

動的環境における頑健な背景差分の実現法

波部 齊 大矢 崇 松山 隆司

京都大学大学院情報学研究科 知能情報学専攻

Abstract: In this paper, we propose a robust background subtraction method for non-stationary scenes. The non-stationarities modeled by the method are (1) variations of overall lighting conditions and (2) local image pattern fluctuations caused by soughing leaves, fluttering flags, flickering CRTs and so on. First we propose a novel correlation measure between two blocks in images (SMNVD) to realize the robust background subtraction against varying illuminations. Then, to characterize local image pattern fluctuations, we propose a two dimensional histogram (TNVDCM), where the distribution representing the temporal fluctuation pattern in a block is recorded. Experimental results of the background subtraction using SMNVD and TNVDCM demonstrate their robustness and effectiveness for real world scenes.

1 はじめに

動画からシーン内の移動物体の検出や追跡を行う際に用いられる手法の一つとして、背景差分が挙げられる。背景差分は観測画像と背景画像とを比較することで、効率よく移動物体を検出する手法であるが、背景シーンが静止していなければならないという厳しい制約が課せられる。それに対して、現実世界では屋内、屋外を問わず背景の物体が完全に静止していることはなく、日光などの照明条件も刻々と変化するので、常にこの制約を満すことは困難である。

そのような動的に変化する環境において背景差分を実現するために、これまで次のような方法が提案されてきた。1) 各画素における輝度値の変化を確率分布でモデル化し、その分布に対する尤度を求めることで、移動物体に相当する画素を検出する手法 [1]–[4]。2) 照明条件の変化や新たな背景物体の出現を検知し、適応的に背景を更新する手法 [5]–[7]。前者では確率分布を仮定するために、それに当てはまらないような変化に対応できないという問題があり、後者では背景の変化が検知されるまでは誤検出が多くなってしまいうという問題が生じる。これらの要因により、動的環境に幅広く適応可能な検出法の実現には至っていない。

本論文では、背景の変動を

- 照明条件の変化による変動。ここでは照明条件の変化は画素値の様な変化を引き起こすものとする。
- 木の葉の揺らぎ、旗のはためき、ディスプレイのちらつきなどによる画像上の変化。

の二つに類別し、前者に対しては様な画素値の変化に不変な特徴量を導入し、後者に対しては局所的な時間的変動をモデル化することによって、様々な種類のシーンの変化に対して、安定に移動物体を検出する手法を提案する。

様な画素変化の影響を受けず、頑健な侵入物体検出を実現するためには、画像ブロックの空間的な構造特徴に着目する。[6]では画像ベクトル間の正規化距離を用いることで画像の構造変化と様な明度分布変化との区別を試みているが、照明が暗いときにはノイズの影響が大きくなってしまい、安定な検出ができないという問題がある。それに対し我々は、正規化距離に画像の空間的な構造特徴を加味した特徴量として、*spatially modulated normalized vector distance* (以降では SMNVD と略する) を提案する。実験結果からも照明が暗いときの SMNVD の優位性が示されており、画像の特徴を評価する上で、空間的な構造特徴が有効であることが分かる。

画像上の局所的な時間的変動を特徴づけるためには相関共起行列を提案する。相関共起行列は各ブロックに対して定義される 2次元のヒストグラムであり、その (i, j) 要素は、時間間隔 Δt 離れた観測データ間の正規化距離が i と j となる共起頻度を表している。行列の頻度分布を相関共起分布と呼び、その形状を解析することで、ブロックにおける時間変動パターンを特徴づけることができる。本論文では、時間変動パターンを 5つのクラスに分け、それぞれのクラスについて最適な差分演算を定義することにより、動的なシーンに対して移動物体の検出が行えることを示す。現段階では実験結果の中でクラスの類別が不十分な部分があるが、それでも変動する背景に対して十分効果的に対象が検出できることが示されている。

A Robust Background Subtraction Method for Non-Stationary Scenes

Hitoshi Habe, Takashi Ohya, Takashi Matsuyama
Department of Intelligence Science and Technology, Graduate School of Informatics, Kyoto University

