

2値線画像解析のための2値方向パターンヒストグラム 特徴量HBOP

Histograms of Binary Orientation Pattern Features for
Binary Line Image Analysis

石井 大祐 渡辺 裕
Daisuke ISHII Hiroshi WATANABE

あらまし 本論文では、2値細線画像の取り扱いに優れた、新しい画像特徴量Histograms of Binary Orientation Pattern (HBOP)を提案する。近年マンガは日本文化として国内外で広く知られており、この内容解析のために、一般物体認識技術の利用が期待される。一般物体認識に利用される画像特徴量は、その主たる処理対象を自然画像としている。一方でマンガは主に2値線画により描かれるが、その線数は多く複雑である。このような複雑な線画像の構造を解析するための手法として2値細線化処理がある。しかし2値細線画像の特性は多値画像と大きく異なるものである。従来の多値画像向け画像特徴量の一手法であるHOGは2値細線画像からは線の方向を正しく得ることができないものの、グレースケールの線画像に対して一定の解析が可能である。本論文で提案するHBOPは局所的な2値細線の方向ヒストグラムを用いる事で、2値細線画像の解析を可能とする。2値細線画像に対する評価実験から、HOGと比較して学習速度の向上および検出精度の向上を確認した。また、本研究で想定する処理対象であるマンガ画像に対して、2値細線を扱う場合にHOGと比較して高い識別性能を確認した。今回実施した実験において、HOGでは2値細線化を行わない場合に最も良い識別精度が得られた。一方提案手法は、HOGにおける最良の結果と比較して、0.052のTrue Negativeの低下があるものの、0.044高いTrue Positiveが得られた。

1 はじめに

近年コンピュータビジョンやITSなどの幅広い分野において画像解析技術を利用した物体検出、認識手法が提案されている[1, 2]。物体検出は任意の対象物と同一の物体を検出する特定物体認識と、検出対象カテゴリに属するものを画像の中から発見もしくは選別する一般物体認識に分けられる。特に、一般物体認識の一処理である人物顔検出は、今日ではほぼすべてのデジタルカメラに搭載されている。このように画像解析技術は今日の技術発展の中で大きな役割を担っている。

これまでアプリケーションや、その目的および検出対象に応じてさまざまな画像特徴量が提案されてきた。一般物体認識に利用される画像特徴量として、Haar-like [3], HOG [4-6], Edgelet [7]などが提案

されている。また、検出処理ウィンドウ内において、複数のローカル領域から得られる画像特徴量の組み合わせを学習することで、より高精度な識別器を構成する手法が提案されている[8-11]。

ところで、画像の一形態として線画像がある。線画像の例としては、画像からのエッジ検出処理結果、文字、マンガをはじめとするイラスト等がある。エッジ検出処理は、特徴抽出処理の一つとして、古くから画像解析分野において利用されている。また、マンガは近年国内外に広く注目されるコンテンツとなっている。マンガは主に2値線画により描かれるが、その線数は多く複雑である。

一方で、近年提案された画像特徴量は主に輝度情報に基づく特徴を扱っており、主に写真や動画像などの自然画像を主眼に検討がなされてきた。自然画像では多大なバリエーションの濃淡変化が存在しこの情報を

もとに識別処理が実現されている。

複雑な線画像の構造を解析するための手法として2値細線化処理がある。しかし、2値細線画像の画像特性は通常の多値画像と大きく異なり、従来の画像特徴量ではその解析が困難である。

我々はマンガ画像解析を主目的として2値細線画像解析性能の向上を試みる。マンガ画像は多数の線により描かれる白黒の画像である。この画像の特徴的な点として、一般的な人物検出対象画像等と比較して濃淡情報が白黒2値に基づく点、多数の線からなるエッジが存在する点が挙げられる。人物認識や人物認証を、マンガ画像でも実施可能とするためには、線画像の構造を認識することが有効であると考えられる。また、対象カテゴリ内の変動を許容し、形状変化に対応する特徴記述が必要である。

本論文ではマンガ等を初めとする複雑な線画像において、その構造を解析するために、2値細線画像に適用可能な画像特徴量Histograms of Binary Orientation Pattern (HBOP) を提案する。HBOPは、形状の変化を許容しつつ、画像のエッジ特性を捉えることができ、線画像に対して一定の識別性能をもつHOG特徴量を基に、局所的な2値直線成分の方向分布を画素の並びからパターンに分類しヒストグラム化することで、2値細線画像の特性をとらえる。

本論文の構成を以下に示す。2章では、画像解析の従来手法について、その特性および2値細線画像に対する適性を述べる。3章では、本論文で提案するHBOPの詳細を示す。4章では、HBOP特徴量とHOG特徴量との比較実験結果を示し考察を行う。5章は本論文のまとめである。

2 画像解析の従来手法

2.1 近年の画像解析技術の概要

画像解析の分野では、近年の機械学習の発展とともに目的に応じた多様な画像特徴量が提案されている。一般物体検出に用いられる画像特徴量としてHaar-like [3], HOG [4][5] とその派生であるEHOG [6], Edgelet [7] などをはじめとする様々な手法がある。さらに、これら特徴量を用い、その共起性を捉える手法等が提案されている [8-11]。

Haar-like特徴量は局所領域間の平均輝度差を求めるもので、人物顔検出等において高速かつ高精度な検出が可能である。HOG特徴量は局所的なエッジ方向の分布に着目した特徴量であり、セルと呼ばれる小領域の位置情報と共に、その中に含まれるエッジ成分を取得できるため、人物検出や車検出等に有効な特徴量である。HOG特徴量の派生版であるEHOG特徴量は、HOG特徴量に対して方向成分の単純化およびセルの除去により特徴量の次元数を削減している。Edgelet特徴量はエッジの出現位置とその形状から検出対象の特定を行う。Edgelet特徴を用いることで、

人物検出において、人間の両肩に起因するエッジ等を捉えることが可能である。

上記画像特徴量に加え、複数箇所の画像特徴量について、その共起性や、出現位置を組み合わせることで、より高精度な検出を行う手法も提案されている。Joint-HOG [8], CoHOG [9], Deformable Parts Model [10] では、検出領域内において、複数箇所の小領域に対しHOG特徴量を用いた検出結果を組み合わせ、再度学習を行うことにより、従来よりも高い精度をもつ強識別器を構成する。Shapelet [11] は複数の小領域内におけるエッジ方向の共起性をとらえるものである。

以下ではこれら一般物体認識向け画像特徴量の概要と2値細線画像および本論文の処理ターゲットである2値細線化したマンガに対する適応性について述べる。特に、Joint-HOG, CoHOG, Deformable Parts Modelの共起特徴量にも用いられ、また本論文で提案するHBOP特徴量の基礎となるHOG特徴量およびHaar-like特徴量のIntegral-Image計算方法については詳細を示す。

