

因子分解法による連続画像中の物体認識手法の改良

Improving Factorization Method for Recognizing Object in Sequential Images

高野秀也†

Hideya Takano †

†NTTデータ通信(株)

†NTT DATA Communications Systems Corp.

アラン・ブリック‡

Allan Brik ‡

‡マサチューセッツ工科大学

‡Massachusetts Institute of Technology

Abstract An improvement on the factorization method for computing three dimensional shape and motion from an object in a sequence of images is discussed. The original method requires correspondence between the feature points over all frames in the sequence. When feature points can only be detected in some frames (as in occlusion for example), a severe limitation is imposed to the original algorithm. We propose in this article an improvement on the factorization method that overcomes the above limitation by adapting the original method on the partial frames in order to compute 3D shape and motion and integrating them into the final result.

1 はじめに

画像認識、理解において3次元情報の果す役割は重要であり、その抽出方法について近年盛んに研究されている。その中で、単一のカメラから撮影した連続画像を用いる方法は、データ取得の簡便性や特徴の対応付けが比較的容易であるなどの理由から、有力なアプローチであり、さまざまな手法がこれまで提案されてきた[1]。

それらのいくつかは、物体上の特徴をフレーム間で追跡した情報からカメラと特徴の間の距離を求めるというものである。すなわち、得られる3次元情報はカメラ中心の座標系にて表わされていることになる。この表現の欠点は、カメラから各点までの距離の差が、物体の形状の凹凸から生じた有意なものなのか、それとも計算誤差の累積や雑音によるものなのかははっきりとしないことである。この傾向は、物体がカメラから遠くにある程、そして物体の形状の変化が小さい程、より強く表われる。

これに対して、フレーム間で追跡された点から、物体上の点とカメラとの距離の算出を経ずに、物体中心の座標系における点の座標を算出する方式[2]が提案された。この方式では、カメラ中心の座標系による表現固有の欠点を回避することができる。また、この方式を用いると、カメラ中心の座標系と物体中心の座標系の間の座標変換行列も算出される。この方式の要は、追跡された特徴点の2次元座標から成る行列を特異値分解することにある。そこで、以下、この方法を因子分解法と呼ぶこととする。最近では、この因子分解法を応用して、シーン中で別個の動きをする複数物体を分離する研究[3]も発表されている。

本稿では、まず2節でこの因子分解法の概要の説明を行なう。その後、3節にて因子分解法の問題点を指摘し、それを解決するための手法を4節にて提案する。

2 因子分解法の概要

2.1 因子分解法の前提

因子分解法では、以下の4つの前提を置いた上で話を進めている[3]。

- 3次元世界から2次元画像への投影モデルとして正射影を用いていること
- シーン中に対象物体は一つしかなく、かつその対象物体は剛体であること
- 連続画像の各フレーム画像の特徴として点を用い、その対応付けは全フレームに渡りできていること
- 連続画像は最低3フレームあり、全フレームにおいて最低3点が観測されていること

2.1.1 正射影による座標変換

ここでは、以下の記述の準備として正射影による座標系の変換式を記す。

一般的に、3次元空間内の2つの座標系の変換は、回転を表わす3次正方行列 R と、並進を表わす3次縦ベクトル T により記述される。3次元空間内に n 個の点がある場合に、 j 番目の点 (x_j, y_j, z_j) が座標変換により (X_j, Y_j, Z_j) へ移るとすると、座標変換の式は

$$P = RS + T \cdot \mathbf{1} \quad (1)$$

ここで、

$$\begin{aligned} P &= \begin{pmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \\ Z_1 & Z_2 & \cdots & Z_n \end{pmatrix} \\ S &= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ y_1 & y_2 & \cdots & y_n \\ z_1 & z_2 & \cdots & z_n \end{pmatrix} \\ T &= (t_i), i = 1, 2, 3 \end{aligned} \quad (2)$$

であり、また $R = (r_{ij})$, $i, j = 1, 2, 3$ は

$$RR^T = I \quad (3)$$

を満たす。ここで、 R^T は R の転置行列を表わし、 I は単位行列である。そして $\mathbf{1}$ は、要素すべてが1である n 次行ベクトル、すなわち、

$$\mathbf{1} = (\underbrace{1 \ 1 \ \cdots \ 1}_{n \text{ 個}})$$

である。

この変換後の点 (X_j, Y_j, Z_j) を正射影して X, Y 成分だけ取り出すと、式 (1) は、

$$P_{\text{proj}} = R_{\text{proj}} S + T_{\text{proj}} \cdot \mathbf{1} \quad (4)$$

