

# 藤吉研究室 卒業研究発表 アブストラクト

工学部情報工学科 藤吉研究室



Convolutional Neural Networkの特徴抽出過程における不変性獲得の調査 福井 宏

Random Forestを用いた能動学習における有効なサンプル選択 村田 隆英

決定木を用いた距離画像からの高速なエッジ検出 金子 将也

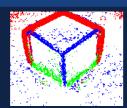
直線成分に基づくHOL特徴量による誤検出の抑制 宮下 卓也

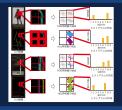
多クラス識別のための行列分解と早期棄却による物体検出の高速化 黒川 貴都













# FLAB.

E-mail: hf@cs.chubu.ac.jp

URL: http://www.vision.cs.chubu.ac.jp

#### 指導教授:藤吉弘亘

#### 1.はじめに

Deep Learning は、従来のパーセプトロンの中間層を複数用いた構造を持つニューラルネットワークであり、最適な特徴量を学習により獲得することができる。本研究ではConvolutional Neural Network(以下 CNN) を対象とし、CNN における特徴量の自動獲得と、その有効性を調査することを目的とする。

# 2.多層パーセプトロンと CNN

# 2.1.多層パーセプトロン

多層パーセプトロン (MLP) は,入力層,中間層,出力層で構成される非線形な多クラス識別器である.各階層間の重みは,誤差逆伝播法による確率的勾配降下法で求める. 識別時は,特徴量を順伝播し,出力層の各ユニットの出力を各クラスの事後確率として識別する.

# 2.2. Convolutional Neural Network

CNN は Deep Learning の一種であり、中間層で畳み込みとプーリングの処理を繰り返し行うニューラルネットワークである。CNN の構造は、図1に示すように特徴抽出部と識別部の2つから構成される。

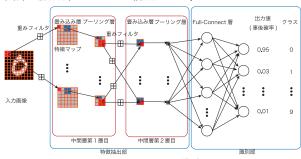


図 1: CNN の構造

特徴抽出部 特徴抽出部では、重みフィルタの畳み込みとプーリングを繰り返すことで特徴を抽出する。入力画像に対して、重みフィルタをラスタスキャンさせ、繰り返し畳み込みを行うことで特徴マップを得る。その後、特徴マップに対してプーリングを行う。プーリングとは、特徴マップの小領域から値を出力して新たな特徴マップに変換する処理であり、識別部に入力する特徴次元削減の働きがある。識別部 識別部は、従来の多層パーセプトロンと同じ構造であり、特徴抽出部にて取得した特徴を入力して識別を行う。CNN の学習は、誤差逆伝播法によるミニバッチ学習により各層の重みを更新する。

## 3.評価実験

CNN が学習により獲得した特徴と、その効果を調査するために2つの実験を行う。まず、獲得した特徴量を調査するために,各層の重みフィルタを可視化する。次に、幾何変化に対する不変性について CNN+ランダム学習を多層パーセプトロンと比較する。ランダム学習は学習サンプルにランダム性を持たせ、汎化性能を向上させる手法である。

## 3.1. 実験概要

本実験では、0から9の文字認識を対象とする。データセットにはMNIST Dataset を使用する。学習サンプルは50,000 枚、評価サンプルに10,000 枚を使用する。CNN の重みフィルタは $5\times5$  のサイズ、1 層目に6 枚、2 層目に14 枚使用する。

# 3.2.重みフィルタの可視化

学習前後の各重みフィルタの可視化例を図2に示す.

