

マルチプロセッサシステムを用いた複数物体の運動検出

中井 宏章 福井 和広 久野 義徳
(株) 東芝 総合研究所

概要

本報告では、2D 画像上での歩行者などの運動物体を、動画像処理によって自動抽出する一手法を提案する。この手法は、通常の室内シーン中を通行する複数の歩行人物を同時に検出し、その移動経路を抽出できるものである。変化検出・追跡・運動解釈の3段階の処理によって構成されており、各々が並列同時に処理を行なうことにより互いの処理の失敗を訂正し、人物が交差して行きかう場合などにおいても安定に移動経路を検出することができる。また、当所で開発した並列画像処理システム“ビジョンプロセッサ”を用いて、リアルタイム処理が可能な動物体検出システムを試作した。その処理構成と、室内を通行する複数人物の移動経路を抽出する実験の結果を報告する。

Detection of Moving Objects Using A Multi Processor System

Hiroaki NAKAI Kazuhiro FUKUI Yoshinori KUNO
Research and Development Center, TOSHIBA Corporation.

Abstract

In this paper we propose a method for detecting human like moving objects in time-varying images. We show how it is possible to detect the image loci of people moving in ordinary indoor scenes. The method consists of three processes: changing region detection, moving object tracking and movement interpretation. The processes are executed in parallel so that each one can recover the others' errors. This ensures the reliable detection of the loci in complicated cases such as movement across complicated backgrounds. We have built a trial detection system using a parallel image processing system, the Vision Processor. The details of the trial system and experimental results of walking person detection are described.

1 まえがき

撮影した画像中から人物などの動物体領域を抽出して運動を理解するという事は、コンピュータビジョンの課題として大変興味深いものの1つである。と同時に、近年、TVカメラを使って歩行人物や自動車を自動検出したいというニーズが高まるにしたいが、動画像処理技術の研究が再び盛んになってきている。動画像シーケンスから運動物体を検出し運動を解釈しようという試みは、かなり以前から数多く行なわれている [1][2] が、TVカメラで通路を撮影し通行する人物の数を計測する、といったごく普通に実現が望まれる課題でさえも、真に実現することは大変難しい。計測対象とする通路の照度などの撮影条件がほぼ一定に保たれており、撮影された人物像の大きさや色などの属性にばらつきが少なく、通行方向もほぼ決まっているといった条件が成立する場合においては、フレーム間差分といった単純な画像処理手法を用いるだけでも実現可能であろう。だが、これらの条件が成立し得るのは、実験のため意図的に作られた環境内のみである。

連続画像の中から動物体を抽出する手法が、最近多く報告されている。例を挙げれば、背景画像との差分を用いて動物体領域を抽出する手法 [3]、微分画像を用いて人物を検出する手法 [4]、2次元時空間画像をから歩行人物を計数する手法 [5]、オプティカルフローなどの motion 情報から動物体領域を抽出する手法 [6]、active contour を用いる手法 [7]、時空間フィルタを使った領域分割により動領域を抽出する手法 [8] など、そのアプローチの方法も実にさまざまである。ところが、複雑な処理を用いても、歩行人物などの動物体領域の抽出に常に成功するとは限らない。さらに、複数の動物体が互いに遮蔽したり、照度変化などに起因する種々の雑音が撮影画像に重畳されるような一般のシーンを対象とした場合、動物体領域を正確に抽出することは大変困難なものとなる。

