Computations of top-down attention by modulating V1 dynamics

David Berga & Xavier Otazu
Computer Vision Center, Autonomous University of Barcelona, Barcelona, Spain

The human visual system processes information defining what is visually conspicuous (saliency) to our perception, guiding eye movements towards certain objects depending on scene context and its feature characteristics [1]. However, attention has been known to be biased by top-down influences (relevance), which define voluntary eye movements driven by goal-directed behavior and memory. We propose a unified model of the visual cortex able to predict, among other effects, top-down visual attention and saccadic eye movements. First, we simulate activations of early mechanisms of the visual system (RGC/LGN), by processing distinct image chromatic opponencies with Gabor-like filters. Second, we use a cortical magnification function to reproduce foveation towards V1 retinotopy. Third, we feed these signals to an excitatory-inhibitory neurodynamic model (Figure 1) of lateral interactions in V1 [3, 4] as a saliency mechanism [2]. Fourth, projections towards the SC (modeled as WTA-like computations) determine the targets of fixations and saccade sequences. Fifth and last, we integrate a top-down inhibition process by simulating retrieval of visual representations as goal-directed selection processes (DLPFC/FEF), later projected towards V1/SC. These top-down representations will modulate the prediction of visual relevance during visual search tasks, where its weights (orientation, scale and opponency) are mapped as cortical signals from early visual areas for each exemplar/category. Our results show (Figure 2) that our model predictions of eye movements improve by including the aforementioned top-down computations. In addition, our model has previously seen to simultaneously reproduce visual discomfort, brightness and chromatic induction effects.

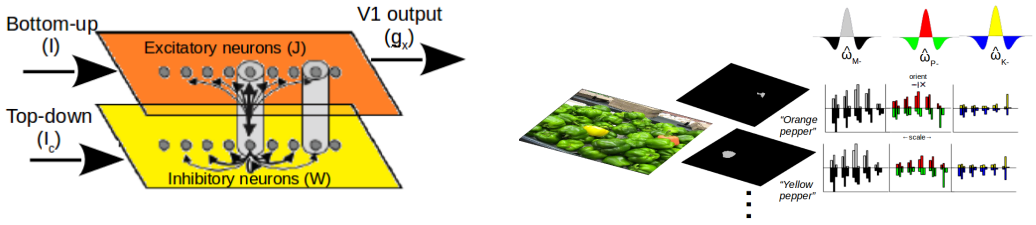

Figure 1: Left: Our excitatory-inhibitory model of lateral interactions of V1 cells (see [4], Supplementary Material). Right: Example of weights of memory formations as the average of wavelet coefficients at distinct Magno-,Parvo- and Konio- pathways.

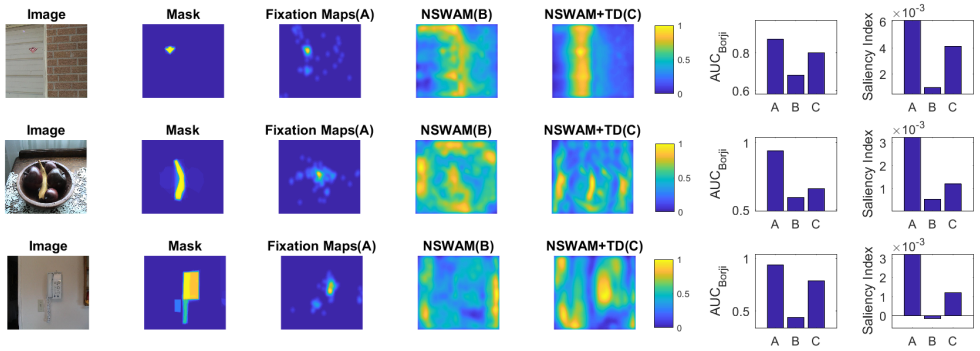

Figure 2: Examples of model output performance given bottom-up only (NSWAM) and with top-down selection (NSWAM+TD). Area Under ROC (AUC_{Borji}) compares predictions with human fixations, and the Saliency Index (SI) is calculated as the energy of the a map inside the mask (St) compared to the one outside the mask (Sb), SI=(St-Sb)/Sb.

References

[1] D. Berga, X. R. Fdez-Vidal, X. Otazu, V. Leborán, and X. M. Pardo. Psychophysical evaluation of individual low-level feature influences on visual attention. *Vision Research*, 154:60–79, jan 2019.

[2] D. Berga and X. Otazu. A neurodynamic model of saliency prediction in v1. *in Review, arXiv:1811.06308*, 2018.

[3] Z. Li. A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10(4):903–940, may 1998.

[4] O. Penacchio, X. Otazu, and L. Dempere-Marco. A neurodynamical model of brightness induction in v1. *PLoS ONE*, 8(5):e64086, may 2013.

This work was funded by the Spanish Ministry of Economy and Competitivity (DPI2017-89867-C2-1-R), Agencia de Gestió d'Ajuts Universitaris i de Recerca (AGAUR) (2017-SGR-649), and CERCA Programme / Generalitat de Catalunya.