

## 特徴選択と Soft-Margin SVM の Boosting を用いた歩行者検出

西田健次<sup>†</sup> 栗田多喜夫<sup>†</sup>

<sup>†</sup> 産業技術総合研究所 脳神経情報研究部門  
305-8568 茨城県つくば市梅園 1-1-1 中央第二  
E-mail: †{kenji.nishida,takio-kurita}@aist.go.jp

あらまし ブースティングによる歩行者識別器を構成する際に、弱識別器に対して最適な特徴、および、局所領域を選択することによって識別性能向上を図る手法を提案する。本稿では、弱識別器にソフトマージン SVM を採用し、ブースティングの各段で最適な特徴、および、局所領域を選択する。実験には MIT CBCL の歩行者画像データベースの画像を用い、特徴としては入力画像のヒストグラム均一化特徴とエッジ特徴の二種類を採用し、一つの入力画像から自動的に 100 個の局所領域を生成した。実験結果は、局所領域の選択を行う事で単一特徴でも高い識別率を示し、二つの特徴から選択する事による識別率の改善はわずかであった。しかし、二つの特徴を用いる事によって、より効果的な局所領域が選択されており、更なる性能向上の可能性が示唆されている。

キーワード AdaBoost, サポートベクターマシン (SVM), 特徴選択, 歩行者検出

## Pedestrian Detection by Boosting Soft-Margin SVM with Feature Selection

Kenji NISHIDA<sup>†</sup> and Takio KURITA<sup>†</sup>

<sup>†</sup> Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology (AIST)  
Central 2, 1-1-1 Umezono, Tsukuba, IBARAKI 305-8568 JAPAN  
E-mail: †{kenji.nishida,takio-kurita}@aist.go.jp

**Abstract** We present an example-based algorithm for detecting objects in images by integrating component-based classifiers, which automatically select the best feature for each classifier and are combined according to *AdaBoost* algorithm. The system employs soft-margin SVM for base learner, which is trained for all features and the optimal feature is selected at each stage of boosting. We employed two features such as Histogram-equalization feature and Edge feature for our experiment. The proposed method is applied to the MIT CBCL pedestrian image database, and 100 sub-regions are extracted from each image as local-features. The experimental result shows fairly good classification ratio with single feature, while the improvement on classification ratio with the combination of two feature is small. However, the combination of features effects to select *good* local-features for base learners.

**Key words** AdaBoost, Support Vector Machine (SVM), Feature Selection, Pedestrian Detection

### 1. はじめに

本稿では、AdaBoost [1] の各段の弱識別器に対して最適な特徴量と局所領域を自動的に選択する手法を提案する。提案手法は、顔や車のような複数の部品が組み合わされた対象に適用可能であるが、本稿では歩行者の検出を例題として用いる。歩行者の検出は、歩行者の姿勢のバリエーションが豊富なため単一のモデルでは対応し難いと考えられ、顔や車の検出よりも難しいと考えられている。

Gavrila [3] は、階層的テンプレートマッチングにより、入力

画像から歩行者の候補を見付ける手法を提案した。この手法では、典型的な歩行者の外形のエッジ画像を用意し、入力画像のエッジ特徴との類似性（差異）を chamfer 距離によって定義する。歩行者の姿勢のバリエーションにはテンプレートのバリエーションによって対応する事になり、十分な数、種類のテンプレートを用意できる事が、この手法では重要となる。

Mohan [2] らは、Adaptive Combination of Classifiers (ACC) を歩行者検出に適用した。この手法は、2 段階の階層的識別器により構成される。一段目は、頭部、脚部、右腕、左腕の部品を検出する四つの識別器が訓練される。二段目は、

一段目の結果を組み合わせる「人」と「人以外」を区別する識別器が構成される。各識別器にはサポートベクターマシン (SVM) [5], [6], [7] が用いられる。彼らの手法により、全身の識別器よりも部品ごとの識別器を組み合わせの方が識別性能が高い事が示されている。一方、この手法においては、部品があらかじめ定められているため、必ずしも識別に最適な部品となっていない。

