使用連接圖搜索演算法搭配親合矩陣改善靜態影像之人員分群準確率

指導教授:范倫達　學生: 黃梓晏

Graduate Degree Program of Cyber Security
National Yang Ming Chiao Tung University

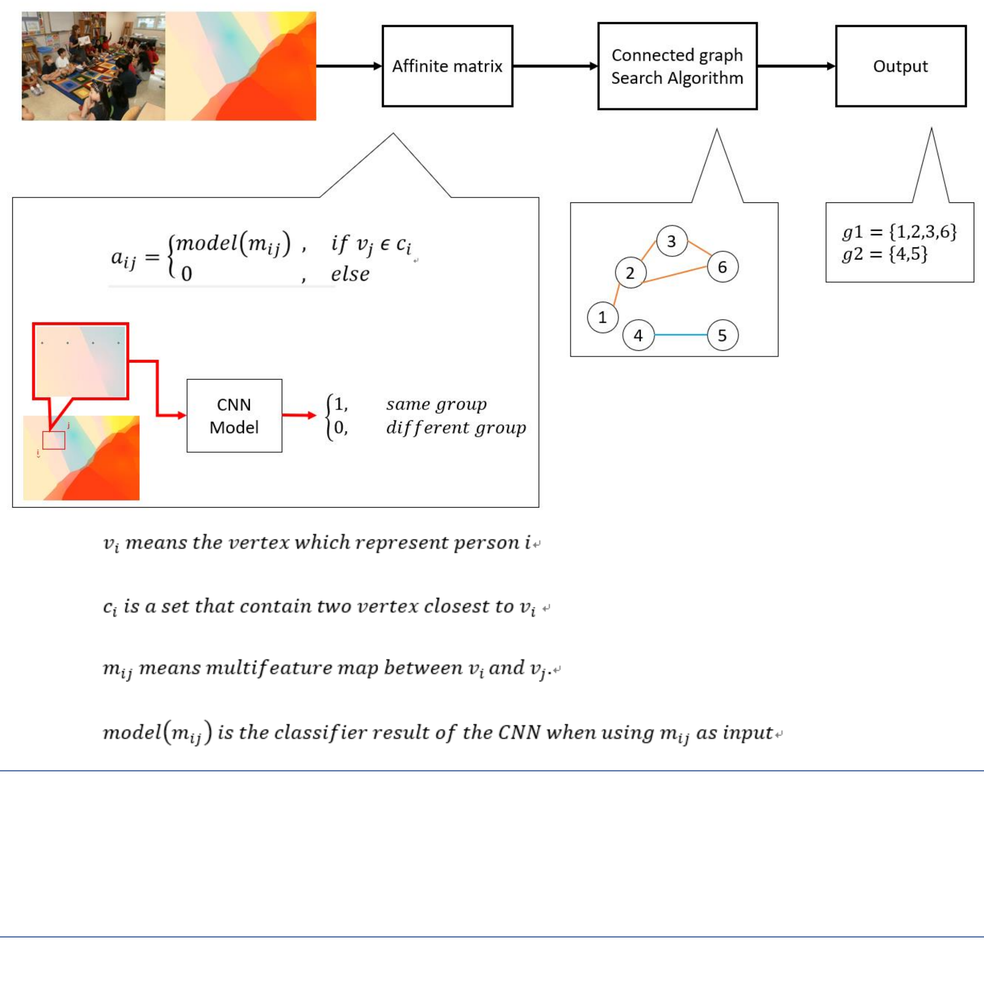
introduction

Group detection has been researched in recent years. It is an important topic, which can group people into groups by analyzing complex social relationships of people. Some studies use specific methods to understand these social relationships, such as F-formation. We can use group detection to help our lives, for example, use it to help with elevator scheduling and to optimize elevator traveling time. Or applied to robots, allowing robots to analyze social relationships and make more complex decisions.

