Modeling Perceptual Attention in Virtual Humans

Randall W. Hill, Jr.
University of Southern California
Information Sciences Institute

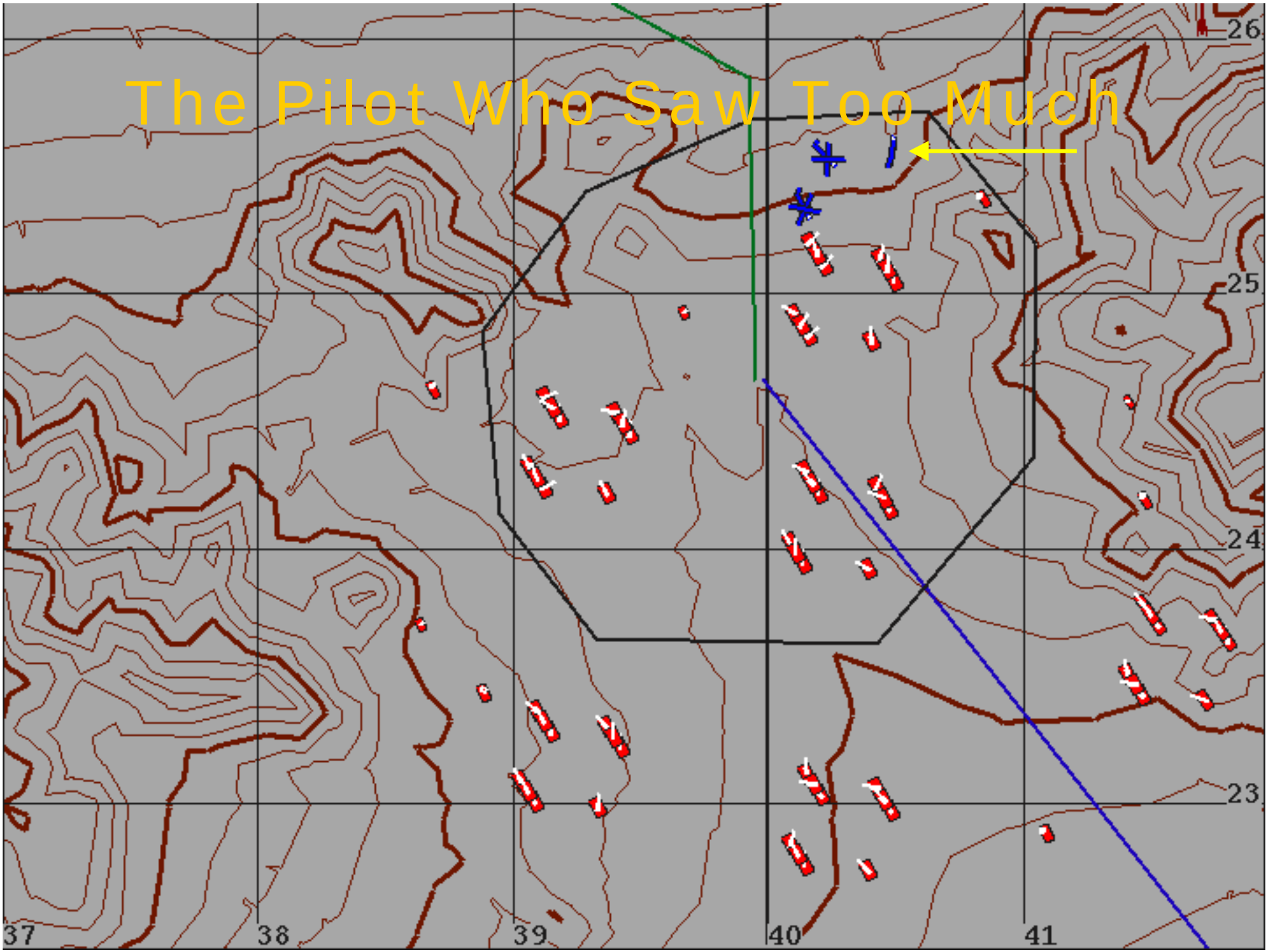
What We Want From Virtual Humans

- **Realistic behavior**
  - Perform within range of human capability
    - Superman need not apply ....
    - No drones, either!
- **Believability**
  - Does the agent's behavior enable me to believe that the agent could be a human?

