

H-016

勾配方向差分特徴量を用いた物体検出 An object detection method based on difference of directional gradient

細井 利憲†
Toshinori Hosoi

石寺 永記†
Eiki Ishidera

1. はじめに

画像からの物体検出技術は、サーバ上の膨大なデータからの情報検索や、組込機器でのリアルタイム映像解析への応用が期待され、高速かつ省メモリ処理の実現が望まれている。この物体検出分野で有効な特徴量として輝度の勾配に基づく特徴が知られている[1][2]。HOG[1]では、勾配方向を bin とした勾配強度ヒストグラムを局所領域別に求め、それを特徴とする。さらに多重解像度化して特徴間の関係を特徴量とした Joint-Hog[3]では、より効率良く物体検出できる事が報告されている。

一方、高速な物体検出の代表的手法である Viola ら[4]の手法では、膨大な特徴プールから物体・非物体識別に向く特徴を学習時に選択し、カスケード型の識別器によって少数の特徴のみで非物体と判定することで高速検出する。このような方式では、より効率的に識別できる特徴をプールに追加することで検出精度・処理時間の改善を図れる。しかし、特徴の候補数をむやみに増やすと学習時間が増大するし、識別に有効な特徴は一般に抽出コストが高く、識別計算する特徴数を削減できても全体の演算コストが増加してしまう。そこで、識別に有効かつ低演算コストの追加型特徴として、勾配方向差分特徴量を提案する。主に勾配強度ヒストグラム特徴と併用する事を想定した特徴量であり、勾配方向の鮮明さを表現する。

2. 提案特徴量と物体検出

2.1 勾配方向差分特徴量

勾配方向差分特徴量の抽出手順を説明する。まず、入力された輝度画像を局所領域のセルに分割し、各領域で勾配方向を bin とした勾配強度ヒストグラムを作成する。図1左は、ヒストグラムを方向番号1から8の8方向の bin 毎に分解し画像状に可視化した図である。この後、HOG のように部分領域を設定して正規化しても良いが、ここでは省略する。次に、方向番号が近いもの同士をまとめた「方向セット」を設定する。2方向分をまとめた「方向セット」の例を図1の右図に帯状の円弧で示す。方向セットAは方向番号2と3、Bは4と5をまとめた例である。次に、方向セットのペアを設定する。図1の例では方向セットAとBをペアとする。そして、セル座標 (x,y) 、勾配方向 d の勾配強度ヒストグラム値を $H_d(x,y)$ と表記すると、勾配方向差分特徴 $F_{AB}(x,y)$ は式(1)により計算される。ここで、方向セットのペアが方向セットAとBのような隣接関係になるという制約を付与することで、ヒストグラムの勾配を抽出する処理になり、勾配方向の鮮明さを表現する特徴を得られる。尚、この例の2方向分をまとめた「方向セット」の配置パターン数は、方向の bin 数8と同じになる。

† (株) NEC 情報システムズ 先端技術ソリューション事業部, NEC Informatec Systems, Ltd.

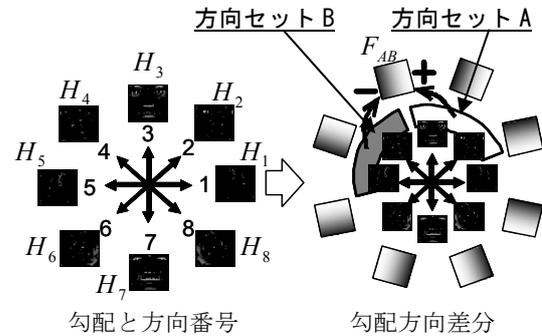


図1 勾配方向差分特徴量

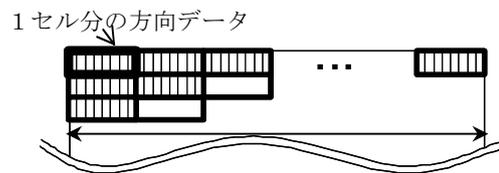


図2 勾配のデータ構造

$$F_{AB}(x,y) = \sum_{d \in \text{setA}} H_d(x,y) - \sum_{d \in \text{setB}} H_d(x,y) \quad (1)$$

2.2 データ構造

提案特徴量の抽出は和差演算のみで実行できるため、メモリアクセスによる遅延の影響が相対的に大きくなる。そこで、勾配ヒストグラムのデータ構造を図2のようにセル単位で方向番号順に並べることでランダムアクセスを抑制する。

2.2 物体検出時の特徴抽出

識別に使用する特徴数(次元)を動的に調整するカスケード型の識別器を用いる場合、通常、特徴抽出処理は必要に応じて逐次実行される。一方、勾配系の特徴抽出処理は比較的高コストであるため、入力画像全体に対して事前に実行してメモリに記憶しておく方が速い場合がある。本提案特徴量は勾配ヒストグラムからの演算コストが低いため、逐次実行すれば良く、本特徴を追加しても必要なメモリ量は殆ど増えない。

3. 物体検出実験

3.1 学習・識別手法

本実験では、学習・識別方式として GLVQ[2]と全探索型特徴選択処理の組合せを用いた。学習処理は基本的には GLVQ に基づくが、調整用データセットを用いて識別に有効な特徴を全探索によって選択する。識別時には、1次元ずつ逐次追加しながら識別スコアを算出し、特徴選

