

SVM と k-NN を組み合わせたジェスチャ認識手法の提案

1. はじめに

カメラ等を用いてジェスチャを認識する技術は、ゲームやインタフェースなど幅広い分野に応用されている。機械学習手法としては SVM が有名であり、ジェスチャ認識以外の分野[1]でも利用されているが、特にジェスチャ認識においては十分な学習データを用意することが難しい。そこで、本研究では SVM と k-NN を併用することで未学習のジェスチャにも対応できる認識手法を提案する。

2. 従来の機械学習手法とその問題点

SVM(support vector machine)は、各クラスからのマージンが最も大きくなるように境界線を設定する(図 1)。安定かつ高速な識別が可能であるが、境界のどちらとしか判別できず、negative データも予め学習させる必要がある。

k-NN(k-nearest neighbor)は入力から k 個の近傍点を求め、その多数決で判定を行う(図 2)。処理が単純ではあるが、距離計算のコストが大きく、学習の密度にも影響される。

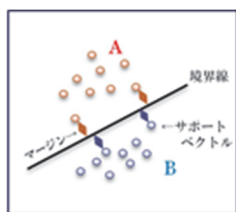


図 1. SVM の判定方法

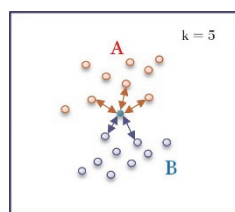


図 2. k-NN の判定方法

3. SVM と k-NN を併用したジェスチャ認識

あらゆるジェスチャを想定し、予め認識対象外のクラスも含めて十分に学習させることは難しい。そこで、SVM と k-NN を組み合わせた認識手法を提案する。それぞれの判定結果(ある認識対象のジェスチャ、もしくは認識対象外)をもとに次の 5 つの場合に分類する。

- ① SVM と k-NN が同じ判定の場合：最近傍点までの距離に閾値を設け、閾値以上の場合は対象外のジェスチャとする。閾値未満では SVM と k-NN の判定結果を採用する。
- ② SVM と k-NN が異なる判定の場合：近傍点の所属クラスを求め、SVM が判定したクラスの割合が 4 割以上ならその結果を採用し、4 割未満なら k-NN の結果を採用する。
- ③ SVM のみが対象外の判定の場合：最近傍点までの距離が閾値未満の場合には k-NN の結果を採用する。閾値以上の場合には SVM の結果を採用し、対象外と判定する。
- ④ k-NN のみが対象外の判定の場合：信頼性に優れている SVM の判定結果を採用する。
- ⑤ SVM と k-NN が対象外の判定の場合：SVM と k-NN の判定結果を採用し、対象外のジェスチャと判定する。



図 3. ジェスチャ例

4. 性能評価実験

性能評価のため、SVM、k-NN、SVM+k-NN(提案手法)の 3 つの学習器を用いて認識実験を行った。認識対象のジェスチャ A,B,C,D(図 3 参照)と、対象外のジェスチャ E の 5 クラスを用い、学習データには各クラス 100 回ずつ計 500 回を使用した。テストデータ 1 として各ジェスチャ A~D を 50 回ずつ計 200 回用意し、テストデータ 2 として各ジェスチャに近い未登録のジェスチャ A', B', C', D' を 50 回ずつ計 200 回用意した(対象外の E と判断されれば正解)。識別に利用する特徴空間は 72 次元(=3 次元座標×3 点(肩・肘・手)×両腕×4 種類の特徴)であり、各点の移動幅、前半の動き、後半の動き、各点の平均の 4 種類の特徴を用いた。認識対象ジェスチャの認識率はやや下がったものの(表 1)、未登録のジェスチャに対しては認識性能が向上したことが確認できた(表 2)。

表 1. テストデータ 1(認識対象ジェスチャ)の認識結果

ジェスチャ	SVM	k-NN	SVM+k-NN
A	100%	100%	90%
B	96%	100%	88%
C	98%	100%	82%
D	98%	98%	90%

表 2. テストデータ 2(未登録ジェスチャ)の認識結果

ジェスチャ	SVM	k-NN	SVM+k-NN
A'	94%	2%	98%
B'	96%	6%	98%
C'	92%	28%	96%
D'	76%	4%	82%

5. おわりに

本研究では SVM と k-NN を組み合わせたジェスチャ認識手法を提案した。単体の学習器に比べて認識率は若干落ちているが、未登録のジェスチャに対して認識性能が向上したことが確認できた。今後は判定を細分化し、さらなる正確性の確保を目指す。

[参考文献]

- [1] 若山・白井, “質問応答システムにおける詳細な質問タイプの分類”, 自然言語処理研究会報告, 2014.