

主成分分析を用いた移動物体認識法の検討

渡辺 睦、武田 信之、小野口 一則

(株) 東芝 関西研究所

概要

動画像から計算されるオプティカルフローを主成分分析することにより得られる情報を用いて移動物体の種別の認識を行う手法について述べる。

このMOROFA法 (Moving Object Recognition by Optical Flow Analysis) においては、まず移動観測系 (カメラ) の動画像から計算されるオプティカルフローに対し、画像を分割して得られた各ブロック毎に動きの消失点を推定する際の残差値を用いてしきい値処理することにより、移動物体の候補領域を検出する。次にこの候補領域に含まれるオプティカルフローが求まった各点の明度、オプティカルフロー方向値を10フレーム分蓄積した結果に主成分分析を行い、得られた第一主成分方向、寄与率を標準値と比較することにより、移動物体の種別の認識を行う。

屋外 (2工場) シーンの画像を用いて乗用車 / 自転車 / 歩行者の識別実験を行い、本手法の有効性を確認した。

A Moving Object Recognition Method by the Principal Component Analysis

Mutsumi WATANABE, Nobuyuki TAKEDA and Kazunori ONOGUCHI

Toshiba Kansai Research Laboratories

This paper presents a new method which can effectively recognize moving objects by analyzing optical flow information acquired from dynamic images. In this MOROFA (Moving Object Recognition by Optical Flow Analysis) method, at first, the optical flow field is detected in image sequences from a camera on a moving observer and moving object candidates are extracted by using the residual error values calculated in the process of estimating the FOE (Focus of Expansion) point. Next, the optical flow directions and intensity values are stored for the pixels involved in each candidate region to calculate the directions and the proportion values of the principal components. Finally, each candidate is classified into a category of object that is expected to appear in the scene by comparing the direction and the proportion values with standard data ranges for the objects which are determined by preliminary experiments.

Experimental results of car/bicycle/walker recognition in real outdoor (two factories) scenes have shown the effectiveness of the proposed method.

1 はじめに

自律移動車を画像認識により制御する移動視覚技術の研究は、1970年頃から人工知能研究の一環として行われてきたが、近年TVカメラが小型かつ安価になり、また、計算機パワーが増大し画像処理が高速化するに伴い、より活発化している。この中で、移動中に出現する障害物を検出、回避するための画像処理は特に重要性が高く[1]高性能な方式が強く求められている。

従来のロボットシステムにおいては、超音波センサーやレーザーによる近接センサを用いて障害物の検出を行っている。この障害物検出が画像処理により実現でき、この種類が認識できれば、例えば人間や自転車が横断しているなら大幅に徐行または一時停止する、乗用車が反対車線を走っているなら警笛を鳴らす、乗用車が反対車線を走っているなら少し減速するに留める、といった柔軟な制御が実現できる。

また、画像を用いた侵入者監視などの産業応用においても、検出した物体の人間か否かの判別など、移動物体の種類を認識するニーズは大きい。

本稿では、動画像から得られるオプティカルフローを主成分分析して得られる第一主成分方向、寄与率を用いて移動物体の種類を認識する手法 MOROFA (Moving Object Recognition by Optical Flow Analysis) について述べる。

2 移動物体の検出手法

2.1 従来の研究動向

従来の画像を用いた物体認識は、形状情報を用いて固定カメラで静止物体の種類、位置姿勢を認識するものが主であり、移動物体の種類を認識する例は少ない。

従来の画像による物体認識の研究は、予め準備したテンプレート(辞書)との照合を行なうことにより認識を行なう手法と、画像から抽出した輪郭などの特徴を基に、物体形状などの構造解析を行なうことにより認識する手法の2つに大別される[2]。

前者は、2次元パターンの認識では標準的な手法であるが、3次元物体の認識においては、視点や物体位置・姿勢の変化に応じて膨大な辞書を準備せねばならない、照合の計算量が多い、などの理由により、これまでほとんど研究されてこなかった。しかし最近、固有空間法を用いたアピアランス空間の3次元物体の位置、姿勢認識法が提案され[3]、注目を集めている。本稿で提案する MOLOFA 法も、基本的にはこの分類に属する方式である。

これに対し、後者の特徴に基づく構造解析による認識においては、見え方の変化に対し認識性能を向上するために3次元形状モデルを利用するモデルベーストビジョンの手法[4]、[5]が一般的である。

