

ステレオ動画像解析の一手法

正 員 山本 正信[†] 非会員 ロベルト チポラ^{††}

A Cooperative Analysis of Stereo and Moving Images

Masanobu YAMAMOTO[†], Member and Roberto CIPOLLA^{††}, Nonmember

あらまし ステレオ動画像の解析には、ステレオ画像と動画像の2種類の画像間対応付けを決定しなければならず、それぞれ多重対応問題とオクルージョン問題を含んでいる。運動軌跡の画像化手法⁽¹⁾を使えば、物体の動きに制約があるが、動画像の対応付け問題を回避でき、直接運動軌跡が抽出される。本論文では、左右のステレオ動画像からそれぞれ抽出された運動軌跡のステレオ対応付けが、エピソード拘束のみを用いて多重対応の数を著しく減少させることができ、場合によっては一意に決定可能であることを示す。この対応付けは、結果的にステレオのオクルージョン問題も解決している。

1. ま え が き

本論文では、ステレオカメラから得られた動画像(ステレオ動画像)の新しい対応付け手法を提案している。

ステレオ画像間の対応点探索においては、ステレオカメラの配置と3次元空間中の点との幾何学的関係から導かれるエピソード拘束が対応点決定のための必要条件となる。カメラ間の基線長を小さくすると微小視差ステレオ⁽²⁾ではエピソード拘束のみでほぼ一意に対応先が定まる。しかし、高い測定精度は望めない。測定精度を向上させるためには基線長を長くする。その場合、エピソード拘束を満たす対応候補点は一般に数多く存在し、真の対応付けを決定するためには、更に別の拘束条件を必要とする。左右の対応点上の明るさの類似性、あるいは物体面上の視差分布の滑らかさ⁽³⁾などは新たな拘束条件として有用である。相関法、DP マッチング⁽⁴⁾あるいは疎密法⁽⁵⁾、し緩法などはこれらの条件を満たす画像間の対応付けを効率的に探索する。最近では、画像上から領域やその境界線の幾何学的特徴量を求めそれらの類似性をたよりに対応相手を探索する手法も研究されてきている^{(6),(7)}。

これらの拘束条件はいずれもその適用には前提条件

が存在し、エピソード拘束ほどの一般性はない。また、上述した多重対応問題のほかに、オクルージョン問題もステレオの対応付けを難しくしている大きな要因である。

ステレオ動画像では更に、動画像間の対応付けにも多重対応問題やオクルージョン問題が加わっている。Jenkin と Tsotsos⁽⁸⁾はステレオ画像と動画像の対応付けを同時に行い、数多くの可能な対応付けの中から、物理的に妥当な3次元移動ベクトルを与える対応付けを選択した。しかし、対応付けの可能性は時空間的に爆発的に広がり、実画像向きではない。動画像が物体の動きに比べて時間的に密に得られるとすれば、動画像間の多重対応の可能性は少なくなる(ここでは窓枠問題は扱わないことにする)。Waxman⁽⁹⁾は左右の動画像上のオプティカルフローをまず求め、物体面上では視差とその時間変化率の比は滑らかに分布するという仮定のもとに対応関係を決定した。また、左右のカメラを独立に移動させ、微小視差ステレオが成り立つ瞬間を利用して、左右動画像のオプティカルフローから距離測定を行う手法⁽¹⁰⁾を提案した。しかし、前者は物体面が細かなテクスチャで覆われた滑らかな面で近似できることが前提であり、後者は瞬間的な動きの解析に便利ではあるが、長時間の挙動の解析には使えない。

本論文では、最初に物体の運動軌跡を左右の動画像上でそれぞれ求め、運動軌跡のステレオ対応付けがエピソード拘束のみを用いて多重対応の数を著しく減少させることができ、場合によっては一意に決定可能で

[†] 電子技術総合研究所視覚情報研究室, つくば市
Electrotechnical Laboratory, Tsukuba-shi, 305 Japan

^{††} オックスフォード大学, 英国
Department of Engineering Science, University of Oxford, Oxford
OX1 3PJ, U. K.

あることを示す。この手法は運動軌跡の画像化手法⁽¹⁾に基づいている。時間的に密に撮られた動画像を時間軸方向にたい積し、3次元時空間画像を構成したとき、この時空間断面に物体の見掛けの運動軌跡が画像化される。物体の運動軌跡はこの時空間断面画像上のエッジ検出とその追跡により、たとえ一時的に動画像上にオクルージョンが起こったとしても、抽出が可能である。抽出された運動軌跡のエピポーラ拘束に基づく対応付けは、ステレオの多重対応問題ばかりではなくオクルージョン問題も解決可能である。

本論文の構成を以下に記す。2.では運動軌跡の画像化手法を使って、エピポーラ拘束のみでステレオ動画像の対応付けが可能であることを示す。3.では、ステレオ動画像の入力装置について述べる。4.では、本手法を使って歩行者の動きを解析する。結論を5.で述べる。

2. ステレオ動画像の対応付け

左右カメラの光軸が平行な標準ステレオカメラモデルを考えよう(図1)。シーンはそれぞれの画像平面に焦点距離 f で中心投影される。左右投影中心を結ぶ線をステレオの基線と呼ぶ。更に、この基線を含む任意の平面をエピポーラ面、この平面と各投影面との交

線をエピポーラ線と呼ぶ。今、3次元シーン中の点がこのエピポーラ面内を運動しているとする。そのとき、この点の投影点は、各々左右のエピポーラ線上を移動している。

左右それぞれのカメラから得られた動画像を時間の順番に並べて3次元の時空間画像を構成する。エピポーラ面を固定したとき、1枚1枚の画像上でのエピポーラ線はこの時空間画像を平面で切った一つの断面上に載っている。もし、動画像のサンプリング時間が対象物体の動きに比べて十分に速いならば、投影点はこの断面上で連続して動き、もしその点がエッジに載っているならば、その軌跡が画像化される。

