

## Object Detection based on Non-parametric Adaptive Background and Shadow Modeling

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

<http://hdl.handle.net/2324/6161>

---

出版情報：情報処理学会研究報告. CVIM, [コンピュータビジョンとイメージメディア]. 2007 (42), pp.105-112, 2007-05. Information Processing Society of Japan

バージョン：

権利関係：ここに掲載した著作物の利用に関する注意 本著作物の著作権は（社）情報処理学会に帰属します。本著作物は著作権者である情報処理学会の許可のもとに掲載するものです。ご利用に当たっては「著作権法」ならびに「情報処理学会倫理綱領」に従うことをお願いいたします。



# ノンパラメトリックな動的背景・影モデルに基づいた 映像からの物体抽出

田中 達也† 島田 敬士‡ 有田 大作\*† 谷口 倫一郎‡

九州大学大学院システム情報科学府†

九州大学大学院システム情報科学研究所‡

〒 819-0395 福岡県福岡市西区元岡 744

財団法人九州システム情報技術研究所\*

〒 814-0001 福岡県福岡市早良区百道浜 2 丁目 1-22

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

あらまし 本稿では、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、新たな観測画素値を基準にその周辺の確率密度のみを局所的に更新することで、背景モデルの更新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が確認できた。また、背景モデルを影モデルへ応用することで、背景差分によって得られた画像から物体によってできた影領域を正確に除去することができた。

キーワード 背景差分, 物体抽出, 動的背景モデル, 動的影モデル, 背景変動, 照明変動, Parzen 推定

## Object Detection based on Non-parametric Adaptive Background and Shadow Modeling

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

Department of Intelligent Systems, Kyushu University†‡

744, Motooka, Nishi-ku, Fukuoka, 819-0395, Japan

Institute of Systems & Information Technologies/KYUSHU\*

2-1-22, Momochihama, Sawara-ku, Fukuoka, 814-0001, Japan

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

**Abstract** We propose a fast algorithm to estimate background models using Parzen density estimation in non-stationary scenes. Each pixel has a probability density which approximates pixel values observed in a video sequence. It is important to estimate a probability density function fast and accurately. In our approach, the probability density function is partially updated within the range of the window function based on the observed pixel value. The model adapts quickly to changes in the scene and foreground objects can be robustly detected. And applying our approach to the shadow models, it enables detecting and modeling moving cast shadows. Several experiments show the effectiveness of our approach.

**Key words** Background Subtraction, Object detection, Shadow model, Non-stationary background, Illumination change, Parzen density estimation

## 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。

また、背景差分法の一つの問題は、対象物によってできる影も物体領域として検出されてしまうことである。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必要である。そこで、様々な影領域の除去法が提案されている [7-11]。その中で Martel-Brisson らは、動的背景モデルと同様に、

確率密度関数により影を表す画素値の変化を近似することで動的な影モデルを構築し、それにより物体によってできた影の領域を除去する方法を提案している [11]。しかしこの手法では動的背景モデルおよび影モデルの構築に混合ガウス分布を用いているため、上記の混合ガウス分布による手法で述べたような問題が残る。

そこで、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行った。また、提案する Parzen 推定を利用した動的背景モデルと Martel-Brisson らの手法を比較し、影除去に対して有効に働くモデルの検証を行った。

以下では、2 節で Elgammal らが提案している Parzen 推定を用いた従来の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。また、4 節では 3 節で提案する背景モデル推定法を応用した動的影モデルによる影領域の除去法について述べる。提案手法を用いた実験結果については 5 節で報告する。

## 2. Parzen 推定による背景モデル推定

この章では、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の構築法について述べる。

### 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  $X$  で表すことにすると、過去に観測された  $N$  個の画素値  $\{X_1, \dots, X_N\}$  の確率密度関数  $P(X)$  は Parzen 推定を用いて、

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

と推定する。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。以下では、 $K$  を  $N(0, \Sigma)$  で表されるガウス分布とする。その次元数を  $d$  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。共分散成分を無視することに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** Step1 で求めた  $P(\mathbf{X}_{N+1})$  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、背景モデル

の推定には多くの計算が必要である。Elgammal らは、全ての場合の  $(\mathbf{X} - \mathbf{X}_i)$  について、窓関数  $K$  の値をあらかじめ計算しておき、ルックアップテーブルを用いることで、処理時間を削減している。しかし、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、背景モデルを推定するのに時間がかかる。そこで、我々が提案する手法では、以下に示す方法でより高速に確率密度関数の推定を行う。

