V1 is an optimal saliency detector

Dashan Gao (dgao@ucsd.edu) Nuno Vasconcelos (nuno@ece.ucsd.edu)

Statistical Visual Computing Laboratory, Department of Electrical and Computer Engineering, University of California San Diego

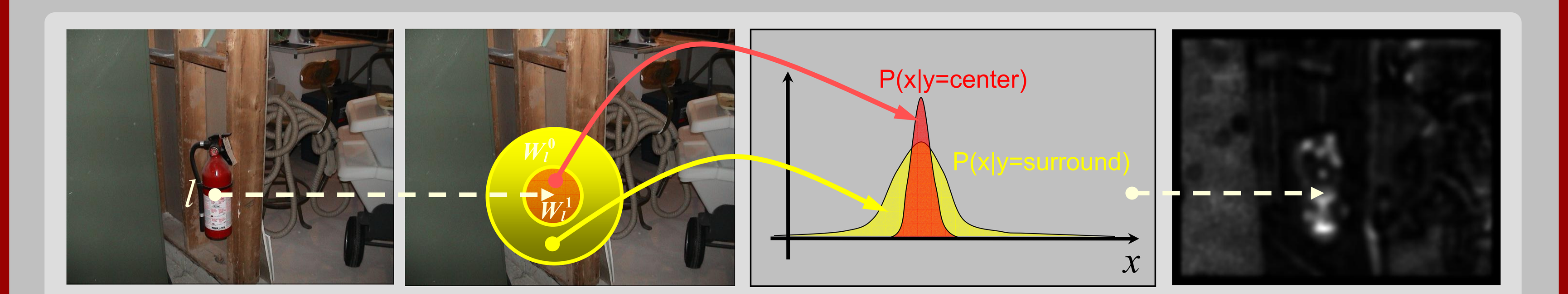
Abstract

- Saliency mechanisms play an important role in directing visual attention

- We propose a new hypothesis: **saliency is a discriminant process**
  - combined with three commonly used principles for neural organization
    - that each layer of perception maximizes the information available to the next (infomax) (Linsker, 1988; Nadal & Parga, 1994; Bell & Sejnowski, 1995)
    - that simple solutions are preferable (computational parsimony) (Occam)
    - that vision is tuned to the statistics of the natural world (Barlow, 1961, 2001; etc.)
- The proposed ***discriminant saliency*** is shown to
  - be consistent with the standard architecture of V1: simple/complex cell, etc.
  - replicate psychophysics of human saliency perception

Bottom-up discriminant saliency

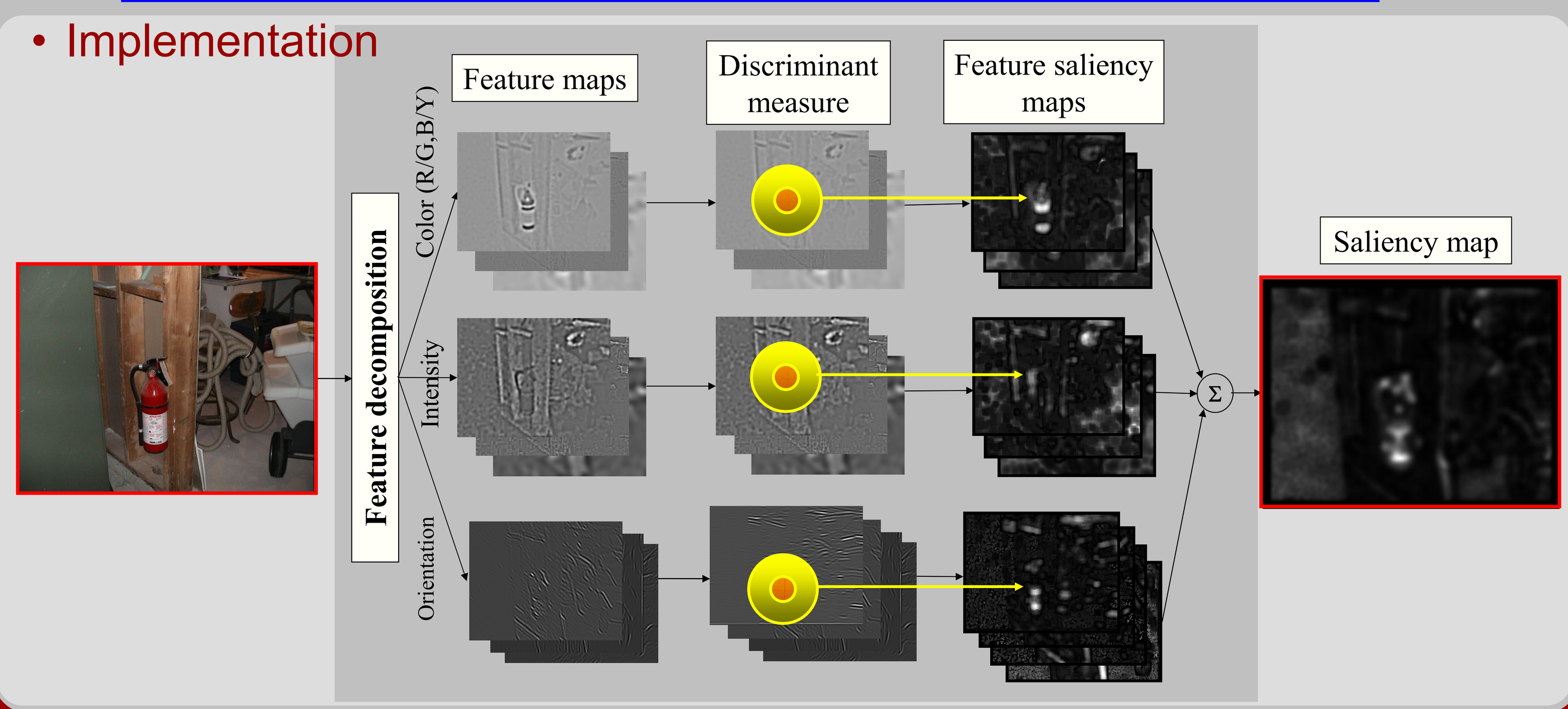
- Discriminant saliency is rooted on a decision-theoretic interpretation of perception
- **Saliency is a *discriminant process***
  - defines, at each image location, a binary classification between the stimulus in the center and that in the surround



