

時空間特徴を考慮した動的背景モデルによる背景変動に頑健な物体検出

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府**

〒 819-0395 福岡県福岡市西区元岡 744

財団法人九州先端科学技術研究所***

〒 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 本稿では時空間特徴を考慮した動的背景モデルの構築法について述べる。提案手法は画素・領域・フレームレベルの背景モデルから成り、画素レベルの背景モデルでは、ノンパラメトリックな確率密度の推定法である Parzen 推定に基づき、背景の変動を画素毎に確率密度関数でモデル化する。領域レベルの背景モデルでは、Radial Reach Correlation に基づき、注目画素とその周辺画素の明度の大小関係をモデル化する。フレームレベルの背景モデルでは、観測画像の照明条件に合わせた背景画像と観測画像との差分をとることで対象検出を行う。これらの背景モデルの統合により、背景変動に頑健な物体検出を実現する。

キーワード 動的背景モデル, 照明変動, Parzen 推定, Radial Reach Correlation

Object Segmentation Based on Adaptive Background Model Considering Spatio-temporal Features

Tatsuya Tanaka*, Atsushi Shimada**, Daisaku Arita*** and Rin-ichiro Taniguchi**

Department of Intelligent Systems, Kyushu University* **

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

Institute of Systems, Information Technologies and Nanotechnologies***

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 new method for background modeling considering spatio-temporal features. Our method consists of three complementary approaches. The pixel-level background model uses the probability density function to approximate background model. The PDF is estimated non-parametrically by using Parzen density estimation. The region-level background model is based on the evaluation of the local texture at pixel-level resolution while reducing the effects of variations in lighting. And the frame-level background model detects sudden, global changes in the image and estimates a present background image from input image and original background image. Then, the moving object is extracted by background subtraction. Fusing their approaches realizes robust object detection under varying illumination. Several experiments show the effectiveness of our approach.

Key words Adaptive background model, Illumination change, Parzen density estimation, Radial Reach Correlation

1. はじめに

近年、背景差分法を用いた物体検出の研究が盛んに行われている。背景差分法は、事前に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検出することができる。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

けでは、木々や水面の揺らぎ、天候による照明条件の変化の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検出されてしまう。そこで、このような背景の変化に柔軟に対応するための様々な背景のモデル化手法が提案されている [1-5]。

背景モデル構築法は、モデル構築の方法に応じて、画素レベル・領域レベル・フレームレベルの三

つに大別することができる。画素レベルの背景モデルでは、画素単位で画素値の変化を確率的にモデル化する方法がよく利用される [2, 3]。Elgammal らは Parzen 推定を用いて背景画素値の変化を精度よく推定する方法を提案している [3]。画素レベルの背景モデルでは、過去に観測された情報に基づいて、確率モデルによりモデル化を行うため、草木や水面の揺らぎ、照明条件の微小変動のような周期的な画素値の変化に柔軟に対応することができる。しかし、確率モデルを用いた動的背景モデル構築法は、過去に観測された画素値の統計情報により、その変動をモデル化するため、急激な背景変動が起こった場合には本質的に対応することができない。

領域レベルの背景モデルについて、佐藤らは、Radial Reach Correlation (RRC) と呼ばれる、注目画素とその周辺画素との明度の大小関係により局所的なテクスチャを評価する値に基づいて、画素毎に前景・背景を判定する手法を提案している [4]。領域レベルの背景モデルは、注目画素を中心とする小領域において、テクスチャ情報を基に背景モデルを構築するため、画素レベルの背景モデルに比べ、照明変動に頑健であるが、一方で木の揺らぎといったテクスチャの変化を引き起こすような背景の微小変動に対して対応することができないという問題がある。

