Eye tracking evidence for V1 Saliency Hypothesis (V1SH) from an anomalous visual search behavior: central-peripheral dichotomy in top-down versus bottom-up processes

Junhao Liang, Severin Maher and Li Zhaoping

Eberhard Karls University of Tübingen, Max Planck Institute for Biological Cybernetics

junhao.liang@uni-tuebingen.de, li.zhaoping@tuebingen.mpg.de / www.lizhaoping.org

A search image

Visual items: disk-pairs tilted 45 degree from vertical in a grey background.

Non-targets: uniformly-oriented hetero-pairs of disks

- Target, non-targets can be clockwise or anticlockwise tilted from vertical.
- A target parallel, rather than perpendicular, to non-targets is easier to find [2]

V1 saliency hypothesis (V1SH): Bottom-up saliency of a visual fraction is signaled by the highest activated V1 neuronal response relative to the highest V1 neuronal response to other locations.

Background and motivation

A novel framework to study biological vision [1] proposes that vision is a three-stage process of encoding, selection, and decoding, with two important theories: V1SH and CPD.

Three stages of vision against our visual search task

Encoding

Selection / looking / searching

Decoding / seeing / recognizing

we can recognize whether a disk pair is in the target:

Top-down selection / bottom-up attention

A counter-intuitive V1SH prediction

A target parallel, rather than perpendicular, to non-targets is easier to find [2]

Condition

Parallel

Perpendicular

Target among non-targets

V1 neuronal response

Higher

Lower

Saliency suppression

Stronger

Weaker

Most activated V1 neurons

Saliency enhancement

Example receptive fields of V1 neurons more activated by disk pairs (superposed)

Orthogonalized V1 neurons

Attentional selection

Behavioral outcome

Attraction of bottom-up attention

Shorter RT

More salient

Less salient

Bottom-up attentional selection

Central-peripheral dichotomy (CPD)

Top-down feedbacks from higher cortex are weaker in peripheral vision. Thus, central-peripheral vision is more involved in top-down/-up processing.

Central-peripheral dichotomy (CPD)

How to measure RTsel / RTelec

- RTsel: the reaction time (RT) to select the target. Approximating RTsel / RTelec by RTelec, we approximate RTsel / RTelec by RTelec is better.

- RTtele: the reaction time (RT) to decode/recognize the target. We can approximate RTtele / RTelec by the post-saccadic recognition time. RTtele

How to measure RTtele / RTelec

- RTtele: the reaction time (RT) to decode/recognize the target. We can approximate RTtele / RTelec by the post-saccadic recognition time. RTtele

A counter-intuitive V1SH prediction

A target parallel, rather than perpendicular, to non-targets is easier to find [2]

Condition

Parallel

Perpendicular

Target among non-targets

V1 neuronal response

Higher

Lower

Saliency suppression

Stronger

Weaker

Most activated V1 neurons

Saliency enhancement

Example receptive fields of V1 neurons more activated by disk pairs (superposed)

Orthogonalized V1 neurons

Attentional selection

Behavioral outcome

Attraction of bottom-up attention

Shorter RT

More salient

Less salient

Bottom-up attentional selection

Central-peripheral dichotomy (CPD)

Top-down feedbacks from higher cortex are weaker in peripheral vision. Thus, central-peripheral vision is more involved in top-down/-up processing.

Central-peripheral dichotomy (CPD)

How to measure RTsel / RTelec

- RTsel: the reaction time (RT) to select the target. Approximating RTsel / RTelec by RTelec, we approximate RTsel / RTelec by RTelec is better.

- RTtele: the reaction time (RT) to decode/recognize the target. We can approximate RTtele / RTelec by the post-saccadic recognition time. RTtele

How to measure RTtele / RTelec

- RTtele: the reaction time (RT) to decode/recognize the target. We can approximate RTtele / RTelec by the post-saccadic recognition time. RTtele

Conclusions and outlook

1. RTgaze is shorter in parallel condition – very significant among the instructed subjects, but not significant among the non-instructed subjects.

2. RTreport is shorter in parallel condition – significant among the instructed subjects (but not among the non-instructed subjects), which replicated the finding in [2].

3. RTgaze is shorter when the saccade to the target has a smaller amplitude, suggesting a better saccadic integration for decoding/recognition. The pre-saccadic preview can start to aid spatial information when the target is more central, within about 10 visual degree in eccentricity, before the saccade.

4. Targets parallel to non-targets are more likely to be vetoed maybe because their orientation is more similar to non-targets in decoding (rather than in selection).

5. Subsequent learn to make decisions with less top-down processing after experience of sufficient number of trials.

References


Eye tracking evidence for V1 Saliency Hypothesis (V1SH) from an anomalous visual search behavior: central-peripheral dichotomy in top-down versus bottom-up processes

Junhao Liang, Severin Maher and Li Zhaoping

Eberhard Karls University of Tübingen, Max Planck Institute for Biological Cybernetics

junhao.liang@uni-tuebingen.de, li.zhaoping@tuebingen.mpg.de / www.lizhaoping.org

A visual search task

Visual items: disk-pairs tilted 45 degree from vertical in a grey background.

Target: a home-pair of disks

Non-targets: uniformly-oriented hetero-pairs of disks

Condition A: parallel (to the target) Condition B: perpendicular (to the target)

Most activated V1 neurons

Saliency enhancement

Example receptive fields of V1 neurons more activated by disk pairs (superposed)

Orthogonalized V1 neurons

Attentional selection

Behavioral outcome

Attraction of bottom-up attention

Shorter RT

More salient

Less salient

Bottom-up attentional selection

Central-peripheral dichotomy (CPD)

Top-down feedbacks from higher cortex are weaker in peripheral vision. Thus, central-peripheral vision is more involved in top-down/-up processing.

Central-peripheral dichotomy (CPD)

How to measure RTsel / RTelec

- RTsel: the reaction time (RT) to select the target. Approximating RTsel / RTelec by RTelec, we approximate RTsel / RTelec by RTelec is better.

- RTtele: the reaction time (RT) to decode/recognize the target. We can approximate RTtele / RTelec by the post-saccadic recognition time. RTtele

How to measure RTtele / RTelec

- RTtele: the reaction time (RT) to decode/recognize the target. We can approximate RTtele / RTelec by the post-saccadic recognition time. RTtele

Conclusions and outlook

1. RTgaze is shorter in parallel condition – very significant among the instructed subjects, but not significant among the non-instructed subjects.

2. RTreport is shorter in parallel condition – significant among the instructed subjects (but not among the non-instructed subjects), which replicated the finding in [2].

3. RTgaze is shorter when the saccade to the target has a smaller amplitude, suggesting a better saccadic integration for decoding/recognition. The pre-saccadic preview can start to aid spatial information when the target is more central, within about 10 visual degree in eccentricity, before the saccade.

4. Targets parallel to non-targets are more likely to be vetoed maybe because their orientation is more similar to non-targets in decoding (rather than in selection).

5. Subsequent learn to make decisions with less top-down processing after experience of sufficient number of trials.

References


