# 映像からの移動対象検出のための適応的背景映像モデリング

# Adaptive Background Scene Modeling for Detection of Moving Objects

１ＸＸＸＸＸＸ（登録番号）

研究者代表　　奈良先端科学技術大学院大学 情報科学研究科 助教　　波部　斉

［研究の目的］

本研究では，映像中に存在する検出対象（人物・障害物・車両など）を精度よく検出する手法の確立を目指す．自律的に動作するロボットや，人間に適切な情報を提供する情報システムを構築する際には，機械が周囲の状況を把握することが不可欠である．そこで，映像からの対象検出は基本となる機能であり，現在でもセキュリティシステムや移動ロボットに広く用いられているが，動的環境（特に屋外）においては性能に限界がある．

それに対して本研究では，動的環境における高精度な対象検出の実現を目指し，背景映像の特徴に応じた適応的なモデリング手法を提案する．

［研究の内容、成果］

本研究では,「背景映像のモデリング」による移動対象検出を行う．一般に移動対象検出は，画像を検出対象とそれ以外（背景）との二つの領域に分別する操作と考えられ，その方法には対象モデルに基づくものと背景モデルに基づくものがある．対象モデルに基づく対象検出は対象が既知でなければならないという制約があるが，背景モデルに基づく対象検出は背景についてのみ学習するので，対象は未知であって構わない．そのため，実世界における対象検出の中で，対象に関して未知である場合にはこの方法が有効である．本研究も後者に属するものである．

背景モデルに基づく対象検出では，背景に関する情報を学習する必要がある．通常，実世界では背景の状態は不変でないため，単純な背景差分に基づく対象検出は有効ではない.そこで，実世界における対象検出を実現するには，背景の変動を適切にモデル化する必要がある．背景変動のモデル化の方法としては，これまで様々なものが提案されてきた．それらは以下の３つに大別される．

（１） 背景変動に対して不変な特徴を利用する方法.

（２） 与えられた背景画像系列より計算される統計量を用いる方法.

（３） 背景変動の時間的, 空間的な相関関係に注目する方法.

（１）は時間変化に対して不変な特徴を単独の画像から抽出するものであるが，現実的には完全に不変な特徴を抽出することは不可能であるので，その時間変化を（２）のような統計モデルで記述するものが用いられている．さらに，単純な統計量ではなく，時間的あるいは空間的な変化・相関関係を利用したものが（３）であるといえ，順をおってモデルの複雑さは増している．

このような中で，本研究では線形システムを用いて背景画像の時間的変化を記述し，背景変動のモデル化を行う．ここでいう線形システムとは，内部状態を示す状態ベクトルの変化が線形な方程式で記述できるシステムのことを指し，時刻 t での状態ベクトルを, 時刻 t-1 から t-R までの状態ベクトルから推定するというもので

ある．背景変動の要因として，実環境では（１）照明条件の変化，（２）背景に存在する物体の移動・出現，（３）木の葉の揺らぎの，周期的・規則的とみなせる変化，など様々なものが挙げられるが，本研究ではこのうち（３）に着目して背景変動をモデル化する．（３）の変動は，過去から連続して発生するものであり，それを表現するものとして線形システムは妥当なものであると考えられる．さらに線形システムでは，演算がすべて線形な行列演算で記述できるため単純かつ高速であり，実時間処理が必要とされるシステムへの適用にも有利であるといえる．

線形システムを用いて背景映像の変動を記述する場合には，観測画像データから状態ベクトルへの変換が非常に重要になる．本研究では，（１）観測画像に対して平滑化による解像度変換を適用し，（２）線形システムのモデル化誤差と状態ベクトルの変動を解析することで，対象検出に最適な状態ベクトルを選択する．細かい解像度の観測画像をそのまま状態ベクトルとしたときには，それぞれの要素の変動が複雑になって線形システムで表現したときの予測誤差が大きくなるが，それを平滑化して得た状態ベクトルでは大域的な変動のみが表現されているので，線形システムでの予測誤差の低減が期待できる．しかし，極端に平滑化した場合には予測誤差が小さくなっても検出対象と背景の識別が不可能になる可能性があるので，（１）のように状態ベクトルの変動を解析して最適な解像度を求める．以下に，その具体的な手法を述べる．