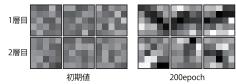


図 2: 重みフィルタの可視化

学習によって重みフィルタの濃淡が,入力画像の特定のエッジに反応するフィルタを学習過程で自動的に獲得していることがわかる

#### 3.3. CNN の幾何変化に対する不変性の実験結果

識別サンプルに平行移動と回転を与えて幾何変化に対する不変性を調査する。このとき、MLP、CNN、CNN+ランダム学習の誤識別率を比較する。ランダム学習では、0から 20°の回転と、0から 5pixel の平行移動により毎回学習サンプルを生成する。図3に、平行移動と回転に対する各手法の誤識別率の推移を示す。図3から、CNNの識別率は MLP より高いことがわかる。また、ランダム学習を導入することで、精度をさらに向上させることができた。

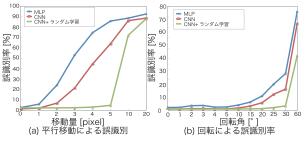


図 3:幾何変化による誤識別率の推移

図4にランダム学習を導入した CNN と通常の CNN の学習誤差の推移を示す。ランダム学習では毎回異なるパターンを教示するため、学習誤差は振動していることがわかる。しかし、epoch 数の増加に伴い学習誤差は収束していくため、汎化性能が高くなる。

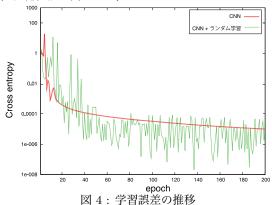
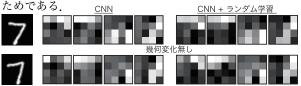


図5に、入力サンプルを10°回転した場合としない場合のプーリング層の可視化例を示す。ランダム学習したCNNの特徴マップは、回転の影響を受けず変化が少ないことがわかる。これは、プーリングが微小な変化を吸収しているためである。



10<sup>\*</sup>回転 図 5: ランダム学習による特徴マップの変化

# 4.おわりに

本研究では、CNN の特徴量の学習が識別に有効であることを示した。また、ランダム学習の導入により不変性の向上が確認できた。CNN は複数の重みフィルタの畳み込みをするため識別に時間を要する。そこで、今後は識別の高速化の研究を行う予定である。

# 参考文献

Y.Lecun, L.Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition", Proceedings of the IEEE, pp. 2278-2324, 1998.

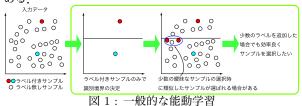
#### 指導教授:藤吉弘亘

#### 1.はじめに

大量の学習データを用いて識別器を構築する際,どのサンプルに教師ラベルを付けることで、効率良く識別境界を決定できるかという問題がある。能動学習では、一度に追加するサンプルが少数の場合、類似したサンプルが選択され効率が悪いという問題がある。そこで、本研究では、Random Forest(RF)の密度推定を用いて適切なサンプルを選択してラベル付けすることで、効率良く識別境界を決定する手法を提案する.

#### 2.問題設定

能動学習 [1] では、図 1 に示すように、識別境界の定まらない領域にある曖昧なサンプルを選択してラベル付けを行い、再学習することでより良い識別境界を求める。従来、追加サンプルの選択には、複数の識別器の出力に統一性の無いサンプルを選択する Vote Entropy が利用されている。従来法では、サンプルの分布を考慮しないため、類似したサンプルを選択することがあり、効率が悪いという問題がある。



# 3.提案手法

本研究では、Density Forest による密度推定結果に着目したサンプル選択を行う。ラベルを追加すべきサンプルは、密度推定にばらつきがある曖昧な領域と、ばらつきは無いがラベル付きサンプルが周囲に存在しない領域にあると考え、この二つの領域に同時にラベルを追加するサンプル選択法を提案する。以下に提案手法の流れを述べる。

# Step1:密度推定とラベル伝播による学習

ラベル付きサンプル集合  $S^{(s)}$  とラベル無しサンプル集合  $S^{(u)}$  を用いて Density Forest を構築する。学習後の各密 度木の密度分布を図 2(a) に示す。次に,各密度分布が連結している方向にラベルを伝播することで,ラベル無しサンプルにラベルを付与する。伝播結果により末端ノードのクラス分布を作成する。









