講演番号	007

文献番号	20225007

Utilizing Human Social Norms for Multimodal Trajectory Forecasting via Group-based Forecasting Module

Hiroaki Minoura [1)] **Tsubasa Hirakawa** [1)] **Takayoshi Yamashita** [1)] **Hironobu Fujiyoshi** [1)]

1) Chubu University
1200 Matsumotocho, Kasugai, Aichi, Japan (E-mail: himi1208@mprg.cs.chubu.ac.jp)

KEY WORDS: **Safety, Avoidance collision, Image processing, Pattern recognition,** Trajectory Forecasting [A1]

Background

Trajectory forecasting is a task to predict how and on what trajectory a target object such as pedestrian or automobile will move. While trajectory forecasting considering human-human social interactions predict the path that avoids collisions with others, people's behavior in a crowded scene is more complex due to the wide variety of behavior pattern changes that occur.

People's behavior in crowded scenes generally follows follow implicit social rules such as avoiding collisions or matching with the movement of surrounding groups, and in most scenes, the majority of pedestrians walk in groups. For accurate trajectory forecasting, it is critical to correctly model such social norms. However, the conventional works consider only limited aspects of social norms, and it is still difficult to forecast realistic and plausible trajectories. For example, they capture individual-level interactions and interactions within the group in a crowded space. In addition, while pedestrians avoid a collision by predicting the future position of people coming from the opposite direction, previous methods only make use of past trajectories and do not model interactions based on the target locations in the future.

Here we aim to capture the social norms of humans for trajectory forecasting in crowded scenes. To this end, we propose a group-based forecasting module for modeling the complex inter- and intra-group interactions between pedestrians. In addition, to reflect the future positions of other pedestrians, we introduce another prospection module. Experiments on two datasets show that our method outperforms several state-of-the-art forecasting methods.

Summary

Our finding is that it is effective for human-human interactions to use prospective trajectory information to avoid collisions with people with a high risk of collision in the future. In addition, we can see that the group-level interaction models had better accuracy than the individual models. We believe that future trajectory information and group-level interaction is important for trajectory forecasting considering human interactions in computer vision.

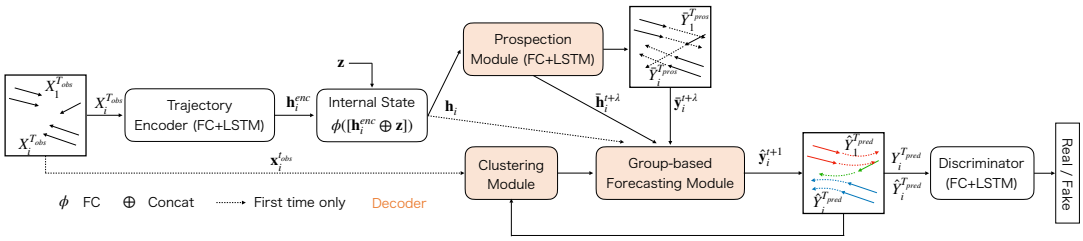

Fig.1 Proposed network architecture

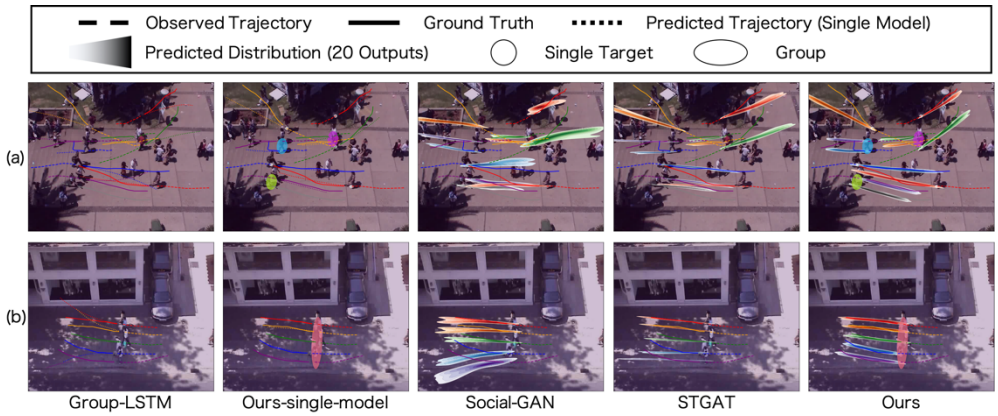

Fig.2 Visual examples of several methods in two different scenes on ETH/UCY