図2に左右の時空間断面画像の例を示す。図には、左右に行き交っている物体の輪郭エッジの動きが画像化されている。点線部は見掛けの重なり合いにより隠されていることを示している。これらの軌跡が動画像上のオクルージョンにもかかわらずそれぞれ抽出され追跡できたとしよう。初期時刻における左時空間断面画像上の軌跡の位置を l 、右時空間断面画像上の軌跡の位置を r とする。標準ステレオカメラの配置から、左画面上で見えるシーン中の点の位置は、右画面上で見える位置よりも必ず右方向に寄って見える。すなわち、

$$l > r \tag{1}$$

これが、エピポーラ拘束である。

図2では、点 l_1 に対してエピポーラ拘束を満たす点は r_1, r_2, r_3 の3点である。点 l_2 に対しては r_2, r_3 の2点、 l_3 に対しては r_3 の1点、 l_4 に至っては対応点は見つからない。このように対応先が視野の外に飛び出している場合や他の物体に隠されている場合、あるいは背景との明るさのコントラストが不十分な場合には対応先を見つけることができない。

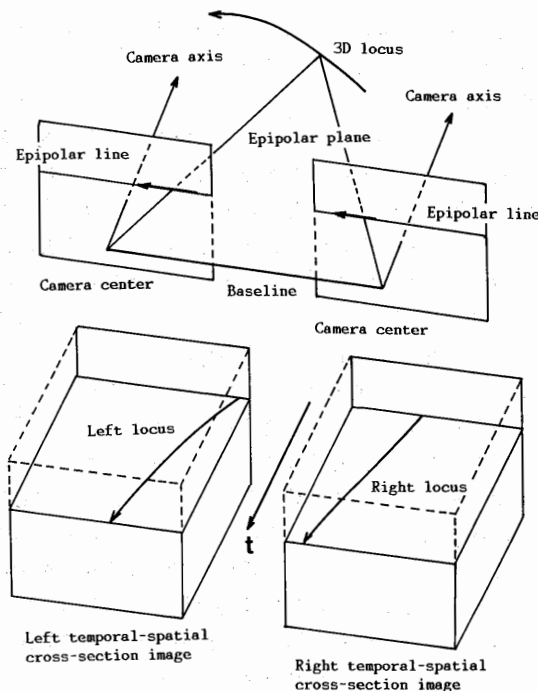


図1 運動軌跡の画像化手法を用いたステレオ動画像解析
Fig. 1 Stereo visualized locus method.

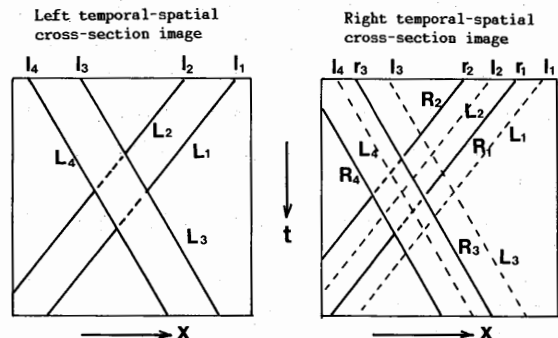


図2 エピポーラ拘束による運動軌跡の対応付け
Fig. 2 Stereo match based on temporal epipolar constraint.

さて、エピソード拘束はステレオカメラ対の幾何学的配置のみにより決まる拘束である。それは、同一の点に対して何時の時刻でも成り立つ拘束である。従って、それを強調する意味で時間 t を入れて、エピソード拘束を、

$$l(t) > r(t) \quad (2)$$

と書き直し、時間エピソード拘束と呼ぶことにする。

この拘束は左右の時空間断面画像上の運動軌跡の対応拘束となる。再度、この時間エピソード拘束を対応付けに適用してみよう。軌跡 L_1 に対応可能な軌跡は当初 R_1, R_2, R_3 であった。しかし、時間の経過と共に R_3 は L_1 と交差しエピソード拘束を満たさなくなる。 R_4 に対しても同様である。従って、 L_1 に対応可能な相手は R_1 と R_2 になる。 L_2 に対しては R_2, L_3 に対しては R_3 と R_4, L_4 に対しては R_4 ということになる。更に、 L_2 と L_4 の対応相手がそれぞれ R_2, R_4 と一意であることから、 L_1, L_3 の対応相手もそれぞれ R_1, R_3 となる。すなわち、すべての軌跡の対応付けがエピソード拘束のみで一意に決まったことになる。また、静止ステレオ画像上でのオクルージョン問題も解決されていることがわかる。

ここで注意しておくべきことは、時間エピソード拘束を使えば必ずしも対応付けが一意に定まるとは限らないことである。上記の例のように運動軌跡が互いに交差している場合はともかく、交差が起こらずほとんど平行に移動している物体に対しては多重対応問題は

解決されない。

本章での議論を成り立たせるために二つの重要な仮定をした。その一つは物体上の点の動きが1枚のエピソード面上に拘束されなければならないこと、もう一つは断面画像上の運動軌跡が動画上のオクルージョンにもかかわらず追跡可能なことである。この二つの仮定の実現性は4.で詳しく論じる。

3. ステレオ動画画像入力システム

ステレオ動画画像を計算機に入力するための装置のブロック線図を図3に示す。ぶれのない映像を得るために、可変電子シャッター付きのビデオカメラ (PJLG-25, 立石電機) を使用した。但し、フリッカー効果を避けるため高周波点灯 (45,000 Hz) の蛍光灯照明下で撮影している。通常の人動きは1/1,000秒程度のシャッタースピードで十分明らかにとらえることができる。

