

## EPI 解析におけるカメラ運動の制限緩和

南田 幸紀<sup>†</sup>      新谷 幹夫<sup>†</sup>      納富 幹人<sup>††</sup>

An Image Stabilization Method for Epipolar Plane Image Analysis

Kouki MINAMIDA<sup>†</sup>, Mikio SHINYA<sup>†</sup>, and Mikito NOTOMI<sup>††</sup>

あらまし 近年、コンピュータグラフィックス (CG) コンテンツ制作の需要が高まっているが、複雑な形状の CG モデル作成は大変人手のかかる作業である。そこで、写真やビデオなどから自動的に CG モデルを抽出する画像ベースモデリング手法が注目を集めている。この中で、エピポーラ画像 (EPI) 解析はビデオ画像から 3 次元形状を抽出する手法として有力であるが、カメラの運動に厳しい制限があり、実施が困難であった。本論文では、このカメラ運動の制限を緩和する手法を提案する。本手法では、入力画像系列中の特徴点軌跡からカメラの変動を推定し、カメラが直線運動したときに得られるはずの画像系列を作成する。また、移動速度を推定することで、カメラの速度変化の影響を吸収する。以上の手順を繰り返し適用することで、モデル抽出の精度を向上させる。提案手法に基づいて実景観で実験を行い、良好な結果を得た。

キーワード エピポーラ画像解析、画像ベースモデリング、カメラキャリブレーション

### 1. ま え が き

近年、映画やゲーム、コマーシャルフィルムなどでコンピュータグラフィックス (CG) を頻繁に目にするようになってきた。しかしながら、3 次元 CG の素材となる 3D モデルの作成は、大変手間のかかる職人的作業であり、この CG モデル作成コスト削減が課題となっている。そこで、実在する物体については、写真やビデオ映像から自動的にモデルを作成する、画像ベースモデリング手法が注目されている [10]。

エピポーラ画像 (Epipolar Plane Image, EPI) 解析 [2] は、ビデオ映像から被写体の 3 次元形状を抽出する手法である。エピポーラ拘束が画面系列を通じて一定であるため、対応点探索、特徴点間の接続の決定が容易という長所がある [8], [9]。しかし一方で、カメラが光軸に垂直に等速直線運動することが前提とされている。このような撮影条件を満たすには、モーションコントロールカメラ等の特殊な機材を用いなければ

ならず、実用化にあたり大きな障害となっていた。

本論文では、カメラ運動の制限を緩和するため、

- 画像系列からカメラの位置、姿勢の変位を推定する、
  - 特徴点の奥行、カメラの移動速度を推定する、
  - 上の結果を用いて画像を変形し、直線運動からの変動により発生するずれを補正する、
- という戦略をとり [4], [5]、具体的アルゴリズムを提案するとともに、有効性を実証的に示す。

なお、同様のアプローチは筆者らとは独立に Zhu [13] らが提案しているが、後述のように、カメラ移動速度の扱いや位置の推定処理において提案手法に優位性がある。

### 2. EPI 解 析

本章では、EPI 解析の基本原理を記述する。図 1 に示すように、カメラが光軸に垂直かつ CCD に対し水平方向に等速運動し、静止物体を撮影しているとする。入力画像列を  $i(x, y; t)$  とする。 $y = y_0$  でのスライス

$$I(x, t) = i(x, y_0; t)$$

をエピポーラ画像 (EPI) と呼ぶ。このとき、物体上の 1 点の像点の軌跡は同一 EPI 上の線分となり、その傾きは物体点の奥行に比例する。したがって、EPI

<sup>†</sup> 日本電信電話株式会社 NTT サイバーソリューション研究所, 横須賀市

NTT Cyber Solutions Labs., NTT Corporation, Yokosuka-shi, 239-0847 Japan

<sup>††</sup> 日本電信電話株式会社 NTT 通信機器事業部, 東京都

NTT Customer Premises Equipment Business Division, NTT Corporation, Tokyo, 163-1441 Japan

上で直線を抽出し、その傾きを求めれば、対応する特徴点の 3 次元座標が得られる。

さて、ここでのカメラ運動の制約をまとめると、

- 速度の一定性
- 姿勢の一定性
- 位置の直線性

である。

この中で、等速性の制限緩和は特に実用上重要である。というのは、カメラの姿勢を固定し、平行移動させることは、レールを敷設しカメラを動かすことで実現できる。レールの敷設は商用撮影では珍しいことではなく、コストもさほどかさまない。しかし、レールを敷設しても等速性を実現することは極めて困難である。

姿勢の変位は、変位が既知であれば、平行化したり [1]、ガウス球を用いる [12] などの手法がとれるが、未知の場合は推定する必要がある。ジャイロや GPS 等の機材を使う試みもあるが [6]、現状では精度的に不

十分であり、画像列から推定することが必要であると考えられる。

位置の非直線性は、姿勢の変動より扱いが難しい。位置の変動による画像点の変位は、対応する物点の奥行に依存するからである。すなわち、単純に平行化するわけにはいかない。このため、カメラの位置のみならず、物体点の奥行も同時に推定することが必要となる。

### 3. 速度変化の許容

速度変化は EPI の見方を少し変えるだけで、容易に対処することが可能である。時刻  $t$  におけるカメラの位置を  $x_e(t)$  とする。これを用いて、

$$I'(x, x_e(t)) = I(x, t) \tag{1}$$

と新たに  $I'(x, x_e)$  を定義すると、移動速度の如何を問わずに  $I'$  において像点の軌跡は直線となり、傾きは奥行に比例する。 $I, I'$  はともに離散画像であるが、 $I$  は等間隔のサンプルで、 $I'$  は  $x_e$  に関し不等間隔のサンプルとなる (図 2)。

