Subject: Reply to doctoral thesis review

Dear Prof. Dr. Zarándy Ákos,

I would like to thank you for taking the necessary time and effort to thoughtful and thorough review of my dissertation. I sincerely appreciate all your previous valuable comments and suggestions, which helped me in improving the quality of my dissertation.

I hope you will be present at my oral defense soon.
Additionally, please find my point-by-point responses below.

Q1) Did the system propose always single steps (either 5 degrees to the left or 5 degrees to the right)?

Yes, the system suggests single steps to left or right based on multiple views, but it is not mandatory to be 5 degrees. It can be any degree and the IMU sensor with tell us about the actual orientation.

Q2) The color based CEDD method is expected to be sensitive for introducing a large foreign color object, which partially covers the target object. How do you comment on it? When you used partially occluded images, did you use the same occlusion object in the entire sequence?

CEDD is a global descriptor. It incorporates color and texture features of the image. If the new color object occluded the target object. The feature vector still contains texture and color features about the target objects but loaded with noise (originating from the foreground occlusion). The HMM is designed to solve noisy observation problems. Additionally, the main reason for using CEDD descriptor is to have a very lightweight solution, but we can use other descriptors for object recognition in our proposed approach.

I used about 100 objects for occlusion in my experiments and naturally in each individual sequence the occlusion object remained the same in the entire sequence, but with random positions.

Q3) The LSTM is designed for long sequences. Did the advantage of the LSTM come out in short (4 steps) sequences?

LSTM (Long Short-Term Memory Network) doesn't have any problem with short sequences. The shortcoming of RNN is, it cannot remember long term dependencies due to vanishing gradients. LSTM was invented to solve this problem. Additionally, LSTM cannot solve the problem of very long term dependencies, In [1], they created a new architecture to solve this problem. Generally thinking, since LSTMs can incorporate CNN structures, they should not perform worse. Consequently, I thought that LSTM will work fine with sequence of length 4. Actually, I didn't test single CNNs, like in Figure 2.7 Chapter 2 in other experiments, for comparisons.

[1] Xia, Wei and Zhu, Wen and Liao, Bo and Chen, Min and Cai, Lijun and Huang, Lei, Novel architecture for long short-term memory used in question classification, Neurocomputing, 2018.

Sincerely,
Amr M. Nagy

Veszprem, 14/11/2022

Amr Abdo