(a) 各木の密度分布 (b) 密度 (b) 密度 (c) 名 (c) 名 (c) 名 (d) 2 (d

Step2: サンプルの曖昧さと密度分布の類似度の算出 学習結果から,ラベル無しサンプル集合  $S^{(u)}$  に属している  $x_i$  が入力されたときの Vote Entropy の値  $VE(x_i)$  と  $x_i$  が到達した末端ノードが持つ密度分布の類似度  $D(x_i)$  を算出する.複数の密度分布間の類似度 D は,シャノンの情報量を用いた JS-Divergence により式 (1) を用いて算出する

$$D(\mathcal{N}_1, \mathcal{N}_2, \cdots, \mathcal{N}_T) = H\left(\sum_{t=1}^T \mathcal{N}_t\right) - \sum_{t=1}^T H\left(\mathcal{N}_t\right) \quad (1)$$

ここで、 $\mathcal{N}_t$  は t 本目の決定木におけるサンプル  $\boldsymbol{x}_i$  の密度分布、 $H(\cdot)$  はシャノンの情報量を示す。図 2(b) に各サンプルの密度分布の類似度を示す。

#### Step3:密度分布の類似度を考慮したサンプルの選択

Vote Entropy の値  $VE(x_i)$  と、密度分布の類似度  $D(x_i)$  を用いてサンプルの選択を行う。ここでは、 $D(x_i)$  から類似度が高いサンプル集合と低いサンプル集合に分け、それぞれ  $VE(x_i)$  の値が最大となるサンプルを選択する。

# Step4:ラベルの追加

選択されたサンプルに対して人手によりラベル付けを行う.

本手法では、各推定クラスに対して1度に2個のサンプルが選択される。

#### Step5:ラベルの再伝播によるクラス分布の更新

ラベルの再伝播を行い、各木の末端ノードのクラス分布を 更新する。提案手法では、決定木を再構築することはしな い。Step2~Step5 を一定の条件に達するまで繰り返すこ とで識別境界を決定していく。

#### 4.評価実験

評価実験では、提案手法と従来法の Vote Entropy を比較する.

## 4.1.実験概要

従来法と提案手法のラベルの追加回数を比較する. 両手法においてラベル伝播の際に用いる密度推定の結果は同じものを用いる. 実験には,スパイラルデータ(2次元)を使用する. RFのパラメータ木の数は400本,木の深さは10とする. ラベルの追加の終了条件は識別率が一定の値に達した場合とする.

#### 4.2.実験結果

図3に識別率が99%に達するまでのラベルの追加回数とラベル再伝播後の識別率を示す。提案手法は、従来法よりラベルの追加回数を削減することができた。

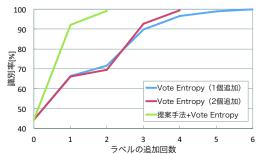
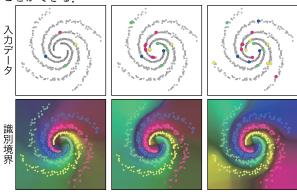


図 3: 追加学習の回数の比較

図4に、従来法と提案手法よりラベルを追加(2回)した際の入力サンプルと識別境界を示す。従来法は、密度推定にばらつきがある曖昧な領域の類似したサンプルが選択されるため、2個追加した場合でも識別境界が大きく変化しない。提案手法は、ラベル付きサンプルが周囲に存在しない領域のサンプルを追加することができるため、類似したサンプルの選択を抑制し従来法と比べて識別境界が大きく変化する。これにより、少ない追加回数で高い識別率を得ることができる。



初期 (44.4%) 図 4: 各手法の入力データと識別境界

#### 5.おわりに

本研究では、密度分布の類似度を考慮したサンプル選択法を提案した。提案手法を導入することで能動学習における追加回数を削減することができた。今後は大規模なデータセットに提案手法を適用する予定である。

#### 参考文献

 B. Settles, "Active Learning Literature Survey", Computer Sciences Technical Report 1648, University of Wisconsin-Madison, 2009.