Challenges to Realistic Behavior

- **Synthetic worlds are information rich**
  - 100's of other entities
  - Vehicle instruments
  - Terrain, weather, buildings, etc.
  - Communications (messages)
  - Amount of information will continue to increase ....
- **Perceive, understand, decide and act**
  - Comprehend dynamic, complex situations
  - Decide what to do next
  - Do it!


Tale of the Apache Pilot


The Pilot Who Saw Too Much
26
25
24
23
37
38
39
40
41

Roots of the Problem

- **Naïve vision model**
  - Entity-level resolution only
  - Unrealistic field of view (360^{o}, 7 km radius)
- **Perceptual-Cognitive imbalance**
  - Too much perceptual processing
  - Cognitive system needs inputs, but ...
  - It also needs time to respond to world events

How Do Humans Do It?

- **Employ perceptual attention**
  - Filter Metaphor
    - Discard excess information
  - Spotlight metaphor
    - Focus perceptual processing on limited regions / objects
  - Zoom lens or Gradient metaphor
- **Process percepts in stages**
  - Preattentive stage: segment, group, filter
  - Attentive stage: search, fixate, track
- **Control the focus of attention**
  - Goal-directed versus stimulus-driven

What exactly is attention?

- **Many metaphors for attention**
  - **How do we operationalize attention in a Soar agent?**
- **ATTENTION is at the nexus between cognition and perception**
  - Consists of a set of mechanisms
  - Spans perceptual and cognitive systems
  - Cognition orients and controls perceptual processing
  - Perception influences cognition

