

HMMに基づく交通監視映像の背景・物体・影の分離手法

加藤 ジェーン[†] 渡邊 豊 英[†] 米田 政 明^{††}

自動車追跡システムのロバスト性向上の主な障害として、移動する物体（自動車）の影がある。精度の高い物体モデルを用いたとしても、追跡プロセスは影によって混乱される。本論文では、背景・物体領域と同様、影もモデル化することが可能な、HMMに基づく領域分離手法を提案する。他の多くの確率的背景モデルと異なり、本モデルにおいては学習データとして特定のものを選択する必要はない。複数のカテゴリ（背景、物体、影）を含む映像シーケンスから自動的にこれらのカテゴリ確率分布を学習することが可能だからである。各カテゴリに関する時間的持続性を利用するだけでなく、輝度値と空間的情報に基づく Sobel 値を観測シンボルとして用いることによって、カテゴリ間の曖昧性を減少させる。本手法は、ローレベルの追跡プロセスとして、または動画像による輪郭追跡システムのコンポーネントとして使われるのに適する。実際に高速道路を撮影した映像シーケンスに対して本手法を適用した結果、各画像において、背景、移動物体、および移動物体の影の3種類をリアルタイムで高精度に分離できた。

HMM-based Segmentation of Background, Object and Shadow from Traffic Monitoring Movies

JIEN KATO,[†] TOYOHIDE WATANABE[†] and MASAOKI YONEDA^{††}

The main obstacle to the robustness of car tracking is large shadows of vehicles. Even with a good foreground model, the tracking process is liable to be disrupted by the shadows. This paper proposes an HMM-based foreground-background segmentation method which is capable of modeling shadow as well as foreground and background regions. Unlike many other probabilistic background models, it is not necessary to select the training data since the distributions for different regions can be learnt from an ordinary video sequence. The ambiguity among different regions (categories) is reduced by not only using the temporal continuity constraint for each category, but using jointly intensity and Sobel values, which measure the homogeneity of a region in space, as the observation symbols. This method itself functions a low level tracker. It can be also used as a low level process for an active contour based car tracker. Results on real-world motorway sequences show that using the proposed method, it is possible to accurately segment the image into background, cars and the shadow of cars in real time.

1. はじめに

画像処理を用いた移動物体の追跡は様々な用途において必要とされ、特に交通監視映像における自動車追跡は高度道路交通システム（ITS）を実現するための重要な技術として期待されている¹⁾。ビデオ・ストリー、中の移動物体を追跡することは、基本的には移動物体の軌跡をたどることである²⁾。困難なタスクであるが、いくつかの関連する問題を解決することによって達成可能である。

最初の問題は、前景である移動物体のモデル化である。例として、計算機が高速道路上の自動車の軌跡をたどる場合を考える。自動車の外形として、ある程度の範囲の形として想定できれば、道路標示や木の影と本物の自動車を区別するためにその情報を用いることができる。また、形だけでなく、動きも予測可能である。自動車の動きはスムーズにはぼ一方方向に向かうものと考えられる。これらの情報を移動物体モデルとして保持することにより、追跡プロセスは物体付近のノイズに惑わされることなく、自動車を捕捉しつづけることが可能となる³⁾。

2番目の問題は、前景である移動物体を背景から分離することである。自動追跡の精度をさらに向上させるためには、前景画素と背景画素を区別する必要がある。物体の輪郭による追跡システムの場合、背景に存

[†]名古屋大学大学院工学研究科情報工学専攻
Department of Information Engineering, Graduate
School of Engineering, Nagoya University

^{††}富山大学工学部知能情報工学科
Faculty of Engineering, Toyama University

在するエッジや影といった強い特徴によって対象物体から外れ、追跡ミスに陥ることがある。各画素が前景、背景のいずれに属するかの情報により、対象物体と背景の特徴を区別することが可能となる。

前景画素と背景画素を分離する方法として、いくつかの異なるアプローチが提案されている。なかでも、画像差分法は、追跡のロバスト性を確保するために、多くの研究者によって使われた^{4)~8)}。この方法によれば、ある画素が前景に属するか背景に属するかは、現在の観測画像と静的な背景静止画像との差分、または現観測画像と直前の観測画像との差分（フレーム間差分）に基づいて決められる。まず、前者によれば、観測画像と背景画像の差分は前景を近似できる。よって、交通監視の場合、自動車の影が差分画像にも現れることが問題となる。他方、後者のフレーム間差分によれば、画像の時間微分を近似できる。そのため、影は差分画像上で輪郭部分しか現れないというメリットがあるが、フロントガラスのように物体内の様な領域は、差分値が小さいため背景と見なされてしまう。したがって、画像差分法では、影を除いた移動物体の真の輪郭を抽出することは非常に困難である。

画像差分とまったく異なる方法として、背景をモデル化することによって背景物体の影響を抑止する方法がある^{9)~11)}。すなわち、モデル化された背景の上で前景イベントが起こった場合、追跡プロセスは背景モデルにマッチする特徴を無視すればよい。Nakai⁹⁾は、過去の定常状態における特徴値や動き情報を学習して確率モデルを生成し、得られた特徴値が定常ではない確率を導く手法を提案した。この手法は、変動する背景の動きに対する反応はある程度抑えられるが、特徴値が広範囲にわたって変動する場合、有効なモデルの生成は困難である。Haritaogluら¹⁰⁾は、各画素が3通りの値（最小輝度値、最大輝度値、連続するフレーム間の輝度差分の最大値）によって表される単純な輝度値分布モデルを自動追跡のために構築した。これら3通りの値で記述される背景シーン・モデルによって、各画素が前景画素、背景画素のいずれかが決定される。しかし、Haritaogluらの方法は前景領域を明瞭に抽出するには不十分で、この方法で抽出された前景画像は、追跡のためには直接使われていない。

Roweら¹¹⁾は、背景画素の輝度値分布の統計的モデルを提案している。彼らは、パン・チルト用回転台を備えたビデオカメラ特有の場合について、以下のように2つの正規分布の混合モデルを用いた。均質な領域に属する画素に対しては1つの正規分布関数を、背景のエッジ上の画素に対しては、輝度値の頻度を適当

にモデル化するため、2つの正規分布関数を用いている。そして、物体の輪郭による追跡システムの実際的な検証では、差分法よりも良い結果を得ている。しかし、そのアルゴリズムは計算コストが高いうえ、物体とともに移動する影の問題を解決していない。

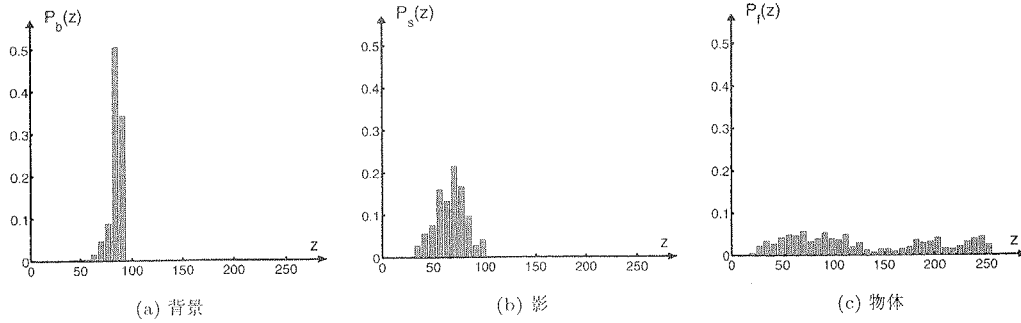
本論文では、交通監視システムにおける自動車の自動追跡の観点から、映像シーケンスを背景、移動物体、移動物体の影の3カテゴリに分離する新しい手法を提案する。影の問題を扱いたいのであるから、画像差分法や既存の確率的輝度モデルは適していないと考える。我々は、より強力な確率的モデルである、隠れマルコフモデル（Hidden Markov Model）をこの目的のために用いた。各画素は、HMMによって3カテゴリのうちのいずれかに分類される。また、HMMの各内部状態での観測シンボルの確率分布を、正規分布と一様分布関数を用いてモデル化することとした。観測シンボルとして、輝度値のみではなく、空間的情報に基づくSobel値も用いた。未知のモデル・パラメータは、E-M（Expectation-Maximization）型学習アルゴリズムによって、数十秒程度の映像から学習される。実際に高速道路を撮影した映像シーケンスに対して本モデルを適用した結果、各画像において背景、物体、影の3種類に高精度に分離できた。

以下、2章において基本コンセプトについて、3章において交通監視における自動車追跡のためのモデル設計と具体的な学習アルゴリズムを述べる。4章で、学習されたモデルに基づく、画素領域の状態の見積りについて説明する。続いて、5章では、学習の評価について述べ、6章で実映像シーケンスを用いた実験結果と評価について、7章でまとめと今後の課題について述べる。

2. 基本コンセプト

画像の前景-背景分離においては、移動物体の影も物体の一部と見なすことが多い。しかし、自動車の自動追跡の場面では、背景、物体（自動車）、影（以下 B , F , S と略す）をそれぞれ異なる3つのカテゴリと考える。背景のモデルを用いてこれらを区別するために有用と考えたので、まず、これら3つの各カテゴリについて輝度値の変化を調べることにした。図1に示したのが、ある画素位置における、それぞれ背景、物体、影と分類された場合の輝度値ごとの頻度プロファイルである。

まず、背景（図1(a)）は、かなり幅が狭く対称な分布となっている。交通が混雑した場合を検討対象から除くとすれば、各画素における頻度分布の大部分を背

図 1 $x \in \{B, S, F\}$ の輝度の頻度プロファイルFig. 1 Frequency profiles of the intensity of $x \in \{B, S, F\}$.