2台のステレオビデオカメラは同じ外部同期信号により駆動されている。従って、各々のビデオ信号は全く同相である。それぞれのビデオ信号はビデオディスクレコーダー (VM-1000M, JVC) に録画される。録画開始信号はコントロールボックスから同時に左右のビデオディスクに送られ、同時に録画が開始される。左右のビデオディスクに録画された画像はそのコマ番号 (フレームあるいはフィールド番号) が同じであれば、全く同じ時刻に異なる位置で撮られたステレオ画像対ということになる。このビデオディスクはそれぞれ1,

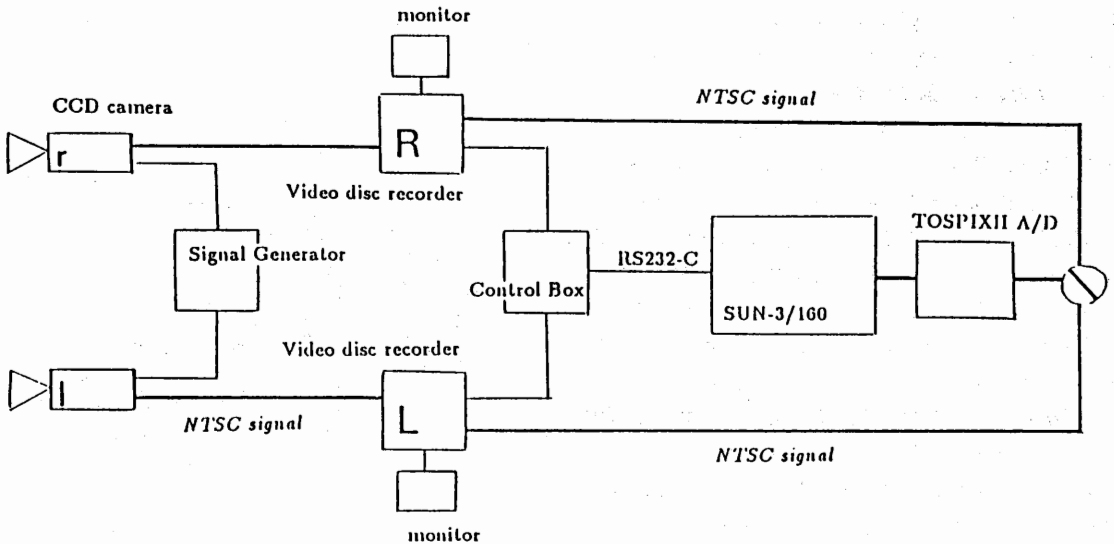


図3 ステレオ動画画像入力システム
Fig. 3 Stereo image sequence input system.

200 フィールド(約 20 秒)の動画を実時間で録画可能である。

ビデオディスクに蓄えられた動画は A-D 変換器(TospixII, 東芝)で 1 フィールドごとに 512×512, 8 ビットの粗さでデジタル化される。断面画像はエビポラ線上の 1 次元画像をフィールド順に並べることにより作成される。その際、エビポラ線を中心として上下 7 画素分の幅の帯画像を上下方向にスムージングして入力している。そのため、インタレースの影響はほとんどない。入力の全フィールド数は 512 フィールド、時間にして約 8.5 秒である。従って、横 512 画素×縦 512 フィールドの大きさの時空間断面画像が得られる。左右の時空間断面画像の生成にかかる時間は約 2 分であった。

4. 歩行シーンの解析

2. で提案したステレオ動画の解析法の有効性を実際のシーンを使って実証しよう。複数の歩行者が行き交っているシーンを対象に、歩行者の位置を追跡する。

本手法の前提である二つの仮定の実現性を 4.1 と 4.2 で述べる。4.1 では、歩行者の動きが時空間断面画像上に画像化されるための条件を、エビポラ面の配置、歩行者の動きと衣服の表面パターンについて検討した。4.2 では、時空間断面画像から運動軌跡を抽出する。前論文⁽¹⁾では、物体の動きをカメラ光軸に垂直かつ等速な直線運動に限ったため直線軌跡の抽出で十分であったが、歩行者の動きはかなり複雑であるため、歩行者の運動モデルを使った運動軌跡の抽出法を示す。4.3 では、実際のステレオ動画を使って歩行者の追跡例を示す。

4.1 運動軌跡の画像化条件

歩行者は水平面上を歩くものとする。エビポラ面を歩行面上に平行にとり、その高さを 120 cm に設定する。この位置はほぼ大人の歩行者のひじより上肩より下に相当する。この高さでは、人体の水平断面径が最大であり、手や足などの部分的な動きの影響が少なく人体の主要な動きが観測できる。歩行者の動きに上下動がなければ、このエビポラ面と歩行者との交点は常にエビポラ面上にある。しかし、実際には歩行には上下動がつきものである。歩行時の上下動が及ぼす影響を歩行モデルを使って調べてみよう。

歩行者の上半身とステレオ撮像系のモデルを図 4 に示す。シーン座標系 (x_0, y_0, z_0) の x_0 軸上にカメラ中心を据え、光軸は z_0 軸方向とする。エビポラ面を x_0z_0 平面とし、歩行者は歩行面 $y_0 = H(120 \text{ cm})$ 上を

歩くとする。上半身を胴体と腕のリンク結合モデルで表す。胴体部の座標系 (x_1, y_1, z_1) は、原点 (x_c, y_c, z_c) を首、 y_1 軸を身体を中心軸、 z_1 方向を身体の正面方向とし、その z_0 軸方向との角度を θ とする。両腕は x_1 軸上の位置 $w(>0)$ または $-w$ にリンク結合され、その x_1 軸回りの回転角を ψ とする。ここで簡単のため、投影面 $z=f$ のカメラ中心をシーン座標系の原点に置く。このとき、右腕の上の点 (x, y, z) の動きは投影面上の点 (X, Y) へ次式で中心投影される。すなわち、

$$X = f \cdot \frac{\begin{Bmatrix} x \cos \theta + y \sin \theta \sin \psi \\ + z \sin \theta \cos \psi + w \cos \theta + x_c \end{Bmatrix}}{\begin{Bmatrix} -x \sin \theta + y \cos \theta \sin \psi \\ + z \cos \theta \cos \psi - w \sin \theta + z_c \end{Bmatrix}} \quad (3)$$

$$Y = f \cdot \frac{y \cos \psi - z \sin \psi + y_c}{\begin{Bmatrix} -x \sin \theta + y \cos \theta \sin \psi \\ + z \cos \theta \cos \psi - w \sin \theta + z_c \end{Bmatrix}} \quad (4)$$