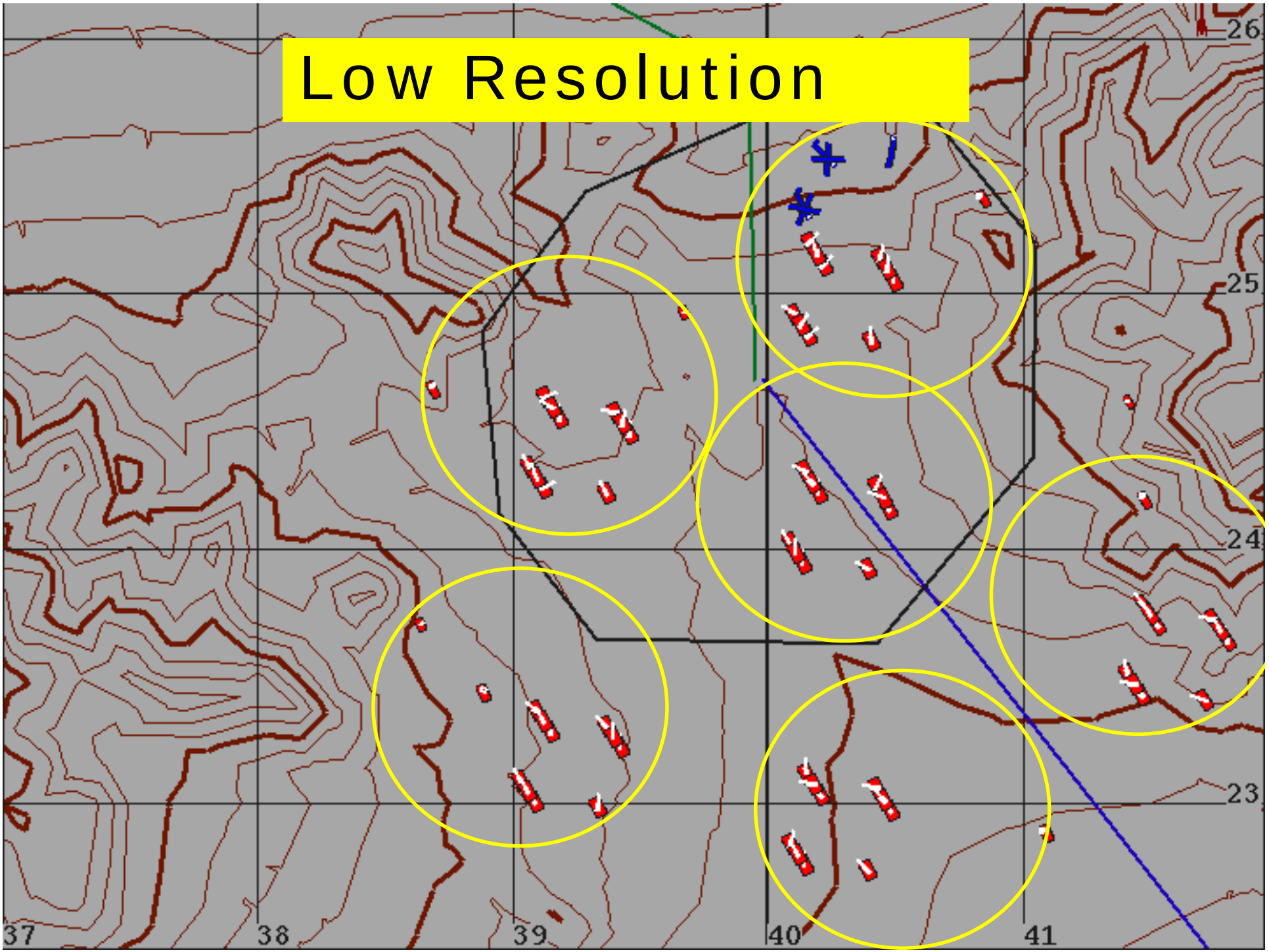
Approach

- **Create a focus of attention**
  - Apply attention mechanisms to entity perception initially
  - Incorporate filters
  - Implement a zoom lens model (covert attention)
- **Stages of perceptual processing**
