

量子化残差に基づく特徴量の遷移尤度モデルを導入した識別器の提案

山内 悠嗣[†] 金出 武雄^{††} 山下 隆義^{†††} 藤吉 弘亘[†]

[†] 中部大学 〒 487-8501 愛知県春日井市松本町 1200

^{††} カーネギーメロン大学 〒 15213 米国ペンシルバニア州ピッツバーグ

^{†††} オムロン株式会社 〒 525-0035 滋賀県草津市西草津 2 丁目 2-1

E-mail: [†]tyuu@vision.cs.chubu.ac.jp, ^{††}tk@cs.cmu.edu, ^{†††}takayosi@omm.ncl.omron.co.jp,

^{††††}hf@cs.chubu.ac.jp

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、量子化残差を用いた 2 つのアイデアに基づき遷移を予測する遷移尤度モデルを識別器へ導入する。1 つ目は、学習サンプルから観測される 2 値符号列と量子化残差に基づいて作成される遷移尤度分布に従い、2 値符号列が他の 2 値符号列へ遷移する可能性を予測する。2 つ目は、識別時に観測される量子化残差から他の 2 値符号列への遷移を直接的に予測する。これにより、観測された 2 値符号列が他の 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 値符号列までも考慮した識別が可能となる。実験の結果より、提案手法である量子化残差に基づく遷移予測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, 遷移尤度モデル, 歩行者検出

1. はじめに

画像から歩行者を高精度に検出する研究は、近年盛んに行われており、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することが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勾配情報を、局所領域において勾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関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。

さらに、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り組みも盛んである。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 値にした符号を複数組み合わせることで 1 つの特徴量としている。松島らは 1 次元ヒストグラムで表現される HOG 特徴量を、閾値処理や 2 つのヒストグラムの大小関係により 2 値化した 2 値符号列として扱うことで、特徴量に使用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持つメリットもある。

しかし、特徴量を 2 値に量子化することは、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生する。欠落した情報量に 2 クラスを分類する高い表現力が含まれていると、2 値化した特徴量を用いた際の識別能力は低下することになる。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されている。

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言えるが、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、例えば実数で表わされる特徴量の値を閾値処理により 2 値化した際、実数値と閾値の差のことを指す。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は反転することのない安定した 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

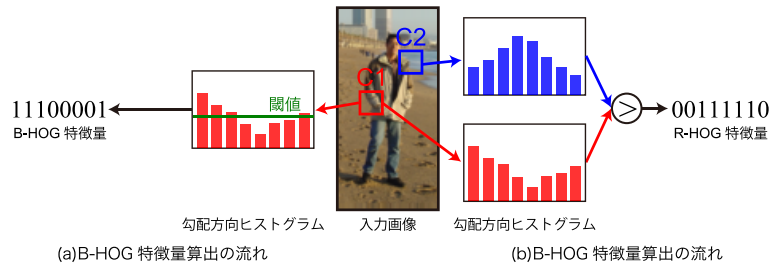


図 1 2 値符号列算出の流れ . (a)B-HOG 特徴量 , (b)R-HOG 特徴量 .

が得られる . 2 値符号を複数組み合わせることで表わされる 2 値符号列は , 1 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尤度モデルを識別器へ導入する . 遷移尤度モデルは , 観測された 2 値符号列 x が他の 2 値符号列 x' へどの程度遷移するかを表現する . 我々は , 量子化残差の特徴を生かした 2 つのアイデアにより遷移尤度モデルを作成する . 1 つ目のアイデアは , 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いて作成される遷移尤度分布に従い , 観測される 2 値符号列が他の 2 値符号列への遷移を予測する . 2 つ目のアイデアは , 識別時に観測される量子化残差から他の 2 値符号列へ遷移する可能性を予測する . 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 , その得たい 2 値符号列までを考慮した識別が可能となる .

本稿では , まず 2 章にて従来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 章にて遷移尤度モデルと , このモデルを導入した識別器について述べ , 4 章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 章にてまとめる .