まず、提案手法では式 (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)。また、 $|\mathbf{u}|$  は  $d$  次元のチェスボード距離、 $\alpha$  は  $K$  の積分値を 1 に正規化するための係数である。

$$K(\mathbf{u}) = \begin{cases} \alpha \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2} \mathbf{u}^T \Sigma^{-1} \mathbf{u}} & \text{if } |\mathbf{u}| \leq \frac{h}{2} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

これは、過去の観測画素値  $\mathbf{X}_i$  が  $\mathbf{X}$  を中心とする窓関数内にある場合のみガウス分布によって決まる値をとり、それ以外の場合には 0 となる関数である。ただし、背景モデルの推定のために十分な数のサンプルが得られる場合には、式 (6) に示す矩形の窓関数を用いても十分な推定精度が得られると考えられる。

$$K(\mathbf{u}) = \begin{cases} \frac{1}{h^d} & \text{if } |\mathbf{u}| \leq \frac{h}{2} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

確率密度関数の推定にこのような窓関数を用いることで、

$\mathbf{X}_{N+1}$  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  $X$  における確率密度関数の値  $P(X)$  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  $X_{N+1}$  が得られた時、 $P(X)$  は以下のようにして更新することができる。

$$P_t(X) = P_{t-1}(X) + \frac{1}{N}K(X - X_{N+1}) - \frac{1}{N}K(X - X_1) \quad (7)$$

これは、新たな観測画素値  $X_{N+1}$  を中心とする窓関数の範囲内の確率を  $K(X - X_{N+1})$  だけ高くし、 $N+1$  フレーム前の観測値  $X_1$  を中心とする窓関数の範囲内の確率を  $K(X - X_1)$  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. Parzen 推定を用いた動的影モデル

この節では、Parzen 推定を用いた動的影モデルの構築法について述べる。画素ごとに影モデルを構築するため、以下ではひとつの画素  $(x, y)$  に注目して考える。影領域の除去には、影の画素値は背景の画素値と比べ色相が変化せず、明るさ成分だけが小さくなっているという性質を利用する。そこで、YUV 表色系において観測画素値  $X$  と背景の画素値  $B$  を比べ、観測画素値の  $Y, U, V$  値が背景の画素値の  $Y, U, V$  値からそれぞれ同じ割合で減少していた場合、その画素を影候補の画素と判定し、その画素値を用いて3節で述べたものと同様の手順で影モデルを更新する。ただし、背景の画素値は、最も観測される確率が高い画素値とする。式(8)~(11)に判定式を示す。

$$\alpha_{min} < \alpha_Y < 1 \text{ with } \alpha_Y = \frac{X_Y}{B_Y} \quad (8)$$

$$\min\{|X_U|, |X_V|\} > \epsilon \quad (9)$$

$$|X_U - \alpha_Y B_U| < \Lambda \quad (10)$$

$$|X_V - \alpha_Y B_V| < \Lambda \quad (11)$$

$\Lambda, \epsilon$  は閾値であり、式(9)が成り立たなかった場合、式(8)のみを用いて判定を行う。

次に、具体的な影領域除去の手順について述べる。

Step1 3節で述べた動的背景モデルにより、背景差分を行う。

Step2 Step1で物体として抽出された場合、影モデルを用いて影領域の除去を行う。影モデルにおいて観測画素値  $X$  が生起する確率  $P(X)$  を求め、 $P(X)$  が閾値以上であれば影、閾値以下であれば物体とする。ただし、影モデルにおいては、観測画素値が影候補であると判定された場合にのみモデルの更新を行うため、画素によっては影モデルの推定に用いる影候補の画素値データが十分でない場合がある。そのため画素値データが不十分な画素では式(8)~(11)によって影領域の除去を行う。

Step3 新たな観測画素値  $X$  が、式(8)~(11)を満たしている場合、その画素値を影候補の画素値とし、3節と同様の手順で影モデルを更新する。

以上の処理を画素毎に繰り返し行うことで影領域の除去を行う。

## 5. 実験

### 5.1 実験1：動的背景モデルに関する実験

実験にはビデオサーベイランスの評価で良く用いられるPETS2001のデータセット(PETS2001)<sup>1</sup>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、分布数可変型の混合ガウス分布による手法[4]、Elgammalらが提案している手法[5]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。 $R, G, B$ それぞれの成分において、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、その画素を抽出対象とする。なお、実験に用いた計算機は、Pentium4 3.2GHzのCPUおよび2.5GBのメモリを搭載している。

