

# 動画画像処理技術を用いた移動体監視システムの開発

Visual Surveillance System by Using Dynamic Scene Analysis

技術本部 見持圭一\*1 塘中哲也\*2  
 名古屋誘導推進システム製作所 日浦誠司\*3 富田文博\*3  
 井上政雄\*4

対象物の形だけでなく動きを抽出する動画画像処理技術により、高度な“機械の眼”の機能を実現できる。本研究ではテレビカメラの時系列画像から屋外シーンの中に含まれる人や車などの移動体を検知し、追跡しながら識別する画像処理アルゴリズムを開発した。背景差分法により背景と移動体領域を分離し、その領域の大きさ、形状、エッジ分布、移動ベクトルのフロー分布と、それらの時間変化に基づく評価関数により移動体を人と車に識別する。本アルゴリズムを並列画像処理ボードに実装した移動体監視システムを実用化し、リアルタイム検知・識別（処理時間0.2s以下）とカメラ旋回による自動追尾機能を実現した。

Recognizing object shape and movement from image sequences is basic to many computer vision applications. We designed a way to detect moving objects and classify them into human beings, vehicles, and other types of interference. Candidate object regions are detected by background image subtraction. Regional image features such as area, shape, edge direction, and optical flow are obtained while tracking the region in the image sequence. An object corresponding to each region is recognized as human being, vehicles, or other interference based on image feature evaluation. This application runs on a parallel image processing board having four digital signal processors. Moving object detection and classification cycle time is within 0.2 s in visual surveillance that tracks moving objects automatically.

## 1. ま え が き

テレビカメラの画像を処理し、シーンに含まれる対象物の種類や位置・姿勢を決定する視覚認識技術は、従来人間が行ってきた様々な作業を自動化する上で必要不可欠である。例えば生産ラインで活躍する産業用ロボットの視覚センサや各種製品の検査装置など工場内の生産設備の高度化に寄与している。これらは静止している対象物の形状を空間的に捕えて画像処理するのに対し、最近では移動する対象物に対しても、空間的かつ時間的に捕えてその動きも認識することが可能な動画画像処理技術の研究開発が活発化している。この動画画像処理の応用分野として注目されているものに画像監視システムがある<sup>(1)</sup>。近年、港湾、鉄道、空港やプラント等の重要施設における監視業務の必要性が重要視され、多くの監視員がテレビカメラの映像を四六時中注視している。しかしこのような監視方法は作業負担が大きくコストも掛かるので、24時間自動監視による省力化が望まれている。

そこで動画画像処理技術を用いてテレビカメラの時系列画像の中から移動する物体を検知し、人や車を識別する移動体監視システムを開発した。本報では、移動体監視システムの構成、移動体検知と識別のための動画画像処理アルゴリズムとシステムの適用事例について述べる。

## 2. システム構成

画像による監視システムの基本的な構成を図1に示す。テレビカメラからの映像はサンプリング時間  $\Delta t$  秒ごとに画像処理装置に入力される。シーン中に人や車などの移動体が存在すると、画像上ではその部分の明るさが時間的に変化する。現在の画像と少し前の画像の明るさの差から、画像が変化している部分のみを移動体候補領域として検出する。この候補領域の位置は時々刻々変化するので、 $\Delta t$  秒前の候補領域と常に対応を付けながらその領域を

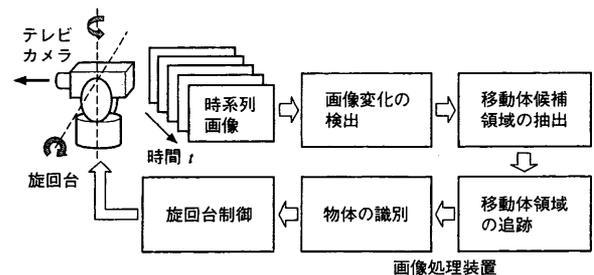


図1 システム構成 テレビカメラからの映像は画像処理装置に入力され、移動体の検知・識別が行われる。  
 System configuration and image processing scheme

