

ノンパラメトリックな動的背景モデルを用いた対象抽出

田中, 達也
九州大学システム情報科学研究院知能システム学部門

島田, 敬士
九州大学システム情報科学研究院知能システム学部門

有田, 大作
九州大学システム情報科学研究院知能システム学部門

谷口, 倫一郎
九州大学システム情報科学研究院知能システム学部門

<https://hdl.handle.net/2324/6163>

出版情報 : 画像センシングシンポジウム予稿集. 13, pp. IN3-20-1-IN3-20-5, 2007-06. 画像センシング
技術研究会
バージョン :
権利関係 :

ノンパラメトリックな動的背景モデルを用いた対象抽出

Object Detection based on Non-parametric Adaptive Background Modeling

田中達也 † 島田敬士 † 有田大作 ‡ 谷口倫一郎 †

Tatsuya Tanaka†, Atsushi Shimada†, Daisaku Arita‡, Rin-ichiro Taniguchi†

†九州大学 ‡財団法人九州システム情報技術研究所

†Kyushu University, ‡Institute of Systems & Information Technologies/KYUSHU

E-mail: {tatsuya, atsushi, rin}@limu.is.kyushu-u.ac.jp, arita@isit.or.jp

Abstract

確率密度関数をノンパラメトリックに推定する方法のひとつである Parzen 推定を利用した高速な動的背景モデル構築アルゴリズムについて述べる。動的背景モデル構築では画素単位で画素値の変化を確率密度関数で近似する方法が一般的であり、その近似を毎フレーム高速かつ精度良く行うことが重要である。提案手法では、新たに観測された画素値を基準にその周辺の確率密度関数のみを局所的に更新することで、ノンパラメトリックな背景モデルを更新するために必要な処理時間を大幅に短縮している。実験では、混合ガウス分布を利用したセミパラメトリックな手法に比べ、高速かつ高精度に背景モデルの更新および対象物体の抽出を行えることを確認した。

1 はじめに

観測画像中から背景以外の物体を検出する手法は、視覚システムの基盤となる技術の一つである。背景差分法は、あらかじめ用意した背景画像と観測画像との差分を計算することにより、物体に関する事前知識を必要とせずに物体領域を簡単に抽出することができるため、よく利用される物体検出法である。しかし、屋外などを観察する場合には、単純に背景画像を用意しただけでは、木々や水面の揺らぎなどの微小な変化や、天候による照明条件の変化による影響を受けるために、対象物体以外に様々なノイズが前景として抽出されてしまう。そこで、このような背景の変化に対応するため、動的な背景モデルとして、確率モデルを利用した動的背景モデル構築手法が用いられている [1-6]。

動的背景モデルとして良く用いられる手法は混合ガウス分布による背景モデルであり、そのパラメータを動的に更新することで、照明変動に対応した背景モデルを構築する [2-4]。しかし混合ガウス分布はあくまでも近似であり、複雑な照明変動に対処するためには、よ

り複雑な分布を記述できる枠組みが必要である。

一方で Elgammal らは、ノンパラメトリックな確率分布を推定する手法である Parzen 推定を用いて背景モデルを推定する手法を提案している [5]。この手法では、パラメトリックな形式を仮定せずに確率密度関数を推定することができるため、真の分布がどのような形であっても精度よく背景モデルの推定を行うことができる。しかし、精度よく推定を行うためには十分な数のサンプル数が必要であり、確率密度関数の推定に多くの計算が必要となうえ、その計算量はサンプル数に比例するため、実時間処理には向いていない。

そこで我々は、Elgammal らによる従来手法を改良し、Parzen 推定を利用した動的背景モデル構築とその高速化手法について研究を進めた。以下では、2 節で Elgammal らが提案している Parzen 推定を用いた従来背景モデル推定法を紹介する。3 節では、我々が提案する高速な背景モデル推定法について述べる。提案手法を用いた実験結果については 4 節で報告する。

2 Parzen 推定による背景モデル推定

この節では、まず 2.1 節で背景モデルの構築に用いる Parzen 推定について述べ、2.2 節で Elgammal らが提案している Parzen 推定を用いた動的背景モデル構築の従来法について述べる。

