

## ステレオ視による位置姿勢の変化する 3次元物体の追跡

石山 豊\*、角 保志\*\*、富田 文明\*\*

\* スタンレー電気 (株) 技術研究所

\*\* 電子技術総合研究所

ステレオ視によって3次元空間中の運動物体を追跡する手法を提案する。まず、ステレオカメラから得られるシーンの3次元情報と、追跡物体の幾何モデルから追跡物体を認識することにより、物体の位置姿勢を検出し、これを物体追跡の初期値とする。物体追跡はエッジ上の追跡点に注目し、その局所の特徴から追跡点のフレーム間対応を求め、対応点の移動前後の3次元情報から運動パラメータ(回転成分、並進成分)を算出することにより、運動物体の位置姿勢を随時検出する。本手法は簡潔であるため、汎用ワークステーションでも実時間で追跡処理が可能であることを実験結果と共に示す。

## Motion Tracking of 3-D Objects Using Stereo Vision

Yutaka Ishiyama\*, Yasushi SUMI\*\*  
and Fumiaki TOMITA\*\*

\*Stanley Electric Co., Ltd., R&D

\*\*Electrotechnical Laboratory

We propose the algorithm for model-based motion tracking of 3-D objects. A geometric model of an object is compared with a 3-D description of the scene which is generated with stereo vision. The algorithm detects the tracking points from frame-to frame using local edge features in the images. The motion parameter is calculated from 3-D tracking points detected in each frame. Since, this algorithm is simple, it can track moving 3-D objects in real-time by usual work-station.

## 1 はじめに

これから到来する21世紀に向けて、車社会の日本では、車の安全性、快適性に対する要望が強くなっている。このような背景から、車の走行環境を理解し、ドライバーの運転をサポートするシステムの開発が、必要であると考えられる。それには、まず第一に、人間の視覚機能を有するシステムの開発が、必要不可欠である。そこで、現在、我々の研究グループでは、ステレオビジョンによる実用化システム [1] の開発を進めている。本システムは、ステレオカメラによる3次元情報の入力から、物体認識や運動物体の追跡までを行うトータルの3次元視覚システムである。本論文では、その中の物体追跡に関する新たな手法を提案する。

物体の動きを認識する場合には、視覚センサから連続的に得られる時系列画像を用いる。その時系列画像から物体の動きを検出する基本的な方法には次の3通りある。

第1は、2画像間で物体が微小な場合に利用できる方法である。画像各点に対して、画像上の空間的な明るさの変化と連続する画像間の時間的な明るさの変化から、物体の動きを決定する方程式を導くことができる。しかし、この方法は、雑音に弱く、実データを用いた例は少ない。

第2は、2画像間で物体が大きな場合に利用する方法である。画像間で物体の同じ部分を表す点の対を求め、その変位から物体の動きを計算することができる。しかし、この方法は、ステレオより難しい対応問題を抱えている。但し、物体の動きが予測できる場合には、その対応探索は比較的容易である。

第3は、多数の時系列画像を時間の順に重ねてできる3次元時空間画像を利用する方法である。データ量は非常に大きくなるが、隣接する画像間の物体の動きは微小であるので、対応点を求めるのは容易であり、それを時間順に追跡していけば、結果として変移の大きな対応点を求めることができる。

本論文ではステレオ視によって3次元空間中の運動物体を追跡する手法を提案する。図1に示したシステム構成において、まず、ステレオカメラから得られるシーンの3次元情報と、追跡物体の幾何モデルから追跡物体を認識することにより、物体の位置姿勢を検出し、これを物体追跡の初期値とする。物体追跡はエッジ上の追跡点に注目し、その局所的特徴から追跡点のフレーム間対応を求め、対応点の移

動前後の3次元情報から運動パラメータ(回転成分、並進成分)を算出することにより、運動物体の位置姿勢を随時検出する。本手法は簡潔であるため、汎用ワークステーションでも実時間で追跡処理が可能であることを実験結果と共に示す。