これら従来の形状情報に基づく手法に共通する問題点として、認識に用いる特徴抽出、記述、モデルとの照合処理を行わねばならず計算コストが高いことが挙げられる。特徴抽出段階が成功していないと、認識性能は大幅に低下するため、認識対象とする物体がほぼ完全に抽出されている状況でないとなし難い。又、人間のように形状が変化する非剛体の場合は、モデルの記述、照合が困難であるという問題点がある。

これに対し、固定カメラを用いて移動物体を観測、識別する例として、背景画像と時系列的に得られる画像の差分ヒストグラムの変化により移動物体の出現を検知し、2値化した結果の面積、外接長方形の形状により人物かどうかを判定する手法が提案されている[6]。また、画像の下部に設けた部分領域内の明度変化を解析することにより歩調、歩幅を推定し、移動物体が歩行者かどうかを識別する手法[7]や、時系列画像から抽出されたエッジセグメントから時空間ベクトル空間を構成し、標準パターンとDPマッチングを行なうことによりジェスチャの認識を行なう手法[8]などが提案されている。

これらの手法においては、視点固定で定常背景であるという条件に基づき、差分、処理領域の限定などによる計算コストの低減を行なっている。これは実用化の際には有利であるが、屋外などの複雑環境や、観測系が移動し定常背景の仮定が成立しない状況では十分な性能が確保できなくなる危険性がある。

そこで、カメラが移動する場合でも出現する物体の認識が可能である手法として、動画像から得られるオプティカルフローを主成分分析して得られる第一主成分方向、寄与率を用いて移動物体の種類を認識する手法 MOROFA (Moving Object Recognition by Optical Flow Analysis) を考案し、検討を行なった。

2.2 移動物体検出方式

本節では、MOROFA法の処理内容について述べる。全体は図3に示すように、オプティカルフロー計算部、移動物体候補検出部、移動物体認識部の3要素から構成される。

まず、オプティカルフロー計算部では動画像系列より各画素単位の動きベクトル(オプティカルフロー)を求める。次に移動物体候補検出部では、このオプティカルフロー情報を用いて画像中の移動物体の投影領域を検出す

る。最後に移動物体認識部では、この検出された領域に含まれるオプティカルフロー情報を解析することにより、予め与えられた移動物体の種別のどれに該当するかを認識する。

以下この各構成要素の処理内容について説明する。

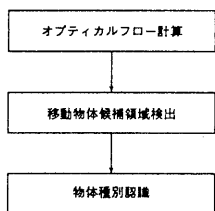


図 1: MOROFA 法の全体構成

2.2.1 オプティカルフローの計算

オプティカルフローの計算法として、Weber-Malik 法 [9] を用いた。これは Horn らにより提案された勾配法の基本式を複数の時空間フィルタ出力を連立させて解くことにより、オプティカルフローを推定する手法である。具体的には、以下の方程式 (1) を解く。

$$\mathbf{A} \cdot \mathbf{v} + \mathbf{I}_t = 0 \quad \mathbf{A} = \begin{bmatrix} I_{1x} & I_{1y} \\ I_{2x} & I_{2y} \\ \vdots & \vdots \\ I_{nx} & I_{ny} \end{bmatrix}, \quad \mathbf{I}_t = \begin{bmatrix} I_{1t} \\ I_{2t} \\ \vdots \\ I_{nt} \end{bmatrix} \quad (1)$$

ただし、 I_{ix} は i 番目の時空間フィルタ出力の微分係数、 $\mathbf{v} = (u, v)$ はオプティカルフローベクトルである。

この解は、特異値分解を用いて直交回帰形の推定を行うことにより、式 (2) で与えられる。

$$\mathbf{v} = -(\mathbf{A}^T \mathbf{A} - \lambda_3^2 \mathbf{I})^{-1} \mathbf{A}^T \mathbf{I}_t \quad (2)$$

ここで、 λ_3 は観測行列の最小の特異値である。現在の実装では、原画像に加え 4 個の時空間フィルタを用い、スケールを 3 段階変化させつつ施して得られる 15 個の連立方程式を解いて $\mathbf{v} = (u, v)$ を求めている。