景の分布が占める。そして、背景の分布はほぼ正規分布の形となっている。この現象は、輝度値の変化の原因として、たとえば、ある時間帯にわたる照度の変化や、ビデオカメラが完全に静止しているわけではないことから説明可能である。他方、影は比較的光量が少ない領域であるから、測定した画素においても背景に比べて低い輝度値を持ち、影の分布は背景の分布と部分的に重なっている。影の分布もまた正規分布でモデル化できる。自動車に対応する輝度値変化の分布は、単純に一様分布で近似することが妥当である。自動車の輝度値の変化は、観測画素の位置にかかわらず、輝度値レベルの大きな範囲にわたっているからである。

図 1 に見られるように、異なるカテゴリ (B, S, F) の分布は互いに重なっている。そのために、これらの確率密度を個別に学習するのは、ロバストな方法ではない。我々は、輝度値の分布に加え、高速道路の映像シーケンスから得られる時間的な情報を利用できると考えた。すなわち、各カテゴリの出現は時間的にある程度連続性を持つ。これは、各画素がある程度の時間連続して、1つのカテゴリに属していることを意味する。

定常状態においては、ある画素が移動物体に属する確率は、背景に属する確率と比較すると、かなり小さいのが普通である。しかし、画素の輝度値が、時間的に連続する 2 枚のフレーム間において独立ではない事実を考慮に入れると、実際には異なる。あるタイムステップにおいて画素が物体 (自動車) に属しているならば、次のタイムステップでもその画素が物体 (同じ自動車) に属する可能性は高い。

隠れマルコフモデル (以下、HMM と略す) はこの種の問題に適している。HMM は、観測不可能で (隠れた) 確率的な内部プロセスと、観測可能なシンボルの列を生成する確率的なプロセスを持った、2 重に確率的なプロセスである¹²⁾。映像シーケンス中のある画

素について考えてみる。任意のタイムステップにおいて、輝度値が観測可能であり、その値は観測不可能な内部プロセスに依存している。この内部プロセスの振舞いによって、各タイムステップにおける隠れたカテゴリ (B, S, F) 間の遷移が説明される。よって、映像シーケンスを 3 カテゴリに分離する問題の解は、観測された輝度値を説明でき、特徴づける信号モデルを構築することになる。そして、このようなモデルが得られたならば、後で他の観測輝度値を分析、認識するために使うことができる。

我々の手法をさらにロバストにするため、特に自動車を確実に分離するため、カテゴリ間、特に自動車と影の分布の重なりを減少させることを考えた。この目的のために、空間的な情報に基づく第 2 の観測値、Sobel 値を導入した。影と背景はほぼ輝度が均質な空間なのに対し、物体は空間的に変化が多いという特徴があるからである。Sobel 値は、学習の対象領域の輝度値の変動を示すものとする。この変動は、影や背景に対しては小さい値となり (境界を除く)、自動車については、内部に存在するエッジのため大きい値となるべきである。輝度値と同様、我々は実データへの観測に基づき、背景と影の Sobel 値の確率密度は正規分布の形式、前景の Sobel 値の確率密度は一様分布になるとそれぞれ想定した。

本研究の目的である画像分離のためには、以下の 2 つの問題を解かなければならない。

- (1) 観測シンボル値の列 (学習シーケンス) が与えられたとき、これらが観測される可能性を最大にするためには、どのようにモデル・パラメータを調整すればよいのか。この問題については、3 章に記述する。
- (2) 別の観測シンボル値の列 (テスト・シーケンス) が与えられたとき、対応する最善な状態 (カテゴリ) の列をどのように選べばよいのか。この

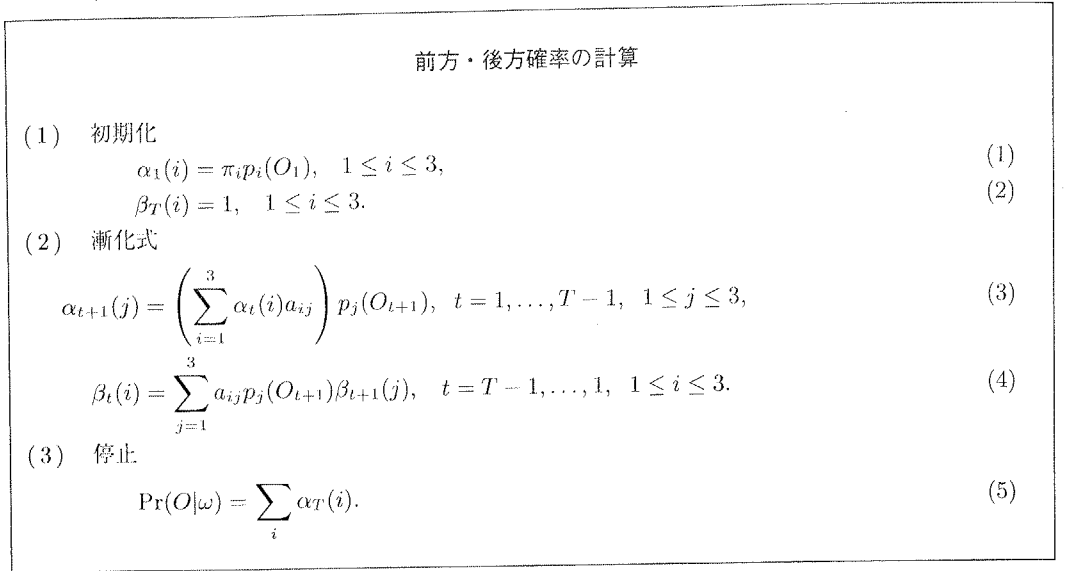


図2 補助確率を求めるアルゴリズム

Fig. 2 The algorithm to compute the auxiliary probabilities.

問題については、4章に記述する。

3. モデル・パラメータの学習

3.1 特定の問題のためのパラメータ

我々の問題においては、輝度値と Sobel 値を観測シンボルとしてよい。また、各タイムステップにおける実際のカテゴリ (B , F , S) が隠れた状態に相当する。すなわち、モデル・パラメータは状態遷移確率行列、観測シンボル確率分布、初期状態分布から構成される。したがって、HMM の各パラメータと表記は以下のように記述される。

- $I = \{i_1 = B, i_2 = S, i_3 = F\}$, 状態,
- $T =$ 観測シーケンスの長さ (フィールド数),
- $M =$ 観測可能なシンボルの数 (輝度値については $M = 256$, Sobel 値については $M = 512$),
- $O = \{O_1, \dots, O_T\}$, 観測シンボル値の列,
- $A = \{a_{ij}\} = \begin{pmatrix} a_{bb} & a_{bs} & a_{bf} \\ a_{sb} & a_{ss} & a_{sf} \\ a_{fb} & a_{fs} & a_{ff} \end{pmatrix}$,
 $a_{ij} = \Pr(i_j \text{ at } t+1 | i_i \text{ at } t)$,
 状態遷移確率行列,
- $P = \{p_j(y)\}$, $p_j(y) = \Pr(y \text{ at } t | i_j \text{ at } t)$,
 状態 j における観測シンボル y の確率分布,
- $\pi = \{\pi_b, \pi_s, \pi_f\}$,
 $\pi_i = \Pr(i_i \text{ at } t = 1)$, 初期状態確率.