Viola [4] らは、動き情報と画像の輝度情報を統合する事で、歩行者検出を行う手法を提案した。この手法では、検出器は二つの連続するビデオフレームを走査し、AdaBoost を用いる事で、動きと姿勢の両者の利点を生かすようにした。二つのモダリティを組み合わせることにより、高速 (約 4 フレーム/秒) で false positive 率の低い検出器を構成する事ができた。

Viola により、動き情報を組み合わせる事の利点が示されたが、カメラ自身が固定されていない場合には、この手法を用いる事は難しい。そこで、我々は、静止画での姿勢を元にした歩行者検出手法をとることとした。本稿での提案手法では、Adaboost の弱識別器としてソフトマージン・サポートベクターマシン (SVM) を採用し、boosting の格段で最適な特徴、局所領域の選択を行う。実験では、単一の特徴で局所領域の選択を行う事によって高い認識性能が示された。二つの特徴を組み合わせる事による認識率の改善は大きくなかったが、単一の特徴のみの場合に比べて「良い」局所領域が選択されるようになっている。

次節で、我々の手法について述べ、続いて実験結果を示す。

## 2. 特徴選択と Boosting による歩行者識別

### 2.1 アルゴリズム

図 1 に、特徴選択と AdaBoost を組み合わせた歩行者識別手法のアルゴリズムを示す。本稿では、「特徴」はエッジ特徴などのように画像情報に何らかの関数を作用させる事で得られるもの、「局所領域」は画像の一部とする (図 2)。Boosting の各段で最適な特徴と局所領域を選択していく。Boosting の  $i$  段目では、 $i-1$  段目で定められた入力重みで、全ての特徴、局所領域に対する弱識別器を訓練する。そして、その中で最もエラー率の低い弱識別器を  $i$  段目の弱識別器とする。次段の入力重みと識別器の重みは、その識別器のエラー率を元に計算される。

歩行者画像として MIT CBCL の歩行者画像データベースの 926 枚の  $128 \times 64$  画素の画像を用い、歩行者以外の画像は我々が収集した 2,000 枚の画像を用いた。この画像を 64 times 32 画素に低解像度化し、特徴抽出して識別器のデータとした。特徴量として、ヒストグラム均一化特徴 (コントラストが強調される) とエッジ特徴の二つを採用した。それぞれの特徴画像を  $4 \times 8$  画素の領域から順次領域を拡大しながら画像全体を走査し、それぞれ 100 個の部分画像を生成し、これを局所領域とした。Fig. 2 に、原画像、特徴抽出画像 (二枚)、局所領域の例を示す。歩行者画像 700 枚、非歩行者画像 700 枚を訓練データとし、歩行者 200 枚、非歩行者 200 枚のテストデータで汎化誤差の評価を行った。

### 2.2 ソフトマージン・サポートベクターマシンの Boosting

本節では、サポートベクターマシン (SVM) の概略を述べ、次いでソフトマージン・サポートベクターマシン (ソフトマージン SVM) およびカーネル・サポートベクターマシン (カーネル SVM) について補足する。

入力ベクトルに対し、識別関数

$$y = \text{sign}(\mathbf{w}^T \mathbf{x} - h) \quad (1)$$

により 2 値 ( $\pm 1$ ) を出力するモデルを考える。ここで、 $\mathbf{x}$  は入力ベクトル、 $\mathbf{w}$  入力ベクトルに対する重み、 $h$  は閾値である。関数  $\text{sign}(u)$  は、 $u > 0$  のとき 1 をとり、 $u \leq 0$  のとき  $-1$  をとる符号関数である。