フレームレベルの背景モデルについて、福井らはフレーム間差分等により大まかに求めた仮の背景領域に基づいて、背景画像から入力画像への画素値の変換テーブルを作成し、画素値の正規化を行う手法を提案している [5]。この手法は、観測画像の照明条件に合わせて背景画像を生成するため、急激な照明変化に頑健であるという特徴を持つ。しかし、画素値の変換テーブル生成の際に、画像全体で一律な照明変化を仮定しているため、非一律な照明変化に対しては対応できない。また、この手法ではあらかじめ用意した背景画像を用いて画素値の正規化を行うため、想定外の背景変動に対応できない。

以上のように、各背景モデル構築法にはそれぞれ、長所・短所が存在するため、頑健な対象検出の実現にはこれらの性質の異なる背景モデルを効果的に統合することが重要であると考えられる。そこで、本論文では画素レベル・領域レベル・フレームレベルの背景モデルによる、統合型の背景モデル構築法を提案する。

2. 時空間特徴を考慮した動的背景モデル

2.1 画素レベルの背景モデル

画素レベルの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的にモデル化する方法が一般的であり、その処理を毎フレーム高速かつ精度

良く行うことが重要である。確率密度関数を高精度に推定するという点では、Parzen 推定を利用した動的背景モデル構築法が有効であり、Elgammal らは Parzen 推定を用いて背景画素値の変化を精度よく推定する方法を提案している [3]。しかし、精度良く推定するためには十分な数のサンプル数が必要であり、Elgammal らによって提案されている手法では確率密度関数の推定に多くの計算が必要である。更に、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、提案手法では、Elgammal らによる手法を改良した、高速な Parzen 推定に基づく動的背景モデル構築法を画素レベルの背景モデルとして採用する。背景モデル構築の詳細については、文献 [6] を参照していただきたい。

2.2 領域レベルの背景モデル

領域レベルの背景モデルでは、RRC [4] を改良した、更新型の背景モデルにより対象検出を行う。以下では、2.2.1 節で RRC について述べ、2.2.2 節で、RRC を改良した動的更新型 RRC による対象検出法について述べる。

2.2.1 Radial Reach Correlation (RRC)

画素ごとに RRC を構成するため、以下では特に断りがない限りひとつの画素 (x, y) に注目して述べるものとする。RRC は、注目画素を中心とする 8 方向の放射状の延長腕 (リーチ) 上において、各方向で注目画素と閾値以上の明度差をもつ画素を探索し、明度値ペアを構成する。また、各ペアにおいて明度差の正負極性を 2 値符号列として求め、入力画像と背景画像におけるその符号列間の相関値を注目画素の類似度とする。

画素 (x, y) の位置をベクトル $\mathbf{p} = (x, y)$ 、放射状のリーチを延長する方向を表す方向ベクトル $\mathbf{d}_k (k = 0, 1, \dots, 7)$ を、 $\mathbf{d}_0 = (1, 0)^T$, $\mathbf{d}_1 = (1, 1)^T$, $\mathbf{d}_2 = (0, 1)^T$, $\mathbf{d}_3 = (-1, 1)^T$, $\mathbf{d}_4 = (-1, 0)^T$, $\mathbf{d}_5 = (-1, -1)^T$, $\mathbf{d}_6 = (0, -1)^T$, $\mathbf{d}_7 = (1, -1)^T$ とする。そうすると、これらの各方向へのリーチ長 $\{r_k\}_{k=0}^7$ は以下の式により求められる。

$$r_k = \min\{r \mid |f(\mathbf{p} + r\mathbf{d}_k) - f(\mathbf{p})| \geq T_p\} \quad (1)$$

ここで、 $f(\mathbf{p})$ は背景画像の位置 $\mathbf{p}$ における画素値、 T_p は明度差分の閾値である。式 (1) により求まるリーチ群によって選択された画素と注目画素との各ペア間の明度差に基づいて、明度差分の極性符号である増分符号の系列 $b_k(\mathbf{p})$ は次式のように定義される。

$$b_k(\mathbf{p}) = \begin{cases} 1 & \text{if } f(\mathbf{p} + r_k\mathbf{d}_k) \geq f(\mathbf{p}) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

