

尤度ベースの背景モデルを用いた物体検出手法 An Object Detection Method Based on a Likelihood Background Model

池田 浩雄†
Hiroo Ikeda

石寺 永記†
Eiki Ishidera

1. はじめに

近年、監視映像を用いたセキュリティへの期待が高まっている。特に、映像の中から人/物の存在や動きを理解する技術が注目されており、その要素技術として物体検出がある。

物体検出は、日照等の照明変動や木揺れ等の環境変動の影響によらず、静止/移動する物体を安定して検出する必要がある。従来手法として背景差分法があるが、照明変動や環境変動に対する安定性に問題がある。その問題に対処する為に、統計的なモデルを用いる手法があり、著者らは、マハラノビス距離を用いた手法[1]を提案した。しかし、[1]では背景の輝度分布を正規分布で仮定していた為、正規分布でない輝度差の大きな環境変動等に対して、未検出の問題がある。他にも、背景の分布に混合正規分布を用いた手法[2]が提案されている。環境変動や照明変動に有効であるが、分布推定の為の分布数の決定が難しく、その数によっては背景が適切に表現されずに誤検出/未検出する問題がある。

本稿では、表現能力の高いヒストグラム法で背景の輝度分布を推定し、これを背景モデルとして用いる物体検出手法を提案する。さらに、検出精度を向上させる為、時空間方向における事象のマルコフ性を用いて尤度ベースのモデルを拡張することを提案する。

2. 尤度ベースのモデルに基づく物体検出

2.1. 尤度ベースの物体検出と輝度分布の推定

背景の輝度分布は、正規分布で仮定できることが多いが、木揺れや水面の揺れ等の環境変動においては複雑な分布を示す場合がある。その為、マハラノビス距離を用いた従来手法[1]では、正規分布でない輝度差の大きな環境変動の場合、物体が検出されにくくなる問題が発生する。これは、背景の変化を単純な正規分布で仮定していることが原因で、背景の分布には高い表現能力が要求される。そのための手法として、混合正規分布を用いた手法[2]がある。しかし、分布推定の為の分布数の決定方法が難しく、その数によっては背景の分布が適切に表現されずに誤検出/未検出することがある。

そこで、本稿では、表現能力の高いヒストグラム法で背景の輝度分布を推定し、尤度ベースのモデルによって、画素を背景と前景に識別する。尤度ベースのモデルには、式(1)で示すベイズの定理に基づく事後確率を用いる。

$$P(\omega_i | I) = \frac{p(I | \omega_i) P(\omega_i)}{p(I)} \quad (1)$$

ただし、 ω_i は、事象 i ($i=0$:背景、 $i=1$:前景)を示し、 I は輝度を示す。また、事前確率 $P(\omega_i)$ について背景と前景の事象で等しいことを仮定する。

式(1)の $P(I | \omega_i)$ は、背景と前景における輝度の生起確率で、背景については、過去に背景と判定された輝度の時系列を用い

て逐次、ヒストグラム法で推定する。前景については、検出対象を特定しないので一様と仮定する。

2.2. マルコフ性を用いたモデルの拡張

式(1)の尤度ベースのモデルは、背景の輝度分布のみに注目しモデル化している。そこで、それとは別の時空間方向における事象のマルコフ性に着目しモデルを拡張する。

ここで示す時空間方向における事象のマルコフ性とは、図1に示すように、注目画素とその周辺画素で同じ事象(背景/前景)が起こりやすいという性質と、画素の時系列で背景の出現頻度が高いという性質である。

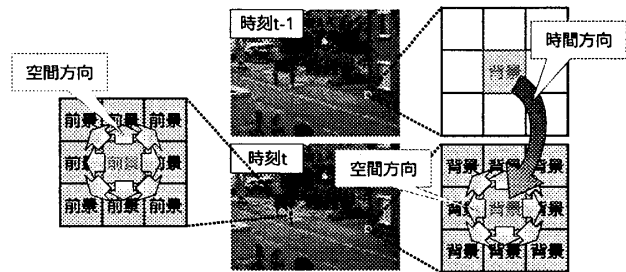


図1. 時空間方向における事象のマルコフ性

これらの性質を事象間の関係を表す確率(条件付確率)として表現し導入することで、モデルを拡張する。

まず、観測できる注目画素の輝度、その4近傍(周辺画素)の輝度、注目画素の過去の事象を条件として、注目画素の背景/前景の事後確率を求める尤度ベースのモデルを設定する。4近傍の事象が変数として表れていないので積分を用いて変数を加え、事象間の関係を表す条件付確率が出現するようベイズの定理と変数の独立を用いて、式(2)のように変形する。

