DeepSocNav

Social Navigation by Imitating Human Behaviors

Authors: Juan Pablo de Vicente and Alvaro Soto [1]

1: J. de Vicente and A. Soto, Dept. of Computer Science, Pontificia Universidad Católica de Chile, (jpdevicente@uc.cl, asoto@ing.puc.cl)

For social navigation first person view is more informative

Information
Frontal view of pedestrians.
Possible access to their intentions and behaviors.

Implementation
Mounted camera on robot.

Problem

Most large scale datasets are in bird's eye view perspectives

Sample	Name	Pedestrian Trajectories	Perspective
	ETH	750	Bird's eye view
	UCY	786	Bird's eye view
	SSD	5232	Bird's eye view

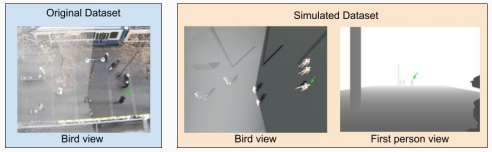
Source: OpenTraj: Assessing Prediction Complexity in Human Trajectories Datasets

Solution

Recreate bird's eye view datasets on a simulated environment

unity

Recreate bird's eye view datasets on a simulated environment

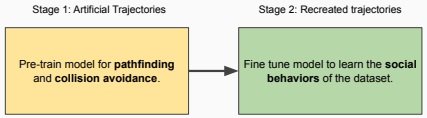


Ability to generate artificial scenes

Ability to generate artificial scenes

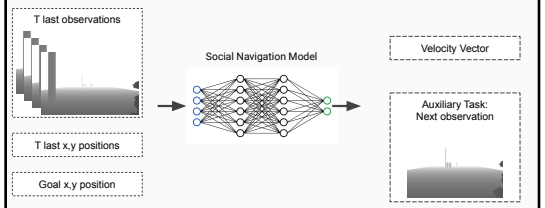
These do **not** have social behaviors!

Two stage training

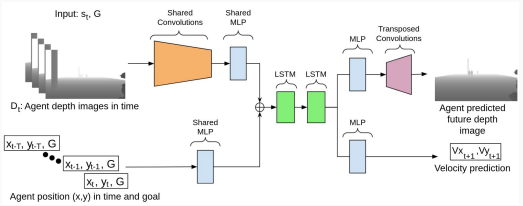


Model: DeepSocNav

Model: DeepSocNav



Model: DeepSocNav



Model: DeepSocNav

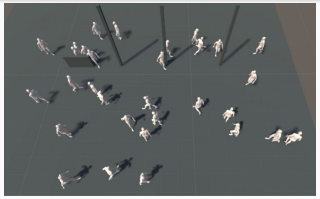
We used **T=10** observations as history

Experiments

Social trajectories: ETH Dataset

Replicated ETH and Hotel trajectories for a total of 750 social trajectories.

Artificial trajectories



We generated a total of 6000 artificial trajectories.

Baselines

- Reciprocal Velocity Obstacles (RVO)
- Social Force Model (SFM)
- NaviGAN

Ablation study

- DeepSocNav no Aux: No auxiliary task
- DeepSocNav T=1: No history
- DeepSocNav half pretraining
- DeepSocNav no pretraining

Metrics

- SocialScore: Score distances to other agents
- Collisions
- Success rate
- Average Distance Error (ADE)
- Final Distance Error (FDE)

Results

Results

Model	Social Score	ADE [m]	FDE [m]	Collisions	Target
GT	0.034	-	-	-	-
SFM	0.054	**0.20**	**0.16**	0.037	**1**
NaviGAN	-	0.43	0.74	-	**1**
DeepSocNav	**0.040**	0.24	0.34	**0.018**	**1**
DeepSocNav no Aux	0.042	0.24	0.20	0.037	**1**
DeepSocNav T=1	0.051	0.32	0.52	**0.018**	**1**
DeepSocNav no Pre-train	0.047	0.33	0.41	0.094	**1**

Conclusions

Analysis and conclusion

- Competitive performance based on an input simpler to obtain (Depth First Person View).
- Pretraining on artificial trajectories lessen the required amount of social data.
- Predicting future observations boosts navigational performance.

Future work

- Add more pedestrian datasets.
- RGB image and more detail pedestrians models to exploit intention of other agents.

Thanks!

Contact us:

Juan Pablo de Vicente at

jpdevicente@uc.cl