本研究で採用する線形システムは，具体的に

$\underline{x}\vee\underline{t} = \underline{F}\,[\,\underline{x}\vee\underline{t}\text{-}1\wedge T, \underline{x}\vee\underline{t}\text{-}1\wedge T, \cdots, \underline{x}\vee\underline{t}\text{-}R\wedge T]\wedge T$

で定義する．ここで，$\underline{x}$ は各時刻での状態を表す状態ベクトルであり，$\underline{F}$ は過去のデータと現在のデータの関係を記述する状態遷移行列である．先述のとおり，観測された画像データから状態ベクトルへの変換が非常に重要になり，本研究では，以下のように変換を行う．

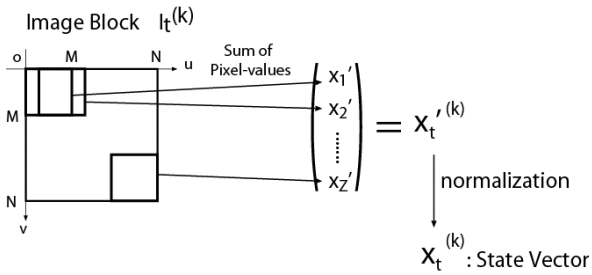


図１　状態ベクトルの生成

状態ベクトルとして最も単純なものとしては，画像ブロックをそのままベクトル表記したものが考えられるが，この場合状態ベクトルの各要素はもとの画像の画素を表しており，それぞれの要素の変動が非常に複雑なものとなる．そのため状態ベクトルの変動を線形な変換で記述するのはほぼ不可能である．そこで，本研究では画像の平滑化を行った後に状態ベクトルに変換する．得られた状態ベクトルは背景の大域的な変動のみを表現しているため，線形システムに適用したときに誤差の小さい予測が行えると考えられる．以下，画像ブロック $\underline{I}\vee\underline{t}\wedge(\underline{k})(\underline{t}=1,\cdots,\underline{T})$から状態ベクトル $\underline{x}\vee\underline{t}\wedge(\underline{k})$を生成するアルゴリズムを示す（図１）．

（１）　画像ブロック $\underline{I}\vee\underline{t}\wedge(\underline{k})$ ($\underline{N}\times\underline{N}$ 画素)内に，$\underline{M}\times\underline{M}$画素の小ブロックを考える．小ブロックの始点を画像ブロックの左上端（座標を(0,0)とする）とし，小ブロック内の画素値の和を $\underline{x}'\vee 1$ とする．なお，M の決定方法については後述する．

（２）　小ブロックの始点を(0,0),(0,1),… ,(1,0),(1,1),… ,$(\underline{N\text{-}M},\underline{N\text{-}M})$と移動させていき，それぞれの始点の場合における小ブロック内の画素値の和を $\underline{x}'\vee 1,x'\vee 2,\cdots,x'\vee Z$ とする(ここで，$\underline{Z} = (\underline{N}\text{-}\underline{M}+1)\wedge 2$).

（３）　ベクトル $\underline{x}'\vee\underline{t}\wedge(\underline{k})=[\underline{x}'\vee 1\wedge T, \underline{x}\vee\underline{2}\wedge T, \cdots, \underline{x}\vee\underline{Z}\wedge T]\wedge T$ を，時刻 $\underline{t}$ における状態ベクトルとする．状態ベクトルの各要素が小ブロック内の画素値の和であるので，以上の操作は

画像の解像度を小ブロックサイズのレベルまで粗くすることと同等である．

（4） ベクトル x’ ∨t∧(k)を以下のように正規化する．

$$\boldsymbol{x}_t^{(k)} = \frac{\boldsymbol{x}_t'^{(k)}}{\|\boldsymbol{x}_t'^{(k)}\|}$$

得られた x ∨t∧(k)を状態ベクトルとする．このように正規化することにより，一様な画素値の変化の影響を受けなくなる．照明条件の変化が画素値の一様な変化を引き起こすと近似できるとすると，正規化により照明条件の変化の影響を削減することができる．

上記の方法により得られる状態ベクトルは，小ブロックのサイズ M，すなわち画像の解像度に依存する．M のとりうる値の範囲は 1≦M≦N であり，各ブロックにおいて対象検出を行う際に最も適切な解像度が存在すると考えられる．最適な M を選択するために，本研究では以下の２つの指標を用いる．

■状態ベクトルの予測誤差

まず，直観的に考えられる指標として，与えられたデータからある M を用いて導出した状態遷移行列 F から予測した状態ベクトルと観測データの間の誤差が考えられる．この誤差は，現在の解像度 M でのデータが線形システムで表現できるか否かを表している．以下ではこの誤差を E と定義する．

