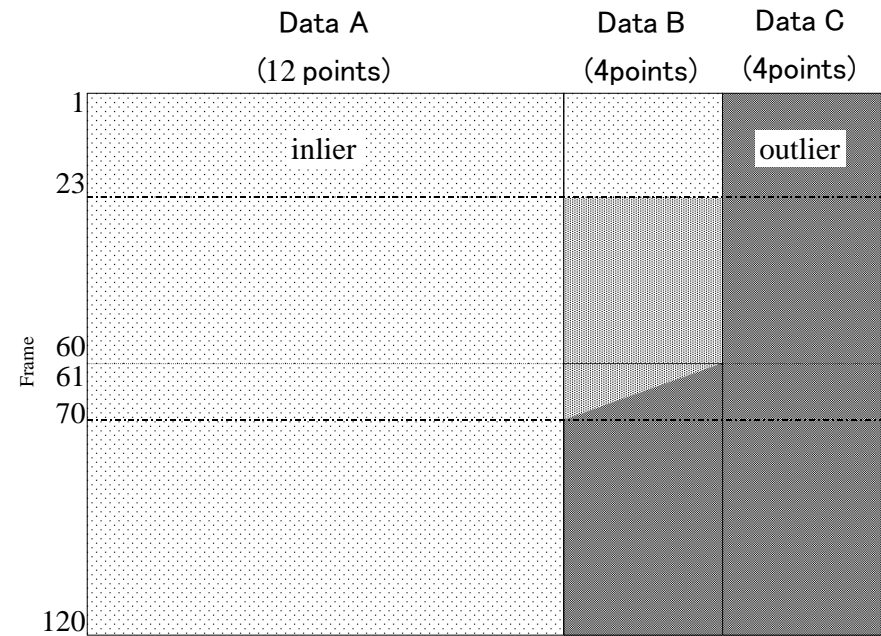
$$C_f^T = R_x\left(160 - \frac{4}{3}f\right) R_y(0) R_z\left(\frac{1}{3}f\right) \quad (f = 61, \dots, 120)$$

- 奥行きは $z_1 = 2000\text{mm}$, $z_{120} = 1600\text{mm}$ となるように等速移動
- $u_1 = 115\text{pixel}$, $u_{60} = 0\text{pixel}$, $u_{120} = 101\text{pixel}$ となるように滑らかに移動

特徴点の選択結果



観測誤差と外れ値の設定



特徴点選択結果

観測誤差と外れ値の設定：特徴点を 3 組に分けそれぞれ以下のように設定

データセット A [12 点]：第 1 から第 120 フレームまでの正しい追跡結果に平均 0，標準偏差 1pixel のガウシアンノイズを加えた点

データセット B [4 点]：第 1 から第 60 フレームまでの正しい追跡結果に平均 0，標準偏差 3pixel のガウシアンノイズを加え，第 61 フレームからは最大 2pixel/frame でランダムに動く点

データセット C [4 点]：第 1 から第 120 フレームまで最大 2pixel/frame でランダムに動く点

- paraperspective 射影モデルで復元．試行回数 J は 100 回．
- 誤差評価の比較対象：データセット A の 12 点のみを用いた通常の因子分解法
- 120 フレーム全てによる LMedS 推定では，データセット A の 12 点が選択された

特徴点選択結果

初期段階 ($k = 23$) データセット C を外れ値として除去

更新段階

60 フレームまで

- 残りの 16 点のうち 1, 2 点を頻繁に外れ値として除去．特に標準偏差 3pixel の誤差を持つデータセット B を外れ値として選ぶ傾向が強かった
- データセット A を外れ値として除去することもあったが，直前のフレームの復元結果を用いて，除去した点をその後のフレームで復帰できた

60 フレーム以降

- データセット B を徐々に外れ値として除去．第 69 フレーム以降はすべて除去．

むすび

- 成果

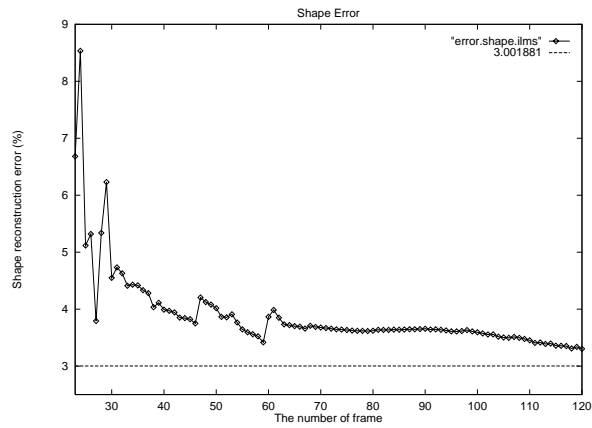
- LMedS 推定を用いて特徴点を選択し, 外れ値を含むデータに対してロバストに運動と形状を推定することができる因子分解法を提案
- 実時間で適用するために本手法を逐次型因子分解法へ拡張