SVM は、最大マージンによる分離超平面を求める。ここで、マージンとは分離超平面に最も近いサンプルまでの距離である。訓練サンプルが線形分離可能ならば、

$$t_i(\mathbf{w}^T \mathbf{x}_i) \geq 1, \quad i = 1, \dots, N. \quad (2)$$

を満たすパラメータが存在する。これは、 $H1: \mathbf{w}^T \mathbf{x} - h = 1$  と  $H2: \mathbf{w}^T \mathbf{x} - h = -1$  の 2 枚の超平面で訓練サンプルが完全に分離されており、2 枚の超平面の間にはサンプルがひとつも存在しないことを示している。これらの超平面との距離 (マージンの大きさ) は、 $\frac{1}{\|\mathbf{w}\|}$  となる。したがって、マージンを最大とするパラメータ  $\mathbf{w}$  と  $h$  を求める問題は、結局、制約条件

$$t_i(\mathbf{w}^T \mathbf{x}_i - h) \geq 1, \quad (i = 1, \dots, N) \quad (3)$$

の下で、目的関数

$$L(\mathbf{w}) = \frac{1}{2} \|\mathbf{w}\|^2 \quad (4)$$

を最小とするパラメータを求める問題と等価になる。

ソフトマージン法では、マージン  $\frac{1}{\|\mathbf{w}\|}$  を最大としながら、幾つかのサンプルが超平面  $H1$  あるいは  $H2$  を越えて反対側に入ってしまうことを許す。反対側にどれくらい入り込んだかの距離を、パラメータ  $\xi_i (\geq 0)$  を用いて、 $\frac{\xi_i}{\|\mathbf{w}\|}$  と表すとすると、その和

$$\sum_{i=1}^N \frac{\xi_i}{\|\mathbf{w}\|} \quad (5)$$

はなるべく小さいことが望ましい。これらの条件から最適な識別面を求める問題は、制約条件

$$\xi_i \geq 0, \quad t_i(\mathbf{w}^T \mathbf{x}_i - h) \geq 1 - \xi_i, \quad (i = 1, \dots, N) \quad (6)$$

の下で、目的関数

$$L(\mathbf{w}, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i \quad (7)$$

を最小とするパラメータを求める問題に帰着される。ここで、あらたに導入したパラメータ  $C$  は、第 1 項のマージンの大きさと第 2 項のはみ出しの程度とのバランスを決める定数である。

カーネルトリックは SVM の識別性能を劇的に向上したが、

(1) Let  $N$  be the number of samples,  $M$  be the number of boosting steps,  $L$  be the number of local-features,  $K$  be the number of features. Thus,  $K \times L$  be the total number of local features in Local-Feature pool.

(2) Generate Local-Feature pool for all the features from input samples  $x$ , such as  $x \rightarrow x^{11}, \dots, x^{1L}, x^{21}, \dots, x^{k1}, \dots, x^{KL}$ .

(3) Initialize the observation weights  $w_i = 1/N, i = 1, 2, \dots, N$ .

(4) For  $m = 1$ , to  $M$ :

(a) For  $k = 1$  to  $K$ , for  $l = 1$  to  $L$

i. Fit a classifier  $G_m^{kl}(x^{kl})$  to the training samples of local-feature  $x^{kl}$  that are randomly selected depending on the weights  $w_i$  from all the training samples

ii. Compute  $err_m^{kl} = \frac{\sum_{i=1}^N w_i I(y_i \neq G_m(x_i^{kl}))}{\sum_{i=1}^N w_i}$ .

(b) set  $err_m$  with the smallest  $err_m^{kl}, l = 1, 2, \dots, L, k = 1, 2, \dots, K$ .

(c) set  $G_m(x) \leftarrow G_m^{kl}(x^{kl})$  with  $k$  and  $l$  in above step.

(d) compute  $\alpha_m = \log((1 - err_m)/err_m)$ .

(e) set  $w_i \leftarrow w_i \cdot \exp[\alpha_m \cdot I(y_i \neq G_m(x_i))], i = 1, 2, \dots, N$ .

(5) Output  $G(x) = \text{sign}[\sum_{m=1}^M \alpha_m G_m(x)]$ .