2.1 Parzen 推定

Parzen 推定では、過去に得られた画素値を基に、観測された画素値の確率密度関数をノンパラメトリックに推定する。画素 (x, y) における画素値を d 次元ベクトル \mathbf{X} で表すことにすると、過去に観測された N 個の画素値 $\{\mathbf{X}_1, \dots, \mathbf{X}_N\}$ の確率密度関数 $P(\mathbf{X})$ は Parzen 推定を用いて、

$$P(\mathbf{X}) = \frac{1}{N} \sum_{i=1}^N K(\mathbf{X} - \mathbf{X}_i) \quad (1)$$

と推定することができる。ここで、 K は窓関数と呼ばれ、その積分値は 1 であり、適切な広がりを持ってい

る関数である．滑らかな確率密度関数を推定するため，一般的にガウス分布が窓関数としてよく用いられている．窓関数 K を $N(0, \Sigma)$ で表される d 次元のガウス分布とすると，式 (1) は，

$$P(\mathbf{X}) = \frac{1}{N} \sum_{i=1}^N \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{(\mathbf{X}-\mathbf{X}_i)^T \Sigma^{-1} (\mathbf{X}-\mathbf{X}_i)}{2}} \quad (2)$$

となり，この式により確率密度関数の推定を行う．

2.2 背景モデル構築アルゴリズム

画素ごとに背景モデルを構築するため，以下では特に断りがない限りひとつの画素 (x, y) に焦点を絞って考える．まず，式 (2) において，共分散行列を次式で近似する．

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_d^2 \end{pmatrix} \quad (3)$$

この式は，例えば \mathbf{X} が RGB 値で構成される場合，各色成分は独立であることを意味する．これにより，確率密度関数の推定において，共分散成分を考慮する必要がなくなり，逆行列などの複雑な計算を回避できるため，計算量を削減することができる．以上により，画素値 \mathbf{X} が生起する確率は式 (4) を用いて推定することができる．

$$P(\mathbf{X}) = \frac{1}{N} \sum_{i=1}^N \prod_{j=1}^d \frac{1}{(2\pi\sigma_j^2)^{\frac{1}{2}}} e^{-\frac{(\mathbf{X}-\mathbf{X}_i)_j^2}{2\sigma_j^2}} \quad (4)$$

具体的な背景モデル構築の手順は以下の通りである．

Step1 新たに画素値 \mathbf{X}_{N+1} が観測されたとき，式 (4) を用いて \mathbf{X}_{N+1} が生起する確率 $P(\mathbf{X}_{N+1})$ を求める．

Step2 $P(\mathbf{X}_{N+1})$ が閾値以上であれば背景，閾値以下であれば前景と判定する．

Step3 観測画素値 \mathbf{X}_{N+1} を保存し，最も古い画素値データ \mathbf{X}_1 を削除する．

以上の処理を画素ごとに繰り返し行うことで，背景モデルの構築を行う．

3 確率密度関数推定の高速化

2.2 節で述べた式 (4) を用いて画素値 \mathbf{X} が生起する確率を毎フレーム求める場合，窓関数 K により求まる値を N 回足し合わせる計算をしなければならず，背景モデルである確率密度関数を求める処理時間が増大してしまう．そこで提案手法では，高速な確率密度関数推定を可能にすべく，式 (5) に示す窓関数を用いる．ただし，

h は窓関数の幅を表すパラメータである (図 1 参照)．

$$K(\mathbf{X} - \mathbf{X}_i) = \frac{1}{h^d} \psi\left(\frac{|\mathbf{X} - \mathbf{X}_i|}{h}\right) \quad (5)$$

ただし，

$$\psi(u) = \begin{cases} 1 & \text{if } u \leq \frac{1}{2} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

ここで， $|\mathbf{X} - \mathbf{X}_i|$ は d 次元のチェスボード距離である．そうすると，式 (1) は以下のように変形できる．

$$\begin{aligned} P(\mathbf{X}) &= \frac{1}{Nh^d} \sum_{i=1}^N \psi\left(\frac{|\mathbf{X} - \mathbf{X}_i|}{h}\right) \\ &= \frac{k(\mathbf{X})}{Nh^d} \end{aligned} \quad (7)$$