- 課題

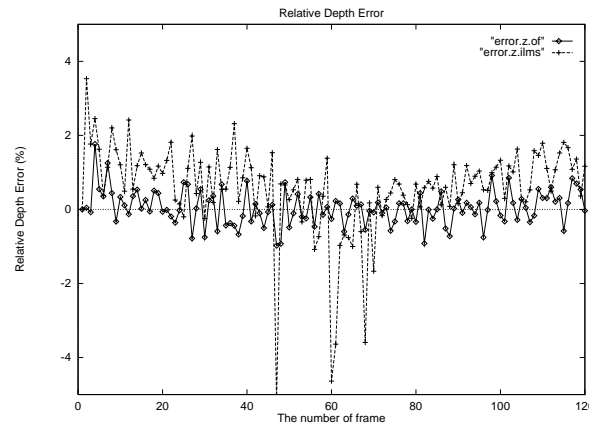
- 更新段階の時点で新たに追跡される特徴点についての扱い
- 提案手法の breakdown point についての評価
- 動きによる領域分離や特徴点追跡処理などと提案手法との協調処理による, 複数物体への対応や追跡精度の向上

各アフィン射影モデルの具体的なパラメータ

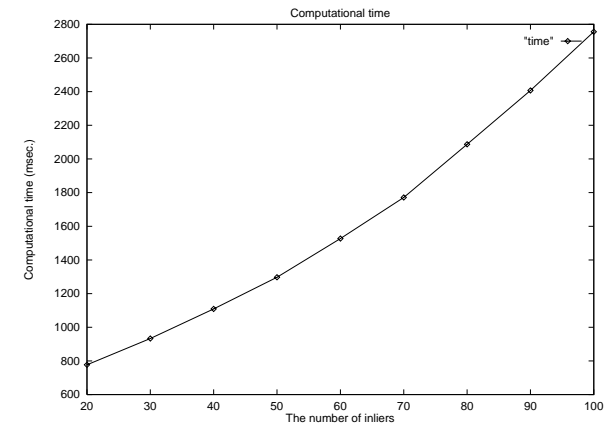
scaled orthographic 射影モデル (正射影モデル $\frac{l}{z_f} = 1$)	paraperspective 射影モデル
$A_f = \frac{l}{z_f} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	$A_f = \frac{1}{z_f} \begin{pmatrix} l & 0 & -x_{f.} \\ 0 & l & -y_{f.} \end{pmatrix}$
$R_f = I_2$	$R_f = \frac{1}{u_f} \begin{pmatrix} x_{f.} & -y_{f.} \\ y_{f.} & x_{f.} \end{pmatrix}$
$\Sigma_f = \frac{l}{z_f} I_2$	$\Sigma_f = \frac{1}{z_f} \begin{pmatrix} \sqrt{l^2 + u_f^2} & 0 \\ 0 & l \end{pmatrix}$
$z_f = \frac{l}{ \mathbf{m}_f } = \frac{l}{ \mathbf{n}_f }$	$\begin{aligned} z_f &= \frac{\sqrt{l^2 + x_{f.}^2}}{ \mathbf{m}_f } = \frac{\sqrt{l^2 + y_{f.}^2}}{ \mathbf{n}_f } \\ &= \sqrt{\frac{x_{f.} y_{f.}}{\langle \mathbf{m}_f, \mathbf{n}_f \rangle}} \end{aligned}$