ここで、 $k = 0, 1, \dots, 7$ である。同様に入力画像 g の各画素についても増分符号列 $b_k'(\mathbf{p})$ を求める。ここで、 $\{r_k\}_{k=0}^7$ は背景画像に基づいて定義されていることに注意する。

$$b_k'(\mathbf{p}) = \begin{cases} 1 & \text{if } g(\mathbf{p} + r_k \mathbf{d}_k) \geq g(\mathbf{p}) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

以上により得られた $b_k(\mathbf{p})$, $b_k'(\mathbf{p})$ を基に、両増分符号列間の一致数 $B(\mathbf{p})$ を求める。

$$B(\mathbf{p}) = \sum_{k=0}^7 \{b_k(\mathbf{p}) \cdot b_k'(\mathbf{p}) + \overline{b_k(\mathbf{p})} \cdot \overline{b_k'(\mathbf{p})}\} \quad (4)$$

ここで、 $\bar{x} = 1 - x$ はビットの反転を示す。この $B(\mathbf{p})$ は画素 $\mathbf{p}$ 周辺の明度分布の類似度 (相関値) を表すものであり、RRC と呼ばれる。この値を閾値 T_B と比較することにより、類似あるいは非類似を判定する。ここで、その判定結果を示した 2 値画像 $C(\mathbf{p})$ は RRC 画像と呼ばれ、次式で定義される。

$$C(\mathbf{p}) = \begin{cases} 1 & \text{if } B(\mathbf{p}) < T_B \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

2.2.2 動的更新型 RRC による動的背景モデル構築

RRC において、式 (1) により求まる各ペア点は、1 枚の背景画像から明度差分の閾値 T_p に基づいて生成されるものであり、全てのペアの中にはノイズや照明条件の変化等の影響を受けた際に、明度の大小関係が安定に保たれないものが含まれている可能性がある。また、ペア点・符号列は初期フレームにおいて決定され、その後更新されることはないため、時々刻々と変化する背景に適応することができない。そこで、動的更新型 RRC では各画素におけるペア点および符号列を背景の変化に応じて動的に更新する。式 (6) に符号列の更新式を示す。

$$b_k^{t+1}(\mathbf{p}) = (1 - \alpha) \cdot b_k^t(\mathbf{p}) + \alpha \cdot b_k'^t(\mathbf{p}) \quad (6)$$

ここで、 $b_k^t(\mathbf{p})$ は時刻 t において、画素 $\mathbf{p}$ が持つ増分符号列の値であり、 $k = 0, 1, \dots, 7$ である。また、 α は学習率であり、この値をある程度大きな値に設定しておくことで、急な背景変化に対して素早く対応できる。この符号列の更新により $b_k^t(\mathbf{p})$ は $[0, 1]$ の値をとる。従って、この符号が 0 もしくは 1 に近い値である程、安定に明度の大小関係が保たれていることを示す。一方で、0.5 に近い値である程、そのペア間の大小関係が安定に保たれないことを示す。そこで、 $[T_r, 1 - T_r]$ の符号を持つペアに関してはペア点の再選択を行う。ここで、 T_r はペア点再選択の閾値である。

次に、入力画像と各画素が持つ増分符号列間の類似度を以下のように定義し直す。

$$B(\mathbf{p}) = \sum_{k=0}^7 |b_k(\mathbf{p}) - b_k'(\mathbf{p})| \quad (7)$$

この値を閾値 T_B と比較することにより、類似あるいは非類似を判定する。

具体的な背景モデル構築の手順は以下の通りである。

Step1 入力画像から増分符号列を求め、式 (7) により背景モデルの符号列との類似度を計算し、その類似度に応じて背景・前景を判定する。

Step2 各画素の符号列を式 (6) により更新する。

Step3 Step2 において、符号列の値が $[T_r, 1 - T_r]$ になったペアに関しては入力画像を基にペア点の再選択を行う。

2.3 フレームレベルの動的背景モデル

福井らが提案している、背景画像の正規化に基づく対象検出法は、照明変化に頑健な対象検出が可能であるが、本質的には画像上で一様な照明変化に対応することを目的としており、非一様な照明変化に対しては、安定に対象検出を行えないという問題がある。また、背景画像を事前に用意しておくため、想定外の背景変動に対応することができない。