追跡する。ある一定時間経過の後、追跡している移動体領域に含まれる画像の形状や特徴を用いて、その移動体を人、車又はそれら以外に識別し、カメラで注視すべき対象を選択する。旋回台の方位角とふ角を制御して選ばれた移動体が画像の中心に来るようにカメラを旋回させる。カメラ静止後は再び画像による移動体検知・追跡とカメラ旋回を続け、一度選択した移動体を見失わないように追尾する。

## 3. 動画画像処理アルゴリズム

### 3.1 移動体検知

画像の中から移動体を含む領域を検出するために、図2に示すような現在の画像と過去の画像を比較する必要がある。その手法として時間差分法と背景差分法がある<sup>(2)(3)</sup>。前者は現在の画像と  $n\Delta t$  秒前 ( $n$  は任意) の画像の明るさの差を求め、その絶対値が大きい領域を移動体領域とするものであるが、移動体領域の中で明るさが一定の部分では検出できない問題がある。一方、背景差分法は移動体を含まない背景だけの画像を記憶しておき、入力画像との明るさの差を求め、その絶対値が大きい領域を移動体領域と

\*1 高砂研究所電子技術研究室主務  
 \*2 高砂研究所電子技術研究室

\*3 誘導・電子機器部電子システム設計課  
 \*4 誘導・電子機器部主査

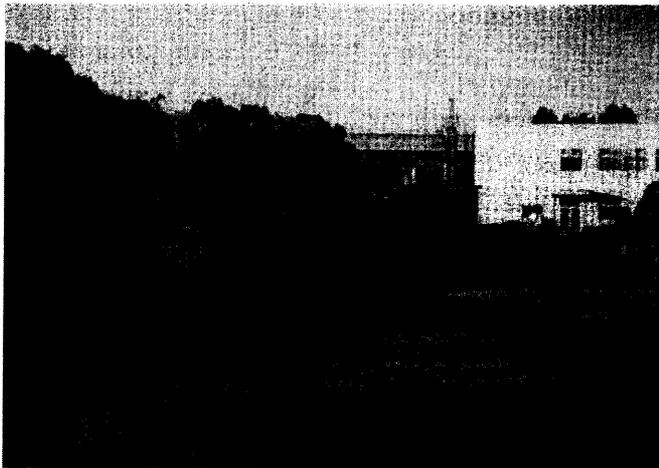


図2 入力画像の例 時々刻々変化するシーンが時系列画像として画像処理装置に入力される。  
Input image



図3 背景差分 ブロック単位で入力画像と背景画像の明るさの差を求める。  
Background image subtraction

するので検知感度が高い。しかし、背景は日照変化やカメラの旋回により時々刻々変化するので背景画像を適宜更新する必要がある<sup>(4)(5)(6)</sup>。

そこで、背景画像は移動体領域と判定された領域以外は画像の明るさ変化の度合に応じた更新間隔で更新する。移動体領域と判定された領域の中では、時間差分の値が十分小さい領域だけ更新を行う。その結果、以前の移動体領域が背景画像に残って、現在の移動体領域が見掛け上、尾を引いたようになる不具合を防ぐことができる。

背景差分の絶対値が大きい領域は移動体以外にも様々な外乱（例えば日照変化、鳥や犬などの小動物、草木の揺れ、波などの自然現象による画像変化）により生じる。また画像ノイズにより画素単位で明るさに変化する。そこで背景差分は一つ一つの画素ごとに扱うのではなく、図3に示すように画像を小さなブロックに分割して行う。ブロック内での背景差分の平均値があらかじめ設定したしきい値より大きい場合にそのブロックの値を“1”とし、それ未満を“0”として2値化する。こうして生成された2値画像に対して、“1”が隣接している領域ごとに分割する。分割された領域はその領域をすべて含む最小の長方形で近似され、その面積があらかじめ設定した範囲内に入る領域だけを移動体候補領域

