

## 事例ベース背景モデリング

島田 敬士<sup>†</sup> 長原 一<sup>†</sup> 谷口 倫一郎<sup>†</sup>

<sup>†</sup>九州大学 大学院システム情報科学研究所

〒 819-0395 福岡市西区元岡 744 番地

E-mail: †{atsushi,nagahara,rin}@limu.ait.kyushu-u.ac.jp

あらまし 映像サーベイランスなどにおいて物体検出の基盤技術として利用される背景モデルは、高性能化と低コスト化について議論がなされてきた。様々な背景変動に対応するために背景モデルの高性能化を行うと計算コストとメモリ使用量が増えてしまう。当然、計算コストを軽減させる場合には性能が犠牲になる。本稿ではこのようなトレードオフの問題に対するブレイクスルーを事例ベース背景モデリングという枠組みで提案する。画素単位でモデルを保持する代わりに、画像全体で事例化された代表の背景モデルを必要数だけ保持する。各画素は高速に事例を参照することで背景差分が実行される。本研究の貢献は、事例の構築を高速かつ省メモリで行うことと、事例の選択方法に工夫を施すことにより事実上背景モデルを高性能化していることにある。本フレームワークを統計的背景モデルに適用し、従来法よりもメモリ使用量を半減でき、処理速度が約 50% 向上し、さらに物体検出性能も高くなることが確認された。キーワード 背景モデリング, 事例ベース, 物体検出, 低コスト化, 高性能化

### 1. はじめに

背景モデリングは、観測シーン内で起こる背景変動をモデル化する手法であり、出現する物体に関する事前知識を必要とすることなくその領域を検出することができるため、映像解析の基盤技術として幅広く研究されている。従来、背景モデリングは、様々な背景変動に対応するためのモデルの高性能化と、モデルを少ないメモリ使用量や短時間の計算で実現する低コスト化の点で議論されることが多いが、その両立を実現する画期的な手法は未だ確立されていない。

背景モデルの高性能化は、背景モデルの表現能力を向上させることで実現されるケースが多い。統計的背景モデルでは、混合ガウス分布 [1], [2] あるいはカーネル密度推定 [3] により、過去に観測された画素値の出現確率をモデル化している。画素 1 点のみの情報を利用するのではなく、局所領域内の特徴を利用する方法 [4] ~ [8] では、各画素の表現できる画像特徴を増やすことで背景変動への頑健性を向上させている。さらには、複数の背景モデルをハイブリッド的に組み合わせることで、画素の持つ情報量を増やしている [9], [10]。

上記のような背景モデルの高性能化を行うと、背景モデル構築に必要なメモリ量や計算コストが飛躍的に増加するという問題が生じる。混合ガウス分布による統計的背景モデル [1], [2] を例に取ってみても、過去に観測した多くの背景変動をモデル化するために混合数を増加させるという措置がなされる。すると、ガウス分布を表現するために必要なメモリ量が増え、さらにモデルを更新する際の手続きにかかる時間も増加することになる。一方で、混合ガウス分布の要素数を減らすなどの低コスト

化を行うと、モデルの表現能力が低下して、結果として背景変動への頑健性が低下するという問題が起こってしまう。

そこで本研究では、背景モデルの高性能化と低コスト化のトレードオフの問題に対するブレイクスルーを「事例ベース背景モデリング」という枠組みで提案する。従来のように画素単位で背景モデルを保持する代わりに、画像全体で事例化された代表モデルを必要数だけ保持することで、モデル構築に必要なメモリ使用量を削減させる。また、計算コストを減らすために、背景モデルを事例として高速に登録ならびに参照する仕組みを実現する。さらに、事例を選択する際に利用する画素特徴に工夫を施すことで、背景モデルの表現能力を事実上高めることが可能になり、結果として背景変動への頑健性が増す。提案する事例ベース背景モデリングのフレームワークは、本稿で紹介する統計的背景モデルへの適用に留まらず、これまでに提案されている多くの背景モデルへ応用可能であり、背景モデリングに対する事例ベースアプローチの幅広い普及が期待できる。