ここでは、動物体領域を正確に切り出すということには拘らず、個々の運動物体の存在を検知し、物

体全体の大まかな移動経路を抽出するという事に問題を限定する。よって、動領域分割の成否や精密さを問題とするのではなく、撮影環境条件や物体の運動形態によって左右されることのない、安定な動物体検出アルゴリズムについての検討を行なう。しかし現在のところ、どのように複雑な画像処理手法を用いても、ただ一つの処理で自由に運動する複数物体の移動経路を検出することは困難である。したがって処理を複数に分け、これらを絶えず動かす、すなわち見続けることにより、全体として正しい結果を出す方式を提案する。ある時点の個々の処理には、失敗する可能性がかなりあるかも知れないが、常に処理を続け、そして他の部分と相互に補い合うことにより、全体として失敗を少なくしようという考えである。本報告ではこの方針に基づき、処理を変化検出・追跡・運動解釈の3段階に分け、それらを絶えず実行することにより、安定した運動物体検出を実現した。また、当所で開発した並列画像処理システム“ビジョンプロセッサ” [9][10] を用いて、歩行人物などの複数物体の運動を実時間に検出できるシステムを試作した。ここにそのシステムの構成と一般の屋内シーンでの複数の歩行者を同時に抽出する実験の結果を報告する。

2 システム構成と処理方法

2.1 処理の構成

本報告では厳密な動物体領域抽出を実現するのではなく、複数の動物体を同時に検出でき、その大まかな移動経路を抽出し得ることを目的とする。これを、変化検出・追跡・運動解釈の3段階の処理構成によって実現する。具体的には、まず変化検出処理によって撮影画像中のすべての変化を検出し、追跡処理によって検出された画像領域が運動物体か否かを検証し、さらに運動解釈処理を行なうことによって動物体の適正な移動経路を得る、というものである。ただし、各々の段階の処理が常に成功するものとは期待しない。というのも、上の各処理にどのような手法を用いても「完全」な結果は望めないから

である。しかし、変化検出・追跡・運動解釈の3つの処理を常に同時に実行することにより、一部の処理の失敗を他の処理で訂正することが可能である。例えば、ある時点では物体を追跡する処理に失敗しても、その直後に変化検出処理でその物体を検出できれば、さらに追跡処理を行なうことが可能であり、追跡経路が途中で途切れてしまっても、全体の追跡結果から改めてその物体に応じた移動経路を抽出し直すことができる。また別の時点では、変化検出処理によって動物体領域の一部しか検出できなくても、引続き変化検出処理を行なうことによって全体を検出できる可能性があり、また追跡処理の結果から動物体の領域を推定することも可能である。このように、複数の処理を常時独立に実行することにより、一般の複雑な条件下のシーンに対しても、互いの処理の失敗をカバーできる安定な動物体検出処理を実現できる。

変化検出と追跡の各々の画像処理は画像処理計算機上に、画像処理制御部と運動解釈処理はワークステーション上にインプリメントし、実時間検出が可能な動物体検出システムを試作した。以下の項でハードウェアと3段階処理の詳細を述べる。

2.2 ハードウェア

画像処理には、当所で開発した画像処理用並列計算機“ビジョンプロセッサ”（以下、VPと略す）を用いた。この計算機は、画像入出力モジュール、ローカルプロセッサ (MC68030) 16台、高速画像バスで構成されており、動画像の実時間処理を目指して開発された。また、各々のローカルプロセッサはホストコンピュータ (東芝 AS3160) から制御されると同時に、ホストからプリロードされたソフトウェアにより画像処理を行なう。このほかにも、画像処理から得られた結果の解釈処理にワークステーション (東芝 AS4075) を用いている。全体のハードウェア構成を図1に示す。

