

Mašīnmācīšanās iekštelpu simulātoros

Guntis Vilnis Strazds

gs15014@lu.lv

Darba vadītājs: Profesors Guntis Barzdīņš, Dr.comp.sc.

Mērķis

Projekta mērķis bija padziļināti iepazīties ar metodēm un simulācijas vidēm, kas tiek izmantotas, lai apmācītu mākslīglus agentus ar elementāru sapratni par objektiem un telpisko vidi kurā tie atrodas – piemēram, lai tie varētu atbildēt uz jautājumiem vai izpildīt navigācijas uzdevumus:

- Kurā istabā ir televizors šajā dzīvoklī?
- Kādā krāsā ir krēsls dzīvojamā istabā?
- Aizej līdz galdam, pagriezies pa labi, tad pa gaiteni līdz guļamistabai

Izdarītais

Tika veikta literatūras izpēte, kā arī instalētas un izmēģinātas dažādas iekštelpu simulācijas platformas. Tika veikti praktiski eksperimenti ar mērķi atkārtot publicētos rezultātus no Das et al. 2017 [1] un Gordon et al. 2018[2].

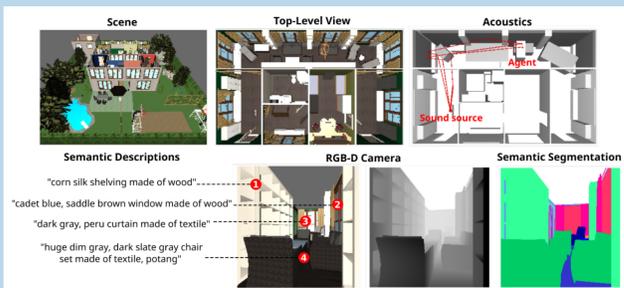
Literatūras un eksperimentu analīzes rezultātā nonācu pie slēdziena, ka būtu lietderīgi nofokusēties uz specifiskāku (ne tik vispārīgu, ne tik ambiciozu) uzdevumu par tiem, kas tiek proponēti [1] un [2], un noslēgumā proponēju definēt un risināt jaunu (nav pārstāvēts apskatītajā literatūrā) uzdevuma veidu mākslīgiem agentiem (MINOS simulātorā).

Piemēri

Skatpunkti Matterport3D vidē:



HoME platforma: [3]



Atsauces

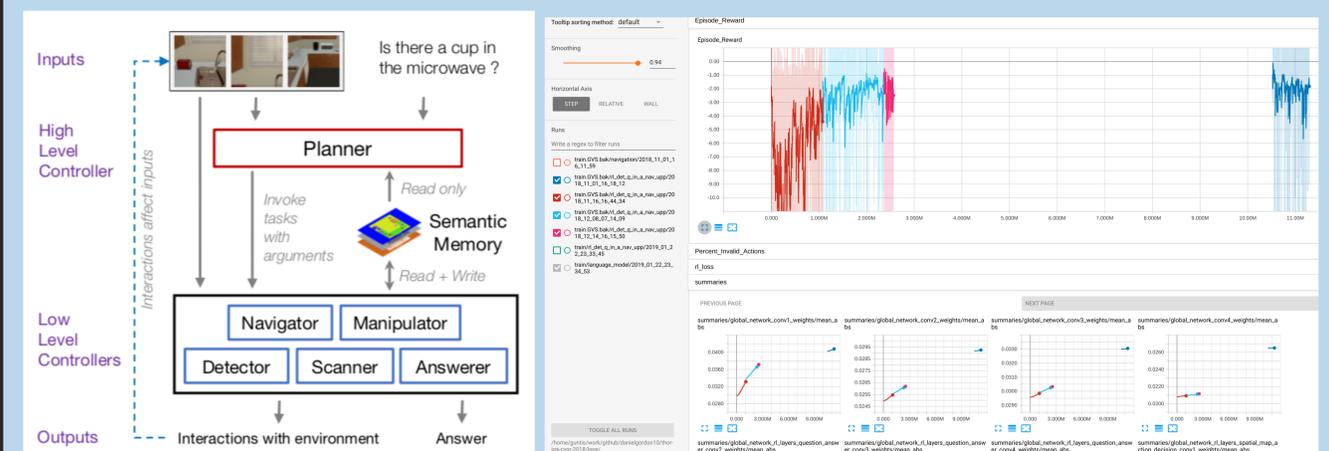
- [1] A. Das, S. Datta, G. Gkioxari et al.: *Embodied Question Answering*, IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., June 2018, pp. 2135–2144
- [2] D. Gordon, A. Kembhavi, M. Rastegari et al.: *IQA: Visual Question Answering in Interactive Environments*, Dec 2018. [Online] <http://arxiv.org/abs/1712.03316>
- [3] S. Brodeur, E. Perez, A. Anand et al.: *HoME: a Household Multimodal Environment* Nov 2017 [Online] <http://arxiv.org/abs/1711.11017>