となる。ここで、

$$\begin{aligned} P_{\text{proj}} &= \begin{pmatrix} X_1 & X_2 & \cdots & X_n \\ Y_1 & Y_2 & \cdots & Y_n \end{pmatrix} \\ R_{\text{proj}} &= \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \end{pmatrix} \\ T_{\text{proj}} &= \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \end{aligned}$$

この式 (4) は、以下では、あるフレームにおいて、物体上の n 個の点の3次元座標が正射影によって画像平面上に変換される関係を意味する。

2.2 因子分解法の概要

因子分解法の概要を順を追って説明する。

1. 観測値行列の作成

f 枚のフレームに渡り n 個の点 $P_1, \dots, P_n$ の対応が付けられているものとする。そして、 i 番目のフレームにおける j 番目の点 P_j の座標が (X_{ij}, Y_{ij}) であるとする。この時、次のように行列 $X = (X_{ij})$ と行列 $Y = (Y_{ij})$ を縦に並べて行列 W を作るのが第1段階である。ここで、第2.1節で述べたように、 f も n も3以上である。

この行列 W を、以下、観測値行列と呼ぶ。

$$W = \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{pmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ X_{f1} & X_{f2} & \cdots & X_{fn} \\ Y_{11} & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ Y_{f1} & Y_{f2} & \cdots & Y_{fn} \end{pmatrix} \quad (5)$$

2. 観測値行列の正射影による変換式

対象物体が剛体であるという2.1節第2の前提の下に、式 (4) を導出したのと同じようにして f フレームの特徴点を正射影した場合の式を導出できる。それは以下になる。

$$W = R_{\text{frames}} S + T_{\text{frames}} \cdot \mathbf{1} \quad (6)$$

ここで、 W は式 (5) で作成される観測値行列である。また、 S は物体上の n 個の点の3次元座標を表わす行列で、式 (2) の S と同じ成分を持つ。この行列を以下、形状行列と呼ぶ。 T_{frames} は要素 $2f$ 個の列ベクトルで、

$$T_{\text{frames}} = (t_{x1} \ \cdots \ t_{xf} \ t_{y1} \ \cdots \ t_{yf})^T$$

以下 S の座標系を、この行列が表わす n 個の点の重心を原点としたものにとる。すると、 T_{frames} は物体中心の座標系からカメラ中心座標系への変換における並進成分を表わす。また、

$$R_{\text{frames}} = \begin{pmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ \vdots & \vdots & \vdots \\ R_{f1} & R_{f2} & R_{f3} \\ R_{(f+1)1} & R_{(f+1)2} & R_{(f+1)3} \\ \vdots & \vdots & \vdots \\ R_{(2f)1} & R_{(2f)2} & R_{(2f)3} \end{pmatrix}$$

であり、各フレームにおける回転行列の要素を並べた行列である。すなわち、この行列の第 i 行をベクトル $\mathbf{r}_i$ と表わすと、 $\mathbf{r}_i$ と $\mathbf{r}_{i+f}$ は第 i フレーム目における正射影の関係式 (4) における回転行列 R_{proj} の2つの行ベクトルに相当する。よって、式 (3) の条件より、

$$\|\mathbf{r}_i\| = 1 \quad (7)$$

$$\|\mathbf{r}_{i+f}\| = 1 \quad (8)$$

$$\mathbf{r}_i \cdot \mathbf{r}_{i+f} = 0 \quad (9)$$

が成り立つ。

3. 並進ベクトル T_{frames} の算出

T_{frames} は物体中心座標系の並進を表わすことから、その各要素は以下のように求まる。

$$t_{xj} = \frac{1}{n} \sum_{k=1}^n X_{jk}, \quad t_{yj} = \frac{1}{n} \sum_{k=1}^n Y_{jk}$$

4. R_{frames} と S の算出

T_{frames} が求まると、式 (6) より、

$$W_r = W - T_{\text{frames}} \cdot \mathbf{1} = R_{\text{frames}} S \quad (10)$$