類似した画素値が観測される画素をクラスタリング手法により発見し、各クラスタ内のメンバ（画素群）でひとつの背景モデルを共有する手法 [11] が提案されている。画像全体で保持する背景モデルの数がクラスタの数で制限されるので、モデル構築に必要なメモリ使用量を節約できる。また、モデルの更新にかかる時間もモデル数が削減された分だけ短縮されるという利点がある。しかし、クラスタ内での背景変動は同一であることを仮定しているため、メンバの一部の画素のみに背景変動が起こるような場合では、検出精度が低下するという問題が生じる。そのような場合、画素の再クラスタリングが必要になる

が、クラスタリングには計算時間を要するため、頻繁に再クラスタリングが必要となる背景変動の激しいシーンへの適用は難しい。画像全体の画素をクラスタリングの対象とするのではなく、道路や建物の壁のように単一色で表現できる領域のみをクラスタ化して代表モデルを保持し、それ以外の画素は個別に背景モデルを保持する手法 [12] も提案されている。このような方法であっても、クラスタリングを再度実行する必要があるような照明変動の激しいシーンでは、クラスタの再構築に時間を要するため実時間性を必要とする環境での実用化は難しい。

## 2. 事例ベース背景モデリング

### 2.1 フレームワーク

本節では、事例ベース背景モデリングのフレームワークについて述べる。提案手法には、大きく 2 つの重要なコンセプトが存在する。

- 各画素が利用する背景モデル (事例) はフレーム時刻毎に再選択を行う
- 必要に応じて新たな事例の追加と不要な事例の削除を行う

第 1 に、フレーム時刻毎に各画素に最も適切な事例を選び直すことで、クラスタ内では同一の背景変動が起ると仮定していた従来研究の問題点を解消する。現在までに作成されている背景モデルを  $M_j$  ( $j$  は事例の番号) とすると、事例  $M_j$  を利用する画素特徴の代表ベクトル  $p_j$  と類似した画素特徴を持つ画素は、同じ事例  $M_j$  を共有する (図 1 では、7 つの画素の画素特徴 (図左の黒い点) が代表ベクトル  $p_1$  あるいは  $p_2$  と類似しており、その代表ベクトルに紐付けされた背景モデルを利用している例を示している)。したがって、フレーム時刻毎に各画素でその画素特徴  $q$  に対して類似した代表ベクトル  $p_j$  を探索する必要があるが、本研究ではハッシュに基づく代表ベクトルの選択により、計算コストの削減を行っている。これについては、2.2 節で詳しく述べる。

第 2 に、画素特徴  $q$  に類似する代表ベクトル  $p_j$  が見つからない場合は、 $q$  を新たな代表ベクトル  $p_{j+1}$  として、それに対応する背景モデルを新たな事例  $M_{j+1}$  として追加する (図 2 は、ひとつの画素特徴に類似する代表ベクトルが見つからなかったため、代表ベクトル  $p_3$  を新たに追加し、それに対応する背景モデルの事例を構築している例である)。このようにあらかじめ考えられる事例を登録しておくのではなく、登録されている事例では対応できない場合に逐次的に事例を増やすことにより、新たな画素特徴が得られた場合に対応する仕組みを提供する。一方で、提案手法では背景モデルを適宜削除する仕組みも提供する。背景モデルの削除方針とその方法については、2.4 節で詳しく述べる。