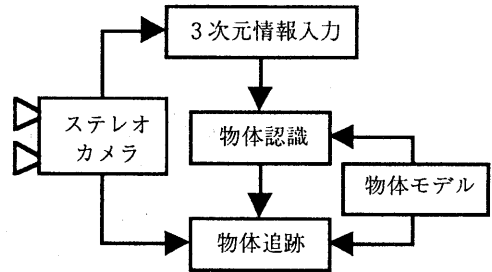


図1：システム構成

## 2 3次元情報入力

3次元情報入力には、現在、開発中のステレオビジョンシステムPATAPATA [2] を使用した。このシステムにより入力した図2のステレオ画像(640x480pixels, 256gray-levels)から、まず、物体の境界線(エッジ)を抽出し、特徴点(分岐、変曲、屈曲、遷移)により分割して、セグメント化する。その結果を境界表現(B-REP)データ構造 [3], [4] で保持する。これは、“領域”、“境界線”、“セグメント”、“点”を要素として関係付けるグラフ構造で、トポロジー的に表現する。このデータ構造は、セグメントベースステレオにおいて、有益な情報を効率良く蓄えることができる。図3がB-REPの点列を表示したものである。そして、左右画像のB-REPから、セグメントをステレオの対応単位として得られた3次元情報を図4に示す。

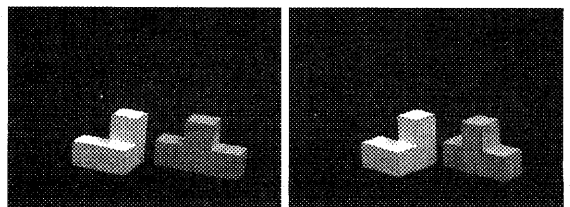


図2：ステレオ画像

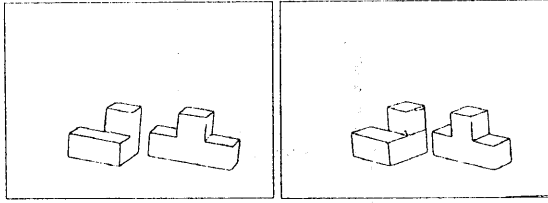


図3：境界表現

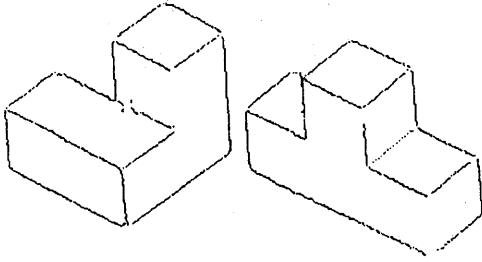


図4：3次元情報

### 3 物体認識

物体追跡を行うためには、追跡対象物体を特定し、その物体の初期フレームにおける位置姿勢を知る必要がある。これには、同時に開発を進めている、物体認識手法 [5] を使用した。

#### 3.1 物体モデル

物体モデルは、物体認識と物体追跡で共通したデータ構造を持つようにした。この物体モデルは、各セグメント上に等間隔にサンプル点を持つ。そのサンプル点には、物体の幾何特徴に関する情報が蓄えられている。そして、そのモデル生成のために2通りの方法を開発した。

幾何モデラ：

EusLispのモデリング機能を利用して、生成した物体オブジェクトのデータ構造を、本システムにおける物体モデルのデータ構造に変換する。

対話型モデラ：

EusLispにより、記述が困難な対象物体場合に、その物体のモデルを、実シーンからステレオカメラを用いた3次元計測で入力する。そして、ユーザとの対話形式で必要なセグメントを選び、モデルを構築させていく。また、一度に入力できるセ