■状態ベクトルの変動

しかし，対象検出に最適な解像度を決定するための評価値として E だけでは不十分である．なぜなら，E が小さければ状態ベクトルの変動の予測精度は高いといえるが，極端に平滑化されているために状態ベクトルの変動がほとんど見られずほぼ不変であり，そのために予測精度が高いという可能性もありうる．

仮に，ある画像ブロックに対し異なる２つの解像度において状態ベクトルを得たときに E が同程度であった場合，状態ベクトルの変動が小さいほうは単に極端に平滑化されているために E が小さいと考えられる．他方，変動が大きい状態ベクトルは細かい背景変動まで表現しているといえるので，E が同程度の場合は変動の大きい状態ベクトルを用いるのが適切と考えられる．また，極端に平滑化したために状態ベクトルの変動が小さい場合には，対象が現れた場合に背景との識別が困難になり，対象検出の精度が低下する可能性がある．そのような観点からも，状態ベクトルの変動に注目することは重要といえる．状態ベクトルの変動を評価するものとして，下式で表される変動値 V を用いる

$$V(\boldsymbol{x}^{(k,M)}) = \frac{V_0(\boldsymbol{x}^{(k,M_{min})})}{V_0(\boldsymbol{x}^{(k,M)})}$$

ここで M ∨{min}は M の最小値であり，V∨0 は

$$V_0(\boldsymbol{x}^{(k,M)}) = \frac{1}{T-1}\sum_{t=2}^{T} Sim(\boldsymbol{x}_{t-1}^{(k,M)}, \boldsymbol{x}_t^{(k,M)})$$

で定義される．ここで，Sim(・)は二つのベクトルの類似度を示し，この値が小さいほど時刻 t-1 から時刻 t での状態ベクトルの移動が大きい．そのため V∨0 が小さいほど状態ベクトルの変動が大きいものとみなせる．したがって V は，同一ブロックにおける，解像度 M での状態ベクトルの変動と解像度 M∨min での状態ベクトルの変動との比を示している．

■二つの指標の統合

以上のように求まった，状態ベクトルの予測誤差 E，および状態ベクトルの変動値 V を用いて，最適な解像度 M を求める．言うまでもなく，E が大きいと線形システムは精度よく機能しない．また，E が小さくても変動値 V が小さければ単に画像が極端に平滑化されているだけの可能性があるので，E が小さいものの中で V は大きいほうが好ましい．このような観点から最適な解

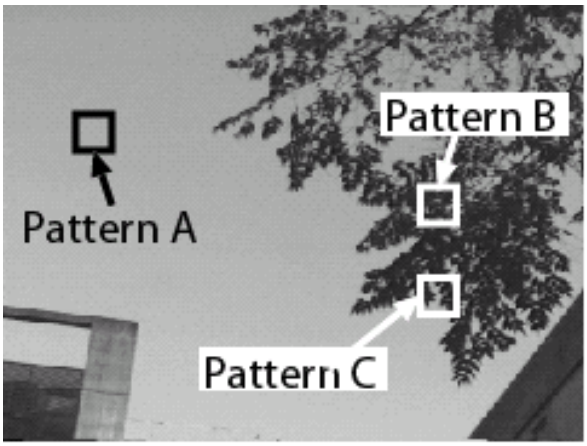

図２　実験で使用した映像

像度 M を決定する.
以上の手法の有効性を確認するための実験結果を以下に示す.
実験では，図２に示す屋外で撮影した画像を使用した．その中で特徴的なパターンＡ，Ｂ，Ｃを選び出し，それぞれに対して解像度を変更させて行った対象検出の性能評価結果を図３から図４に示す．それぞれの図で，横軸は対象を誤って検出した割合を示しており，縦軸は対象を正しく検出した割合を示している．すなわち，図上で左上に近いほど，性能が良いことがわかる．この図をみると，解像度 M を大きくすると性能が向上する傾向がみられるが，極端に大きくすると逆に悪化することがわかる．紙面の都合上省略するが，先日の評価値EおよびVが最適な解像度を選択する指標として有効であることも確認でき，本手法の妥当性が確認できた.
［今後の研究の方向、課題］
今後は，多数のデータを用いて詳細な実験を行い, アルゴリズム改良を行う予定である. また，背景の動きだけでなく，屋外シーンのさまざまな構造を認識し，対象検出に応用することも考えられる（科学研究費補助金の研究課題として実施予定である）.
［成果の発表、論文等］
情報処理学会コンピュータビジョンとイメージメディア研究会や電子情報通信学会パター認識とメディア理解研究会などの国内学会や，