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 Histograms of Oriented Gradients(HOG) 特徴量は , 局所領域における勾配方向毎に勾配強度を累積した特徴量 [1] 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勾配強度と勾配方向を計算し , セルと呼ばれる領域毎に勾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 特徴量は局所領域毎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 方向の勾配方向ヒストグラム V_c を式 (1) に示すように閾値処理し , 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 $x_c^B = \{x_c^B(1), x_c^B(2), \dots, x_c^B(8)\}$ 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 , 最適な閾値が必要となる問題がある . 閾値を必要としない 2 値化手法として , 図 1(b) に示すように , 2 つの勾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) に示すように , 2 つのセル領域から得られる 8 方向の勾配方向ヒストグラム $V_{c_1}^H, V_{c_2}^H$ の大きさを比較することで 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 $x_{c_1 c_2}^R = \{x_{c_1 c_2}^R(1), x_{c_1 c_2}^R(2), \dots, x_{c_1 c_2}^R(8)\}$ 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 , 2 つの勾配方向ヒストグラムの情報を用いているため , 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可能である .

(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) で表現するために勾配方向ヒストグラムの量子化数を 8 とする .

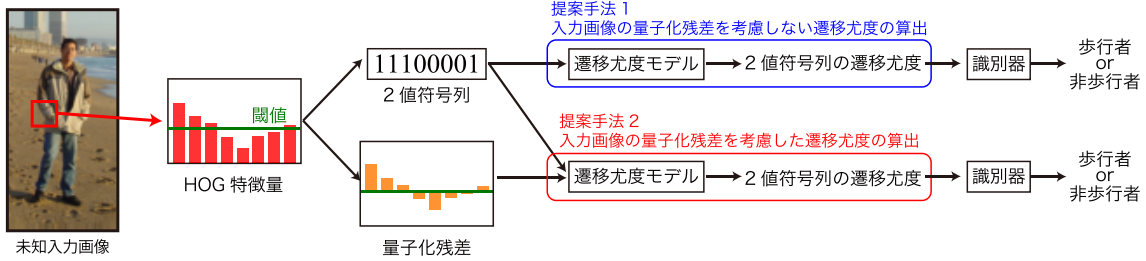


図 2 提案手法による遷移尤度を用いた識別の流れ。

2.4 従来法の問題点

実数で表現される特徴量を 2 値化することは、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がある。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からだと考えられる。従来法の方法では、Boosting の学習アルゴリズムに基づいて特徴選択をしているため、陰に量子化残差が小さくなるように 2 値符号列を選択している可能性があるが、量子化残差に識別に有効な情報が含まれていれば、検出性能の低下は避けられない。

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、1 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

本研究は従来法では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尤度モデルを識別器へ導入する。遷移尤度モデルは、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わす遷移尤度を出力する。提案手法ではこの遷移尤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することが可能となる。

図 2 に提案手法による識別の流れを示す。図 2 に示すように、提案手法は 2 つの方法により遷移尤度を算出する。1 つ目は遷移尤度モデルとして、学習サンプルの量子化残差に基づき作成した遷移尤度分布を用いて遷移尤度を算出する。2 つ目は遷移尤度モデルとして、遷移尤度分布と識別時に観測した量子化残差に基づき、2 値符号列の遷移尤度を算出する。

本章では、まず 2 値符号列の量子化残差について定義



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。青字は 2 値符号の遷移がないこと、赤字は 2 値符号の遷移があることを表わしている。

し、量子化残差に基づいた遷移尤度分布の作成について述べる。そして、2 つの提案手法による 2 値符号列の遷移尤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。

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 の勾配方向ヒストグラムの勾配強度 $v_c(n)$ と閾値 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。

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 2 つの勾配方向ヒストグラム $v_{c_1}(n), v_{c_2}(n)$ の差となる。

$$q_{c_1, c_2}^R(n) = v_{c_1}(n) - v_{c_2}(n) \quad (4)$$

式 (4) により 8 方向全てに対して量子化残差 $q_{c_1, c_2}^R = \{q_{c_1, c_2}^R(1), q_{c_1, c_2}^R(2), \dots, q_{c_1, c_2}^R(8)\}$ を得る。

3.2 2 値符号列の遷移尤度分布

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 値符号列の遷移尤度分布を作成する。遷移尤度分布は、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

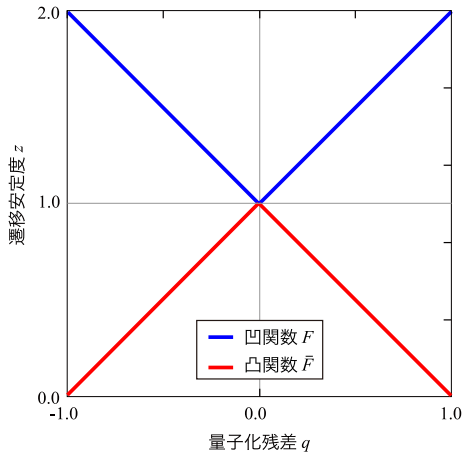


図4 2値符号の非反転度を求める凹関数 $F()$ と凸関数 $\bar{F}()$.