$$P(\omega_{i_0}^{(t)} | I_0, \dots, I_t, \omega_{i_0}^{(t-1)}) \\ = \frac{p(I_0 | \omega_{i_0}^{(0)})}{p(I_0, \dots, I_t)} \left[\sum_{k_1=0k_2=0k_3=0k_4=0}^1 \sum_{j=1}^4 \prod_{j=1}^4 p(I_j | \omega_{k_j, j}^{(t)}) P(\omega_{k_j, j}^{(t)} | \omega_{i_0}^{(t)}) \right] P(\omega_{i_0}^{(t)} | \omega_{i_0}^{(t-1)}) \quad (2)$$

ただし、 $\omega_{a,b}^{(t)}$ は、時刻 t における画素位置 b ($b=0$:注目画素、 $b \neq 0$:4近傍)の事象 a ($a=0$:背景、 $a=1$:前景)を示す。 I_b は、時刻 t における画素位置 b ($b=0$:注目画素、 $b \neq 0$:4近傍)の輝度を示す。また、仮定した変数の独立とは、注目画素の過去の事象は注目画素の現在の事象以外と独立、事象と輝度において同じ画素位置以外は独立、4近傍の事象は互いに独立、4近傍の輝度は互いに独立というものである。

式(2)において、 $P(\omega_{k_j, j}^{(t)} | \omega_{i_0}^{(t)})$ は空間方向における事象間の関係を表す条件付確率を示し、注目画素とその周辺画素で同じ事象が起こりやすいという性質を利用して、同じ事象 ($k_j = i$)の場合、異なる事象より大きく設定する。 $P(\omega_{i_0}^{(t)} | \omega_{i_0}^{(t-1)})$ は時間方向における事象間の関係を表す条件付確率を示し、画素の時系列で背景の出現頻度が高いという性質を利用して、背景

† 日本電気(株) 共通基盤ソフトウェア研究所

($i=0$) の場合、0.5 より大きく設定する。 $p(I_0 | \omega_{i,0}^{(t)})$ や $p(I_j | \omega_{k,j}^{(t)})$ は、注目画素とその4近傍の事象における輝度の生起確率を示し、2.1項で述べたように背景はヒストグラム法で求め、前景は一様分布とする。

式(2)で拡張した尤度ベースのモデルを用いて、画素を背景と前景に識別し、物体検出を行う。

3. 実験

提案手法の性能を評価する為、静的な背景と動的な背景の上を人が通過する評価DBを用い、実験を行った。静的な背景とは、道路や建物などの固定された背景であり、動的な背景とは木や草むらなどの環境変動を含む背景である。実験に用いた評価画像は、69シーンの映像(すれ違いなども含む)から抽出された動画像データで、静的な背景44840枚、動的な背景14043枚である。また、それらの画像サイズは320×240のQVGAである。

評価指標としては、図2(a)に示す検出対象の物体数に対する正解数の割合(リコールレート)と、検出された物体数に対する正解数の割合(プレジジョンレート)を用いる。よって、両レートが高くなることが望ましい。正解であるという判定は、検出された物体矩形が人手で与えられた正しい位置(検出対象の物体矩形)と近いかを判断することで行う。つまり、図2(b)に示すように、検出対象の物体矩形と検出された物体矩形において、矩形幅の差分絶対値、矩形高さの差分絶対値、矩形中心の距離、全てが10ピクセル以下を正解とする。検出された物体矩形とは、画素単位で検出された物体(前景)をラベリングし、距離の近いラベル同士を結合して矩形で括ったものである。

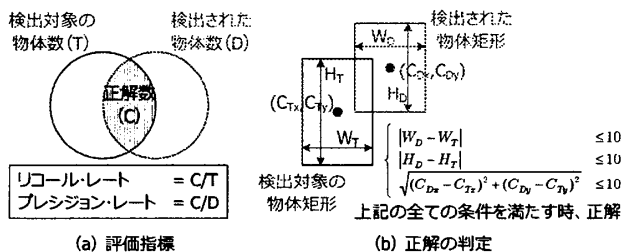


図2. 評価指標と正解の判定

3.1. ヒストグラム法を用いた手法の性能評価

ヒストグラム法で推定された表現能力の高い背景の分布を用いた尤度ベースのモデル(式(1))による手法と、背景の分布を正規分布と仮定した従来手法[1]を比較評価した。図3に、性能を比較したROC曲線を示す。