2.2 Haar-like 特徴量

ViolaとJonesによって提案されたHaar-like特徴量を利用した顔画像検出 [3] は、顔画像検出分野において一般的に使用される手法である。Haar-like特徴量は画像内局所領域の輝度の分布特徴を捉える。特に人物画像では、人物の瞳部分は暗く、周囲の肌部分では明るい輝度が存在する。Haar-like特徴により人間の顔に特有な輝度分布を捉えることができる。Haar-like特徴量は2つの局所領域に含まれる合計輝度の差から特徴量を求める。この局所領域の形状は検出対象に合わせ多数のバリエーションをもつ。顔検出では瞳、鼻、口などに対してそれぞれ異なる形状のHaar-like特徴量を用いた検出処理を行い、その結果をカスケード接続することで、高い精度で人物の顔検出が可能である。

Haar-like特徴量の処理過程では、多数の局所領域内の合計輝度を算出する必要があるが、Integral Image

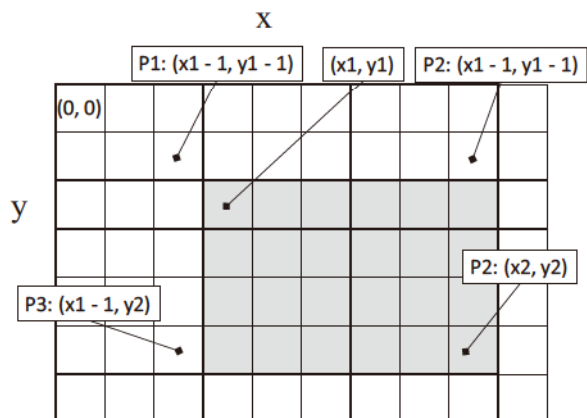


図1 Integral Imageにおける対象領域内参照箇所
Fig. 1 Reference points of target area in Integral Image

を利用することで計算の高速化が図られている。Integral Imageの各画素の値は、画像の起点(0, 0)から当該画素(x, y)を対角点とする矩形領域内に含まれる輝度の合計値として算出される。画像の左上を起点とした場合、画像の右下に向けてIntegral Imageの値は徐々に大きくなる。矩形領域内の輝度合計値の計算には、局所領域内に含まれる画素数分の参照が必要であったが、Integral Imageを用いることで、4点の参照および計算のみで実現可能となる。図1に示される(x1, y1)と(x2, y2)を対角とする矩形領域内の輝度合計値を求める場合、Integral Image上のP1からP4で示された4点からP4-P2-P3+P1を計算することでその値が得られる。Integral Imageを利用することで多数の矩形領域内の輝度合計値を得る場合に計算コストを削減可能である。

Haar-like特徴量を用いた顔検出処理は高速かつ頑健であるため、非常に多く利用されている。2値細線画像を対象とした処理では、Haar-like特徴量の計算方法から、局所領域内に存在する線の密度についてのみ検知可能である。しかし円形や四角形などの構造および線の方向等については検知不可能である。

2.3 HOG特徴量

2.3.1 HOGの概要

HOG [4] は画像内の輝度勾配方向の分布に着目した特徴量である。人物輪郭などのエッジを捉えることができ、学習により、同一の検出対象がバリエーションを持つ場合にも対応可能である。

HOG特徴量の計算では、各画素について、上下方向の輝度差 V_d と左右方向の輝度差 H_d から下記の式により勾配ベクトルの大きさと m その向き θ を計算する。

$$m(x,y) = \sqrt{V_d(x,y)^2 + H_d(x,y)^2} \quad (1)$$

$$\theta(x,y) = \tan^{-1} \frac{V_d(x,y)}{H_d(x,y)} \quad (2)$$

このようにして求めた係数について、画像を複数の小さなセルに分け、そのセル内においてヒストグラム化する。ヒストグラム化にあたり、方向成分 θ については、 $[0^\circ, 180^\circ)$ の範囲に射影し、その中で任意の N 方向に量子化する。この方向に基づいて、 $m(x, y)$ の値を投票しヒストグラムを生成する。また、複数セルをまとめてブロックとし、ブロック単位にて正規化を行う。

HOG特徴量の次元数は下記の式により求める。

$$\begin{aligned} \text{HOG次元数} &= \text{ブロック内のセル数} \\ &\times \text{方向量子化数}(N) \\ &\times \text{総ブロック数} \end{aligned} \quad (3)$$

ここで、画像内において、一セルずつ順番にブロックを移動し正規化を行う。総ブロック数は、画像のセ

ル数が縦6、横5、ブロック内に含まれるセル数が縦3セル、横3セルであった場合に12となる。このとき、ブロック内のセル数は9であり、方向量子化数 N を9とすると、HOG特徴量の次元数は972となる。

HOG特徴量においても、画像をラスタスキャンしつつ対象物体を探索する場合には、セル内のヒストグラム計算を繰り返し実施することとなり、膨大な計算コストが生じる。この計算コスト削減のために、Haar-like特徴量と同様に、方向成分ごとにIntegral Imageを作成する手法 [5] が提案されており、現実的な処理時間で画像内ラスタスキャンによる人物検出処理等が可能である。

2.3.2 EHOg特徴量

HOG特徴量の派生であるEHOg特徴量 [6] では、主要方向成分エッジ量の相対表現と、ブロックにセルを定義しないことにより、特徴量の次元数が削減される。HOG特徴量ではセルごとにヒストグラム化した方向成分を、ブロック単位で正規化し、これをすべてHOGの特徴量の各ベクトルに割り当てていた。EHOgでは、以下の式により各ブロックにつき $F_b(D)$ を求め、これを特徴量とする。

$$F_b(D) = \sum_{i \in D} g_b(i) / Z_b \quad (4)$$

ここで、 D はブロック内で着目する主方向を表す。 D は9方向に量子化された方向成分から、連続した1から3方向の組み合わせによる27パターンの中から選択される。このとき、 D の含有する角度は 20° から 60° となる。

また、ブロック内にセルを定義しないため、EHOg特徴量の次元数は総ブロック数となる。したがって特定ブロック内に特定方向のエッジが統計的に得られる場合には、HOG特徴量同等の検出精度を保ちつつ、大幅な次元数の削減が可能である。

2.3.3 HOG特徴量の2値線画像およびマンガ画像への適用性

HOG特徴量は、セル単位でヒストグラムが作成されるため、若干の形状変化が許容される。また、方向成分の分布を捉えるため、現在一般に使用されている画像特徴量の中では線画像に対しても高い識別性能がある。一方で、2値線画像に対しては、方向および勾配強度の計算方法に問題がある。以下ではHOG特徴量の2値線画像に対する問題点を示す。

HOG特徴量では2.3.1に示した通り、注目画素の上下左右の輝度差からエッジ強度とその方向が算出される。図2にHOG特徴量計算時の参照画素を示す。図2において、丸で示された画素は縦方向、十字で示された画素は横方向の参照画素である。(A)–(D)に示される様に45度方向に線を構成する画素が存在する場合や、垂直、水平に直線画素が並ぶ場合に、式(1)、