<sup>1</sup>International 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。

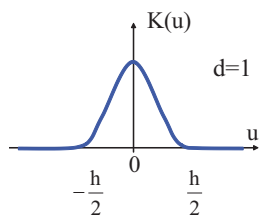


図 1: 提案手法で用いる窓関数

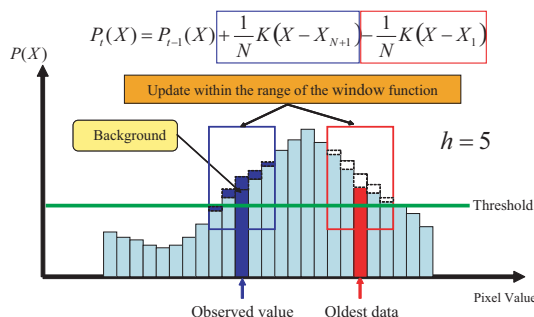


図 2: 背景モデル更新アルゴリズム

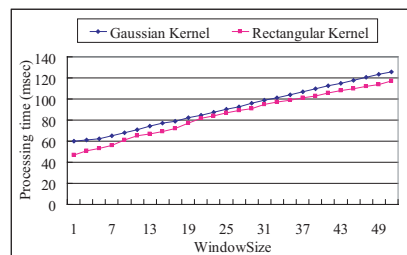


図 3: 窓関数の大きさと処理時間の関係

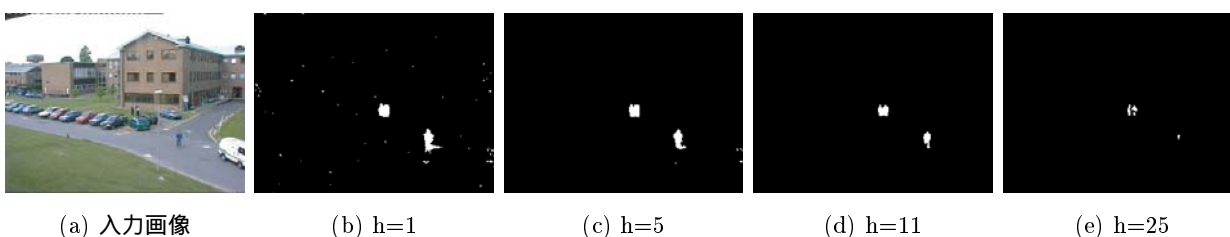


図 4: 窓関数の大きさと背景差分の精度の関係

まず、Parzen 推定に用いる窓関数の大きさについて検証を行った。Parzen 推定において窓関数は重要な役割を持っており、窓関数の大きさによって、推定する確率密度関数の分布の滑らかさを調整することができる。推定される分布の滑らかさが大きすぎる場合は、過度の平滑化が起こり良い推定結果が得られなくなる。一方、滑らかさが十分で無い場合には、確率密度関数が個々の画素値データに強く依存するようになり、観測頻度が低い画素値周辺では、データ過少によるノイズが発生してしまう。従って、良い推定結果を得るためには、窓関数の大きさを適切な値に決めることが重要となる。図 3 に窓関数の大きさ  $h$  と処理時間の関係を示す。横軸は  $h$ 、縦軸は処理時間である。図より、 $h$  の値が大きくなるにつれて処理時間が増加していることがわかる。提案手法では新たな画素値が観測されるたびに、窓関数内の確率密度関数を更新している。そのため、 $h$  の値が大きくなると確率密度関数の更新にかかる時間が増加し、それに伴って処理時間も増加している。実時間性を考慮すると、窓関数の大きさは 25 以下であることが望ましい。

図 4 は式 (5) に示す窓関数において、 $h$  の値を

それぞれ 1, 5, 11, 25 にした場合の背景差分の結果を表す画像である。窓関数の大きさを 1 にした場合、物体領域以外のノイズも前景として抽出されている。これは窓関数を小さくしすぎたために、滑らかな確率密度関数の推定が行えなかったためであると考えられる。一方で、 $h$  を 25 にした場合は物体領域を正しく抽出できていない。これは窓関数を大きくしすぎたために、確率密度関数の過度の平滑化が起こったためと考えられる。以上のことから、処理時間と物体領域抽出の精度のバランスを考えると、窓関数の大きさは 5 ~ 11 程度が妥当であると考えられる。