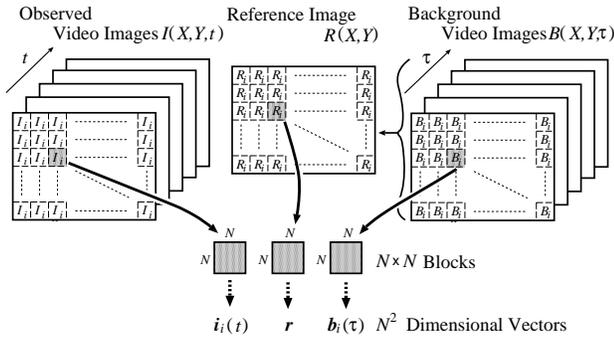


図 1: 検出に用いると画像とベクトルへの変換

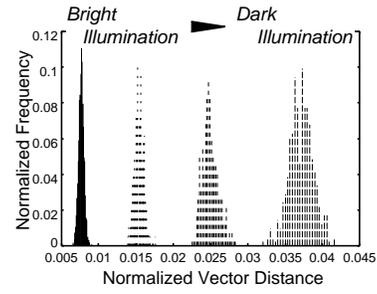


図 2: 静止背景での正規化距離のヒストグラム

2 空間的な構造特徴の利用

本節では、照明条件の変化に代表されるような、背景の画像値が一様に変化する環境で、移動物体を検出する手法について述べる。

2.1 正規化距離

本論文で述べるすべての手法は、あらかじめ背景画像系列 $B(X, Y, \tau)$ ($\tau = 1, \dots, T_B$) から基準画像* $R(X, Y)$ をはじめとする背景モデルを生成しておき、処理の対象となる観測画像 $I(X, Y, t)$ ($t = 1, \dots, T$) と $R(X, Y)$ とを比較することで移動物体の検出を行う。

以降では以下のような表記を用いる。まず、各画像 B, R, I を $N \times N$ のブロックに分割したものをそれぞれ $B_i(x, y, \tau), R_i(x, y), I_i(x, y, t)$ とする。添字 i はブロックの番号を表し、ブロック番号と画像内のブロックの位置の関係はすべての画像で同一とする。画像間の比較はブロック単位で比較を行うこととし、画像領域間の構造的な変化を検出できるようにする。

さらに、画像ブロックの各画素をベクトルの各要素としたときのベクトルをそれぞれ $b_i(\tau), r_i, i_i(t)$ とする (図 1)。このベクトルは画像内の局所的なパターンを表す。ベクトル表記を用いた場合、ブロック間の類似度を求める問題は、ベクトル間の距離を求める問題に帰着できる。[6] では距離の尺度として、以下で定義される正規化距離 (normalized vector distance; NVD) を用いている。

$$NVD(a_i, b_i) = \left| \frac{a_i}{|a_i|} - \frac{b_i}{|b_i|} \right|. \quad (1)$$

ここで、 a_i, b_i は共に画像ブロックを表すベクトルとする。正規化距離はベクトルの大きさの変化には不変な量であるので、一様な画素値の変化の影響は受けない。照明条件の変化が画素値の一様な変化を引き起こすとモデル化できるならば、正規化距離を用いて照明条件の変化以外の移動物体による画素変化の検出を行うことができる[†]。

*本論文の実験では背景画像系列の各ピクセル毎のメディアンを計算し、基準画像としている。

[†]ベクトル同士がなす角 θ も大きさに独立であるが、実際に計算に用いることになる $\cos \theta$ には θ が小さいときに分解能が悪くなるという問題がある。

以上の事項はノイズがない理想的な環境では明らかであるが、実際には撮像系に起因する様々なノイズが $I(X, Y, t)$ に含まれることになる。各画素に含まれるノイズは互いに独立に正規分布に従うとすると、それらの影響について以下のような定理が得られる (証明過程は略)。

[定理 1] $I(X, Y, t)$ として $B(X, Y, \tau)$ と同じ画像系列を用いたときに、 $NVD(i_i(t), r_i)$ の平均値は零ではなく、近似的に $|i_i|$ に反比例する。 i_i は $i_i(t)$ の各要素の平均をとったベクトルである。