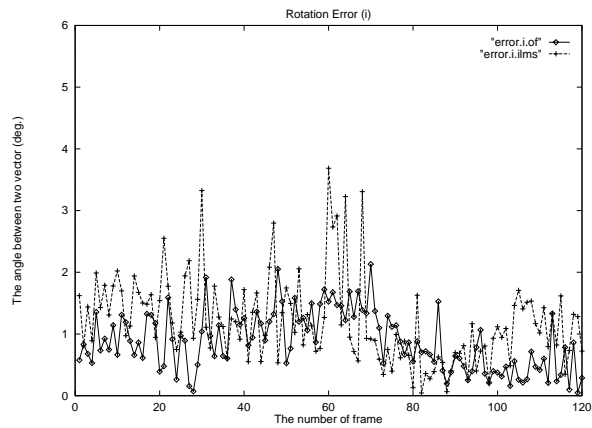
形状復元誤差



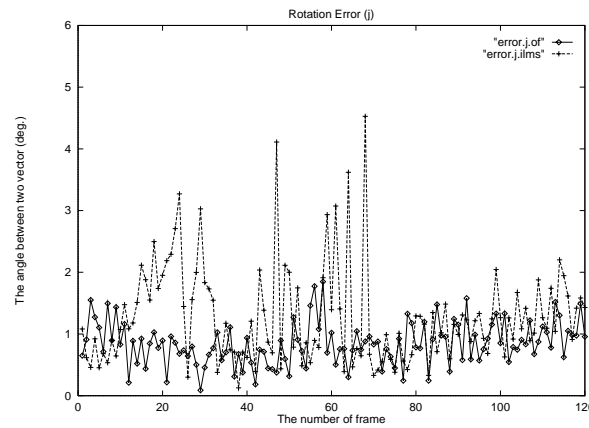
平均奥行き復元誤差



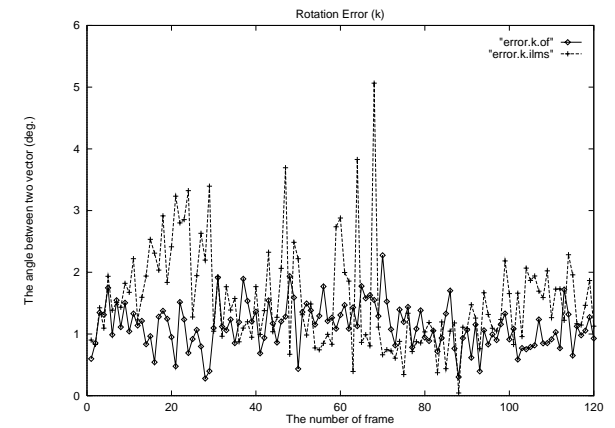
特徴点数と計算時間



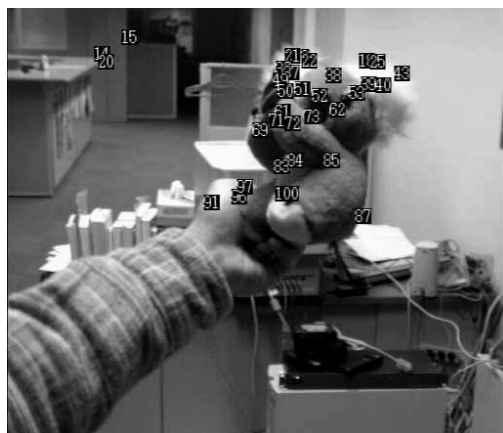
運動復元誤差 (i_f)



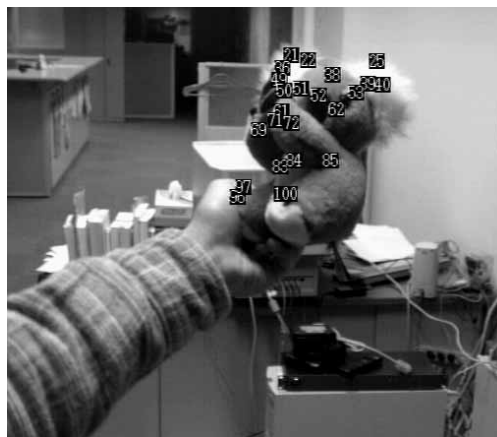
運動復元誤差 (j_f)



運動復元誤差 (k_f)



