

量子化残差に基づく遷移ゆう度モデルを導入した識別器の提案

山内 悠嗣^{†a)} 金出 武雄^{††} 山下 隆義^{†††} 藤吉 弘亘^{†b)}

A Proposal of Classifier Introducing Transition Likelihood Model Based on Quantization Residual

Yuji YAMAUCHI^{†a)}, Takeo KANADE^{††}, Takayoshi YAMASHITA^{†††},
and Hironobu FUJIYOSHI^{†b)}

あらまし 近年、特徴量の低メモリ化やノイズに対する頑健性を得るために、実数で表現される特徴量を2値化した2値符号列が物体認識の分野で利用されている。しかし、実数で表現される特徴量を2値化することで、特徴量に含まれる多くの情報が欠落する問題がある。そこで、我々は特徴量を2値化する際に欠落する情報である“量子化残差”に着目する。本研究では、画像から観測される2値符号列が他の2値符号列へ遷移する可能性を考慮するために、量子化残差に基づき遷移を予測する遷移ゆう度モデルを識別器へ導入する。これにより、観測された2値符号列が何らかの影響により真に得たい2値符号列とは異なっても、本来得たい2値符号列への遷移を考慮した識別が可能となる。実験の結果より、提案手法である遷移予測モデルを導入した識別器は、導入する前の識別器と比較して高精度な歩行者検出が可能であることを確認した。

キーワード 量子化残差, 2値符号列, 遷移ゆう度モデル, 歩行者検出

1. ま え が き

画像から歩行者を高精度に検出する研究は、ドライバーの運転支援や安心安全のためのセキュリティシステムを実現するために精力的に取り組まれている。検出対象である歩行者は、向きや姿勢などの様々な要因により大きく見え方が変化するため、多数の歩行者の画像から極力共通な要素を抽出できる特徴量が必要である。そのため、これらの要因に対して頑健となるような特徴量が多数提案されている。その最も成功したアプローチとして、Dalalらが提案したHistograms of Oriented Gradients (HOG) 特徴量[1]がある。HOG特徴量は各画素から得られるこう配情報を、局所領域においてこう配方向ヒストグラムとして表現した高次

元な特徴量である。歩行者の形状を捉えることができるHOG特徴量は、比較的簡単な処理にもかかわらず検出性能が高く、多くの歩行者検出法[2]~[4]に用いられている。

更に、近年では人検出技術の実用化に向けてモバイル端末[5]やFPGAのようなハードウェア上[6]で動作することを目的とした取組みも盛んである。これらのデバイス上で動作させるためには、低メモリかつ高精度な検出を可能とする特徴量が必須とされている。その一つの方法として、2値化された特徴量が提案されている[7]~[10]。一般的に特徴量は実数を要素とする特徴ベクトルとして表現されるが、上記の特徴量はこれを何らかの指標により2値化した符号を複数組み合わせ一つの特徴量としている。松島らは、HOG特徴量をしきい値処理や二つのヒストグラムの大小関係により2値化することで特徴量を表現するメモリを削減した[10]。また、Muらは近傍画素との大小関係を表現するLocal Binary Pattern (LBP) [11]を特徴量として歩行者検出に利用した[7]。これらの2値化した特徴量は、ノイズに対する頑健性をもつメリットもある。

しかし、特徴量を2値に量子化することで、特徴量に含まれる多くの情報量が欠落するという問題が発

[†] 中部大学大学院, 春日井市

Graduate School of Engineering, Chubu University, 1200 Matsumoto, Kasugai-shi, 487-8501 Japan

^{††} カーネギーメロン大学, 米国

The Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213 USA.

^{†††} オムロン株式会社, 草津市

OMRON Corporation, 2-2-1 Nishikusatsu, Kusatsu-shi, 525-0035 Japan

a) E-mail: yuu@vision.cs.chubu.ac.jp

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

生する。欠落した情報に2クラスを分類する情報が含まれていると、2値化した特徴量を用いた際の識別能力は低下する。実際に[10]では、実数で表現されるHOG特徴量を単純に2値化しただけでは検出性能が低下すると報告している。

本研究では特徴量を2値化した際に欠落する情報量に着目する。この欠落する情報量は量子化前後の値の差と定義されている量子化誤差ともいえるが、本研究では量子化後の値を“0/1”とするため欠落する情報量を“量子化残差”と定義する。量子化残差とは、例えば実数で表される特徴量の値をしきい値処理により2値化した際、実数値としきい値の差を指す。量子化残差の特徴として、量子化残差が大きい場合には容易に反転しない安定した2値符号が得られる。一方、量子化残差が小さい場合には、容易に反転してしまう不安定な2値符号が得られる。2値符号を複数組み合わせることで表される2値符号列は、一つの2値符号が反転するだけで別の特徴量となる離散変数である。そのため、実数で表現している際には類似した特徴量であっても、2値化することで同一の2値符号列として観測できない場合がある。本来、高精度に歩行者を検出するためには、検出対象に共通する要素を抽出する必要があるため、このような特徴量の表現は適していない。