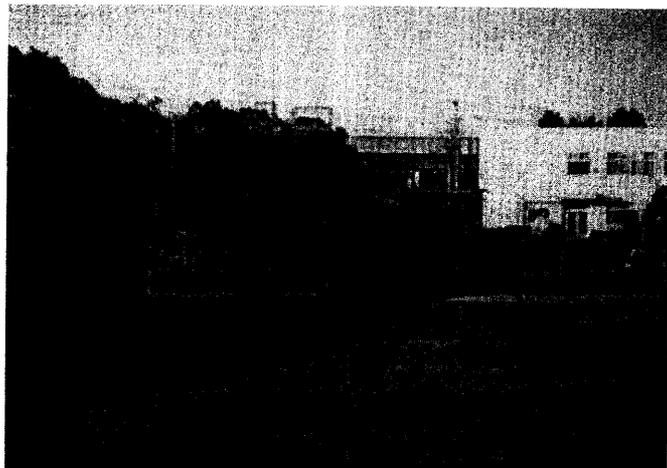


図4 移動体候補領域の抽出 背景差分の絶対値の大きい領域をひとかたまりとして長方形近似し移動体候補領域とする。  
Detecting candidate regions of moving objects

とする。図4に検出した移動体候補領域を示す。人や車以外に草木の揺れによる外乱も検知している。

### 3.2 移動体追跡

真の移動体に対応する候補領域の位置は、ある軌跡を描きながら時間的に変化する。視野から外れるか静止するまである時間存在し続けるはずである。

このような移動体に関する前提条件を利用して、検出された複数の移動体候補領域の中から真の移動体に対応する領域だけを追跡する。具体的な追跡条件を以下に示す。

- (1) 時刻  $t_i$  において検出した  $j$  番目の移動体候補領域  $S_j(t_i)$  が、時刻  $t_{i-1} = t_i - \Delta t$  での移動体候補領域の一つ  $S_k(t_{i-1})$  と一定面積以上の割合で重なっているとき、 $S_j$  の追跡回数  $M_j$  を1増やす。そうでないならばその領域は新しい領域として  $M_j = 1$  とする。
- (2) 時刻  $t_i$  において  $j$  番目の移動体候補領域  $S_j(t_i)$  の追跡回数  $M_j$  が設定した値  $N$  と等しくなれば、その領域は移動体を追跡できなかったとして3.3節で述べる識別を行う。ただし、過去  $N$  回の移動体候補領域の面積とその中心軌跡が時間的に極端に変動するのは外乱であるとして除外する。
- (3) 時刻  $t_{i-1}$  における  $k$  番目の移動体候補領域  $S_k(t_{i-1})$  が時刻  $t_i$  において対応する移動体候補領域が存在しないときは、次の時刻  $t_{i+1} = t_i + \Delta t$  に再度候補領域を探すが、このときも発見できなければこの移動体領域は除外する。

### 3.3 識別のための画像特徴

前節に示したように、時刻  $t_i$  において過去  $N$  回追跡できた移動体候補領域  $S_j (j=1, 2, \dots, M)$  には以下のデータが付随している。

- (1) 移動体候補領域の面積（領域は長方形で近似するので、その縦と横の積）： $A_{jk} (k=1, 2, \dots, N)$
- (2) 移動体候補領域の形状（縦と横の長さの比）： $R_{jk} (k=1, 2, \dots, N)$

これらの値を利用し、抽出した移動体候補領域に含まれている物体を人、車、又は外乱にファジィ手法を用いて識別することが試みられている<sup>(7)</sup>。

しかし抽出した領域が対象物体すべてを含んでいなかったり、逆に背景の一部が含まれていると、正しく識別できない場合がある。そこで、移動体候補領域を追跡中にさらに次に示すような画像特徴を抽出する。

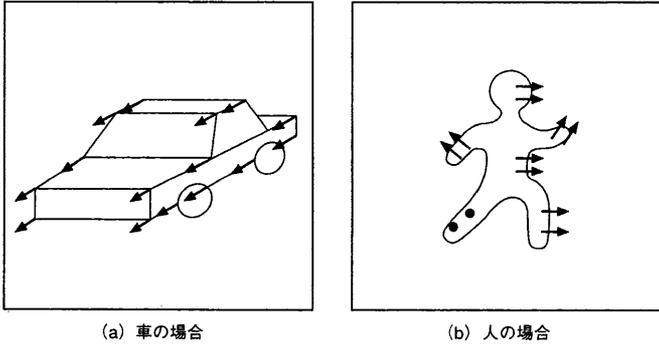


図5 移動体のエッジとフロー分布 車と人ではその輪郭エッジの方向分布や移動ベクトルのフロー分布が異なる。  
Edge direction and optical flow for moving objects