グメントは、カメラから観測可能な部分だけである。そこで、観測方向を変えながら、上記の処理を繰り返し、モデルを更新させる。

図5は、幾何モデラにより、生成した積み木の物体モデルである。そして、図6は、シーンにおいて、積み木を認識した結果である。

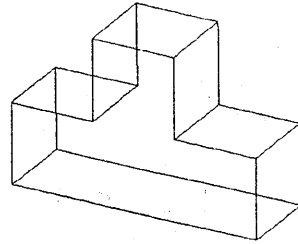


図5：物体モデル

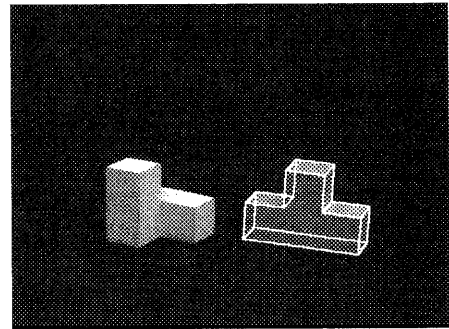


図6：物体認識

### 4 物体追跡

物体追跡はエッジ上の追跡点に注目し、その局所の特徴から追跡点のフレーム間対応を求め、対応点の移動前の3次元座標  $P(t)$  と、移動後の3次元座標  $P(t+1)$  から運動パラメータを次式より算出し、運動物体の位置姿勢を随時検出する。

$$P(t+1) = RP(t) + T$$

なお、Rは回転成分を示す  $3 \times 3$  行列、Tは並進成分を示す  $3 \times 1$  行列である。

以下、物体追跡の手法について詳しく説明する。

### (1) 初期設定

初期フレーム (frame (0)) において、物体認識結果から得られた、追跡物体の位置姿勢に関する以下の情報を、初期値として蓄える。

・頂点情報：

頂点の3次元座標と法線方向、及び、その頂点を形成する2セグメントのセグメント番号

・サンプル点情報：

図7に示すように、物体認識で使われた各セグメント上のサンプル点を、間引きしながら等間隔に選び、これを代表点とする。その代表点は、3次元座標と左右のカメラ投影面における2次元座標とその点における接線方向をそれぞれ持つ。

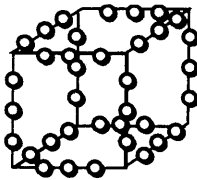


図7：代表点

### (2) 追跡点の選択

各頂点の3次元座標とその法線方向から得られる頂点ベクトル $V_n$ と、カメラパラメータから得られる左右カメラそれぞれの視線ベクトル $C_L$ 、 $C_R$ から、各々の内積を次式から計算し、左右カメラから観測可能な頂点を選ぶ。

$$\cos^{-1} \left( \frac{V_n \cdot C_L}{|V_n| |C_L|} \right) > \frac{\pi}{2}$$

$$\cos^{-1} \left( \frac{V_n \cdot C_R}{|V_n| |C_R|} \right) > \frac{\pi}{2}$$

図8に示した、頂点ベクトル $V$ は、内積計算により観測可能な頂点として判定された頂点である。ここで、その頂点を形成するセグメント番号を調べる。頂点 $V$ の場合は、セグメント $S1$ 、 $S2$ となる。この処理を物体モデルの全ての頂点に対して行い、観測可能なセグメントをリストアップする。そして、この選択されたリストの中で、セグメントの表と裏の両方がある場合は、表裏一体化により、どちらか的一方のみを選び、他方はリストから削除する。

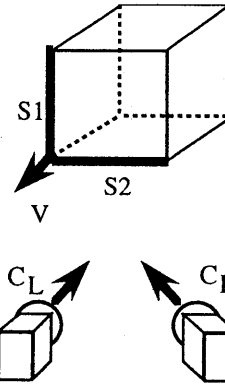


図8：観測可能な頂点

図9は選択されたセグメント上の代表点を示している。以降、この代表点を追跡点として処理を行う。また、この時、各追跡点は、形成されている面の輝度情報を入力する。

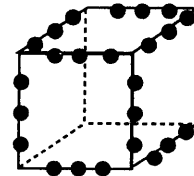


図9：選択された追跡点

### (3) 追跡点のフレーム間対応

左画像上の各追跡点において、frame (t) の追跡点 $P_{n_L}(t)$ からframe (t+1) の追跡点 $P_{n_L}(t+1)$ を検出することにより、各追跡点のフレーム間対応をとる。なお、ここでは、frame (t) の追跡点をM (モデル)、frame (t+1) の追跡点をD (データ) と置き換えて説明する。