そこで、本研究では量子化残差に基づく2値符号列間の関係性を表現する遷移ゆ度モデルを識別器へ導入する。遷移ゆ度モデルは、観測された2値符号列 $\mathbf{x}$ が他の2値符号列 $\mathbf{x}'$ へどの程度遷移するかを表現する。本研究では遷移ゆ度モデルとして、学習サンプルから得られる2値符号列と量子化残差に基づき作成した遷移ゆ度分布を利用する。識別時には作成した遷移ゆ度分布に従い、観測される2値符号列から他の2値符号列への遷移を予測する。この2値符号列の遷移予測を識別器に導入することにより、観測された2値符号列が何らかの影響によって真に得たい2値符号列とは異なっても、本来得たい2値符号列への遷移を考慮した識別ができる。

本論文では、まず2.にて従来の歩行者検出に用いられる2値符号列と量子化残差について述べる。そして、3.にて遷移ゆ度モデルと、このモデルを導入した識別器について述べ、4.にて提案手法の有効性を確かめるために評価実験を行う。最後に5.にて本論文をまとめる。

2. HOG特徴量と2値化

本研究では、人の形状を捉えることができるHOG特徴量[1]と、これを2値化した2値符号列を用いる。ここでは、HOG特徴量と2値符号列としてBinarized HOG特徴量とRelational HOG特徴量[10]について述べる。

2.1 HOG特徴量

Dalalらによって提案されたHOG特徴量[1]は、局所領域におけるこう配方向ごとにこう配強度を累積した特徴量である。HOG特徴量は物体検出や認識などの研究に幅広く利用され、近年ではHOG特徴量を拡張した手法も多数提案されている。

HOG特徴量を求めるためには、はじめに入力画像の各画素からこう配強度とこう配方向を計算し、セルと呼ばれる領域ごとにこう配方向ヒストグラムを作成する。そして、ブロックと呼ばれる領域に含まれる各セルのヒストグラムを結合し、正規化した多次元ベクトルがHOG特徴量となる。HOG特徴量は局所領域ごとに特徴量を正規化するため、照明変動に対して頑健な特徴量となる。

2.2 Binarized HOG

Binarized HOG (B-HOG)特徴量は、2.1にて述べた要素が実数で表現されるHOG特徴量のメモリを削減するために提案された特徴量である。B-HOG特徴量は図1(a)に示すように、あるセル c における8方向のこう配方向ヒストグラム $\mathbf{V}_c$ を式(1)に示すようにしきい値処理し、2値符号を得る。

$$x_c^B(n) = \begin{cases} 1 & \text{if } v_c(n) > th \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

式(1)により8方向^(注1)全てに対して2値符号を計算し、得られた2値符号列 $\mathbf{x}_c^B = \{x_c^B(1), x_c^B(2), \dots, x_c^B(8)\}$ をB-HOG特徴量とする。

2.3 Relational HOG

B-HOG特徴量は、メモリを削減する有効な手法であるが、最適なしきい値が必要となる問題がある。しきい値を必要としない2値化手法として、図1(b)に示すように、二つのこう配方向ヒストグラムを比較することで2値化するRelational HOG (R-HOG)特徴量[10]が提案された。R-HOG特徴量は式(2)に示す

(注1): 文献[10]と同様に、1セルの特徴量を1バイト(符号なし整数型)で表現するためにこう配方向ヒストグラムの量子化数を8とする。

ように、二つのセル領域から得られる 8 方向のこう配方向ヒストグラム $\mathbf{V}_{c_1}^H, \mathbf{V}_{c_2}^H$ の大小を比較することで 2 値化する。

$$x_{c_1, c_2}^R(n) = \begin{cases} 1 & \text{if } v_{c_1}(n) > v_{c_2}(n) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

得られた 2 値符号列 $\mathbf{x}_{c_1 c_2}^R = \{x_{c_1 c_2}^R(1), x_{c_1 c_2}^R(2), \dots, x_{c_1 c_2}^R(8)\}$ を R-HOG 特徴量とする。R-HOG は、B-HOG と同じメモリ量で特徴量を表現しているが、二つのこう配方向ヒストグラムの情報を用いるため、B-HOG 特徴量よりも多くの情報量を含んでいる。そのため、R-HOG は B-HOG と比べて高精度な歩行者検出ができる。

2.4 従来法の問題点

実数で表現される特徴量を 2 値化することで、特徴量を表現するメモリ量を 1/64 に削減することができる。また、特徴量を 2 値化する際の指標を覆す要因以外には一切の影響を受けないなどのメリットもある。しかし、特徴量を表現するメモリ量を削減するに伴い、当然ながら特徴量の表現能力が大幅に低下することがある。これは、2 値化した際に欠落する情報量である“量子化残差”に識別に有効な情報が含まれているためである。現に、文献 [10] では実数として表現される HOG 特徴量をしきい値処理することにより 2 値化した場合、検出性能が低下したと報告されている。