### 2.2 ハッシュに基づく背景モデルの登録と参照

背景モデルの事例  $M_j$  を利用する画素特徴の代表ベク

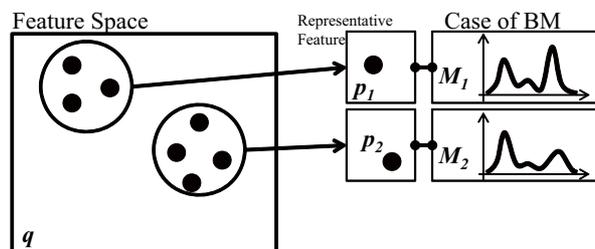


図 1 フレーム時刻  $t$  での背景モデル選択

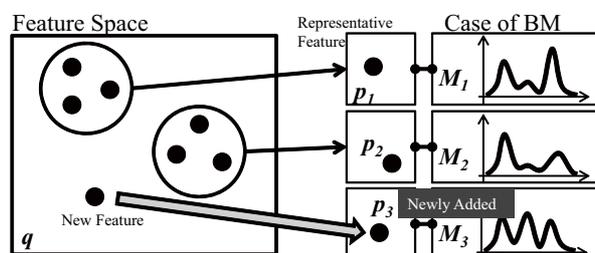


図 2 フレーム時刻  $t+1$  での背景モデル再選択と新規登録

トル  $p_j$  を高速に探索するために、ハッシュに基づく探索手法を導入する。LSH (Locality Sensitive Hashing) は、ハッシュに基づく近似最近傍探索手法 [13], [14] のひとつであり、局所鋭敏 (Locality Sensitive) なハッシュ関数を用いることで、距離の近い特徴ベクトル同士が同じハッシュ値を取る確率が高くなり、距離の遠い特徴ベクトル同士は同じハッシュ値を取る確率が小さくなる。LSH では、このような条件を満たすハッシュ関数を  $k$  個生成して、それらをひと組のハッシュ関数群として、最近傍候補を抽出する。さらにこのような処理を  $L$  個のハッシュ関数群で行い、各ハッシュ関数群から得られる最近傍候補の和集合を最終的な最近傍候補をする。つまり、LSH では  $k \times L$  個のハッシュテーブルを作成して最近傍候補を抽出していることになる。本研究では、ベクトル空間で利用可能な文献 [13] の LSH で利用されている次式で表される局所鋭敏なハッシュ関数を利用している。

$$h(p) = \left\lfloor \frac{a \cdot p + b}{w} \right\rfloor \quad (1)$$

$p$  は特徴ベクトル、 $a$  は各次元の要素の値をガウス分布から独立に取得したベクトル、 $b$  は区間  $[0, w]$  からランダムに選ばれた実数、 $w$  はハッシュ幅である。

本研究では、クエリとなる画素特徴ベクトル  $q$  に対して、上記のハッシュ関数群により最近傍候補となる代表ベクトルが見つかった場合は、その候補に割り当てられている背景モデルの事例を利用する。一方で、最近傍候補が見つからない場合は、 $q$  を新たな画素特徴の代表ベクトルとし、それに対応する新しい背景モデル作成し、それを事例として登録する。

### 2.3 背景モデルの更新

背景モデルの事例  $M_j$  は複数の画素から参照されることがある。このような場合、背景モデルの更新は事例を

利用した画素の中からランダムに選択したひとつの画素を利用して行う<sup>(注1)</sup>。一方で、どの画素からも参照されなかった事例については、背景モデルの更新を行わない。

## 2.4 背景モデルの削除

背景モデルの事例  $M_j$  は、生存時間 TTL を保持する。最後に事例が参照されたフレーム時刻から TTL の間、利用されない状態が続いた場合には、当該背景モデルをメモリから削除する。

## 3. 事例ベース背景モデリングの統計的背景モデルへの適用

本節では、事例ベース背景モデリングのフレームワークを、混合ガウス分布による統計的背景モデルへの適用する方法について述べる。

### 3.1 処理手順

図 3 に処理の手順を示す。フレーム時刻  $t$  における入力画像  $I$  の注目画素  $i$  についての処理に注目して説明する。なお、以下の説明では事例の検索に利用する画素  $i$  の特徴ベクトルを  $q_i$  とし、画素  $i$  の画素値を  $X_i$  とし、敢えて区別して表記している。これは、本研究では特徴ベクトルに画素値  $X_i$  を利用するのではなく、3.2 節で述べるような特徴ベクトルを利用するためである。その理由と利点については 3.2 節を参照されたい。

**事例の探索** 代表ベクトルには背景モデルの事例が紐付けされているため、事例の探索は画素特徴の代表ベクトルの探索により行われる。画素  $i$  から得られる特徴ベクトル  $q_i$  に類似する画素特徴の代表ベクトルを近似最近傍探索により得る。