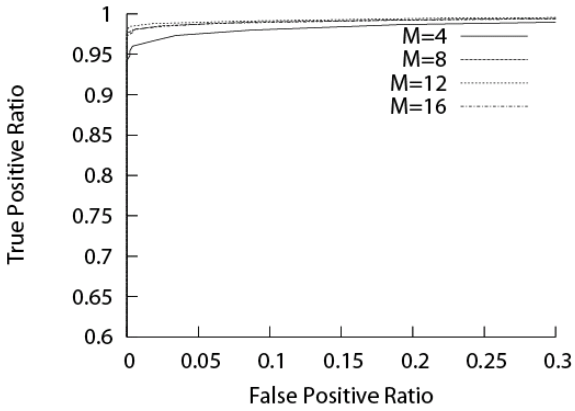

図３　パターンＡでの評価結果

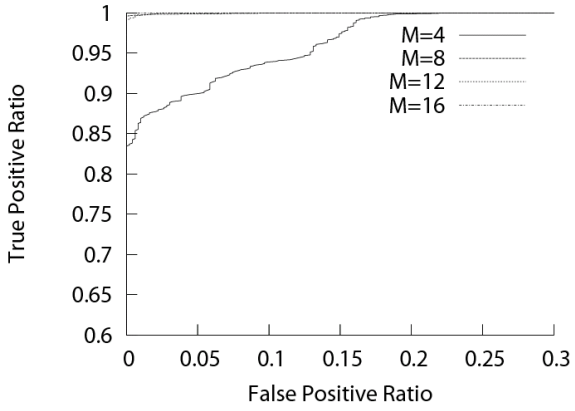

図４　パターンＢでの評価結果

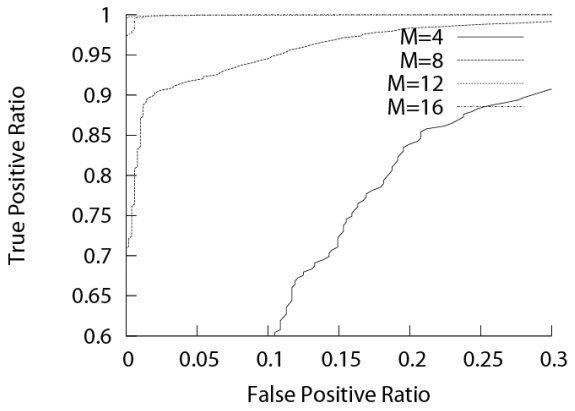

図５　パターンＣでの評価結果

CVPR, ICPR などの国際学会に投稿する予定である.
［謝辞］
この研究を行う機会を与えていただいた立石科学技術振興財団に心から感謝いたします. また，日頃から議論を行っていただいている京都大学大学院情報学研究科松山隆司教授はじめ松山研究室の各位に感謝いたします.