カメラの位置  $x_e(t)$  は、例えば、以下のように求めることができる。コーナ点などの特徴点を追跡し、 $x_i(t)$  とする。点群  $(x_i(t), x_e(t))$  は  $i$  に対して決まる直線上に乗るので、これを利用して  $x_e(t)$  を決める。すなわち、最適な  $x_e(t)$  の組を求めるために、以下の線形最小 2 乗法を各  $t$  に対して用いた。

$$\min_{x_e(t)} \sum_i |L_i x_e(t) + x_i(0) - x_i(t)|^2, \tag{2}$$

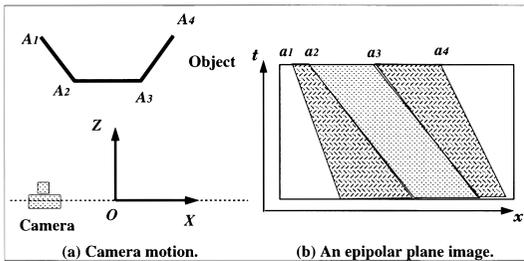


図 1 EPI 解析  
Fig. 1 Epipolar plane image analysis.

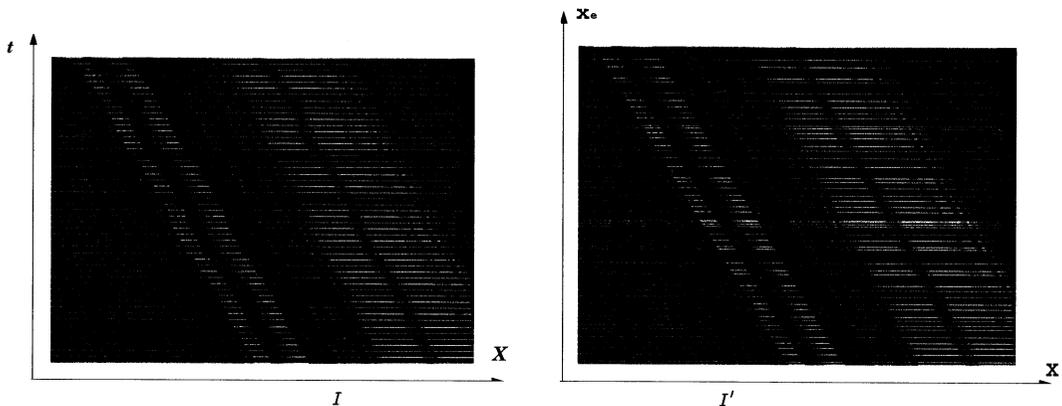


図 2 速度変化の吸収  
Fig. 2 Compensation of camera speed variation.

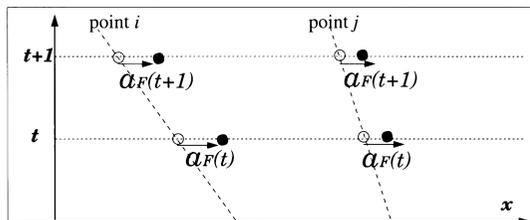


図 3 Zhu らの速度変化吸収

Fig. 3 Zhu's compensation method of camera speed variation.

$$L_i = x_i(T - 1) - x_i(0).$$

ただし、 $t = 0$  は最初のフレームを、 $t = T - 1$  は最後のフレームを表す．また簡単のため  $x_e(T - 1) = 1$  とおいた．式 (2) から、

$$x_e(t) = \sum_i L_i(x_i(t) - x_i(0)) / \sum_i L_i^2 \quad (3)$$

を得る．

直線  $i$  の傾き  $L_i$  から対応する点の奥行  $z_i$  が得られる． $z_i$  は、後述する繰返し手法で改善することができる．

本手法と Zhu らの手法 [13] との比較を以下に述べる．Zhu らは、隣接するフレーム間の像点  $i$  の移動量  $a_i(t)$  の平均  $\langle a_i \rangle_i(t)$  を一定値  $a_S$  で近似し、カメラの等速直線運動による像点の移動量の平均  $a_S$  と、カメラの速度変化による像点の移動量の平均  $a_F(t) = \langle a_i \rangle_i(t) - a_S$  を求めている．我々の手法が  $I(x, t)$  を  $t$  方向に動かそうとする (図 2) のに対し、この手法では  $x$  方向に動かそうとしている (図 3)．図 3 に示すように、本来、像点  $i, j$  のずれは、物点の奥行により異なるので、点について平均した  $a_F(t)$  で速度変化を吸収するのは理論的に無理がある．一方、これまでに見たように、提案手法ではこのように不必要な制約はない．また、計算量などのペナルティもないので、優位性をもっているといえる．

#### 4. 位置・姿勢変動の許容

本章では、カメラ運動の非直線性、姿勢の変動を許容するための処理を議論する．まず、これらの変動による像点の変位を求める．次いで、像点の変位からカメラの位置、姿勢を推定する．やっかいな点は、像点の変位が物点の奥行にも依存する点である．そこで、奥行値とカメラ位置、姿勢を繰返し処理により求める

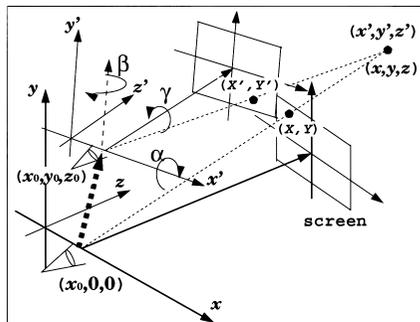


図 4 ピンホールカメラモデル

Fig. 4 Pin-hole camera model.

