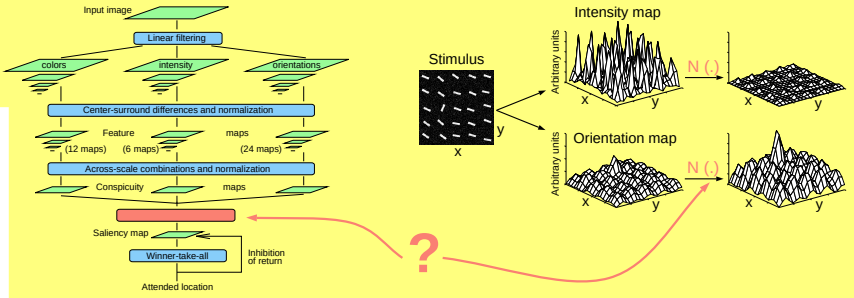
A Simple Model of Long-Range Interactions for the Computation of Salience

Laurent Itti, Christian Scheier, Beena Khurana and Christof Koch
California Institute of Technology, Computation and Neural Systems, Pasadena, California 91125

1. Introduction

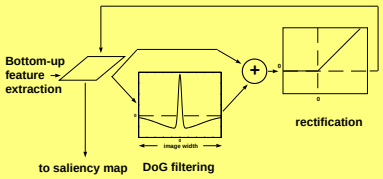
We propose a simple model for the control of visual attention, based on a single topographic "saliency map". Using this model, we specifically address the problem of combining several types of visual features, such as information about object color, intensity and orientation, into a unique scalar measure of salience. A simple read-out of the saliency map then allows the model to predict scanpaths followed by the attentional spotlight.

2. A model of bottom-up attention

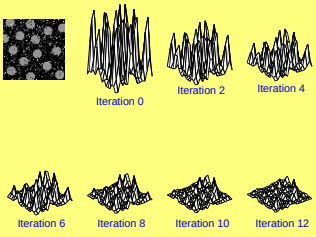


3. Long-range competition

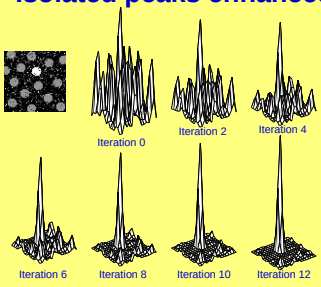
3.1. Principle



3.2. Uniform activity suppressed

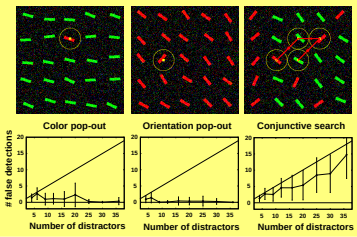


3.3. Isolated peaks enhanced

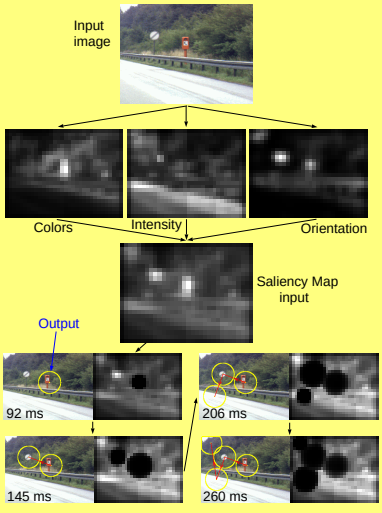


4. Results

4.1. Synthetic images



4.2. Examples of natural scenes

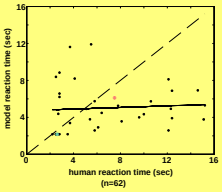


The model was tested on a database of 44 test images provided by Dr. A. Toet from the TNO Human Factors Research Institute, The Netherlands. This database was collected in order to measure human search times on a difficult task, and to compare models.

To allow comparison between model and human data, we scaled model time such that:

- the the model shifts towards no more than 2 to 4 locations per second
- a delay of 1.5 sec. is added to model time to account for motor response

With such scaling, the model found the target faster than humans in 3/4 of images.



5. Conclusion

Despite its simplicity, this model demonstrates strong performance at detecting target in complex natural scenes. The success of this approach suggests that a unique topographic saliency map is an efficient way to control, in a purely bottom-up manner, where the attentional spotlight is to be deployed. Our conclusion is supported by recent electrophysiological evidence for neurons which specifically encode for salience irrespectively of the type of simple visual target used, in particular in monkey area LIP.

Supported by the National Science Fundation (Caltech Neuromorphic Engineering ERC) and the Office of Naval Research.