そこで提案手法では事前学習を必要とせずに、画像上で非一様な照明条件の変化に対しても対応可能な手法により、背景画像の画素値を照明変化に対して正規化し、背景差分を行う。以下では、フレームレベルの背景モデルで用いる背景画像の推定方法と画素値の正規化に基づく背景差分法について述べる。

提案手法では、画素値の正規化に多層パーセプトロンを用いる。これにより、入力ベクトル (背景画像の画素値) と出力ベクトル (入力画像の画素値) の間の線形および非線形の関係を学習することができる。近傍画素間の共起性を考慮するため、入力ベクトルに各 (x, y) 座標値とその座標における画素値 (R, G, B) を合わせた 5 次元ベクトルを用い、ネットワークの出力として現在時刻におけるその画素の画素値 (R, G, B) を得る。色情報に位置情報を加えたベクトルを用いることで、画像上の位置を考慮した画素値の変換テーブルを構築することができ、画像上の非一様な画素値の変化に対しても対応できる。多層パーセプトロンを用いた背景画像推定の流れは以下の通りである。

Step1 背景画像の座標 (x, y) における画素値 (R, G, B) と、観測画像における同じ座標での画素値 (R, G, B) のペアによる学習データを背景候補領域の画素毎に作成する。また、作成した学習データ

を用いて多層パーセプトロンの学習を行う。

Step2 Step1 で学習した多層パーセプトロンを用いて、背景画像に対し、画素値の正規化を行い、現在時刻における背景画像を推定する。

Step3 推定された背景画像と入力画像の差分をとり、画素値の差が閾値以上の画素を対象領域として検出する。

2.4 時空間特徴を考慮した動的背景モデル構築

この節では、画素・領域・フレームレベルの背景モデルによる、時空間特徴を考慮した動的背景モデル構築について述べる。具体的な処理の流れは以下の通りである。

Step1 画素レベルの動的背景モデルで対象検出を行う。

Step2 領域レベルの動的背景モデルで対象検出を行う。

Step3 Step1, Step2 の対象検出結果を統合する。また、統合結果に基づいて各背景モデルが持つパラメータを更新する。ここでは、両モデルにおいて前景と判定された画素のみ、統合結果を前景とする。それ以外の画素については背景と判定する。なお、各背景モデルが持つパラメータについては、統合結果により背景と判定された画素のみ、両モデルのパラメータを更新し、それ以外の場合は、画素レベルの背景モデルのみパラメータを更新する。

Step4 現時刻の入力画像と前フレームでの入力画像でフレーム間差分を行い、画像上である一定以上の割合の画素に変化があった場合、フレームレベルの背景モデルに対して、TTL(Time To Live)を設定する。TTLはフレームレベルの背景モデルを使用する期間を表わし、TTLを設定することで、急激な照明変化があった場合のみフレームレベルの背景モデルを利用することになる。

Step5 TTLが0でなければ、フレームレベルの背景モデルにより対象検出を行い、得られた結果を最終的な対象検出結果とする。また、TTLを減少させる。(フレームレベルの背景モデルを利用しない場合は、Step3 で得られた対象検出結果を最終的な対象検出結果とする。) なお、フレームレベルの背景モデルで用いる背景画像については、画素レベルの背景モデルにより、最頻値画像を作成し、これを背景推定用の背景画像とする。また、Step3 で得られた対象検出結果を背景候補領域とする。