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わし，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わす．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，4つのパターンに分けて考える．まず，2値符号の反転が無く，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．

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

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

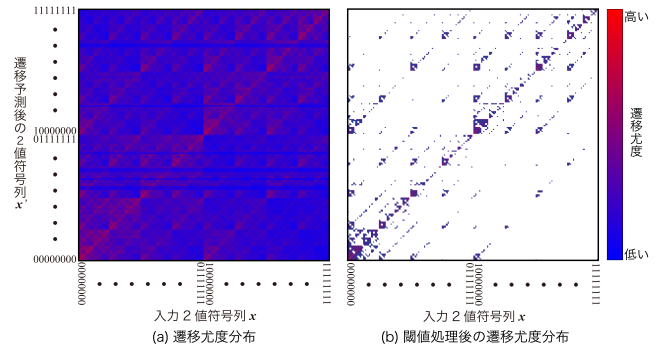


図5 2値符号列の遷移尤度分布 $E(x'; x)$ の可視化画像．

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 x が2値符号列 x' に遷移すると仮定した際の2値符号列の遷移スコア $e(x'; x^i)$ 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．

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

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値符号列の遷移尤度分布 E を作成する．

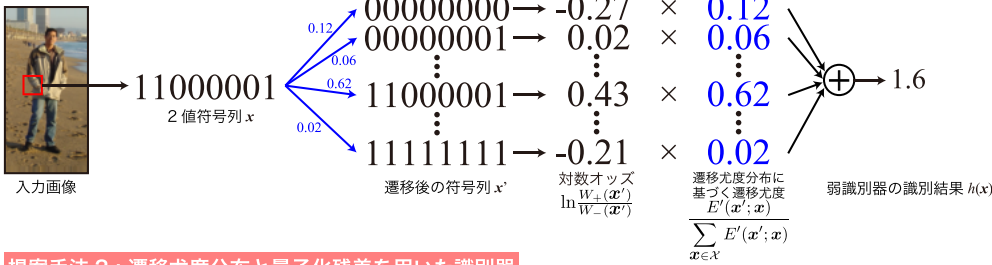
$$E(x'; x) = \sum_{i=1}^I \sum_{x' \in \mathcal{X}} e(x'; x, q^i) \delta[x^i - x'] \quad (7)$$

ここで， \mathcal{X} は2値符号列の全体集合を表わす． $\delta[\cdot]$ はクロネッカーのデルタ関数であり， $x^i - x'$ が0の場合には1，それ以外の場合には0を出力する．

作成した遷移尤度分布の例を図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値化に使用する閾値 th は，文献[10]と同様に0.09とした．この遷移尤度分布は，入力2値符号列に似た符号列ほど高い遷移尤度を持つ．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 $\{00000000\}$ が最も尤度が高い．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，低い遷移尤度となることがわかる．

このように，基本的に遷移尤度分布の尤度は，離散変数を扱う距離の指標として利用されるハミング距離に似た値となるが，学習サンプルの量子化残差に基づき遷移尤度を決定する．そのため，評価用サンプルに似た画像が学習サンプルに含まれているならば，提案手法はハミング距離と比べて実際に起こりえる2値符号列間の遷移を表現することができる．

提案手法 1 : 遷移尤度分布を用いた識別器



提案手法 2 : 遷移尤度分布と量子化残差を用いた識別器

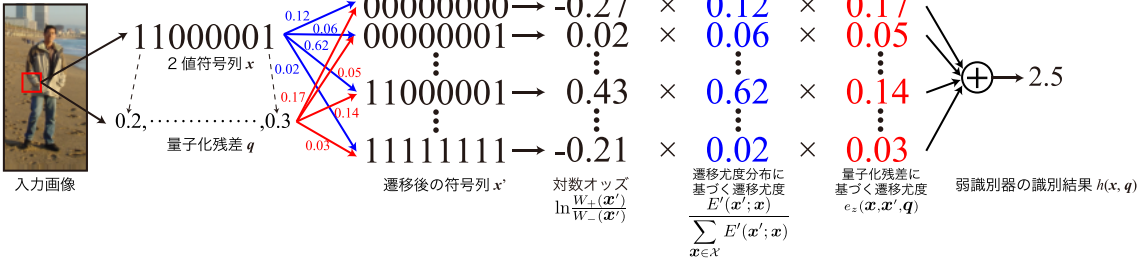


図 6 2 値符号列の遷移予測を導入した識別器 .

