

画像処理による手話認識のための手形状識別

Hand Shape Classification for Sign Language Recognition by Image Processing

山田寛¹
Yutaka Yamada

島田伸敬¹
Nobutaka Shimada

白井良明¹
Yoshiaki Shirai

立命館大学 理工学研究科¹

Ritsumeikan University, The School of Science and Engineering

1 まえがき

画像処理による手話認識は、認識に必要な特徴を抽出する必要がある。手の形状特徴として、手領域の主成分スコアを用いる手法が提案されている [1]。しかし、手話では顔と手や手同士の隠蔽がしばしば起こり、正確な手領域が得られず特徴が不正確となる。さらに、照明変化によっても正確な手領域が得づらい。

そこで、隠蔽にも頑健な手の形状を取得する方法を提案する。エッジベースの特徴で局所的な照明変化に強い HOG 特徴 [2] を入力特徴として識別を行う。識別に効果的な特徴を使うため Boosting 学習によって識別器を作成する。

2 手の形状の識別

手領域に対し、手の形状毎に作成した識別器にかける。

2.1 識別器

識別器を AdaBoost により構築する。識別器 $H(x)$ を下式に示す。

$$H(x) = \sum_t \alpha_t h_t(x) \quad (1)$$

ここで x は入力、 α_t は t 番目の弱識別器 $h_t(x)$ の重みである。 $h_t(x)$ は弱識別器の候補 $g_i(x)$ の内、最も識別率の良いものとする。 $g_i(x)$ は下式により定義される。

$$g_i(x) = \text{sign}\{\theta_i(x_i - \text{th}_i)\} \quad (2)$$

ここで、 θ_i は 1, -1 のいずれかの値で閾値以上か以下かを決定する。 x_i は x の i 次元目、 th_i は閾値を表す。閾値は学習サンプルの x_i を最も識別できる値を設定する。

学習時は、まず入力ベクトルの要素数分 $g_i(x)$ を作成し、それぞれの識別率を求める。次に識別率が最も良いものを弱識別器として $H(x)$ に追加する。次に、各サンプルに対する重みを更新し、同様の作業を繰り返す。 $H(x)$ の誤識別率が 10^{-6} より小さくなれば弱識別器の追加を終了する。

識別時は、スコアが最大かつ、正の識別器に対応する手形状を識別結果とする。最大スコアが正でない場合は手が見つからなかったとする。識別器への入力はエッジベースの特徴である HOG 特徴 [2] を用いる。

2.2 HOG 特徴

HOG 特徴は、局所領域における輝度の勾配方向のヒストグラム化した特徴ベクトルである。まず輝度の勾配画像を求める。次に勾配画像から勾配ヒストグラムを 6×6 ピクセルの小領域 (セル) 毎に作成する。勾配方向

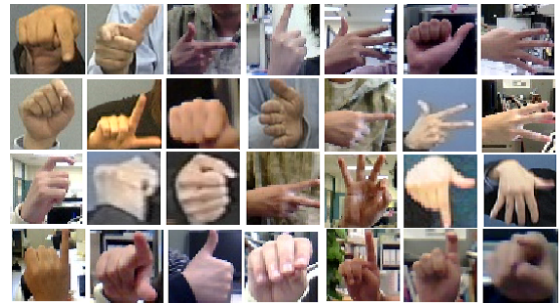


図 1 分類する手の形状

表 1 実験結果

TP: 認識成功 FN: 正解クラスが負のスコア
FP: 正解クラスが正のスコアだが 2 位以下

	TP	FP	FN
隠蔽なし	0.888	0.011	0.101
隠蔽あり	0.823	0.014	0.163

は 0° - 360° であるが、向きを考慮する必要がないので 0° - 180° とする。 20° ずつに分けるため 9 段階の勾配方向ヒストグラムとなる。作成したヒストグラムを 3×3 セル単位で正規化を行う。局所領域でヒストグラム化するため照明や影に影響を受けにくい特徴となる。

3 識別実験

提案手法の有効性を確かめるため識別実験を行った。図 1 に示す 28 種類の手形状の識別器を作成し、手話単語に対して手の形状の識別を行った。各識別器はポジティブサンプル 300 枚、ネガティブサンプル 600 枚を用いて学習をした。今回の実験では識別器の有効性を確認するため、手領域は手動で取得した。

実験結果を表 1 に示す。隠蔽なしと隠蔽ありで結果に大きな差が出なかった。これより、提案手法が隠蔽にも頑健であるといえる。

4 あとがき

HOG 特徴によって手形状の識別をした。今後はこの識別器を用いて手話単語の認識を目指す。

参考文献

- [1] 柳哲, 柳生雄午, 徳田恵一, 北村正: 手の動作と形状を用いた HMM 手話認識; 電子情報通信学会総合大会講演論文集, D-12-119, pp.285, (2004)
- [2] N. Dalal, B. Triggs and C. Schmid: "Histograms of Oriented Gradients for Human Detection"; CVPR, vol.1, pp. 886-893, (2005)