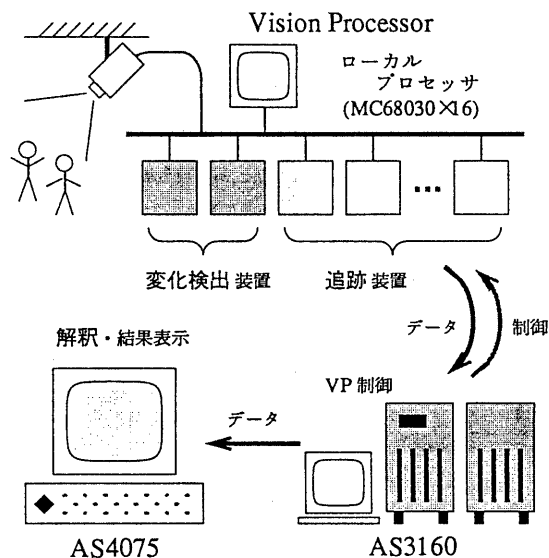


図1: ハードウェア構成

2.3 変化検出部

画像シーケンスの中の動物体による変化を検出する方法としては、フレーム間の単純な差分を用いる方法、微分画像の差分を用いる方法、オプティカルフローを算出する方法などが挙げられる。後者の方法ほど高い検出精度を期待できるが、計算コストが高くなる反面、変化検出処理だけで動物体に正確に対応した画像領域を得るのは困難である。ここでは、処理の失敗が起こることを当然の仮定としているので、簡便だが高速に実行できるフレーム間差分をベースとした方法を変化検出処理として用いる。

前述のVPのローカルプロセッサ2つを用いて、画像全画面中 (512×512画素) の変化検出処理を行なう (図2)。各々のプロセッサは画面の左右半面を担当する。連続するフレーム間での差分の後に二値化処理を行ない、次に画像のxおよびy軸方向に投影処理を行い、これによって得られたヒストグラムをさらに閾値処理することによって、変化領域の位置を得る。投影処理のみでは検出の正確さに欠けるが、このように高速に行なえる方法の方が、後述の追跡処理での失敗を少なくすることができる。

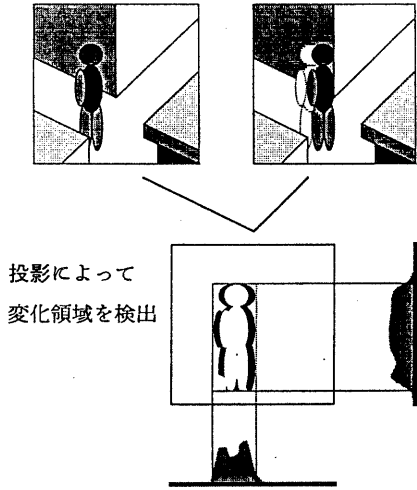


図 2: 変化検出処理

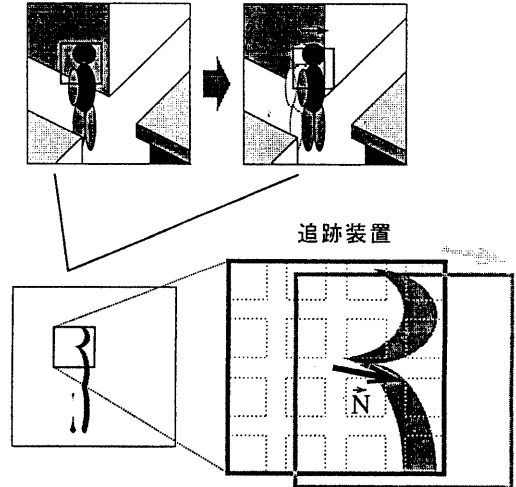


図 3: 追跡処理

2.4 追跡部

追跡部は、動物体を追跡する装置と、これら追跡装置を制御する処理とから成っている。

2.4.1 追跡装置

画像中の動物体を追跡する方法として、テンプレートマッチング法、相関法、特徴追跡法、active contourを用いる方法、ニューラルネットを用いる方法など、実にさまざまな方法が報告されている。本報告ではフレーム間差分を用いた簡略な方法を用いて、普通の演算装置 (MC68030) で動物体の実時間追跡を行なう (図 3)。VP ローカルプロセッサの残り 14 台を用いて、検出された変化を追跡する処理を行なう。ローカルプロセッサは各々独立に追跡処理を実行し、全画面の中から制御部によって指定された数十画素四方の追跡領域内の動物体を追跡する。追跡領域内にはさらに数画素四方の小領域が 16 箇所、格子状に設定されている。小領域 R_i 内で、前時点と現時点との画素輝度差の総計を S_i 、追跡領域中心からその小領域中心までの位置ベクトルを \vec{P}_i とした場合、追跡装置の移動ベクトル \vec{N} を次式によって決定する。

$$\vec{N} = k \sum_{i=1}^{16} (S_i \cdot \vec{P}_i) \quad (k \approx 1)$$

この式の右辺では追跡領域内の変化部分の重心位置をサンプリングして求めており、その位置に追跡装置の中心を移動させることにより変化領域の追跡が行なえる。また、この小領域内で総和を求めることによって画像の平滑化を行なうとともに、重心計算のコストも削減することができる。