処理時に決定された閾値に基づいて非物体と判定された場合、その次いで処理を打ち切る。

3.2 実験条件

物体・非物体識別用の学習／調整／評価セットには、それぞれ FERET 顔 DB[5]等を含む 12,000 個の正面顔画像を用い(図 3 参照)、物体検出の評価には、学習／調整セットではない FERET 顔 DB を利用した。実験に用いた特徴量は、勾配ヒストグラムのみ(凡例: Org)、勾配ヒストグラムとそれを多重解像度化した特徴(凡例: +MultiRes)、勾配ヒストグラムと提案手法の勾配差分特徴(凡例: +DiffDir)の 3 種類である。「Org」は、ヒストグラム作成対象のセルサイズが 2×2 、セル数が 11×11 、方向 bin 数が 8 の 968 次元特徴量とした。「+MultiRes」は、セル数が 10×10 、 8×8 、 4×4 の勾配ヒストグラムを追加した 968+1440 次元特徴量とし、「+DiffDir」は、2 方向分をまとめた方向セットのみを採用し追加した 968+968 次元特徴量とした。尚、実行環境は Pentium4 3.4GHz の PC である。

3.3 実験結果

まず、学習により選択された上位 5 次元の特徴を図 4 に示す。本実験では、上位に両目周辺の勾配差分特徴が選択された。上位 5 つとも勾配差分特徴であったことから、少ない特徴数では勾配ヒストグラムよりも識別に有効だといえる。

次に、識別時の使用特徴数(特徴ベクトルの次元数)と非物体データの累積棄却率の調査結果(図 5) から、多重解像度化した特徴が 3 種中で最も効率良く低次元で非物体を棄却できることがわかる。勾配差分特徴は若干劣るが、勾配ヒストグラムのみの場合よりは明らかに良好であり、より低次元での棄却を期待できる。

次に、物体検出全体の処理時間を図 6 に示す。勾配差分特徴量の場合が最も高速であり、効率よく非物体を棄却しつつ低演算コストで処理できたことがわかる。多重解像度化した特徴については、効率よく棄却できるものの特徴抽出に時間がかかるため、全体としては速くならなかった。

最後に、物体検出率を図 7 に示す。図 7 の横軸は適合率、縦軸は再現率であり、勾配差分特徴量と多重解像度化した特徴量の場合の検出率が良好であることがわかる。

4. おわりに

勾配特徴量を用いたカスケード型の物体検出手法を高速化する勾配方向差分特徴量を提案した。勾配方向の鮮明さを表現する追加型の特徴量であって、実験によって、処理が低コストかつ識別に有効である事、また物体検出処理を高速化できる事を確認した。今後の課題・展望としては、広範な DB による評価、及び共起性を利用した特徴の評価等が挙げられる。

参考文献

- [1] N.Dalal, B. Triggs, "Histograms of oriented gradients for human detection", CVPR2005, vol. 1, pp. 886-893, 2005
- [2] 細井, 佐藤, "一般化学習ベクトル量子化による顔検出", 信学技報, Vol. 102, No. 651, pp. 47-52, 2003
- [3] 三井, 山内, 藤吉, "Joint HOG 特徴を用いた 2 段階 AdaBoost による人検出", SSI08, IN1-06, 2008

- [4] P. Viola, M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", CVPR2001, pp. 511-518, 2001
- [5] P. J. Phillips, et. Al, "The FERET database and evaluation procedure for face recognition algorithms", IJCV Vol. 16, No. 5, pp. 295-306, 1998



図 3 実験用画像の例

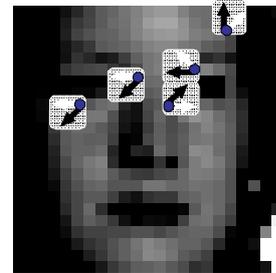


図 4 学習により上位に選ばれた特徴

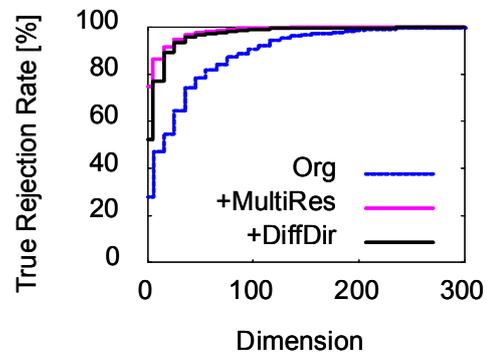


図 5 非物体クラスの累積棄却率

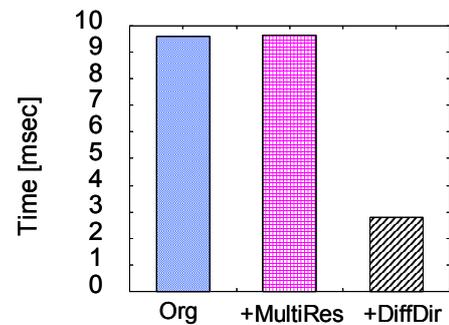


図 6 物体検出の処理時間

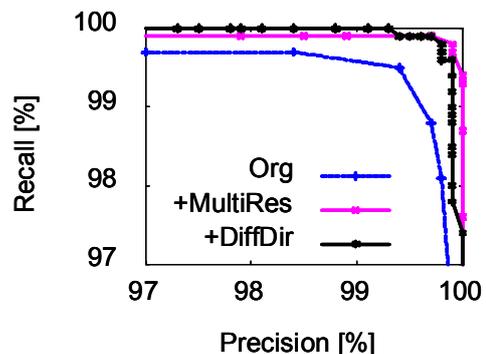


図 7 物体検出率