#### (3-1) 仮の追跡点の検出

図10 (a) に示すように、frame (t+1) に対して、frame (t) の追跡点 $M_L$ の位置からエッジの垂直方向に、追跡探索範囲の画素輝度値を入力する。この画素輝度値から輝度差分を求め、各点の微分値を計算し、微分値のピークとなっている部分を検出する。そして、各ピーク点において、エッジの方向を検出する。これは近傍の $5 \times 5$ 画素の輝度画素値から、16方向に分割した方向コードで表す。そして、各

ピーク点において、frame (t) の追跡点での接線方向 (傾き) と面の輝度値 (明るさ) の相関からフレーム間の類似性を評価し、評価値の良いピーク点を仮の追跡点  $D'_L$  とする。

### (3-2) 仮の追跡点の3次元計測

仮の追跡点の3次元計測は、ステレオ視による計測と、単眼視による計測の2通りの方法を開発した。

#### (a) ステレオ視

図11に示すように、ステレオ対応は、まず、左画像の仮の追跡点  $D'_L$  に対応する、右画像でのエッジ線をカメラパラメータから求める。そして、エッジ線上で、frame (t) での右画像の追跡点  $M_R$  に一番近い点からエッジ線に添って、先ほど、処理(3-1)で左画像に対して行った同様の処理でフレーム間対応をとり、右画像での仮の対応点  $D'_R$  を探す。この点と左画像での仮の対応点  $D'_L$  がステレオ対応の対応点となり、3次元座標が求まる。

#### (b) 単眼視

図12に示すように、単眼では左カメラのレンズ中心  $(x_{CL}, y_{CL}, z_{CL})$  と、カメラ投影面上の仮の追跡点  $D'_L$  を結ぶ直線の式をカメラパラメータから計算し、frame (t) での追跡点の3次元座標  $(x_M, y_M, z_M)$  から、さきほどの直線への垂線の足を frame (t+1) における追跡点の3次元座標として、近似する。

$$\frac{x_{D'L} - x_{CL}}{u} = \frac{y_{D'L} - y_{CL}}{v} = \frac{z_{D'L} - z_{CL}}{w}$$

$$x_{D'} = x_{CL} + us, \quad y_{D'} = y_{CL} + vs, \quad z_{D'} = z_{CL} + ws$$

$$s = \frac{u(x_M - x_{CL}) + v(y_M - y_{CL}) + w(z_M - z_{CL})}{u^2 + v^2 + w^2}$$

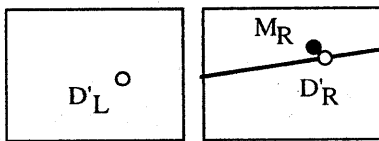


図11：ステレオ対応

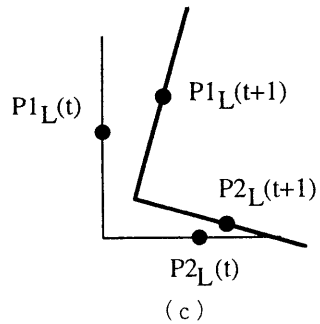
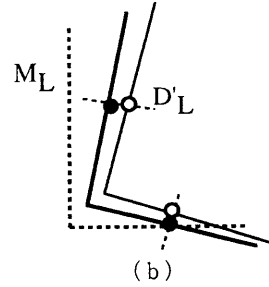
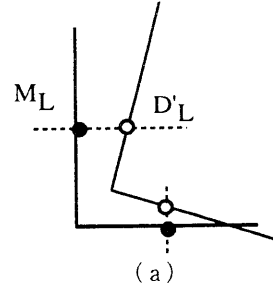