[定理 2] 定理 1 と同じ条件で $NVD(i_i(t), r_i)$ の分散は近似的に $|i_i|^2$ に反比例する。

$B(X, Y, \tau)(= I(X, Y, t))$ として、照明条件を 4 段階に変化させて撮影した動画像を用意し、その中のあるブロック (静止背景) に着目して、正規化距離のヒストグラムを作成すると図 2 のようになる。この図からも上に挙げた正規化距離の性質を確かめることができる。

2.2 空間的構造特徴との統合

定理 1 および 2 により、正規化距離を画像ブロック間の相関値として用いれば、照明が暗いときに背景を移動物体と誤検出することが多くなり、信頼性が低下することになる。これは正規化距離だけの問題ではなく、色の情報を用いて照明条件に依存しない検出や認識を行う際 [8] にも、本質的に同じ要因によって照明が暗いときに信頼性が低下することが知られている。

我々は、この問題が生じる原因として、画像が本来持っている空間的な構造特徴を無視していることに着目する。つまり、 $N \times N$ の画像ブロックを N^2 次元のベクトルとするとときに失われている情報に着目するのである。実際、移動物体による画素値の変化は画像ブロック内で集中して起こることが多く、ノイズや照明変化による画素値の変化は一様に起こることが多いとすることができる。しかし、正規化距離では空間的な構造特徴を反映できないのでこれらを区別することができず、誤検出が多くなってしまう。

そこで、正規化距離に画素変化の空間的構造を表す特徴量 $SP_i(t)$ を加味することで頑健な検出の実現を試みる。つまり、 a_i と b_i の間の SMNVD を、

$$S_i(t) = F(NVD(a_i, b_i), SP_i(t)), \quad (2)$$

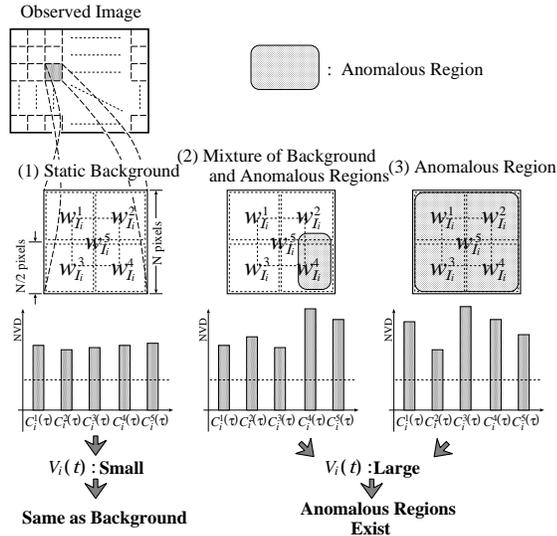


図 3: ブロック i の空間的な特徴

と定義する。ここで、 F は正規化距離に空間的構造特徴量を統合する関数である。以降に SMNVD を具体的に実現するアルゴリズムの一例について述べる。

2.3 SMNVD の実現例

ここでのアルゴリズムは非常に直感的で簡単なものであるが、4.1 節 に示す実験結果ではこのアルゴリズムでも十分有効であることが示されている。

1. 背景画像系列と基準画像の各ブロック間の正規化距離 $D_i(\tau) = NVD(b_i(\tau), r)$ を求め、その平均 \overline{D}_i と分散 σ_{D_i} を背景モデルとして保持する。正規化距離は正規分布に従うとすると、その確率密度関数 $P_{NVD}(D_i(\tau))$ は以下ようになる。

$$P_{NVD}(D_i(\tau)) = \frac{1}{\sqrt{2\pi}\sigma_{D_i}} \exp\left\{-\frac{(D_i(\tau) - \overline{D}_i)^2}{2\sigma_{D_i}^2}\right\}. \quad (3)$$