胴体上の点の投影式は上式で $w = \psi = 0$ と置くことにより得られる。また、左腕上の点の投影式も w を $-w$ で置き換えることにより容易に得られる。

さてここで、歩行者の動き方に重要な制約を設ける。歩行者はその正面方向に直線的に移動し、決して、向きを変えたり回転したりはしないものとする。この仮定の理由は次のような状況を避けるためである。図 5 に示されるように、同じ明るさの服を着た 2 人の歩行者が、互いにすれ違った場合と向きを途中で変えた場合を考える。このとき時空間断面画像のパターンはほとんど変わらず、このパターンのみから両者の違いを

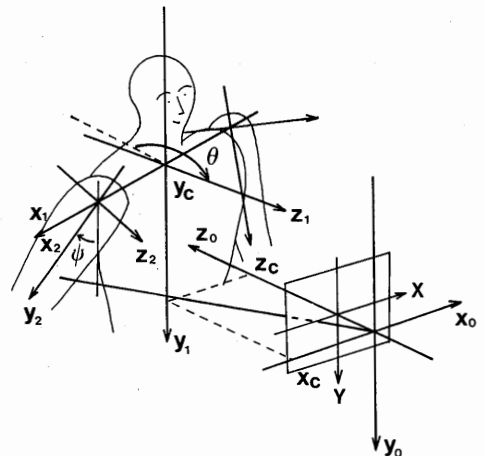


図 4 歩行者のモデル

Fig. 4 An articulated body model of a walking man.

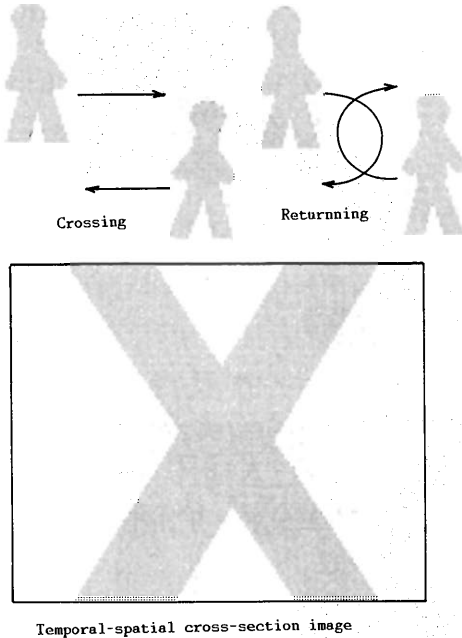


図5 時空間断面画像の解釈におけるあいまいさの例
 Fig. 5 An ambiguity of the interpretation of the temporal-spatial cross-section image.

見分けることは難しい。このあいまいさを避けるために、便宜的に上記の仮定を設けた。短い観測時間(約8.5秒)と固定されたカメラの視野内では、この動きの制約はそれほど厳しくはないと考えられる。

歩行者の向き θ は一定、歩行速度 $V(t)(\geq 0)$ とする。そのとき、首の位置、すなわち胴体座標系の原点 (x_c, y_c, z_c) は $V(t), \theta$ 、歩行に伴う左右の揺れ $a(t)$ および上下動 $\beta(t)$ を使って次式で与えられる。

$$\begin{cases} x_c = d(t)\sin\theta + c_x + a(t)\cos\theta \\ y_c = \beta(t) - h \\ z_c = d(t)\cos\theta + c_z - a(t)\sin\theta \end{cases} \quad (5)$$

但し、移動量 $d(t) = \int V(t)dt$ 、初期位置 $(c_x, -h, c_z)$ 、 $a(0) = \beta(0) = d(0) = 0$ とする。更に腕の振り ϕ は小さいとすれば、

$$\sin\phi \approx \phi, \cos\phi \approx 1 \quad (6)$$

また、初期時刻には胴体や腕の表面上の点がエピポーラ面に載っているとす。すなわち、

$$y = h \quad (7)$$

式(5)~(7)を式(3), (4)に代入して整理すれば次式を得る。

$$X = f \tan\theta + \frac{f}{\cos\theta}$$

$$\frac{a(t) + w + c_x \cos\theta - c_z \sin\theta}{\left\{ \frac{d(t)\cos\theta + \phi(t)h \cos\theta - a(t)\sin\theta}{+ z \cos\theta - (w+x)\sin\theta + c_z} \right\}} \quad (8)$$

$$Y = f \cdot \frac{\beta(t) - \phi(t)z}{\left\{ \frac{d(t)\cos\theta + \phi(t)h \cos\theta - a(t)\sin\theta}{+ z \cos\theta - (w+x)\sin\theta + c_z} \right\}} \quad (9)$$