図10：フレーム間対応

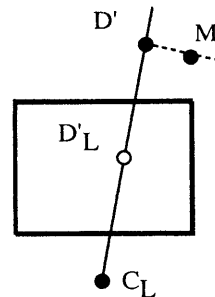


図12：単眼視による距離の近似

### (3-3) 運動パラメータの算出

frame (t+1) での仮の追跡点の3次元座標  $D_n'$  と、それに対応する frame (t) の追跡点の3次元座標  $M_n'$  から、次式を最小にする運動パラメータ (R:回転成分、T:並進成分) を最小自乗法により求める。なお、 $W_n$  は追跡点の3次元計測誤差に対する重み付けで、追跡点でのエッジ方向から重み付けを決め、エッジ線に垂直な方向ほど高くし、平行になるにつれて低くする。

$$\sum_{n=1}^k W_n (RM_n + T - D_n')^2$$

### (3-4) 仮の追跡点の更新

frame (t) の追跡点に対し、求めた運動パラメータから frame (t+1) における追跡点の3次元座標を計算する。そして、3次元計測により得られた3次元座標と比べて、3次元距離誤差が極端に大きい追跡点は、誤対応の可能性があるので、その点を省いて、再び運動パラメータを計算し直す。この運動パラメータにより、物体モデルの位置姿勢を変換し、次に、その点のカメラ投影点をカメラパラメータから計算する。これを新たな仮の追跡点として更新する(図10(b)参照)。この処理(3)を3回繰り返す。これで、運動パラメータによる座標変換が求まり、図10(c)に示すように、追跡点  $P_{1L}$ 、 $P_{2L}$  のフレーム間対応がとれた。

### (4) 追跡点情報の更新

物体の運動により、位置姿勢が変化したので、物体モデルが持つ頂点ベクトル情報を、運動パラメータから更新する。更に、処理(2)と同様にして、移動後 (frame (t+1)) の観測可能な頂点を選び直し、追跡点を更新する。また、この時、面の輝度情報も更新させる。

この一連の処理(2)~(4)を繰り返すことにより、運動物体の位置姿勢を随時検出し、物体が追跡できる。また、本手法は直線のみならず、曲線部分を含む物体に対しても、同様の処理で追跡できる。

## 5 実験

### 5.1 実験環境

ワークステーション:

SPARC Station10 Model41 [SUN]

画像入力ボード:

Frame Grabber Model S2200 [DATA CELL]

CCDカメラ:

XC-75 [SONY] 2台

### 5.2 実験結果

図2のシーンにおいて得られた、追跡対象物体の認識結果(図6)から物体追跡を行った。図13は、左画像上のあるフレーム間での追跡点の様子を示した。また、図14は、5フレーム毎の物体追跡の様子を、追跡結果より得られた運動パラメータから、物体モデルを座標変換した結果を、シーンの左画像と重ねて表示した。なお、図中の実線部分はカメラから観測可能なセグメントを表し、点線部分は観測不可能なセグメントである。回転により、物体の姿勢が反転しても、追跡対象とするセグメントを変えながら、確実に物体追跡を行っていることがわかる。さらに、図15は、曲線図形に対して、追跡して結果である。曲線図形でも良好な結果が得られている。また、これらの処理時間については、追跡点の数により変化するが、実験したシーンでは、約3フレーム/秒という結果であった。特別な専用プロセッサを使わずに、シングルCPUの汎用ワークステーションでも、実時間で物体追跡処理が可能なが確認できた。また、更に、1フレームあたりの処理時間が、短縮されれば、フレーム間の移動量も少なくなり、より安定した結果が得られると考えられる。

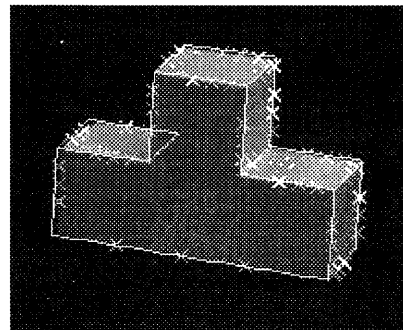
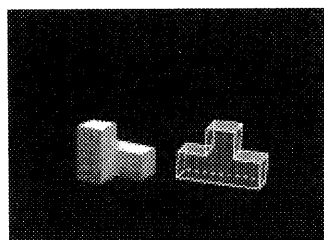
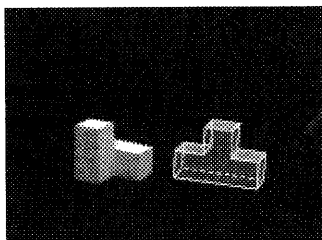