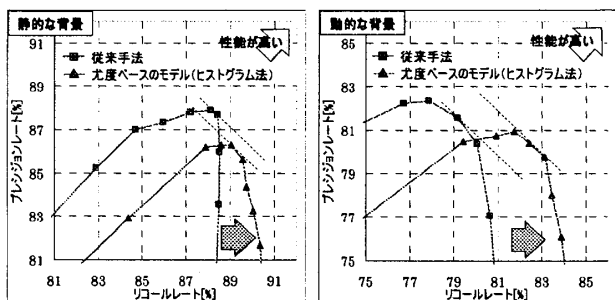


図3. 背景の分布形状別のROC曲線

図の縦軸はプレジジョンレート、横軸はリコールレートを表し、右上に向かう程、性能が高いことを示している。

図より、動的/静的な背景の両方において、ヒストグラム法を適用することで、リコールが改善され未検出が低減できることが分かる。しかし、プレジジョンが少し低下する。従来手法は、背景の分散が大きい場合に検出感度を低下させるという誤検出の低減に強く重みをおいた手法で、その効果が強すぎたと思われることができる。

3.2. 拡張モデルの性能評価

拡張モデル(式(2))を用いた手法について、事象間の関係を表す条件付確率を変化させ予備実験を行ったところ、事象のマルコフ性を示す範囲で性能が改善されることを確認した。

そこで、従来手法[1]と拡張前の尤度ベースのモデル(式(1))を用いた手法を対象に性能を比較評価した。図4に、性能比較の結果であるROC曲線を示す。式(2)で用いた事象間の関係を表す条件付確率は、 $P(\omega_{0,0}^{(t)} | \omega_{0,0}^{(t-1)}) = 0.8$, $P(\omega_{1,0}^{(t)} | \omega_{1,0}^{(t-1)}) = 0.5$, $P(\omega_{k,j}^{(t)} | \omega_{i,0}^{(t)}) = 1.0$ ($k_j = i$ 時)である。

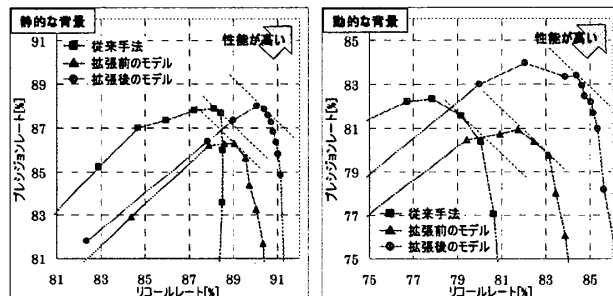


図4. 手法別のROC曲線

図の縦軸はプレジジョンレート、横軸はリコールレートを表し、右上に向かう程、性能が高いことを示している。

図より、拡張モデルを用いた手法の性能が優れていることが分かる。これにより、事象のマルコフ性が性能改善に有効であることが分かる。特に、厳しい環境である動的な背景下で性能は大きく改善される。

4. おわりに

本稿では、表現能力の高いヒストグラム法で背景の輝度分布を推定し、これを背景モデルとして用いる物体検出手法を提案した。さらに、検出精度を向上させる為、時空間方向における事象のマルコフ性を用いて尤度ベースのモデルを拡張することを提案した。実験の結果、背景の分布に表現能力の高い分布(ヒストグラム法)を用いることで、未検出を低減できることを示した。さらに、時空間における事象のマルコフ性が性能改善に有効であることを示した。今後の課題としては、急激な照明変動への対応や、事象間の関係を表す条件付確率を各画素で適応的に決定して性能を向上させることなどが挙げられる。

参考文献

[1] 池田,石寺, "輝度情報の時間的性質を用いた背景のモデル化とそれに基づく物体検出", FIT2006 一般講演論文集, 第3分冊, I_072, pp.175-176, 2006
 [2] Stauffer, Grimson, "Adaptive background mixture models for real-time tracking", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, pp.246-252, 1999.