$k(\mathbf{X})$ は \mathbf{X} を中心とする一辺 h の超立方体 R 内にある \mathbf{X}_i の個数を表す．式 (7) を用いて確率密度関数の推定を行う場合， $P(\mathbf{X})$ の値は過去に観測された N 個の画素値 \mathbf{X}_i のうち， \mathbf{X} を中心とする超立方体内にいくつ存在するかを数えることになる．この $P(\mathbf{X})$ を求める処理において，従来手法と同様に \mathbf{X}_{N+1} の周辺の \mathbf{X}_i を数えることによって $P(\mathbf{X}_{N+1})$ を求める方法では， N が大きくなると処理時間が長くなってしまおうという問題がある．そこで本研究ではこれを改良し，常にすべての \mathbf{X} について $P(\mathbf{X})$ を求めておき，毎フレーム $P(\mathbf{X})$ を更新していくことにより，処理時間がより短く，しかも N の大きさに影響されない手法を提案する． $P(\mathbf{X})$ は \mathbf{X} を中心とする超立方体内の \mathbf{X}_i の個数によってきまるので， \mathbf{X}_{N+1} が観測された場合にその影響を受ける $P(\mathbf{X})$ は \mathbf{X}_{N+1} を中心とした窓関数内，および \mathbf{X}_1 を中心とした窓関数内に限られる．言い換えると，新たな画素値 \mathbf{X}_{N+1} が観測されたときに確率密度関数の更新が必要となるのは，観測画素値を中心とする窓関数内と，最も古い画素値を中心とする窓関数内だけである．最も古い画素値を中心とする窓関数内の確率密度関数も更新しなければならない理由は，確率密度関数の推定に用いるデータ数は N 個なので，新たな画素値が観測された場合，最も古い画素値のデータを削除する必要があるからである．提案手法では，このことを利用し，全ての画素値 \mathbf{X} において確率密度関数を推定し直すのではなく，局所的に確率密度関数を推定しなおすことによって背景モデルの更新を行う．

まず，画素値は離散値であるのですべての \mathbf{X} における確率密度関数の値 $P(\mathbf{X})$ はあらかじめ計算し，保持しておくことが可能である．ここで，新たな観測画素値 \mathbf{X}_{N+1} が得られた時， $P(\mathbf{X})$ は以下のようにして更新することができる．

$$\begin{aligned} P_t(\mathbf{X}) &= P_{t-1}(\mathbf{X}) + \frac{1}{Nh^d} \psi\left(\frac{|\mathbf{X} - \mathbf{X}_{N+1}|}{h}\right) \\ &\quad - \frac{1}{Nh^d} \psi\left(\frac{|\mathbf{X} - \mathbf{X}_1|}{h}\right) \end{aligned} \quad (8)$$

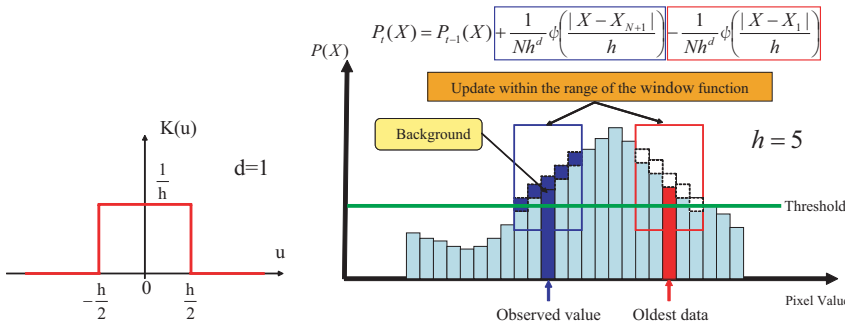


図1 提案手法で用いる窓関数

図2 背景モデル更新アルゴリズム

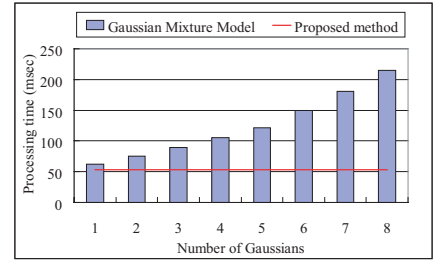


図3 ガウス分布の数と処理時間の関係

これは、新たな観測画素値 X_{N+1} を中心とする窓関数内の確率を $\frac{1}{Nh^d}$ だけ高くし、 N フレーム前の観測値 X_1 を中心とする窓関数内の確率を $\frac{1}{Nh^d}$ だけ低くする操作である。これにより、新たな画素値が観測されるたびに局所的に $P(X)$ を更新するだけでよく、処理時間の大幅な短縮が可能である。図2は $d = 1$ の場合の背景モデル更新の様子を図で表したものである。なお、確率密度関数の初期値はすべての X について等確率、つまり一様分布としている。