手法を提案する．

##### 4.1 カメラ変動と像点の変位

本手法ではカメラモデルとしてピンホールカメラモデル (図 4) を用いる．カメラの光軸を  $z$  軸とし、視点  $x$  軸上の  $x = x_0$  の位置にあるとすると、透視変換により、物体点  $(x, y, z)$  の像点  $(X, Y)$  は、

$$X = a(x - x_0)/z \quad (4)$$

$$Y = ay/z, \quad (5)$$

となる． $a$  はカメラ内部パラメータであり、画角によって決まる．

カメラが視点を中心とし  $x, y, z$  軸方向周りに角度  $\alpha, \beta, \gamma$  回転し、カメラ視点  $(x_0, y_0, z_0)$  になるように平行移動したとすると、像点は  $(X', Y')$  に変移する．

$$X' = ax'/z',$$

$$Y' = ay'/z',$$

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = R_x(-\alpha)R_y(-\beta)R_z(-\gamma) \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix}$$

ただし、 $R_*$  は、 $*$  軸周りの回転行列を表す． $\alpha, \beta, \gamma, y_0$  及び  $z_0$  が小さいとき、次の近似式が成り立つ．

$$X' \cong X + \alpha XY/a - \beta(a + X^2/a) - \gamma Y + (z_0/z)X - ax_0/z \quad (6)$$

$$Y' \cong Y + \alpha(a + Y^2/a) - \beta XY/a + \gamma X + (z_0/z)Y - ay_0/z. \quad (7)$$

式 (6) の右辺第 5, 6 項及び式 (7) の右辺第 5, 6 項から、像点の変移は物体点の奥行  $z$  に依存することがわかる．

## 4.2 推定処理

式 (5) からわかるように、像点の  $Y$  座標はカメラ位置  $x_0$  に依存しないので、理想的には、像点の軌跡はあるエピポーラ画像  $I(X, x_0; Y)$  上に乗る。しかしながら、カメラの運動が変動する場合はこの条件は満たされず、エピポーラ画像解析を困難にしていた。そこで、画像中の特徴点の軌跡からカメラの変動を推定し、これをもとに、入力画像をカメラが理想的な動きをした場合に得られるはずの画像列に変形する。変形後の画像列では、像点の軌跡は  $Y = \text{一定}$  なる直線上に乗るようになる。

画像中の特徴点  $i$  の軌跡を  $(X_{i,t}, Y_{i,t})$  と表記する。カメラを  $x, y, z$  軸周りに微小な角度  $\alpha, \beta, \gamma$  回転し、カメラ視点を微量量  $(0, \delta_y, \delta_z)$  だけ平行移動させると、式 (6), (7) より、特徴点の像点は、

$$X'_{i,t} \cong X_{i,t} + \alpha X_{i,t} Y_{i,t} / a - \beta(a + X_{i,t}^2 / a) - \gamma Y_{i,t} + (\delta_z / z_{i,t}) X_{i,t}, \quad (8)$$

$$Y'_{i,t} \cong Y_{i,t} + \alpha(a + Y_{i,t}^2 / a) - \beta X_{i,t} Y_{i,t} / a + \gamma X_{i,t} + (\delta_z / z_{i,t}) Y_{i,t} - a \delta_y / z_{i,t}, \quad (9)$$

へ移動する。ここで、パラメータ  $\alpha(t), \beta(t), \gamma(t), \delta_y(t), \delta_z(t)$  を理想的カメラパスからの変移であるとすれば、移動した点の座標  $(X'_{i,t}, Y'_{i,t})$  は、理想的カメラパスから撮影した場合の像点の座標である。ここで、 $t$  によらず  $Y'_{i,t} = \text{一定}$  を満たすような  $\alpha(t), \beta(t), \gamma(t), \delta_y(t), \delta_z(t)$  を求めれば、カメラの変動を得ることができる。

カメラの平行移動による像点の動きは物体点の奥行  $z_{i,t}$  に依存するが、最初はこの値は未知であるので、 $z_{i,t} = \text{一定}$  と近似し、 $\delta_y/z, \delta_z/z$  を未知数とみなすことにより、近似解を求める。パラメータ推定には、最小 2 乗法を用いる。

$$\min \sum_i |Y'_{i,t} - \bar{Y}'_i|^2. \quad (10)$$

なお、筆者らは、 $\bar{Y}'_i = Y_{i,0}$  と置き、式 (10) を線形化した。

入力画像  $i(X, Y; t)$  の各点  $(X, Y)$  を式 (8), (9) により点  $(X', Y')$  に移せば、補正後の画像  $i'$  を得る。

$$i'(X', Y'; t) = i(X, Y; t) \quad (11)$$

$i'$  上では、特徴点が走査線方向に移動するようになり、エピポーラ画像解析を容易に適用することができる。

## 4.3 繰返し処理

いったん被写体のモデルが得られたら、視点を撮影時と同じカメラパスに置いてレンダリングすることで、入力画像と同じ画像列及び奥行マップを作成することができる。つまり、入力画像中の全可視点の奥行を得ることができる。この奥行を式 (8), (9) の  $z_{i,t}$  に代入することにより、最初の近似解より精度良くカメラパラメータ変動を推定する。

なお、モデルの抽出には、EPI ベースの手法で、物体の形状とテクスチャを復元する手法である Strip-DP 法 [8], [9] を用いた。

この処理を繰り返すことによって、カメラ変動パラメータ推定とモデル作成の精度を上げることができる。これは、文献 [13] の手法にない特徴である。

## 5. 実装

筆者らの実装のアウトラインを図 5 に示す。

a) Input image sequence

ビデオカメラを光軸とほぼ垂直方向に、速度一定で

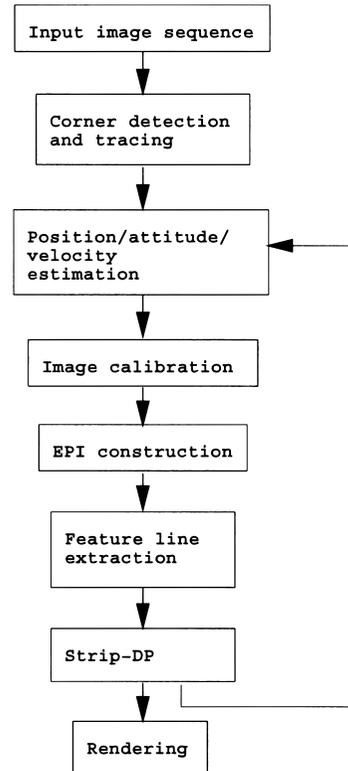


図 5 本アルゴリズムの概要  
Fig. 5 The outline of the algorithm.