また、従来の歩行者検出法に用いられている 2 値符号列は離散変数であるため、一つの符号が異なるだけで全く別の特徴量として表現されていた。そのため、実数で表現している際には類似した特徴量であっても、2 値化することで同一の 2 値符号列として観測できない場合がある。本来、高精度な歩行者検出を行うためには、特徴量は検出対象に共通する要素を抽出する必要があるため、このような特徴量の表現は適していない。

3. 提案手法

本研究は従来全く利用されていなかった“量子化残差”に基づき作成した遷移ゆう度モデルを識別器へ導入する。遷移ゆう度モデルは、観測された 2 値符号列が他の 2 値符号列へ遷移する可能性を表す遷移ゆう度を出力する。提案手法では、この遷移ゆう度を識別時に考慮することで、離散変数でありながらも 2 値符号列間の関係性を表現することができるため、より信頼度の高い識別結果を出力できる。

図 2 に提案手法による識別の流れを示す。提案手法は、学習サンプルから得られる 2 値符号列と量子化残差から遷移ゆう度モデルを作成する。そして、識別時には未知入力画像の 2 値符号列と遷移ゆう度モデルから得られる遷移ゆう度を識別器に入力することで、歩行者と歩行者以外を識別する。

本章では、まず 2 値符号列の量子化残差について定義し、遷移ゆう度モデルとして量子化残差に基づき作成する遷移ゆう度分布について述べる。そして、2 値符号列の遷移ゆう度モデルを導入した識別器について述べる。

3.1 2 値符号列の量子化残差

提案手法は、実数で表現される特徴量を 2 値化した際に欠落する情報量である量子化残差に着目する。本研究で用いる B-HOG 特徴量と R-HOG 特徴量の量子化残差の算出を以下に述べる。

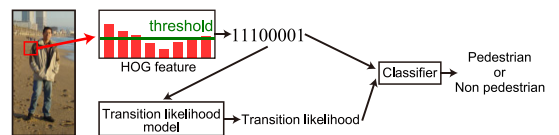


図 2 提案手法による識別の流れ
Fig. 2 Our classification process.

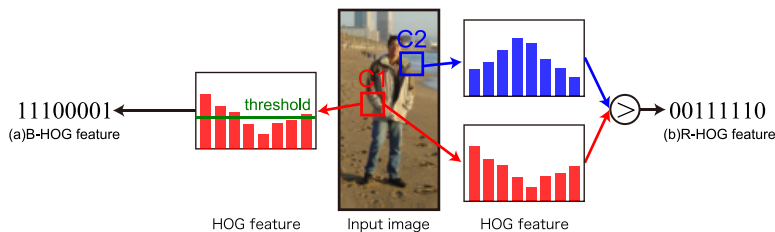


図 1 2 値符号列算出の流れ。(a) B-HOG 特徴量、(b) R-HOG 特徴量。
Fig. 1 Binary codes. (a) B-HOG feature, (b) R-HOG feature.

(1) B-HOG 特徴量

B-HOG 特徴量の量子化残差は、式 (3) に示すようにある方向 n のこの配方向ヒストグラムはこの配強度 $v_c(n)$ としきい値 th との差である。

$$q_c^B(n) = v_c(n) - th \quad (3)$$

式 (3) により 8 方向全てに対して量子化残差 $q_c^B = \{q_c^B(1), q_c^B(2), \dots, q_c^B(8)\}$ を得る。

(2) R-HOG 特徴量

R-HOG 特徴量の量子化残差は、式 (4) に示すように二つのこの配方向ヒストグラム $v_{c1}(n), v_{c2}(n)$ の差となる。

$$q_{c1,c2}^R(n) = v_{c1}(n) - v_{c2}(n) \quad (4)$$

式 (4) により 8 方向全てに対して量子化残差 $q_{c1,c2}^R = \{q_{c1,c2}^R(1), q_{c1,c2}^R(2), \dots, q_{c1,c2}^R(8)\}$ を得る。

3.2 2 値符号列の遷移ゆ度分布

2 値符号列が他の 2 値符号列にどの程度遷移する可能性があるかを表現するために、2 値符号列の遷移ゆ度分布を作成する。遷移ゆ度分布は、全学習サンプルの量子化残差に基づいて算出される 2 値符号列の遷移スコアの累積となる。図 3 に 2 値符号列の遷移スコアの算出の流れを示す。図 3(a) のように観測された 2 値符号列が他のあらゆる 2 値符号列へ遷移する場合を考える。まず、図 3(b) に示すように量子化残差に基づき、各 2 値符号の非反転度を算出する。2 値符号の非反転度は、2 値符号の遷移（ビット反転）の起こりにくさを表し、非反転度が高い場合には 2 値符号が遷移しにくいことを表す。そして、それらの 2 値符号の非反転度から 2 値符号列の遷移スコア（図 3(c)）を求める。以下に 2 値符号列の遷移スコアの算出について述べる。

3.2.1 2 値符号の遷移スコア

2 値符号列の遷移スコアを算出するために、まず 2

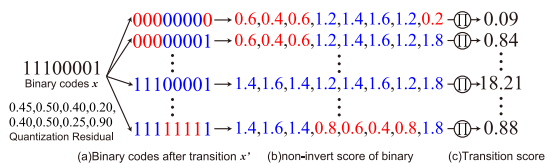


図 3 量子化残差に基づき 2 値符号列の遷移スコアを算出する流れ

Fig. 3 Flow to compute the transition score of binary codes based on quantization residuals. Blue represents that a binary code not invert. Red represents to invert of a binary code.