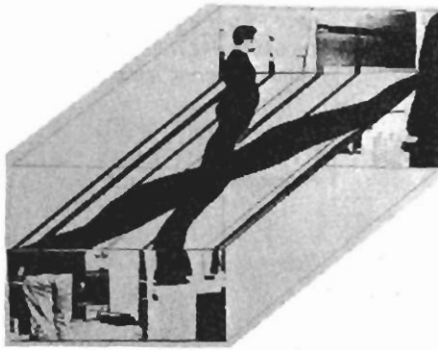
歩行者上の点がエピポーラ面に載っているという条件は $Y=0$ であるが、この条件を満たすためには、 $\beta(t) - \phi(t)z = 0$ でなければならない。通常の歩行時における上下動幅 $\beta(t)$ は約4cmであると報告されている⁽¹¹⁾。また、腕の最大振幅角を10度、径をたかだか10cmとすれば、 $\phi z = \pi/18 \times 10/2 \approx 0.87$ cmとなり、腕の上下動幅も胴体の上下動幅とほぼ同じである。従って、 $Y=0$ とはなりがたいが、歩行者上の点が垂直なパターン上に載り、しかもそのパターンの長さが4cm以上ならば、多少の上下動があっても、時空間断面画像上に連続した運動軌跡を描き、この軌跡はこのパターン上の任意の点の動きとみなすことができる。

図6に時空間断面画像の例を示す。背広や模様の少ないセーターなどでは、運動軌跡が安定に画像化されている。しかし、衣服に横縞模様のある場合には、軌跡の連続性は必ずしも保たれない。

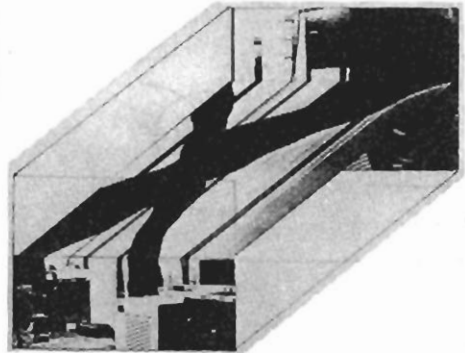
4.2 運動軌跡の抽出

時空間断面画像から運動の軌跡を実際に抽出してみよう。まず、画像を微分してエッジを求め、明るさのコントラストが大きなエッジを残す。次いで細線化を行いエッジ線画を作る。更に、分岐点の除去、時間軸方向の1価関数化を行う。歩行者の動きは滑らかであるとして、エッジセグメント上の K 曲率($K=3$)が大きな個所でエッジセグメントを分断する。ここでは K 曲率のしきい値を17度とした。更に実際にはあり得ない速度無限大に相当する水平セグメントと、静止物体に対応する垂直エッジセグメントを処理の簡便化のため削除する。このようにして得られたエッジセグメントは歩行者の運動軌跡の断片である。図6の右時空間断面画像から運動軌跡の断片を抽出した結果を図7に示す。

残った断片的な軌跡をつなぎ合わせて運動軌跡を求める。その際、次の2種類の手順を使い分けた。一つは、局所的結合と呼ばれ、ノイズなどにより誤って分断された比較的接近した運動軌跡をつなぐためのものである。もう一つは、大域的結合と呼ばれ、オクルージョン等のため互いに大きく離れた運動軌跡をつなぐ



Left 3D temporal-spatial image



Right 3D temporal-spatial image



Left temporal-spatial cross-section image



Right temporal-spatial

図6 ステレオ動画像とステレオ時空間断面画像
Fig. 6 An example of temporal-spatial cross-section image.

ためのものである。

その準備として、断面画像上の運動軌跡の性質を先に述べた歩行モデルを使って調べてみよう。腕の振り、身体の左右の揺れがないとすれば、 $\alpha(t) = \psi(t) = 0$ と置き、歩行者の運動軌跡式(8)は

$$X = f \tan \theta + \frac{f}{\cos \theta} \cdot \frac{w + c_x \cos \theta - c_z \sin \theta}{d(t) \cos \theta + z \cos \theta - (w + x) \sin \theta + c_z} \quad (10)$$

となる。これは歩行者の主要な動きである。移動量 $d(t)$ は時間に関して単調に増加する。従って、主要な動きの軌跡も図8の太い実線に描かれるようにと時間に関して単調関数となる。一方、腕の振り $\psi(t)$ と左右の揺れ $\alpha(t)$ はともに歩行の周期をもつ正弦波状の周期関数である。故に上式の分母の中の項 $\psi(t) \cos \theta - \alpha(t) \sin \theta$ も同じ歩行の周期をもつ周期関数である。従って歩行者の運動軌跡はその主要な動きの軌跡が歩行の周期をもつ周期関数で変調された関数と考えてよ

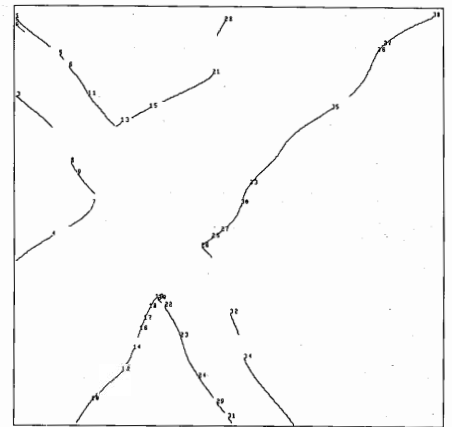


図7 抽出された運動軌跡の断片
Fig. 7 A fragmentary loci.

い。その軌跡を図8の細い実線で表す。この軌跡は歩行者の向きや腕の振りの大きさなどにより、主要な運動軌跡と違って局所的に単調関数とはならない場合が

ある。また、図6の運動軌跡の例にも見られるように、最も運動軌跡がはっきりと観測されるところは身体の輪郭部であることが多い。その場合、軌跡は必ずしも身体上の同じ点に対応しているとは限らず、腕と胴体が交互に現れ両者の軌跡の合成である可能性もある。しかし、その合成関数もやはり歩行の周期を局所的な揺らぎにもち一つの点に対応する軌跡とほぼ同じ軌跡を描く。