Datu Kopas un Simulāciju Platformas

Platforma	# Vides	# Istabas	m ²	
2D-3D-S	3 ēkas	270 viewpts	6000	
SUNCG	45,622	404,058		Manually authored layouts of simulated living-spaces, containing objects instantiated from 2644 unique 3D meshes and 84 object categories.
Matterport3D	90 ēkas	2056	46,561	Photo-realistic panoramic RGB+D images captured from real locations + reconstructed textured 3D meshes. Annotated with room and object semantic segmentation data
ScanNet	1513	707	34,453	RGB+D mages from real locations. Manually annotated with semantic segmentation data, plus 3D object models
HoME	45,622	404,058		Based on SUNCG, plus physics (Bullet3) and 3D audio (EVERT), integrated with OpenAI Gym.
House3D	45,622	404,058		Based on SUNCG. Efficient rendering - up to 1800 fps at 120x90 resolution with multiple renderers sharing one GPU
MINOS	suncg+Mp3D			Either Matterport3D or SUNCG scenes at 100s of fps. Configurable environments (furniture, textures) and agents (sensors, continuous or discrete navigation)
AI2-THOR		120		Near photo-realistic 3D models, with physics. Much richer interactivity: many objects are actionable (agent can open/close them, turn them on/off)
CHALET	10 houses	58		Supports creation of custom house and room layouts. Rich interactivity (like AI2-THOR)
Gibson Env	572 buildings	1447 floors	211,000	Photo-realistic panoramic RGB+D images from real locations + textured 3D meshes. Physics engine. Explicit support for virtual-to-real-world transfer ('Goggles').

Eksperimenti, Rezultāti

Tika trenēts modelis no *IQA: Visual Question Answering in Interactive Environments*[2]^a



Rezultāti, salīdzinājumā ar publicētajiem:

Eval set	type	%	ep.len	type	%	ep.len	type	%	ep.len
Random guess	Exist	50		Count	25		Rel	50	
Repl. Unseen	Exist	72.0	354	Count	33.1	930	Rel	57.5	553
Publ. Unseen	Exist	68.5	318	Count	30.4	926	Rel	58.7	516
Repl. Seen	Exist	70.4	303	Count	36.1	940	Rel	59.5	542
Human	Exist	90	58.4	Count	80	81.9	Rel	90	43.0

Tika trenēts arī modelis no *Embodied Question Answering*[1]

Experiment	Eval set	acc 50	acc 10	d_D_50	d_D_30	d_D_10	stop_50
Repl.EQA	Val	39.3%	49.4%	2.27	1.15	-0.056	69%
Repl.EQA	Test	35.3%	46.6%	1.64	1.14	-0.85	72%
EQA[1]	Test	40.8%	50.2%	1.51	0.65	0.10	24%

^a<https://github.com/danielgordon10/thor-iqa-cvpr-2018>