  - Attention in different stages**:** preattentive & attentive
- **Control the focus of attention**
  - Goal-driven
  - Stimulus-driven

Zoom Lens Model of Attention

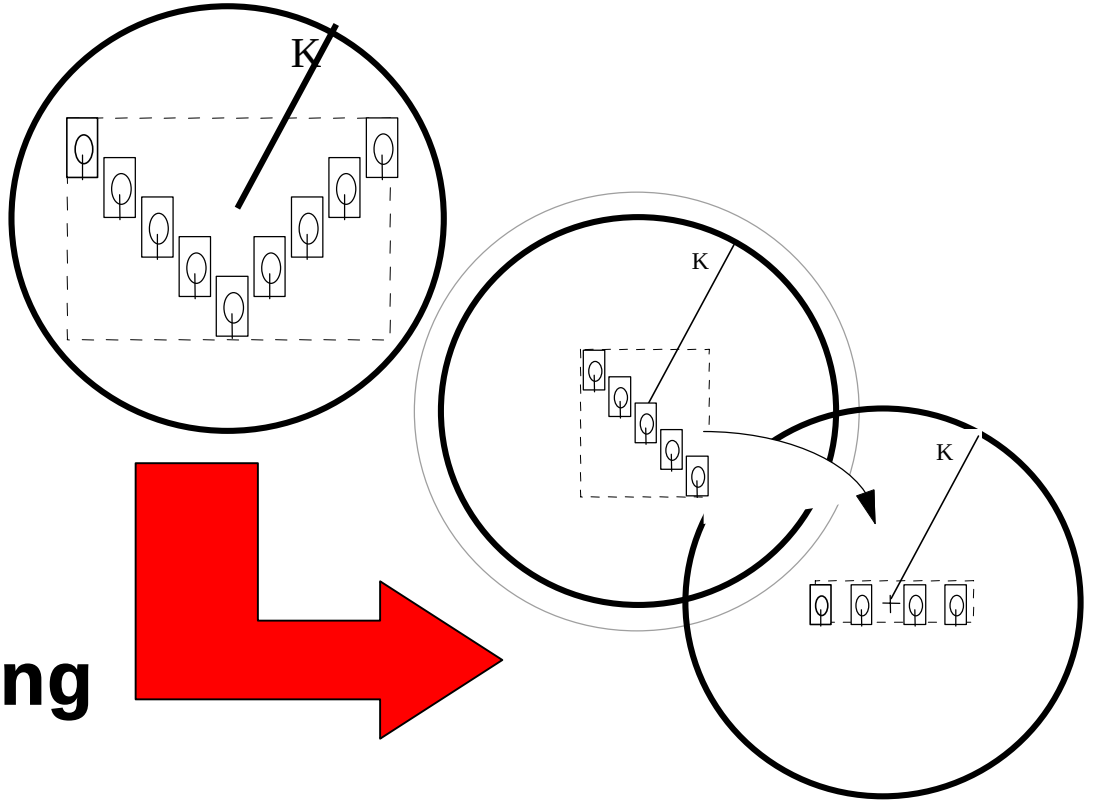
(Eriksen & Yeh, 1985)

