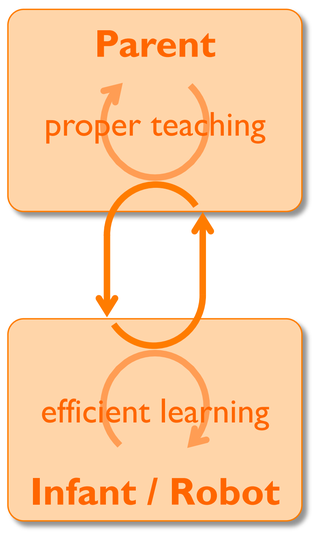
Bottom-up Mechanism to Extract Key Actions from Parental Task Demonstration

Yukie Nagai (Osaka University)

Motivation

- Parents exaggerate their body movement when teaching tasks to infants. [Brand et al., 2002; Rohlfing et al., 2006]

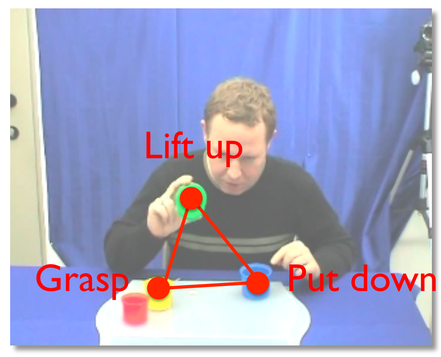


Open Questions

- Does parental teaching facilitate infant/robot's learning?
- What can infants/robots learn from parental demonstration?

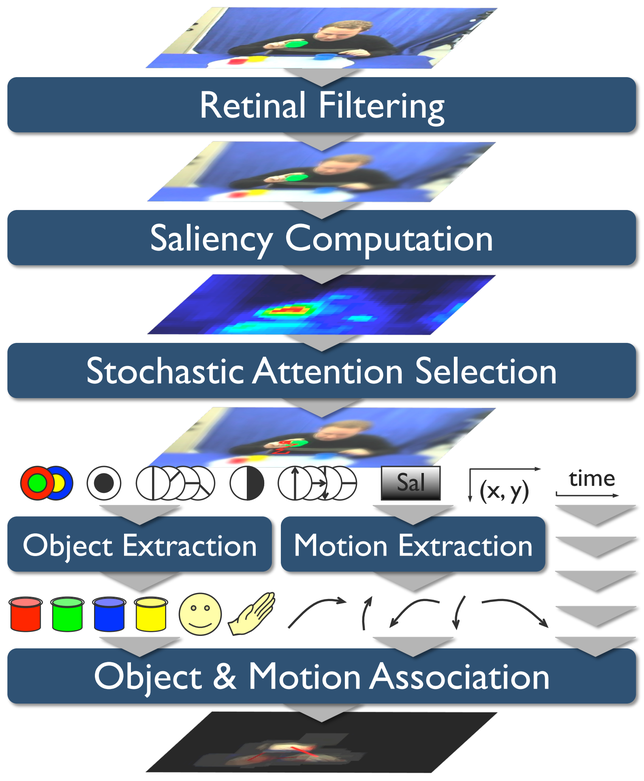
Our Challenge

- To develop a bottom-up mechanism to extract key actions from parental task demonstration
  e.x.) Cup-stacking task:
  grasp a cup, lift it up, put it down

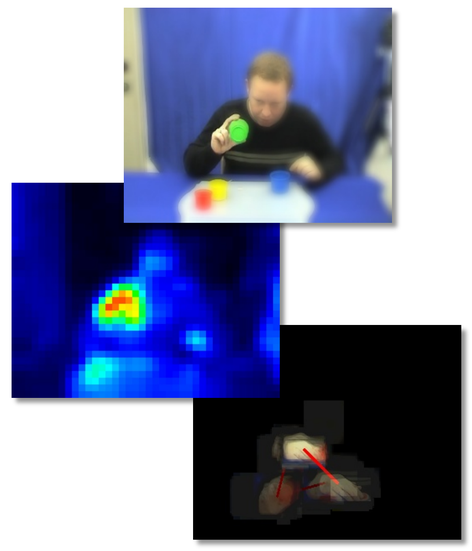


Assumption:
- Infants have *little knowledge* about context.