**背景差分** 事例が発見された場合は、その事例に割り当てられている背景モデルを利用して背景差分を実行する。この結果、画素には前景か背景のラベルが与えられる。

**事例の新規追加** 事例が発見されなかった場合は、特徴ベクトル  $q_i$  を新たな画素特徴の代表ベクトルとして登録し、それに対応する背景モデルを画素  $i$  の画素値  $X_i$  を用いて構築し、新規事例として登録する。また、このとき画素  $i$  は対応する背景モデルが存在しなかったため前景画素として判断する。

**背景モデル更新** すべての画素について上記処理を終えた後、背景モデルの更新を行う。1 回でも参照された事例は背景モデルの更新を行う。一方、1 回も利用されなかった背景モデルについてはモデルの更新を行わない。

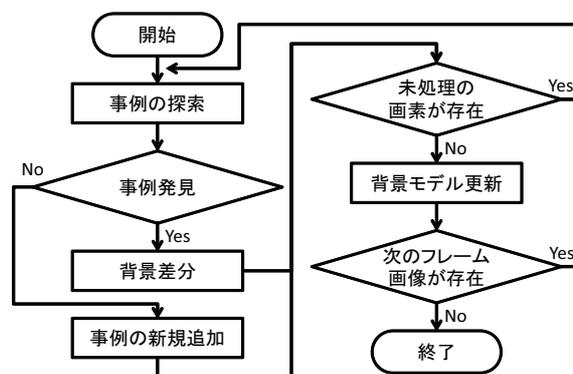


図 3 事例ベース統計的背景モデルの処理の流れ

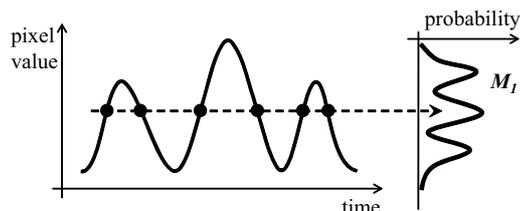


図 4 画素の特徴ベクトルに現在の画素値  $X^t$  を利用する場合

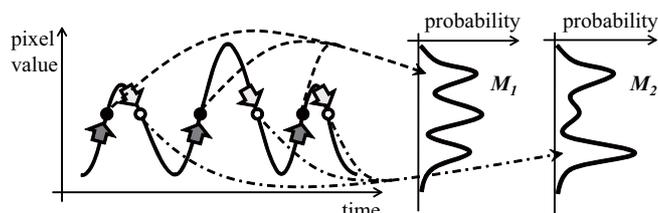


図 5 画素の特徴ベクトルに画素値  $X^{t-1}$  と現在の画素値  $X^t$  を利用する場合

### 3.2 事例探索に利用する画素の特徴ベクトル

事例検索には画素  $i$  の画素値  $X_i$  を直接利用せずに、画素値  $X_i$  を含む特徴ベクトル  $q_i$  を利用する。具体的には、フレーム時刻  $t$  と画素  $i$  の画像座標  $(u_i, v_i)$  を導入し、

$$q_i = (X_i^{t-1}, X_i^t, u_i, v_i)^T \quad (2)$$

としている。これは、画像内の位置が近くて連続するフレームでの画素値の変化の傾向が類似した事例を検索することになる。こうすることにより次のような利点が得られる。尚、以下の説明では表記を簡単にするために画素  $i$  の添え字は省略している。

まず、前フレーム時刻での画素値  $X^{t-1}$  と現在の画素値  $X^t$  を利用する場合と、現在の画素値  $X^t$  のみを利用する場合での違いを図 4 と図 5 を利用して説明する。各図において、左側のグラフは、注目画素で各時刻で得られた観測画素値の時間遷移を表している。また、同図右側は、輝度値に対する過去の観測頻度である。つまり、各図の右側は統計的背景モデルを示している。図 4 は、現在の画素値  $X^t$  のみを利用する場合の例である。図中

