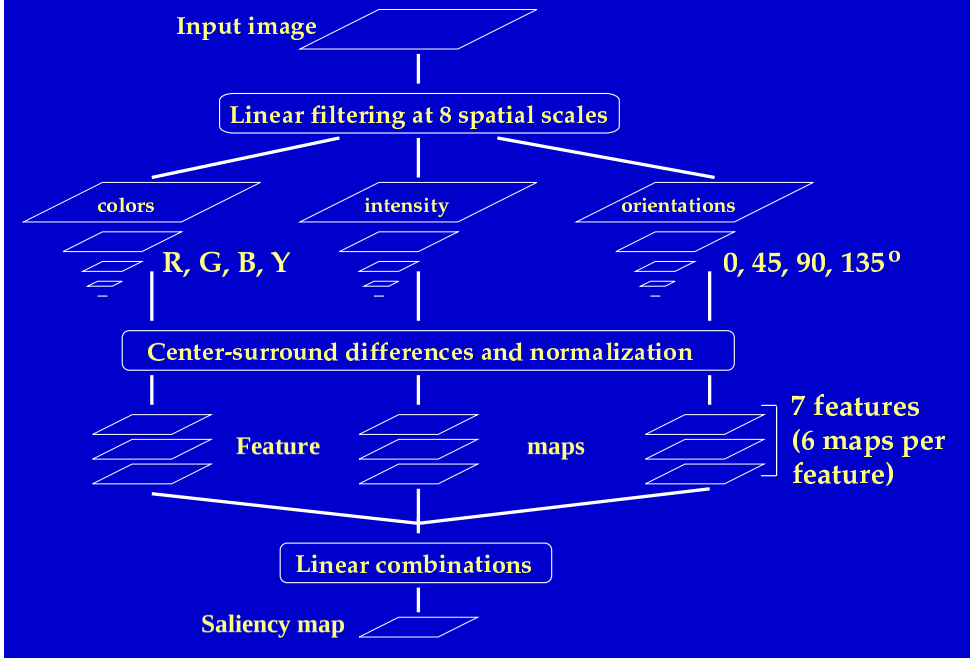
A Trainable Model of Saliency-based Visual Attention

L. Itti, E. Niebur, J. Braun and C. Koch

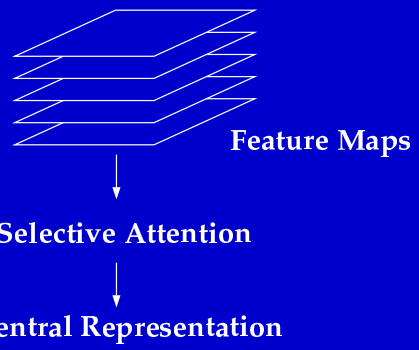
California Institute of Technology and
Johns Hopkins University

Extraction of Visual Features



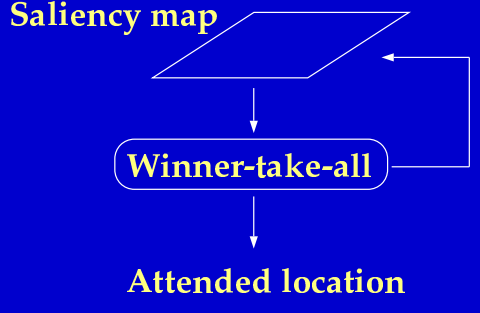
Overview

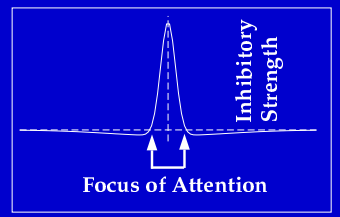
Model Architecture

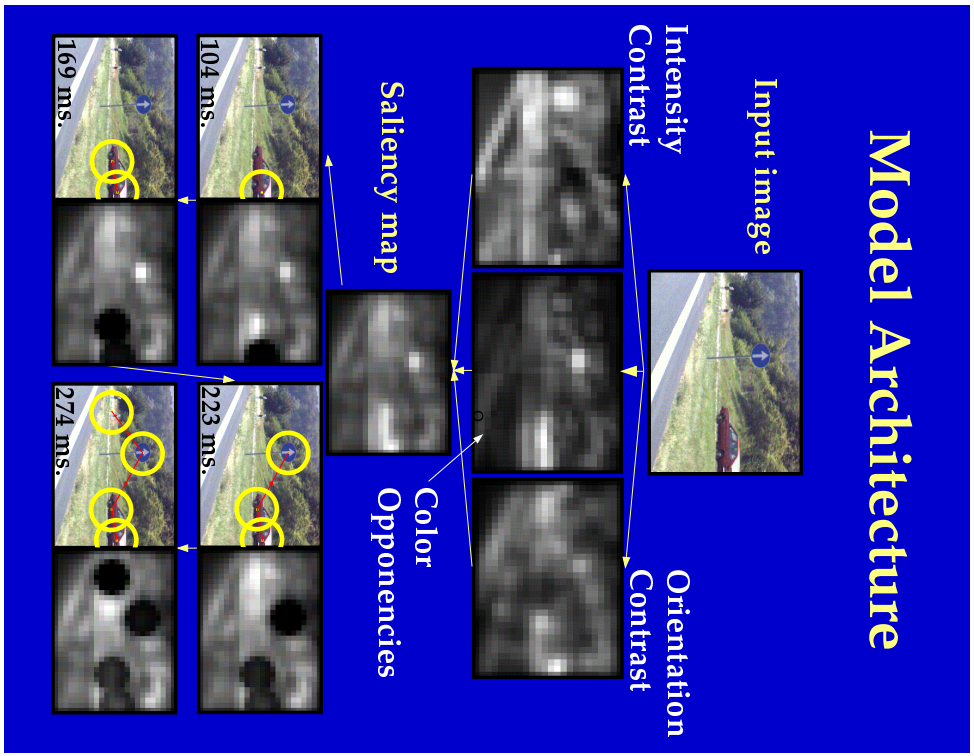


- Extraction of early visual features
- saliency map
- dynamical attention shifting