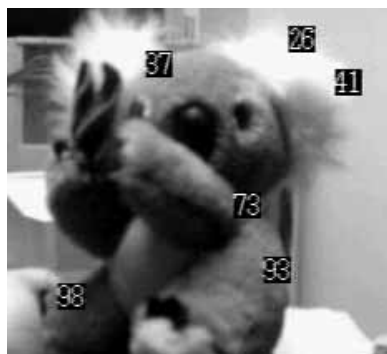
[frame 99: 35 points]
全特徴点



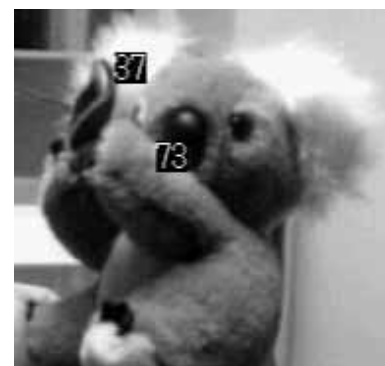
[frame 99: 23 points]
選択された点



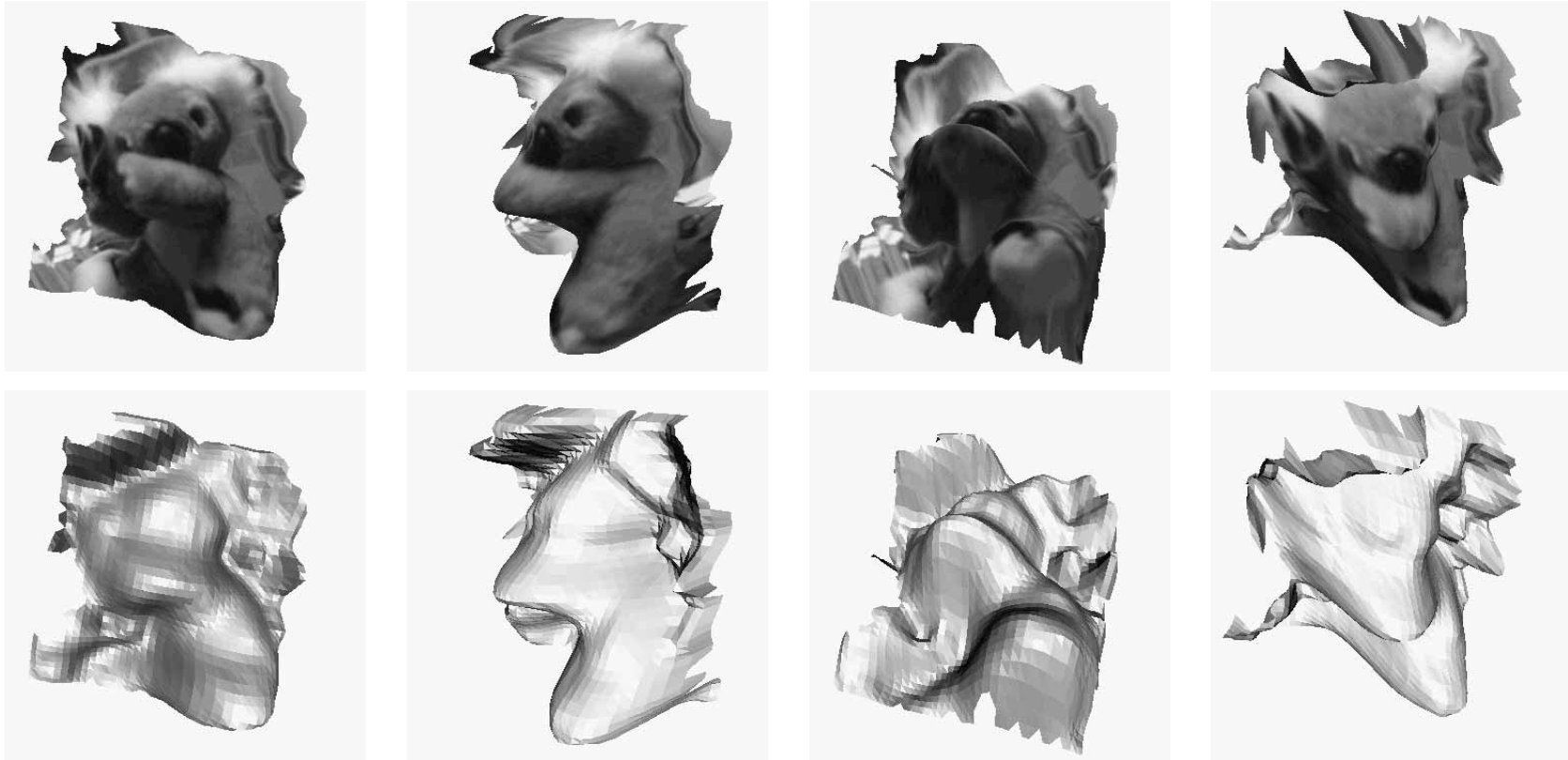
[frame 99: 12 points]
外れ値



外れ値 [frame 33: 6 points]



外れ値 [frame 99: 2 points]



第 64,65,66 フレームを用いた仮想 3 眼ステレオにより復元された
対象物体の密な 3 次元形状