2. 基準画像中のブロック R_i の中に小ブロック $W_{R_i}^k$ ($k = 1, \dots, n$) を設け、そのベクトル表現を $w_{R_i}^k$ とする。ここで、 n は小ブロックの数である (図 3 では 5 つの小ブロックを考えている)。背景画像系列中にも同様に小ブロック $W_{B_i}^k(\tau)$ を設ける (ベクトル表現を $w_{B_i}^k(\tau)$ とする)。この小ブロック同士で正規化距離 $C_i^k(\tau) = NVD(w_{B_i}^k(\tau), w_{R_i}^k)$ を求め、その平均 \overline{C}_i^k を背景モデルとして保持する。ここまでの背景画像系列に対して行う前処理である。
3. 観測画像系列 $I(X, Y, t)$ と基準画像の間で、各ブロック i の正規化距離 $D_i(t) = NVD(i_i(t), r_i)$ とブロック内の小ブロック $W_{I_i}^k$ (ベクトル表現を $w_{I_i}^k$ とする) の正規化距離 $C_i^k(t) = NVD(w_{I_i}^k(t), w_{R_i}^k)$ ($k = 1, \dots, n$) を求める。

4. 図 3 に示すように、ブロック i における背景と移動物体の相互関係は (1) 静止背景のみの場合、(2) 背景と移動物体が混在する場合、(3) 移動物体のみの場合の 3 つに場合分けすることができる。これらを区別するために正規化距離の空間的な分散 $V_i(t)$ を定義する。

$$V_i(t) = \frac{1}{n} \sum_{k=1}^n (C_i^k(t) - \overline{C}_i^k)^2. \quad (4)$$

- (1) 静止背景のみの場合 通常の照明条件では、各 $C_i^k(t)$ ($k = 1, \dots, n$) の値は零に近くなり互いにほぼ等しくなる。よって空間的分散 $V_i(t)$ も同様に零に近くなる。照明を変化させたときに $V_i(t)$ が受ける影響については、各画素のノイズが独立な正規分布に従うとし、 $|w_{I_i}^k(t)|$ の平均値を $|w_{I_i}^k|$ とおくことで、以下の定理が証明できる (証明過程は略)。

[定理 3] $V_i(t)$ の平均値は近似的に $\frac{1}{n} \sum_{k=1}^n \frac{1}{|w_{I_i}^k|^2}$ に比例する。

[定理 4] $V_i(t)$ の分散は近似的に $\frac{1}{n^2} \sum_{k=1}^n \frac{1}{|w_{I_i}^k|^4}$ に比例する。

これらの定理より、照明が暗くなったときの $V_i(t)$ の分散の増加の度合いは $D_i(t)$ の分散の増加の度合いより小さく、その信頼性が高いことが分かる。

- (2) 背景と移動物体が混在する場合 移動物体が存在する小ブロックでは正規化距離 $C_i^k(t)$ の値は際立って大きくなる。よって、 $V_i(t)$ は大きくなる。
- (3) 移動物体のみの場合 移動物体が背景と全く同じテクスチャを持っていない限り、各々の $C_i^k(t)$ はランダムな値をとる。従って $V_i(t)$ は大きくなる。

以上のことをまとめると、空間的分散 $V_i(t)$ はブロック i の変化の空間的な散らばり方を表し、これによって移動物体の検出がある程度可能であることが分かる。よって、式 (2) 中の $SP_i(t)$ を $V_i(t)$ で定義する。

5. $SP_i(t)$ と $D_i(t)$ を統合させることで、 $S_i(t)$ が得られる。ここでは、式 (2) 中の F は次式で定義するものとする。

$$F(D_i(t), SP_i(t)) = \begin{cases} 1 & \text{if } P_{NVD}(D_i(t)) < Th1 \\ & \text{and } SP_i(t) > Th2, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

$Th1$ と $Th2$ は閾値である。

6. $F(D_i(t), SP_i(t))$ が 1 であるならば、ブロック i には移動物体が検出されたとする。

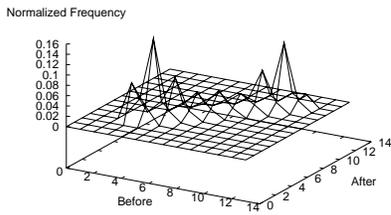


図 4: 相関共起分布の一例

3 背景の変動のモデル化

実世界のシーンでは、木の葉の揺らぎ、旗のはためき、ディスプレイのちらつきなどの様々な種類の局所的な変動がある。それらは画像上の画素の変動を引き起こすが、移動物体の検出の際には背景として処理されなくてはならない。従って、背景差分を行うための背景モデルと差分演算を、それらの局所的変動と移動物体を区別できるように拡張する必要がある。我々はそのような局所的な変動を、各ブロックにおける時間間隔 Δt をもつ正規化距離対の共起頻度の分布でモデル化し、動的に変動する背景に対しても適応できる背景差分を実現することを試みる。