- **Attention limited in scope**
  - Multi-resolution focus
  - Magnification inversely proportional to field of view
- **Low resolution**
  - Large region, encompassing more objects, fewer details
  - Perceive groups of entities as a coherent whole
- **High resolution**
  - Small region, fewer objects, more details
  - Perceive individual entities (e.g., tank, truck, soldier)


Low Resolution
26
25
24
23
37
38
39
40
41

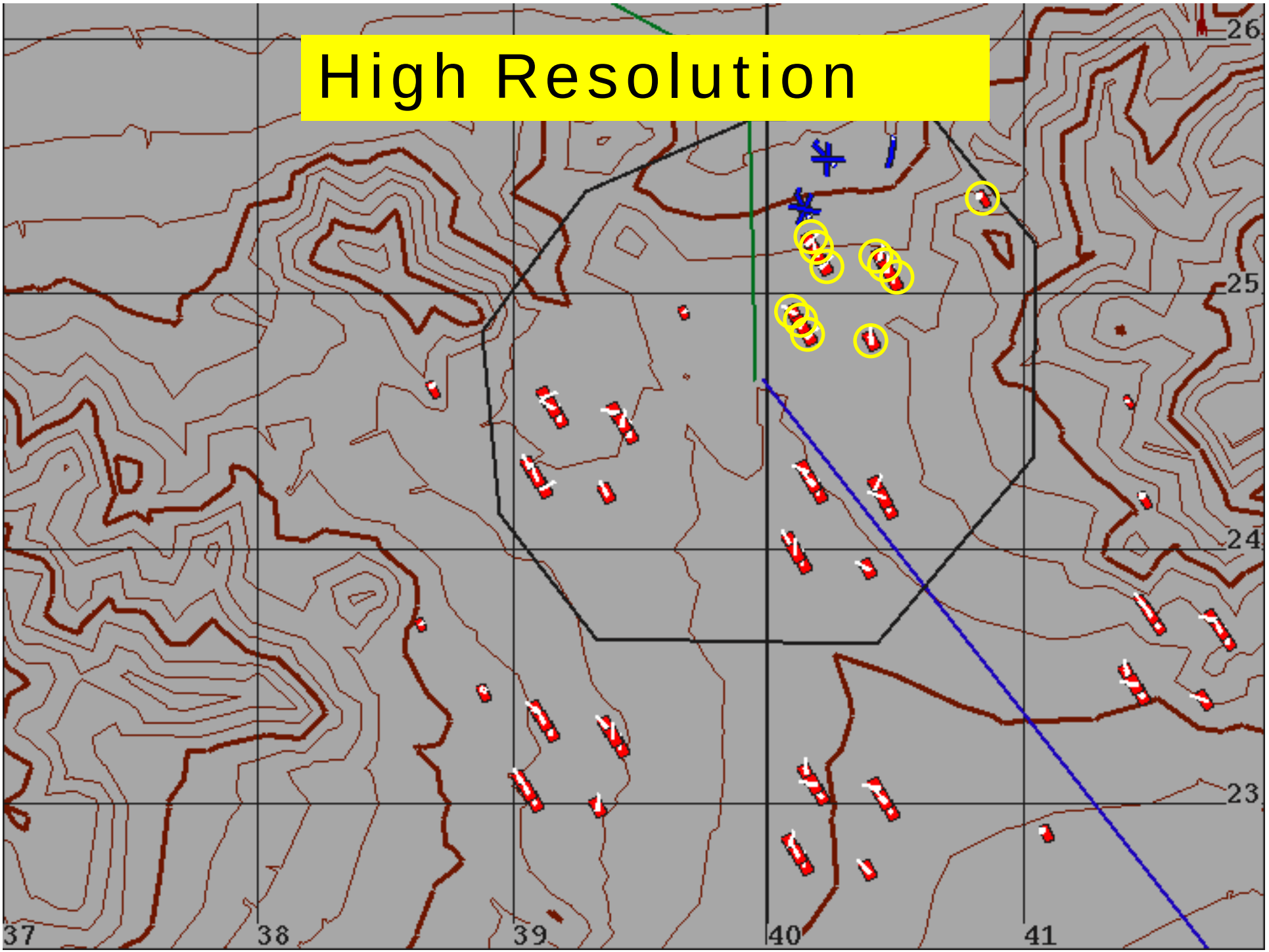
Perceptual Grouping

- **Preattentive**
- **Gestalt grouping**
  - Involuntary
  - Proximity-based
  - Other features
- **Dynamic**
- **Voluntary grouping**



Group Features

- **Quantity and composition**
- **Activity**
  - Moving
  - Shooting
- **Location**
  - Center-of-mass
  - Bounding-box
- **Geometric relationships wrt pilot**
  - Slant-range, azimuth, etc.


High Resolution
26
25
24
23
37
38
39
40
41

Entity Features

- Location (GCS)
- Speed
- Velocity
- Orientation
- Slant Range
- Force
- Object, Object Type
- Vehicle Class
- Function
- Sense Name
- Altitude
- Angle Off
- Target Aspect
- Magnetic bearing
- Heading
- Status
- Lateral Range
- Lateral Separation
- Closing Velocity
- Vertical Separation