さて、最初に局所的結合を行う。時間間隔の短い運動軌跡同士を接続する場合、その間は直線で近似できると考えられる。先に、軌跡が歩行の周期の揺らぎをもつ関数であることを見てきた。通常の歩行の周期(歩行の2ステップにかかる時間)は約1.4秒であると実験で報告されている⁽¹¹⁾。この周期時間は我々の時空間断面画像上では約80フィールドに相当する。本論文では1/4周期に当る20フィールド以内を直線で近似接続した。接続条件であるが、時間間隔が20フィールド以内にある二つの運動軌跡の最も近い端点を線分で仮につなぐ。つなぎられた運動軌跡も運動軌跡の滑らかさを保たなければならない。つなぎ目で K 曲率を計算し、その値が先の分割の際に与えたしきい値を超えなければ二つの軌跡を線分でつなぐ。接続可能な運動軌跡が複数個存在する場合にはその中で K 曲率値が最も小さな接続を選択する。

次に、大域的結合を行う。時間間隔が20フィールド以上開くと直線接続は難しい。軌跡のモデルを使っても式の複雑さからその手続きの難しさが予想される。ここでは、次の簡単な二つの接続基準を用いて運動軌跡の接続を行う。第1の基準は、対象物体の明るさが時間に関してほぼ一定であることを利用する。運動軌跡の右側あるいは左側の少なくとも一方の領域

は同一物体上のほぼ同じ面に対応している。従って、観測中の明度はほぼ一定と見てよい。もう一つの基準は、身体的主要な動きは時間に関して単調であることに基づくものである。しかし、先の運動軌跡の解析で見てきたとおり、実際の軌跡には歩行時の各種の揺れにより局所的に単調とならない場合もある。けれども、その非単調さは歩行の周期と一致しているので、運動軌跡を時間軸方向に歩行周期(約80フィールド)で移動平均してやれば、その平均値は時間と共に単調となるはずである。この二つの基準を満たす軌跡が複数存在する場合には、局所接続と同様に線分接続したときに最も K 曲率の小さな軌跡を選択する。

以上述べた、明るさの不変性と運動軌跡の単調性を

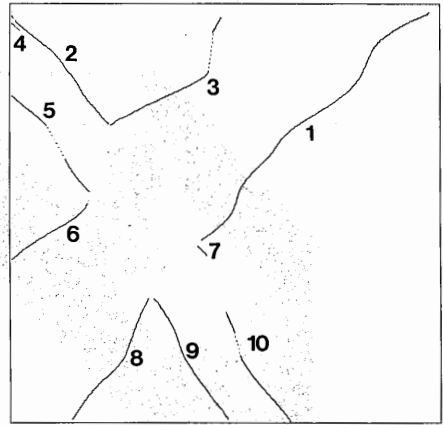


図9 断片的な運動軌跡の接続
Fig. 9 Global connection of loci.

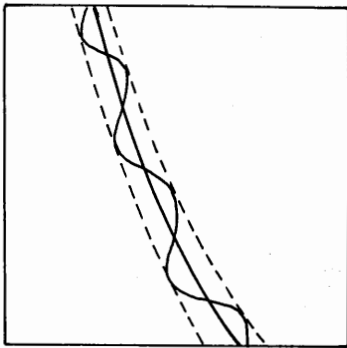


図8 時空間断面画像上における歩行者の運動軌跡のモデル
Fig. 8 Locus of the passerby model on the temporal-spatial cross-section image.

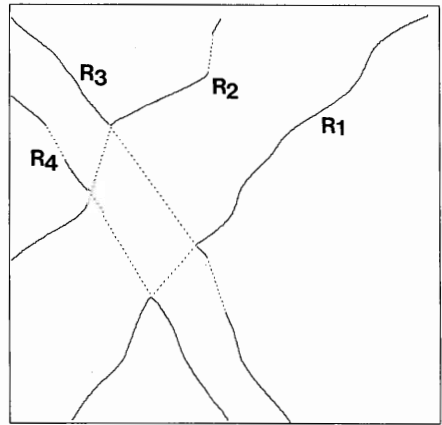


図10 右時空間断面画像上の運動軌跡
Fig. 10 Loci on the right temporal-spatial cross-section image.

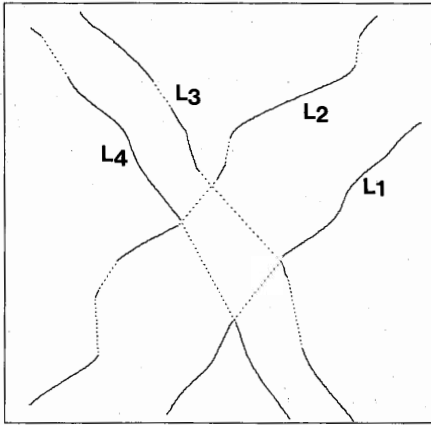


図 11 左時空間断面画像上の運動軌跡
Fig. 11 Loci on the left temporal-spatial cross-section image.

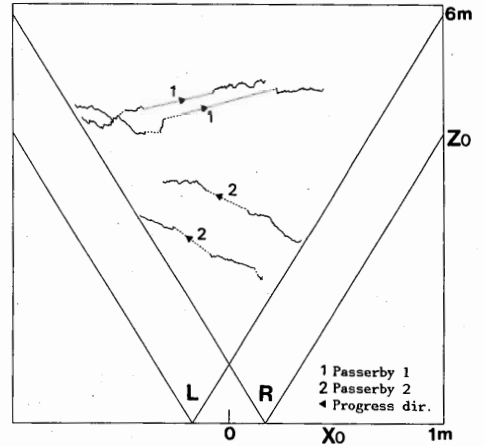
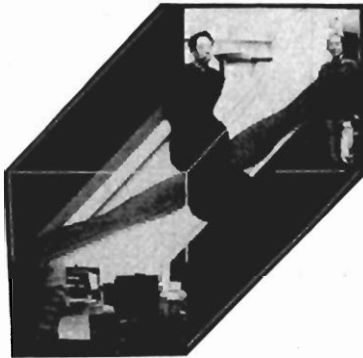
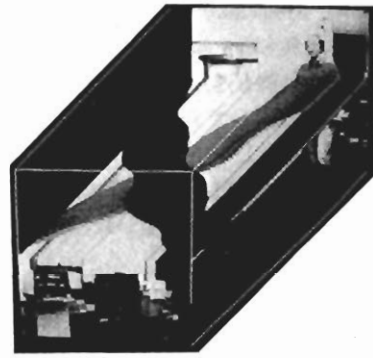