図1 特徴選択を用いた AdaBoost のアルゴリズム

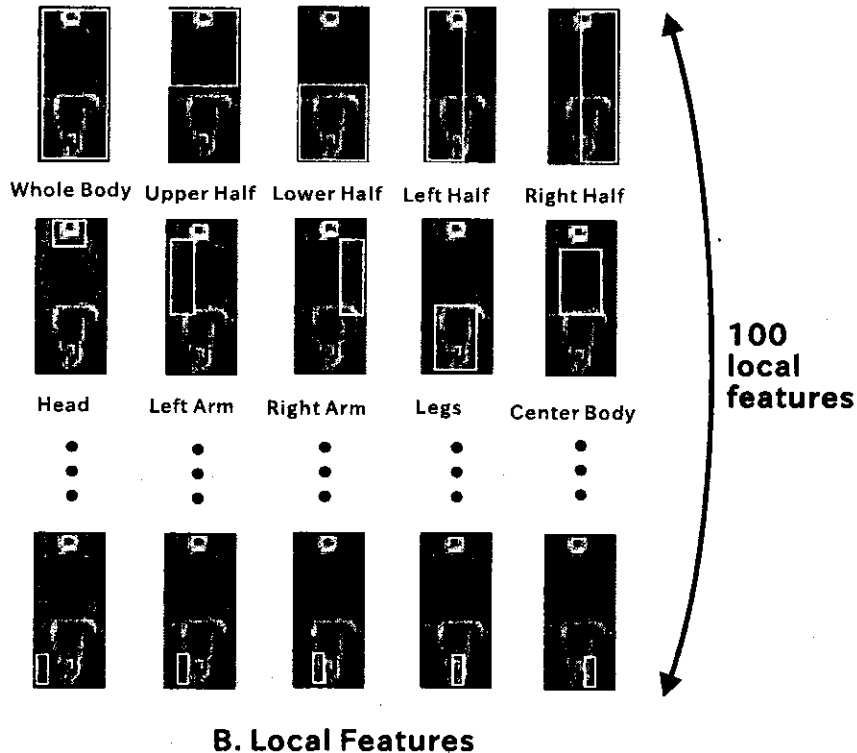
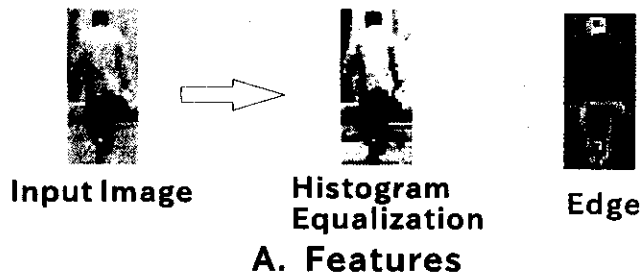


図2 サンプル画像と特徴、および、局所領域の例

この手法はソフトマージン SVM にも適用できる。カーネルトリックでは、元の特徴ベクトルは非線形写像  $\phi(\mathbf{x})$  により変換され、写像先で線形識別が行われる。SVM が二つの入力ベクトルの内積のみ依存するため、非線形写像を直接計算しなくとも、写像先での入力ベクトルの内積を用いる事ができる。つまり、もし非線形に写像した空間での二つの要素  $\phi(\mathbf{x}_1)$  と  $\phi(\mathbf{x}_2)$  の内積が

$$\phi(\mathbf{x}_1)^T \phi(\mathbf{x}_2) = K(\mathbf{x}_1, \mathbf{x}_2) \quad (8)$$

のように、入力特徴  $\mathbf{x}_1$  と  $\mathbf{x}_2$  のみから計算できるなら、非線形写像によって変換された特徴空間での特徴  $\phi(\mathbf{x}_1)$  や  $\phi(\mathbf{x}_2)$  を陽に計算する代わりに、 $K(\mathbf{x}_1, \mathbf{x}_2)$  から最適な非線形写像を構成できる。ここで、このような  $K$  のことをカーネルと呼んでいる。

