

Human Activity Action Recognition Approach

Leonardo Queiroz, Xingdong Cao

I. INTRODUCTION AND OBJECTIVES

One of the many applications for a smart system is the ability to provide an automated assessment of health. In the current aging population, “the challenges of maintaining mobility and cognitive function make it increasingly difficult to remain living alone therefore forcing many people to seek residence in clinical institutions”.

II. FRAMEWORK

This section describes the proposed approach of using only skeleton points in recognizing a human activity. Fig. 1 illustrates the overall design of the recognition system. The system consists of 4 main components: database, skeleton parser, classifier, and output.

A. Databases

B. Skeleton Parser

C. Classifier

In this project, we use two kinds of classifiers, support vector machine(SVM) and random forest. A SVM classifier projects the data onto a higher dimension so that it is possible to divide the data into separable groups[1]. Random forest generates many classifiers and aggregates their results[2].

1) Training: We have extracted two sets of features, the first set contains 44 subjects and their corresponding estimated breathing rate, apnea and motion difference features, the second set contains the same 44 subjects and face position information of each frame of the video. We will train the classifiers using these two sets separately and together, to see if there is any improvement. The labels we are to predict are Motion and Face. Motion Label has 3 classes, no moving, one side moving, and both sides moving. Face label has 2 classes, dropping face and no dropping face.

2) Testing: Due to the lack of data, we leave 1 sample as testing set, use the other 43 samples as training set, and iterate this process for 44 times, to get to average of accuracy.

D. Performance Metric

III. EXPERIMENTAL RESULTS

The results are shown in Table 1, 2 and 3. Table 1 is the result by using only feature set 1, Table 2 is the result by using only feature set 2, and Table 3 is the result by using the combined feature set. By using each feature set, we use

Table 1: Result using feature set 1

Acc Clf	Label	Motion	Face
SVM		17/44	30/44
Random forest		27/44	25/44

Table 2: Result using feature set 2

Acc Clf	Label	Motion	Face
SVM		17/44	26/44
Random forest		19/44	36/44

Table 3: Result using combined feature

Acc Clf	Label	Motion	Face
SVM		17/44	29/44
Random forest		26/44	36/44

SVM and random forest to predict motion class and face class.

By observing the results, We can get the following conclusions:

- 1) SVM doesn't work well in motion class prediction, the accuracy is 39%, just a little higher than random guess(33%);
- 2) Feature set 2 works well in face class prediction, as the features are more related with face characteristic;
- 3) Table 3 has almost the best result, so using combined features can help improve classifier performance.

IV. FUTURE WORK

V. CONCLUSIONS

REFERENCES

- [1] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.
- [2] Andy Liaw, Matthew Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

APPENDIX

```
# -*- coding: utf-8 -*-  
"""
```

```
Created on Thu Apr 2 21:02:25 2020
```

```
@author: xdcao@zju.edu.cn  
"""
```

```
print("Pythoncode")
```