動かすように配慮して、静止シーンを撮影する。このとき、カメラを厳密にコントロールする必要はない。撮影した画像列をデジタル化し計算機に取り込む。

#### b) Corner detection and tracking

入力した時系列画像上で特徴点検出を行い、時間軸に沿って追跡する。ここでは比較的ラバストに検出可能な局所の特徴であるコーナを特徴点とする。コーナ検出アルゴリズムとしては Plessey アルゴリズム [3] を用いる。コーナ検出後、コーナ近傍で correlation matching を行い、同一点を追跡する [11]。

#### c) Position/attitude/velocity estimation

特徴点の軌跡から、式 (10) によってカメラの姿勢、位置の変動を推定する。コーナの奥行も推定する。

また、特徴点の座標を式 (8), (9) に基づいて変換した後、式 (3) によって  $x$  軸上のカメラ位置  $x_e$  を推定する。

#### d) Image calibration

各入力画像を式 (11) に基づいて変形する。変形後の画像  $i'$  は、カメラが理想的カメラパス上を動きながら撮影された画像とみなせる。最初のループでは、被写体の平均的な奥行を用いて変形する。2 回目以降のループでは、後述する Strip-DP ステップで作成された奥行を用いて画像変形の精度を上げる。

#### e) EPI construction

補正後の画像からエピソード画像  $I$  を作成する。

$$I(x, t; y) = i'(x, y; t)$$

#### f) Feature line extraction

EPI ごとに特徴線検出を行う。物体点が多数であるため、Hough 変換を行ってもシャープなピークが得られない。このため、ピーク探索誤りによる誤検出が多くなる。そこで筆者らは、dynamic programming ベースの方法を用いてエッジを追跡する方法を採用した。また、途切れたエッジからも直線を検出できるように、追跡されなかったエッジに対して Hough 変換を適用する 2 段階の処理として実装した。

(1)  $x$  軸方向に 1 次元微分フィルタをかけ、エッジを抽出する

(2) DP を用いてエッジを時間軸方向に追跡する [7]。

(3) 追跡された点列に対して最小 2 乗法を適用し、直線を抽出する。

(4) 追跡処理に漏れたエッジに対して Hough 変換を適用し、直線を抽出する。

上のステップ (3) (4) においては、時刻の代わりにカメラ位置  $x_e$  を用いて速度変化の影響を吸収する。

#### g) Strip-DP

抽出された特徴線間の最適な接続関係とテクスチャを求める。接続された特徴線は、3 次元の線分としてモデル化される。

ここで最適という意味は、抽出されたモデルを使って再生した画像列と、入力画像列との差が最小であるということである。最適な接続関係はオクルージョンを回避しながら手前から奥の被写体へ、繰り返し dynamic programming を適用することによって安定的かつ大局的に求めることができる。また、この段階で不必要な特徴線 (ノイズ) も除去される。詳細は [8], [9] を参照されたい。

#### h) Rendering

抽出された線分による 3D モデルを使って、ユーザの求める任意視点画像を生成する。モデル抽出処理はスキャンラインごとに独立して行われるため、スキャンライン間の接続関係は復元されないが、描画処理において、その間を塗りつぶし、すき間のない画像を生成する。同時に Z バッファ処理を行うことによって、奥行画像も生成される。

#### i) Iterative refinement

入力画像系列の各画像の奥行をステップ c) に入力し、カメラ変動を再推定し、続く処理を再実行する。この繰返し処理でモデルの抽出の精度を向上させる。

## 6. 実 験

### 6.1 非等速運動の補正

本速度補正手法の効果を評価するため、定量実験を行った。実験のために、次のような速度変動を含むカメラパラメータ系列を生成した。

$$v(t) = 1.0 + a \sin(2\pi t / (T - 1)).$$

これを用いて図 11 と同様な CG 画像系列を生成した。sin の項は、速度変動をフーリエ級数展開したときの 1 次の項であり、速度の変動のモデルとして適当であると考えた。この画像系列に対して提案手法を適用し、本来の速度と本手法により得られた補正後の速度の差の絶対値平均

$$v_{err} = 1/N(T-1) \sum_{t=1}^{T-1} \left\{ \sum_i |1.0 - v_i(t)/L_i| \right\}$$

$$v_i(t) = (x_i(t) - x_i(0)) / (x_e(t) - x_e(0))$$

を評価した。  $N$  は特徴点数とする。  $a$  と  $v_{err}$  の関係を図 6 に示す。速度変動は 0.01 程度にまで低減されており、本アルゴリズムの有効性を示している。本実験での主な誤差の要因は、特徴点座標の量子化誤差と考えられる。  $a = 0.7$  の場合の  $x-t$  平面での特徴点軌跡を図 7 に示す。

次いで、屋外の景観を対象にした実験を行った。天守府天満宮の御協力のもと、天満宮の敷地内にレールを敷設し、35 mm フィルムで撮影を行った。本実験では、非等速運動の補正の効果を実証するため、レールによって、できるだけカメラの運動を直線上に制限した。ただし、地面の凹凸のため、完全な直線とはならない。