コストパラメータ  $C$  を導入する事により、AdaBoost でのサンプル重みを実現する方法が二つ考えられる。一つは、コストパラメータ  $C$  を AdaBoost でのサンプル重みとして SVM を記述する方法。もう一つは、 $C$  を重みとして入力サンプルをサンプルしなおす方法である。Schwenk らは [8]、AdaBoost において擬似損失関数を定義する事と再サンプリングが同様の効果を持つ事を示しており、我々は再サンプリング法を用いて、LIBSVM [9] ライブラリを利用する事とした。訓練サンプル 1400 枚 (歩行者 700 枚、非歩行者 700 枚) から、1000 枚の再サンプリングを行った。

### 3. 実験結果

まず、SVM のコストパラメータ  $C$  の効果を単一の特徴量、10 個の局所領域で検証した。この実験では、入力画像の輝度値をそのまま入力ベクトルとして用い、10 個の局所領域とは全身、上半身、下半身、右半身、左半身、右腕、左腕、頭部、脚部、胴体中央である。図 3 に Boosting 段数に対するエラー率 (テストエラー) を示す。コストパラメータ  $C$  は 0.1, 0.7, 100 の三種である。この実験では、10 個の局所領域はほぼ均等に選択された。 $C = 0.7$  および  $C = 100$  において、訓練エラーは早い段階 ( $C = 0.7$  で 10 段、 $C = 100$  で 5 段の Boosting) でゼロに収束しているが、テストエラーはその後も漸次減少を続け、100 段の Boosting により  $C = 0.7$  で 4%、 $C = 1$  で 4.5%、 $C = 0.1$  で 5.25% となった。この結果は、Shapire [10] の Boosting では過学習が起りにくいという実験結果と一致する。

図 4 に Boosting50 段でのコストパラメータ  $C$  に対するエラー率を示す。テストエラーは、 $C = 0.7$  で最小値 4% を示した。この結果は、通常の SVM よりもソフトマージン SVM の方が識別率の点で有利である事を示す。また、これは、ソフトマージン SVM の評価式 (7) を最小化する際に、第 1 項のマージンと第 2 項のみ出しの影響のバランスをとることで、最適値が得られる事を示している。

図 5 に、ヒストグラム均一化特徴、エッジ特徴、両特徴の組み合わせによる Boosting100 段までのエラー率を示す。先の実験結果に基づき、コストパラメータ  $C$  は 0.7 とした。訓練エ

ラーで見るとヒストグラム均一化特徴がエッジ特徴よりも低い値を示しており、訓練エラー率が零に収束するのもし早い。一方、テストエラーで比較すると、エッジ特徴の方がエラー率が低く Boosting100 段の後では 3% となっているのに対し、ヒストグラム均一化特徴は 3.5% である。

二つの特徴の組み合わせでは、100 段中 84 段がヒストグラム均一化特徴を選択し、エッジ特徴を選択したのは 16 段であった。これは、訓練エラーによって特徴を選択している事が原因と考えられる。Boosting100 段後のテストエラーは 3% で、16 回しか選択されなかったエッジ特徴でのエラー率と同等である。また、91 段目では、2.5% とエッジ特徴よりも良い成績を示していた。

局所領域として 100 個の部分画像を用意したが、そのうち 50 個程度しか選択されなかった。図 6 は、選択された局所領域の例を示す。この結果は、単一特徴を使用した場合よりも複数 (二つ) の特徴を組み合わせの方が、小さな局所領域からある程度大きな局所領域までが満遍なく選択されて傾向があることを示す。特徴量の組み合わせにより、より「良い」局所領域が選択されると考えられる。

### 4. おわりに

特徴選択とソフトマージン SVM の Boosting を用いた歩行者検出手法を提案した。実験では、テストエラーで 3% と高い汎化能力を示した。本稿では、ヒストグラム均一化特徴とエッジ特徴の二つを組み合わせしたが、これによる識別性能の改善は大きくはなかった。しかし、より「良い」局所領域が選択されるようになっており、汎化性能向上への可能性を示している。