Stages of Attention

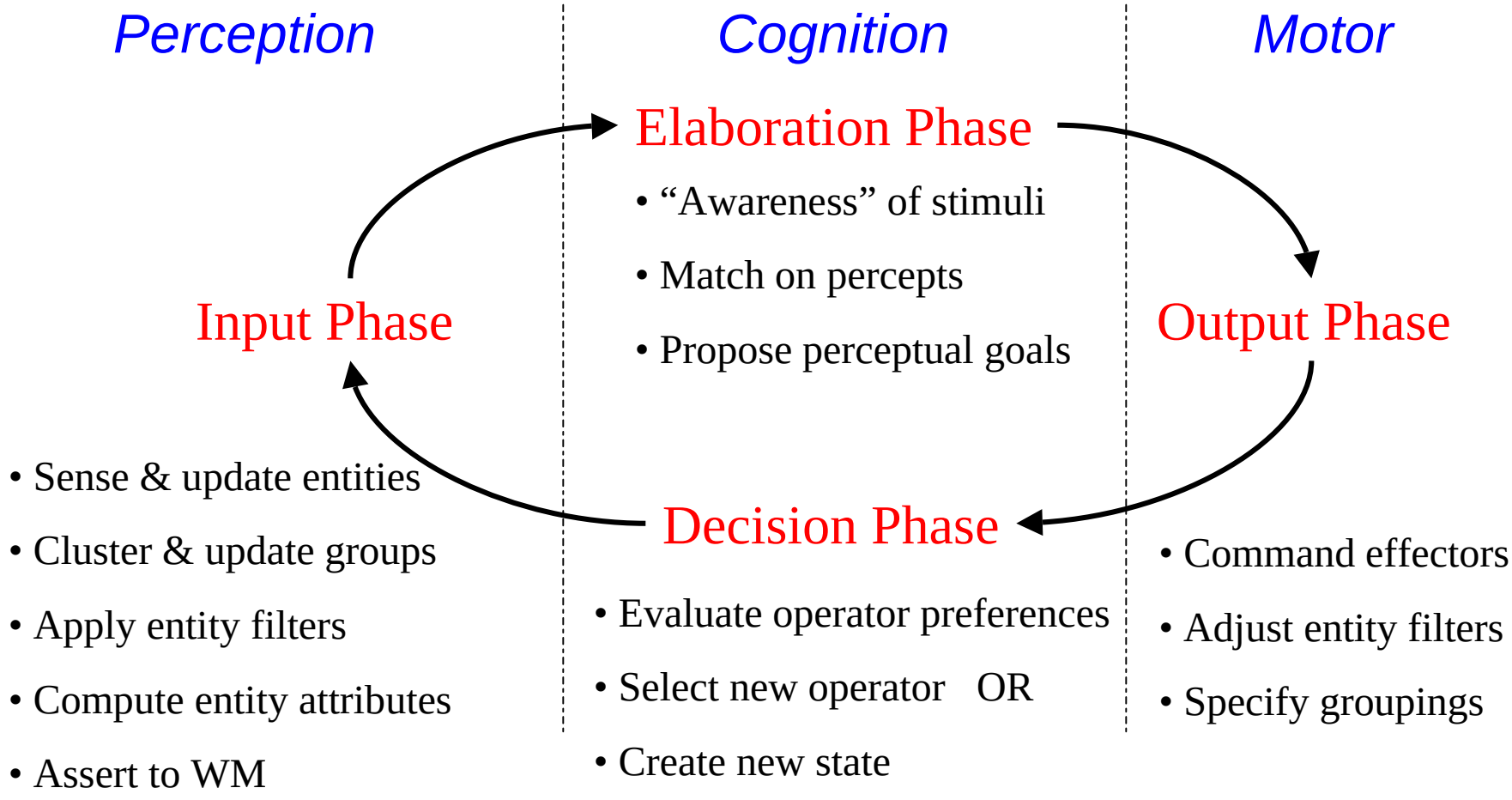
Preattentive

- Sense entities with ModSAF visual sensor
- Update location, range, and speed features of entities
- Cluster newly sensed entities into groups
- Update group features

Attentive

- Apply selective filter to entities
- Compute and assert features to Working Memory
- Match Soar productions on entity/group features