# 決定木を用いた距離画像からの高速なエッジ検出

EP10035 金子 将也

#### 1.はじめに

距離画像におけるエッジには、ジャンプエッジ、凸ルーフエッジ、四ルーフエッジが存在し、これらを検出する手法としてリングオペレータが提案されている [1]. しかし、リングオペレータは注目画素周辺の画素値に対してフーリエ変換を施しエッジ検出するため、処理速度が遅いという問題がある。そこで、本研究では決定木を導入した距離画像の高速なエッジ検出法を提案する.

# 2. 距離画像におけるエッジとリングオペレータ

距離画像におけるエッジは、物体と背景の境界であるジャンプエッジ、物体の面と面の境界である凹ルーフエッジ、凸ルーフエッジの3種類である。この、3種類のエッジと平面を分類する手法としてリングオペレータが提案されている(図1). リングオペレータは、距離画像からリング状に抽出した画素値を、周期  $2\pi$  の周期関数としてフーリエ変換する.フーリエ変換により得られた振幅スペクトルをしきい値処理することにより、ジャンプエッジ、凸ルーフエッジ、凹ルーフエッジと平面に分類する.

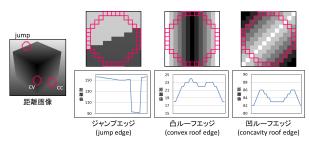


図1:リングオペレータ

リングオペレータは, 距離画像の各画素に対してフーリエ変換を施すため, 処理時間を要するという問題がある.

#### 3.提案手法

本研究では、リングオペレータによるエッジ検出問題に、FAST と同様に決定木を導入し、分類問題として高速化する.注目画素の距離値  $D_p$  とその周囲 32 画素の距離値  $D_{p\to x}$  を、式 (1) により Far, Similar, Near の 3 値にした特徴ベクトルを抽出する.

$$S_{p\to x} = \begin{cases} Far & D_p + t \le D_{p\to x} \\ Similar & D_p - t < D_{p\to x} < D_p + t \\ Near & D_{p\to x} \le D_p - t \end{cases}$$
(1)

ここで、 $S_{p\to x}$  は三値化した円上の画素を、t はしきい値を表す。この三値化した 32 次元の特徴ベクトルを入力として機械学習により決定木を構築する。学習に用いるサンプルは、OpenGL で生成した距離画像に、リングオペレータを適用し、面、ジャンプ、凸、凹のエッジに分類した結果を正解ラベルとして用いる。各ノードの分岐関数では、式 (2) から求められる情報利得が最大となる周囲の画素を選択する。

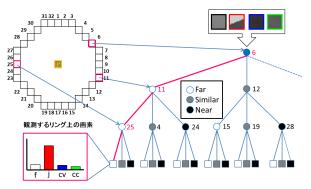
情報利得 = 
$$H(P) - H(P_f) - H(P_s) - H(P_n)$$
 (2)

ここで, P は親ノードにあるサンプル集合を表し,  $P_f$ ,  $P_s$ ,  $P_n$  はそれぞれ, 子ノードにおいて Far, Similar, Near と分類された各集合である. H は情報エントロピーであり式(3) より求められる.

$$H(P) = (f + j + cv + cc) \log_2(f + j + cv + cc)$$

$$-f\log_2 f - j\log_2 j - cv\log_2 cv - cc\log_2 cc \qquad (3)$$

ここで、f, j, cv, cc はそれぞれ、面 (f), ジャンプエッジ (j), 凸ルーフエッジ (cv), 凹ルーフエッジ (cc) の各ラベルに属するサンプル数を表している. 対象のノードの情報利得が0になるまで、子ノードを生成する. リーフノードには、4つのラベルのうち最も多く辿り着いたクラスを記録しておく (図 2). エッジ検出時は、距離画像をラスタスキャンして、注目画素を決定木に入力する. 決定木をトラバーサルして辿り着いたリーフノードにより、面、ジャンプエッジ、凸ルーフエッジ、凹ルーフエッジを判定する.