#### 文 献

- [1] T.Hastie, R.Tibshirani, J.Friedman, *The Elements of Statistical Learning - Data Mining, Inference, and Prediction*, Springer-Verlag, 2001.
- [2] A.Mohan, C.P.Papageorgiou and T.Poggio, "Example-Based Object Detection in Images by Components," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.23, No.4, pp.349-361, 2001.
- [3] D.M.Gavrira, "Pedestrian Detection from a Moving Vehicle", *Proc. of European Conference on Computer Vision*, pp.37-49, 2000.
- [4] P.Viola, M.J.Jones and D.Snow, "Detecting Pedestrians Using Patterns of Motion and Appearance", *Proc. of Int'l Conf. on Computer Vision*, pp.734-741, 2003.
- [5] V.N.Vapnik, *Statistical Learning Theory*, John Wiley & Sons (1998).
- [6] B.Scholkopf, C.J.C.Burges, A.J.Smola, *Advances in Kernel Methods - Support Vector Learning*, The MIT Press, 1999.
- [7] N.Cristianini, J.S-Taylor, *An Introduction to Support Vector Machines and other kernel-based learning methods*, Cambridge University Press, 2000.
- [8] H.Schwenk and Y.Benjio, "Boosting Neural Networks", *Neural Computation*, Vol.12, pp.1869-1887, 2000.
- [9] C.C.Chung and C.J.Lin, "LIBSVM: a library for support vector machines", *Software available at* <http://www.csie.ntu.edu.tw/>, 2001.
- [10] R.E.Schapiro, "The Boosting Approach to Machine Learning An Overview", *MSRI Workshop on Nonlinear Estimation and Classification*, 2002.