3. 実験

提案手法の有効性を示すため、ビデオサーベイランスの評価で良く用いられるPETS(PETS2001)¹ および、解像度 320×240 、15fps で撮影した屋

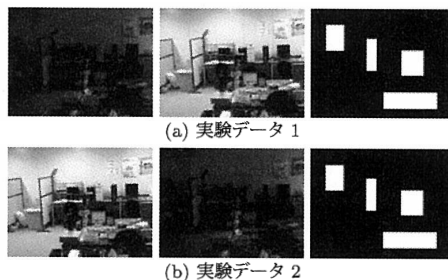


図 1: 実験データ

内映像に対し、実験を行った。なお PETS2001 の映像については、 320×240 の画像サイズに変換して利用した。PETS2001 の映像には通りを行き交う人々や急激な天候の変化などが含まれている。また屋内映像には、電灯の点灯・消灯による激しい照明条件の変化が含まれている。なお、実験に用いた計算機は、Intel Core2 3.2Hz および 4.0GB のメモリを搭載している。

3.1 背景画像の推定精度に関する検証

背景画像の推定精度について実画像を用いて検証した。図 1 に実験データを示す。図中の左列は観測画像、中央の列は背景画像、右列は観測画像中の背景領域を示している。この実験データの入力画像には、照明の点灯・消灯による、画像上で非一様な照明条件の変化が含まれており背景画像と観測画像の背景領域の情報をもとに、観測画像と同じ照明条件の背景画像を推定する。

これらの実験データを用いて、ニューラルネットワークおよび福井らの変換テーブルによる手法 [5] により背景画像の推定を行い、その結果を比較した。なお、提案手法において、5 入力、3 出力、中間層に 3 つのニューロンを持つ 3 層のパーセプトロンを用いた。なお、活性化関数にはシグモイド関数を用い、学習アルゴリズムには弾力性バックプロパゲーションアルゴリズムを用いた。また、画素間の近傍性を考慮することの有効性を示すため、入力ベクトルを (R, G, B) の 3 次元ベクトルとしたニューラルネットワークについても実験を行った。

表 1, 2 に推定精度を示す。結果から、明度の高い画像から、明度の低い画像を推定する場合の方が精度よく推定できていることがわかる。また、変換テーブルを用いた手法や、入力として (R, G, B) の 3 次元のみを用いたニューラルネットワークを用いた場合に比べ、ニューラルネットワークの入力として (X, Y, R, G, B) の 5 次元を用いた場合の方が推定誤差が小さくなっている。変換テーブルや、色情報のみを用いたニューラルネットワークでは、

¹Benchmark data of International Workshop on Performance Evaluation of Tracking and Surveillance.
From <ftp://pets.rdg.ac.uk/PETS2001/> available.

表 1: 背景画像の推定精度 (実験データ 1)

	平均誤差	標準偏差
変換テーブル	9.5	13.9
NN (R,G,B)	8.5	10.4
NN (X,Y,R,G,B)	5.4	9.3

表 2: 背景画像の推定精度 (実験データ 2)

	平均誤差	標準偏差
変換テーブル	41.5	40.4
NN(R,G,B)	44.9	28.6
NN(X,Y,R,G,B)	28.6	30.4

位置情報をまったく考慮していないため、画像上で非一様な変化が起こった場合に対応できなかったと考えられる。

3.2 対象検出精度による評価

対象検出精度に関しては、図 2 に示す 2 通りの映像を用い、提案手法、RRF、Parzen 推定を利用した高速な動的背景モデル構築法 [6]、分布数可変型混合ガウス分布による手法 [2] を用いて実験を行い、人手で与えた ground truth² を基に映像を解析したときの再現率・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。

$$\text{再現率 (\%)} = \frac{\text{正しく検出した画素数}}{\text{検出すべき画素数}} \times 100 \quad (8)$$

$$\text{適合率 (\%)} = \frac{\text{正しく検出した画素数}}{\text{検出した画素数}} \times 100 \quad (9)$$