(3) 移動体候補領域のエッジ分布

図5に示すように、車の輪郭を画像の明るさが変化するエッジとして抽出すると、そのエッジは同一方向の直線状のエッジが多く含まれる。一方、人の輪郭には様々な方向のエッジが多数含まれている。そこでエッジを構成する画素のエッジ方向に対するヒストグラム（頻度分布）を作成し、その最大値を与える方向  $\theta_0$  を中心として  $\pm\theta^\circ$  の範囲内に限定したヒストグラムの標準偏差  $E$  を求める。この値はエッジ方向分布のばらつきを定量化するものであり、特定の方向に多くのエッジが存在するほど小さい値を取る。結果として、追跡した過去  $N$  回の移動体候補領域  $S_j$  のエッジ分布  $E_{jk}$  ( $k=1, 2, \dots, N$ ) を得る。

(4) 移動体候補領域のフロー分布

さらに図5に示すように、車と人では、画像の中で明るさ変化が際立った特徴点の移動方向の分布も異なる。車の場合はその領域内のすべての特徴点が同じ方向に移動するが、人の場合は必ずしもすべて同じ方向を向くとは限らない。そこで移動体候補領域内の画像に含まれる多くの特徴点の移動方向（オプティカルフロー）を抽出し、その平均移動方向に対する標準偏差  $U$  によってフロー分布の均一性を定量化する。

このフロー分布を抽出する方法としては様々な方法が提案されているが<sup>(8)</sup>、ここでは計算負荷が大きい方法として時空間こう配法を用いる。時刻  $t$  において位置  $(x, y)$  に存在する画像特徴点を含む小領域  $W$  の画像  $I(x, y, t)$  は、 $\Delta t$  秒後に位置は変化してもその明るさは変化しないという拘束条件を仮定する。

$$I(x, y, t) = I(x + u\Delta t, y + v\Delta t, t + \Delta t) \quad (1)$$

ここに、 $(u, v)$  は求めるべき特徴点の移動ベクトルであり、式(1)をテイラー展開した一次近似式に最小二乗法を適用して次式を得る。

$$\begin{bmatrix} \sum_W I_x^2 & \sum_W I_x I_y \\ \sum_W I_x I_y & \sum_W I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} -\sum_W I_x I_t \\ -\sum_W I_y I_t \end{bmatrix} \quad (2)$$

ここに、添字  $x, y, t$  はそれぞれ画像の明るさ  $I$  の  $x$  方向、 $y$  方向の空間微分と時間微分であり、各特徴点の移動ベクトルはこの連立方程式を解くことで得られる。追跡した過去  $N$  回の移動体候補領域  $S_j$  に含まれる特徴点の平均移動ベクトルに対する標準偏差からフロー分布の均一性  $U_{jk}$  ( $k=1, 2, \dots, N$ ) を得る。

3.4 識別のための評価関数

時刻  $t$  において  $N$  回の追跡に成功した  $M$  個の移動体候補領域  $S_j$  ( $j=1, 2, \dots, M$ ) に対し、上述の(1)~(4)の特徴量に基づきその領域に含まれる物体の識別を行うが、そのための識別評価関数を以下に定義する。

(1) 面積評価関数： $f_1(A)$

過去  $N$  回の追跡結果により得られた移動体候補領域の面積  $A_{jk}$  の平均値を  $A$  とし、人と想定される大きさの場合は正の値を、車と想定される大きさの場合は負の値を取る面積評価関数  $f_1(A)$  を設定する。カメラの設置高さ、カメラ光軸の方向とレンズの焦点距離が既知であり対象物体が同一平面上に存在する、という仮定を設けることができるならば、画像上での面積は実空間上での面積に変換できる<sup>(9)</sup>ので、より正確な識別が可能となる。

(2) 形状評価関数： $f_2(R)$

人の場合の画像領域は縦長になりやすく、車の場合は横長になりやすい。そこで移動体候補領域の縦と横の比  $R_{jk}$  の平均値を  $R$  として、 $R$  が大きいほど（縦長なほど）正の大きな値を取り、 $R$  が小さい横長の場合、負の大きな値を取る形状評価関数  $f_2(R)$  を設定する。

