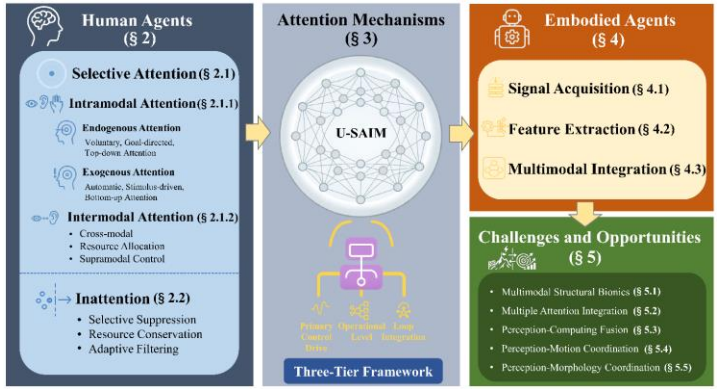
Selective Attention or Inattention? Enabling Efficient Perception in Embodied Intelligence Agents

Ruonan Xu, Bin Guo, Ke Ma, Lei Wu, Yasan Ding, Yao Jing, Zhiwen Yu

Frontiers of Computer Science, DOI: 10.1007/s11704-026-60382-7

Problems & Ideas

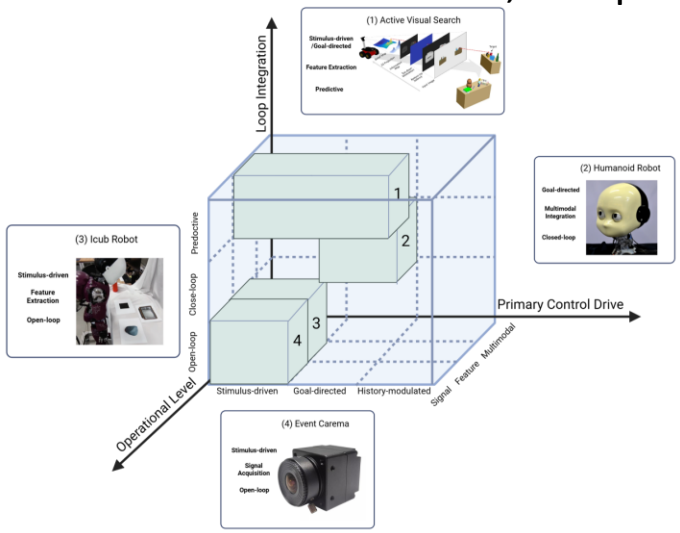
- Problems of conventional stereo matching approaches:
  - Embodied agents often process all sensory inputs uniformly, causing redundant computation.
  - Existing attention-inspired methods are fragmented across sensing, feature extraction, and multimodal fusion.
  - The positive role of inattention in suppressing irrelevant information remains underexplored.
- Ideas:
  - Treat embodied perception as a resource-aware process combining selective enhancement and active suppression.
  - Propose U-SAIM to unify selective attention and inattention for efficient embodied perception.



Overall structure of attentional embodied perception, progressing from human attention mechanisms to U-SAIM, embodied implementations, and future challenges.

Main Contributions

- Contributions:
  - A systematic taxonomy of human attention mechanisms is established, covering intramodal attention, intermodal attention, and inattention.
  - A unified conceptual model, U-SAIM, is proposed to couple selective attention and inattention for resource-efficient embodied perception.
  - A three-tier framework is developed to map biological attention principles to embodied systems across signal acquisition, feature extraction, and multimodal integration.
  - Key challenges and future opportunities are identified, including multimodal structural bionics, multiple attention integration, perception–computing fusion, perception–motion coordination, and perception–morphology coordination.



Representative embodied attention systems organized according to primary control drive, operational level, and loop integration.