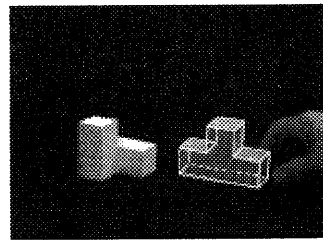
図13: 追跡点の検出



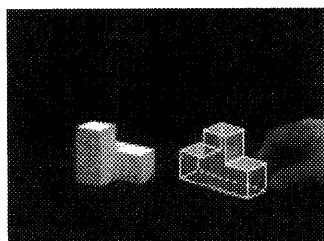
frame (0)



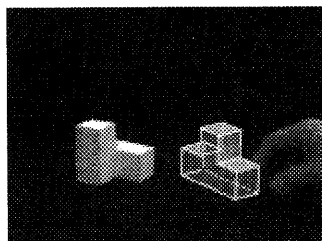
frame (5)



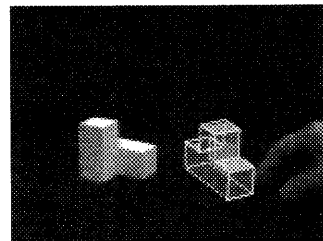
frame (10)



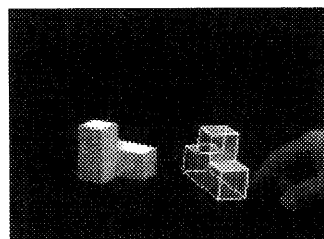
frame (15)



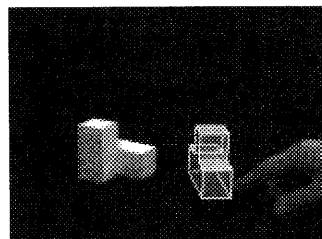
frame (20)



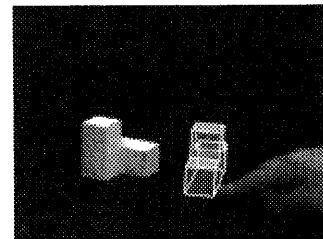
frame (25)



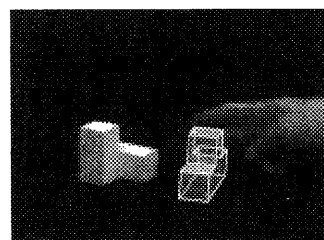
frame (30)



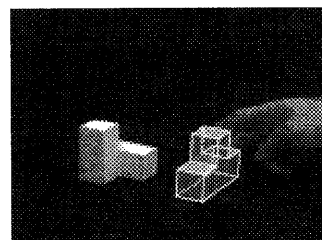
frame (35)



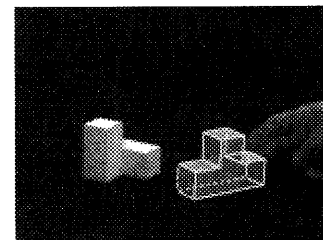
frame (40)



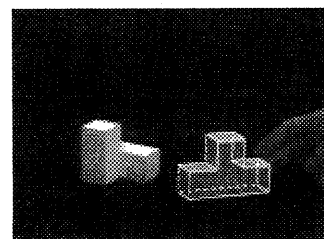
frame (45)



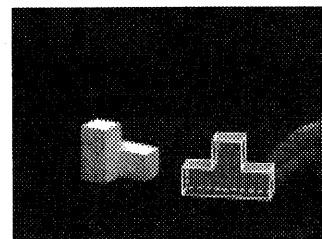
frame (50)