となる $2f$ 行 n 列の行列 W_r を求めることができる。

式 (10) の行列 W_r は、 $2f \geq n$ であれば特異値分解 [7] により、 $2f$ 行 n 列の直交行列 Q_1 、 n 次元直交行列 Q_2 、 n 次元対角行列 Σ を用いて以下のように分解される。

$$W_r = Q_1 \Sigma Q_2^T$$

もし $2f < n$ であれば、 W_r の転置行列を特異値分解した上で、後述と同じ手順を適用すれば良い。

ここで、 R_{frames} と S はそれぞれ $2f$ 行 3 列、3 行 n 列の行列であるから、 $\text{rank}(W_r) \leq 3$ である。よって、 Σ の中の特異値

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$$

のうち有意な特異値は $\sigma_1, \sigma_2, \sigma_3$ の 3 つしかない。そこで、この 3 つの特異値だけから成る 3 次元対角行列 $\hat{\Sigma}$ を作る。また、その 3 つの特異値に対応する Q_1 の列を抜き出して $2f$ 行 3 列の行列 $\hat{Q}_1$ を、同様にして対応する Q_2 の行を抜き出して 3 行 n 列の行列 $\hat{Q}_2$ を作る。この行列 $\hat{Q}_1$ に対して $\hat{Q}_1 A$ が式 (7) - (9) を満たすような 3 次正方行列 A を求める。この時、

$$R_{\text{frames}} = \hat{Q}_1 A, \quad S = A^{-1} \Sigma \hat{Q}_2$$

により R_{frames} と S が求まる。

2.3 因子分解法の特質

因子分解法には以下のような性質があるものと考えられる。

- カメラが動いている場合でも固定の場合と同様に適用できる

カメラが動いている場合でも、特徴点の対応さえ取れば、物体中心の座標系における 3 次元座標値を表わす行列 S が求まる。カメラの位置や運動が既知でなくはいけないなどの制約はない。しかしこの場合には、回転行列 S と並進行列 T はカメラの運動による影響を含んだものとなり、物体自身の運動を解析するためには使えない。以上の性質は、行列 S と観測値行列の間の座標変換の関係のみに基づいている手法であることから言える。

- カメラから物体までの絶対的な距離は求まらない

この性質は、投影モデルとして正射影を用いていることから来る [4]。ただしカメラが固定の場合には、 R_{frames} を式 (3) の条件の下に 3 次元化することにより、

物体上の各点とカメラとの距離の差は求めることが可能である。

- 観測されるフレームや特徴点の数が少ない場合やフレーム間での回転成分が少ないシーンに対しては、結果に誤差が生じやすい

これは因子分解法が、特異値分解により観測値行列中の有効な成分を抽出する方式であるからである。実際、(10) の行列 W_r が $\text{rank}(W_r) \neq 3$ である場合には、

R_{frames} や S は求まらない。

3 解決すべき問題点

上述したように、因子分解法は強い前提の上に成立している手法である。ここではまず、これらの前提について検討する。

ビデオカメラなどの撮像機器による撮影で得られる画像は透視変換であるから、その点で正射影という前提は実際とそぐわないものである。しかし、物体からカメラまでの距離が遠ければ遠いほど透視変換は正射影に近づくので、遠くの物体を扱う場合には問題ないと思われる。また、この正射影という前提により、数学的な扱いが楽になっているという利点もある。[4]

シーン中に一つの物体（運動）しか存在しないという前提も、物体が剛体であるという前提も、どちらも式 (6) を成立させるために必要である。すなわち、物体が非剛体であったり、シーン中に複数物体が存在すると、シーン毎に行列 S が変化することになってしまう。これに対する改善は、シーン中の複数の運動を解析しようという試みとして研究されている [3]。そこで、本稿では取り上げない。

また、フレーム中の特徴点数やフレーム数に対する前提は、あくまで特異値分解を行なって 3 次元情報を得るために、数学的に必要とされているものである。よって、この前提を変えることは不可能である。

最後に、全フレームにわたり特徴点の対応付けがされているという前提について述べる。この前提は以下の理由により、無理があるように考える。すなわち、実際のシーンでは、すべての特徴点が全フレームに渡り対応付けできるわけではない。他物体とのオクルージョンや視点の変化による自己オクルージョンなどの原因により一部のフレームにて隠れて見えない点が存在する場合や、または特徴点自体は全フレームにて見えていても対応付けに失敗する場合が十分起り得る。