次に、それぞれの手法を用いたときに 1 フレームを処理するのに必要な計算時間について検証を行った。ここで、提案手法において、式 (5) に示す窓関数を用いている。また、 $h = 5, N = 500$  としている。まず図 5 に提案手法と分布数可変型の混合ガウス分布による手法との比較について示す。横軸は処理したフレーム番号で、左の縦軸はそのフレームを処理するのにかかった時間、右の縦軸はそのフレームでの 1 画素あたりのガウス分布の平均の数を表している。同図中の 2500 フレーム辺りからガウス分布の数が増えている。

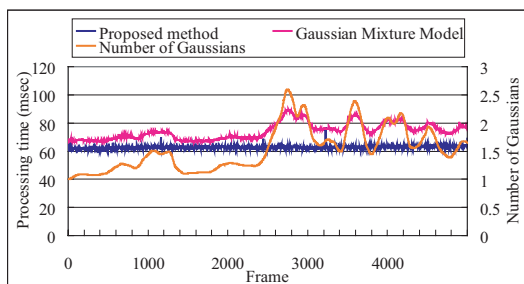


図 5: 分布数可変時の処理時間および 1 画素あたりの平均分布数

このときの映像を見ると、映像に人や車などの物体が現れている時間帯であった。ガウス分布による手法では、これらの物体による映像の変化に対応するため、ガウス分布の数を増やして背景モデルを構築している。しかし、分布の数が増加するに従って処理時間も増加している。一方で提案手法では、どのフレームにおいても高速かつ安定した速度で処理が行えていることがわかる。このことから、処理時間、またその安定性という面から見て非常に優れた結果が得られていることがわかる。

次に、図 6 に Elgammal らの手法との処理時間の比較について示す。Elgammal らの手法では、確率密度関数の推定に用いる画素値データの数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、一定の処理速度で計算を行うことができる。

最後に、物体抽出に関する精度を比較するため、映像を解析したときの再現率と適合率について検証を行った。ここで、再現率、適合率は以下の式で定義される。

$$\text{再現率 (\%)} = \frac{\text{正しく抽出した物体数}}{\text{抽出すべき物体数}} \times 100 \quad (12)$$

$$\text{適合率 (\%)} = \frac{\text{正しく抽出した物体数}}{\text{抽出した物体数}} \times 100 \quad (13)$$

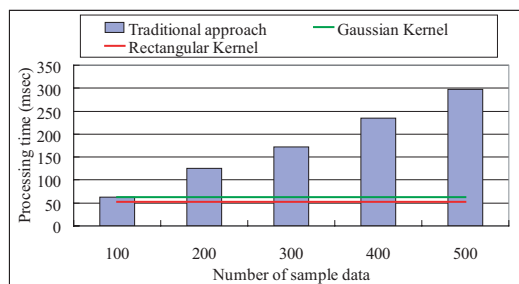


図 6: 確率密度関数推定に用いる画素値データ数と処理時間の関係

なお、提案手法と Elgammal らの手法において、 $N = 500$  とし、提案手法において  $h = 5$  としている。

図 7 に映像を解析したときの再現率と適合率の関係を示す。縦軸は再現率、適合率である。まず、分布数可変型の混合ガウス分布による手法と比較すると、提案手法は非常に優れた結果であることがわかる。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。このときの再現率は 97.2%、適合率は 84.9% であった。しかし、矩形の窓関数を用いた場合にも同程度の精度が出ており、処理時間とのバランスを考えると、矩形の窓関数を用いた場合の方が良いということもできる。

また、提案手法による画像処理結果を図 8 に示す。図 8(a) はあるフレームにおける入力画像である。図 8(b) はそのフレームにおける背景モデルであり、観測される確率が最も高い画素値を表している。図 8(c) は背景差分により得られた物体領域を示す画像である。比較のため、図 8(d) に分布数可変型の混合ガウス分布による手法を用いた場合の結果を示す。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に人物や車の領域を抽出できており、結果の画像からも、提案手法が非常に良い結果が得られていることがわかる。

## 5.2 実験 2: 動的影モデルに関する実験

解像度  $320 \times 240$  で撮影した屋内画像を使用して実験を行った。映像には、複数光源によって歩行者からできた様々な濃さの影が含まれており、単一光源による影に比べ、より複雑なモデルを構築する必要があると考えられる。この映像に