(3) エッジ分布評価関数： $f_3(E)$

前節で述べたように、人の場合はエッジ分布の値  $E_{jk}$  は大きく、車の場合は小さい。したがってその平均値を  $E$  として、エッジ分布評価関数  $f_3(E)$  は人と想定される場合は正の値を、車と想定される場合は負の値になるように設定する。

(4) フロー分布評価関数： $f_4(U)$

この場合も同様に、人の場合はフロー分布の値  $U_{jk}$  は大きく、車の場合は小さい値を取る。したがって平均値  $U$  に対してフロー分布評価関数  $f_4(U)$  を、人と想定される場合は正の値に、車と想定される場合は負の値になるように設定する。

これら4個の評価関数の識別能力は必ずしも同一でない<sup>(10)</sup>ので、次式(3)に示す重み係数  $w$  を用いた総合識別関数  $F$  の値とあらかじめ設定したしきい値  $F_{th}$  とを比較して最終的な判定を下す。

$$F = w_1 f_1(A) + w_2 f_2(R) + w_3 f_3(E) + w_4 f_4(U) \quad (3)$$

$F > F_{th} \rightarrow$  人と判定

$F < -F_{th} \rightarrow$  車と判定

$|F| \leq F_{th} \rightarrow$  外乱と判定

図6に過去  $N$  回（ここでは6回）の移動体候補領域の追跡結果に基づいた識別結果を示す。人と車を正しく識別し、草木の揺れは除外されている。

人や車を含む昼夜の様々なシーンに対して試験を行い、検知・識別率 90% 以上を達成した。

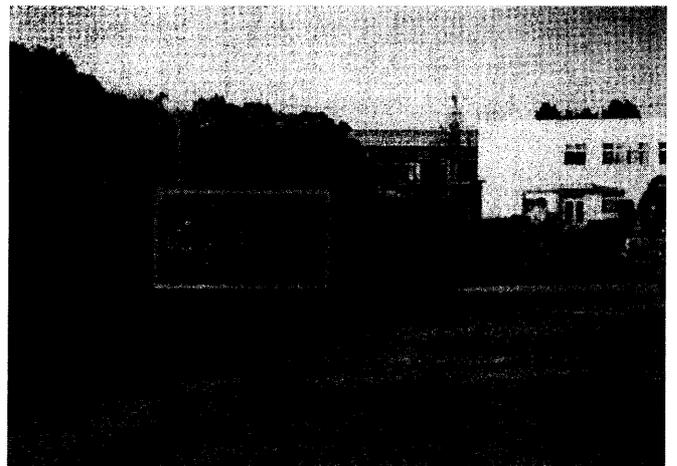


図6 識別結果 移動体候補領域に含まれる物体を人、車と外乱に識別している。  
Moving object recognition result

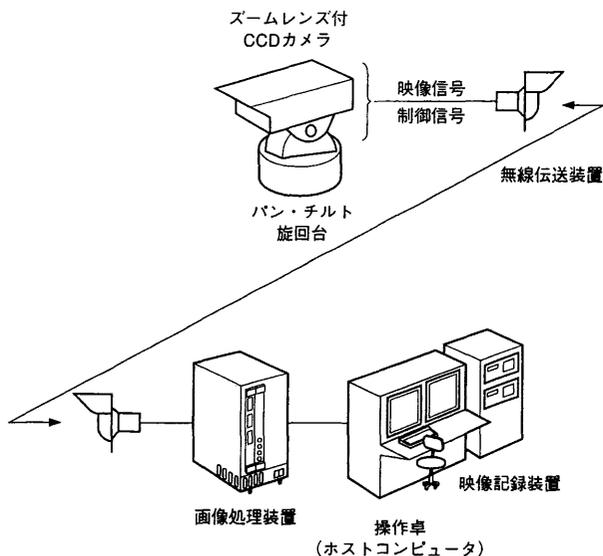


図7 移動体監視システムのハードウェア構成 監視カメラの映像は無線で操作卓側の画像処理装置へ送られ移動体検知・識別と映像記録が行われる。  
System configuration of visual surveillance system