4 実験

提案手法の有効性を示すため、計算時間と対象抽出精度について評価した。実験データはビデオサーベイランスの評価で良く用いられる PETS (PETS2001)¹ を用いた。映像には通りを行き交う人々や車の移動、木々の揺らめきなどが含まれている。この映像に対し、提案手法、混合ガウス分布による手法、Elgammal らによる従来手法を用いて実験を行い、その結果を比較した。ここで、提案手法では各色成分は独立であると仮定し、 R, G, B それぞれの軸で1次元の確率密度関数を推定する。 R, G, B それぞれの軸において、観測画素値における確率が閾値を下回るものがひとつでもあれば、その画素を前景と判定する。なお、実験に用いた計算機は、Pentium4 3.2GHz および 2.5GB のメモリを搭載している。

4.1 実験1: 計算時間による評価

まず、混合ガウス分布による手法と比較した。ここで、提案手法において $h = 5, N = 500$ としている。図3に示す結果からわかるように、混合ガウス分布による背景モデルの場合には、ガウス分布の分布数によって計算時間が異なってくる。分布数を2以上(混合ガウス分布)にすると、提案手法の高速性が顕著に表れる。また、実際の応用では、事前に分布数を決めておくこと

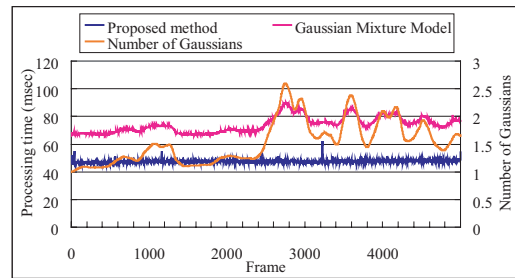


図4 分布数可変時の処理時間および1画素あたりの平均分布数

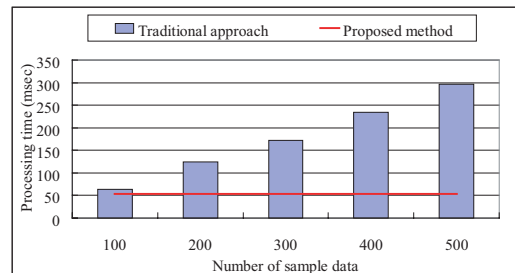


図5 確率密度関数推定に用いる画素値データ数と処理時間の関係

は難しく、動的に分布数を変更することが望ましい [4]。しかし、図4に示す結果からわかるように、分布数を可変にするとその変動に伴って1フレームあたりの計算時間が変動するため、安定した実時間処理が行えないという問題がある。一方、提案手法は計算時間を左右するパラメータを含んでいないため、安定した実時間処理が可能になる。

次に、図5に Elgammal らの従来手法との比較について示す。ここで、提案手法において $h = 5$ としている。Elgammal らの従来手法では、確率密度関数の推定に用いる画素値データ数 N が増加するにつれて処理時間も増加している。実時間性を考慮すると推定に用いることのできるデータ数は150個程度までということ

¹International Workshop on Performance Evaluation of Tracking and Surveillance でのベンチマークデータ。ftp://pets.rdg.ac.uk/PETS2001/ より入手可能。

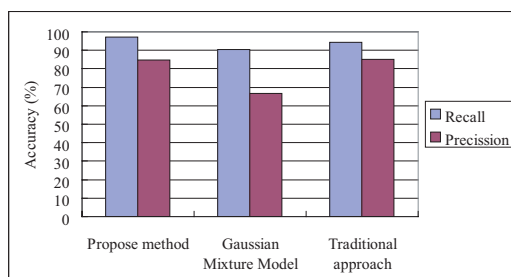


図 6 再現率・適合率のグラフ

になる．一方で，提案手法ではデータ数にかかわらず，一定かつ短い処理時間で計算が行えていることがわかる．提案手法では，新たに観測された画素値および最も古い画素値を中心とする窓関数内の確率密度関数だけを更新しているので， N の大きさにかかわらず一定の処理速度で計算を行うことができる．