本節では、そのための手法として相関共起分布を導入する。相関共起分布は相関値としての正規化距離の時間的な共起確率を表しており、その分布形状を解析すれば画像上の局所的な変動を 5 つのクラスに分類できる。各々のクラスに適応した背景差分を用いれば動的に変化するシーンに対しても移動物体検出が可能になることが期待される。

3.1 相関共起分布

背景画像系列と基準画像の間の正規化距離 $D_i(\tau) = NVD(b_i(\tau), r_i)$ は連続的な値であるとみなしてよいが、適当な量子化間隔で量子化し、1 から M までの整数値をとる $D'_i(\tau)$ を得る。その上で、 $M \times M$ の行列 $A_i = \{a_{jk}^i\} (j, k = 1, \dots, M)$ を以下のように定義する。

$$a_{jk}^i = \frac{1}{T_B - \Delta\tau} \sum_{\tau=1}^{T_B - \Delta\tau} P(i, j, k, \tau). \quad (6)$$

ここで、

$$P(i, j, k, \tau) = \begin{cases} 1 & \text{if } D'_i(\tau) = j \text{ and } D'_i(\tau + \Delta\tau) = k, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

であり、 $\Delta\tau$ は時間間隔である。我々はこの行列を相関共起行列 (temporal NVD co-occurrence matrix; TNVDCM) と呼び、行列内の頻度分布を相関共起分布と呼ぶ。相関共起分布の一例を図 4 に示す。

3.2 相関共起分布のパターン

相関共起分布のパターンは動的な背景の変動パターンを反映している。そのパターンは以下に示す 5 つに分類することができる (図 5)。

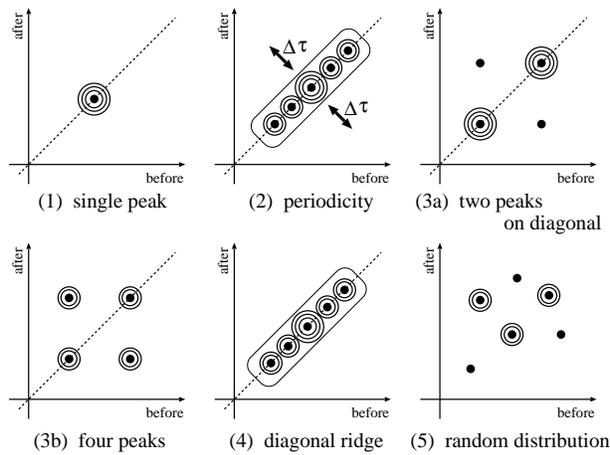


図 5: 相関共起分布のパターン

- (1) 単一の峰 相関共起分布が単一の峰のみを含んでいる場合、そのブロックは正規化距離が変化しない静止背景領域に相当する。
- (2) 周期的 ブロック中の変動が周期的である場合、相関共起分布の対角線まわりのモーメント M_d †は $\Delta\tau$ の変化に伴い変化する。 $\Delta\tau$ が周期の整数倍に等しいときにモーメントは極小になる。この性質を使えば、変動が周期的であるかどうかを確かめることができ、もし周期的であればその周期 ω_i も求めることができる。
- (3) 複数の峰 画像のパターンの変動がコントラストの高いエッジの振動のように、状態遷移モデルによってモデル化できるとき、相関共起行列には複数の峰があらわれる。図 5 (3a)(3b) は 2 状態の状態遷移モデルから得られる分布パターンを表している。
- (4) 対角線上の尾根状の分布 揺らいでいる木の葉や枝のような場合、正規化距離は徐々にだが大きな範囲で変化する。したがって相関共起分布には対角線上に滑らかな尾根があらわれる。
- (5) 不規則な分布 風にはためく旗のように、不規則に動く物体の場合、相関共起分布は他のどのクラスにも分類できない不規則な分布になる。