(注1): 同一事例を利用した画素の平均を利用する方法や、代表的なものを利用する方法など、選択方法には様々考えられるが、本研究ではランダムに画素を選択する手法を採用した。これは平均の計算などに要するコストを減らすためである。

では、ある画素値を取る時刻を黒い丸で表しており、そのような画素値を特徴ベクトルとして事例を探索して得られた背景モデルを  $M_1$  としている。背景モデル  $M_1$  には、観測画素値がどのような変化の経緯によって観測された画素値であるのかという情報は含まれない。そのため時間変化の情報を考慮した背景モデルを構築するためには、時間遷移をモデル化する手法を別途導入して統計的背景モデルと組み合わせて利用されていた。

これに対して、図 5 に示すように前フレーム時刻での画素値  $X^{t-1}$  と現在の画素値  $X^t$  を利用すれば、フレーム時刻間での画素値の変化が類似した事例を探索できる。これにより、たとえ同じ輝度値が観測されていても、その経緯が異なれば異なる背景モデルを選択できることになる。図では、画素値が上昇傾向にある場合は背景モデル  $M_1$  が選択され、画素値が下降傾向にある場合は背景モデル  $M_2$  が選択されている。このように提案手法では、時間的な変化の傾向を画素特徴の代表ベクトルで事例化し、そのときの観測画素値を背景モデルとして事例化していることになる。同じ画素値であっても変化の傾向が異なれば、異なる背景モデルを準備する必要がありメモリ使用量の増加が懸念されるが、実際には同様の变化傾向が見られる複数の画素で背景モデルを共有するため、さほど問題にはならない。これについては実験結果でも実証する。

次に、画像座標  $(u, v)$  を導入する利点について述べる。まず、画像座標間の距離が小さい画素同士では、類似した画素値の変化が観測されやすいことは想像するに難しい。一方で、画像座標間の距離が大きい画素同士で背景モデルを共有することを考える。利点としては、画像全体で利用する背景モデルの数を節約できることが挙げられるが、一方で、距離の離れた画素では偶発的に類似した画素値変化が起こった可能性もある。このような問題を回避するために画像座標  $(u, v)$  により画素の局所性を導入することにした。その有効性については実験結果で詳しく分析する。

### 3.3 統計的背景モデリング

混合ガウス分布を利用した統計的背景モデルは、時刻  $t$  までに観測された画素値  $\{X^1, \dots, X^t\}$  を  $K$  個のガウス分布を用いてモデル化する方法である。すなわち、各画素において画素値  $X^t$  の生起確率は、

$$P(X^t) = \sum_{k=1}^K w_k^t \eta(X^t | \mu_k^t, \Sigma_k^t) \quad (3)$$

として表現される。ここで、 $w_k^t$ 、 $\mu_k^t$ 、 $\Sigma_k^t$  は時刻  $t$  において  $k$  番目のガウス分布の重み、平均値、分散共分散行列である。また  $\eta$  は式 (4) で表されるガウス分布の密度関数である。

$$\eta(X^t, \mu^t, \Sigma^t) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X^t - \mu^t)^T \Sigma^{-1} (X^t - \mu^t)} \quad (4)$$

$K$  はガウス分布の数で、背景変動に対して適応的に増減する。

統計的背景モデルでは、分布のパラメータセット  $w_k^t$ 、 $\mu_k^t$ 、 $\Sigma_k^t$  を利用して、高い頻度で観測される画素値を背景として認識する。また、新たな観測画素値に適合するようにこれらのパラメータは更新される。詳しくは、文献 [2] を参照されたい。

## 4. 実験結果

### 4.1 実験環境

公開画像データセットを利用して提案手法の有効性を検証した。画像サイズは  $320 \times 240$  画素、画像フレーム数は最も長いシーンで 5300 フレームほどのシーンである。計算機は Intel Core i7 2.80GHz の CPU を利用した。LSH のパラメータは事前実験により検証を行い、 $k=3$ 、 $L=1$  として以降の実験で共通して利用した。また、事例の生存時間 TTL については、4.4 節の実験を除いて 10000 とした。これは実験に用いた評価映像のフレーム数を超える値であり、事例は削除されることはない。事例の削除も含めた実験結果を示すと、背景モデルを事例化することの有効性の検証が難しくなるためこのような設定とした。