値符号の非反転度 z を求める。図 3 に示すように、ある学習サンプルから観測される 2 値符号列 x が他の 2 値符号列 x' へ遷移すると仮定した際、2 値符号列 x を構築する各 2 値符号 $x(n)$ の遷移の有無とその量子化残差より非反転度を決定する。本研究では 2 値符号の非反転度を算出する上で、下記の 2 点を考慮する。

(1) 2 値符号の反転の有無

(2) 量子化残差の大きさ

上記の 2 点を考慮し、四つのパターンに分けて考える。まず、2 値符号の反転がなく、量子化残差が大きい場合には非反転度を大きくし、量子化残差が小さい場合には非反転度を小さくする。そして、2 値符号の反転があり、量子化残差が大きい場合には非反転度を小さくし、量子化残差が小さい場合には非反転度を大きくする。非反転度の算出には、図 4 に示すように、本研究では 1 次関数を組み合わせた凹関数 $F(q)$ と凸関数 $\bar{F}(q)$ を用いる。サンプル i から観測された 2 値符号 $x^i(n)$ と、ある 2 値符号列 x' への遷移を考えた際の 2 値符号 $x'(n)$ の遷移の有無により、式 (5) に示すように凹関数 $F(q^i(n))$ と凸関数 $\bar{F}(q^i(n))$ を用いて 2 値符号の非反転度を算出する。

$$z(x^i(n), x'(n), q^i(n)) = \begin{cases} F(q^i(n)) & \text{if } x^i(n) = x'(n) \\ \bar{F}(q^i(n)) & \text{otherwise} \end{cases} \quad (5)$$

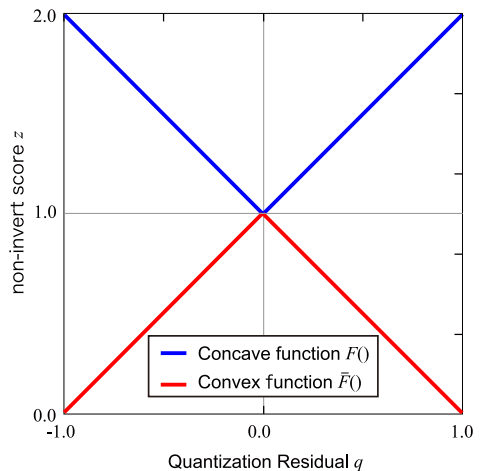


図 4 2 値符号の非反転度を求める凹関数 $F()$ と凸関数 $\bar{F}()$

Fig. 4 Non-invert score of a binary code is computed from concave function $F()$ and convex function $\bar{F}()$.

式 (5) から得られる 2 値符号の非反転度は、図 3(b) に示すように 2 値符号の遷移がない場合（青字で示される符号）には高く、遷移がある場合（赤字で示される符号）には低くなる。そして、これら 2 値符号の遷移の有無に加え、量子化残差の値に応じて遷移スコアは出力される。

3.2.2 2 値符号列の遷移スコア

得られた 2 値符号の非反転度 z から、2 値符号列 $\mathbf{x}$ が 2 値符号列 $\mathbf{x}'$ に遷移すると仮定した際の 2 値符号列の遷移スコア $e(\mathbf{x}'|\mathbf{x}^i)$ を求める。2 値符号列の遷移スコアは 2 値符号の同時性を考慮するために、式 (6) に示すように 2 値符号の非反転度の総積により求める。

$$e(\mathbf{x}'|\mathbf{x}^i, \mathbf{q}^i) = \prod_{n=1}^8 z(x^i(n), x'(n), q^i(n)) \quad (6)$$

これを全ての学習サンプル I から求め、式 (7) に示すように累積を求めることで 2 値符号列の遷移ゆう度分布 E を作成する。

$$E(\mathbf{x}'|\mathbf{x}) = \sum_{i=1}^I e(\mathbf{x}'|\mathbf{x}^i, \mathbf{q}^i) \delta[\mathbf{x}^i - \mathbf{x}'] \quad (7)$$

$\delta[\cdot]$ はクロネッカーのデルタ関数であり、 $\mathbf{x}^i - \mathbf{x}'$ が 0 の場合には 1、それ以外の場合には 0 を出力する。上記より作成する遷移ゆう度分布 E は、特徴量ごとに作成する。そのため、B-HOG 特徴量の場合では 72 個、R-HOG 特徴量の場合では 8,128 個の遷移ゆう度分布を作成する。