逆に言えば、この前提の下に算出できる 3 次元情報は、対応付けされている特徴点のものだけということになる。これは、得られる 3 次元情報が限られてしまうことを意味し、その 3 次元情報を用いて物体の認識を行なうためには望ましいことではない。

例えば、対応付けされなかった点を含んだ以下のような観測値行列があるものとする。

$$\begin{pmatrix} 106 & 122 & -1 & 141 & -1 & 258 \\ 107 & 120 & -1 & 142 & -1 & 259 \\ 108 & 119 & -1 & 144 & 433 & 260 \\ 110 & 118 & 134 & 146 & 436 & 262 \\ 110 & 118 & 134 & 148 & 437 & 261 \\ -1 & 116 & -1 & 149 & 439 & 259 \end{pmatrix} \quad (11)$$

ちなみにこの例では、観測値行列 W のうちの上半分または下半分に当たる成分、すなわち式 (5) の行列 X （または行列 Y ）の成分のみを掲げている。それは、行列 X と行列 Y とは -1 が存在する位置も含めて同じ形であるからである。

また、 -1 が記入されている要素は対応付けられなかった点を示している。つまり、行列の i 行 j 列の要素が -1 なら、 i 番目のフレームにおいて j 番目の特徴点の対応付けが失敗していることを意味する。

従来の因子分解法を適用して式 (11) のような行列を処理しようとする、対応の付けられなかった点を除く操作をするしかない。すなわち、 -1 である要素を一つでも含む列を削除することになる。すると上の例の観測値行列の場合、以下のような行列に変形される。

$$\begin{pmatrix} 122 & 141 & 258 \\ 120 & 142 & 259 \\ 119 & 144 & 260 \\ 118 & 146 & 262 \\ 118 & 148 & 261 \\ 116 & 149 & 259 \end{pmatrix}$$

このように、対応付けのされていないシーンを扱うと、あらかじめ3次元情報が求まる点の数が限られてしまう。

本稿では、この問題に対して焦点を当てる。

4 解決手法

この節では、前節にて指摘した問題点を解決する具体的なアルゴリズムを示す。

このアルゴリズムの特徴は、すべてのフレームで対応付けられていない特徴点であっても3つ以上のフレームで対応付けされていれば、その3次元位置やの算出に用いることができるように改良したものである。3フレーム以上で対応付けされている必要があるのは、従来の因子分解法での制約と同じ理由からで、算出の途中で行なう特異値分解において分解される行列のランクが3以上である必要があるからである。

具体的には、一部のフレームでしか対応付けされていない特徴点をも含むように、観測値行列を分割して部分行列を作り、それに対して因子分解法を適用する。よって本手法の欠点としては、各々の部分行列のサイズが元の観測値行列よりも小さいことにより、従来の因子分解法に比べて計算誤差が生じる可能性が高いことが考えられる。

以下、前節で例として用いた観測値行列 (11) を例にとり、詳細を説明する。

1. 3フレーム以上に渡り対応付けが取れている特徴点の座標を用いて、観測値行列 W を作る。この行列は $2f$ 行 n 列であるものとする。行列 (11) の場合には、2フレームでしか対応付けされていない3番目の特徴点を除去することで次のようになる。

$$W = \begin{pmatrix} 106 & 122 & 141 & -1 & 258 \\ 107 & 120 & 142 & -1 & 259 \\ 108 & 119 & 144 & 433 & 260 \\ 110 & 118 & 146 & 436 & 262 \\ 110 & 118 & 148 & 437 & 261 \\ -1 & 116 & 149 & 439 & 259 \end{pmatrix} \quad (12)$$

2. 観測値行列 W の1行目から $f-2$ 行目まで各行ごとに以下の作業を繰り返し、行列 W の分割を行なう。これと同じ分割が $f+1$ 行目から $2f-2$ 行目に対しても適用される。

- (a) その行 (i 行目とする) の表わすフレームから初めて観測され始めた特徴点を見付ける。そのような点が複数あったら、対応する列ベクトルの内、長さが3以上で最も長さの短いものを捜す。そのような点がなければ、次の行に移る。
- (b) その最短の列ベクトルの要素と同じフレームにて観測されている点の要素を用いて、 W の部分行列 W'_i を作る。ここで、どの点を選ばれたかを記録しておく。この情報は、次の段階において、各フレーム毎に計算された形状を統合するのに用いられる。
- (c) W'_i から、因子分解法を用いて R'_i, S'_i, T'_i を計算する。