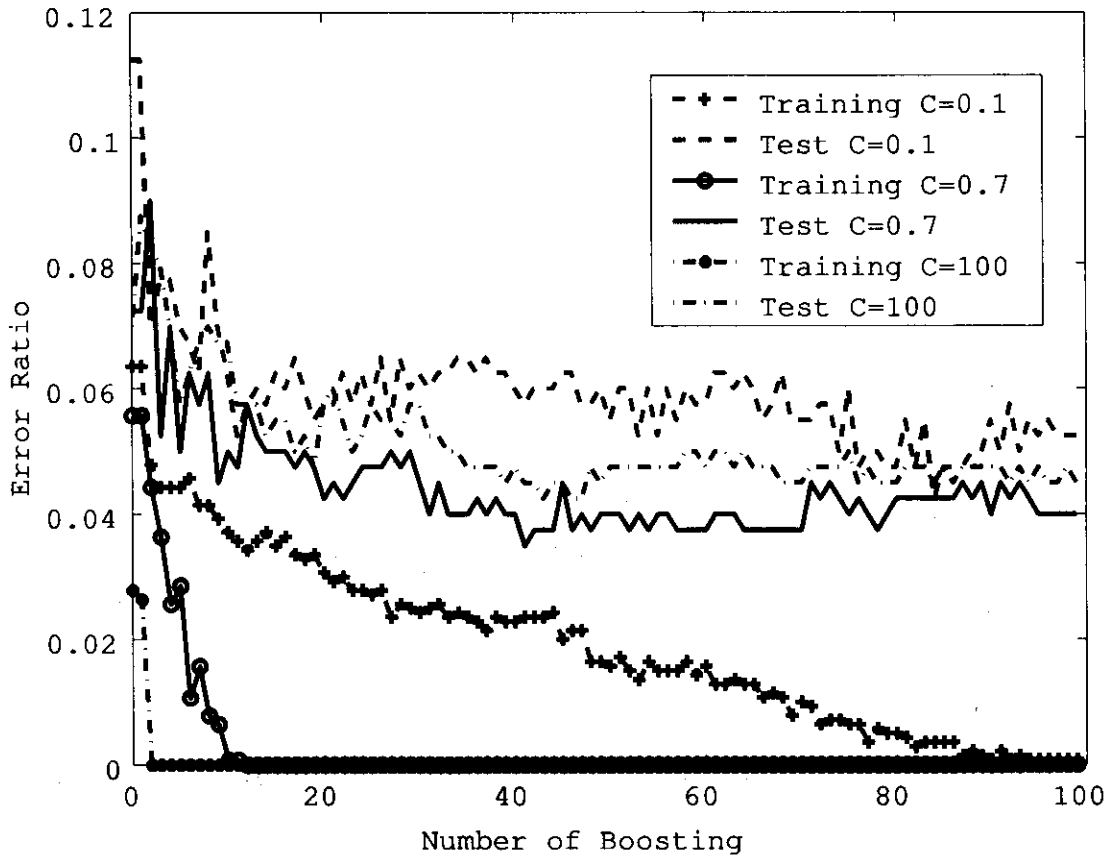


図3 単一特徴、10局所領域でのエラー率

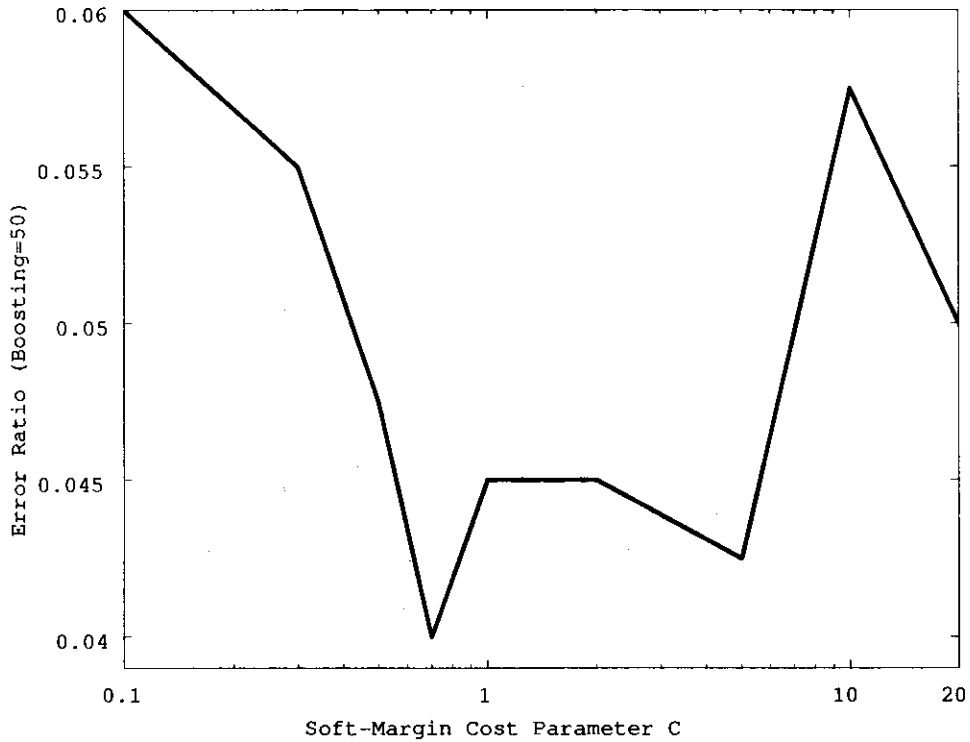


図4 コストパラメータCとエラー率の関係