2.4.2 追跡装置制御部

追跡部には上述の VP の他に、ワークステーション上で処理を行なっている追跡装置制御部がある。ここでは、変化検出処理の結果を受け取って、

1. 未追跡の変化領域に対し、追跡装置を配置して追跡を実行させる
2. 追跡不能の追跡装置の処理を終了させる

という制御処理を行なうと同時に、追跡処理の結果である位置座標を各追跡装置から受け取り、次述の運動解釈部へと転送する処理も行なう。追跡装置の時間平均した移動量が 0 に近い場合には、追跡に失

敗しているか、検出された変化領域が動物体によるものではないとして、追跡不能と判断し、追跡処理を中止する。

2.5 運動解釈部

本報告では、前述の変化検出および追跡処理において失敗が起こるのは当然と仮定し、一部が失敗しても全体として正しい結果を出せるような方式を検討していることは前にも述べた。変化検出・追跡と同時に運動解釈を行なうことにより、物体の移動経路を適正に抽出できる結果解釈方法を考案した。

運動解釈部は、候補抽出部と対応づけ部の2つで構成されている。以降で詳細を述べる。

2.5.1 候補抽出部

追跡装置群の追跡結果は、まず候補抽出部に送られる。ここでは追跡処理結果から運動物体の候補を抽出する。まず、各々2つの追跡装置*i, j*間に仮想的な結合を想定し(図4)、次式を用いてその結合の強度を算出する。

$$s_{ij}(t) = K_p(t) \cdot F\{p_i(t) - p_j(t)\} + K_d(t) \cdot G\{d_i(t) - d_j(t)\}$$

この式において、*s*は結合強度、*t*は追跡時間、*p_i*と*d_i*はそれぞれ追跡装置*i*の位置と移動方向を表す。*F*および*G*は各々、引数である追跡装置の位置と移動方向が近いときほど大きな値となる関数であり、*K_p*および*K_d*は追跡開始からの時間により変化する係数である(図5)。結合強度が算出された後、既定のしきい値以上の強度を持つ結合によって継れている追跡装置群を1つのグループとして分離し、各々のグループを運動物体の候補とみなして、次の対応づけ部へ動物体候補リストを転送する。