作成した遷移ゆう度分布の例を図 5(a) に示す。特徴量には B-HOG 特徴量を用いており、HOG 特徴量の 2 値化に使用するしきい値 th は、文献 [10] と同様に 0.09 とした。この遷移ゆう度分布は、入力 2 値符号列に似た符号列ほど高い遷移ゆう度をもつ。例えば、入力 2 値符号列が {00000000} の場合、当然ながら遷移後の 2 値符号列がハミング距離=0 となる {00000000} が最もゆう度が高い。その次には、ハミング距離=1 となる {00000001}, {10000000} に続く。一方、ハミング距離が大きいような場合は、2 値符号列の遷移は起こりにくいため、低い遷移ゆう度となることが分かる。しかしながら、遷移ゆう度は学習サンプルの量子化残差に基づいて決定されるためハミング距離とは異なる値となる。そのため、2 値符号列の遷移が生じやすい、つまり学習サンプルに量子化残差の小さいサンプルが多く含まれている場合に、提案手法はハミング距離よりも実際に起こりえる 2 値符号列間の遷移を表

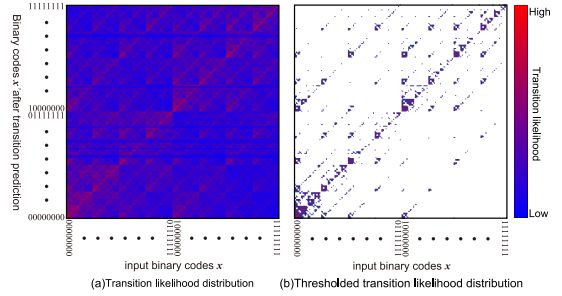


図 5 2 値符号列の遷移ゆう度分布 $E(\mathbf{x}'|\mathbf{x})$ の可視化画像
Fig. 5 Visualization of transition likelihood distribution $E(\mathbf{x}'|\mathbf{x})$ of binary codes.

現することができる。

3.3 学習サンプルの量子化残差に基づいた遷移予測

本節では、遷移ゆう度モデルとして学習サンプルから得られる量子化残差に基づき作成した遷移ゆう度モデルを導入した識別器（図 6）について述べる。

3.3.1 遷移ゆう度分布を用いた識別器

本研究では、統計的学習手法として高精度かつ高速に識別可能な Real AdaBoost [12] を用いる。Real AdaBoost のアルゴリズムにより学習される強識別器 $H(\mathbf{x})$ は、式 (8) に示すように弱識別器 $h_t(\mathbf{x})$ の線形結合により表現される。

$$H(\mathbf{x}) = \sum_{t=1}^T h_t(\mathbf{x}) \quad (8)$$

ここで、 T は組み合わせる弱識別器の個数、 t は弱識別器の番号を表す。以降では、弱識別器の番号は関係しないため、一般的弱識別器として $h(\mathbf{x})$ と表現する。弱識別器 $h(\mathbf{x})$ は、対数オッズとして式 (9) に示すようにポジティブクラスとネガティブクラスの出現確率の対数比によって決定する。

$$h(\mathbf{x}) = \frac{1}{2} \ln \frac{W_+(\mathbf{x})}{W_-(\mathbf{x})} \quad (9)$$

ここで、2 値符号列の確率密度関数 $W_{\pm}$ は、学習サンプルの重み w_i を累積することで式 (10) のように作成する。

$$W_+(\mathbf{x}) = \sum_{\mathbf{x}^i = \mathbf{x} \wedge y_i = +1} w_i, \quad W_-(\mathbf{x}) = \sum_{\mathbf{x}^i = \mathbf{x} \wedge y_i = -1} w_i \quad (10)$$

本研究では、観測された 2 値符号列 $\mathbf{x}$ が他の 2 値符

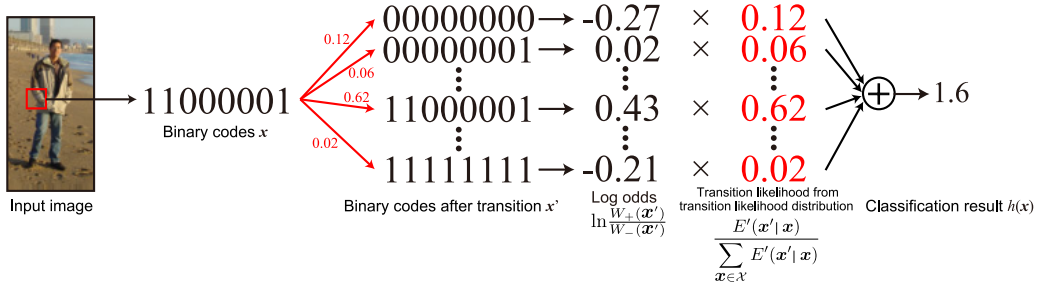


図 6 2 値符号列の遷移予測を導入した識別器
Fig. 6 Classification using transition prediction of a binary codes.