指導教授:藤吉弘亘

図 2: リングオペレータの決定木の識別

#### 4.評価実験

OpenGLで生成した距離画像に正しいラベル付けを行った結果を真値とし、作成した識別器のエッジ検出結果が何画素一致しているかを評価する. 図 3 は各手法におけるエッジ検出結果を示す. 各ラベルの色は、平面が白、ジャンプエッジが赤、凸ルーフエッジが青、凹ルーフエッジが緑である.

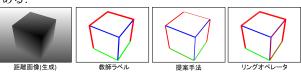


図 3:生成した距離画像とエッジ検出結果 提案手法と従来法 (リングオペレータ) の各ラベルにおける精度と VGA サイズの距離画像における処理時間を表 1 に示す.表 1 より,提案手法は従来法と同等以上の精度で 計算コストを約 1/10 に抑えることができた.

表 1: エッジ検出精度と速度の比較

	エッジ検出精度 [%]				処理時間 [sec]
	flat	jump	cv	cc	CONTROL [SCC]
従来法	99.9	98.3	68.5	61.9	0.98
提案手法	98.7	94.3	97.3	94.1	0.11

また図4に、TOFカメラにより取得した距離画像に対するエッジ検出結果を示す。TOFカメラの距離画像にはノイズを含むため、平面の領域に誤検出が発生しているが、ジャンプエッジ、凸ルーフエッジ、凹ルーフエッジを検出できている。

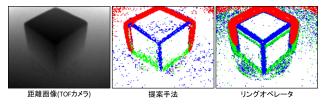


図 4: TOF カメラによる距離画像のエッジ検出結果

#### 5.おわりに

提案手法では、リングオペレータに決定木を導入して高速なエッジ検出法を示した. 従来法と比較して精度と処理時間を向上させることができた. 今後は、距離画像におけるノイズへの対処法と物体の構造理解への応用について検討する.

# 参考文献

- [1] 松田文男, 仁田武志, 井口征土,"リングオペレータによる距離 画像のラベリング",電子情報通信学会論文誌 D, Vol. J66-D, No. 10, pp. 1161-1168, 1983.
- [2] E. Rosten, R. Porter, T. Drummond, "Faster and better: a machine learning approach to corner detection", IEEE Transactions on PAMI, Vol. 32, No. 1, pp. 105 - 119, 2010.

#### 直線成分に基づく HOL 特徴量による誤検出の抑制

EP09130 **宮下 卓也** 

#### 指導教授:藤吉弘亘

#### 1.はじめに

勾配方向ヒストグラムに基づく HOG 特徴量 [1] は, 照明変化や幾何学的変化に頑健であるため人検出に利用されている.しかし, 背景画像に含まれる直線は, 人の輪郭と似た見えを持つため, 誤検出をするという問題がある. そこで, 本研究では人と直線を含む背景の違いを捉える特徴量を提案し, 誤検出の抑制を目的とする.

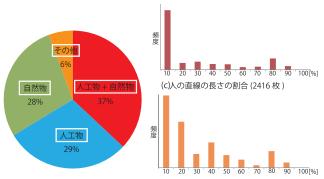
## 2. 誤検出の傾向調査

まず, 従来の人検出法がどのような背景画像を誤検出するか調査する. HOG 特徴量と SVM による誤検出例を図1に示す.



図 1: 従来法で発生する誤検出の例

誤検出した領域は、検出ウィンドウを縦断する直線を含んでおり、特に車両や建物等の人工物を人と誤検出していることがわかる。そこで、誤検出の要因として考えられる背景に存在する「物体の属性」、「直線がどの程度存在するのか」を調査する。車載カメラから撮影されたデータセット KITTI Vision Benchmark Suite から 7,481 枚の画像に対して、人検出器を適用し、その際の誤検出領域の傾向を調べる。図 2 に誤検出した際に含まれる物体の属性と直線の割合を示す。誤検出する傾向として、図 2 の (a) から、ウィンドウ内に人工物が含まれる割合が 20,129 枚あり、全体の約 6 割であることがわかる。また、図 2 の (b) からはウィンドウの 20~90% を占める長い直線が含まれていることがわかる。



(a) 画像内に存在するオブジェクトの属性 (b) 背景の直線の長さの割合 (30341 枚)

図 2: 誤検出の傾向調査

## 3.HOL 特徴量

2. における調査から、誤検出の要因として、直線に反応することが判明した. そこで、本研究では、直線を考慮したHistograms of Oriented Lines(HOL) 特徴量を提案する.