観測値行列 (12) に対しては、

$$W'_1 = \begin{pmatrix} 106 & 122 & 141 & 258 \\ 107 & 120 & 142 & 259 \\ 108 & 119 & 144 & 260 \\ 110 & 118 & 146 & 262 \\ 110 & 118 & 148 & 261 \end{pmatrix}$$

$$W'_2 = \begin{pmatrix} 107 & 120 & 142 & 259 \\ 108 & 119 & 144 & 260 \\ 110 & 118 & 146 & 262 \\ 110 & 118 & 148 & 261 \end{pmatrix}$$

$$W'_3 = \begin{pmatrix} 108 & 119 & 144 & 433 & 260 \\ 110 & 118 & 146 & 436 & 262 \\ 110 & 118 & 148 & 437 & 261 \end{pmatrix}$$

$$W'_4 = \begin{pmatrix} 118 & 146 & 436 & 262 \\ 118 & 148 & 437 & 261 \\ 116 & 149 & 439 & 259 \end{pmatrix}$$

となる。ここで W'_2 は、 W'_1 と同じ特徴点で構成されかつ W'_1 より少ないフレーム数しか含まない。このような場合には、 W'_2 は W'_1 と比較して正確さにおいて同じかそれより劣る結果しかもたらさない。よって、このような部分行列は以後の計算

から除外される。

このような行列の分割によって、特徴点と部分行列の対応表 1 が作られる。

表 1: 部分行列における特徴点の位置

W	P_1	P_2	P_3	P_4	P_5
W'_1	1	2	3	×	4
W'_3	1	2	3	4	5
W'_4	×	1	2	3	4

3. S'_i は、 W'_i に含まれる特徴点の 3 次元位置を表わす行列である。しかし、その座標系は特徴点の重心を原点としたものであるので、各 S'_i の座標系は異なる。

そこで、以下のようにして適当な座標系に統一する。ここでは簡単のために、 S'_1 の座標系に統一する方法を述べる。まず、 S'_1 と S'_2 で共通する特徴点を選び、その間の座標変換のパラメータを求める。すなわち、 S'_1 と S'_2 の両方に含まれる点から成るそれぞれの部分行列を S''_1, S''_2 と記す時、 $S''_1 = AS''_2 + B \cdot \mathbf{1}$ となるような座標変換行列 A と B を以下のようにして求める。

- S''_1 と S''_2 のそれぞれの重心 g''_1 と g''_2 を求める。
- $S''_1 - g''_1 \cdot \mathbf{1} = A(S''_2 - g''_2 \cdot \mathbf{1})$ となる回転行列 A を最小二乗法を用いて求める。[5][6]
- A が求まれば B は g''_1, g''_2 を用いて、 $B = (g''_1 - Ag''_2) \cdot \mathbf{1}$ となる。

この A と B により S'_2 のすべての点を S'_1 の座標系に変換し、それらの座標値から成る行列 S' を作る。それ以降の S'_i ($i \geq 3$) に対しては、上述と同じ手順（共通な点から成る部分行列間の変換行列を求めることで、次々と座標変換を行なってゆくこと）により S' に情報を統合してゆく。このようにして、すべての S'_i の情報を統合した結果得られる S が、最初の観測値行列 W に現われたすべての特徴点の 3 次元位置を表わす行列である。

5 実験結果と考察

5.1 実験内容と結果

実験に用いた原画像は、1 台のミニチュアの模型車両が左から右へ進行しているシーンを CCD カメラで撮影したものである。実験では、15 フレームを用いた。それぞれのフレームの画像のサイズは 512 ピクセル $\times$ 480 ピクセルで、1 ピクセルあたり 256 階調の濃淡画像である。このシーケンシャル画像のうちの最初のフレームを図 1 に、最後のフレームを図 2 に示す。

