

# マンガの複数キャラクターに対する顔検出率について

## On face detection rate for characters in comic images

柳澤 秀彰<sup>\*1</sup>石井 大祐<sup>\*2</sup>渡辺 裕<sup>\*1,2</sup>

Hideaki Yanagisawa

Daisuke Ishii

Hiroshi Watanabe

<sup>\*1</sup> 早稲田大学大学院基幹理工学研究科<sup>\*2</sup> 早稲田大学大学院国際情報通信研究科<sup>\*1</sup> Graduate School of Fundamental Science and Engineering, WASEDA University<sup>\*2</sup> Graduate School of Global Information and Telecommunication Studies, WASEDA University

### 1. まえがき

近年、電子書籍市場の拡大に伴い、電子コミックに関するより高度な検索システムを提供するために、マンガ画像からメタデータを自動抽出する技術について研究が行われている。

本稿では、マンガ作品において重要な要素である登場人物の顔領域検出について、学習サンプルの違いによる検出率の変化を検討する。マンガの登場人物の顔領域は現実の顔画像と比較して、個々の特徴変化が大きい傾向にある。そこで、顔検出器の学習に、特定の登場人物の顔画像のみを使用した場合と、複数の登場人物の顔画像を使用した場合について、検出率の比較を行った。

### 2. 顔領域の検出手法

本稿では、マンガ画像から顔領域を検出するための手法として、Felzenszwalb らによって提案された物体検出手法である Deformable Part Model [1]を使用した。Deformable Part Model の検出モデルは物体の大まかな形状を捉えるルートフィルタと、物体の各パーツの形状を捉える移動可能なパートフィルタの 2 種類から構成されている。具体的な検出手順は以下ようになる。

1. 複数の解像度の画像（画像ピラミッド）についてそれぞれ HOG（Histograms of Oriented Gradients）特徴量を計算し、HOG ピラミッドを求める。
2. HOG ピラミッドに対するルートフィルタと各パートフィルタの応答を計算する。ここでパートフィルタではルートフィルタの 2 倍の解像度の画像に対する応答を計算する。
3. 各パートフィルタの応答から移動コストを減算した値を最終的なパートフィルタの応答とする。
4. 全てのフィルタの応答の和を取り、評価関数を計算する。
5. 評価関数の値の大きい場所が物体として検出される。

従来の物体検出手法は物体について各パーツの位置が固定のため、物体の姿勢が変化した場合に正しく検出できないといった問題があった。本手法は各パーツの位置についてある程度可変であるため、物体の姿勢変化について頑強であるという利点を持っている。

この手法はマンガ画像に対しても有効であり、HOG 特徴量と SVM を用いた手法よりも高い精度でマンガ画像内の顔領域の検出を行えることが示されている[2]。

### 3. 実験

マンガ 1 作品における 4 種類の登場人物 A~D の正面正立画像をそれぞれ 70 枚ずつ切り出し、1 種類の登場人物のみをポジティブサンプルとした場合と、4 種類の登場人物全て

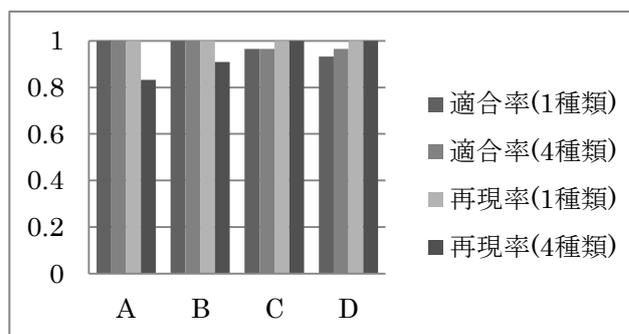


図 1 各キャラクターに対する検出結果

をポジティブサンプルとした場合について検出精度を調べた。ネガティブサンプルはマンガ画像から登場人物の顔領域を含まない領域を切り出した画像 800 枚を使用し、認識対象とする入力画像は A~D の正面正立画像をそれぞれ含むマンガ画像 30 枚ずつとした。

図 1 は登場人物 A~D についてそれぞれ該当する登場人物のみを学習した検出器と、4 種類の登場人物を学習した検出器における顔検出結果の適合率、再現率を表している。4 種類の人物を学習した検出器では、D について約 3.3% の適合率の上昇が見られた。一方、再現率は A について約 16.7%、B について約 9% の減少が見られた。この結果から、複数の登場人物を学習することで、変化が大きい顔領域を認識することが可能になったが、同時に顔以外の領域を誤検出する確率も増加したことが分かる。

### 4. まとめ

本稿では、マンガ画像からの顔検出に関して、複数の人物の顔画像を学習した場合の検出率の変化について検討を行なった。今後は、異なる作品間における Deformable Part Model の検出率の変化についても検討を行ないたい。

謝辞 本研究は JSPS 科研費 25330137 の助成を受けたものである。

### 参考文献

- [1] P. Felzenszwalb, D. McAllester, D. Ramanan, "A Discriminatively Trained, Multiscale, Deformable Part Model," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008.
- [2] H. Yanagisawa, D. Ishii, H. Watanabe: "Face detection for comic images with deformable part model," The 4th International Workshop on Image Electronics and Visual Computing 2014 (IEVC2014), 4A-1, Oct. 2014.