号列 x' への遷移を考慮するために、式 (9) へ遷移ゆ度モデルを導入し、式 (11) のように弱識別器 $h(x)$ を定義する。

$$h(x) \triangleq \frac{1}{2} \sum_{x' \in \mathcal{X}} \left(P(x'|x) \ln \frac{W_+(x')}{W_-(x')} \right) \quad (11)$$

ここで、 $P(x'|x)$ は 2 値符号列 x が 2 値符号列 x' に遷移する確率を表す。しかし、実際には観測する 2 値符号列は他の 2 値符号列へは遷移しないため、 $P(x'|x)$ を求めることができない。そこで、本研究では式 (12) に示すように 2 値符号列間の遷移の可能性を表す遷移ゆ度分布 $E(x'|x)$ で代用する。遷移ゆ度分布 $E(x'|x)$ は、実際には観測できない 2 値符号列の遷移確率 $P(x'|x)$ を擬似的に表現することができる。

$$P(x'|x) \approx \frac{E(x'|x)}{\sum_{x' \in \mathcal{X}} E(x'|x)} \quad (12)$$

このとき、学習サンプルにおいて 2 値符号列 x が観測されなかった場合には、2 値符号列の遷移もないと考える。この場合、式 (12) の分母及び分子は 0 となるため、 $P(x'|x) = 0$ とする。

上記に示したように、式 (11) と (12) により、観測した 2 値符号列が他の 2 値符号列に遷移する可能性を考慮した識別ができる。しかしながら、式 (11) は全ての 2 値符号列に遷移した場合を考えているため、観測した 2 値符号列と大きく異なる 2 値符号列の出現確率の対数比も弱識別器に影響を与える。その際に、確率密度関数 $W_{\pm}$ の出現確率がほぼ 0 に等しいような場合でも対数比を求めるため、極端に大きな値、若しくは小さな値を加算することになり、弱識別器 $h(x)$ の値に悪影響を及ぼす。これを解決するために、観測した 2 値符号列から遷移する可能性が低い 2 値符号列に

ついては、識別結果への影響を抑制するために式 (13) のように遷移ゆ度をしきい値処理する。

$$E'(x'|x) = \begin{cases} E(x'|x) & \text{if } E(x'|x) > \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

このしきい値 ϵ は、極端に高い値に設定すると 2 値符号列の遷移を考慮しないモデルとなる。本研究では、予備実験から $\epsilon = 0.005$ とした。しきい値処理する前の遷移ゆ度分布を図 5(a)、しきい値処理した後の遷移ゆ度分布を図 5(b) に示す。

提案する 2 値符号列の遷移ゆ度分布を用いた弱識別器 (式 (11)) は、一般的に用いられる Real Adaboost の弱識別器 (式 (9)) と比較した際に、一見すると計算コストが大きいように見える。しかし、実際には観測する 2 値符号列ごとの出力をあらかじめ計算し、ルックアップテーブルとして保持することで同じ計算コストで済む。

3.3.2 識別

ここでは、学習した識別器を用いて未知入力画像を識別する際の流れについて述べる。図 6 に示すように、まず未知入力画像から 2 値符号列を算出する。そして、式 (11) に示すように観測した 2 値符号列が他の 2 値符号へどの程度遷移しやすいかを遷移ゆ度分布から求め、対数オッズとの積の総和を弱識別器の出力とする。最終的な識別結果は、式 (8) に示すように複数の弱識別器 $h(x)$ の和によって算出する。

4. 評価実験

提案手法の有効性を評価するために、評価実験を行う。

4.1 データセット

評価実験に用いるデータセットとして INRIA person dataset [1] を用いる。INRIA person dataset は、

学習用ポジティブサンプル 2,415 枚, ネガティブ画像 1,218 枚, 評価用ポジティブサンプル 1,126 枚, ネガティブ画像 453 枚で構成されている. 学習用ポジティブ画像及び評価用ポジティブ画像は, 人領域に合わせて切り出されており, 画像サイズが 64×128 画素に正規化されている. 学習用ネガティブ画像及び評価用ネガティブ画像は, 人が存在しない画像である. 学習用ネガティブサンプルは, 1 枚の画像からランダムで 10 枚を切り出し, 合計 12,180 枚を用いる. 評価用ネガティブサンプルは, 評価用の画像中を網羅的にラスタスキャンし, 約 200 万枚を用いる.

図 7 に INRIA person dataset の一部を示す. ポジティブサンプルに含まれる画像は, 様々な人の姿勢や向きを含んでおり, 見えが大きく異なる画像で構成されており, 非常にチャレンジングなベンチマークデータセットである.

4.2 実験概要

実験では, [10] で提案された 2 値化した特徴量である B-HOG 特徴量, R-HOG 特徴量, 提案手法である学習サンプルの量子化残差に基づいた 2 値符号列の遷移予測モデルを導入した識別器を比較する.

比較には Detection Error Tradeoff (DET) カーブを用いる. DET カーブは, 横軸に False Positive Per Window (FPPW), 縦軸に miss rate を表し, 左下の原点に近いほど検出性能が高いことを示す.

