V1 serves as a motor cortex for visually guided saccades: evidence from human and monkey behavioral and neural data

Poster presentation at MütTü 2020 oculomotor research meeting, Tübingen, Germany Feb 14-15, 2020

Li Zhaoping, University of Tübingen, Max Planck Institute for Biological Cybernetics, Germany. www.lizhaoping.org, l_zhaoping@tuebingen.mpg.de

Abstract

I proposed in the late 1990s and early 2000s that the primary visual cortex, V1, creates a bottom-up saliency map from visual inputs (Li, 1999, 2002). This is the V1 Saliency Hypothesis (V1SH). According to V1SH, the highest V1 response to any visual location represents this location’s saliency, defined as the strength of exogenous attentional attraction. Hence, the visual location evoking the largest V1 response is most likely to attract visual attention or evoke a gaze shift driven by bottom-up visual inputs. For example, a vertical bar in a background of horizontal bars is salient; so is a red dot among green dots, because the feature singletons evoke higher V1 responses than the homogeneous background items. Contextual influences via the intracortical horizontal connections make V1 neurons’ responses dependent on contextual inputs outside the classical receptive field (e.g., Knierim & van Essen 1992). These contextual influences are such that neurons tuned to the same or similar features suppress each other, so that the neurons preferring and responding to the homogenous background inputs suppress each other’s responses. The neurons responding to the feature singletons escape such suppression and the singletons are thus salient. Here I review some of the supporting evidence in human and monkey data from past works by my collaborators and myself. In lower vertebrates, the saliency map resides in the superior colliculus/optic tectum (Zhaoping 2016). More info at http://www.lizhaoping.org/zhaoping/V1Salency.html

V1 Saliency Hypothesis (V1SH, Li 1999, 2002): V1 creates a bottom-up saliency map to guide attentional/gaze shifts

V1SH explains why color, orientation, and motion feature singletons attract attention

V1 neurons are tuned to orientation, color, or motion direction, and some of them are tuned simultaneously to color and orientation, or to motion direction and orientation

V1SH prediction: Attention capture by an eye-of-origin singleton — because V1 neurons are tuned to eye-of-origin, and V1 has iso-eye-of-origin suppression — confirmed experimentally (Zhaoping 2008, 2012)

V1 neurons respond to the search target were measured while the monkey was doing the task. Usually it took more than 200 ms after stimulus onset before the saccadic onset. At around 40–50 ms after the stimulus onset, given the same visual inputs, trials with the initial V1 neural responses to the target were more likely to evoke faster and accurate subsequent saccades to the target. This latency, at 40–60 ms, is so short that higher cortical areas or superior colliculus could not have provided this saliency signal to V1. This latency agrees with the time course of the contextual influences observed in Knierim and van Essen 1992. White et al. 2017 did not see this saliency signal at this short latency, but they saw a signal at a latency longer than that in superior colliculus, in their monkeys doing a similar task. Yan et al. 2018 showed that a signal at such a longer latency arises from top-down factors.

References


V1SH fully predicts reaction times (RTs) using zero parameters, suggesting that V2 and other visual cortical areas do not contributing to saliency computation

Koene & Zhaoping 2007, Zhaoping & Zhe 2019

Mathematically, V1SH predicts the RT, \( RT_{\text{CMO}} \), of detecting singleton pop out by triple-feature (CMO) uniqueness in color (C), motion direction (M) and orientation (O) from RTs for single- and double-feature uniqueness (RT\(_C\), RT\(_M\), RT\(_O\), RT\(_{CM}\)) by equation

\[
\min(RT_{CM}, RT_{CO}, RT_{MO}) = \min(RT_{CM}, RT_{CO}, RT_{MO})
\]

This equation uses no parameters, and is derived using the observation that V1 has no cells simultaneously tuned to color, orientation, and motion direction. V2 and other extrastriate cortical areas do have such cells, so if they were involved in the saliency computation, the equation above would become an inequality.

ERT data is consistent with the V1 prediction suggesting that V2 and above do not contribute to the saliency computation (though they could contribute to guide attention by top-down endogenous control).