### 4.2 物体検出精度

図 6 に示す 3 つの異なるシーン<sup>(注2)</sup> について物体検出精度を検証した。提案手法と精度を比較するために、2 つの統計的背景モデルを比較手法として採用した。ひとつは、提案手法でも利用している混合ガウス分布の要素数を適応的に決めることができる背景モデル [2] (以下では、Adaptive GMM と呼ぶ) である。もうひとつは、要素数があらかじめ設定されている背景モデル [1] (以下では、GMM と呼ぶ) である。これらの手法の学習率はそれぞれの文献に書かれているものを採用した。評価には、以下の式で計算される Precision (適合率) と Recall (再現率)、それらの調和平均の F 値を利用した。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$F = 2 / \left( \frac{1}{\text{Precision}} + \frac{1}{\text{Recall}} \right) \quad (7)$$

ここで、TP (True-Positive) は正しく検出した画素数、FP (False-Positive) は誤って検出した画素数、FN (False-Negative) は誤って検出できなかった画素数である。

表 1 は、各シーンでの物体検出精度である。Scene 1 と Scene 2 は屋外を撮影した映像で、天候の変化により背

(注2): 画像と Groud Truth は <http://limu.ait.kyushu-u.ac.jp/dataset/> より入手



図 6 物体検出精度の評価に利用したシーン

表 1 各シーンでの物体検出精度

		Scene 1	Scene 2	Scene 3
GMM [1]	Precision	0.26	0.23	0.07
	Recall	0.74	0.68	0.90
	F 値	0.39	0.34	0.13
Adaptive GMM [2]	Precision	0.21	0.16	0.04
	Recall	0.75	0.75	0.87
	F 値	0.33	0.26	0.07
Proposed Method	Precision	0.60	0.43	0.16
	Recall	0.69	0.69	0.82
	F 値	0.64	0.52	0.27

景変動が頻繁に生じるシーンである。GMM と Adaptive GMM では、再現率が高いが適合率は低い。これは背景変動に追従できずに背景領域を誤検出してしまうシーンが多く含まれていたためである。これに対して提案手法では、GMM と Adaptive GMM と同程度の再現率でありながらこれらよりも高い適合率が得られている。提案手法では、初めて観測される背景変動には従来手法同様に対応できないが、一度事例として登録されれば、次回以降は同様の背景変動が起これば即時に対応可能であり、この特長が高い適合率に寄与したと考えられる。そのため、F 値も比較手法の中で最も高い数値になっている。どのような事例がどのようなタイミングで利用されているかについては次節で詳しく検証する。

Scene 3 は屋内を撮影した映像で、天井にある複数の照明の点灯・消灯が繰り返し起こり、画像全体の輝度値が短期間で急激に変化するシーンである。統計的背景モデルはそのような急激な背景変動にはほとんど対応することができないため、非常に低い適合率となっていることがわかる。提案手法も決して高い適合率とは言えないが、一度生じた背景変動と同じ変動が起こるシーンでは適合率が増し、結果として GMM や Adaptive GMM より高い適合率と F 値が得られた。以上このことから、提案する事例ベースの背景モデルは、一度生じた背景変動を事例として記憶することができるため、従来の統計的確率モデルよりも背景変動への頑健性が増すことが確認できた。また、クラスタリングに基づく低コスト化のアプローチ [11] でも、GMM や Adaptive GMM と同程度の性能であることが報告されていることから、提案手法の物体検出に関する優位性は高いといえる。

### 4.3 メモリ使用量と処理時間の検証

提案手法と従来の統計的背景モデル [1], [2] を利用して

表 2 各手法におけるメモリ使用量と計算時間

	# of Dist.	Memory(KB)	Time(fps)
GMM [1]	230,400	54,152	10
Adaptive GMM [2]	115,200	35,632	10
Proposed	32,450	16,228	16

メモリ使用量と処理時間について比較を行ったところ、表 2 に示すような結果が得られた。メモリ使用量については、背景モデルの構築に必要な混合ガウス分布の画像全体での要素数に加え、実際に計算機上で実行した際に使用された物理メモリ量について調べた。