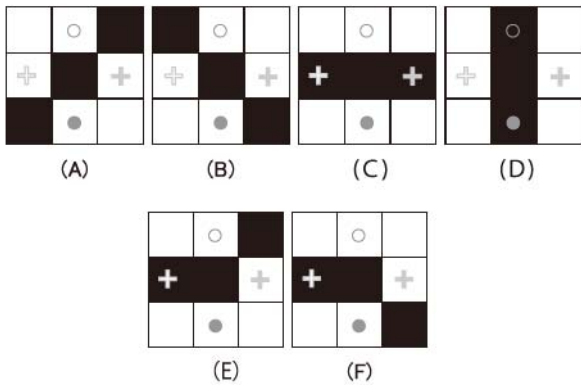


図2 2値細線上のHOG特徴量参照画素
Fig. 2 Reference pixels of HOG feature on binary thin line

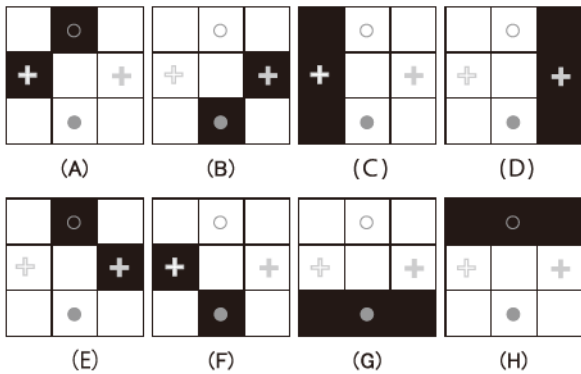


図3 2値細線周囲のHOG特徴量参照画素
Fig. 3 Reference pixels of HOG feature around binary thin line.

(2)により勾配強度を求めると V_d , H_d 共に差分を検出できないため、線を構成する画素上では特徴量が得られない。また、(E), (F)の場合には左右方向の差分は存在するが、上下方向の差分を取得できないため、(E), (F)の線の状況を区別できず、さらに直線自体の方向とは異なる90度としてエッジ方向を観測する。

直線上の画素では多くの場合上記のように特徴量が得られないため、基本的には図3に示す様に線と隣接する点において得られる値から間接的に直線を表現することとなる。ただし、この場合においても、観測点が4点でかつその値が1か0のみに限られる。HOG特徴量では方向を、 $[0^\circ, 180^\circ)$ の区間に射影し、その中で任意の方向数のbinを生成する。よってここで得られる方向は0度(C, D), 45度(A, B), 90度(G, H), 135度(E, F)の4種類のみとなる。また、2線が平行して存在する場合には参照点間の輝度差が0となるため、やはり方向の測定が不可能となる。

HOGの方向算出方法は通常の濃淡画像では有効な数値を得られる。しかし2値細線画像に対しては直線を構成する画素の上では方向成分を取得できない点、周囲の画素からの影響により本来出現しない方向成分が現れる点が問題となる。

2値細線化されたマンガ画像への適用という観点から評価した場合、マンガ画像では、多数の線により絵が描かれるため、人体検出と比較すると、多数の方向

にエッジ成分が出現する。また、対象により描かれる線分のパターンが変化すると考えられるため、対象領域内に存在する線の方向分布に特徴が現れる。HOGでは複数方向の分布を表現可能であるが、直線方向検出処理時に、隣接する線や孤立点からの影響を受ける。この影響を除外するためには、線を構成する画素上にて線の方向が記述可能である必要がある。

一方EHOGの主要方向算出アルゴリズムは、対象領域内にシンプルなエッジが存在し、その角度が一定方向に集まるという前提が成り立つ場合に有効とされる。しかしマンガ画像から得られる2値細線画像では、出現する線の方向は多岐にわたり上記条件を満たさない。

2.4 Edgelet特徴量

Edgelet [7] はline, arcとして直線と弧形状、そしてこれを対象に配置したsymmetric-pairなどの形状を、あらかじめ定義する。そして、定義された形状上の、重み付き評価値を基に学習を行う。検出時には、学習された形状に対して高い相関を示すエッジ成分を探し、その類似度を測る。このため、人物検出等において、人間の両肩や人間の頭部等に起因するエッジをその位置と共に得ることができる。この特徴から、Edgelet特徴量は人物検出処理など、位置と共に特徴的なエッジが計測されやすい状況であれば良好な識別性能を持つ。

-1	-2	-1	-1	1
			-2	2
1	2	1	-1	1

Vertical Horizontal

図4 Sobelオペレータ

Fig. 4 Sobel Operator

Edgeletでは、line, arc等の定義された形状を構成する画素上において、それぞれ重み付き評価値を計算する際に、 3×3 画素のSobelオペレータを使用する。評価値の計算では、 $[0^\circ, 180^\circ)$ の区間に射影された角度成分を6方向に量子化し、そのインデックス差に応じて重みづけを行う。

Sobelオペレータは2方向の線形フィルタからなり、周囲画素値に対して図4に示される重みを掛けて計算される。一方2値細線画像に対する勾配方向の算出ではHOG特徴量と同様に、着目点が直線上に存在し、図2(A)-(D)の配置となる場合には縦方向横方向双方の差分が0となり、直線の方向が算出できない。図2(E), (F)の場合にはそれぞれ45°, 135°となるが、直線上にて得られる方向成分だけでは直線を表現できない。したがって線形状は近隣画素から得られる値により、間接的に表現する必要がある。SobelオペレータではHOG特徴量計算時と比較して参照画素

数が増えるため、線の近隣にノイズや他の直線が存在する場合など、実際の線とは関係のない方向成分が算出されやすい。また、マンガのように、様々な方向の線分が多数存在する場合には、高い相関を持つ Edgelet 特徴量が対象領域内に無数に観測されるため、検出精度が著しく低下する。

2.5 複数局所領域の関連性を用いる手法

Joint-HOG [8], CoHOG [9], Deformable Parts Model [10] などに代表される共起特徴量では、検出領域内において、複数箇所の小領域に対して、従来の特徴量による検出結果を取得し、これを学習することで、単一の特徴量を用いるよりも高い精度をもつ強識別器を構成するものである。上記3手法では、先に示したHOG特徴量による弱識別機を利用している。Shapelet [11] は複数の小領域内におけるエッジ方向の共起性に着目した特徴量である。Shapeletではその低レベル特徴として、 $[-1, 0, -1]$ のカーネルを用いて0, 45, 90, 135度の4方向について差分値を取得する。一般的に用いられるHOG特徴量と比較して、方向の分割数が少ないため人物検出精度は低い。また、2値細線画像においても方向取得性が低い。

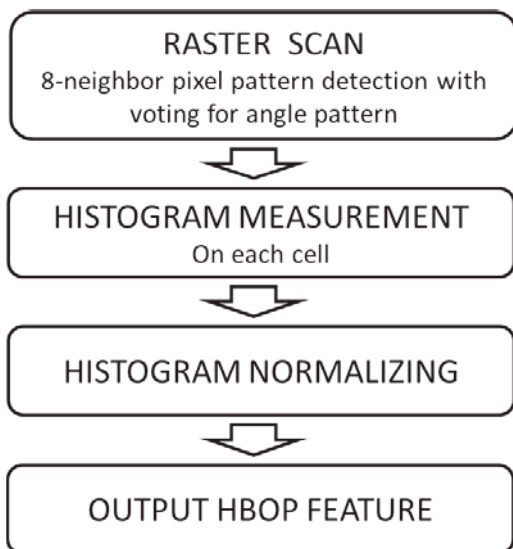


図5 HBOP特徴量の計算手順
Fig. 5 Calculate procedure for HBOP feature

線画像はエッジ抽出処理後の画像とほぼ同等のものである。一般的に使用されている画像特徴量の方向成分取得アルゴリズムは、線そのものの方向を記述する

為に設計されていない。特に2値細線画像からの特徴量計算では、内部で利用する低レベル特徴量の出力に依存するものであり、これまでに述べた2値画像およびマンガ画像に対する問題点が基本的にそのまま残る。一方で、共起性を用いるという戦略は利用する特徴量に左右されない。したがって、共起特徴量の手法において、その内部で利用する特徴量を、2値細線画像に有効な手法に置き換えることで、全体としてより高精度な強識別器の構成が可能であると考えられる。