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、遷移尤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尤度分布を導入した識別器 (図 6(提案手法 1)) について述べる .

3.3.1 遷移尤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 $H(x)$ は、式 (8) に示すように弱識別器 $h(x)$ の線形結合により表現される .

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

弱識別器 $h(x)$ 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 .

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

ここで、 W_{\pm} は 2 値符号列 x の確率密度関数を表わし、式 (10) により求める . 確率密度関数は、学習サンプルの重み w_i を累積することで作成する .

$$W_+(x) = \sum_{i: x' \in \mathcal{X} \wedge y_i = +1} w_i, W_-(x) = \sum_{i: x' \in \mathcal{X} \wedge y_i = -1} w_i \quad (10)$$

式 (9) は、画像から観測された 2 値符号列 x に基づき、対数オッズを出力する . 本研究では観測された 2 値符号列 x が他のあらゆる 2 値符号列 x' にどの程度の確率で遷移するかを予測し、各クラスの出現確率の対数比である対数オッズに対して重み付けするために式 (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わす . 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わす遷移尤度分布 $E(x'; x)$ で代用する . 遷移尤度分布 $E(x'; x)$ は量子化残差に基づき作成しているため、実際には観測できない 2 値符号列の遷移を擬似的に表現することができる .

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

上記に示したように、式 (11)、(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、図 5(a) からわかるように、2 値符号列が遷移する可能性がある 2 値符号列は多くの場合において限られている . 遷移する可能性が低い場合には、識別結果に与える影響は少ない . そこで、式 (13) に示すように、遷移尤度が閾値 ϵ 以上の場合のみを考慮する .

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

閾値 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 . 閾値処理をした遷移尤度分布を図 5(b) に示す .

提案する 2 値符号列の遷移尤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 値符号列毎の出力を予め計算し，ルックアップテーブルとして保持することで一切変わらない計算コストで済む．

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(提案手法 1) 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 値符号列への程度遷移しやすいかを遷移尤度分布から求め，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の和によって算出する．

3.4 評価サンプルの量子化残差に基づいた遷移予測

3.3 では，学習サンプルから得られる量子化残差に基づいて作成した遷移尤度分布を遷移尤度モデルとして識別器へ導入した．この識別器は，遷移尤度分布のみに従い遷移尤度を求めるため，学習サンプルから観測される 2 値符号列の遷移の傾向に基づいて予測する．そのため，未知入力画像の量子化残差が小さいために 2 値符号列の遷移が容易に起こるような場合でも，これを考慮した遷移の予測はできない．そこで，本節では図 6(提案手法 2) に示すように，遷移尤度モデルへ新たに識別時に観測した量子化残差に基づいて算出する 2 値符号列の遷移尤度の追加について述べる．識別時に観測された未知入力画像の量子化残差から得られる遷移尤度をモデルに追加することで，より直接的に 2 値符号列の遷移を予測することができるため，検出器の高精度化が期待できる．

3.4.1 量子化残差に基づく遷移尤度を追加した遷移尤度モデル

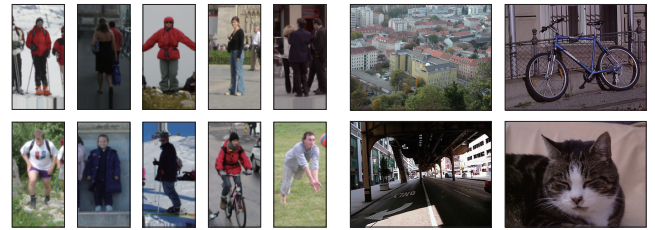
2 値符号列の遷移の可能性を考慮するために遷移尤度モデルとして導入した識別器を式 (11) に示した．この遷移尤度モデルに対して，新たに量子化残差を考慮するように定義した識別器を式 (14) に示す．

$$h(\mathbf{x}, \mathbf{q}) \triangleq \frac{1}{2} \sum_{\mathbf{x}' \in \mathcal{X}} \left(P(\mathbf{x}' | \mathbf{x}, \mathbf{q}) \ln \frac{W_+(\mathbf{x}')}{W_-(\mathbf{x}')} \right) \quad (14)$$