2.2.2 消失点推定残差法による移動物体候補領域の検出

観測系が静止している場合は、画像中の変化は主に移動物体の出現によると仮定できるため、背景画像との間の

空間差分、または時間差分により移動物体領域を検出することができるが、観測系が移動する場合は背景も同時に変化するため、単純な差分では検出することはできない。

本節では、観測系が直進移動する場合における移動物体候補領域の検出方式について述べる。

移動物体が直進移動する場合、このオプティカルフローを延長した直線は一点で交わる。この点を動きの消失点 (Focus of Expansion; FOE) と呼び、移動方向に応じて定まる定点となる。

例えば図 2 に示す例では、車の移動による FOE が画像の中央付近にあるのに対し、道路左から出現した歩行者による FOE は画像中央左端付近に存在する。つまり、移動物体毎に独立の FOE を有する。従って、移動障害物を含む領域において推定した FOE は、背景のみを含む領域で推定した FOE (車の移動による FOE) と大きく変化することになる。

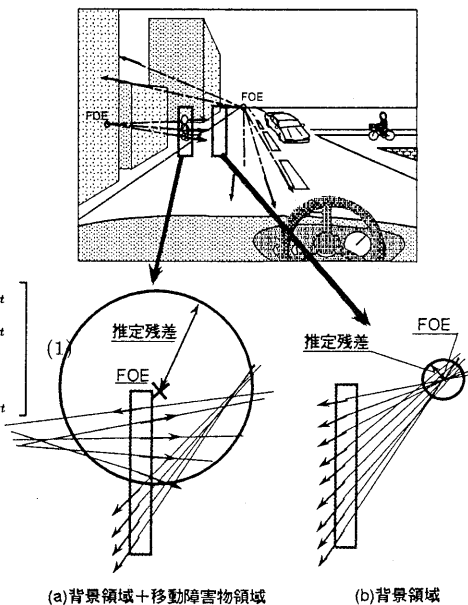


図 2: FOE の推定とその残差

消失点推定残差法は、局所領域毎に FOE を推定する際の推定残差を加算、しきい値処理することにより移動物体を検出する手法である。

図 2 に示す例では、歩行者を含む領域における推定残差が背景における推定残差に比べ大きな値となることを模式的に示している。

具体的には、以下に述べる手順で推定残差を計算する。

各フローベクトルから各フローの直線の方程式を求める。点 (x_i, y_i) においてフローベクトル (u_i, v_i) が得られたとする。この直線の方程式は、

$$a_i x + b_i y + c_i = 0 \quad (3)$$

として与えられる。ここで

$$\begin{aligned} a_i &= -\frac{v_i}{\sqrt{u_i^2 + v_i^2}} \\ b_i &= \frac{u_i}{\sqrt{u_i^2 + v_i^2}} \\ c_i &= \frac{v_i y_i - u_i x_i}{\sqrt{u_i^2 + v_i^2}} \end{aligned} \quad (4)$$

である。係数行列 A 、定数ベクトル c を次式のように定める。

$$A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \end{bmatrix} \quad (5)$$

$$c = [c_1 \quad c_2 \quad \dots \quad c_n]^T \quad (6)$$

ただし、 n は FOE の推定に用いるフローベクトル数である。次に、方程式

$$Ax + c = 0 \quad (7)$$

を解き、

$$\bar{x} = -A^+ c \quad (8)$$

により求まる \bar{x} が FOE 推定結果となる。ここで、 A^+ は A の疑似逆行列である。

局所領域内のフローベクトルに対し、式 (8) を適用することにより推定した FOE を用い、

$$E = \|A\bar{x} + c\|^2 \quad (9)$$

を求めることにより、FOE 推定における推定残差を計算する。

以上述べた処理により求めた推定残差を、局所領域を移動しながら加算していく。更に、局所領域の大きさを変えて加算処理を行なう。現在の実装では、局所領域として 32×32 、 64×64 、 96×96 の 3 通りを用い、移動量は 4 画素に設定している。各局所領域毎にフローベクトルの抽出間隔を変え、各々最大 64 本のフローベクトルを用いて推定残差を計算するように設定した。この結果得られた値を、ノイズ除去の目的で 6 フレーム長のガウシアンフィルタで平均化し、実験的に定めたしきい値により移動物体候補領域を抽出する。