まず、画像全体で背景モデル構築に利用されたガウス分布の数について考察する。表中の GMM では、すべての画素で決められた数の分布を利用するため、解析画像期間を通して常に同じ数の分布が利用される。これに対して、画素単位で適応的に分布数を増減させる Adaptive GMM では、GMM と比較して利用するガウス分布の数を約半分に抑えることができている<sup>(注3)</sup>。提案手法では、Adaptive GMM と比較して、3 分の 1 以下まで分布数を削減できている。実機でのメモリ使用量についても、提案手法が最も少ないメモリ消費量で背景モデルを構築できていることが確認できた。メモリ消費量が Adaptive GMM の 3 分の 1 まで減っていないのは、提案手法では事例の検索に利用するハッシュテーブル用のメモリが必要になるからである。それを加味しても、従来手法に比べて少ないメモリ消費量で背景モデルを構築できることは提案手法の大きな貢献のひとつである。

次に計算時間について考察する。GMM と Adaptive GMM では、ともに 10fps で処理可能であるのに対して、提案手法では 16fps で処理可能である。従って、提案手法は処理時間を短縮するうえでも有効であることが確認できた。

### 4.4 事例削除の影響検証

Scene 1 の映像に対して、各フレーム時刻において登録されている事例の総数とその時刻で実際に利用された事例の数の関係を調べた。まず、TTL=10000 (解析映像では事例が削除されない) とした場合の結果を図 7 に示す。横軸はフレーム時刻、縦軸は背景モデルの数である。背景モデルの総数は徐々に増えて最終的には 21000 余りの事例が登録された。一方で、実際に各フレーム時刻で利用された事例の数は、約 10000 を上限に平均で 7500 程度であった。したがって、実際に利用されている事例の約 3 倍もの事例が登録されていることになる。登録事例の総数が増えるとメモリ使用量に直に影響するため、どの程度の事例を保持しておくことが適切であるかを検証する必要がある。

(注3): 表中の数値は解析期間を通しての利用された分布数の平均値である。

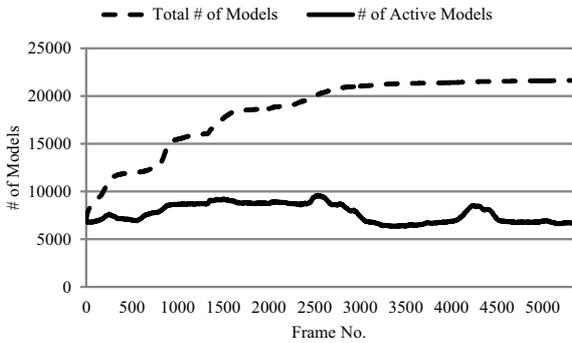


図 7 各フレーム時刻での背景モデル総数（事例の総数）と利用された背景モデル数

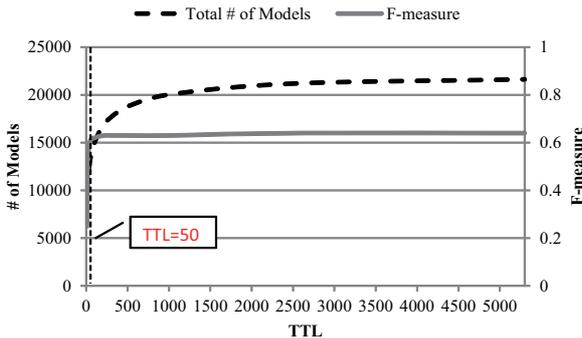


図 8 TTL に対する背景モデル総数と物体検出性能

そこで次に、TTL を徐々に短くして事例の削除による影響を調べた。図 8 では、横軸に TTL、対応する TTL のときの解析シーンでの最大の背景モデル数を左縦軸、そのときの物体検出精度として F 値を右縦軸として、TTL を変化させた際の各数値を表示している。TTL が小さいほど事例は削除されやすくなる。物体検出精度の F 値がほぼ横ばいになるのは、TTL を 50 に設定したときであった。このときの背景モデルの総数は 13000 余りであった。事例を全く削除しない場合と比較して、保持している事例を約 60% に削減しても、精度にはさほど影響を及ぼしていないことが確認できた。このことから、TTL を設定することは背景モデルの総数（事例の総数）を減らすことに貢献でき、それによる物体検出性能低下への影響は小さいといえる。