2.5.2 対応づけ部

候補抽出処理から得られた動物体の候補同士を時間軸方向に対応づけることにより、各時点での候補の位置情報から物体の移動経路を表すデータ列に変

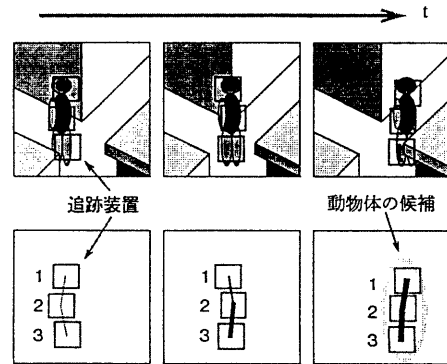


図4: 運動解釈処理1: 動物体候補抽出

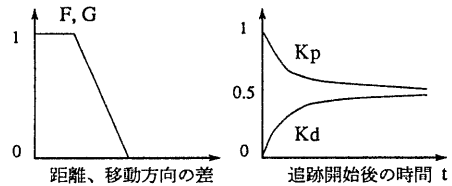


図5: 関数 *F, G* と係数 *K_p, K_d* の変化曲線

換する処理が必要である。この処理を対応づけ部で行なっている。

各フレームで得られている複数点の位置情報から、運動の滑らかさを用いて推測や対応づけを行ない、点の移動経路を得るという手法がいくつか報告されている[11][12]。本報告でも同様に、運動の滑らかさを用いて動物体候補の対応づけを行ない、候補の移動経路を得るが、方法は異なる。前述の処理から得られた動物体候補には複数の追跡装置が含まれている。まず個々の追跡装置同士で対応得点を計算し、その合計から動物体候補同士の対応値を求める。異なる時点での候補中の追跡装置が滑らかに繋がられるようであると、その候補同士の対応値に大きな点数を得点として与える。候補*x*と*x'*に含まれる追跡装置を各々*m_i*、*m'_j*、候補*x*と*x'*との対応値を*M_{x,x'}*とすると(図6)、

3 実験結果

3.1 WS 上での予備実験

実際のシステムを試作する前に、建物内の通路を2人の人物が交差して通行する15フレームの連続画像を用いて、追跡および運動解釈アルゴリズムの妥当性を検証するための予備実験を行なった。結果を図7に示す。上段の a) ~ c) には、開始・交差・分離の3時点での追跡装置の位置を、画像中の矩形で表している。下段 1) ~ 5) は運動解釈結果であり、同じ印 (○, □, △) の追跡装置グループが各々の時点での運動物体の候補である。追跡処理が行なわれると、運動によってグループ分けされ2)、さらに候補同士が交差し3)、新たに追跡装置が配置されても4) △印)、結果として適正な候補のみが抽出される5)。最終的には、左から右へ移動する○印のグループと、右から左へと移動する□印のものとの、各々2人の人物に応じた候補が抽出された。

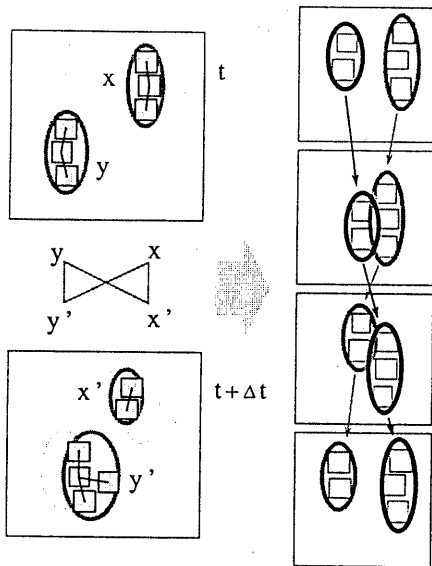


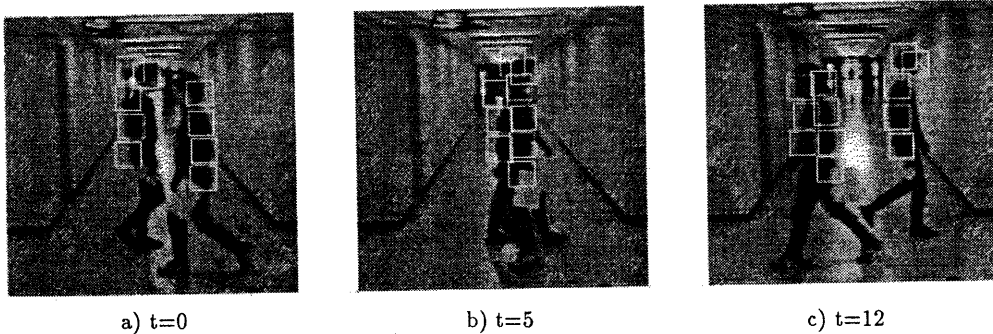
図 6: 運動解釈処理 2: 候補間対応づけ

$$M_{x,x'} = \sum_{i,j} H(m_i, m'_j)$$

という単純な加算式により候補同士の対応得点を得る。関数 H は追跡装置間の位置や運動方向に応じて、それらが近い場合には正、遠い場合には負の値となるステップ関数である。全ての候補同士で対応得点を算出した後、大きな得点の順に、正の得点を持っている候補同士の対応を取って行く。このようにして対応づけられた時間連続の候補群を同じ1つの運動物体候補と記録し直して、その候補の移動経路を抽出する。このとき同時に、長時間対応の取れない孤立した候補や、ある時間範囲での移動経路が極端に小さい候補などは、雑音などの環境変動に起因するものが多いので、検出結果として出力しないようにここで取り除く。