2章で述べたように、背景と影の観測確率は正規分布密度関数でモデル化し、物体の観測確率は一様確率

密度関数でモデル化する。したがって、前2者の観測確率は、異なる輝度値と Sobel 値それぞれに対応する確率値の代わりに、4つの平均値 $\mu_{\kappa\iota}$ と4つの分散値 $\sigma_{\kappa\iota}^2$ のみによって決まることになる。すなわち、

$$p_{\kappa\iota}(y) = \frac{1}{\sqrt{2\pi\sigma_{\kappa\iota}^2}} e^{-\frac{(y-\mu_{\kappa\iota})^2}{2\sigma_{\kappa\iota}^2}}, \quad \kappa \in \{1, 2\}, \quad \iota \in \{B, S\} \quad (6)$$

である。ここでは、 $\kappa = 1$ と $\kappa = 2$ はそれぞれ輝度値と Sobel 値を、 $\iota = B$ と $\iota = S$ はそれぞれ背景と影を意味する。前景の観測確率は領域全体にわたり一様で、次のように定義される。

$$p_{1f}(y) = \frac{1}{256}, \quad p_{2f}(y) = \frac{1}{512}. \quad (7)$$

3.2 学習アルゴリズム

モデルは $\omega = \{A, P, \pi\}$ で表すことができる。ある時間区間の観測シンボル列が与えられた場合に HMM モデルを学習する問題は、与えられた観測データの集合 x に対し、尤度 $L(x, \omega) = \log[p(x|\omega)]$ を最大にする ω を求めることである。隠れた内部変数があるために、この ω の厳密解法は一般には知られていないので、Baum-Welsh 法のような期待最大化アルゴリズム (以下、EM アルゴリズムと略す) を用いるべきである。EM アルゴリズムは、それぞれ対応する L が順に大きくなるような ω の近似解を順に生成することができる¹³⁾。

Baum-Welsh の再評価公式¹²⁾ を我々の目的に合わせて記述するため、いくつかの補助確率を導入してお

Baum-Welsh の再評価公式

(1) 初期状態確率

$$\bar{\pi}_i = \gamma_1(i), \quad i \in \{B, F, S\}. \quad (8)$$

(2) 状態遷移確率行列

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad i, j \in \{B, F, S\}. \quad (9)$$

(3) 観測シンボル値の確率分布

$$\bar{\mu}_{\kappa j} = \frac{\sum_{t=1}^T O_t \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)}, \quad \kappa \in \{1, 2\}, \quad j \in \{B, S\}, \quad (10)$$

$$\bar{\sigma}_{\kappa j}^2 = \frac{\sum_{t=1}^T \gamma_t(j) (O_t - \mu_{\kappa j})^2}{\sum_{t=1}^T \gamma_t(j)}, \quad \kappa \in \{1, 2\}, \quad j \in \{B, S\}. \quad (11)$$

図 3 HMM パラメータを再評価するアルゴリズム

Fig. 3 The algorithm to reestimate the parameters of the HMM.

く. なお, モデル $\omega = \{A, P, \pi\}$ は与えられているとする. 以下の式で定義される前方確率 $\alpha_t(i)$ と

$$\alpha_t(i) = \Pr(O_1, \dots, O_t, i_t = i | \omega), \quad (12)$$

以下の式で定義される後方確率 $\beta_t(i)$ を考える.

$$\beta_t(i) = \Pr(O_{t+1}, \dots, O_T | i_t = i, \omega). \quad (13)$$

前方, 後方確率は, 図 2 に示すように, 再帰的に求めることが可能である. これらは, 次のように Baum-Welsh 再評価公式で必要となる 2 つの確率 $\gamma_t(i)$, $\xi_t(i, j)$ を計算するために有効である.

$$\gamma_t(i) = \Pr(i_t = i | O, \omega), \quad (14)$$

$$\xi_t(i, j) = \Pr(i_t = i, i_{t+1} = j | O, \omega). \quad (15)$$

前方確率と後方確率を用いて表すと, 以下のようになる.

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\Pr(O|\omega)}, \quad (16)$$

$$\xi_t(i, j) = \frac{\alpha_t(i)a_{ij}p_j(O_{t+1})\beta_{t+1}(j)}{\Pr(O|\omega)}. \quad (17)$$

パラメータ π , A , P を計算する Baum-Welsh 再評価公式を図 3 に示す. Baum-Welsh アルゴリズムでは, モデル・パラメータを見積もるため, 補助確率とイベント生起推計という概念が使われている*. B と S の平均と分散を再評価するためには, 通常の統計計算を使うが, 対応するカテゴリに所属する確率に心じて重み付けを行う. 式 (10) と (11) において, 各タイムステップで, あるカテゴリの確率が 1, 他のカ

テゴリの確率が 0 となる状況を想定すれば, 平均と分散を求める通常の統計計算とまったく等しくなることが分かる.

3.3 輝度値と Sobel 値の相関関係

カテゴリ間, 特に自動車と影の分布の重なりを減少させるため, 輝度値と空間的な情報に基づく Sobel 値を観測シンボルとして併用した. ノイズを減少させる視点より, 画像輝度値をそのまま使うのではなく, $n \times n$ の平均フィルタを計算してその結果を観測シンボル値とした. 一方, Sobel 値は, 学習の対象領域 (平均フィルタと同サイズを用いる) の輝度値の変動を示す. この $n \times n$ の領域を HMM 領域と呼ぶ. 平均フィルタと Sobel フィルタを用いるもう 1 つ重要な利点は, HMM モデルの数を減らすことである. $n = 3$ の場合, HMM モデルの数は 9 倍縮小するので, 学習時間も 9 倍ほど減少できる.

輝度値の平均フィルタと Sobel フィルタを併用するにあたって, 2 種類の観測値が確率的に独立であることを確かめておく必要がある. 良いフィルタは新しい情報をもたらすべきである. 我々はまず, N 個の画素を画像中からランダムに選択する. これらの画素すべてについて, 映像シーケンス中の M 枚の画像に対して 2 種類の観測値を計算し, $(O_1^{(n)}, O_2^{(n)})$, $n \leq N \times M$ を得る. そして, 以下の式で定義される線形相関係数を求める.

* Baum-Welsh 再評価公式は参考文献 12) に詳しい.

スケーリング率の再帰的定義とスケーリングされた前方・後方確率

(1) 初期化

c_1 と $\hat{\alpha}_1(i)$ を以下の式によって計算する.

$$c_1 = \frac{1}{\sum_j \alpha_1(j)}, \quad (18)$$

$$\hat{\alpha}_1(i) = c_1 \alpha_1(i). \quad (19)$$

(2) 前方確率

$\hat{\alpha}_{t+1}(j)$, c_{t+1} , $\hat{\alpha}_{t+1}(j)$ は再帰的に定義される.

$$\hat{\alpha}_{t+1}(j) = \left(\sum_i \hat{\alpha}_t(i) a_{ij} \right) p_j(O_{t+1}), \quad (20)$$

$$c_{t+1} = \frac{1}{\sum_j \hat{\alpha}_{t+1}(j)}, \quad (21)$$

$$\hat{\alpha}_{t+1}(j) = c_{t+1} \hat{\alpha}_{t+1}(j). \quad (22)$$

(3) 後方確率

同じスケーリング率を用いて再帰的に定義される.

$$\hat{\beta}_T(i) = c_T \beta_T(i), \quad (23)$$

$$\hat{\beta}_{t-1}(i) = \sum_j a_{ij} p_j(O_t) \hat{\beta}_t(j), \quad (24)$$

$$\hat{\beta}_{t-1}(i) = c_{t-1} \hat{\beta}_{t-1}(i). \quad (25)$$

図 5 スケーリング技術

Fig. 5 Scaling technique.

$$r = \frac{\sum_n (O_1^{(n)} - \bar{O}_1)(O_2^{(n)} - \bar{O}_2)}{\sqrt{\sum_n (O_1^{(n)} - \bar{O}_1)^2} \sqrt{\sum_n (O_2^{(n)} - \bar{O}_2)^2}}. \quad (26)$$

ここで、 \bar{O}_1 と \bar{O}_2 は、それぞれ 2 種類の観測値の平均である。線形相関係数の値 r は -1 以上 1 以下になる。1 の値をとるのは、各データ点が正勾配の直線上に完全に並ぶ場合である (O_1 と O_2 がともに増加する)。もし、各データ点が負勾配の直線上に完全に並ぶのであれば、 r は -1 となる (O_1 が増加すると O_2 は減少する)。 r の値がゼロに近ければ、変数 O_1 と O_2 は相関していないことを示す。 $N = 20$, $M = 50$ のとき、式 (26) によって計算される線形相関係数は、

$$r = 0.008$$

であった。この値から、輝度値と Sobel 値は確率的独立であると考え、そして、

$$O = \{O_1, O_2\} \quad (27)$$

と定義すると、

$$\Pr(O) = \Pr(O_1)\Pr(O_2) \quad (28)$$

を得ることができる。この独立の性質を前方、後方確

率の計算に適用する。

輝度値のみを用いる場合と比べて、輝度値と Sobel 値を併用する手法は状態間の分布の重なりを大きく減少させる効果がある。図 4(a) (13 ページ参照) は、この改良について概念的に示したものである。2 次元平面で見たとき (2 つの観測値を用いた場合) の、前景と影、前景と背景が重なる領域部分は、1 次元における重なり (輝度値のみの場合) よりも小さくなっている。このように、ベイズの危険が減少したため、各カテゴリ間の曖昧性は少なくなる。