## 5. おわりに

事例ベース背景モデリングのフレームワークと、統計的背景モデルへの適用方法について提案した。背景モデルに要求される高性能化と低コスト化を両立することが可能な新たなブレークスルーとして事例ベース背景モデリングは有効な手法であり、その汎用性も高い。今後の課題としては、事例ベースアプローチを他の背景モデリング手法へ実際に導入してその性能を検証することや、超長時間映像に対する性能評価などが挙げられる。

- [1] Chris Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 2, pp. 246–252, 1999.
- [2] Atsushi Shimada, Daisaku Arita, and Rin ichiro Taniguchi. Dynamic Control of Adaptive Mixture-of-Gaussians Background Model. *CD-ROM Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance*, 2006.
- [3] Ahmed Elgammal, Ramani Duraiswami, David Harwood, and Larry Davis. Background and Foreground Modeling using Non-parametric Kernel Density Estimation for Visual Surveillance. *Proceedings of the IEEE*, Vol. 90, pp. 1151–1163, 2002.
- [4] Heikkila Marko and Pietikainen Matti. A Texture-Based Method for Modeling the Background and Detecting Moving Objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 4, pp. 657–662, 2006.
- [5] Yutaka Satoh, Shun'ichi Kaneko, Yoshinori Niwa, and Kazuhiko Yamamoto. Robust object detection using a Radial Reach Filter (RRF). *Systems and Computers in Japan*, Vol. 35, No. 10, pp. 63–73, 2004.
- [6] YOKOI Kentaro. Probabilistic bprcc: Robust change detection against illumination changes and background movements. *IEICE Transactions on Information and Systems*, Vol. 93, No. 7, pp. 1700–1707, 2010.
- [7] LingFeng Wang, HuaiYu Wu, and ChunHong Pan. Adaptive LBP for Background Subtraction. In *Asian Conference on Computer Vision 2010*, 2010.
- [8] Satoshi Yoshinaga, Atsushi Shimada, Hajime Nagahara, and Rin ichiro Taniguchi. Object Detection Using Local Difference Patterns. In *Asian Conference on Computer Vision 2010*, 2010.
- [9] Kentaro Toyama, John Krumm, Barry Brumitt, and Brain Meyers. Wallflower: Principle and Practice of Background Maintenance. *International Conference on Computer Vision*, pp. 255–261, 1999.
- [10] Tatsuya Tanaka, Atsushi Shimada, Rin-ichiro Taniguchi, Takayoshi Yamashita, and Daisaku Arita. Towards robust object detection: integrated background modeling based on spatio-temporal features. In *Asian Conference on Computer Vision 2009*, 2009.
- [11] Atsushi Shimada, Tatsuya Tanaka, Daisaku Arita, and Rin ichiro Taniguchi. Spatial-temporal integration of adaptive gaussian mixture background models. *Proceedings of the 14th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, 2008.
- [12] Brian Valentine, Senyo Apewokin, Linda Wills, and Scott Wills. An efficient, chromatic clustering-based background model for embedded vision platforms. *Comput. Vis. Image Underst.*, Vol. 114, pp. 1152–1163, 2010.
- [13] Mayur Datar, Nicole Immorlica, Piotr Indyk, and Vahab S. Mirrokni. Locality-sensitive hashing scheme based on p-stable distributions. In *SCG '04: Proceedings of the twentieth annual symposium on Computational geometry*, pp. 253–262, 2004.
- [14] Andoni Alexandr and Indyk Piotr. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions. *Commun. ACM*, Vol. 51, pp. 117–122, 2008.