Motion & Gaze Selection for Atari Games

Krishanu Agrawal
Electrical & Computer Engineering
kagarwal87@gatech.edu

Ezra Ameperosa
Electrical & Computer Engineering
eameperosa3@gatech.edu

Mudit Gupta
College of Computing
mgupta303@gatech.edu

Himanshu Vairagade
Electrical & Computer Engineering
himanshuv@gatech.edu

I. Research Goals / Problems

The primary goal is to produce a Behavior Cloning (BC) agent with better than state-of-the-art performance on playing Atari games. The agent will be trained on trajectories and frame-by-frame gaze information from human game-play on 23 Atari games [1]. The data was obtained by pausing at each frame to allow the human to look around the scene and make optimal decisions (scores comparable to known human records [1]). A secondary goal with this work is to perform a results-based analysis of game characteristics which cause a game to necessitate gaze- versus motion-based attention mechanisms.

II. Existing Literature Insufficiency

The existing works (Selective Eye-gaze Augmentation (SEA) [2] and Coverage-based Gaze Loss (CGL) [3]) use only gaze to augment the attention of the agents being trained on the Atari-HEAD dataset [1]. The SEA approach uses a recurrent gating-network which learns when to use or not use predicted gaze information. Their results show that the trained agents converged to only use the gaze information roughly 40% of the time. The CGL approach defines a custom loss function which reflects feature counting over the regions with high gaze density. The authors of the CGL approach attempted training with motion heatmaps but saw little performance increases. Both of these works made advances in taking advantage of gaze/motion information to better train an agent, but neither makes sufficient steps towards combining the information in unique proportions for different games.

III. Algorithm Approach: Mo-GaS

The proposed approach does Motion and Gaze Selection. The new process will have both predicted gaze heatmaps along with well-defined/computed motion heatmaps. Both of these heatmaps will be conditionally gated before concatenation with an embedding of the input frames. This enables the agent to efficiently learn when to use each of the available heatmaps. Furthermore, the agent will use coverage-based loss to ensure that the heatmap data is prioritized when passed through the gate. This effectively makes the agent learn when to use heatmap information, and when it does so, it should prioritize on only regions with high heatmap values.

Results favoring motion-prioritization may reveal that time spent on gaze data collection could be better spent on more effective use of easily attainable features (such as motion).

IV. Progress

A variety of resources exist on GitHub related to the Atari-HEAD dataset. There is a visualizer for seeing gaze data per frame on recorded trajectories. There is also code for the CGL [3] and SEA [2] approaches on Github. Our code will most closely resemble the SEA code given that the pytorch-based gating architecture is an essential component of our design, and the CGL function can be easily incorporated.

Both related works (SEA and CGL) train an autoencoder that takes in 4 images and predicts gaze heatmaps based on the gaze information from the Atari-HEAD dataset [1]. The pretrained models from the CGL code output seemingly random gaze, so an autoencoder was instead trained from scratch based on the SEA paper and code. Training was successful and the network output usable gaze representations (potentially not at the fidelity of the original work due to code not being immediately reproducible).

V. Remaining Work

The action prediction network needs to be worked on to reproduce the existing work which will serve as the benchmark for our proposed Mo-GaS algorithm. The midterm exam for success would be comparable performance with the reproduced SEA agent. A gating mechanism for motion heat maps needs to be developed to mirror the functionality of the gaze gating. CGL can be implemented in parallel to work with gaze heatmaps and then applied to motion once ready. The final exam of success will be an analysis on the performance gains on Atari gameplay when applying CGL to both gated gaze and gated motion. A supplementary final step would be to do a cross-game analysis on the usage of motion versus gaze information for different games.

VI. Risks / Challenges

The described code references are not immediately reproducible. Additionally, the time to train the autoencoder on one game is more than expected (>8 hours for ∼30 epochs. With these two challenges in mind, we have to budget our time carefully and consider contingencies to ensure we meet our goals. If time, our only cost, becomes a concern, the team will consider constraining the evaluation to fewer games which may weaken our cross-game analysis, but will enable us to finish by mid-December.

References

[1] R. Zhang, C. Walshe, Z. Liu, L. Guan, K. S. Muller, J. A. Whritner, L. Zhang, M. M. Hayhoe, and D. H. Ballard, "Atari-head: Atari human eye-tracking and demonstration dataset," 2019.

[2] C. Thammineni, H. Manjunatha, and E. T. Esfahani, "Selective eye-gaze augmentation to enhance imitation learning in atari games," 2020.

[3] A. Saran, R. Zhang, E. S. Short, and S. Niekum, "Efficiently guiding imitation learning agents with human gaze," 2020.