3.4 初期化

EM アルゴリズムは局所的最大化手法であり、初期化が非常に重要となる。あまりにも大域的な最大値や十分に精度の良い局所的な最大値から離れた初期パラメータを選んでしまうと、精度不十分な解で停止してしまう。

ここで、パラメータの初期値を妥当な値に設定するため、時間に関係する基準値をいくつか定義しておく。ある画素が背景に属している典型的な時間 (以下、代

スケーリングされた確率を用いた Baum-Welsh 再評価公式

(1) 初期状態確率

$$\bar{\pi}_i = \frac{\hat{\alpha}_1(i)\hat{\beta}_1(i)\frac{1}{c_1}}{\sum_j \hat{\alpha}_T(j)}, \quad i \in \{B, F, S\}. \quad (29)$$

(2) 状態遷移確率行列

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \hat{\alpha}_t(i)a_{ij}p_j(O_{t+1})\hat{\beta}_{t+1}(j)}{\sum_{t=1}^{T-1} \hat{\alpha}_t(i)\hat{\beta}_t(i)\frac{1}{c_t}}, \quad i, j \in \{B, F, S\}. \quad (30)$$

(3) 観測シンボル値の確率分布

$$\bar{\mu}_{\kappa i} = \frac{\sum_{t=1}^T \hat{\alpha}_t(i)\hat{\beta}_t(i)\frac{1}{c_t}O_t}{\sum_{t=1}^T \hat{\alpha}_t(i)\hat{\beta}_t(i)\frac{1}{c_t}}, \quad \kappa \in \{1, 2\}, \quad i \in \{B, S\}. \quad (31)$$

$$\bar{\sigma}_{\kappa i}^2 = \frac{\sum_{t=1}^T \hat{\alpha}_t(i)\hat{\beta}_t(i)\frac{1}{c_t}(O_t - \mu_{\kappa i})^2}{\sum_{t=1}^T \hat{\alpha}_t(i)\hat{\beta}_t(i)\frac{1}{c_t}}, \quad \kappa \in \{1, 2\}, \quad i \in \{B, S\}. \quad (32)$$

図 6 スケーリング技術を用いた HMM パラメータの再評価アルゴリズム

Fig.6 Algorithm to reestimate the parameters of the HMM using scaling techniques.

連続時間とする)を τ_b , 同様に影, 自動車に属している典型的な時間をそれぞれ τ_s , τ_f とする. また, B, S, F に属している時間的な割合を, それぞれ λ_b , λ_s , λ_f とする. これらは, $\lambda_b + \lambda_s + \lambda_f = 1$ を満たす. さらに, $\tilde{\lambda}_{ij} = \frac{\lambda_i}{\lambda_i + \lambda_j}$ と定義する. したがって, 状態遷移確率行列の妥当な初期パラメータを,

$$A = \begin{pmatrix} 1 - \frac{1}{\tau_b} & \frac{1}{\tau_b}\tilde{\lambda}_{sf} & \frac{1}{\tau_b}\tilde{\lambda}_{fs} \\ \frac{1}{\tau_s}\tilde{\lambda}_{bf} & 1 - \frac{1}{\tau_s} & \frac{1}{\tau_s}\tilde{\lambda}_{fb} \\ \frac{1}{\tau_f}\tilde{\lambda}_{bs} & \frac{1}{\tau_f}\tilde{\lambda}_{sb} & 1 - \frac{1}{\tau_f} \end{pmatrix} \quad (33)$$

とし, 初期状態確率を,

$$\pi = \{\lambda_b, \lambda_s, \lambda_f\} \quad (34)$$

とした.

正規分布関数で近似された観測確率密度のパラメータのうち, 背景に対応する $\mu_{\kappa b}$ ($\kappa \in \{1, 2\}$) は, 対象画素の一定時間にわたる観測値の最頻値 (モード) で近似できる. $\lambda_b \gg \lambda_s$ かつ $\lambda_b \gg \lambda_f$ が想定されるからである. パラメータ $\sigma_{\kappa b}^2$ は, 経験的に決定した. このパラメータを実データから見積もるのは, たとえ交重量が少なくても, 安定でなかったからである.

影 S のためのパラメータの初期値は, 影の輝度直方が背景よりも小さいということを考慮して定めた. 正規分布の密度関数の支持領域 (危険率 5%) は, $\mu - 2\sigma, \mu + 2\sigma$ と考えられるので, μ_{1s} と σ_{1s} は, 影の分布の支持領域が 0 から背景の分布の支持領域の

上限までをカバーするように選んだ. すなわち,

$$\mu_{1s} = \frac{\mu_{1b} + 2\sigma_{1b}}{2}, \quad \sigma_{1s} = \frac{\mu_{1s}}{2}. \quad (35)$$

式 (35) によれば, $\mu_{1b} > 2\sigma_{1b}$ のとき (背景があまり暗くないとき), $\mu_{1s} < \mu_{1b}$ が保証される. 一方, μ_{2b} と σ_{2b} は実データから見積もる.

以上の方法で得られた初期パラメータは, HMM における確率的制約を満足している. すなわち, $\sum_i \pi_i = 1$, $\sum_j a_{ij} = 1$, $\sum_{k=1}^M p_j(k) = 1$, $i, j \in \{B, S, F\}$.

3.5 実装上の問題点

補助確率 $\alpha_t(i)$, $\beta_t(i)$ は両者とも指数関数的にゼロに近づく. Baum-Welsh アルゴリズムの実装上, 計算機のアンダーフローを避けるため, スケーリング技術を使用する必要がある^{14),15)}. 本論文で述べるスケーリング技術は, 前方確率 $\alpha_t(i)$ をスケーリング率 c_t で乗じ, スケーリング後の $\alpha_t(i)$ を $1 \leq t \leq T$ の間, 計算機のダイナミックレンジ内にとどめるようにする. それぞれのタイムステップ t について, $\alpha_t(i)$ のときと同じスケーリング率を後方確率 $\beta_t(i)$ についても適用する. スケーリング率は, 計算の最後には互いに約され消去される.

α が元の前方確率, $\hat{\alpha}$ がスケーリング前の前方確率, $\hat{\alpha}$ がスケーリング後の前方確率係数と仮定する. スケーリング率を再帰的に算出する手続きを図 5 に示す. 元の前方・後方確率とスケーリングされた前方・

後方確率との関係は、

$$\hat{\alpha}_t(i) = \left(\prod_{s=1}^t c_s \right) \alpha_t(i) = C_t \alpha_t(i), \quad (36)$$

$$\hat{\beta}_t(i) = \left(\prod_{s=t}^T c_s \right) \beta_t(i) = D_t \beta_t(i) \quad (37)$$

となる。図6に、スケーリング後の確率を用いた再評価式を示す。 $\hat{\alpha}_t(i)\hat{\beta}_{t+1}(i)$ または $\hat{\alpha}_t(i)\hat{\beta}_t(i)\frac{1}{c_t}$ は、 $C_t D_{t+1} = C_T$ と書けるから t に依存せず、項 $C_t D_{t+1}$ は、式(29)~(32)の分子と分母から、互いに約され消去される。このようにして、式(8)~(11)は、精度を保持して実現される。

4. 状態の見積り

観測シンボル列に基づいてモデルのパラメータを学習する問題についてすでに検討した。本論文の最大の目的は、画像を異なるカテゴリに分割することである。そして、それはモデルのパラメータが与えられたとき、学習シーケンスとは別の観測シーケンス(テスト・シーケンス)について、最適な状態の列を求めることを意味する。最適化の基準はいくつか考えられる。そのうちの1つは、ベイズ分類法を用いることである¹⁶⁾。しかし、この方法では時間情報を考慮していないため、前景である確率は実際より過小評価されることになる。よって、我々の目的には適さない。また、本研究の基本的な要件として、状態見積りは映像と同速度(リアルタイム)で動作する必要がある。したがって、我々は未来の観測シーケンスを知ることはできず、過去の履歴のみを利用できる。その結果、すべての観測シーケンス値が必要な基準を用いて最大化することは不可能である。たとえば、 $\Pr(O, I | \omega)$ の最大化は Viterbi アルゴリズム¹²⁾ によって可能ではあるが、もちろんリアルタイムではない。

さらに、高速化のため再帰的に記述されるアルゴリズムが必要である。解は、過去の観測シンボル列 $O = \{O_1, \dots, O_t\}$ が与えられたとき、時刻 t におけるその状態の確率を最大にするものである。これは、前方確率の処理のみを用いて再帰的に実現できる。我々は次の式を状態の見積りに用いた。

$$\Pr(i_t | O_1, \dots, O_t, \omega) = \frac{\alpha_t(i)}{\sum_{i=1}^3 \alpha_t(i)}. \quad (38)$$

この処理は非常に軽くリアルタイムに計算可能である。

5. 学習の評価

モデルの評価をテスト用シーケンス(学習シーケンスとは別)で行う前に、学習されたモデルの各パラメータを、学習シーケンスと比較しながら観察してみる。我々は、高速道路を撮影した30秒間の映像シーケ

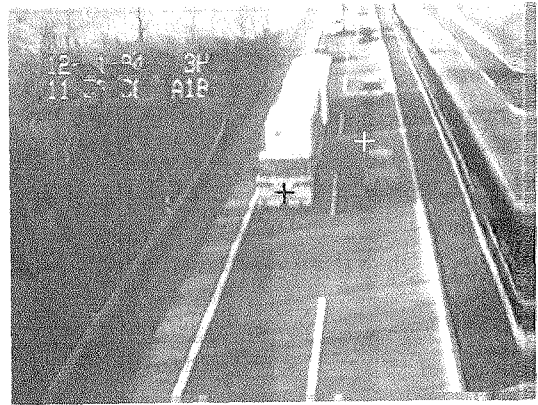


図7 高速道路を撮影した学習シーケンスからの画像
Fig. 7 An image taken from a motorway learning sequence.