3 2値線画像向け画像特徴量HBOP

3.1 提案手法の概要

2値線画像解析処理の高精度化を実現するため、2値細線画像に特化した特徴量であるHBOPについて述べる。2値細線化されたマンガ画像では、一般的な自然画像と比較して多数の線およびテクスチャが存在する。したがって、近隣画素からの影響をなくすため、直線を構成する画素上で、直線方向を記述できることが望ましい。HBOP特徴量ではHOG特徴量を基に、2値細線を構成する画素上で直線方向の記述が可能となるようにアルゴリズムを設計した。

従来の画像特徴量の中で、HOG特徴量は直線近傍画素を用いることで間接的に線分方向を記述できるため、その基本成分を線により構成されたマンガ画像解析においても一定の解析性能を持つことが確認されている [8]。ただし、HOG特徴量において得られる方向成分のヒストグラムは線そのものの方向分布を表現するものではない。

提案するHBOPは2値の線分の取り扱いに特化したものであり、注目点の8近傍に存在する画素のパターンを集約し、領域ごとにその出現量をヒストグラム化する。ヒストグラム化するパターンを線分中出现するものに限ることで、2値線分を取り扱うために冗長な情報を削減できる。

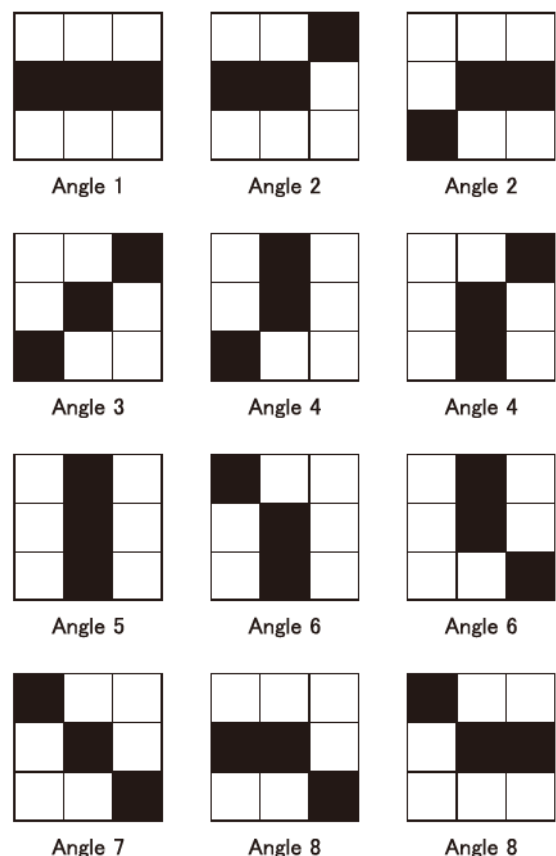


図6 HBOP特徴で用いる8近傍2値線画パターン
Fig. 6 8-neighbor binary line pattern used by HBOP Feature

3.2 HBOP特徴量の計算手順

学習時のHBOP特徴量算出手順について述べる。HBOP特徴量では入力画像を任意数のセルに分割し、セルごとに出現する方向パターン histograms を計算する。ここで、入力画像については2値の線画像とする。学習時には入力画像全体から一つのHBOP特徴量の計算を行う。処理は図5に示される4ステップに分かれる。画像内のラスタスキャンを行い、そこで得られる8近傍画素パターンにより方向パターンへの投票を行う。次に、各セルにおいて出現する方向パターンの histograms を計算する。そして正規化を行うことでHBOP特徴量を得る。各ステップの詳細を以下に述べる。

ステップ1では、画像内に出現する8近傍画素パターンを計測する。ここで利用する8近傍画素パターンと方向パターンを図6に示す。まず、HBOP特徴量では注目する画素自身が2値の線上である場合のみ histograms への追加処理を行う。注目点自身が線上である場合の8近傍画素について、HBOPではこの12種類のみを扱うこととする。実際の処理では、図6に黒色で示される3画素について線上か否かの判定を行う。3画素とも線上であればそのパターンの属する方向カテゴリに対して投票を行う。図6に示されたAngleは投票される方向カテゴリを表す。したがってステップ1が終了した段階で画像内に出現する各方向パターンの出現数が取得される。

ステップ2では、セル内 histograms の計算を行う。HBOP特徴量では画像全体を任意数のセルに分割し、それぞれのセルにて各方向パターンの合計出現回数を histograms 化する。

ステップ3では、得られたセル内の全方向パターンの合計により、ステップ2で得られた histograms を割ることで正規化する。HOG特徴量では照度変化に対応するため、複数セルから構成されるブロックによる正規化が行われていた。2値線画像では、画像内に現れる線はそのままセル内の全方向パターンの合計値として得られ、かつすべての点の強度は一定であるため、ブロック単位での正規化の必要性はない。

ステップ4では、ステップ3までで得られた特徴量をブロックごとに各方向成分を順番に並べたものをHBOP特徴量とし、これを出力する。ここで出力されるHBOP特徴量の次元数は次の式により求められる。

$$\text{Dimension} = \text{Cell Number} \times \text{Angle Pattern Size} \quad (5)$$

セル数が 5×5 、方向パターン数が8の場合、HBOP特徴量の次元数は200となる。

一般的に線の終端や角部分などでは図6に示すHBOP特徴量で扱わないパターンが出現する。出現例を図7, 8に示す。図7(A)および図8(A)に示されるパターンは、図6には出現しない。一方でこれらのパターンは直線上に多数出現することが考えられ

る。ただし、HBOP特徴量計算時には、線を構成する隣接画素において得られるHBOPパターンにより、間接的に線形状を表記できる。具体的には、図7では(B), (C), 図8では(B)–(E)に示されるパターンにより、線の形状が表現される。図中の三角などの形状はそれぞれのパターンにおいて参照する画素を示す。また、この様な重ね合わせによる記述を行うことで、4連結、8連結ともに共通のパターンで記述可能である。HBOPでは、HOG特徴量等における直線の間接表現と比較して、直線を構成する画素上にてパターンを取得するため、近隣に直線がある場合でもその影響を受けない。

HBOPで用いるパターン数を増やし、全ての8近傍パターンを導入することも可能であるが、特徴量の次元数が増加するというデメリットがある。一方で、上記の理由から、図6に示すパターンを用いることで、2値細線形状を表現することは十分に可能である。したがって、本提案手法においては、図6に示したパターンのみを用いることとした。