フィルムを  $2048 \times 1744$  24 bit カラーでデジタル化し、オリジナルの解像度の画像と各辺を  $1/3$  に縮小したものを作成し、それぞれを入力として、提案手法を適用した。その結果を図 8 に示す。処理には、SGI

社製ワークステーション 4 台を用いて、図 8(b) の場合で約 3 時間を要した。

再生画像は写真のようなリアリティがあり、良好な品質であるといえる。また、復元形状を示すために、奥行を濃淡値に置き換えた画像を付す(図 8(a') 及び(b'))。

### 6.2 姿勢・位置の補正

カメラ位置・姿勢の変動補正を検証する実験を行った(図 9)。まず、CG で合成した画像を入力とし、カメラ姿勢推定の精度を定量的に評価した。図 10 中で、 $\alpha, \beta, \gamma, \delta_y, \delta_z$  のグラフは、CG 作成時に与えたカメラの変動(真値)と、提案手法によって推定された値を比較している。また、 $N$  のグラフは、カメラ姿勢推定に用いられた特徴点の数を表している。カメラが移動するにつれ、特徴点が視野から消えていくため、特徴点数は時間が経つに従い減少する。このことによって変動推定の精度が低下する傾向がある。

次に、カメラ位置・姿勢の誤差が画像の補正に与える影響の評価を述べる。図 11(a) に示す条件と式(8)、(9)を用いると、画像補正の誤差は、 $\alpha, \beta, \gamma$  が各々 0.2 度の誤差を含む場合、

$$\begin{aligned} |X'_{i,t} - X_{i,t}| &\leq |\alpha X_{i,t} Y_{i,t}/a| + |\beta(a + X_{i,t}^2/a)| + |\gamma Y_{i,t}| \\ &\cong 2.3 [\text{画素}] \\ |Y'_{i,t} - Y_{i,t}| &\leq |\alpha(a + Y_{i,t}^2/a)| + |\beta X_{i,t} Y_{i,t}/a| + |\gamma X_{i,t}| \\ &\cong 2.3 [\text{画素}], \end{aligned}$$

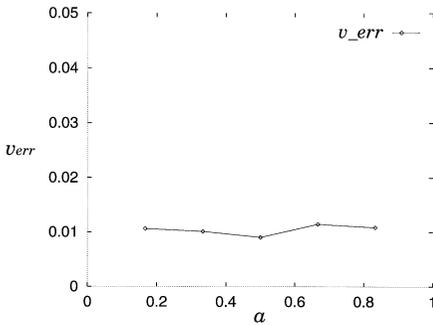


図 6 カメラ速度変動補正の評価

Fig. 6 The evaluation of compensation of camera speed variation.

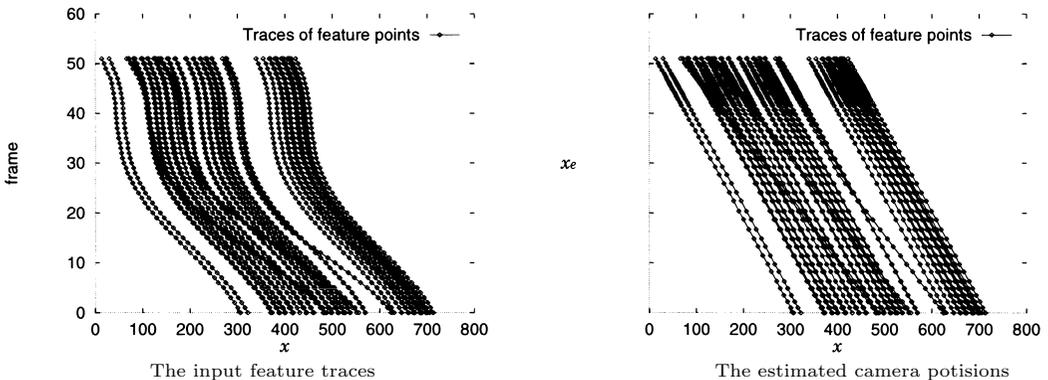


図 7 速度変化吸収の実験

Fig. 7 The experiment for compensation of camera speed variation.