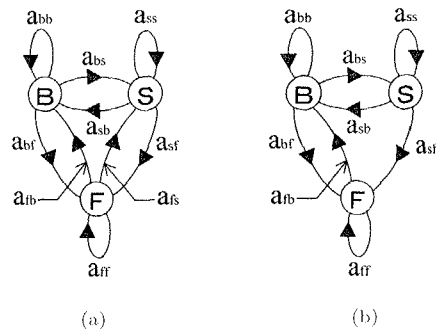


図8 標準エルゴード的モデルと制約付きモデル
Fig. 8 The ergodic HMM and constrained HMM.

ンスを、学習のために用いた(図7)。そして、例として左側車線にある十字マーク位置のHMM領域についてモデル・パラメータを学習させた。この領域は、自動車と影の間の遷移がほぼないエリアに属している。まず、HMMをどの状態間も1ステップで移動可能なエルゴード的モデルと考えて構成した(図8(a))。初期パラメータは3.4節で記述された方法で選択した。問題を単純化するために輝度値のみを用いた。3×3の平均フィルタを計算してその結果を観測シンボル値とした。

学習された統計パラメータは、

$$\mu_{1b} = 80.143, \sigma_{1b}^2 = 10.244,$$

$$\mu_{1s} = 68.2294, \sigma_{1s}^2 = 44.9762$$

であった。学習された観測確率密度がどのように重なり合い、実データとどの程度合致しているかが気になるところである。正確を期すために、学習結果である2つの正規分布関数と1つの一様確率密度の関数形を図4(b)に示す。各カテゴリにおける確率密度が状態

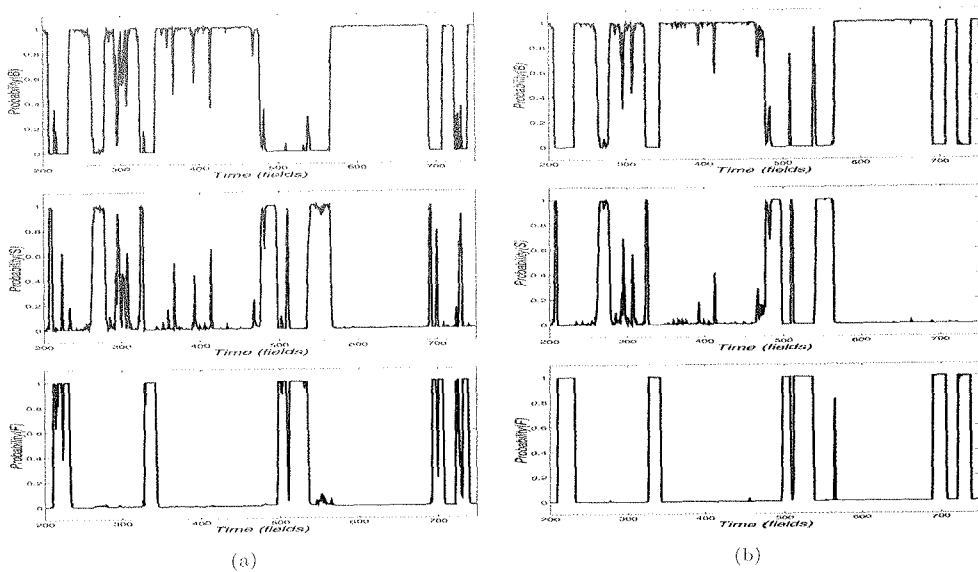


図 9 エルゴード的モデルと制約付きモデルを用いた場合の各状態に属する確率

Fig.9 The probabilities of being in a state using the standard ergodic model and the constrained model.

遷移確率行列の定常状態確率によって重み付けされていることに注意してほしい。同定常確率はそれぞれのカテゴリで費やされた時間の比率を示すものだからである。観測シーケンスの長さ T によって正規化した観測値のヒストグラムを図 4(c) に示す。明らかに、学習されたパラメータ $\mu_{1b}, \sigma_{1b}, \mu_{1s}, \sigma_{1s}$ は、実データとよく合致しているといえる。また、学習された影の分散は十分に小さい。影の分散が小さい場合は、大きい場合よりも前景物体との区別がより容易といえる。

学習された状態遷移確率行列は、

$$A = \begin{pmatrix} 0.9860 & 0.01294 & 0.001024 \\ 0.01388 & 0.8843 & 0.1018 \\ 0.03338 & 0.02537 & 0.9413 \end{pmatrix}$$

である。状態遷移確率行列 A は、時間的持続性を表すので、 $\tau_b \approx 72, \tau_s \approx 9, \tau_f \approx 17$ が導かれる。背景の代表連続時間 τ_b は、実際に背景で費やされる時間の平均である 75 タイムステップ（フィールド数）とほぼ一致している。よって、 B と $\{S, F\}$ 相互間の分離ミスは少ないと考えられる。他方、自動車の影はこの領域には出現しなかったが、影の代表連続時間 $\tau_s \approx 9$ が抽出されている。これは、学習シーケンス映像において自動車の直前に必ず暗い部分が存在していることに原因がある。この部分の画素は影と判定されている。 $a_{bf} \approx 0$ となっている事実から、自動車は通常、影と分類される暗い部分から始まっていることが分かる。ただし、これは重要な問題ではない。

問題は、 $\tau_f \approx 17$ という数値が、実際に前景と連続して分類される時間の平均値 31 タイムステップよりも短いことである。これは、前景が過小に評価されていることを示す。自動車のフロントガラスは、通常自動車の他の部分よりも暗く、学習プロセスにおいて影と見なされる。結局、1 台の自動車は 2 つの前景部分として学習されるため、 τ_f が過小評価されてしまう。この問題を解決するため、我々は HMM の各状態がすべて 1 ステップでつながっているものではないと考えた。すなわち、条件 $a_{fs} = 0$ を付加した (図 8(b))。

そして、再度同じ HMM 領域でモデル・パラメータを学習させた。学習された観測確率密度の統計パラメータは、

$$\begin{aligned} \mu_{1b} &= 80.1187, \quad \sigma_{1b}^2 = 10.5288, \\ \mu_{1s} &= 67.861, \quad \sigma_{1s}^2 = 41.9549 \end{aligned}$$

となり、状態遷移確率行列は、

$$A = \begin{pmatrix} 0.9805 & 0.01571 & 0.003829 \\ 0.01306 & 0.8978 & 0.08918 \\ 0.04800 & 0 & 0.9520 \end{pmatrix}$$

となった。状態遷移確率 a_{ff} が増加したことに注意してほしい。標準的なエルゴード的モデルと、制約 $a_{fs} = 0$ を与えたモデルを用いて状態 (B, S, F) に所属する確率を計算した結果を、図 9(a) と (b) に示す（上から順に B, S, F の確率となる）。この画素位置上を 5 台の自動車が通過した。どちらの方法でもすべての自動車に対して高い確率を得られたが、前

者の方法では、フロントガラスが影と誤って分類されるため、4台が分離された部分として抽出された。後者の方法では、4台の自動車は単体として認識されたが、1台のみフロントガラスの一部が誤って認識された。後者では、前者の方法に基づいて得られた代表連続時間よりも長い値 $\tau_f \approx 21$ が得られた。この問題は、さらに Sobel 値の導入によって改善されるが、これについては6章に述べる。

我々がモデル・パラメータの学習に用いている繰返しアルゴリズムは、3.2節で記述されている。パラメータの再評価は尤度 L が安定するまで、すなわち、現在のステップにおける尤度と前のステップにおける尤度の差が十分小さくなったときとしている。我々の実験によれば、学習アルゴリズムは高速に収束する。実際には一定回数10回の再評価でも十分で、このときは計算時間も妥当である。