3.2 試作システムでの実験結果

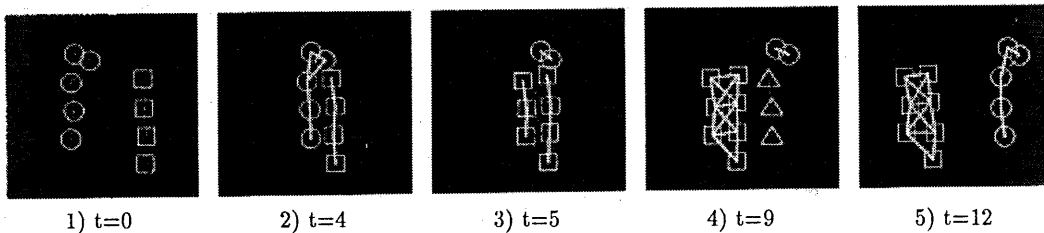
以上で述べてきた処理手法を、図1に示すハードウェア上にインプリメントし、歩行人物の移動経路ををリアルタイムに抽出するシステムを試作した。雑用品を販売する一般的な商店の中を歩行する複数の人物を抽出する実験を行なった結果を図8に示す。図中 1) ~ 3) は、店内を歩行する人物を撮影した連続画像の内の3フレームを抜き出したものである。3) のフレームでの追跡装置の配置状態を 4) 中の矩形に、その時の動物体候補の抽出状態を 5) に、候補の移動経路の抽出結果を 6) に示している。複雑な背景の中を人物が交差して通行したり、人物が物影に隠れたりするような条件で、しかもカメラレンズの焦点距離が短いために人物の大きさがかなり変化するような場合でも、歩行人物の移動経路をほぼ適正に抽出している。