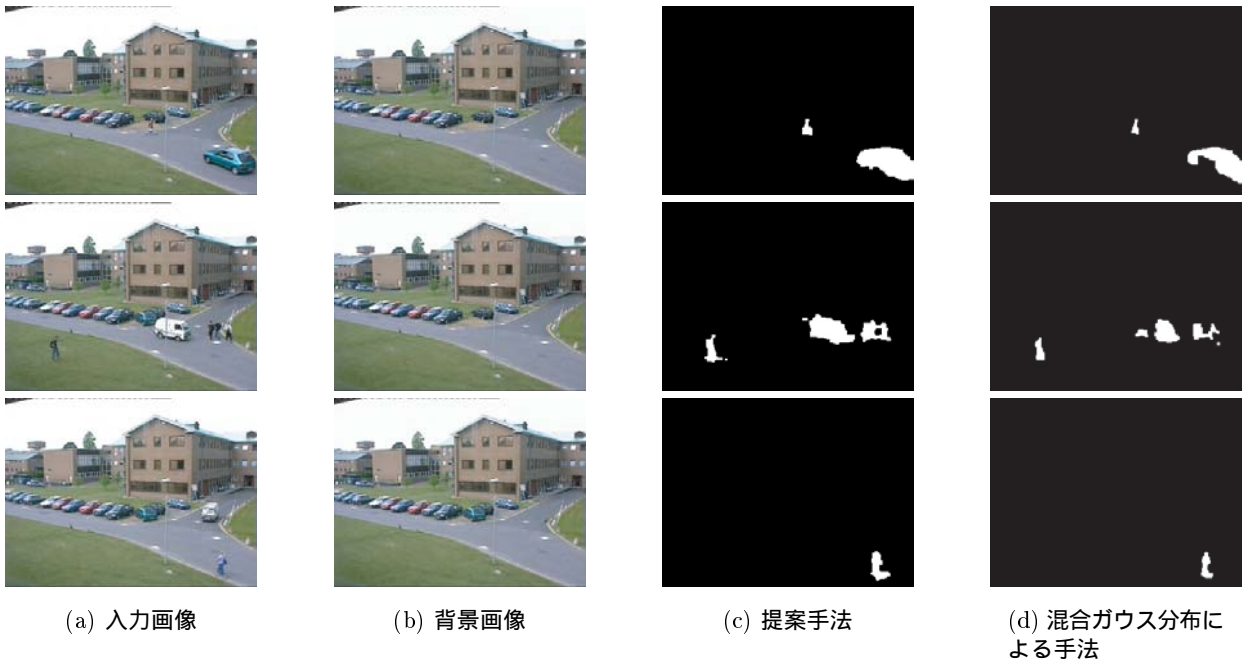


図 8: 背景差分の結果画像．(a) 入力画像，(b) 提案手法における観測される確率が最も高い画素値による画像，(c) 提案手法による前景抽出画像，(d) 混合ガウス分布を用いた手法による前景抽出画像．

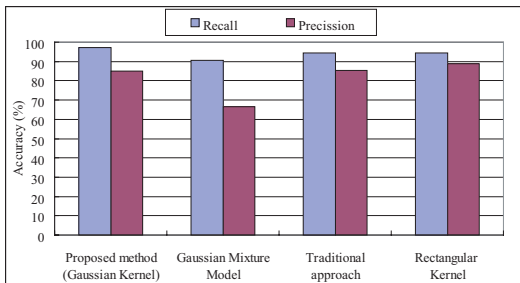


図 7: 再現率・適合率のグラフ

対し，提案する Parzen 推定を用いた動的影モデル，Martel-Brisson らが提案している混合ガウス分布を用いた動的影モデル [11] により影領域の除去を試みた．なお，物体領域抽出にはそれぞれ Parzen 推定を用いた動的背景モデル，混合ガウス分布を用いた動的背景モデルを用いている．

図 9 に，動的影モデルを用いて影領域の除去を行った画像処理結果を示す．図 9(a) はあるフレームにおける入力画像であり，図 9(b) は提案手法により影領域の除去を行った結果の画像である．比較のため，図 9(c) に Martel-Brisson らの手法による影領域除去の結果画像を示す．赤色