6. 実験結果と検討

本章では、学習されたモデルを学習シーケンスとは別の映像シーケンスを用いて評価する。

分類処理では、過去の観測値の履歴が与えられたときの各カテゴリの確率が用いられる(式(38))。我々は、30秒間の映像を3つ使用した実験結果を示す。テスト・シーケンスでは、学習シーケンスと比較して各カテゴリ (B, F, S) で費やされる代表連続時間が変わると想定できる。なお、交通量と照明条件はあまり大きく変化しないものとする。

評価には、2つの方法を用いた。第1の方法は、過去の観測値履歴が与えられたとき、最も確からしいカテゴリ \hat{c} を1つだけ以下の式によって求めることである。

$$\hat{c} = \operatorname{argmax} \{ \Pr(i_t | O_1, \dots, O_t, \omega) \}. \quad (39)$$

これを、画面上のある注目エリア内のすべてのHMM領域について計算した。カテゴリ F, S にはそれぞれ色を割り当て、上式の結果によって得た最も確からしいカテゴリに対応した色を画面上に表示する。この方法は、確率的な情報を伝達できないが、どこで分類に成功しないし失敗したかを確認することができる。第2の方法として、5章で紹介したように、いくつかのHMM領域について、タイムステップごとに全カテゴリの前方確率を観察することができる。この方法では、確率的な情報も検討することができる。

6.1 実験 1

学習シーケンス(シーケンスA)の右車線のHMM領域について考察する。図7にその領域の位置を示す。我々は、 3×3 の平均フィルタとSobelフィルタを用

い、モデルの制約条件として、 $a_{fs} = 0$ を与えた。観測確率密度の統計パラメータと状態遷移確率行列は、以下のように学習された。

$$\begin{aligned} \mu_{1b} &= 106.068, & \sigma_{1b}^2 &= 21.0492, \\ \mu_{1s} &= 76.0468, & \sigma_{1s}^2 &= 323.127, \end{aligned}$$

$$A = \begin{pmatrix} 0.9765 & 0.02347 & 5.921 \times 10^{-7} \\ 0.02827 & 0.9192 & 0.05256 \\ 0.09277 & 0 & 0.9072 \end{pmatrix}.$$

背景分布の分散がやや大きいことに注目してほしい。これは、背景の明るさが時間とともに変化したことを示している。実際にこのシーケンスの中で、4台の明るい色のトラックが通っており、それらが通り過ぎる際に、ビデオカメラの自動絞り機構が働いて背景が一時的に暗くなったのである。

次に、ほぼ同様な交通量のテスト・シーケンスBについて、同じモデル(シーケンスAで学習)を使用して同位置のHMM領域を評価した。図10(a)-(c)に、シーケンスBの全タイムステップにわたって計算した(B, S, F)の確率を示す。人間の目で判定した結果(各カテゴリへの遷移時刻とカテゴリ名が、それぞれ破線と文字で同図に示されている)と比較してみると、それぞれのカテゴリが互いにほぼ完全に分離されていることが分かる。特に、各自動車は、1台を除き(暗い色の自動車)1つの実体として検出されている。広い領域で得られた分類結果を図11(13ページ参照)に示す。図11では、前景画素は赤色、影は緑色、背景は画素濃淡値そのままに表示している。明るい色の自動車に対してはほぼ完全に(図11(a), (b))分類でき、暗い色の自動車に対してもかなり識別できた(図11(c), (d))。

6.2 実験 2

図7のような、自動車が暗い影に覆われてしまう場合に認識されるかを検証した。シーケンスAをテスト・シーケンスとし、別のシーケンス(C)を学習シーケンスとしてパラメータの学習に使用した。学習の結果、統計パラメータは、

$$\begin{aligned} \mu_{1b} &= 109.606, & \sigma_{1b}^2 &= 12.6143, \\ \mu_{1s} &= 83.8244, & \sigma_{1s}^2 &= 290.919 \end{aligned}$$

であり、状態遷移確率行列は、

$$A = \begin{pmatrix} 0.9857 & 0.01426 & 2.140 \times 10^{-22} \\ 0.06976 & 0.8839 & 0.04635 \\ 0.1002 & 1.412 \times 10^{-5} & 0.8997 \end{pmatrix}$$

となった。この場合、自動車は完全に囲まれるため、仮定として $a_{fs} \approx 0$ は強すぎると考えられる。我々

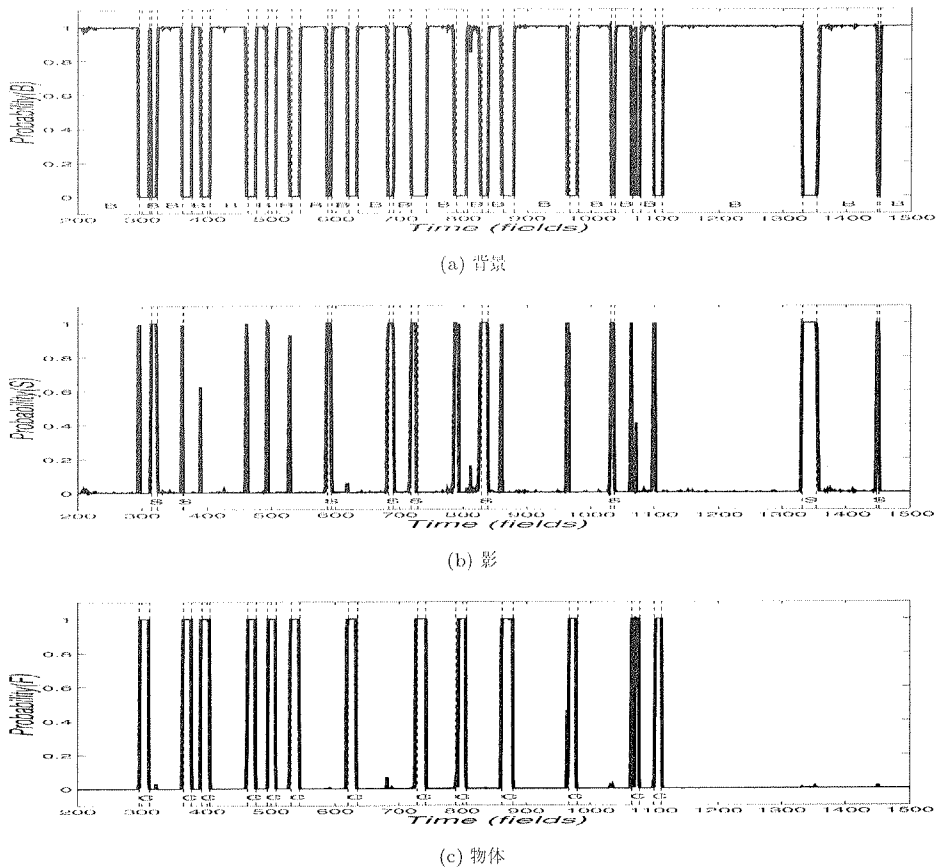


図 10 テスト・シーケンス上で計算された各カテゴリに属する確率

Fig. 10 The probabilities of being in a state for the HMM region in a test sequence.

は、標準的なエルゴード的モデルを用い、エッジ検出の効果を強調するために 5×5 の Sobel フィルタを用いた^{*}。そして、このモデルを用いてシーケンス A を評価した。

3つの各カテゴリに属する確率を計算した結果を、図 12 (a)-(c) に示す。与えられた HMM 領域において、4台の自動車(3番目、6番目、8番目、9番目)がトラックの大きな影に入ったまま通り抜ける。Sobel 値の効果を観察するため、自動車 F の検出については、輝度値と Sobel 値を用いた結果と、輝度値のみを用いた結果を、図 12 (c) と (d) にそれぞれ示した。2種類の観測値を用いた場合、これらの自動車が濃い影に覆われ、影とほとんど輝度値の差がない状態になっているにもかかわらず、良い結果になっている。2台

の自動車(1番目、6番目)に関しては、フロントガラスが部分的に誤って影に分類されているが、それでも車体の輪郭は抽出されている。他方、輝度値のみを用いた場合は、特に暗い色の自動車や影に覆われたものでは確率が低くなっているうえ、1台の自動車(暗い色)はまったく抽出できなかった。Sobel 値により、自動車の検出がよりロバストになったといえよう。図 13 は輝度値のみおよび輝度値と Sobel 値を用いて影の中の自動車の分類処理をビジュアル表示したもの(拡大図)を示す。輝度値のみでは、自動車の形は不明瞭である(図 13 (b))。Sobel 値を用いることによって、形が明瞭になる(図 13 (c))。