図表などの拡大図（それぞれ eps ファイルとして CD-R に格納しています）

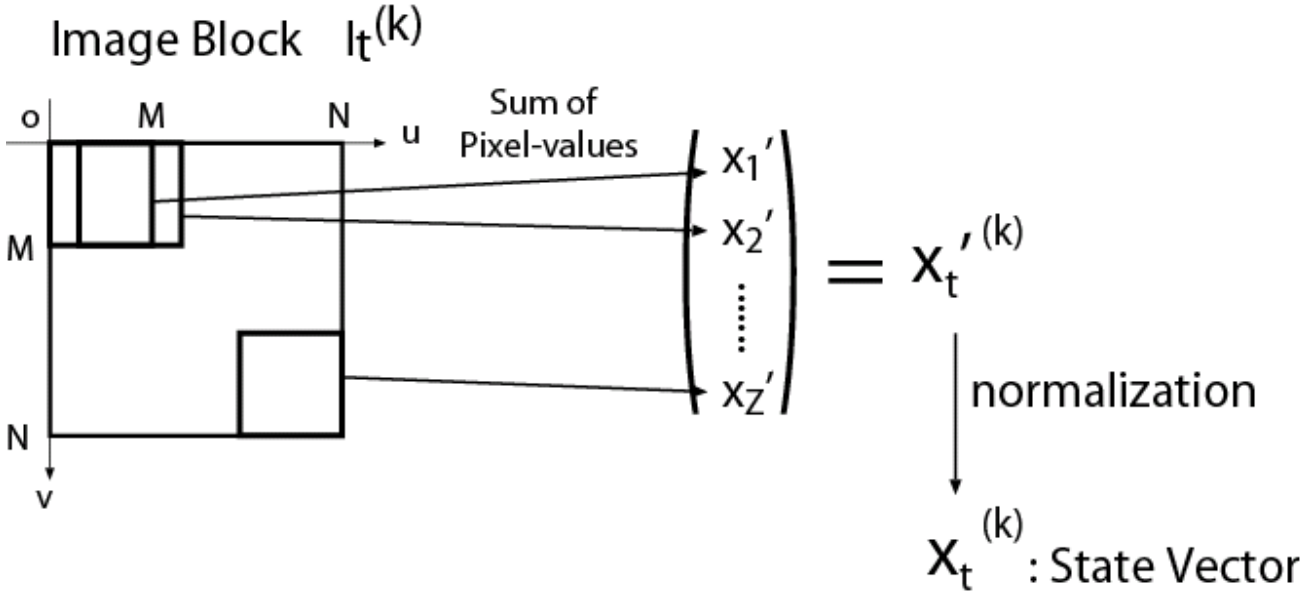


図１

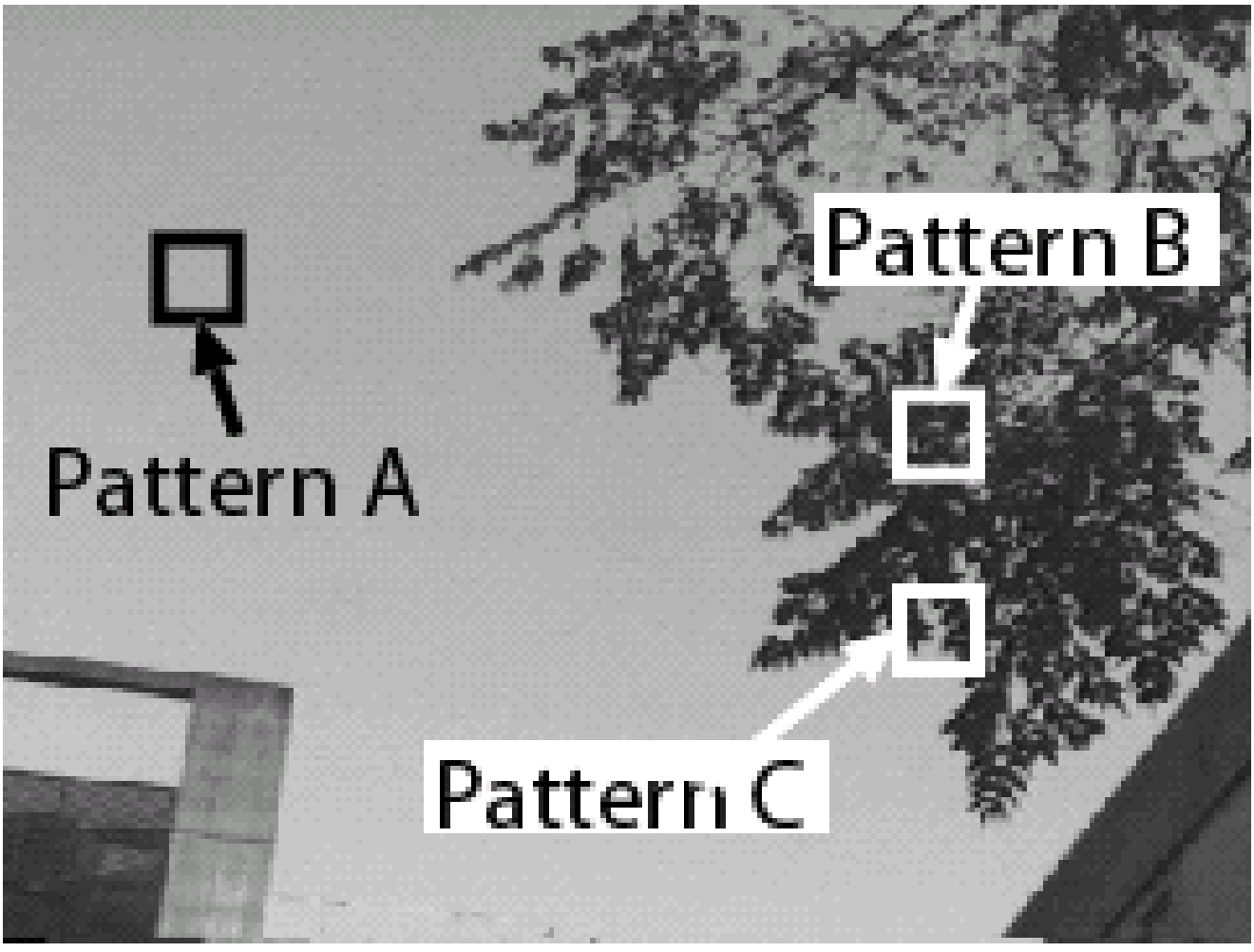


図２

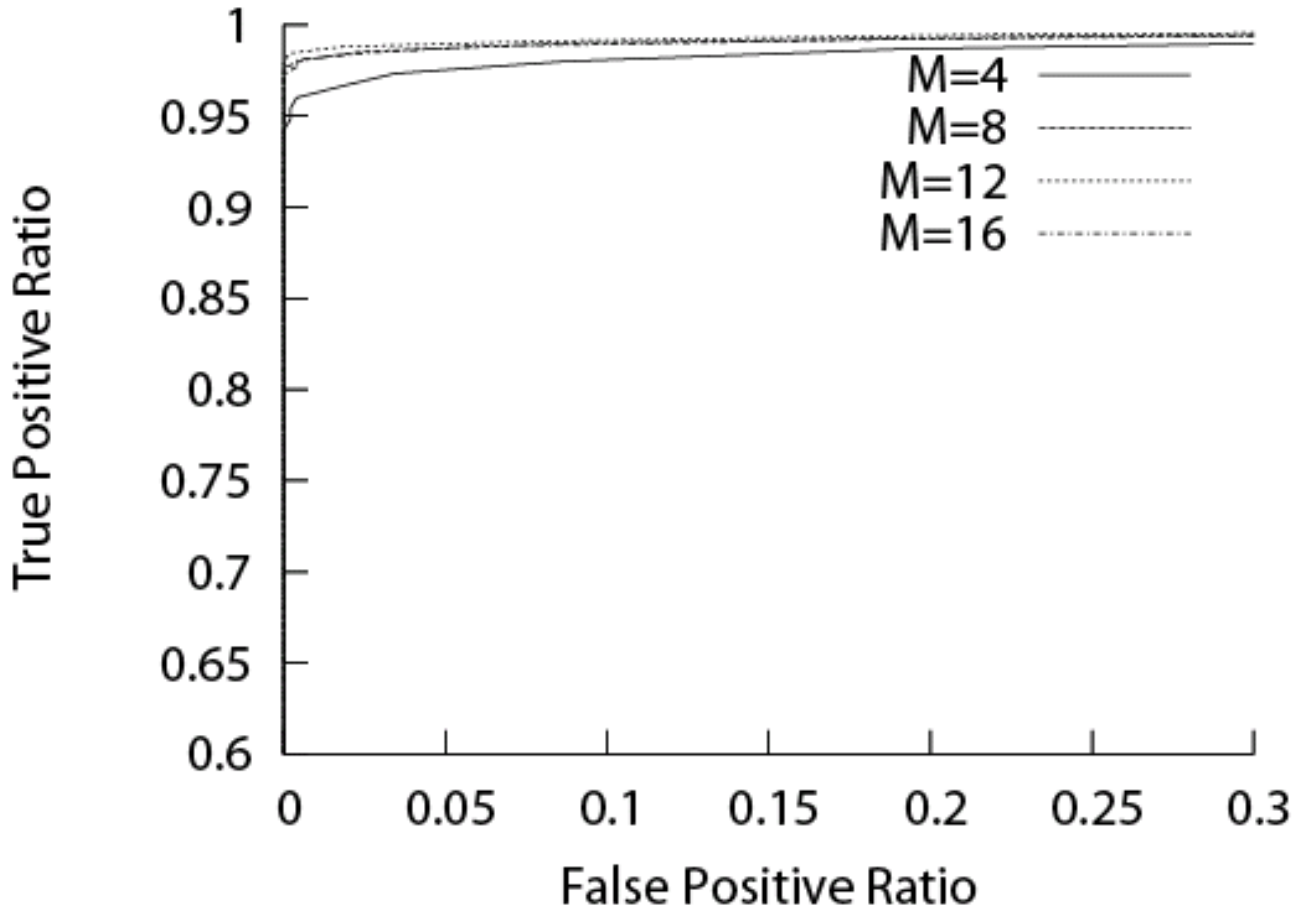
True Positive Ratio
False Positive Ratio
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0
0.05
0.1
0.15
0.2
0.25
0.3
M=4
M=8
M=12
M=16