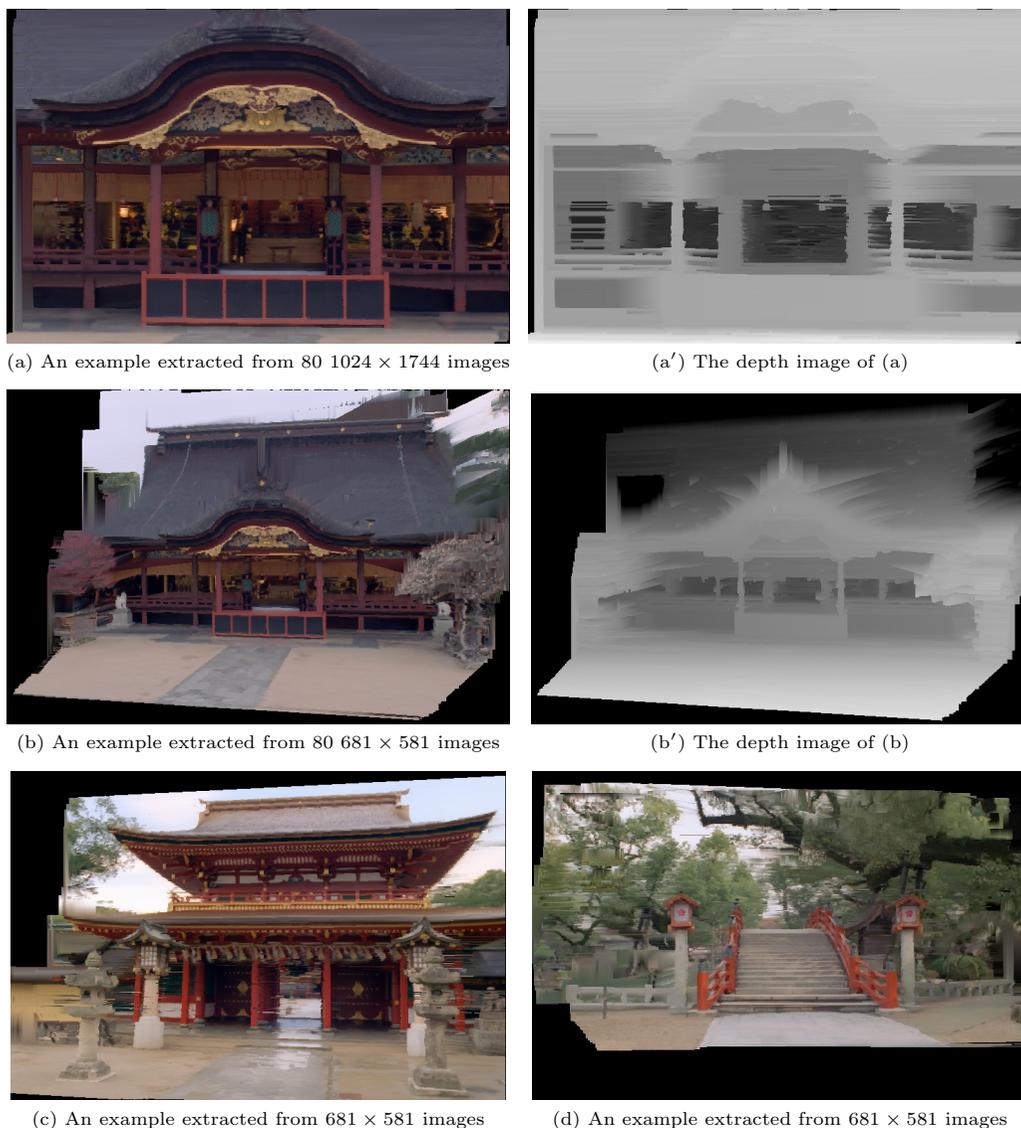


図 8 3-D モデル抽出結果  
Fig. 8 Extracted 3-D models.

また、 $\delta_y, \delta_z$  が 0.1 の誤差を含む場合、

$$\begin{aligned}
 |X'_{i,t} - X_{i,t}| &\leq |(\delta_z/z_{i,t})X_{i,t}| \cong 1.2 \text{ [画素]} \\
 |Y'_{i,t} - Y_{i,t}| &\leq |(\delta_z/z_{i,t})Y_{i,t}| + |a\delta_y/z_{i,t}| \cong 2.4 \text{ [画素]}
 \end{aligned}$$

で抑えられる。実際には  $\pm 1$  画素程度の誤差で補正されており、図 10 の 80 フレーム目までの誤差程度であれば EPI 解析を適用可能である。

図 11 に、既知物体に対する形状復元の定量評価の結果を示す。図 11 (d) は、抽出したモデルの奥行画像を作成し、正しい奥行との差を濃淡値に置き換えた画像である。横線状の誤差の大きい部分は Strip-DP において接続が正しく復元されていないために生じている。このような特に大きく外れている部分があるものの、奥行の差を画像全体で平均すると 0.34、物体の最前面までの距離の 4% に収まる。x, y 方向の形状の誤差は、画像上で 2 画素程度、世界座標で 0.04 程度で

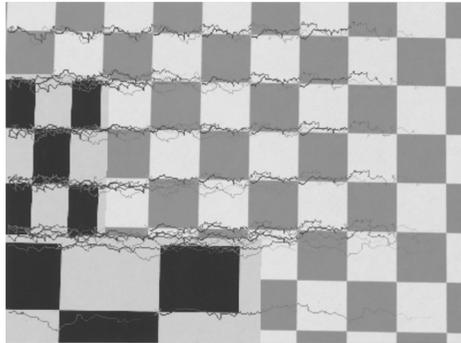
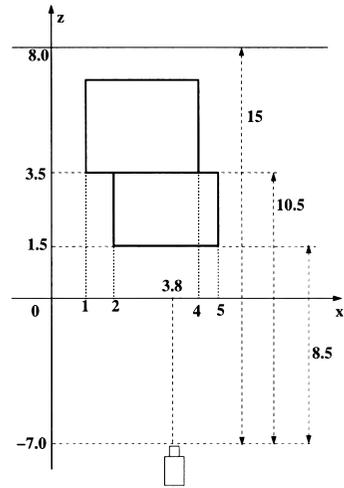
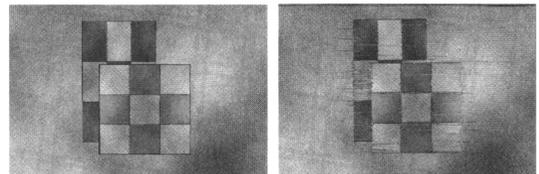


図 9 CG シーケンスより、図中の線は特徴点の軌跡を表す  
 Fig. 9 A source image of a CG sequence. Lines indicate the traces of corner points.