以上述べた移動物体候補領域検出処理の流れを、図 3 に示す。

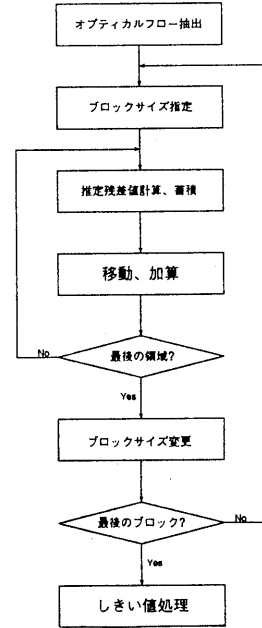


図 3: 移動物体候補領域検出処理の流れ

3 移動物体の認識

前節で述べた状況下で安定に働く移動物体認識法として、抽出された移動障害物候補領域に含まれる画素の明度、及びフローベクトルの方向成分を統計解析して得られる情報を用いて行なう手法を考案した。

ここで、移動物体において同一の運動を行なう部分を運動剛体と呼ぶことにすると、30frame/sec のフレームレートでは、例えば、前方を横切る乗用車は 1 つ（全体のみ、車輪の回転は検出されない）、自転車は 3 つ（自転車、手、足）、歩行者は 5 つ（上半身、両手、両足）の運動剛体が観測される。

検出した移動物体候補領域に移動物体が、ほぼ含まれていれば、主成分分析により得られる第一主成分は移動物体に対応する。また寄与率が大きい程、運動剛体の数が少なく第一主成分で移動物体全体の動きが記述されることを意味する。

主成分分析に用いる説明変数は、処理時間、記憶容量の面から必要最小限のものを選択するべきである。今回は、各運動剛体が独立の明度を持つと仮定し、背景にある道路など非移動物体領域の混入に対処するために、移動障害物候補領域に含まれるオプティカルフローの方向成分、

オプティカルフローが得られた点の明度の2つを用いた。

乗用車のように運動剛体が一つの場合は、蓄積した空間におけるデータは、全体の移動方向にほぼ一致する単一のフロー方向成分を持つ。一方明度は、車体と窓、車輪部分における値をとる。従って、空間内におけるデータの分布は明度成分方向のみに広がり、最大固有値に対応する固有ベクトルの方向である第一主成分の方向は0にほとんど近く、かつ寄与率は大きくなる。これに対し歩行者のように複数の運動剛体から構成される場合、蓄積したデータのフロー方向成分は複数の運動剛体の動きを反映したものととなり、明度は、頭、上半身の服、下半身のズボン、靴など様々な値をとる。従って、データの分布はフロー方向成分、明度成分、両方向に広がるため、第一主成分方向は大きな値となり、逆に寄与率は小さくなる。一方自転車の場合、自転車自体は一つの運動剛体であるが、これを運転する人間はハンドルを操作する腕、ペダルを漕ぐ足など複数の剛体から構成される。従って、蓄積したデータは、乗用車とも歩行者とも異なる独特の分布を示し、第一主成分方向、寄与率に差異が生ずる。以上述べた、(明度、オプティカルフロー方向)特徴空間における分布の様子を、模式的に図4に示す。

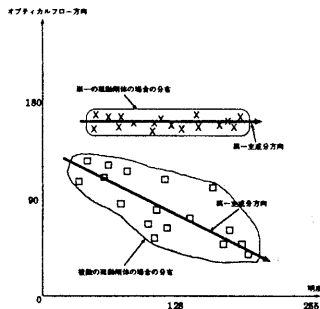


図4: (明度、オプティカルフロー方向)空間におけるデータ分布

図5に、移動物体識別処理の流れを示す。

消失点推定残差法により抽出された移動物体候補領域に含まれる画素のうち、オプティカルフローが求まった点について、このフロー方向と明度値を組にして2次元記憶部に格納する。この処理を既定フレーム数Tについて行う。次に、記憶部に格納されたデータに対して主成分分析を行い、第一主成分に対応する固有ベクトルのx、y成分の比の逆正接値を計算することにより、第一主成分の方向と寄与率を求める。最後に、両者の値が予め定めた範囲に

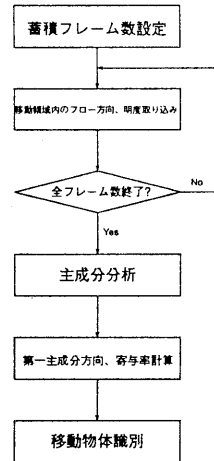


図5: 移動物体の認識処理の流れ

あるかどうか判定することにより、種別の認識を行う。

4 実験

4.1 実験条件

前節で述べた識別方式の有効性を検証するため、2箇所の工場(大阪:シーン1~14、東京:シーン15~25)で収集した移動画像25シーンをを用いて実験を行った。認識する対象は、工場において頻繁に出現する乗用車/自転車/歩行者に未知物体を加えた4種類である。