Soar Decision Cycle



Control of Attention

Goal-driven control

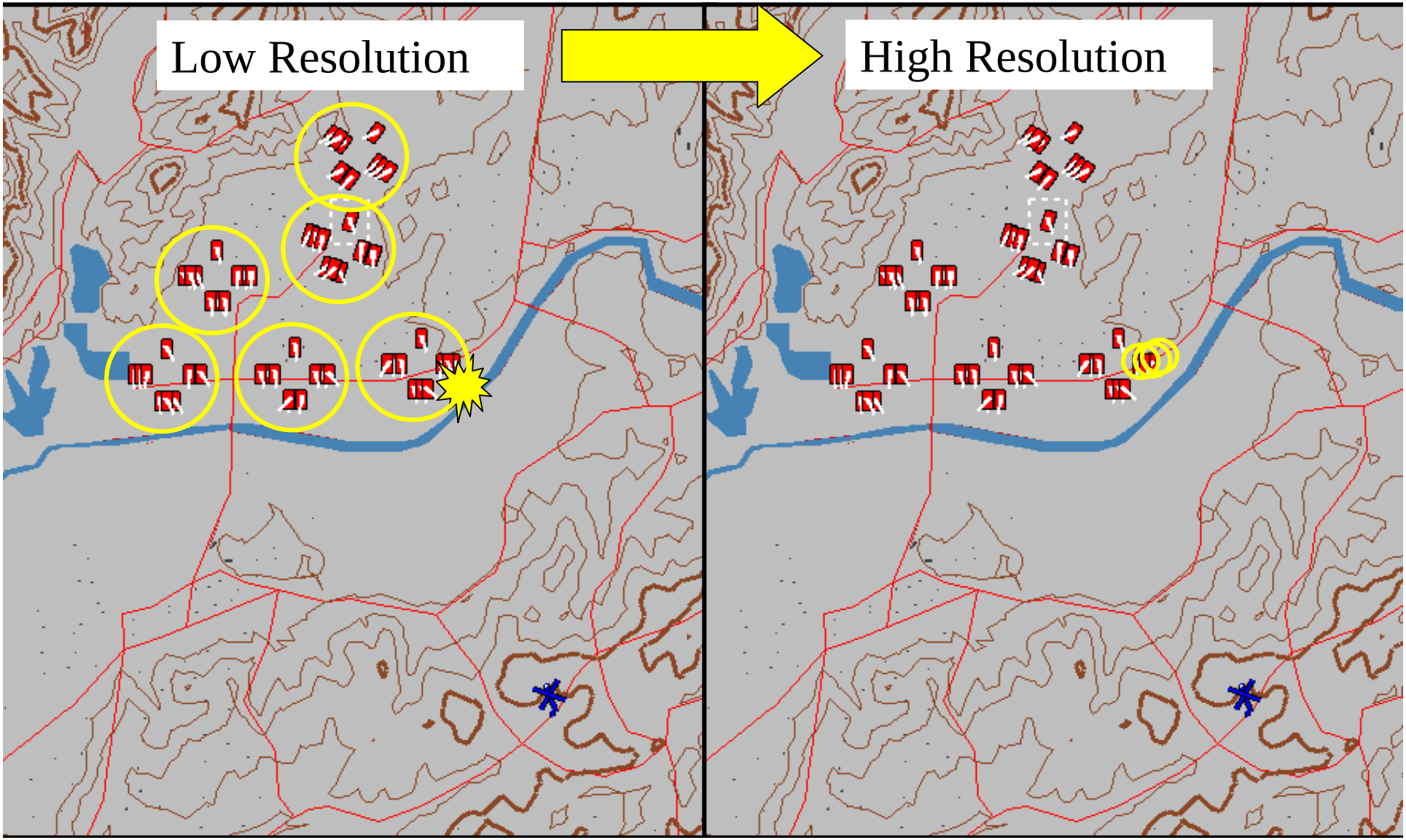
- Agent controls the focus / resolution of attention
  - Low resolution: Scouting groups of enemy; escorting group
  - High resolution: Search for air-defense entities; engage target
- Sets filters that select entities for WM