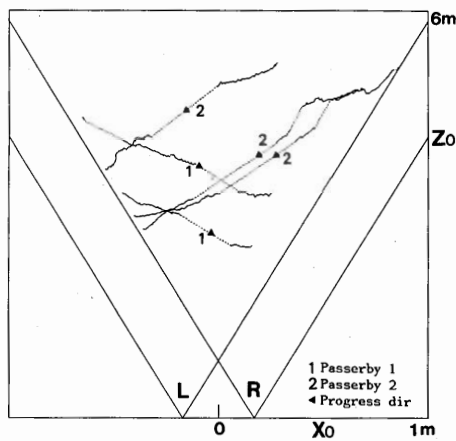
図 12 例 1 における 2 人の歩行者の移動軌跡
Fig. 12 3D loci of passersby(plan view).



Left 3D temporal-spatial image



Right 3D temporal-spatial image



3D loci of passersby (plan view)

図 13 例 2 : 互いに交差する 2 人の歩行者の追跡
Fig. 13 Example 2 : Two passersby cross.

使って断片的運動軌跡の接続を行い運動軌跡を求める。図7の例を対象に処理を行った。途中経過が図9に示されている。既に局所的結合を終え、運動軌跡3と5と10が大域的結合による断片運動軌跡の接続結果である。つなぎ目が点線の線分で示されている。さて、運動軌跡8を延長する場合を考えてみよう。運動軌跡の時間一価性から、候補は1, 2, 3, 4, 5, 6, 7である。明るさの不変性から3, 5, 6は除外される。残りの1, 2, 4, 7のうち軌跡の単調性を満たさないものは2, 4, 7である。従って、運動軌跡8に接続可能な軌跡は1のみということになる。同様に運動軌跡6に対しての候補は2, 3, 4, 5であるが明るさの不変性を満たさない2と4が脱落する。更に単調性により5が除外され、結局6に接続するのは3のみとなる。さて運動軌跡10に接続可能な候補は2, 4,

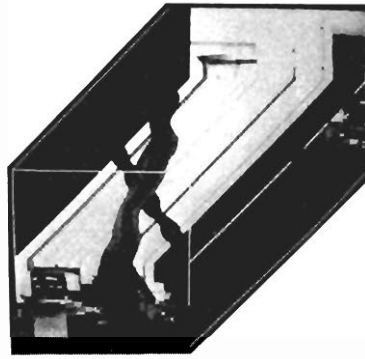
5, 7および3と6の接続軌跡である。ここで、1は既に接続が終わっているので対象とはならない。二つの接続基準を適用すると、2, 4, 7が接続可能となるが、この中で線分接続したときの K 曲率値が最も小さいものとして7を選択する。同様に7は2と接続可能となり10, 7, 2が1本の運動軌跡となる。運動軌跡9は5あるいは4と接続可能であるが、やはり最小 K 曲率値をもつ5と接続される。5の接続候補として4が挙げられるが、これは軌跡の時間単調性を満たさない。接続の可能性がなくなったところで最終的な運動軌跡図(図10)を得る。但し、短い軌跡(20フィールド以下、ここでは4)を除外する。

4.3 運動軌跡の対応付けと歩行者の追跡

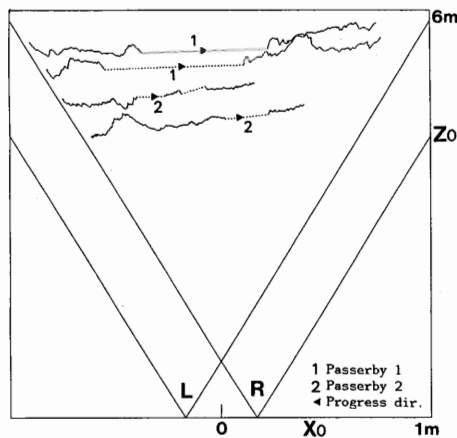
図11と10は、それぞれ図6の左右の時空間断面画像から抽出された運動軌跡である。これらの運動軌跡



Left 3D temporal-spatial image

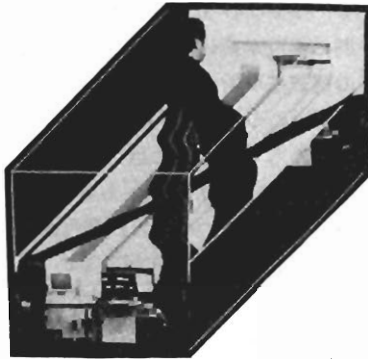


Right 3D temporal-spatial image

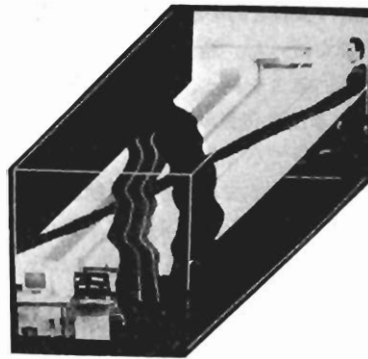


3D loci of passersby(plan view)

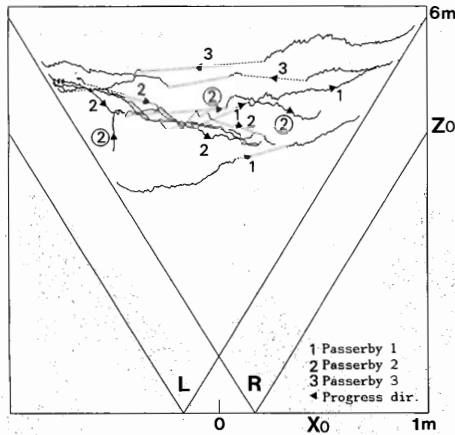
図14 例3: 追越しをする歩行者の追跡
Fig. 14 Example 3: One passerby passes another.