表1に、各シーンに含まれる障害物の移動方向を示す。

まず、移動障害物候補領域が比較的安定に抽出されている5シーン(シーン1、2、4、6、7)を用いて、第一主成分の方向と寄与率の識別範囲を定めた。また、蓄積フレーム数は実験的に10に定めた。

第一主成分方向と寄与率の範囲を同時に満足する場合に該物体と識別し、いずれにも該当しない場合は未知物体と判定した。

図6、図7、図8に、移動物体候補領域抽出結果(黄色の領域)の例を示す。シーン1-8においては、ほぼ正しく移動物体候補領域の検出が行なえるが、シーン9-25においては一部背景領域の誤検出などが発生している。

図9、図10、図11に、特徴空間への蓄積結果の例を示す。一部背景の誤検出などの原因により、図4に示した分布とは若干異なっているが、運動剛体数の違いは反映されており本手法で認識可能であることが示されている。

表 1: 実験シーンの構成

	移動物体の種類	移動方向
シーン1	乗用車	水平 (右から左)
シーン2	乗用車	水平 (左から右)
シーン3	乗用車	右折
シーン4	自転車	横断 (左から右)
シーン5	自転車	斜め横断 (左奥から右手前)
シーン6	歩行者	接近 (道路左側)
シーン7	歩行者	横断 (左から右)
シーン8	歩行者	横断 (右から左)
シーン9	乗用車	先行車 直進
シーン10	乗用車	左折
シーン11	乗用車	右折
シーン12	乗用車	先行車 直進
シーン13	自転車	接近 (道路左側)
シーン14	自転車	斜め横断 (左手前から右奥)
シーン15	自転車	水平 (右から左)
シーン16	自転車	斜め横断 (右手前から左奥)
シーン17	自転車	遠離 (道路左側)
シーン18	自転車	斜め横断 (右奥から左手前)
シーン19	歩行者	遠離
シーン20	歩行者	斜め横断 (左奥から右手前)
シーン21	歩行者	斜め横断 (左奥シーンから右手前)
シーン22	歩行者	水平 (左から右)
シーン23	歩行者	斜め横断 (左手前から右奥)
シーン24	歩行者	接近 (道路左側)
シーン25	歩行者	遠離 (道路左側)

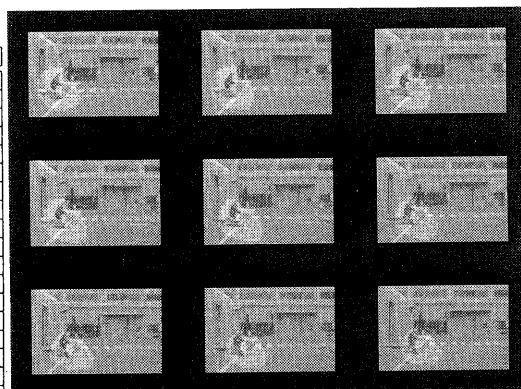


図 7: 移動障害物候補領域抽出結果 (自転車、シーン5)

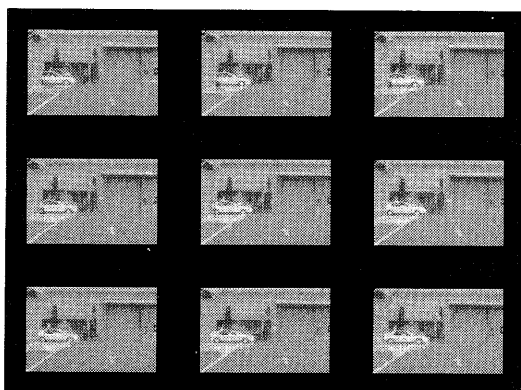


図 6: 移動障害物候補領域抽出結果 (乗用車、シーン2)