4.2 実験 2：対象抽出精度による評価

対象抽出精度に関しては，人手で与えた真値を基に，分布数可変の混合ガウス分布による方法，Elgammal らによる従来手法，および提案手法における再現率と適合率を求めた．ここで，再現率，適合率は以下の式で定義される．

$$\text{再現率 (\%)} = \frac{\text{正しく抽出した物体数}}{\text{抽出すべき物体数}} \times 100 \quad (9)$$

$$\text{適合率 (\%)} = \frac{\text{正しく抽出した物体数}}{\text{抽出した物体数}} \times 100 \quad (10)$$

なお，提案手法では $N = 500, h = 5$ ，Elgammal らの従来手法では $N = 500, \sigma = 1$ としている．図 6 に示すように，抽出精度についても提案手法が混合ガウス分布による手法より優れていることがわかる．また，Elgammal らの手法と比べても，同程度の精度が出ていることがわかる．

最後に，提案手法による背景差分の結果画像を図 7 に示す．また，画像処理結果の動画像を添付する．図 7(a) は四つのフレームにおける入力画像である．図 7(b) はそのフレームにおける背景モデルにおいて，観測される確率が最も高い画素値を表している．図 7(c) は提案手法により得られた物体領域を示す画像である．また，比較のため，図 7(d) に分布数可変型の混合ガウス分布による背景差分の結果，図 7(e) に Elgammal らの従来手法による背景差分の結果を示す．提案手法では，混合ガウス分布を用いた手法よりも高精度，および Elgammal らの従来手法と同程度に，人物や車の領域を抽出できていることが分かる．

5 おわりに

動的背景モデルに基づいた対象抽出は，照明条件の変動が大きい屋外情景の画像解析において必要不可欠

な手法である．本稿で提案した Parzen 推定によるノンパラメトリックな背景モデルによる対象検出は，計算時間，実時間性，抽出精度の点で，従来から良く用いられている混合ガウス分布による方法，および Parzen 推定を用いた従来手法よりも優れていることを確認した．また，本手法はシステムの持つパラメータも少ないため，非常に使いやすい手法ということが出来る．

今後の課題と展開としては以下のことがあげられる．

- 影領域の除去：
今回提案した手法では，物体による影の領域も物体領域として抽出されてしまう．背景差分によって得られた物体領域を画像認識などに利用する場合，影領域は悪影響を及ぼす可能性があるため，影領域を除去する仕組みが求められる．
- 使用するメモリ領域の削減：
提案手法では，過去の画素値データおよび確率密度関数のヒストグラムを保持しておく必要があるため，映像の解析には多くのメモリ領域を確保する必要がある．解像度の高い画像に対し解析を行うことを考えた場合，使用するメモリ領域を減らすための工夫が必要である．

参考文献

- [1] 古山孝好, 向川康博, 亀田能成, 大田友一: “屋外大規模空間における自由視点映像生成のための選手領域抽出法”, 画像の認識・理解シンポジウム (MIRU2005), pp. 1412–1419, 2005.
- [2] Chris Stauffer, W.E.L. Grimson: “Adaptive background mixture models for real-time tracking”, *Proc. of Computer Vision and Pattern Recognition (CVPR)*, vol. 2, pp. 246–252, 1999.
- [3] Jian Cheng, Jie Yang, Yue Zhou: “A Novel Adaptive Gaussian Mixture Model for Background Subtraction”, *Proc. of Conference on Pattern Recognition and Image Analysis (IbPRIA)*, pp. 587–593, 2005.
- [4] 島田敬士, 有田大作, 谷口倫一郎: “混合ガウス分布による動的背景モデルの分布数増減法”, 画像の認識・理解シンポジウム (MIRU2006), pp. 746–751, 2006.
- [5] Ahmed Elgammal, Ramani Duraiswami, David Harwood, Larry S. Davis: “Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance”, *Proc. of the IEEE*, vol. 90, pp. 1151–1163, 2002.
- [6] 島井博行, 栗田多喜夫, 梅山伸二, 田中勝: “ロバスト統計に基づいた適応的な背景推定法”, 電子情報通信学会論文誌 D-II, vol. J86-DII, no. 6, pp. 796–806, 2003.