#### 3.1.線分検出法

HOL 特徴量を算出するために,まず画像全体から直線を検出する. 直線検出には Line Segment Detector(LSD)を用いる. LSD は計算コストが少なく,細かなパラメータ調整が不要であり,エッジの多い画像に対しても良好な直線検出結果が得られる.

# 3.2.HOL 特徴量の特徴記述

HOL 特徴量を算出するために LSD により検出した直線 成分の角度  $\theta$  と長さ l を算出し、式 (1) よりセル  $c(8\times 8$  ピクセル) 毎にヒストグラム  $V_c=\{v_c(0),v_c(1),...,v_c(n)\}$  を作成する.

$$v_n = \sum_{x=0}^{8} \sum_{y=0}^{8} l(x, y) \delta[n, f(\theta(x, y))]$$
 (1)

ここで、n は直線の角度  $\theta$  の量子化数、 $\delta[\cdot]$  はクロネッカーのデルタ関数である。最後に各セルで作成したヒストグラム  $v_c$  を複数のセルからなるブロック  $(2 \times 2$  セル) 領域ごとに正規化する.

図 3 に人と背景画像の HOG 特徴量と HOL 特徴量を示す. HOG 特徴量では, 背景と人では類似したヒストグラムが作成されるが, HOL 特徴量は異なったヒストグラムとなる. このように, HOL 特徴量は HOG 特徴量人と背景の違いを捉えることができる.

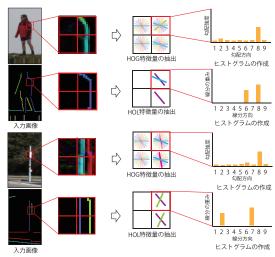


図 3: HOG 特徴量と HOL 特徴量の抽出の流れ

#### 4.評価実験

展案手法の有効性を確認するために、評価実験を行う.実験は、HOG 特徴量(HOG)、HOG 特徴量と HOL 特徴量の併用(HOG+HOL)を比較する.学習用サンプル、評価用ポジティブサンプルに INRIA Person Dataset を用い、学習用ポジティブサンプルには 2416 枚、ネガティブサンプル 12,180 枚に誤検出データ 10,000 枚を追加したものを用いる.評価には、ポジティブサンプル 1,132 枚を用い、ネガティブサンプルには誤検出データ 20,341 枚を用いる.識別結果の DET カーブを図 4(a) に、検出結果を図 4(b) に示す.図 4(a) より、MissRate が 10%のとき、HOG+HOL特徴量の FPPW は 2.5%であり、6.6%の誤検出が抑制されている.これは、HOL 特徴量が人と背景の直線成分の違いを捉えたことにより、誤検出が抑制され、精度が向上した.図 4(b) から、直線を含む物体の誤検出が抑制されていることがわかる.

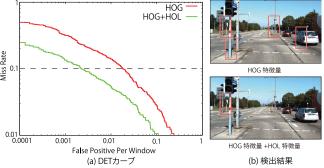


図 4: DET カーブと検出例

#### 5.おわりに

本研究では, 直線成分に基づく HOL 特徴量を提案し, 評価実験によりその有効性を確認した. 今後は特徴量抽出の時間短縮を予定している.

#### 参考文献

N.Dalal, et al, "Histograms of Oriented Gradients for Human Detection", CVPR, Vol 1, pp.886-893, 2005.