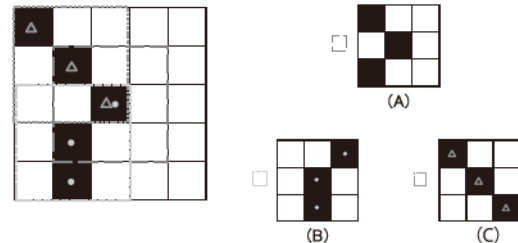


図7 HOBPを用いた2値細線パターンの表現 (1)
Fig. 7 Representation for binary thin line using HBOP pattern (1)

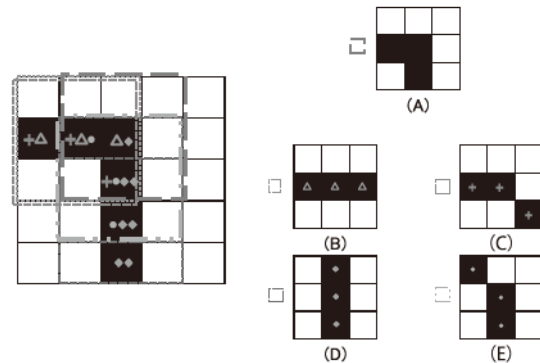


図8 HOBPを用いた2値細線パターンの表現 (2)
Fig. 8 Representation for binary thin line using HBOP pattern (2)

3.3 検出処理時の計算量削減

検出処理における特徴量算出の手順を述べる。検出処理時には、 histograms 計算を多数行う必要があるため、計算量が膨大となる。Haar-like特徴量や改良されたHOG特徴量 [5, 6] では、矩形領域内における輝度合計値の計算コストを削減するためにIntegral Imageを用いる。HBOP特徴量の算出においても方向パターンごとに出現回数をまとめたIntegral Imageを生成することで、計算量の削減が可能である。

方向パターンのIntegral Image化には以下の式を用いる。

$$Integral(x, y) = P_a(x, y) + I_a(x, y - 1) + I_a(x - 1, y) - I_a(x - 1, y - 1) \quad (6)$$

$$I_a(x, y) = \begin{cases} 0 & (x = 0 \wedge y = 0) \\ P_a(x, y) + I_a(0, y - 1) & (x < 1) \\ P_a(x, y) + I_a(x - 1, y) & (y < 1) \\ Integral(x, y) & (otherwise) \end{cases} \quad (7)$$

ここで、 I_a はIntegral Imageのa番目の方向パターンを示す。 P_a は方向パターン計算時に得られたa番目の方向パターンの出現回数が格納されている2次元平面である。各方向パターンにつき、上記x, yを1ずつ変化させ全画素スキャンを行うことで、方向パターンのIntegral Imageを取得する。



図9 マンガ画像の2値細線表現 [14]
Fig. 9 Binary thin line representation of comic image [14]



図10 学習用正例画像例
Fig. 10 Positive sample images for learning

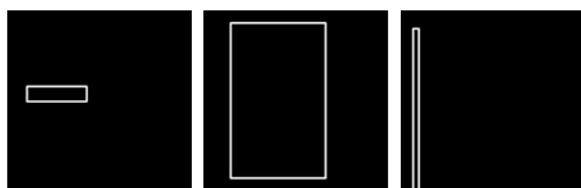


図11 学習用負例画像例
Fig. 11 Negative sample images for learning

3.4 2値化について

HBOP特徴量では2値の線画像をその処理対象とする。入力画像がもともと2値線画像であればそのまま特徴量計算が可能である。また、一般の多値画像においても、一般的な前処理により2値線画像化することで、同様に取り扱うことができる。ただし、得られる画像は前処理に依存するため、用途により有効な2値線画像化を行う必要がある。

ここで、HBOP特徴量におけるパターンのヒストグラム化は画素ごとに行われる。注目する線を構成する画素の近傍画素に存在する全てのパターンに対して投票を行う。このため、4連結、8連結どちらの細線化処理でも扱うことが可能である。ただし、前処理として用いる細線化手法により得られる線の位置は一般的に異なるため、学習時と検出時には同様の手法で細線化処理を行う必要がある。

一方で、マンガ画像に対する予備実験では正規化画像サイズが150×150 [pixel] かつセル数が5×5の場合に良好な動作を示した。図9にこの画像例を示す。この条件では1セルは30×30 [pixel] となる。このセルサイズは通常の人物検出で用いるものと比較して大きい。セルのサイズが大きく、かつ学習、検出で用いる細線化処理が同一という条件の下では、細線化処理アルゴリズムに起因する線の若干の差異については、セル単位のヒストグラム化により十分に吸収可能であると考えられる。

4 HBOP 評価実験

4.1 HBOP 評価実験の概要

本章では、HBOP特徴量の2値細線画像に対する特性を示すため、HOG特徴量との比較実験を行う。単純な2値線画像に対する学習、識別および検出実験を行い、HBOP特徴量の特性を明らかにする。次に、本論文の処理ターゲットであるマンガ画像を用い、学習、識別実験を行い、その性能を示す。

4.2 単純2値線画像に対する識別実験

4.2.1 実験に関する設定および実験環境

HOG特徴量と本提案方式であるHBOPの2値線画像に対する特性を調査するため、各特徴量において次元数ごとの学習の収束状態、学習時間、識別性能について調査を行った。

学習用入力画像はすべて50×50pixelの2値線画像を用いた。学習用の正例として円を、不例として矩形を用いた。また、円については中心座標および直径を、矩形については始点、幅と高さをランダムなパラメータに変更し作成した画像を使用する。正例および不例にはそれぞれ700枚の画像を使用した。学習用画像の例を図10および図11に示す。

実験において、HBOPはセル数を3×3, 5×5, 7×7と変化させた。各HBOP特長量の次元数はそれぞれ72次元, 200次元, 392次元である。以下ではそれぞれの次元数に基づきHBOP72, HBOP200, HBOP392と表記する。HOG特徴量はすべて5×5セル、角度の量子化ステップは9とし、正規化ブロックのサイズをそれぞれ1×1, 2×2, 3×3とした。各HOG特徴量の次元数はそれぞれ225次元, 576次元, 729次元である。以下それぞれの次元数に基づきHOG225, HOG576, HOG729と表記する。

各特徴量の学習にはReal-Adaboostアルゴリズムを使用した。Real-Adaboostの実装には中部大学藤吉研究室で公開されているプログラム[13]を基に、提案手法に合うように一部変更したものを使用した。

実験環境を表1に示す。また、実験のため、学習時間の計測時には並列処理を無効とした。

4.2.2 HOG特徴量およびHBOP特徴量の学習特性

4.1で示した学習用画像を用いた学習結果について調査を行った。実際に学習を行った際の収束状態について図12に示す。今回使用した画像はランダムに生成した円と四角形であるため、学習対象としては単純な部類である。このため収束は早くすべての特徴量において20回程度までの学習で学習に用いた正例に対するTrue Positiveが1.0となった。本結果から、今回利用する特徴量の各パラメータにおいて学習は問題なく収束していることが確認できる。

学習に要した時間の学習回数ごとの変化を図13に示す。学習回数が極端に少ない時点においては、画像からの特徴量計算に関するオーバーヘッドがある。25回以降では基本的に特徴量の次元数に応じて学習にかかる計算時間が増加していることが確認できる。