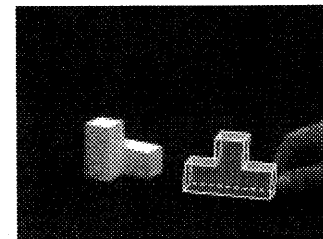
frame (55)



frame (60)



frame (65)



frame (70)

図14：積み木の追跡

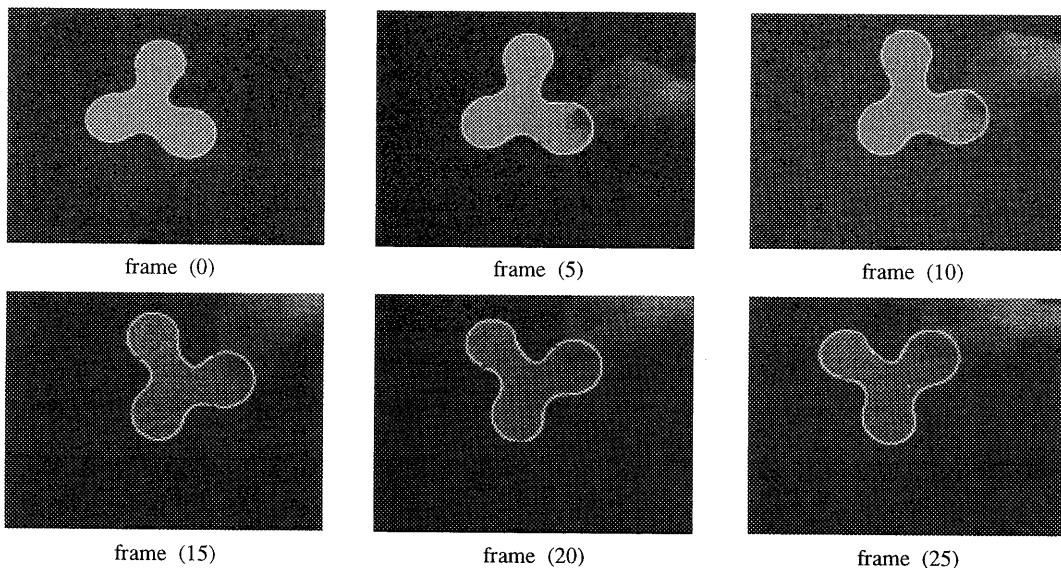


図 1 5 : 曲線図形の追跡

## 6 まとめ

セグメントベースのステレオビジョンシステムにおいて、3次元空間中の運動物体を追跡する手法について述べた。

本手法はエッジ上の追跡点に注目し、その局所の特徴からフレーム間対応を取り、そして、ステレオ対応で追跡点の3次元計測し、フレーム間の3次元運動を求める。このフレーム間対応と、ステレオ対応のそれぞれの対応付けには、類似した手法を使うことができる。また、物体追跡は高速処理が必要となる。そこで、今回は高速化を充分考慮して、開発したため、汎用ワークステーションでも、実時間処理が可能であった。そして、ハードウェア化も容易であり、ビデオレート処理も可能であると考えている。

また、今後は、曲面を含んだ任意形状の物体追跡ができるシステムを開発する。

## 謝辞

日頃有益な助言と討論を頂く、電総研視覚情報研究室、PGV研究会の各位に感謝致します。

## 参考文献

- [1] 富田: 3次元ビジョンの実用化に向けて, 日本ロボット学会誌, Vol.12, No.8, pp.1124-1127 (1994) .
- [2] 河井ほか: ステレオカメラシステム—パタパター, 情処学シンポジウム (MIRU'94) 論文集 No.II, pp.127-134 (1994) .
- [3] 富田, 高橋: 画像のB-REPのためのアルゴリズム, 信学技報, PRU86-87 (1986) .
- [4] 杉本, 富田: 輪郭線の屈曲点, 変曲点, 遷移点の検出, 情処学シンポジウム (MIRU'94) 論文集 No.I, pp.83-90 (1994) .
- [5] 角, 富田: ステレオビジョンによる3次元物体の認識, 情処学CV研, CV-95-2 (1995) .