図 12 から分かるように、背景であるはずの一部画素が影として誤分類されている。これは、テスト・シーケンス A の背景分布の分散値(≈ 21.0492)が、学習シーケンス C のもの(≈ 12.6143)よりかなり大きいためである。というのは、シーケンス A の背景画素のうちいくつかの輝度値が、シーケンス C で学習された影分布の支持領域に含まれることになり、これが

^{*} 実験 1 のように制約付きモデルを用いる問題点としては、自動車が完全に影に囲まれるシーンにおいて、 F から S への直接遷移ができないため、自動車に続く影の一部分は自動車として誤分類されやすい。よって、自動車は実物よりやや長く分類される傾向がある。

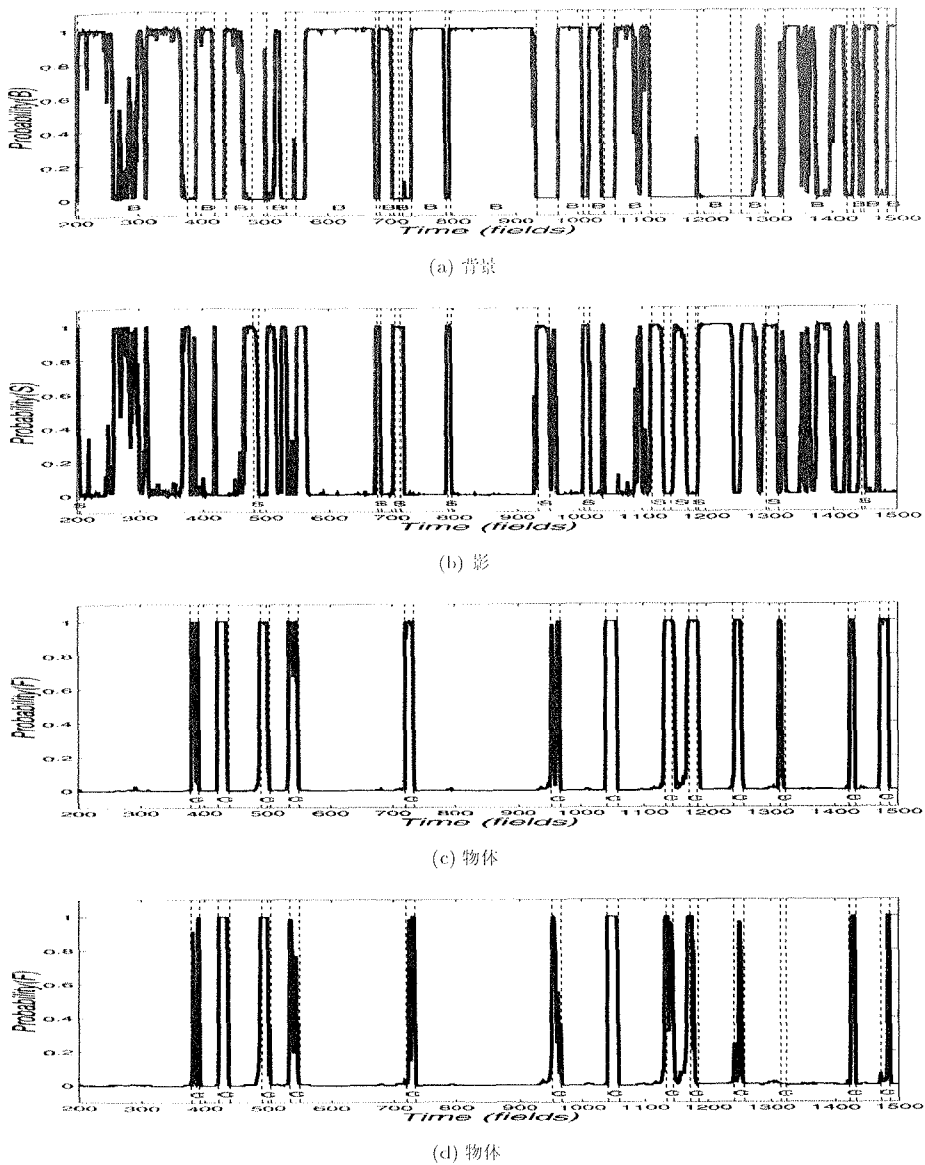


図 12 テスト・シーケンス上で計算された各カテゴリに属する確率および Sobel 値を導入する効果
 Fig. 12 The probabilities of the different categories for a region in a test sequence and the effectiveness of introducing Sobel filter.

誤った分類につながっていると考えられる。

6.3 実験 3

前述のとおり、状態遷移確率行列は時間的持続性を表している。実験 1 で学習されたモデル・パラメータのうち、状態遷移確率行列のみを人工的な値 $a_{ij} = 1/3$ に変更してテスト・シーケンス B を評価した。分類結果を図 11 (e) に示す。図 11 (d) と比べて認識結果が悪くなっていることが明らかである。観測値のみに基づいたこの分類は不十分であり、時間的持続性を表す状態遷移確率行列が重要な役割を果たす証拠となる。

7. 結 論

映像シーケンス中の前景、移動物体、移動物体の影を分離する新しい手法を提案した。特に、自動車の自動追跡の観点から、高速道路を撮影した映像シーケンスの背景モデルの学習に本手法を適用した。実験結果から、本学習手法を用いたシステムは、背景、自動車、自動車の影をそれぞれのカテゴリに高精度に分類できた。自動車追跡プロセスを混乱させる移動物体の影の問題は、ほぼ解決されると思われる。

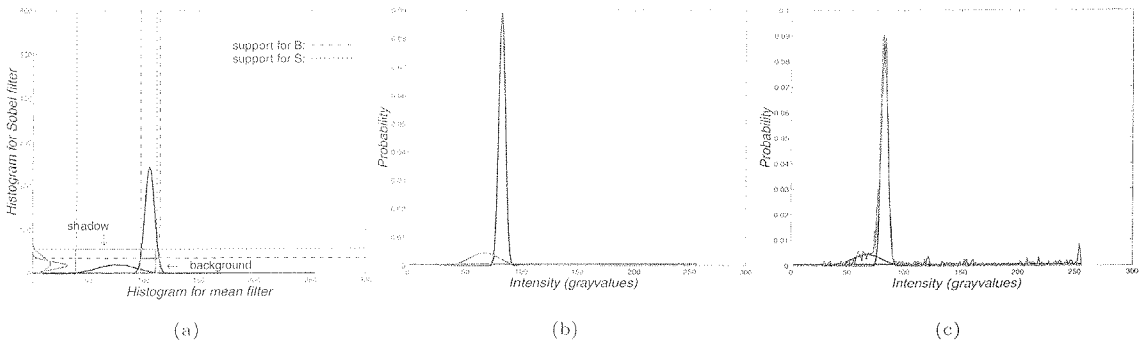


図 4 (a) 2 次元的に示す輝度値と Sobel 値の正規分布の支持領域 (2 つの長方形). (b)-(c) 学習されたモデルおよびそのモデルと実データとの比較
 Fig. 4 (a) Support of Gaussian densities for intensity and Sobel values (two boxes). (b) (c) Learnt emission model and the comparison of it with the raw data.

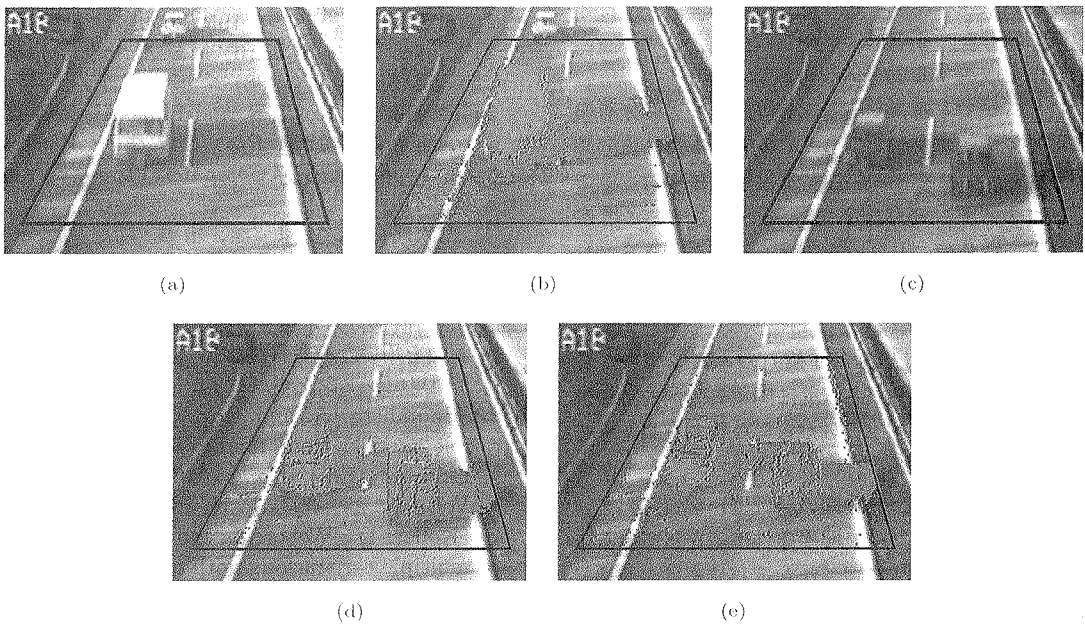


図 11 (a)-(d) 注目エリアに対する結果のビジュアル表示. (c) 時間的持続性に関する制約の重要性
 Fig. 11 (a)-(d) The visualization of the results for a region. (c) Importance of the temporal continuity constraint.