Dynamical control of Attention







Combination of Features

Supervised learning of feature map weights:

$w \leftarrow w + \eta$ (max. inside target - average max. outside)

Naive linear model

Trained linear model (n=32)

Combination of Features

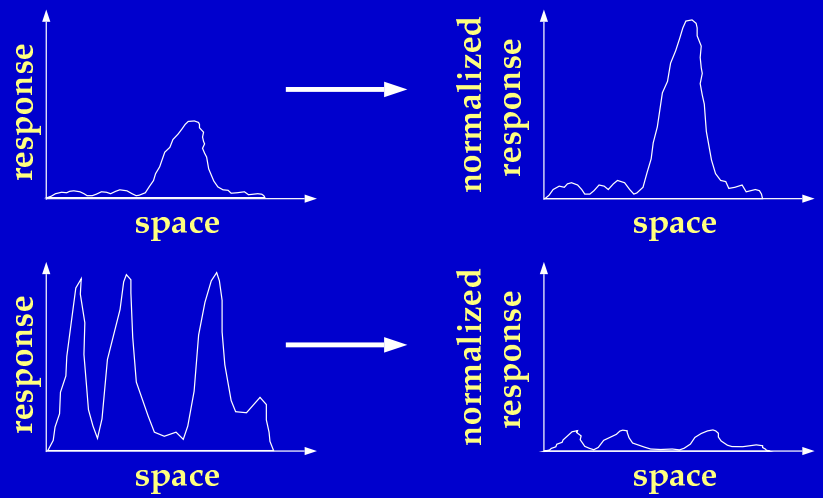
Summation

Color

Intensity

Orientation

Combination of features



Combination of Features

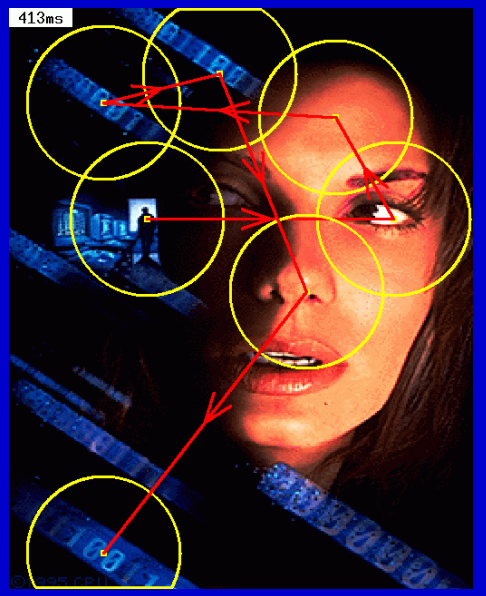
Contents-based non-linear combination:
Map weight = (global max. - average local max.)^2

Color

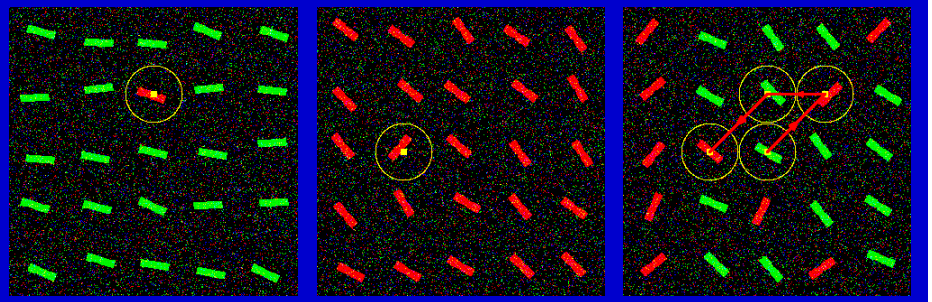
Intensity

Orientation

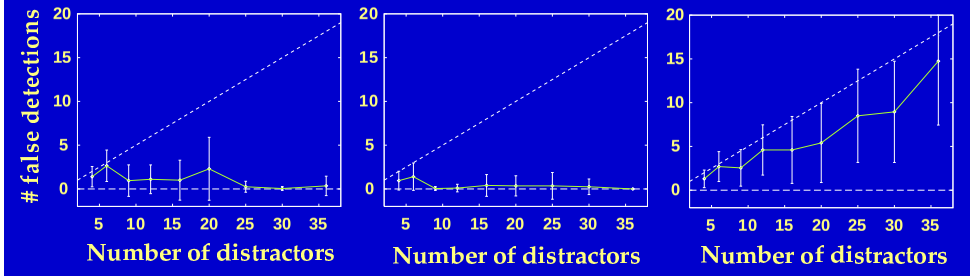
Natural Images



Simulated Psychophysics



Color pop-out | Orientation pop-out | Conjunctive search



Conclusion

Architecture of early primate visual system

Adaptive non-linear feature combination yields best results

"Saliency map" approach works remarkably well on artificial and natural images

1999 Research Portfolio

Laurent Itti California Institute of Technology
Computation and Neural Systems Program
Mail Code 139-74, Pasadena, CA 91125, USA
Tel: +1 626 395-2881 Fax: +1 626 796-8876
Email: itti@klab.caltech.edu
WWW: http://www.klab.caltech.edu/~itti/