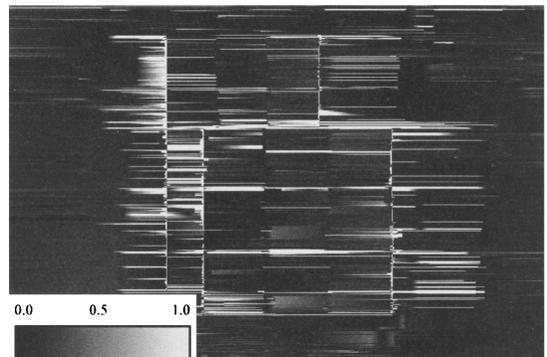


(a) The definition of the CG scene



(b) An input image

(c) A synthesized image



(d) The difference of the shape

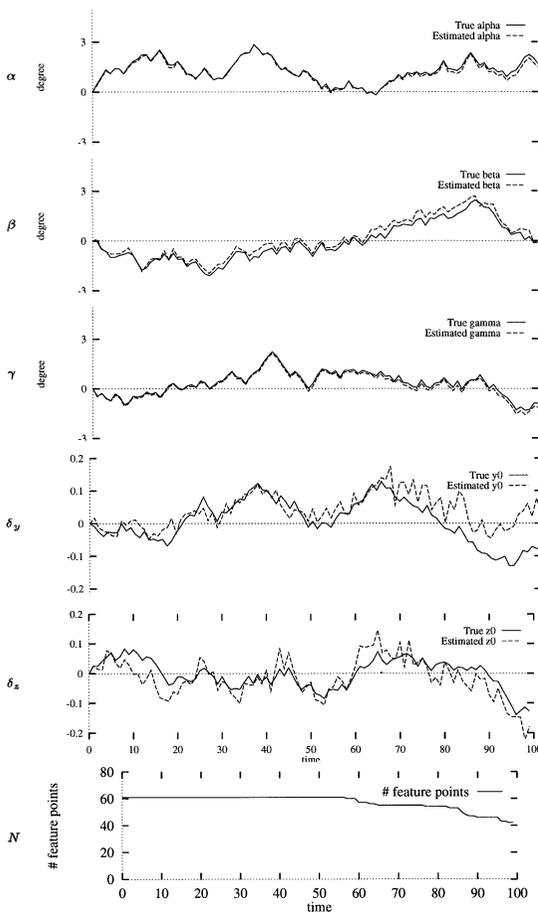


図 10 カメラ姿勢推定．カメラと手前の物体面との距離は 15 である  
 Fig. 10 Camera geometry estimation. The distance from the camera to the nearest surface is 15 units.

図 11 形状復元の定量評価

Fig. 11 The quantitative evaluation of shape extraction.

あり、概形は復元されているといえる。

また、Strip-DP は入力画像と合成画像との差を評価関数としているため、接続を誤った場合でも描画画像の品質は保たれることが特徴である。

図 12 は、屋内景観での実験の入力画像である。DV カメラを手持ちで撮影した。補正後の画像 (図 13) では、特徴点軌跡が水平になっていることが確認できる。

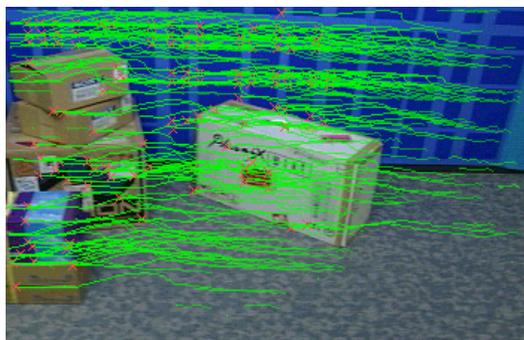


図 12 室内景観：入力画像と特徴点軌跡  
Fig. 12 A BOX SCENE. Corners are tracked.

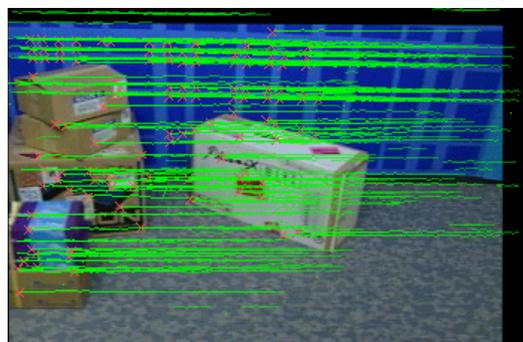


図 13 室内景観：変形後画像と特徴点軌跡  
Fig. 13 The calibrated BOX SCENE.



図 14 室内景観：抽出したモデルから生成した画像  
Fig. 14 An image synthesized from extracted textured line-segments.



図 15 屋外景観入力画像  
Fig. 15 An input image of outdoor scene.



図 16 屋外景観生成画像  
Fig. 16 A synthesized image of the outdoor scene.

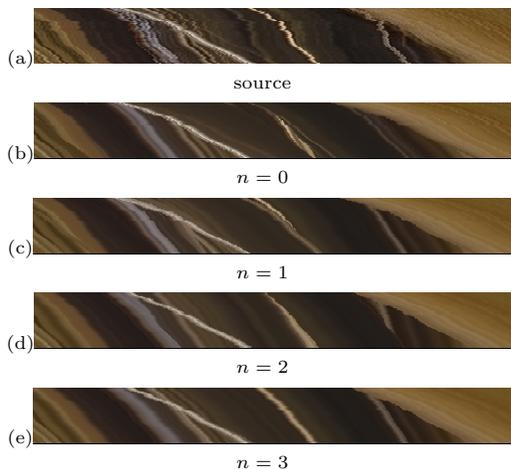


図 17 繰返し手法により EPI が改善される様子．補正後の EPI と、 $n$  回目のループで生成された EPI  
Fig. 17 EPIs are refined by iteration. Calibrated EPI and synthesized EPIs in the  $n$ th loop.