表 3, 4 に映像を解析した時の再現率・適合率を示す。なお、画素レベルの背景モデルにおいて $h = 9$, $N = 500$, 領域レベルの背景モデルにおいて $T_p = 10$, $T_B = 5.5$, $T_r = 0.3$, フレームレベルの背景モデルにおいて、背景差分の閾値を 30, TTL の初期値を 60 とした。TTL の初期値については、急激な画素値の変化が起こった際に、画素レ



(a) PETS2001



(b) Indoor scene

図 2: 実験データ

²実験データを基に作成した、各フレームで物体領域として検出すべき領域を示す画像。この画像系列の ground truth を含め数種類を Web で公開している。http://limu.is.kyushu-u.ac.jp/dataset/

表 3: 物体検出精度 (屋外映像)

	recall	precision
Proposed method	77.9%	69.1%
Radial Reach Filter	37.5%	22.4%
Parzen density estimation	56.3%	51.6%
Gaussian Mixture Model	61.3%	58.2%

表 4: 物体検出精度 (屋内映像)

	recall	precision
Proposed method	76.1%	77.4%
Radial Reach Filter	26.9%	24.9%
Parzen density estimation	37.8%	58.5%
Gaussian Mixture Model	35.6%	46.1%

ベルの背景モデルが対応するまでに必要なフレーム数の 2 倍程度を目安としている。表 3, 4 の結果から、提案手法が RRF, Parzen 推定を利用した高速な動的背景モデル構築法、分布数可変型混合ガウス分布を用いた手法に比べ、検出精度が優れていることがわかる。

次に、PETS2001 の映像に対し、提案手法を用いて対象検出を行った結果画像を図 3 に示す。比較のため、図 3(d)~3(f) にそれぞれ RRF, Parzen 推定を利用した高速な動的背景モデル構築法、分布数可変型混合ガウス分布を用いた手法による対象検出結果を示す。ここで、RRF では初期フレームを背景画像 f として用いている。

まず図 3(d) について見ると、RRF は照明変化には対応できているが、初期フレームにおいてあらかじめ構築しておいた背景モデルを利用しているため、雲の移動等による背景の変化に対応できず、対象領域でない画素が過検出されていることがわかる。また、Parzen 推定を利用した高速な動的背景モデル構築法や、分布数可変型混合ガウス分布による手法では、それらの背景変化には対応できているが、急激な照明変化に対して対応できておらず、地面や建物の壁などが誤検出されている。一方提案手法では、それらの手法に比べ、背景の変化や照明変化に頑健に人物の領域を検出できていることが分かる。

3.3 計算時間についての実験

図 2(b) の映像を用いて、1 フレームを処理するのに必要な計算時間を測定した。図 4 に時空間特徴を統合した動的背景モデルの計算時間の測定結果を示す。横軸が処理したフレーム番号、左の縦軸がそのフレームを処理するのに要した時間を表している。また、提案手法における各パラメータについては、3.2 節の実験と同じ値に設定した。