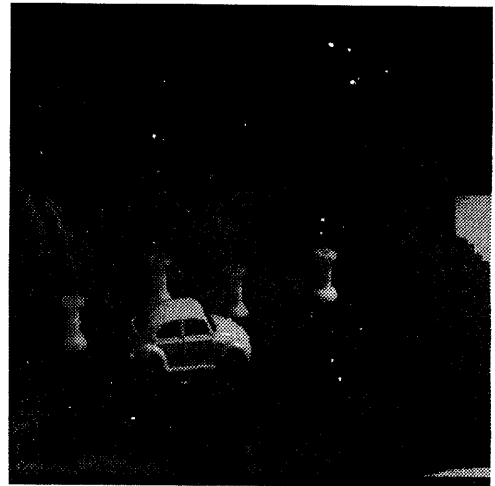


図 1 原画像の最初のフレーム

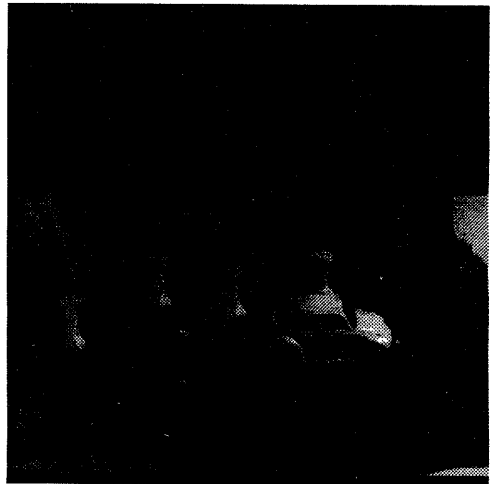


図 2 原画像の最後のフレーム

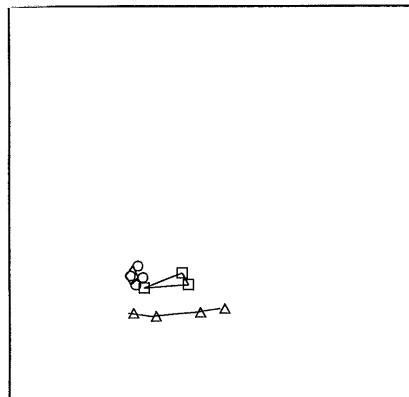


図 3 特徴点の位置

これらの画像に対して車両上の計11点の特徴点の対応付けをマニュアルにて行ない、すべてのフレームにおいて点の対応付けがされている観測値行列 W を作成した。図3に、特徴点の位置を示す。△で示されている4点は車両の下縁の部分であり、□で示されている3点は側面の窓ガラスから、○で示される4点は後部の窓ガラスからそれぞれ取っている。原画像中の物体の形状と対応が取れやすいようにそれぞれの特徴点を線で結んである。

この W から、因子分解法により S を求めた。次に、 W のうち数個の点を任意に選び、その点に相当する要素が欠けている（対応付けされていないことに相当する）行列 W^* を作った。今回提案した改良手法により、 W^* から S^* を求めた。

さらに W^* のうち欠けた要素を含む列を除去した行列 W^{**} を作成し、因子分解法により S^{**} を求めた。これは、隠れのあるシーンを因子分解法により処理したことに相当する。ところで、 W^* からは13個の部分行列 W_i^* , $i = 1, \dots, 13$ が作成されたが、そのうち因子分解法で S_i^* が求まったものは6つであった。また、 S_i^* を座標変換して S^* へ統合する段階において、基準となる座標系として S と同じ座標系を取っている。

このようにして求まった形状行列において示されている各特徴点の座標をプロットしたものが図4と図5である。図4は、原画像と同じように見えるように3次元座標の値を変換し直した上でプロットしたもので、図5は車両をおよそ真上から見た場合に相当するようにプロットしたものである。ただし、比較を容易にするため、両方の図の大きさは原画像に比べ拡大されている。また、○は S を、□が S^* を、△が S^{**} を示している。

5.2 結果の考察

まず、図1の原画像と図4を比較する。 S と S^{**} は概して良好な結果が求まっていると言える。ただ、後部窓ガラスの4点のうち奥にある2点は、ややかけ離れた値を示した。とりわけ、原画像から見て奥行き方向の座標値を見るとこのことがはっきりする。この原因は、これらの点についてはフレームによっては対応点の座標がはっきりしない場合も多いのに、無理に対応付けしたことにあると思われる。

次に、実験で求まった S^* を S^{**} とを比較する。まず、3次元情報が求まった特徴点の数は S^* の方が多い。これにより、本稿で提案した手法の長所が確認されたことになる。

さらに、 S に対してどの程度近い値が求まっているかという観点で比較する。 S^{**} の4点の座標値は S の中の対応する点の座標値とかなり近い値を示しているのに対し、 S^* の10点の座標値はあまり正確な値ではない。