図 14 は、抽出した 3D モデルを描画したものである．繰返し処理の効果を評価するための実験を屋外景観で行った (図 15, 16)．図 17 は、もとの EPI と、 $n$  回目のループで抽出されたモデルで再生した EPI を

示している．改善の尺度として、2 乗誤差  $err$  を次のように定義する

$$err = 1/h \sum_Y \left\{ \sum_{X,t} |i_o - i_s|^2 / (i_o)^2 \right\}. \quad (12)$$

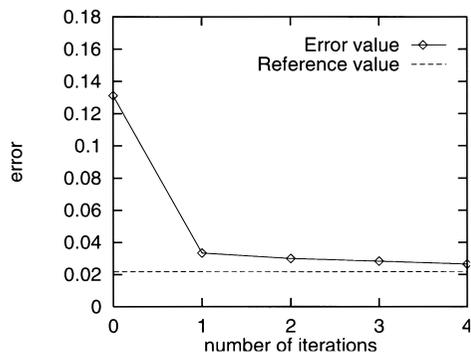


図 18 繰返し回数と 2 乗誤差

Fig. 18 Error evaluation versus iteration number.

ここで,  $i_o$  は補正後の入力画像,  $i_s$  は再生画像,  $h$  はスキャンライン数を表す. 繰返し回数に対する誤差の値をプロットしたグラフを図 18 に示す. 誤差は単調に減少し, 4 回目で 0.026 に達した. 比較のために, 入力画像に 1 画素のランダムジッタを加えた画像  $j$  を作成し, このときのエラー値を参照値とした.

$$j(X, Y; t) = i_o(X - n_x, Y - n_y; t).$$

ここで  $n_x, n_y$  はランダムに +1 か -1 の値をとる. 参照値は 0.021 であった. よって, 本手法によって得られる 3D モデルは, 1 画素程度の誤差で原画像を再現できるといえる.

## 7. むすび

本論文では, ビデオ画像から自動的に 3D CG モデルを抽出する EPI ベースの手法を提案した. この手法では, カメラが非等速直線運動しながら撮影した動画画像を, あたかもカメラが等速直線運動しながら撮影したかのような動画画像に補正する. カメラ姿勢・位置の変動は, 画像中の特徴点軌跡から推定する. 補正後の動画画像には, 従来の EPI 解析を高精度に適用することができる. 繰返し手法によって, 画像変形の精度を改善することもできる.

本手法により, エピポーラ画像解析におけるカメラ運動の制限が緩和され, 低コストで簡単に景観・物体の 3 次元モデル化が可能になった.

## 文 献

[1] H.H. Baker and R.C. Bolles, "Generalizing epipolar-plane analysis on the spatiotemporal surface," International Journal of Computer Vision, vol.3, no.1, pp.33-49, 1989.  
 [2] R.C. Bolles, H.H. Baker, and D.H. Marimont,

"Epipolar-plane image analysis: An approach to determining structure from motion," International Journal of Computer Vision, vol.1, no.1, pp.7-55, 1987.

[3] C. Harris, "Determination of ego-motion from matched points," Alvey Vision Conference, 1987.  
 [4] 南田幸紀, 新谷幹夫, 曾根原登, "エピポーラ画像解析のための画像補正処理," 秋季信学ソ大, D-12-71, p.293, 1998.  
 [5] 南田幸紀, 新谷幹夫, 曾根原登, "エピポーラ画像解析のための撮影時刻補正処理," 信学大全, D-11-95, p.95, 1999.  
 [6] 納富幹人, 小澤史朗, 全 柄東, "移動体観測による都市モデル構築," 信学論 (D-II), vol.J81-D-II, no.5, pp.872-879, May 1998.  
 [7] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," IEEE Trans. PAMI, vol.7, no.2, pp.139-154, 1985.  
 [8] 新谷幹夫, 納富幹人, 南田幸紀, 齋藤隆文, "生成画像誤差最小化による EPI 解析," 信学論 (D-II), vol.J84-D-II, no.9, pp.2020-2028, Sept. 2001.  
 [9] M. Shinya, T. Saito, T. Mori, and N. Osumi, "VR models from epipolar images: An approach to minimize errors in synthesized images," Lecture Notes in Computer Science 1352 Computer Vision - ACCV '98, volume II, pp.471-478, Springer, 1998.  
 [10] 新谷幹夫, 杉山和弘, "実写ベースのコンピュータグラフィックス技術," 情報処理, vol.41, no.6, pp.676-681, 2000.  
 [11] G. Xu and Z. Zhang, Epipolar Geometry Stereo, Motion and Object Recognition — A Unified Approach, Kluwer Academic Publishers, 1996.  
 [12] T. Yasuno and S. Suzuki, "Occlusion analysis of spatio temporal images for surface reconstruction," 4th British Machine Vision Conference, pp.549-558, 1993.  
 [13] Z. Zhu, G. Xu, and X. Lin, "Constructing 3D natural scene from video sequences with vibrated motions," Proc. IEEE VRAIS 98, pp.105-112, 1998.

(平成 13 年 1 月 4 日受付, 4 月 26 日再受付)



南田 幸紀 (正員)

平 6 東大大学院修士課程了. 同年日本電信電話(株)入社. 社内情報システムの開発を経て, 現在, NTT サイバーソリューション研究所でコンピュータグラフィックスの研究開発に従事.



新谷 幹夫 (正員)

昭 56 早大理工修士課程了。同年、日本電信電話公社入社。以来、文字認識の研究、コンピュータグラフィックスの研究に従事。昭 63~64 トロント大学客員研究員。平 7~9 日本電信電話(株)NTT ヒューマンインタフェース研究所カリフォルニアラボ勤務。現在、NTT サイバーソリューション研究所主任研究員。工博。ACM 会員。



納富 幹人 (正員)

平 10 東京商船大大学院修士課程了。同年日本電信電話(株)入社。以来、コンピュータグラフィックスの研究開発に従事。現在、NTT 西日本通信機器事業部に勤務。