3.3 背景の変動パターンに応じた背景差分

時間的に変動するシーンに対して背景差分を行うために、まず相関共起分布を用いてブロック i が先に述べた 5 つのクラスのどれに属するかを判断する。その後で観測画像の各ブロックに対して、クラス分けの結果に最適化された差分演算を適用する。

3.3.1 動的背景のクラス分け

変動する背景のクラス分けは、図 6 に示される決定木にしたがって行われる。その中で、 M_c, M_d, C_t, C_b

†対角線まわりのモーメントは $M_d = \sum_{j,k} a_{jk} |j - k|$ で定義する。

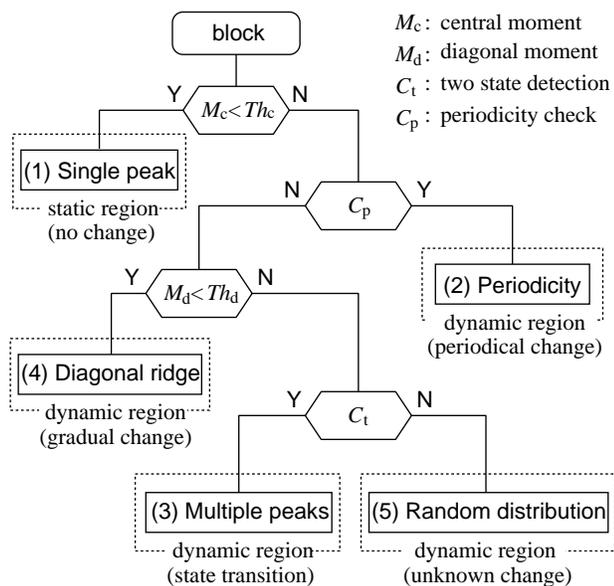


図 6: クラス分けのための決定木

はそれぞれ重心まわりのモーメント[§], 対角線まわりのモーメント, 2 状態の状態遷移の検出結果[¶], 周期性の検出結果を表している. 状態遷移の検出は相関共起分布に対してクラスタリングを行うことで実現し, 周期性の検出は相関共起行列作成時の時間間隔 $\Delta\tau$ を変化させたときの M_d の変化をみることで行っている.

3.3.2 最適化された背景差分

前節のクラス分けの結果の各クラスに最適化された背景差分は以下ようになる.

- (1) 単一の峰 背景画像系列と基準画像よりブロック i の正規化距離が従う確率密度関数が正規分布の形で得られる (式 (3)). それにより, 観測画像のブロック i での正規化距離 $D_i(t) = NVD(i_i(t), r)$ に対する尤度が計算でき, 閾値処理を行うことで移動物体の検出ができる.
- (2) 周期的 クラス分けを行う際に, 変動の周期 ω_i が推定できる. その周期 ω_i だけ離れた時刻での正規化距離の値の差 $D_i(t - \omega_i) - D_i(t)$ は正規分布に従うものとし, その値に対する閾値処理を行うことで検出を行う.
- (3) 複数の峰 状態遷移を隠れマルコフモデルによってモデル化する. 観測画像から計算される, ある時間幅の正規化距離の系列 $D_i(t - l) (l = L, \dots, 0)$ がそのモデルより生成される確率が計算でき, それに対する閾値処理を行えばよい. (現在未実装)

[§]重心まわりのモーメントは, (c_x, c_y) を相関共起分布の重心として, $M_c = \sum_{j,k} a_{jk} (|j - c_x| + |k - c_y|)$ で定義する.

[¶]現在の実装では, 状態遷移モデルは 2 状態のみを持つものとしている.

- (4) 対角線上の尾根状の分布 相関共起分布そのものが一組の正規化距離の生起確率を表しているとし, 観測画像から得られた一組の正規化距離 $D_i(t - \Delta t)$ と $D_i(t)$ の生起確率を求め, 閾値処理を行う.

- (5) 不規則な分布 (4) と同様.

4 実験

本節では, SMNVD と相関共起分布を実装したアルゴリズムによる実験結果を示し, それらの有効性を明らかにする. 用いたアルゴリズムは今までに述べたように非常に簡単なものであるが, 以下に示す実験結果はそれらが十分有効であることを示している.