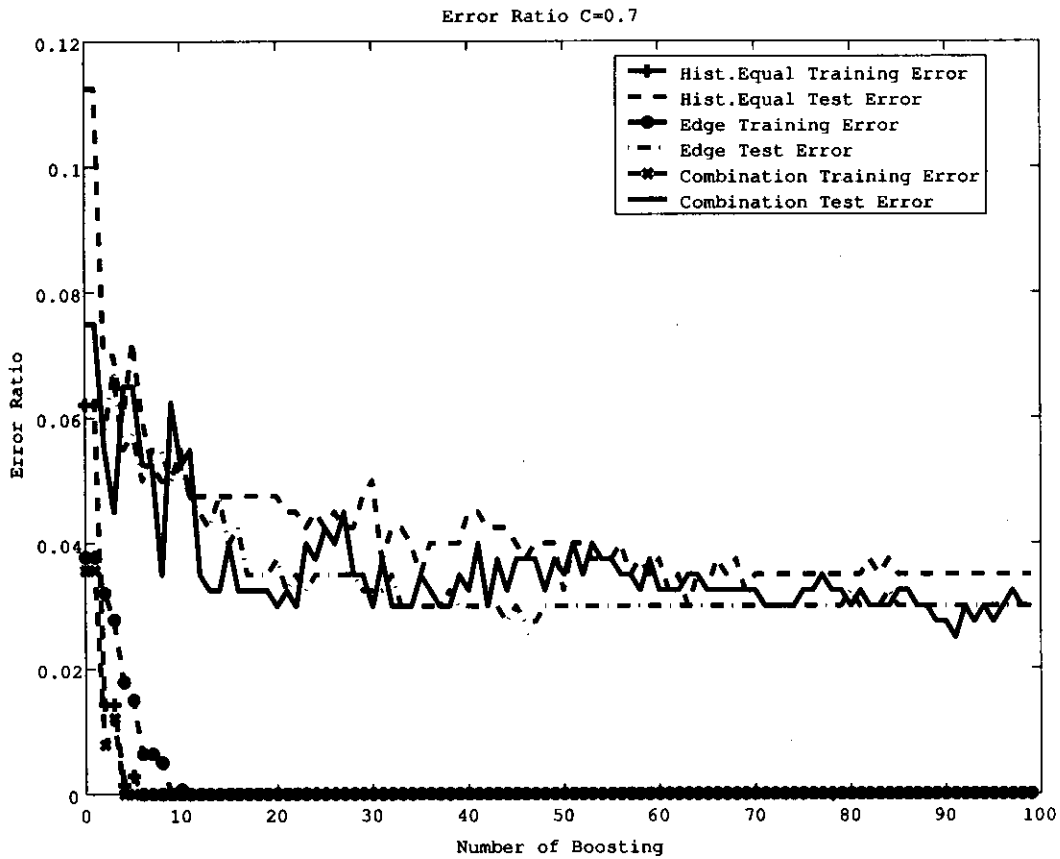


図 5 100 局所領域でのエラー率

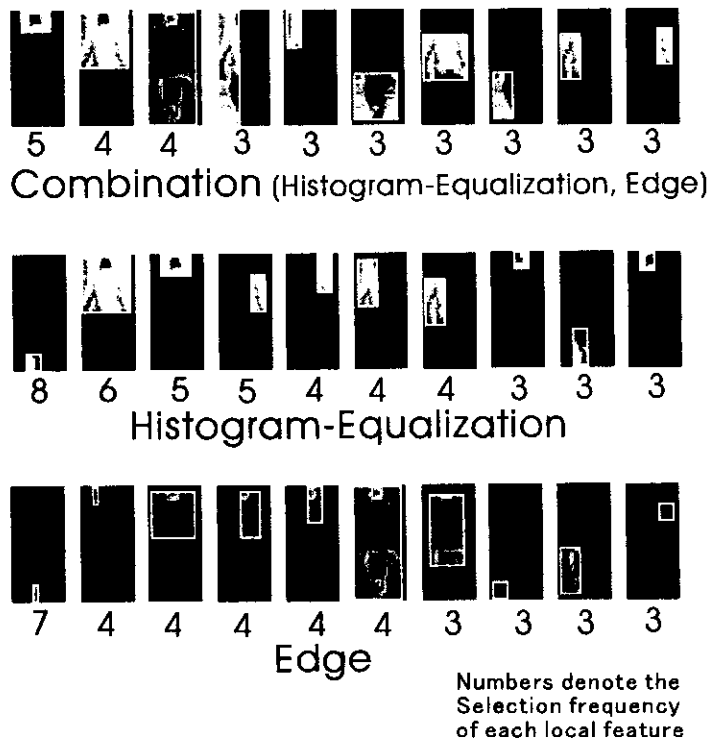


図 6 選択された局所領域の例