HBOP特徴量はHOG特徴量と比較して、ブロック単位での正規化が無いこと、方向成分の量子化パターンが限られることから、基本的に次元数が低い。このため、同程度のセル数を用いた場合には計算量の削減が見込める。学習回数が400回の時点においてもHBOP72およびHBOP200については約1秒までの時間で学習されており、より複雑なモデルを学習する際により多い繰り返し回数が必要となる場合に計算量において利点があると考えられる。

4.2.3 未知画像識別実験

4.2.2において学習を行った結果を用いて、未知画像に対する識別実験を行った。本実験で用いる未知画像は学習用画像と同様の手順により生成し、円および矩形に対する識別成功率を調査した。ここで生成したテスト用画像は、学習用画像と生成時のランダム値が異なる画像である。テスト用画像は正例500枚、負例500枚の合計1,000枚を使用した。各特徴量における正例、負例に対する学習回数ごとの識別正解率を表2および表3に、その平均値をグラフ化したものを図14にそれぞれ示す。図14より、すべての手法で95%を超える正解率を持つことが確認できる。この中で、HBOP200の学習回数50、100、200およびHBOP392の学習回数400の際に最も高い平均識別精度99.8%を達成した。また、HBOP200においては、すべての測定点においてHOGよりも高い識別成功率を得た。

注意点として、すべての手法において、学習が収束していない5回程度の学習時点であっても未知画像に対する識別成功率が95%を超えている。この主な要因は本テストで学習・識別対象とした図形が比較的単純

である点および今回のテスト用画像として学習用画像同様の手順で生成された画像を使用している点によるものであると考えられる。

HBOP200とセルの分割数が同一でブロックにおける正規化を行わないHOG225と比較すると、HOG225では学習が進んでも結果が安定せず、その識別精度もHBOP200より低い値となっている。このことから、HBOPで設定した画像方向パターンの設計は、本論文の目的である2値細線画像を対象とした画像特徴を得るという目的において、HOGで用いる特徴よりも高精度かつ高効率であるといえる。

4.2.4 複雑な画像からの円検出実験

より複雑な状況からの形状検出に対する適応性を調べるため、複数の円と矩形が混在する画像上での円検出実験を行う。実験に用いた画像例を図15に示す。実

表1 実験環境

Table 1 Experimental environment

OS	Ubuntu13.10
Memory	32GB
CPU	Core-i7 3770K
OpenCV	Version 2.4.5
g++	Version 4.8.1

True Positive (%)

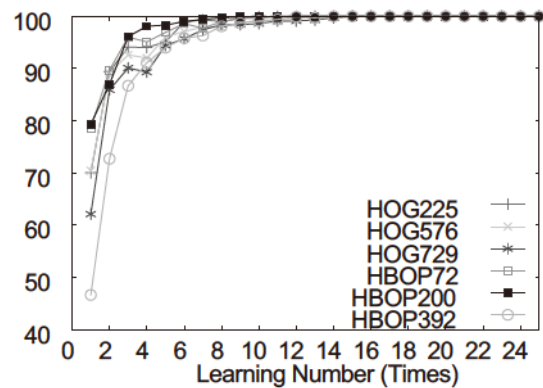


図12 学習の収束状態

Fig. 12 Convergent status of learning

Time (Second)

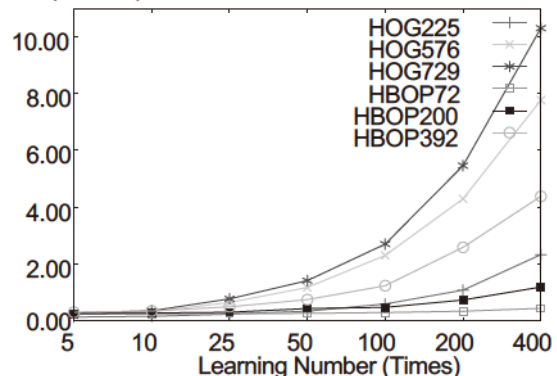


図13 学習処理時間

Fig. 13 Process time for learning

表2 未知入力画像（正例）に対する識別精度

Table 2 Identification rate for unknown image (Positive)

Iteration	HOG			HBOP		
	225	576	729	72	200	392
5	0.974	0.952	0.950	0.980	0.982	0.944
10	0.990	0.984	0.984	0.996	0.992	0.984
25	0.990	0.992	0.982	0.992	0.994	0.980
50	0.982	0.994	0.990	0.990	0.996	0.986
100	0.984	0.996	0.992	0.988	0.996	0.990
200	0.982	0.998	0.990	0.974	0.996	0.992
400	0.984	1.000	0.990	0.982	0.994	0.996

表3 未知入力画像（負例）に対する識別精度

Table 3 Identification rate for unknown image (Negative)

Iteration	HOG			HBOP		
	225	576	729	72	200	392
5	0.960	0.970	0.978	0.996	0.990	0.964
10	0.986	0.980	0.988	0.994	0.998	0.994
25	0.988	0.978	0.992	0.996	1.000	0.988
50	0.984	0.992	0.990	0.994	1.000	0.996
100	0.992	0.988	0.986	0.984	1.000	1.000
200	0.988	0.992	0.990	0.974	1.000	1.000
400	0.990	0.990	0.990	0.974	0.996	1.000

Accuracy (%)

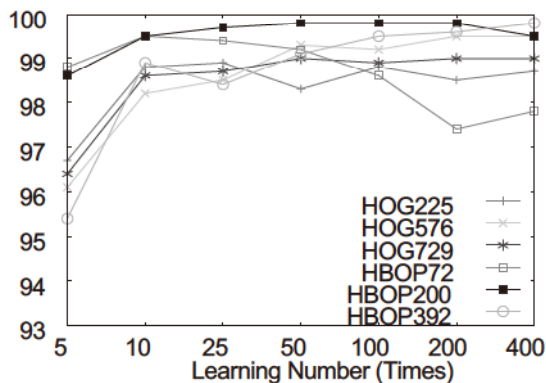


図14 未知入力画像に対する平均識別精度

Fig. 14 Average Identification Accuracy for Unknown Image

験には同様に作成した50枚の画像を使用した。

ここでは、先の実験において識別精度の高かったHOG576の学習回数400の場合と、HBOP200の学習回数200およびHBOP392の学習回数400について検出結果をまとめる。

対象画像に対してスライドウィンドウによるラスタスキャンを用いて検出処理を行う。また、多数の近傍領域に対する検出結果を整理するため、Mean Shift Clusteringによる統合処理を行う。検出ウィンドウは正方形を用い、サイズは25×25画素から125×125画素まで25画素刻みで変化させた。

本検出実験の評価は、円形の検出が主目的であるため、検出された領域内にある円の75%以上が含まれていれば成功、これ以外を失敗として評価した。ここで

表4 複雑な配置における円検出実験結果

Table 4 Circle detection results on complex images

Method	HOG576-400	HBOP200-200	HBOP392-400
True Positive	340	349	345
False Positive	44	12	4
Precision	0.885	0.967	0.989

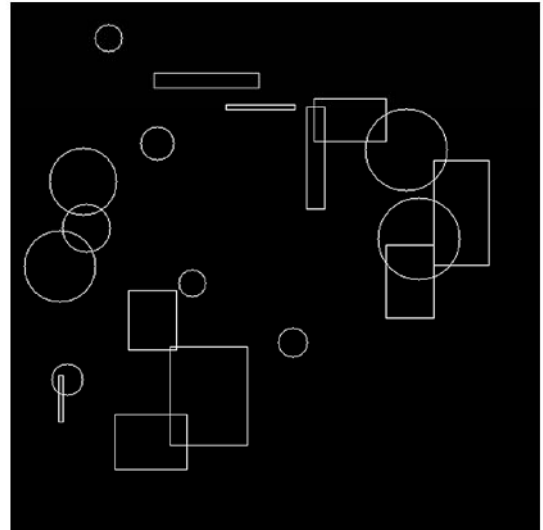


図15 複数の円と矩形が混在する実験画像

Fig. 15 Test image includes multi circle and rectangular.