4.1 実験 1

ここでは, SMNVD を用いると正規化距離に比べて照明条件の変化に対して頑健に移動物体を検出できることを確かめる. 用いた画像は 256 階調の濃淡画像であり, 固定されたカメラから撮影した. まず背景画像系列から基準画像 (図 7 (a)) を求めた. 検出対象の動画は背景画像系列より暗い照明下で撮影し, シーン中に移動物体として人が歩いている (図 7 (b)).

図 7 (c),(d),(e) はそれぞれ通常のピクセル毎の背景差分, 正規化距離の確率密度関数に対する閾値処理, そして 2.3 節に示した SMNVD を用いた検出の結果を示している. さらに閾値の影響を受けない評価を行うために正解領域をあらかじめ与えて求めた ROC 曲線を (f) に示す. これらの結果から, 本論文で提案された SMNVD を用いた背景差分は照明変化に対して頑健であることが分かる.

4.2 実験 2

次に, 3.3 節で述べた相関共起分布を用いた背景差分を実際の映像に適用する. その目的はディスプレイのちらつきや木の葉の揺らぎ, 旗のはためきや電線の揺れなどを含んだ屋外, 屋内を問わないシーンから移動物体を検出することであった. いくつかのシーンに対する解析結果が得られているが, ここではそのうち二つの例を図 8 に示す. (1) では木の枝と電線が揺らいであり, 旗が不規則にはためいている背景の中を車が右から左に移動している. (2) ではディスプレイのフリッカが観測されている中を人が歩いている.

図 8 (1-d), (2-d) は背景に対するブロックのクラス分けの結果である. 左から順に明度が高い部分が “単一の峰”, “対角線上の尾根”, “不規則な分布”, “周期的” に分類された画像ブロックを表している.

図 8 (1-c), (2-c) はブロックの分類に基づく最適化された背景差分の結果を示している. 本来検出されるべき領域が検出されていないことはあるが, その逆はほとんどないことが分かる. このことから, 背景に変動

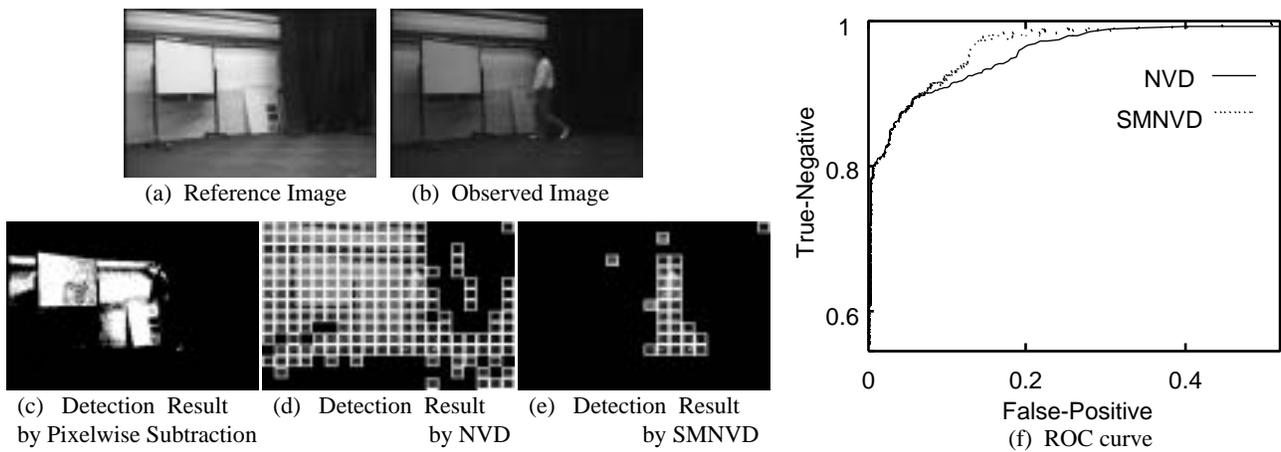


図 7: 実験結果 1. (c) では白い画素が検出された領域を表し, (d) と (e) では四角形の領域が検出されたブロックを表す.