式 (12) と同様に， $P(\mathbf{x}' | \mathbf{x}, \mathbf{q})$ は実際に求めることができないため，式 (15) で代用する．

$$P(\mathbf{x}' | \mathbf{x}, \mathbf{q}) \approx \frac{E(\mathbf{x}'; \mathbf{x}) e_z(\mathbf{x}, \mathbf{x}', \mathbf{q})}{\sum_{\mathbf{x}' \in \mathcal{X}} E(\mathbf{x}'; \mathbf{x}) e_z(\mathbf{x}, \mathbf{x}', \mathbf{q})} \quad (15)$$

ここで， e_z は未知入力画像から観測した量子化残差に基づいて算出した 2 値符号列の遷移スコアを表わし，式 (16) より求める．式 (16) は，式 (6) から得られる 2 値符号列の非反転度 z の量子化ヒストグラム W_z の出現確率の総積により求める．量子化ヒストグラム W_z の量子化数は，実験より 32 に決定した．



(a) 人画像 (b) 背景画像

図 7 INRIA person dataset の画像例

$$e_z(\mathbf{x}, \mathbf{x}', \mathbf{q}) = \prod_{n=1}^8 W_z(z(\mathbf{x}(n), \mathbf{x}'(n), \mathbf{q}(n))) \quad (16)$$

3.4.2 識別

遷移尤度分布のみを用いた識別器は，未知入力画像から観測した 2 値符号列のみを弱識別器へ入力した．提案手法 2 では，図 6(提案手法 2) に示すように未知入力画像から観測した 2 値符号列と量子化残差の両方を弱識別器 (14) に入力し，識別結果を出力する．そして，最終的に式 (8) に示すように複数の弱識別器の和を識別結果とする．

4. 評価実験

2 つの提案手法の有効性を評価するために，評価実験を行う．1 つ目の実験は，提案手法 1 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の有効性を評価する．2 つ目の実験は，提案手法 2 である評価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の有効性を評価する．

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．

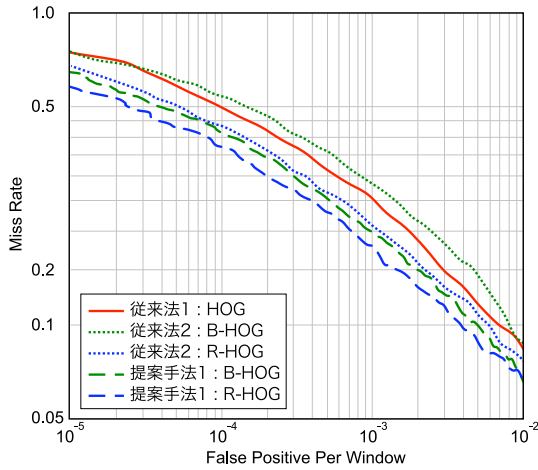


図 8 実験 1 の結果の DET カーブ。

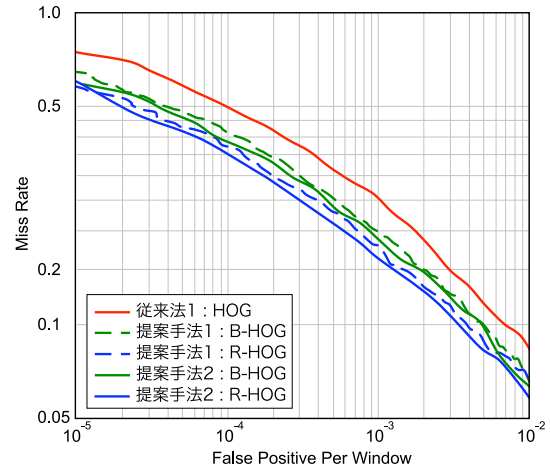


図 10 実験 2 の結果の DET カーブ。

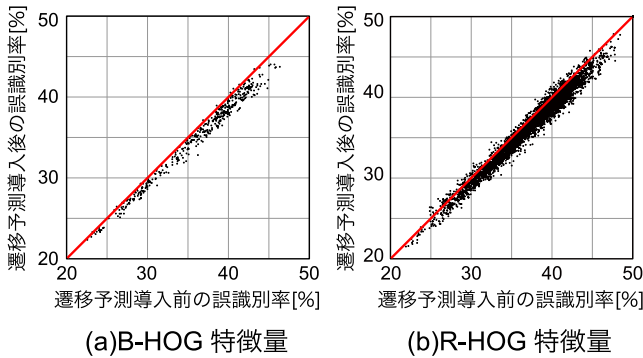


図 9 式 (9) と式 (11) の誤識別率を 2 次元グラフにプロットした結果。プロットされた点は、弱識別器単体の誤識別率を表わす。グラフ中の赤い線は、 $y = x$ を表わしている。(a) は B-HOG 特徴量，(b) は R-HOG 特徴量をベースとしている。