#### 4. 適用事例

前章で述べた移動体検知と識別手法を当社で開発したリアルタイム動画画像処理装置<sup>(9)</sup>に実装した移動体監視システムの構成を図7に示す。24時間監視が可能のように、夜間のわずかな照度の下でも視認性を確保できる高感度CCDカラーカメラと、広範囲を監視するための高倍率ズームレンズを用いている。

またカメラの設置場所と操作卓の設置場所との間の映像信号や制御信号の伝送は無線で行われ、長距離にわたるケーブル敷設は不要である。

画像処理装置で移動体を検知した時点の数秒前の映像からビデオ録画が開始され、移動体を含むシーンだけを記録する。したがってビデオテープを交換することなく長時間の自動記録ができ、また、ビデオ再生による監視員の映像検索の手間を大幅に削減することができる。

風や振動によってカメラが揺れると画像が変化し、静止物を移動体として誤検知する場合がある。そこで3.3節で述べた手法で画像全体のフロー分布を検出し、画像間でのずれを補正するいわゆるぶれ補正の後、背景差分を行うようにしている。画像処理時間としては、移動体検知と識別を0.13~0.2s(移動体領域の大きさや個数に依存)で繰り返す、画像入力から識別までの遅れ時間は最大0.27sである。カメラ旋回時は更に旋回台の応答遅れも含まれるが、実用上問題のない遅れで移動体を自動的に追尾でき、既に実運用されている。

#### 5. ま と め

移動する物体の形状や動きを認識するための動画画像処理手法を開発し、移動体監視システムへ適用した。シーンの中に存在する人と車を検知・識別し、カメラの視線方向を制御して移動体追尾を可能とした。屋外シーンでは検知すべき移動体以外にも様々な外乱により画像が変化したが、移動体領域の面積、形状、エッジ分布や移動ベクトルのフロー分布に基づく評価関数により人と車だけを識別することができる。

従来の人目に頼る作業の軽減と省力化のために、自動監視の必要性は港湾、鉄道軌道、空港、プラントなど様々な重要施設に広がっている。また画像の時間変化を捕えるという意味では、プラント内で発生する水・油漏れや煙発生などの早期異常検知や移動する対象物を扱う実験計測など動画画像処理技術は様々な分野に応用することが可能である。今後、動画画像処理アルゴリズムとそれを高速に実行するリアルタイム動画画像処理装置を組込んだシステムの各種用途への適用を図っていく。さらに、カメラ旋回中の流れる画像からでも移動体を検知・追跡できるアルゴリズム開発や屋外シーンにおいて想定される様々な外乱に対するより一層の信頼性向上を目指したアルゴリズムの改良を行っていく。

#### 参 考 文 献

- (1) 増田 功, セキュリティサービスにおける認識・理解技術への期待, 電子情報通信学会技術報告PRMU 97-54 (1997) p.49~55
- (2) 天本直弘ほか, 画像処理技術による障害物検出と移動物体追跡方法, 電子情報通信学会論文誌 Vol.J 81-A No.4 (1998) p.527~535
- (3) 高橋政雄ほか, 空間微分および差分処理を用いた車両抽出法, 電子情報通信学会論文誌 Vol.J 80-DII No.11 (1997) p.2976~2985
- (4) 川端 敦ほか, 移動物体像の抽出技術, 情報処理学会論文誌 Vol.28 No.4 (1987) p.395~402
- (5) 黒田伸一ほか, 監視システムの画像処理, 映像情報 INDUSTRIAL Vol.24 No.17 (1992) p.31~39
- (6) 影広達彦ほか, 動画像からの背景画像の自動生成と適応的更新, 画像の認識・理解シンポジウム MIRU '94 Vol.II (1994) p.263~270
- (7) 紺野章子, 人物軌跡を用いた不審動作認識, 画像ラボ Vol.8 No.11 (1997) p.45~49
- (8) Barron, J. L. et al., Performance of Optical Flow Techniques, International Journal of Computer Vision Vol. 12 No.1 (1994) p. 43-77
- (9) 塘中哲也ほか, リアルタイム動画画像処理装置, 第4回画像センシングシンポジウム講演論文集 (1998) p.365~370