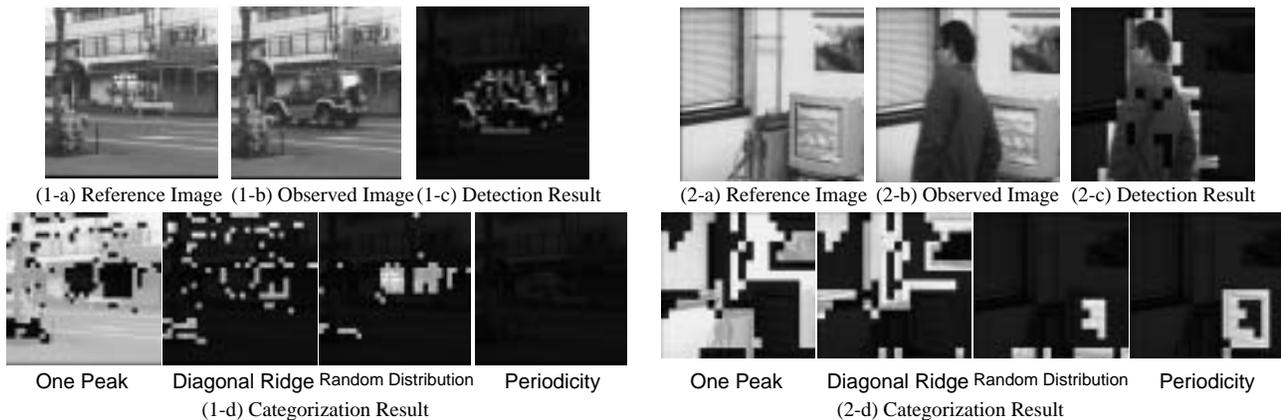


図 8: 実験結果 2

する物体があるシーンに対する背景差分として, この手法が十分有効であることが分かる.

5 まとめ

本論文では, 動的な変動をするシーンに対する背景差分を基本とする移動物体検出法を提案した. 背景の変動を照明変化のような一様な画素変化と, 背景に存在する物体が動くことによる変動に分類し, 前者については SMNVD を, 後者については相関共起分布を用いることで安定な移動物体検出を実現した.

本論文で述べた各々のアルゴリズムは非常に簡単なものであるが, 実験結果はその有効性を十分示している. SMNVD と相関共起分布はそれぞれ別個に実装されているが, それらを統合することによって, 実世界に適用できる背景差分が実現できると考えられる.

今後の課題としては, SMNVD の定義の見直しや相関共起分布のクラス分けの妥当性の検証などがあげられる.

本研究を行うにあたり, 日本学術振興会未来開拓学術研究推進事業 (JSPS-RFTF 96P00501) の補助を受けた.

参考文献

- [1] 中井 宏章: “事後確率を用いた移動物体検出手法”, 情処研報, 94-CV-90, 1994
- [2] 和田 俊和, 松山 隆司: “動的背景モデルを用いた移動領域の抽出”, 情報処理学会第 49 回全国大会講演論文集 (2), pp.141-142, 1994
- [3] E.Grimson: “A Forest of Sensors”, *Proc. of VSAM Workshop*, 1997.11
- [4] L.Davis: “Visual Surveillance and Monitoring”, *Proc. of VSAM Workshop*, 1997.11
- [5] 影広 達彦, 大田 友一: “動画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, pp. 263-270, 1994
- [6] 長屋 茂喜, 宮武 孝文, 藤田 武洋, 伊藤 渡, 上田 博唯: “時間相関型背景判定法による移動物体検出”, 信学論 D-II, Vol. J79-D-II, No. 4, pp. 568-576, 1996
- [7] 高藤 政雄, 北村 忠明, 小林 芳樹: “空間微分および差分処理を用いた車両抽出法”, 信学論 D-II, Vol. J80-D-II, No. 11, pp.2976-2985, 1997
- [8] K.Ohba, Y.Sato, and K.Ikeuchi: “Appearance Based Object Recognition with Illumination Invariance”, *In Proc. DARPA Image Understanding Workshop*, pp. 1229-1236, 1997