図 3

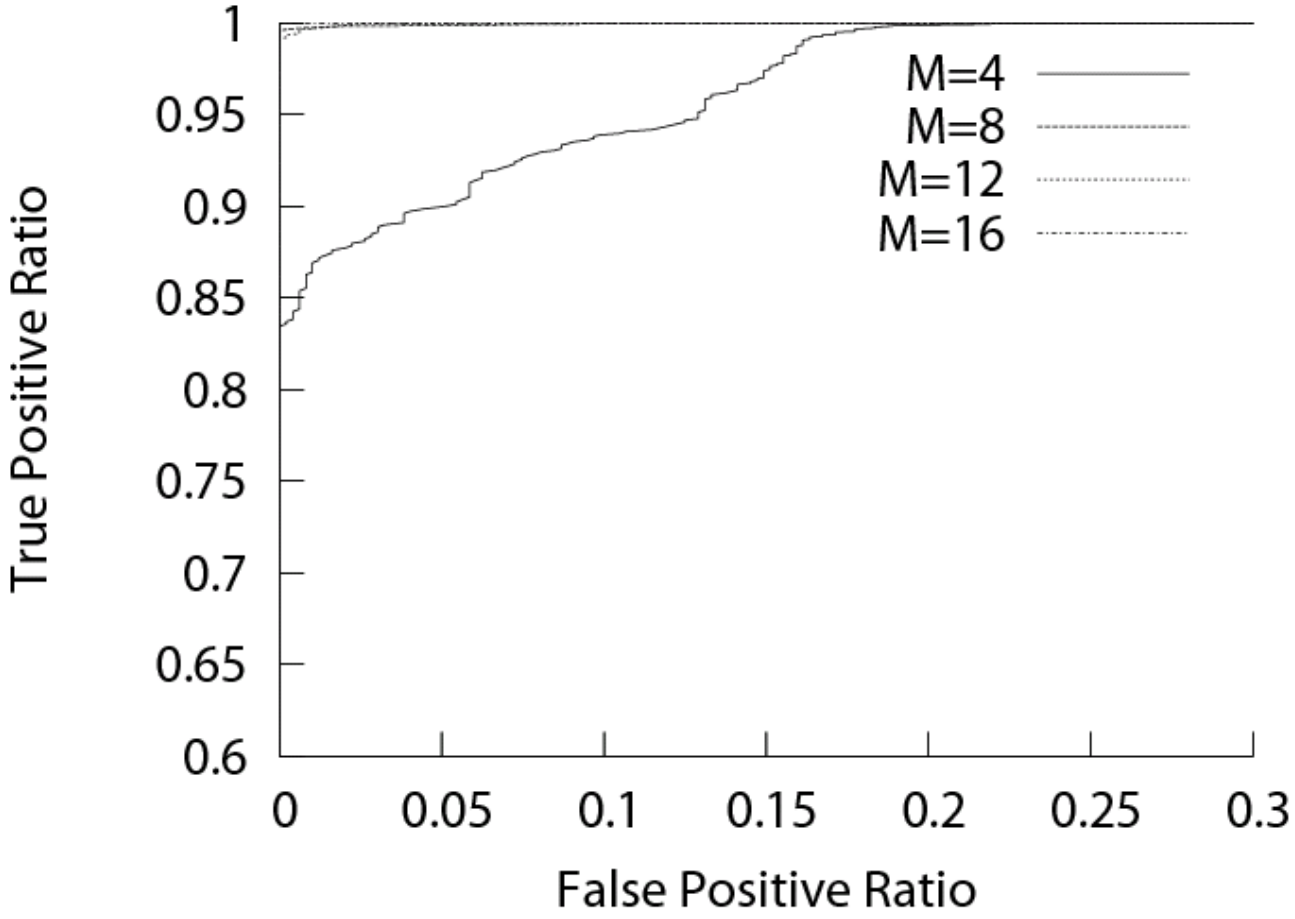
True Positive Ratio
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0
0.05
0.1
0.15
0.2
0.25
0.3
False Positive Ratio
M=4
M=8
M=12
M=16