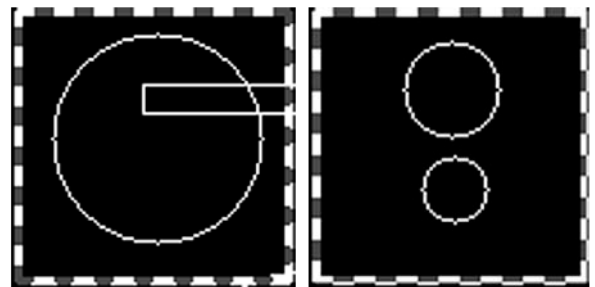


図16 複雑画像上の円検出における正解画像例

Fig. 16 Valid image for circle detection on complex image

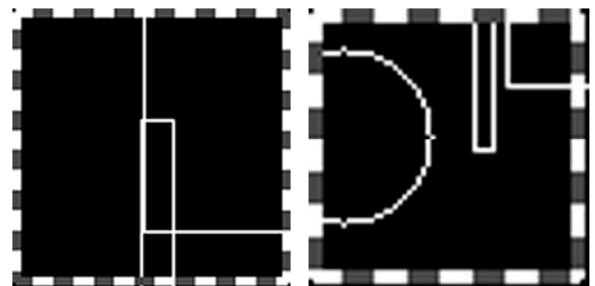


図17 複雑画像上の円検出における不正解画像例

Fig. 17 Invalid image for circle detection complex image

は、円を検出するという目的から、上記成功条件を満たしてさえいれば、検出領域内に複数の円が含まれる場合もしくは検出結果に矩形が含まれている場合も正解としている。

検出成功例と検出失敗例を図16および図17に示す。実際に検出を行った例を図18に、本実験結果の総検出数及び検出成功率を表4に示す。ここで、識別精度は全検出結果中の正解数の割合とした。今回の実験では、検出対象の円には多数の重なりが存在し、また、画像中からの検出アルゴリズム中に統合処理が入ることから、再現率については評価を行わず、正解率についてのみ評価を行った。

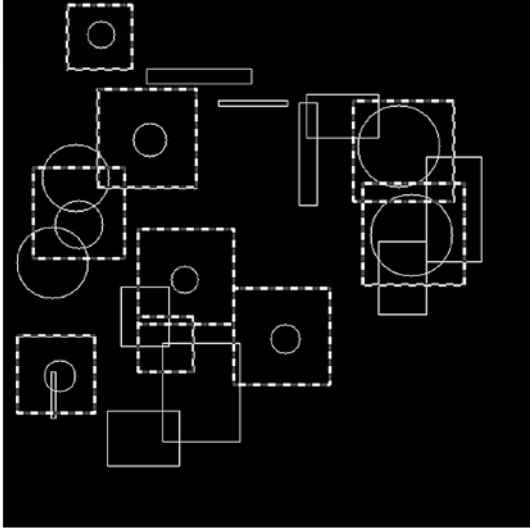


図18 図15に対するHBOPによる円検出結果
Fig. 18 Circle detection result on Fig 15 by HBOP

表4に示された実験結果より、HBOPではHOGと比較してTrue Positiveが高くまたFalse Positiveも少ない結果となった。特にFalse Positiveの減少により、全体としての識別成功率では8%から10%の向上が見られた。画像では、HOGで誤検出している矩形が重なっている部分においてHBOPでは誤検出が減少している。以上より、複雑な状況においてもHOG特徴量と比較して、本提案手法であるHBOP特徴量の線画像解析に対する優位性が確認できる。

4.3 マンガ画像識別実験

本論文において処理の主対象とするマンガ画像に対して、識別実験を行う。マンガ画像において登場人物の抽出は重要な課題である。ここでは、登場人物の形状の中で、特に特徴的に描かれる傾向を持つ瞳部分を学習、識別する。

実験には360dpiで描かれたマンガ画像から人手により切り出した画像を用いる。ここで、今回使用したマンガ画像は完全な2値画像ではなく、高解像度グレースケール画像としてスキャンされたものを360dpiに変換したものである。

HOG特徴量とHBOP特徴量は、若干の位置ずれに関しては許容可能なものの、シフト不変性は持たないため、画像内の同一部位から切り出し領域をさまざまに変更し取得した画像を実験に使用した。学習用の画像には、瞳部分を切り出した画像を正例、瞳以外の領

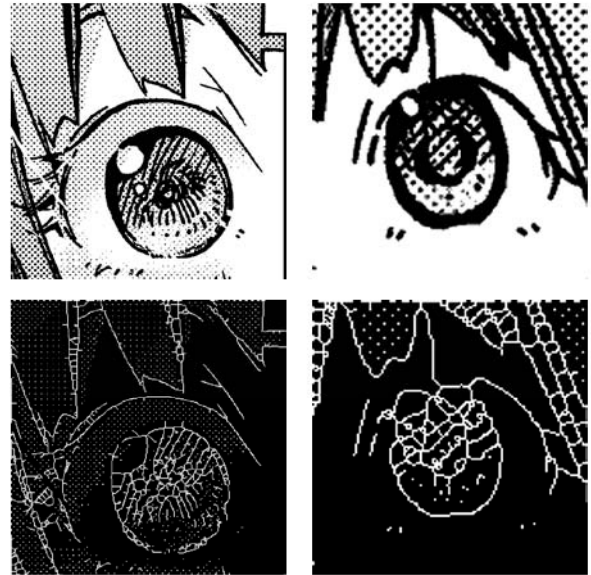


図19 正例画像とその2値細線画像 [14]
Fig. 19 Positive sample images and binary thin line version [14]



図20 不例画像とその2値細線画像 [14]
Fig. 20 Negative sample images and binary thin line version [14]

域を切り出した画像を不例とし、それぞれ368枚、499枚を使用した。識別実験用の未知画像には正例250枚、不例250枚を用いた。実験に使用した画像例を図19、20に示す。

HBOP, HOGともに学習回数は200回とし、学習時のセル数は 5×5 、HOG特徴のブロック数は 3×3 とした。また、HOG特徴量の量子化方向は9とし、特徴量の次元数はHBOPが200、HOGが729である。また、正規化画像サイズは 50×50 , 100×100 , 150×150 , 200×200 の4段階とする。入力画像からの2値細線画像の生成には、2値化後にモルフォロジ処理による細線化を用いた。今回のHBOP特徴量の実装では、線を白、背景を黒と定義しているため、マンガ画像からの2値細線化画像生成時に、画像の色反転