まず、画素レベル・領域レベルの背景モデルについて見ると、それぞれ 30msec, 15msec 程度の計算時間で処理を行うことができています。フレーム

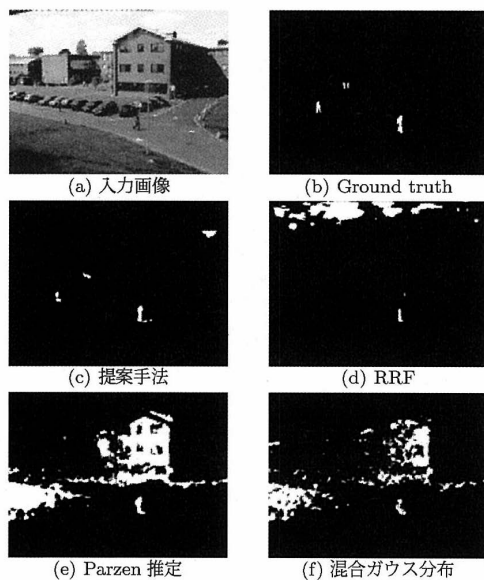


図 3: 背景差分の結果画像

レベルについて見ると、フレームレベルの背景モデルは、急激な画素値の変化があったときのみ利用されるため、基本的には計算時間がほぼ 0 となっている。フレームレベルの背景モデルが利用されたフレームについて見ると、その計算時間は 50～80msec 程度の計算時間を要していることがわかる。これは、全体の計算時間の約 5 割程度を占めており、画素レベル・領域レベルの背景モデルに比べ、計算コストが大きいことがわかる。それに加え、フレームレベルの背景モデルは他の手法と比較して計算時間が不安定であるため、フレームレベルの背景モデルが利用されるフレームでは、計算時間が不安定になる傾向が見られる。したがって、フレームレベルの背景モデルに要する計算時間の削減、また、統合後の背景モデル全体としての計算時間の安定化が今後の課題となる。

4. おわりに

本論文では、時空間特徴を考慮した動的背景モデル構築とそれに基づく物体検出について述べた。性質の異なる複数の背景モデルを統合することで、従来の動的背景モデルに比べ、背景変動に頑健であることが確認できた。

今後の課題としては以下のことがあげられる。

●背景画像の推定精度の向上：

ニューラルネットワークによる背景画像の画素値の正規化の際に、明度の低い画像から、明度の高い画像を推定する場合、背景画像の推定精度が低くなる傾向が見られた。背景画像の推定精度はフレームレベルの背景モデルの対象検出精度に直結するため、この推定精度を向上させるための工夫

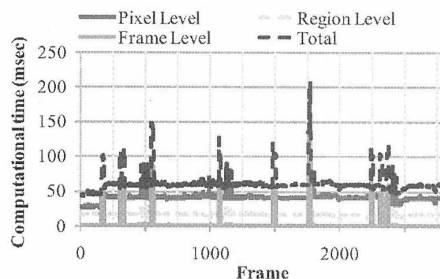


図 4: 時空間統合型背景モデルの計算時間

が必要である。

●計算時間の安定化：

フレームレベルの背景モデルは、画素レベル・領域レベルの背景モデルに比べ、計算コストが大きく、またその計算時間は不安定である。したがって、フレームレベルの背景モデルが利用されるフレームでは、計算時間が不安定になる傾向が見られる。したがって、フレームレベルの背景モデルに要する計算時間の削減、また、統合後の背景モデル全体としての計算時間の安定化が今後の課題となる。

謝辞

本研究の一部は、(株)オムロンとの共同研究により実施されたものである。

参考文献

- [1] K. Toyama, J. Krumm, B. Brumitt and B. Meyers: "Wallflower: Principle and Practice of Background Maintenance", International Conference on Computer Vision, pp. 255-261 (1999)
- [2] 島田, 有田, 谷口: "適応的な分布数の増減法を利用した混合ガウス分布による高速な動的背景モデル構築", 電子情報通信学会論文誌, **J90-D**, 9, pp. 2606-2614 (2007)
- [3] A. Elgammal, D. Harwood and L. Davis: "Non-parametric Model for Background Subtraction", 6th European Conference on Computer Vision (ECCV), **2**, pp. 751-767 (2000)
- [4] 佐藤, 金子, 丹羽, 山本: "Radial Reach Filter(RRF) によるロバストな物体検出", 電子情報通信学会論文誌, 第 86 巻, pp. 616-624 (2003)
- [5] 福井, 石川, 岩堀, 伊藤: "輝度変化に対する頑健な背景差分法", 画像の認識・理解シンポジウム (MIRU2004), **2**, pp. 305-310 (2004)
- [6] 田中, 島田, 有田, 谷口, 富浦: "高速な parzen 推定を用いた動的背景モデルによる映像からの物体検出", 映像情報メディア学会誌, **62**, 12, pp. (141)2045-(148)2052 (2008)