図４

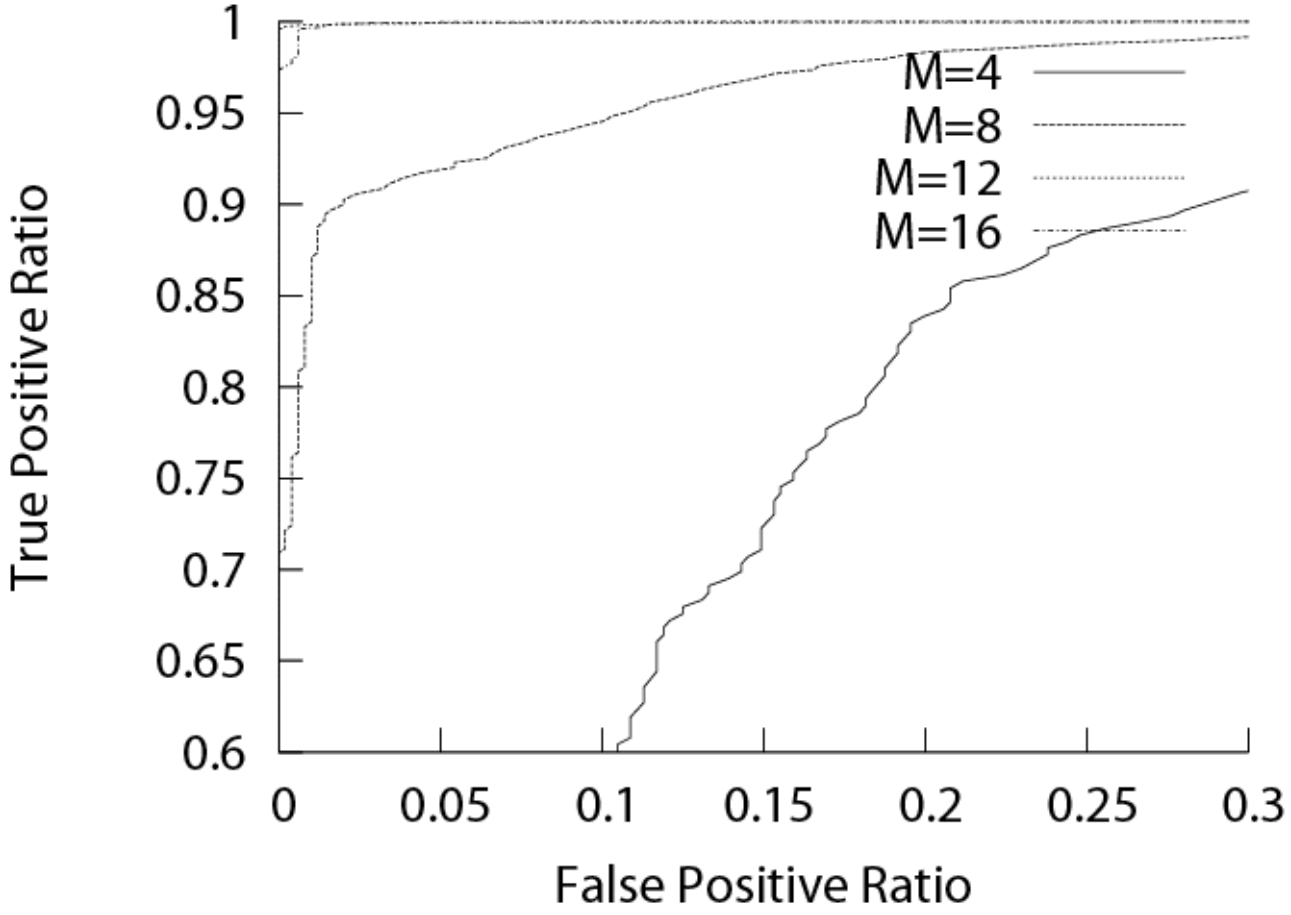
True Positive Ratio
1
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0
0.05
0.1
0.15
0.2
0.25
0.3
False Positive Ratio
M=4
M=8
M=12
M=16

図5