表5 マンガ画像に対する識別実験結果
Table 5 Discrimination results on comic images

	Resolution	TruePositive	TrueNegative
HBOP	50	0.765	0.908
	100	0.880	0.912
	150	0.972	0.880
	200	0.964	0.805
HOG	50	0.729	0.829
	100	0.841	0.813
	150	0.805	0.884
	200	0.833	0.865
HOG(BW)	50	0.896	0.896
	100	0.861	0.912
	150	0.869	0.869
	200	0.861	0.888
HOG(Normal)	50	0.880	0.912
	100	0.928	0.932
	150	0.892	0.936
	200	0.900	0.936

を行っている。HOG特徴量では、細線化されていない画像からも特徴を得ることが可能であるため、細線化、2値化 (BW)、前処理無し (Normal) について、それぞれ実験を行った。

実験結果を表5に示す。2値細線化された画像において、HBOP特徴量では、HOG特徴量と比較して True Positive の値が全パラメータにおいて向上している。True Negative の値は一部下回っているが、その低下率は True Positive の上昇分と比較して小さいものである。

HOG特徴量では、通常の入力画像において識別を行った場合と比較して、2値画像および2値細線画像を入力した場合の True Positive, True Negative が共に低下している。これは、HOG特徴の計算過程において2値細線画像からの方向および勾配値取得の問題点によるものと考えられる。

HOG特徴量では前処理なしで正規化サイズが100の場合に、最も優れた識別性能を得られている。HBOPにおける正規化サイズ150の場合には、これと比較して、True Negative において0.052の低下を伴うものの、True Positive において0.044の向上が見られた。一般的に、Joint特徴量などのカスケード型の識別機を構成する場合には、True Positive の値が高い弱識別器が利用できることが望ましい。この点でHBOP特徴量ではHOG特徴量と比較して良好な結果を得られたといえる。

5 まとめ

本論文では、2値線画像解析に利用可能な新たな画像特徴量として、2値線分の方向ヒストグラムを利用した、Histograms of Binary Orientation Pattern (HBOP) 特徴量を提案した。提案したHBOPの2値線画像特性について、円と矩形の学習および識別において、

HOG特徴量と比較して少ない次元数でより高い検出精度が得られた。また、学習回数を変化させて実施した識別実験においては、200次元のHBOP特徴量はHOG特徴量の識別結果をすべての点で上回る結果が得られた。複数の円と矩形を配置した複雑状況下の円検出実験においては、HBOP特徴量を使用した場合HOG特徴量での実験結果と比較して正解数を減らすことなく、検出結果の成功率が8%から10%高い精度で検出可能であることを確認した。

また、実際のマンガ画像を用いた識別実験においても、HOGを上回る識別精度が得られた。実験結果より、2値の細線画像解析において、HBOP特徴量は低次元かつ高精度であり、HOG特徴量に対し優位性があることが確認された。また、本研究の処理対象であるマンガ画像に対しては、前処理なしのHOG特徴量による識別結果と比較して0.052の True Negative の低下を伴うものの、0.044高い True Positive を得ることが確認された。

今後の検討課題として Joint特徴量やその他カスケード型アルゴリズム等における、HOG特徴量の代替としてのHBOP特徴量の導入が挙げられる。Joint特徴量等では、人物検出等においてHOG特徴量をそのまま利用した場合と比較して、良好な検出精度を実現している。2値細線画像に対する処理においては、利用する低レベル特徴としてHBOPを用いることで、より高精度な解析処理が実現可能であると考えられる。

謝 辞

本研究はJSPS科研費25330137の助成を受けたものである。本稿にて例示したマンガ画像はすべて木野陽様 <http://www.etheric-f.com/> より学術目的の為に使用を許可されているものである。マンガ画像の提供及び原稿への掲載を許可いただいた木野陽様に深く感謝する。

参考文献一覧

- [1] 山下, 藤吉. “特定物体認識に有効な特徴量”, 情報処理学会研究報告 CVIM 2008 (115), pp.221-236, Nov. 2008.
- [2] 山内, 山下, 藤吉, “画像からの統計的学習手法に基づく人検出”, 電子情報通信学会論文誌D, Vol.J96-D No.9, pp.2017-2040, Sept. 2013.
- [3] P. Viola, M. J. Jones, “Robust Real-Time Face Detection”, International Journal of Computer Vision, Vol. 57, No. 2, pp. 137-154, 2004.
- [4] N. Dalal, B. Triggs: “Histograms of Oriented Gradients for Human Detection”, Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 886-893, 2005.
- [5] F. Porikli. “Integral histogram: A fast way to extract higtograms in cartesian spaces”, Coference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp829-836, June 2005.
- [6] C. Hou, H. Ai, and S. Lao, “Multiview Pedestrian Detection Based on Vector Boosting”, Asian Conference on Computer Vision, Vol. Part I, pp. 210-219, Nov. 2007.
- [7] B. Wu, and R. Navatia, “Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of

- Edgelet Based Part Detectors”, International Journal of Computer Vision Vol. 75(2), pp.247-266, Nov. 2007.
- [8] 三井, 山内, 藤吉, “Joint HOG特徴を用いた2段階Adaboostによる人検出”, 画像センシングシンポジウム(SSII), pp. IN1-06, June, 2008.
- [9] T. Watanabe, S. Ito and K. Yokoi, “Co-occurrence histograms of oriented gradients for pedestrian detection”, IPSJ Transactions on Computer Vision and Applications, Vol. 2, pp. 39-47, March, 2010.
- [10] 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, Sept. 2010.
- [11] P. Sabzmeydani, and G. Mori, “Detecting Pedestrians by Learning Shapelet Features”, Computer Vision and Pattern Recognition, 2007. CVPR’ 07 IEEE Conference on, pp. 1-8, June 2007.
- [12] 石井, 渡辺, “マンガからの自動人物検出と識別に関する一検討”, 画像電子学会論文誌, Vol.42, No.4, pp.457-465, July 2013.
- [13] 中部大学藤吉研究室, “CV Tutorial”, FLAB : Computer Vision for Visual Surveillance and Mobile Robotics, <http://www.vision.cs.chubu.ac.jp/CVTutorial/>, 参照 Dec. 2013
- [14] 木野陽, “ベリーベリークリームショコラふたつのベリー”, 自己出版, 2010.



石井 大祐

2006年東京理科大学理学部第1部応用物理学科卒業, 2008年早稲田大学大学院国際情報通信研究科修士課程修了。現在同研究科助手。画像処理および画像解析に関する研究に従事。画像電子学会, 電子情報通信学会, 情報処理学会, IEEE各会員。



渡辺 裕

1980年北大・工・電子卒。1985年同大大学院博士課程修了。同年NTT入社。NTT研究所において画像の高効率符号化及びMPEG標準化に従事。2000年早稲田大学大学院国際情報通信研究科教授。マルチメディアの符号化及び配信に関する研究に興味を持つ。工博。IEEE, 電子情報通信学会, 映像情報メディア学会, 情報処理学会, 画像電子学会各会員。