#### 多クラス識別のための行列分解と早期棄却による物体検出の高速化

EP10040 黒川 貴都

#### 1.はじめに

2 クラス識別の線形 SVM を用いた物体検出の高速化手 法として,ベクトル分解法を用いた早期判定 [1] が提案され ている. 本手法は, 重みベクトルを分解することで二値べ クトル間の演算となり高速化を実現しているが、多クラス 識別に対応していない. そこで, 本研究は行列分解法を提 案し、早期棄却による多クラス識別の高速化を目的とする.

# 2.ベクトル分解法による識別の高速化

2 クラスの線形 SVM の識別関数は、重みベクトルを  $\mathbf{w} \in \mathbb{R}^D$ , 入力特徴ベクトルを  $\mathbf{x} \in \{-1,1\}^D$  とすると,  $f(\mathbf{x}) = \text{sign}[\mathbf{w}^{T}\mathbf{x} + b]$  となる. ここで, 二値ベクトル  $\mathbf{x}$  と 実数ベクトル w 間の内積は計算コストが高い. Hare が提 案したベクトル分解法 [2] を用いると, 重みベクトルを二 値基底行列  $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \cdots, \mathbf{m}_k\} \in \{-1, 1\}^{D \times k}$  とス ケール係数ベクトル  $\mathbf{c} = \{c_1, c_2, \cdots, c_k\} \in \mathbb{R}^k$  に分解する ことができる.

$$f(\mathbf{x}) = \operatorname{sign}[\mathbf{w}^{\mathrm{T}}\mathbf{x} + b] \tag{1}$$

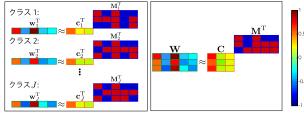
$$\approx \operatorname{sign}[\mathbf{c}^{\mathrm{T}}\mathbf{M}^{\mathrm{T}}\mathbf{x} + b] = \operatorname{sign}\left[\left(\sum_{i=1}^{k} c_{i}\mathbf{m}_{i}^{\mathrm{T}}\mathbf{x}\right) + b\right]$$

ここで k は分解の基底数を表す. これにより. 1.値ベクト ル間の内積計算に置き換えることができ, 高速な演算が期 待できる.

本研究では、多クラスを対象とするため、one-vs.-rest に よる多クラス識別器を、クラス数 J 個の 2 クラス識別器で 実現する. このとき、J個の重みベクトルからなる重み行 列  $\mathbf{W} = \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_J\}^T$  をベクトル分解法で分解する と、 $\boxtimes 1(a)$  のように J 個の二値基底行列  $\mathbf{M}$  とスケール係 数ベクトル c が算出される. 従って、多クラス識別では、特 徴ベクトル $\mathbf{x}$  と $\mathbf{c}$ M の内積の計算回数がJ回に増加する という問題がある

## 3. 実数行列の分解による多クラス識別の高速化

重みベクトル w の相関が高いとき、図 1(a) のように 各二値基底行列 M においても類似した値を持つ. この とき、各二値基底行列 M を 1 つの行列で表現すると、x と CM の内積の近似計算が 1 回に減少する. そこで, 本 研究では図 1(b) のように, 重み行列 W を二値基底行列  $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \cdots, \mathbf{m}_k\} \in \{-1, 1\}^{D \times k}$  と実数スケール 行列  $\mathbf{C} = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_k\}^{\mathrm{T}} \in \mathbb{R}^{J \times k}$  に分解する行列分解 法を提案する. さらに、内積の近似計算処理においてカス ケード構造による早期棄却を導入することで、識別の高速 化を実現する.