a) t=0

b) t=5

c) t=12



1) t=0

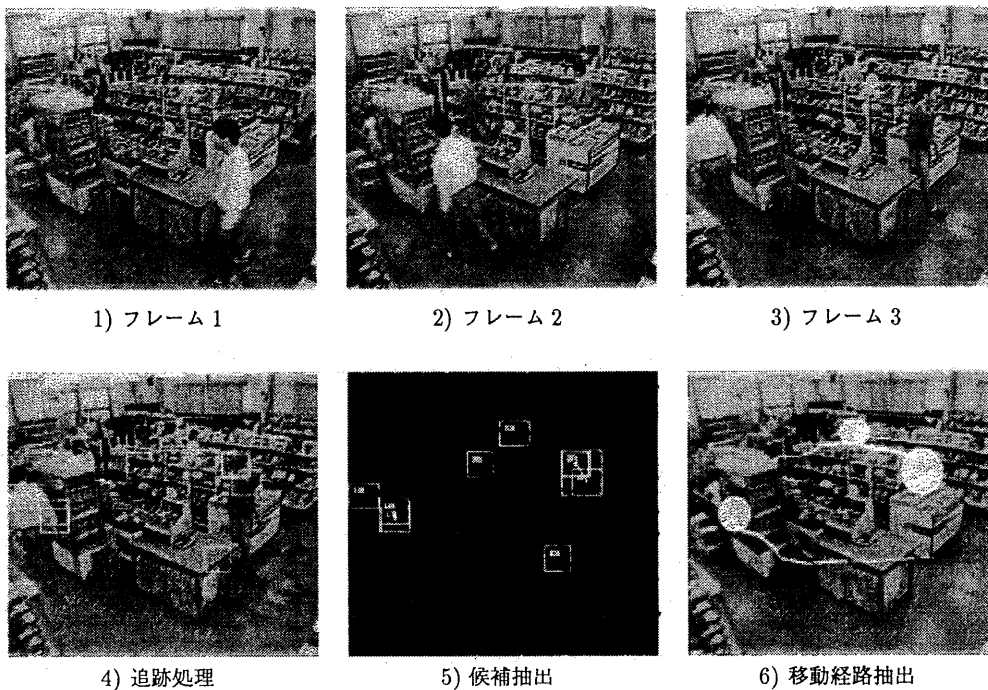
2) t=4

3) t=5

4) t=9

5) t=12

図 7: 予備実験結果



1) フレーム 1

2) フレーム 2

3) フレーム 3

4) 追跡処理

5) 候補抽出

6) 移動経路抽出

図 8: 試作システムでの実験結果

4 まとめ

特に条件を設定せずに、一般の動画像中から個々の運動物体を抽出できる手法を提案し、これを検証するための実時間検出システムを試作した。試作システムを用いて、商店内の客の移動経路を抽出するという実験を行ない、従来法では検出不能のシーンにおいても、本報告の手法が有効であることを確かめた。

本報告で用いている変化検出方法、物体追跡方法は予め失敗が起こることを仮定した簡便なものなので、これらをより信頼性の高い手法に置き換えることにより、システム全体の信頼性が高くなることが見込まれる。具体的には、まず追跡方法として特徴追跡法や active contour を用いることを検討している。また運動解釈処理においても、運動の滑らかさのみを手掛かりにするのではなく、上記の追跡法から得られるであろう追跡領域の特徴を用いて対応づけ処理を行なうことにより、信頼性の向上が期待できる。さらには今後、検出された動物体の存在するらしい領域に対し、何が存在するのかを調べるようなプロセスを起動して、動物体の分類を行なえるようにすると同時に、屋外のような環境変動の激しい環境条件のもとでの適用も検討している。

参考文献

- [1] Martin, Aggarwal: Dynamic Scene Analysis: A Survey, Computer Graphics and Image Processing, Vol.7, No.3 (1978)
- [2] 浅田, 谷内田, 辻: 運動物体の検出と追跡, 信学論, Vol.J62-D, No.6 (1979)
- [3] 川端, 谷藤, 諸岡: 移動物体像の抽出技術, 情処学論, Vol.28, No.4 (1987)
- [4] 佐久間, 伊東, 増田: フレーム間差分を用いた侵入物体検出法, テレビ学技報, Vol.14, No.49 (1990)
- [5] 間瀬: x-t 時空間画像を用いた歩行者の計数, 信学技報, IE90-43 (1990)

- [6] 塩, ス克蘭スキー: 動画像からの歩行者の検出, 情処研報, CV75-5 (1991)
- [7] 高畑, 今井, 辻: アクティブチューブによる非剛体物体の抽出と追跡, 情処研報, CV76-8 (1992)
- [8] 田中, 白井: 時空間フィルタを用いた運動領域の抽出と物体追跡への適用, 情処研報, CV76-20 (1992)
- [9] 久保田, 福井, 石川, 溝口: 物体の認識・識別をめざしたビジョンプロセッサの構想と試作モデルの開発, 信学技報, PRU89-107 (1990)
- [10] Kubota, Okamoto, Mizoguchi, Kuno: Vision Processor for Moving Object Analysis, Proc. of Computer Architecture for Machine Perception '91 (CAMP'91), pp.461-470 (1991)
- [11] Fletcher, Warwick, Mitchell: The Application of a Hybrid Tracking Algorithm to Motion Analysis, Proc. of CVPR'91, pp.84-89 (1991)
- [12] Rangarajan, Shar: Establishing Motion Correspondence, CVGIP Image Understanding, Vol.54, No.1 (1991)