Group detection method usually consists of many different technologies including such as human detection, action posture recognition, human orientation detection, etc. Based on these features, we can then perform group detection. The proposed clustering method applies the connected graph search algorithm with an affinity matrix for clustering. The purpose is to make the algorithm more suitable for multi-feature maps and avoid the influence of false predictions on the group detection results during clustering. This method makes the clustering with some error tolerance, thereby improves the clustering accuracy of group detection for the static images.

Method

We will first encode various features which include orientation, position, and posture into the multi-feature map. According to the formula, if two people are close, use the multi-feature map as model input. Then, use result calculated by the model as a_{ij}. If two people are far, directly let a_{ij} be 0. Finally, draw an unweighted graph according to the affinity matrix. After drawing the unweighted graph, we use the connected graph search algorithm to find the connected graph. Each connected graph represents the detected group.

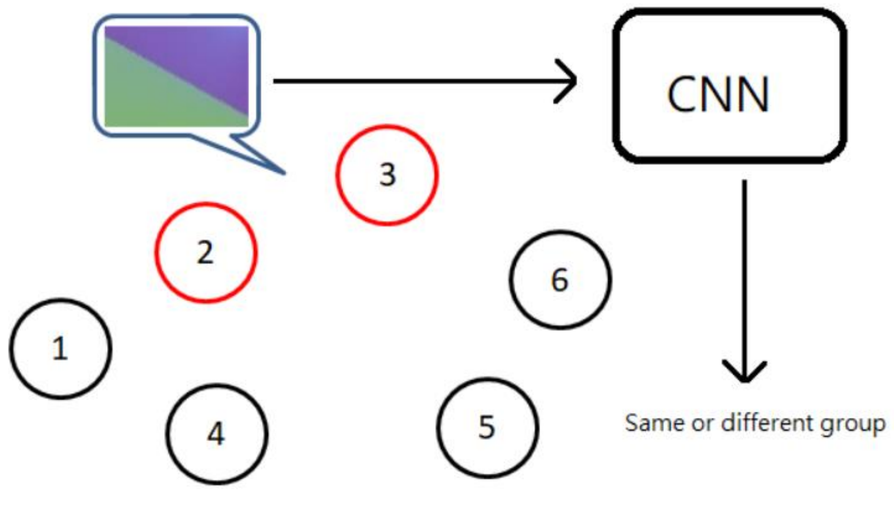


v_i *means the vertex which represent person* i

c_i *is a set that contain two vertex closest to* v_i

m_{ij} *means multifeature map between* v_i *and* v_j.

$model(m_{ij})$ *is the classifier result of the CNN when using* m_{ij} *as input*



The use of the group model here is to calculate the social relationship between two people. We cut out the multi-feature map between the two people and send it to the CNN model. The CNN model is a two-class classifier. It can predict whether $personi$ and $personj$ are in the same group based on the multi-feature map between two people. This map contains rich social relationship information to let the CNN model predict whether they are the same group.

Result

The following table shows the result that uses ground truth data. For analysis, we also list some detail of the experimental and Chen's work which proposed multi-feature maps and group people by the k-means algorithm. We can see that our result is better than Chen's work when we use the same CNN model, so we can say that our method is better than his. We also test our clustering method with different CNN models, the result is shown in the following table.

	Model	Recall	precision	F1-score
Chen's work	VGG16	0.717	0.750	0.733
This work	VGG16	0.755	0.848	0.799
This work	MobileNet	0.761	0.854	0.804
This work	ResNet50	0.802	0.872	0.835

As shown in the following table, we can compare the result between using the multi-feature map with the DS algorithm and using the multi-feature map with our proposed cluster method when we all using the same CNN model. we can see that the experimental results using our proposed method are better than the one using the DS algorithm. We can also observe that, unlike the DS algorithm, the precision of our results is usually higher than the recall.

	Model	Recall	Precision	F1-score
Chen's work	VGG16	0.717	0.750	0.733
Multi-feature map+ DS	ResNet50	0.819	0.777	0.798
This work (proposed method)	ResNet50	0.802	0.872	0.835