4.3 実験結果

実験結果の DET カーブを図 8 に示す. まず, HOG 特徴量と B-HOG 特徴量, B-HOG 特徴量をベースとした提案手法を比較すると, 検出性能が高い順番に提案手法, HOG 特徴量, B-HOG 特徴量となった. FPPW が 10^{-3} において HOG 特徴量をしきい値処理した B-HOG 特徴量は, HOG 特徴量と比較すると検出率が約 2.9% 低下した. 次に, FPPW が 10^{-3} において B-HOG 特徴量をベースとした提案手法と HOG

特徴量, B-HOG 特徴量を比較すると, 提案手法は各手法に対して検出率が約 5.4% と約 8.4% 向上した. 以上のことから, HOG 特徴量を単純にしきい値処理により 2 値化するのみでは検出性能は低下するが, 提案手法のように識別器に 2 値符号列の予測を導入することで HOG 特徴量を上回る検出性能を得ることができた.

そして, この傾向は 2 値化方法が異なる場合においても同様の結果が得られた. FPPW が 10^{-3} において R-HOG 特徴量と R-HOG 特徴量をベースとした提案手法を比較すると, 検出性能が約 3.0% 向上していることが分かる.

以上の結果より, 提案手法は従来法よりも高精度な歩行者検出が可能であることを確認した. 従来の 2 値符号列を用いた手法では, 離散変数である 2 値符号列を互いに独立した値として扱う. そのため, 観測した 2 値符号列が何らかの影響で他の 2 値符号列として観測された場合, 識別時には観測された 2 値符号列のみを考慮するため, 識別結果が大きく変わる可能性があった. 一方, 提案手法は観測した 2 値符号列が全ての 2 値符号列にどの程度の可能性で遷移するかを予測することができるフレームワークとなっている. 2 値符号列の遷移は, 学習サンプルから得られる 2 値符号列と量子化残差から作成した遷移確度分布に基づくため, 実際に起こりやすい 2 値符号列の遷移の可能性を考慮している. そのため, 2 値符号列が何らかの影響で他の 2 値符号列として観測された場合においても, その悪影響を抑制することができる.



図 7 INRIA person dataset の画像例
Fig. 7 Examples of INRIA person dataset.

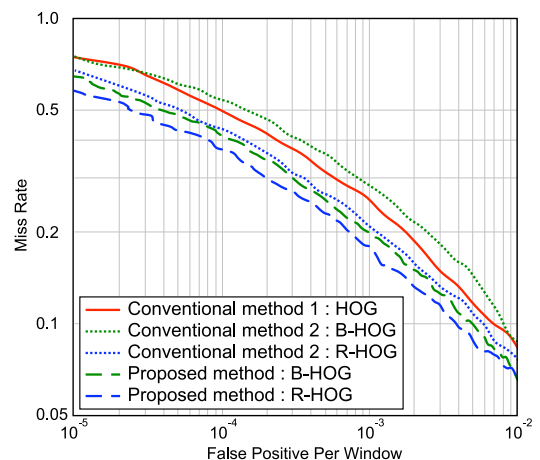


図 8 実験の結果の DET カーブ
Fig. 8 DET curves of experiment.

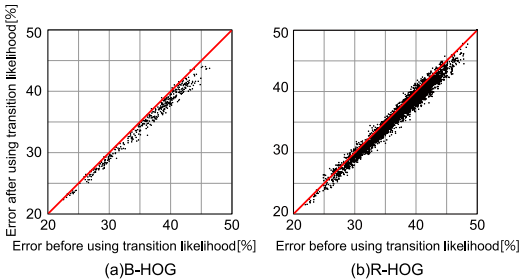


図9 式(9)と式(11)の誤識別率を二次元グラフにプロットした結果。プロットされた点は、弱識別器単体の等価誤識別率を表す。(a)はB-HOG特徴量、(b)はR-HOG特徴量をベースとしている。

Fig. 9 The graph of error computed from equation (9) and equation (11). The plotted points are represented equal error rate. (a) B-HOG feature base. (b) R-HOG feature base.

4.4 考 察

評価実験の結果より、2値符号列の遷移ゆう度モデルを導入した弱識別器は、特徴量の2値化手法にかかわらず検出性能の向上に寄与していることを確認した。ここでは、2値符号列の遷移予測を導入することで、各弱識別器の検出性能がどの程度向上したかについて考察する。

図9に一般的なReal AdaBoostの弱識別器(式(9))の誤識別率と、遷移予測を導入した弱識別器(式(11))の誤識別率を二次元グラフへプロットした結果を示す。プロットされる点は、全く同じ特徴量を使用した際の誤識別率を表す。弱識別器の性能に変化がない場合は、 $y = x$ を表す線上にプロットされる。提案手法の方が検出性能が高い場合には、 $y = x$ よりも右下にプロットされる。

図9から、提案手法の多数の弱識別器の検出性能が向上していることが分かる。B-HOG特徴量をベースとした場合では約96.2%の弱識別器、R-HOG特徴量をベースとした場合では約88.3%の弱識別器の検出性能を向上させることができた。これらの一つひとつの弱識別器は、最大で識別率が約5.0%ほどの向上である。しかしながら、多数の弱識別器を組み合わせるアンサンブル学習法であるReal AdaBoostを学習手法として利用しているため、提案手法で学習する最終識別器はより高い検出性能を得ることができる。