で示した画素が影領域として除去される画素である．また，図 9(b) において緑色で示した画素は影モデルの学習が不十分だったために，式 (8) ~ (11) による閾値処理によって影領域の除去を行った画素である．Parzen 推定を用いた動的影モデルを用いた場合の方が，混合ガウス分布を用いた動的影モデルに比べ，歩行者によってできた影領域を高精度に除去できていることが確認できる．また，1 フレームを処理するのに必要な計算時間に関しても，混合ガウス分布を用いる手法が 97msec 程度であったのに対し，提案手法は 88msec 程度であり，処理時間の点からも提案手法の方が優れていることが確認できた．

## 6. おわりに

本稿では，Parzen 推定を用いた動的背景モデル構築の高速化手法について述べた．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手続きをとることで，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することができる．また，物体によってできた影領域の除去法についても，提案する動的背景モデル構築法を応用す



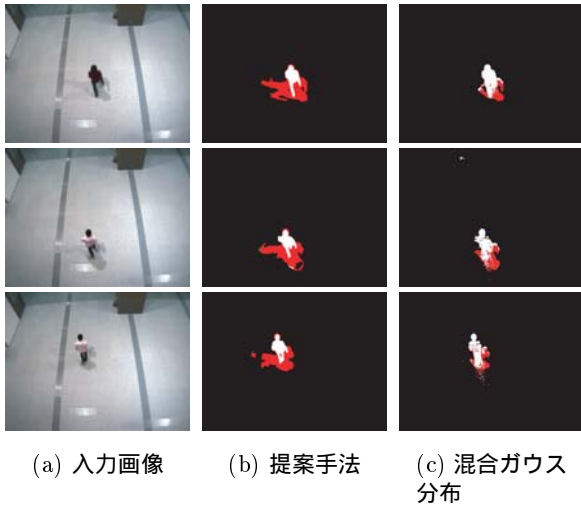


図 9: 影領域除去の結果画像 . (a) 入力画像 , (b) Parzen 推定を用いた動的影モデルによる影領域除去の結果画像 , (c) 混合ガウス分布を用いた動的影モデルによる影領域除去の結果画像 .

ることで、従来の手法と比べ高速かつ高精度に影領域の除去を行うことのできる動的影モデルの構築手法を提案することができた .

実験では、混合ガウス分布を用いた手法や、Parzen 推定を用いた従来の背景モデル推定法に比べ、高速かつ安定した処理時間で背景モデルの更新ができた . また物体領域の抽出精度に関しても、非常に優れていることが確認できた . Parzen 推定を用いた動的影モデルにおいても、混合ガウス分布を用いた動的影モデルに比べ、高速かつ高精度に影領域の除去が行えることが確認できた .

今後の課題と展開としては、影領域除去の精度向上や背景・影モデル構築に使用するメモリ領域の削減などが挙げられる . 前者の問題は、影領域の除去に色情報だけを用いるのではなく、画素間のつながり等の情報を利用することで改善されと考えられる . 後者については、類似性の高いモデルを持つ複数の画素でひとつの背景モデルを構築することで改善されと考えられる .

## 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 and 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 and 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 and Larry Davis: Background and Foreground Modeling using Non-parametric 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-D-II, pp. 796–806 (2003).
- [7] Elena Salvador, Andrea Cavallaro and Touradj Ebrahimi: SHADOW IDENTIFICATION AND CLASSIFICATION USING INVARIANT COLOR MODELS, *Proc. of IEEE International Conference on Acoustics*, Vol. 3, pp. 1545–1548 (2001).
- [8] Rita Cucchiara, Costantino Grana, Massimo Piccardi, Andrea Prati and Stefano Sirotti: Improving Shadow Suppression in Moving Object Detection with HSV Color Information, *2001 IEEE Intelligent Transportation Systems Conference Proceedings*, pp. 334–339 (2001).
- [9] Oliver Schreer, Ingo Feldmann, Ulrich Golz and Peter Kauff: FAST AND ROBUST SHADOW DETECTION IN VIDEOCONFERENCE APPLICATION, *4th IEEE Intern. Symposium on Video Proces. and Multimedia Comm*, pp. 371–375 (2002).
- [10] Alessandro Bevilacqua: Effective Shadow Detection in Traffic Monitoring Applications, *The 11th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision* (2003).
- [11] Nicolas Martel-Brisson and Andre Zaccarin: Moving Cast Shadow Detection from a Gaussian Mixture Shadow Model, *IEEE Computer Society International Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2005).