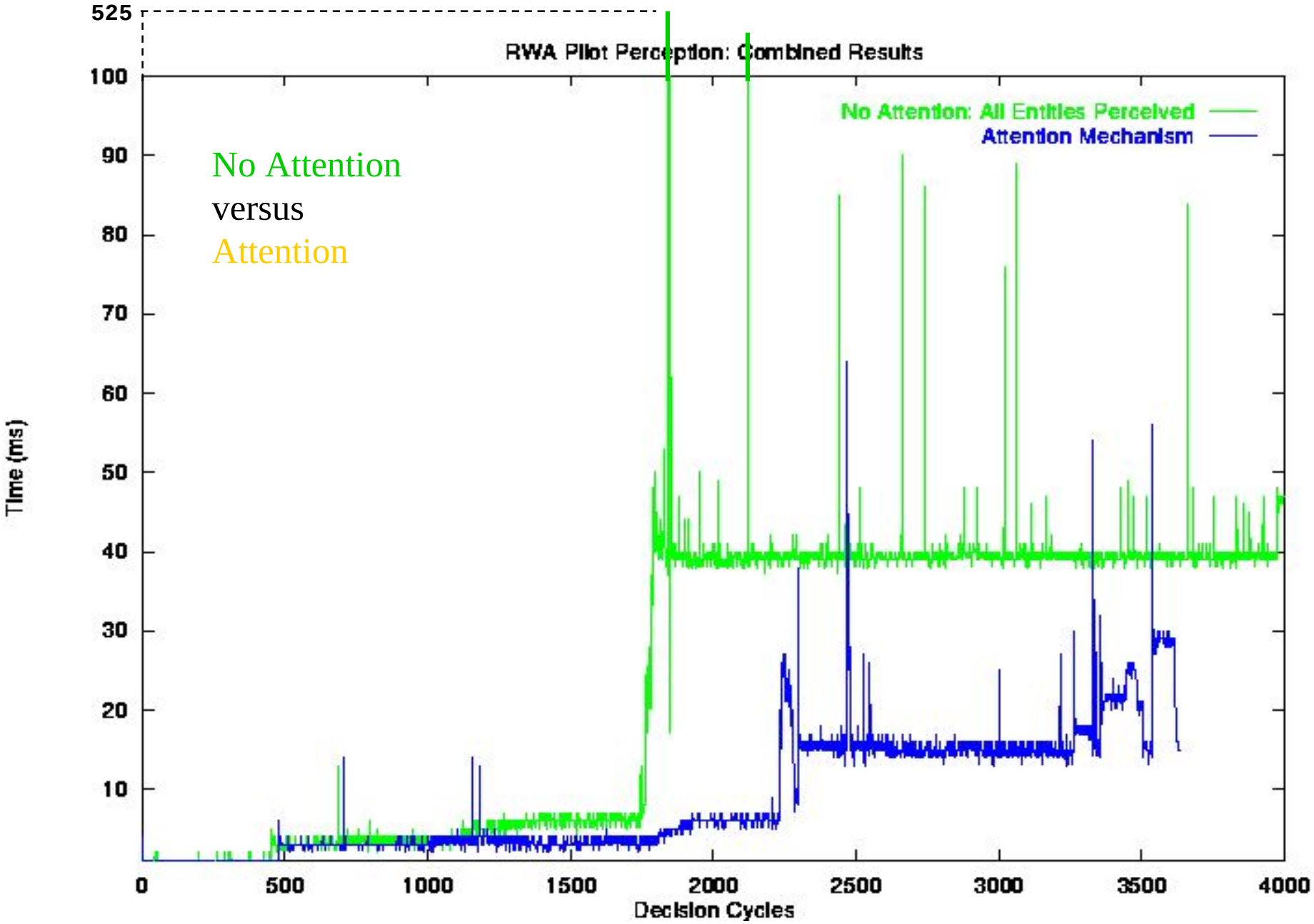
Stimulus-driven control

- Attention can be captured involuntarily by a visual event
  - Muzzle flash (luminance contrast, abrupt onset)
  - Sudden motion (abrupt onset)

Goal-driven Attention

Stimulus-driven Control




RWA Pilot Perception: Combined Results
No Attention: All Entities Perceived
Attention Mechanism
No Attention
versus
Attention
Time (ms)
Decision Cycles
525
100
90
80
70
60
50
40
30
20
10
0
500
1000
1500
2000
2500
3000
3500
4000

Summary

- **Focus of attention**
  - Filters
  - Zoom lens model
  - Production matching
- **Distributed attention over stages**
  - Preattentive grouping and filtering
  - Attentive processing
- **Controlling attention**
  - Goal-driven
  - Stimulus-driven

Current Work

- **More realistic model of vision**
  - Foveal, parafoveal, and peripheral visual fields
  - Better clustering, multi-resolution clusters (see Zhang's talk)
- **Visual search strategies**
  - Tracking, projection, temporal aspects (see Kim's talk)
  - Situation awareness
- **Perceptual learning**
  - Where should I look?
  - How long do I need to look?
  - What do I see? Learning by experience.