図4においては、とりわけ右側の4点は S の位置からかなり離れてしまっている。これは、これら4点を含む S_i^* が少なく、また、その S_i^* を求めるために使われた W_i^* のサイズも小さかったため、計算誤差が出てし

まったものと考えられる。同様のことが、図5においても言える。それ以外の6点については S に近い座標値が求まっている。しかし、一つ一つの S_i^* の値を検討すると、 S の周りになんばらつきが生じていた。すなわち、 S_i^* の統合の際に平均化されて誤差が打ち消されていることにより、 S に近い座標値が求まっていることになる。

ちなみに、 S^{**} が求まっている点において比較すると、 S_i^* のばらつきの方が S^{**} のばらつきよりも大きい。これは、 S^{**} が15フレーム全体に渡る対応から算出された値であることを考えれば、当然のことと思われる。

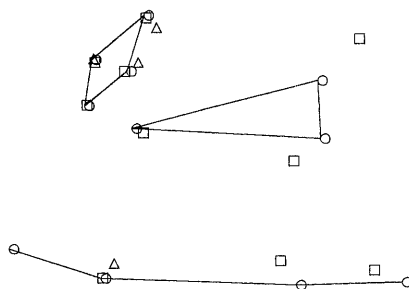


図4 算出された特徴点の位置：原画像と同視点から見たもの ○: S , □: S^* , △: S^{**}

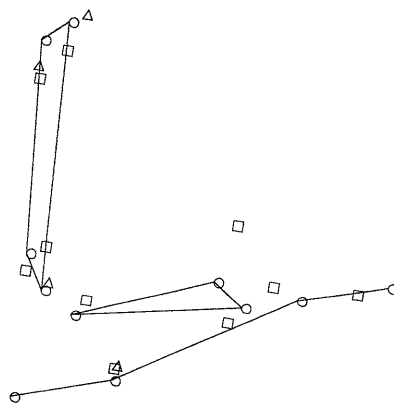


図5 算出された特徴点の位置：車両の上方から見たもの ○: S , □: S^* , △: S^{**}

6 おわりに

本稿では、オクルージョン等の原因により発生する、一部のフレーム間で特徴点の対応が取れない場合にでも、対応の取れているフレームだけの情報からその点の3次元情報が求まるように従来手法を改良した。

この手法を適用した実験により、より多くの点の3次元情報が取れることを確認した。それと同時に、観測値行列を分割することにより部分行列が小さくなりがちなため、正確な座標値が算出できなくなる現象も起きることも実験により示された。

今後は、この問題点の改善につとめたい。とりわけ、

部分行列への分割の仕方には工夫する余地があると思われる。そのためには、さまざまな種類の画像データに対して適用することで、どの程度のフレーム数・特徴点数があればどの程度正確な座標値が算出されるかを定量的に把握する必要がある。

さらに、因子分解法により算出された結果の確からしさを示す指標を検討したい。これが定義可能になれば、各部分行列から得られた結果を統合する際において、確からしさを考慮した重み付けを行なえると考ええる。

謝辞

本研究を進めるにあたり、貴重なアドバイスをいただいたマサチューセッツ工科大学のアレックス・ペントランド教授に、この場をお借りして感謝いたします。

参考文献

- [1] D. W. Murray, B.F. Buxton: *Experiments in the Machine Interpretation of Visual Motion*, MIT Press, 1990
- [2] C. Tomasi, T. Kanade: "Factoring Image Sequences into Shape and Motion", *IEEE Workshop on Visual Motion Proceeding*, pp.21-28, 1991
- [3] T. E. Boulton, L. G. Brown: "Factorization-based Segmentation of Motions", *IEEE Workshop on Visual Motion Proceeding*, pp. 179-186, 1991
- [4] 出口光一郎 「画像と空間 — コンピュータビジョンの幾何学 —」, 昭晃堂, 1991
- [5] K. S. Arun, T. S. Huang, S. D. Blostein: "Least-Squares Fitting of Two 3-D Points Sets", *IEEE Trans. PAMI*, Vol. 9, No. 5, Sep. 1987, pp. 698-700
- [6] 金谷健一 「画像理解 — 3次元認識の数理 —」, 森北出版, 1990
- [7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, 1988