5. む す び

本論文では、従来のパターン認識の分野では利用されていない量子化残差を有効活用する手法を提案した。

提案手法は、画像から観測される2値符号列が他の2値符号列へ遷移する可能性を考慮するために、識別器に遷移ゆう度モデルを導入した。遷移ゆう度モデルとして、学習サンプルから得られる2値符号列と量子化残差により作成する遷移ゆう度分布を利用した。提案手法は、観測された2値符号列が他の2値符号列に遷移する可能性を考慮することができるため、信頼性の高い識別結果を出力できる。実験の結果、提案手法は従来の2値化した特徴量と同じメモリ量と計算コストを維持しながら、検出性能を向上させることができた。

提案手法の量子化残差を有効活用するアプローチは、本研究ではBoostingベースの識別器で実現しているが、マージン最大化を行うSVMや容易に多クラス問題を扱うことができるRandom Forests等の他の識別器への展開が可能だと考えられる。今後はこれらの学習手法への展開について検討する。

文 献

- [1] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," *Comput. Vis. Pattern Recognit.*, pp.886–893, 2005.
- [2] Q. Zhu, S. Avidan, M.C. Yeh, and K.T. Cheng, "Fast human detection using a cascade of histograms of oriented gradients," *Comput. Vis. Pattern Recognit.*, pp.1491–1498, 2006.
- [3] P.F. Felzenszwalb, R.B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.32, no.9, pp.1627–1645, 2009.
- [4] W.R. Schwartz, A. Kembhavi, D. Harwood, and L.S. Davis, "Human detection using partial least squares analysis," *International Conference on Computer Vision*, 2009.
- [5] B. Leibe, K. Schindler, N. Cornelis, and L.V. Gool, "Coupled object detection and tracking from static cameras and moving vehicles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.30, no.10, pp.1683–1698, 2008.
- [6] V. Nair, P.-O. Laprise, and J.J. Clark, "An FPGA-based people detection system," *EURASIP J. Advances in Signal Processing*, 2005, vol.7, pp.1047–1061, 2005.
- [7] Y.D. Mu, S.C. Yan, Y. Liu, T. Huang, and B.F. Zhou, "Discriminative local binary patterns for human detection in personal album," *Comput. Vis. Pattern Recognit.*, pp.1–8, 2008.
- [8] X. Wang, T.X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," *International Conference on Computer Vision*, 2009.
- [9] 阿部 巖, 岡谷貴之, 出口光一郎, "カーネル部分最小二乗法分析を用いた歩行者認識," *情報処理学会コンピュータ*

- タビジョンとイメージメディア研究会 (CVIM), 2010.
- [10] 松島千佳, 山内悠嗣, 山下隆義, 藤吉弘亘, “物体検出のための Relational HOG 特徴量とワイルドカードを用いたバイナリーのマスキング,” 信学論 (D), vol.J94-D, no.8, pp.1172–1182, Aug. 2011.
 - [11] T. Ojala, M. Pietikainen, and D. Harwood, “A comparative study of texture measures with classification based on featured distributions,” Pattern Recognit., vol.29, no.1, pp.51–59, 1996.
 - [12] R.E. Schapire and Y. Singer, “Improved boosting algorithms using confidence-rated predictions,” Mach. Learn., vol.37, no.3, pp.297–336, 1999.

(平成 23 年 6 月 4 日受付, 10 月 3 日再受付)



藤吉 弘亘 (正員)

1997 中部大学大学院博士後期課程了。1997～2000 米カーネギーメロン大学ロボット工学研究所 Postdoctoral Fellow. 2000 中部大学講師を経て 2004 より同大准教授。2006 米カーネギーメロン大学ロボット工学研究所客員研究員。2010 中部大学教授。工博。計算機視覚、動画像処理、パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞, 2009 年度情報処理学会論文誌コンピュータビジョンとイメージメディア優秀論文賞, 2009 年度山下記念研究賞, 情報処理学会, 電気学会, IEEE 各会員。



山内 悠嗣 (学生員)

2009 中部大学大学院工学研究科情報工学専攻博士前期課程了。現在同大学院博士後期課程に在籍。2010 独立行政法人日本学術振興会特別研究員 DC。コンピュータビジョン, パターン認識の研究に従事。



金出 武雄 (正員)

1973 京大工博, 同大学・情報・助手。1980 カーネギーメロン大計算機科学科・ロボット研究所高等研究員, 1985 同教授, 1992～2001 同大学ロボット研究所所長, 2006 より同生活の質工学研究センターセンター長。産業技術総合研究所デジタルヒューマン研究センター兼任。National Academy of Engineering 外国特別会員, IEEE, ACM, AAAI, ロボット学会各フェロー, C&C 賞, Joseph Engilberger 賞, FIT2004 船井業績賞, JARA 賞, 人工知能学会業績賞, Marr 賞など各受賞。



山下 隆義

2002 奈良先端科学技術大学院大学情報科学研究科修士課程了。同年オムロン (株) 入社。2011 中部大学大学院博士後期課程了。顔及び人画像センシング技術の研究に従事。IEEE-CS, 情報処理学会各会員。