Bottom-up Mechanism to Extract Key Actions

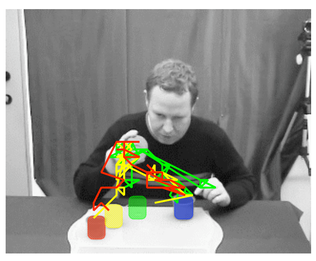


- Retinal filtering
  - High acuity in fovea whereas low in periphery
  - *Stabilize* attention to fovea
- Saliency computation [Itti et al., 1998]
  - Saliency as difference from surroundings
  - Detect likely *important locations*
- Stochastic attention selection
  - More attention shift to stronger saliency while less to weaker
  - Maintain *sensitivity* to new signals in periphery
- Object and Motion Extraction + Their Association
  - Examine *continuity* in extracted features in terms of space and time
  - Extract *key actions* by associating objects and motions

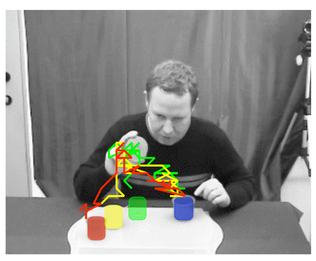


Exp. 1: Stability and Sensitivity of Bottom-up Attention

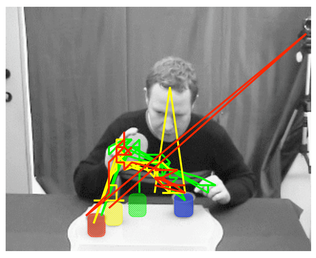
- *w/* retinal filtering + *stochastic* selection
  → Stabilize attention + maintain sensitivity



- *w/* retinal filtering + *deterministic* selection

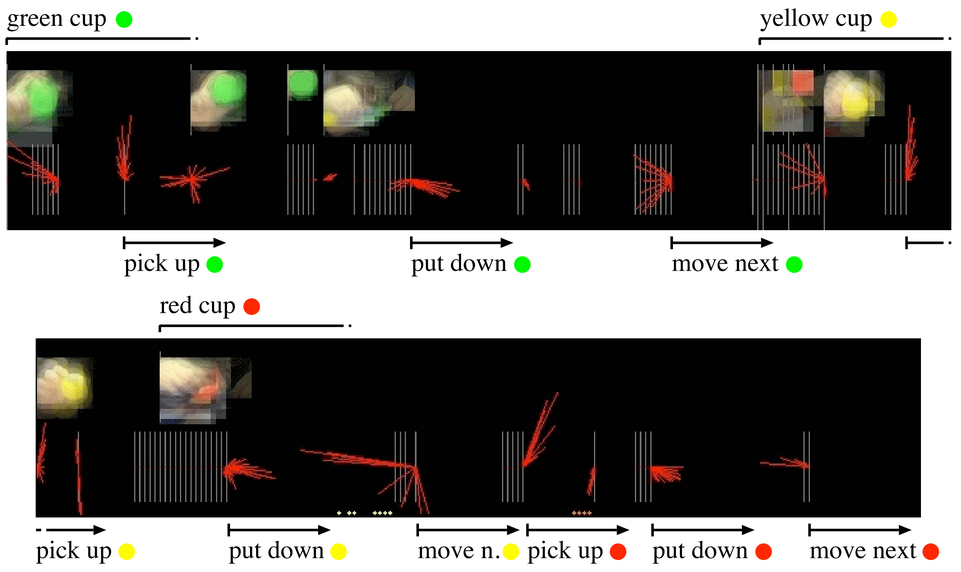


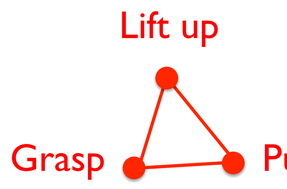
- *w/o* retinal filtering + *deterministic* selection



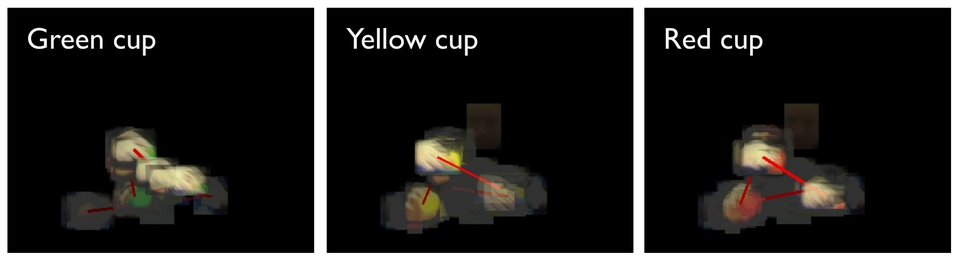
Exp. 2: Extraction of Key Actions

- Create object and motion chunks





- Associate objects and motions
  → Three key actions



Conclusion

- Examining continuity in bottom-up attention extracts key actions from parental task demonstration.
- Key actions are emphasized by parental actions (e.g., make pauses).

Contact: yukie@ams.eng.osaka-u.ac.jp (This work was done at Bielefeld University.)

References

- L. Itti et al., "A model of saliency-based visual attention for rapid scene analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 11, pp. 1254-1259, 1998.
- Y. Nagai and K. J. Rohlfing, "Computational Analysis of Motionese Toward Scaffolding Robot Action Learning," IEEE Transactions on Autonomous Mental Development, vol. 1, no. 1, pp. 44-54, 2009
- .Y. Nagai and K. J. Rohlfing, "From Bottom-Up Visual Attention to Robot Action Learning," ICDL, 2009.