

ランキング学習を用いた手指ジェスチャ識別手法に関する研究

Learning to Rank for FMCW based Hand Gesture Recognition

松本 笙平[†] 古月 夢媛[†] 木村 駿太[†] 三木 大輔[†]

Shohei MATSUMOTO[†] Yue FURUZUKI[†] Shunta KIMURA[†] Daisuke MIKI[†]

[†] 千葉工業大学情報科学部情報工学科

[†] Department of Computer Science, Faculty of Information and Computer Science,
Chiba Institute of Technology

1. はじめに

非接触型操作デバイスは、画面に触れることなく入力が可能であることから、感染症のリスクを軽減できるため衛生的であり、工場や医療現場などへの活用が期待されている。既存の映像解析を用いたジェスチャ認識手法の課題として、撮像機器の前に遮蔽物がある場合、適切に認識できないことが挙げられる。一方で、ミリ波レーダを用いたジェスチャ認識手法は、遮蔽物を透過してジェスチャ認識を行うことができる。教師あり学習では時系列データのジェスチャ動作フレームにアノテーションを付与する作業を要するが、それは煩雑な作業であるため教師あり学習を適用することは現実的ではない。そこで、本研究では弱教師ありデータおよびランキング学習[1]から案出された損失関数を用いることでジェスチャの認識および大きさの異なるジェスチャを識別するためのジェスチャ識別手法を提案した。

2. 方法

本研究では、ランキング学習から案出した損失関数[2]を用いることでジェスチャの検出を行う。まず、それぞれのデータ集合は、 n 個の要素の中に少なくとも一つ以上の要素の教師信号が正であるデータ V_p の集合 B_p 、およびすべての要素の教師信号が負であるデータ V_n の集合 B_n に分割される。正のデータ V_p を DNN モデルで学習させた出力 $f(V_p^i)$ が負のデータ V_n を学習させた出力 $f(V_n^i)$ より高くなるために

$$L = \max\left(0, 1 - \max_{i \in B_p} f(V_p^i) + \max_{i \in B_n} f(V_n^i)\right) + \lambda \quad (1)$$

を最小化するように DNN モデルのパラメータを最適化した。ここで λ は

$$\lambda = \phi_1 \sum_i^{(n-1)} (f(V_p^i) - f(V_p^{i+1}))^2 + \phi_2 \sum_i^n f(V_p^i) \quad (2)$$

で表される正則化項であり、DNN モデルの学習用データへの過学習を防ぎ学習を安定させるために導入した。

複数の大きさの異なるジェスチャの識別は、前述した単一クラスの分類を k クラスだけ並行して行うことで実現される。本研究では、データ集合をジェスチャの大きさの程度に応じた順に並べて、隣り合う大きさのジェスチャとの損失を最小化することにより DNN モデルを最適化した。ここで、正のデータの集合 B_p がジェスチャの数 k だけ必要となるため、正の集合は $\{B_1, \dots, B_k\}$ と表される。これらと負の集合 B_n を合わせた $k+1$ の学習を行うための損失関数は次式のように一般化できる

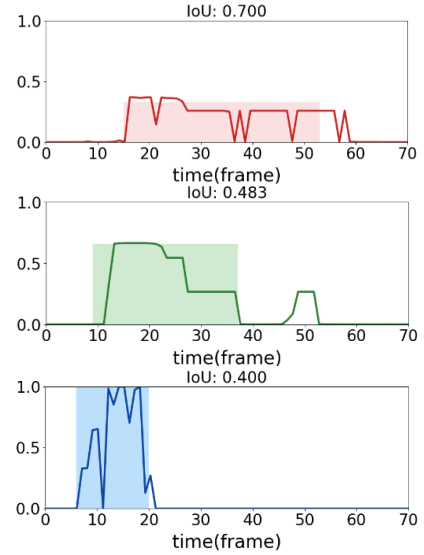


図1 遅く引く、中速で引く、速く引くジェスチャの識別結果

$$L = \sum_j^k \max\left(0, \frac{1}{k} - \max_{i \in B_j} f(V_j^i) + \max_{i \in B_{j+1}} f(V_{j+1}^i)\right) + \lambda. \quad (3)$$

3. 結果および考察

図1に大きさの異なる3種類のジェスチャに対するDNNモデルの出力を示す。ここで、網掛け部分は期待される出力と時間を表している。学習したDNNモデルの出力が期待されている出力に近い値であることから異なる大きさのジェスチャを識別可能であることが示された。さらに、ジェスチャが大きくなるにつれて出力が大きくなる傾向が確認されたため、提案した損失関数を適用したDNNモデルは異なる大きさのジェスチャを学習可能であると考えられる。

4. まとめ

本研究では、弱教師データを用いてジェスチャを行われた時間にアノテーションを付与することなく、異なる大きさのジェスチャを識別するためのDNNモデルの学習手法を提案した。実験では、3種類の大きさの異なるジェスチャの識別が可能であることが示された。

参考文献

- [1] Liu, Tie-Yan and others. Learning to rank for information retrieval. In Information Retrieval, Vol.3, No.3, p.225-331, 2009.
- [2] Sultani, Waqas and Chen, Chen and Shah, Mubarak. Real-world anomaly detection in surveillance videos. In Proceedings of the IEEE conference on computer vision and pattern recognition, p.6479-6488, 2018.