Learning complex visual concepts

Difficult to detect, appear early, important for subsequent learning of agents, goals, interactions
Why are hands difficult to detect?

multiple appearances

Van Gogh
Kirchner
small and inconspicuous

Selectivity to hands appears early in infancy

using a head camera to study visual experience

“overall... hands were in view and dynamically acting on an object in over 80% of the frames”

Yoshida & Smith 2008

What makes hands learnable by humans?
Motion, hand as ‘mover’
(7 months old)

Saxe, Carey (2006) *The perception of causality in infancy*

---

Early sensitivity to special motion types

- high sensitivity to motion in general
  (detecting motion, motion segmentation, tracking)

- specific sub-classes of motion: self-motion, passive, and ‘mover’

A specific motion is highly indicative of hands
Detecting ‘mover’ events
Ullman, Harari, Dorfman (2012)

- mover event = moving image region causing a stationary region to move or change after contact
- simple and primitive, prior to objects or figure-ground segmentation
- ‘mover’ as an innate teaching signal for learning the concept of a hand

Training videos

Movies of scenes, people moving, manipulating objects, moving hands

‘Mover’ events are detected in all movies and used for training
‘Mover’ events extracted from videos

High fraction of detected mover events involve hands
(90% recall, 65% precision)

Extract many images of objects entering cells of mover events & use images to train a hand object classifier

Hand detection in still images

Trained classifier mainly detected hands in object manipulation scenes
Continued learning

Two detection algorithms:

• Detect hands with greater range of appearance by tracking

• Detect more hands by the body context

Hand by surrounding context

face → shoulder → upper-arm → lower-arm → hand

Amano, Kezuka, Yamamoto, 2004
Slaughter, Heron-Delaney, 2010
Slaughter, Neary, 2011
• infants follow the gaze of others
• starts at 3-6 months and continues to develop
• head orientation first, eye cues later
• important in development of communication & language
• modeling mainly head direction

Gaze direction

W.H. Wollaston (1824)
On the Apparent Direction of Eyes in a Portrait
Gaze cues are subtle and inconspicuous

mover events supply the teaching signal

Gaze extraction

Training  Testing

Use facial appearances and gaze directions derived from mover events as a training set to learn direction of gaze