4.2 実験概要

1 つ目の実験では [10] で提案された 2 値化した特徴量である B-HOG, R-HOG 特徴量と、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。2 つ目の実験では、評価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の有効性を評価する。

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 を表わし、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 1

実験 1 では、遷移尤度モデルとして遷移尤度分布を識別器へ導入した提案手法 1 の有効性を確認するために、従来法である HOG 特徴量 [1], B-HOG 特徴量と R-HOG 特徴量 [10] と比較する。

4.3.1 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 特徴量を閾値処理した 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 特徴量を単純に閾値処理により 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わかる。

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。従来の歩行者検出フレームワークでは、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。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、その影響を抑制することができる。また、提案手法は従来の 2 値化した特徴量と一切変わらない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。

4.3.2 考察

評価実験の結果より、2 値符号列の遷移尤度モデルを導入した弱識別器は、特徴量の 2 値化手法に関わらず検

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 2 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、まったく同じ特徴量を使用した際の誤識別率を表わす。弱識別器の性能に変化が無い場合は、 $y = x$ を表わす赤い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、赤い線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わかる。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。これらの一つ一つの弱識別器は、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。

4.4 実験 2

実験 2 では、提案手法 1 の遷移尤度モデルに、識別時に観測した量子化残差に基づいて算出する遷移尤度を加えた提案手法 2 の有効性を確認する。

4.4.1 実験結果

実験結果の DET カーブを図 10 に示す。実験の結果より、B-HOG 特徴量と R-HOG 特徴量の両特徴量をベースとしても、提案手法 1 よりも提案手法 2 の方が高い検出性能であることがわかる。B-HOG 特徴量ベースでは、FPPW が 10^{-3} において提案手法 1 よりも提案手法 2 の検出率が 0.9% が向上した。同様に、R-HOG 特徴量ベースでは、FPPW が 10^{-3} において提案手法 1 よりも提案手法 2 の検出率が 1.4% が向上した。

提案手法 1 では、識別時には未知入力画像の量子化残差は観測しない。そのため、未知入力画像から観測した量子化残差が小さい場合には、本来ならば 2 値符号列の遷移が起こりやすいが、この量子化残差の大きさに基づいて直接的に 2 値符号列の遷移を予測しないため遷移尤度の信頼性は低い。一方、提案手法 2 では識別時に観測した量子化残差に基づいて直接的に 2 値符号列の遷移を予測することが可能となり、信頼性の高い遷移尤度を出力することができる。

5. おわりに

本稿では、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、2 つのアイデアにより 2 値符号列の遷移を予測

し、遷移尤度モデルとして識別器へ導入する。1 つ目は、学習サンプルより得られる量子化残差から 2 値符号列の遷移を予測する遷移尤度分布を作成し、識別時に観測される 2 値符号列に対応した遷移尤度を用いる。2 つ目は、1 つ目の遷移尤度モデルに加え、識別時に未知入力画像から観測される量子化残差に基づいて算出する遷移尤度を用いる。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することが可能である。

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", *Computer Vision and Pattern Recognition*, vol.1, 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", *Computer Vision and Pattern Recognition*, pp.1491–1498, 2006.
- [3] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object Detection with Discriminatively Trained Part Based Models", *Pattern Analysis and Machine Intelligence*, 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", *Pattern Analysis and Machine Intelligence*, 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 Journal on Advances in Signal Processing*, vol.2005, no.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", *Computer Vision and Pattern Recognition*, 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 敏, 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 弘巨, "Relational Binarized HOG 特徴量と Real AdaBoost によるバイナリ選択を用いた物体検出", 画像の認識・理解シンポジウム (MIRU), 2010.
- [11] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions", *Pattern Recognition*, vol.29, no.1, pp.51–59, 1996.
- [12] R. E. Schapire, and Y. Singer, "Improved boosting algorithms using confidence-rated predictions", *Machine Learning*, vol.37, no.3, pp.297–336, 1999.