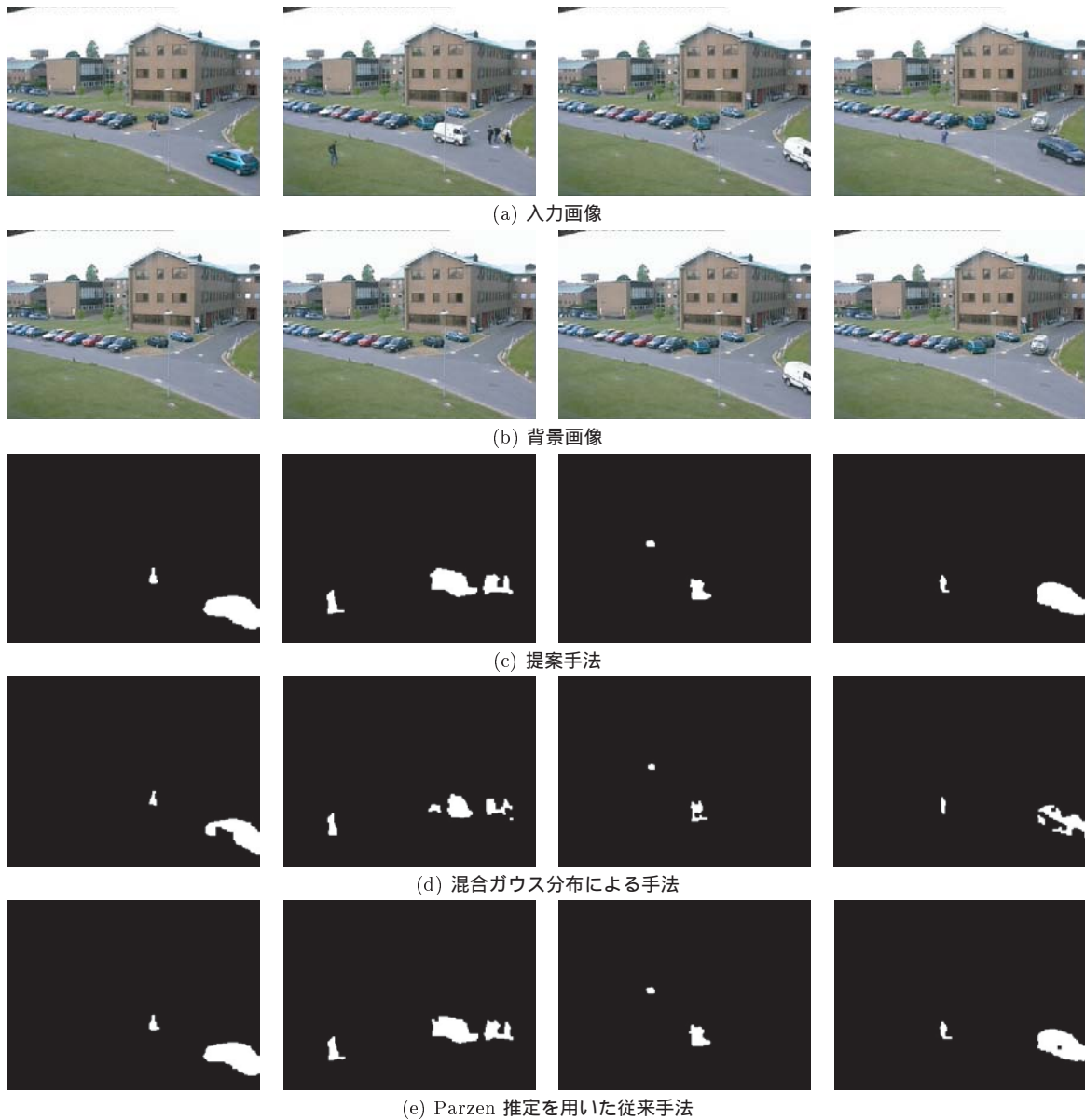


図7 背景差分の結果画像。(a) 入力画像, (b) 提案手法における観測される確率が最も高い画素値による画像, (c) 提案手法による前景抽出結果, (d) 混合ガウス分布を用いた手法による前景抽出結果, (e) Parzen 推定を用いた従来手法による前景抽出結果。