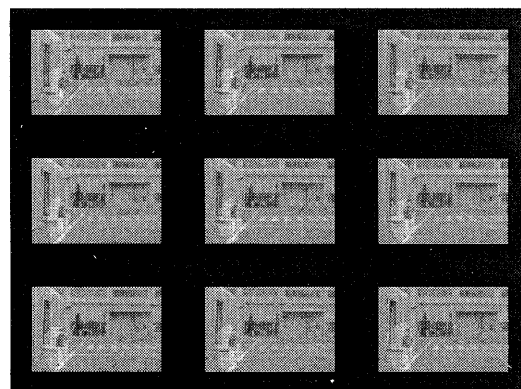


図 8: 移動障害物候補領域抽出結果 (歩行者、シーン6)

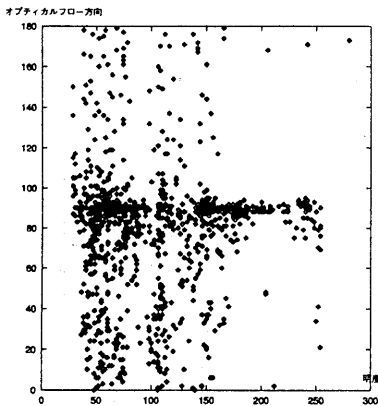


図 9: 特徴空間への蓄積結果例 (乗用車、シーン 2)

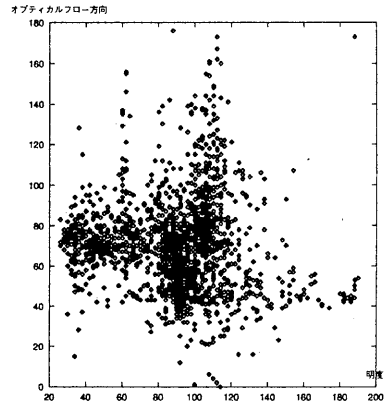


図 11: 特徴空間への蓄積結果例 (歩行者、シーン 6)

表 2 に、25 シーンの識別結果を纏めて示す。予備実験に用いた 5 シーンを含む 21 シーンに関し、正しく識別が行えた。一方、シーン 14、21 は各々、自転車、歩行者が未知物体となり、シーン 13、15 は自転車を歩行者と誤識別した。今回の実験における識別成功率は 84%、誤り率は 8% であった。

5 考察

前章で述べたように、今回実験に用いた 25 シーンにおける 21 シーンで正しく移動物体が識別でき、誤った識別結果となったのは 2 シーンのみであった。また、比較的移動障害物候補領域が精度よく得られたシーンについては、全て正しく識別が行なえ、提案した手法が屋外の移動シーンで適用可能である見通しが得られた。

以下、問題となる可能性のある点について検討を行う。

5.1 観測系の移動軌道上を移動する障害物に対する検出感度低下

消失点推定残差法による障害物検出方式では、原理的に移動物体の消失点と観測系による消失点が近接した場合に検出感度が低下する。つまり、車の正面から対向直進、または前方を先行する移動物体は検出が困難になり、移動物体候補領域の一部欠落、背景領域の混入が発生する。今回の実験におけるシーン 13、15 (自転車) の誤識別は、この原因によるものと考えられる。ただし、観測系

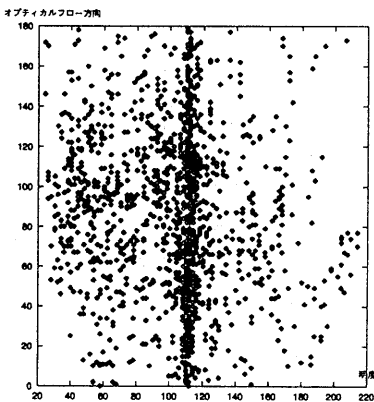


図 10: 特徴空間への蓄積結果例 (自転車、シーン 5)

表 2: 識別実験結果

	移動物体の種類	移動方向
シーン1	乗用車	乗用車
シーン2	乗用車	乗用車
シーン3	乗用車	乗用車
シーン4	自転車	自転車
シーン5	自転車	自転車
シーン6	歩行者	歩行者
シーン7	歩行者	歩行者
シーン8	歩行者	歩行者
シーン9	乗用車	乗用車
シーン10	乗用車	乗用車
シーン11	乗用車	乗用車
シーン12	乗用車	乗用車
シーン13	自転車	歩行者
シーン14	自転車	未知物体
シーン15	自転車	歩行者
シーン16	自転車	自転車
シーン17	自転車	自転車
シーン18	自転車	自転車
シーン19	歩行者	歩行者
シーン20	歩行者	歩行者
シーン21	歩行者	未知物体
シーン22	歩行者	歩行者
シーン23	歩行者	歩行者
シーン24	歩行者	歩行者
シーン25	歩行者	歩行者