Left 3D temporal-spatial image



Right 3D temporal-spatial image



3D loci of passersby (plan view)

図 15. 例 4 : 3 人の歩行者の追跡
Fig. 15. Example 4 : Three passersby cross.

は 2. の例と全く同様に、時間エピソード拘束によりユニークな対応付けが得られる。この対応付けから歩行者のシーン中での移動軌跡を求めたのが図 12 である。L と R はそれぞれ左右カメラの視野を表し、矢印で進行方向が示されている。対応先が隠されている場合には位置が不確定であるので点線で示されている。

図 13 の例では、歩行者 2 の輪郭と表面模様の運動軌跡はエピソード拘束だけでは対応先が一意に定まらず、対応の順序が逆転しないことを前提に一意な対応付けを得ている。

図 14 は歩行者が他の歩行者を追い抜いている例である。

図 15 は 3 人の歩行者の追跡例である。ここでも部分的に、エピソード拘束に加え、対応付け順序の保存を条件に一意な対応付けを得ている。なお、歩行者 2 の移動軌跡②は同一点の動きを表してはならず、初期時

刻では歩行者 1 の軌跡である。これは重なり合っていた歩行者 1 と 2 の分離が緩やかであったため、時空間断面画像上での両者の運動軌跡の区別が難しかったためと考えられる。

5. むすび

ステレオ動画像から生成される左右の時空間断面画像を対応づけることにより、対象物体の 3 次元追跡を行う手法を提案した。結果として、従来の静止ステレオ画像処理で問題となっていた多重対応やオクルージョンなどの問題が(時間)エピソード拘束のみで解決されることが明らかになった。

本手法は時空間断面画像対の一つの解析法であるが、この画像対を 1 対のステレオ画像とみなして、そのパターンの照合を行うことも考えられている⁽¹²⁾。

運動軌跡の画像化手法は、本論文で述べたような静

止カメラで見た物体の追跡⁽¹⁾や移動カメラからの静止シーンの3次元計測^{(13)~(15)}などに応用されてきた。残るのは、移動カメラから見た動きを含むシーンの解析が考えられるが、これについては別の機会に述べる。

謝辞 本研究に御支援頂きました大島正毅視覚情報研究室長ならびに竹内俱佳電通大教授に感謝致します。また、実験に御協力下さいました、横矢直和、喜多泰代、植芝俊夫、杉本和英の諸氏に感謝致します。

文 献

- (1) 山本正信：“画像化された運動軌跡による動画処理”，情報学論，22, 5, pp. 442-449 (昭56).
- (2) B. D. Lucas and T. Kanade：“An iterative image registration technique with an application to stereo vision”，7-th IJCAI, pp. 674-679 (1981).
- (3) D. Marr and T. Poggio：“Cooperative computation of stereo disparity”，Science, 194, 15, pp. 283-287 (1976).
- (4) 大田友一，金出武雄：“走査線間の整合性を考慮した2段の動的計画法によるステレオ対応探索”，情報処理，26, 11, pp. 1356-1363 (昭60-11).
- (5) W. E. L. Grimson：“Computational experiments with a feature based stereo algorithm”，IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-7, 1, pp. 17-34 (1985).
- (6) 富田文明，高橋裕信：“ステレオ画像の境界線表現に基づく対応”，信学論(D)，J71-D, 6, pp. 1074-1082 (昭63-06).
- (7) 関 靖夫：“セグメントの形状を拘束とするステレオ対応付け”，信学論(D)，J71-D, 11, pp. 2357-2369 (昭63-11).
- (8) M. Jenkin and J. K. Tsotsos：“Applying temporal constraints to the dynamic stereo problem”，Comput. Vision, Graphics & Image Process., 33, pp. 16-32 (1986).
- (9) A. M. Waxman and J. H. Duncan：“Binocular image flows：Steps toward stereo-motion fusion”，IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-8, 6, 715-729 (1986).
- (10) A. M. Waxman and S. S. Sinha：“Dynamic stereo：Passive ranging to moving objects from relative image flows”，IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-8, 4, pp. 406-412 (1986).
- (11) 木村 賛，神谷正明：“速度変化に伴いヒトの歩行がどう変わるか”，バイオメカニズム6, pp. 69-79, 東大出版会 (昭57).
- (12) R. Cipolla and M. Yamamoto：“Image sequence analysis of human motion”，ETL Technical Report, TR-88-11 (1988).
- (13) 金子 透：“物体の準静的回転運動による3次元形状入力”，信学技報，PRL85-22 (1985).
- (14) 山本正信：“連続ステレオ画像からの3次元情報の抽出”，信学論(D)，J69-D, 11, pp. 1631-1638 (昭61-11).
- (15) R. G. Bolles, H. H. Baker and D. H. Marimont：“Epipolar-plane image analysis：An approach to determining structure”，Int. J. Comput. Vision, 1, pp. 7-55 (1987).

(昭和63年12月27日受付)



山本 正信

昭48九工大・工・制御卒。昭50東工大大学院修士課程了。同年電子技術総合研究所入所。以来、動画処理、知能ロボットなどの研究に従事。現在、同所知能システム部、視覚情報研究室主任研究官。工博。62年度情報処理学会研究賞受賞。情報処理学会、ロボット学会等各会員。



ロベルト チポラ

1984年英ケンブリッジ大・工卒。1986年米ペンシルバニア大、1988年電気通信大学院修士課程了。現在、英オックスフォード大学大学院博士課程在学中。コンピュータビジョン、ロボティクスなどの研究に従事。IEEE等各会員。