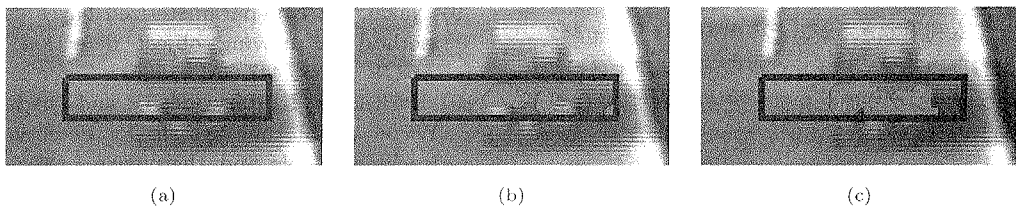


図 13 Sobel 値を用いる効果を示す分類結果のビジュアル表示
 Fig. 13 The visualization of the results for a region to show the importance of using Sobel filter jointly.

既存の確率的背景モデルと比べて、我々の手法は、前景物体が出現しないような、特別な学習データを用意する必要がなく、異なるカテゴリに対応する観測シンボルの確率分布を通常のシーケンスから学習できる。また、本手法は、輝度値、空間的な情報を表す Sobel 値、各カテゴリの時間的持続性を表す状態遷移確率に基づいている。対象物体の形状や運動に関する情報を取り入れていないため^{*}、物体の形状や運動の種類に原則的に制限されないと思われる。日照条件と交通量が変われば、モデルの更新または初期値に関する若干の調整が必要であるが、学習アルゴリズムと状態の見積り手法の有効性が保たれる。

実験結果に基づき、我々は以下の結論に達した。

- 輝度値のみに基づいた HMM では、明るい色の自動車はロバストに分類できるが、暗い色のものや影に覆われたものはロバストに分類できない。空間的情報を考慮する Sobel 値を導入することは、問題を改善する効果的な方法である。
- 物体に属する画素であっても輝度値が影の分布の支持領域に含まれる場合、物体よりも影と誤分類されやすい。F から S への状態遷移を禁じることによって、遷移確率 a_{ff} を増加させることができ、誤分類のリスクを下げるができる。しかし、この手法は自動車が影に覆われてしまうときは妥当でない。
- 観測値（輝度値と Sobel 値）のみに基づいた分類は不十分であり、各状態に関する時間的持続性を表す状態遷移確率行列が重要な役割を果たしている。

実験結果によれば、誤分類の一部は、学習シーケンスとテスト・シーケンスで照度条件が異なっていることに起因する。また、学習された遷移確率は、交通量が大きく異なるテスト・シーケンスには適用できないことも明らかである。今後の課題としては、モデルを再学習することなしに更新することが可能かどうかの検討、再学習が必要であるとすれば、その基準の検討が必要だと考える。

本手法に基づくシステムは、ローレベルの追跡プロセスとしても利用できるが、その出力を粒子フィルタを用いた輪郭に基づくハイレベルの追跡プロセスに提供し、サンプリング空間を有望な領域のみに限定して、追跡をロバストにさせるのにも適している¹⁷⁾。後者の場合は、物体の形状や運動の特徴を前景モデルとして構築する。このようなハイレベルの追跡プロセスは

ローレベルの追跡プロセスより、情報のロスが少なく対象物体の微妙な動きや外形の変化をキャッチすることが可能である。

本研究の目的は、自動車の自動追跡システムのための前景、背景の分離手法の開発であるが、他の映像シーケンス分離問題への HMM の応用も示唆すると考えている。しかし、特定のアプリケーションに対する適切なモデルや、ふさわしい観測シンボルを発見するため、注意深く研究することが重要である。

謝辞 日頃からご指導をいただく中京大学福村晃久教授に、また、本研究を行うにあたり、補助をいただいた栢森情報科学振興財団に心より感謝いたします。

参 考 文 献

- 1) 前田賢一, 小野口一則, 福井和広, 谷口恭弘: 画像認識の ITS への応用, 電子情報通信学会誌 Vol.83, No.3, pp.191-195 (2000).
- 2) Blake, A. and Isard, M.: *Active Contour*, p.352, Springer (1998).
- 3) Ferrier, N., Rowe, S. and Blake A.: Real-time traffic monitoring, *Proc. 2nd IEEE Workshop on Applications of Computer Vision*, pp.81-8 (1994).
- 4) 佐久間喜郎, 伊東 潔, 増田 功: 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, pp.1-6 (1990).
- 5) 嶺 直道, 八木康史, 谷内田正彦: 時系列差分画像を用いた複数移動物体の抽出および追跡, 第3回ロボットシンポジウム, pp.127-132 (1993).
- 6) Baumberg, A. and Hogg, D.: Learning flexible models from image sequences; Jan-Olof Eklundh (Ed.), *Computer Vision-ECCV '9*, Vol.1, pp.299-308, Springer-Verlag (1994).
- 7) Koller, D., Weber, J. and Malik, J.: Robust multiple car tracking with occlusion reasoning; Jan-Olof Eklundh (Ed.), *Computer Vision ECCV '94*, Vol.1, pp.189-196, Springer-Verlag (1994).
- 8) Paragiso, N. and Deriche, R.: A PDE-based Level Set Approach for detection and tracking of moving objects, Technical Report 317, INRIA Sophia Antipolis (1997).
- 9) Nakai, H.: Non-parameterized bayes decision method for moving object detection, *Proc ACCV '95*, pp.447-451 (1995).
- 10) Haritaoglu, I., Harwood, D. and Davis, L.S.: W4 - a real time system for detection and tracking people and their parts, *Proc. Conf. Face and Gesture Recognition*, Nara, Japan (1998).
- 11) Rowe, S. and Blake, A.: Statistical mosaics for tracking, *Image and Vision Computing*, Vol.1

^{*} これらの情報は追跡のための前景モデルとして持つ。

pp.549-564 (1996).

- 12) Rabiner, L.R.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, *Proc. IEEE*, Vol.77, No.2, pp.257-286 (1989).
- 13) Dempster, A.P., Laird N.M. and Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM Algorithm, *J. R. Stat. Soc.*, B39:1-38 (1977).
- 14) Rabiner, L.R. and Juang, B.H.: *Fundamentals of Speech Recognition*, Prentice Hall, Englewood Cliffs, NJ (1993).
- 15) Cohen, M.: Hidden Markov Models II: Implication and Complicated variants, Lecture in Boston University, <http://screwdriver.bu.edu/#cn760-lectures/110/index.htm> (1997).
- 16) 鳥脇純一郎: 認識工学, p.224, コロナ社 (1992).
- 17) Isard, M. and Blake, A.: ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework, *Proc. 5th European Conf. on Computer Vision*, pp.893-908 (1998).

(平成 12 年 1 月 7 日受付)

(平成 12 年 11 月 2 日採録)



加藤ジェーン (正会員)

1990 年名古屋大学大学院工学研究科情報工学専攻博士前期課程修了。1993 年同専攻博士後期課程修了。同年富山大学工学部助手。1999 年オックスフォード大学客員研究員。2000 年名古屋大学大学院工学研究科助教授。工学博士。文書画像理解, パターン認識, コンピュータビジョンに興味を持つ。電子情報通信学会, IEEE 各会員。



渡邊 豊英 (正会員)

1948 年生。1972 年京都大学理学部卒業。1974 年同大学院工学研究科数理工学専攻修士課程修了。1975 年同博士課程中途退学。同年京都大学大型計算機センター助手。1987 年名古屋大学工学部情報工学教室助教授。現在教授。京都大学工学博士。統合化環境, 分散協調環境, データベース環境, 知的 CAI, 文書理解, 地図理解に興味を持つ。電子情報通信学会, 日本ソフトウェア科学会, 人工知能学会, ACM, IEEE-CS, AAAI 各会員。



米田 正明 (正会員)

1967 年東北大学工学部通信工学科卒業。1972 年同大学院工学研究科電気及通信工学専攻博士課程修了。同年富山大学工学部電子工学科講師。現在, 同知能情報工学科教授。この間, オートマトン・言語理論, パターン認識理論, 文字認識・文書画像理解等の研究に従事。工学博士。電子情報通信学会, 人工知能学会各会員。