と同じ方向に移動する場合でも、オプティカルフロー、移動障害物候補領域がほぼ正しく得られている場合(シーン6、9、12、17、19、24、25)には、正しく識別できている。

5.2 処理時間の高速化

現在は図1に示した各3構成要素を独立に作成しており、最適化は行っていない。処理時間は、エンジニアリング・ワークステーション(75MHz, 88 SPECint92, 102 SPECfp92)上に実装した実験プログラム(C言語)にて、オプティカルフロー計算が約13秒、移動物体候補領域検出が約19秒、移動物体の認識が約1.2秒であった。この移動物体認識の処理時間は、移動障害物候補領域の大きさ、密度により変動するが、本実験例では最も処理時間を要したもので1.35秒であった。

現在、Weber-Malik法による処理で15個の方程式を解いてフローベクトルを求めているが、既に2個の方程式を解くことにより実時間でオプティカルフロー検出を得るハードウェア[10]が開発されており、Weber-Malik法の並列ハード化による実時間オプティカルフロー検出も十分実現可能だと思われる。

また移動可能候補領域検出の部分においては、推定残差の計算が支配的であり、推定ブロック単位の並列化により高速化が可能である。

一方移動物体識別処理においては、処理時間の90%以

上をデータの蓄積に費やしている。この部分は移動障害物検出・識別サブシステムに纏める際移動障害物検出処理と接続、効率化を行うことにより更に高速化が実現できる。

6 まとめ

移動物体の認識法として、抽出された移動物体候補領域に含まれるオプティカルフローから得られる動き方向、明度を識別特徴として選択、蓄積、主成分分析し、この第一主成分方向と寄与率の違いを用いて識別する手法(MO-ROFA)を提案し、工場で収集した実画像系列を用いて乗用車/自転車/歩行者/未知物体の識別実験を行なった。

動き情報を用いた物体認識は、従来の物体認識のように輪郭がほとんど完全に抽出されることを前提としない点、移動障害物の検出に用いたオプティカルフローを用いて効率的に移動物体の識別を行なう点、識別に用いるパラメータが2つと少なく調整が容易な点、などの実用上の利点を持つ。

今後は、識別精度、処理速度を向上するための改良、及び移動画像を用いた性能評価実験を行なっていく。

参考文献

- [1] 渡辺、小野口、久野、麻田：'視差予測ステレオを用いた障害物検出'，電子情報通信学会論文誌，Vol. J73-D-II, No.8, pp.862-870, 1990.
- [2] 村瀬：'CVCV特別報告：コンピュータビジョンにおける技術評論と将来展望(VI) - 固有空間法による画像認識 -'，CV研賞，95-CV-97, pp.59-66, 1995.
- [3] Murase, H. and Nayar, S.K.: 'Visual Learning and recognition of 3-D objects from appearance', International Journal of Computer Vision, Vol.14, pp.5-24, 1995.
- [4] Brooks, R.A.: 'Symbolic reasoning among 3-D models and 2-D images', Artificial Intelligence, Vol.17, pp.285-348, 1981.
- [5] 渡辺、南川編著：「知識システム」，電子情報通信学会，1991.
- [6] Kaneta, M. et al.: 'Image Processing Method for Intruder Detection around Power Line Towers', ICICE E76-D, No.10, pp.1153-1161, 1993.
- [7] 安富、森、清弘：'歩行のリズムに基づく歩行者検出の一手法'，電子情報通信学会論文誌，Vol. J78-D-II, No.4, pp.608-617, 1995.
- [8] 高橋、関、岡：'ジェスチャ動画像のスポットティング認識'，信学技報 IE 93-134, pp.9-16, 1993.
- [9] Weber, J. and Malik, J.: "Robust Computation of Optical Flow in a Multi-Scale Differential Framework," in Proceedings of ICCV'93, pp.12-20, 1993.
- [10] 塩原 他：'リアルタイム・オプティカルフロー・プロセッサ I S H T A R'，画像の認識理解シンポジウム(MIRU'94)予稿集，pp.2-295-302, 1994.