(a) ベクトル分解法

(b) 行列分解法

図 1: 重み行列 W の分解

#### 3.1.行列分解法

重み行列  ${f W}$  を二値基底行列  ${f M}=\{{f m}_1,{f m}_2,\cdots,{f m}_k\}$   $\in$  $\{-1,1\}^{D imes k}$  とスケール係数行列  $\mathbf{C} = \{\mathbf{c}_1,\mathbf{c}_2,\cdots,\mathbf{c}_k\}^{\mathrm{T}} \in$  $\mathbb{R}^{J imes k}$  に分解する. 式 (3) を最小化するように二値基底行 列 M とスケール係数行列 C を最適化する. この分解法は、 二値基底行列 M とスケール係数行列 C を交互に最適化す る. 提案する行列分解法の流れを Algorithm1 に示す.

$$||\mathbf{W} - \mathbf{C}\mathbf{M}^{\mathrm{T}}||_{F}^{2} \tag{3}$$

# 3.2.内積の近似計算の早期棄却による識別の高速化

行列分解において、重み行列 W と特徴ベクトル x の内 積計算による識別精度を維持するには、基底数 k の値を大

# Algorithm 1 行列分解のアルゴリズム

Require: W, L

for l = 1 to L do

 $\mathbf{M}_l$  を  $\{-1,1\}$  の乱数により初期化.

C<sub>1</sub> を実数の乱数により初期化.

repeat

 $\mathbf{M}_l$  を固定し, 式 (3) を最小化する  $\mathbf{C}_l$  について最小二乗法によ り求める

指導教授:藤吉弘亘

 $\mathbf{C}_l$  を固定し、式 (3) を最小化する  $\mathbf{M}_l$  について  $\{-1, 1\}$  を総当たりし、最適化する.  $\mathbf{until}$  式 (3) の値が収束

end for

式 (3) が最小となる **M**<sub>l</sub> と **C**<sub>l</sub> を採用

return  $\{M_l\}$ ,  $\{C_l\}$ 

きくする必要がある.しかしkの数が大きくなると $\mathbf{cm}^{\mathrm{T}}\mathbf{x}$ の演算回数が増えるため、高速な識別ができない場合があ る.そこで,図 2 のように行列分解で算出した N 個の二 値基底行列 M とスケール係数行列 C をカスケード状に並 べ、非対象クラスの早期棄却を導入する. 各段で、しきい値 T を基準に非対象クラスを棄却する.

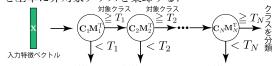


図2:カスケード構造による早期棄却

# 4.評価実験

提案手法の有効性を確認するために評価実験を行う. 多 クラスにおけるベクトル分解法と提案手法の行列分解法に よる誤差を比較する. 実験に用いる識別器は SVM(one-vs.rest) とし, 学習には 4 クラスの標識画像から抽出した 288 次元の B-HOG 特徴量を使用する. また. 17.024 枚の識別 時間を比較する. クラスあたりの基底数に対する二乗誤差 を図3に示す.

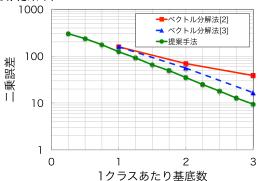


図 3:分解による誤差

図3より、1クラスあたりの基底数が3のとき、提案手法は ベクトル分解法 [3] より誤差を 43.1%小さくすることができ た. 識別時間は, 分解無しが 57.3[ms], 提案手法が 5.2[ms], 提案手法 + カスケードが 3.3[ms] となった. 提案手法は分 解前に比べ約11倍高速に、さらにカスケード構造を導入す ることで約1.6倍高速化することができた.

# 5.おわりに

提案手法は, 行列分解法とカスケード構造による早期棄 却を導入することで、高速な物体検出が可能であることを 確認した. 今後は行列分解法に適した識別の早期判定につ いて検討する予定である.

# 参考文献

- [1] 後藤等、"近似計算を導入した線形識別器の早期判定による高 速な識別"信学論 D, pp. 294-302, vol. 97, No. 2, 2014.
- [2] S. Hare et al., "Efficient Online Structured Output Learning for Keypoint-Based Object Tracking", CVPR, pp.1894-1901, 2012.
- [3] Y. Yamauchi et al., "Distance Computation Between Binary Code and Real Vector for Efficient Keypoint Matching", IPSJ Trans. on CVA, pp.124-128, vol. 5, 2013.