- **Center-surround processing is optimal in a decision-theoretic sense**
  - salient locations: that enable the discrimination between center and surround with minimum expected probability of error
  - **Infomax saliency measure: mutual information**

$$S_{(i,j)}(\mathbf{X};Y) = I_{(i,j)}(\mathbf{X};Y) = \sum_i \int_{\mathbf{x}} P_{\mathbf{X},Y}(\mathbf{x},y)\log\frac{P_{\mathbf{X},Y}(\mathbf{x},y)}{P_{\mathbf{X}}(\mathbf{x})P_Y(y)}dx$$

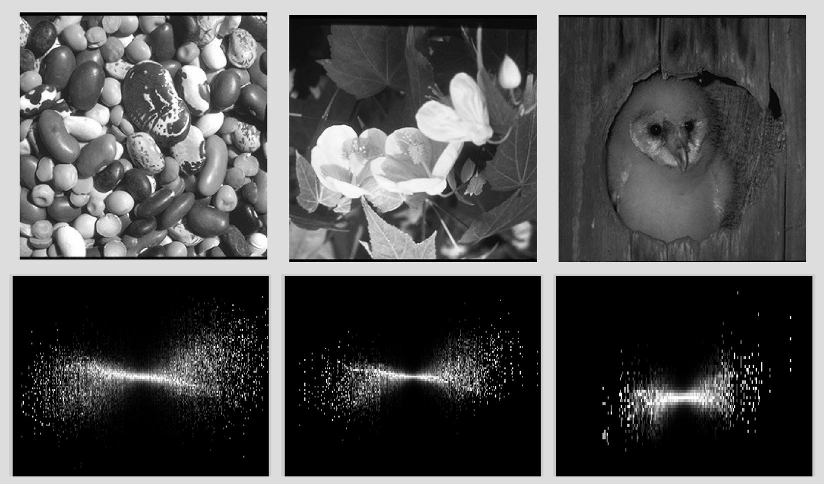
- Implementation



Computational Parsimony and Image Statistics

- Biological visual systems exploit the regularities of the natural stimuli to achieve computational parsimony (Attneave, 1954; Barlow, 1961, 2001; etc.)

➢ Regular pattern of feature dependence (Buccigrossi & Simoncelli,1999; Huang & Mumford, 1999)



Top: three images. Bottom: conditional histogram of the same coefficient, conditioned on its parent.

- bow-tie shape conditional distribution for band-pass image features
- although fine details may vary from scene to scene, coarse structure follows a universal law for all natural scenes
- feature dependencies are not informative about image classes

$$I_{(i,j)}(\mathbf{X},Y) = \sum_{k=1}^{n} I_{(i,j)}(X_k;Y)$$

$$I(\mathbf{X}_{1,n};Y) = \sum_{i=1}^{n} I(X_i;Y) - \sum_{i=2}^{n}[I(X_i;\mathbf{X}_{1,i-1}) - I(X_i;\mathbf{X}_{1,i-1}|Y)]$$

$$\frac{\sum_{i=2}^{n}\left[I_l(X_i;\mathbf{X}_{1,i-1}) - I_l(X_i;\mathbf{X}_{1,i-1}|Y)\right]}{\sum_{i=1}^{n} I_l(X_i;Y)} \approx 0 \qquad \mathbf{X}_{1,i} = \{X_1,\ldots,X_i\}$$

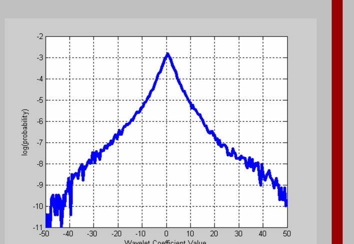
(Vasconcelos & Vasconcelos, 2004)

➢ Generalized Gaussian Density (GGD)

- the marginal distributions of natural image features follow a generalized Gaussian density

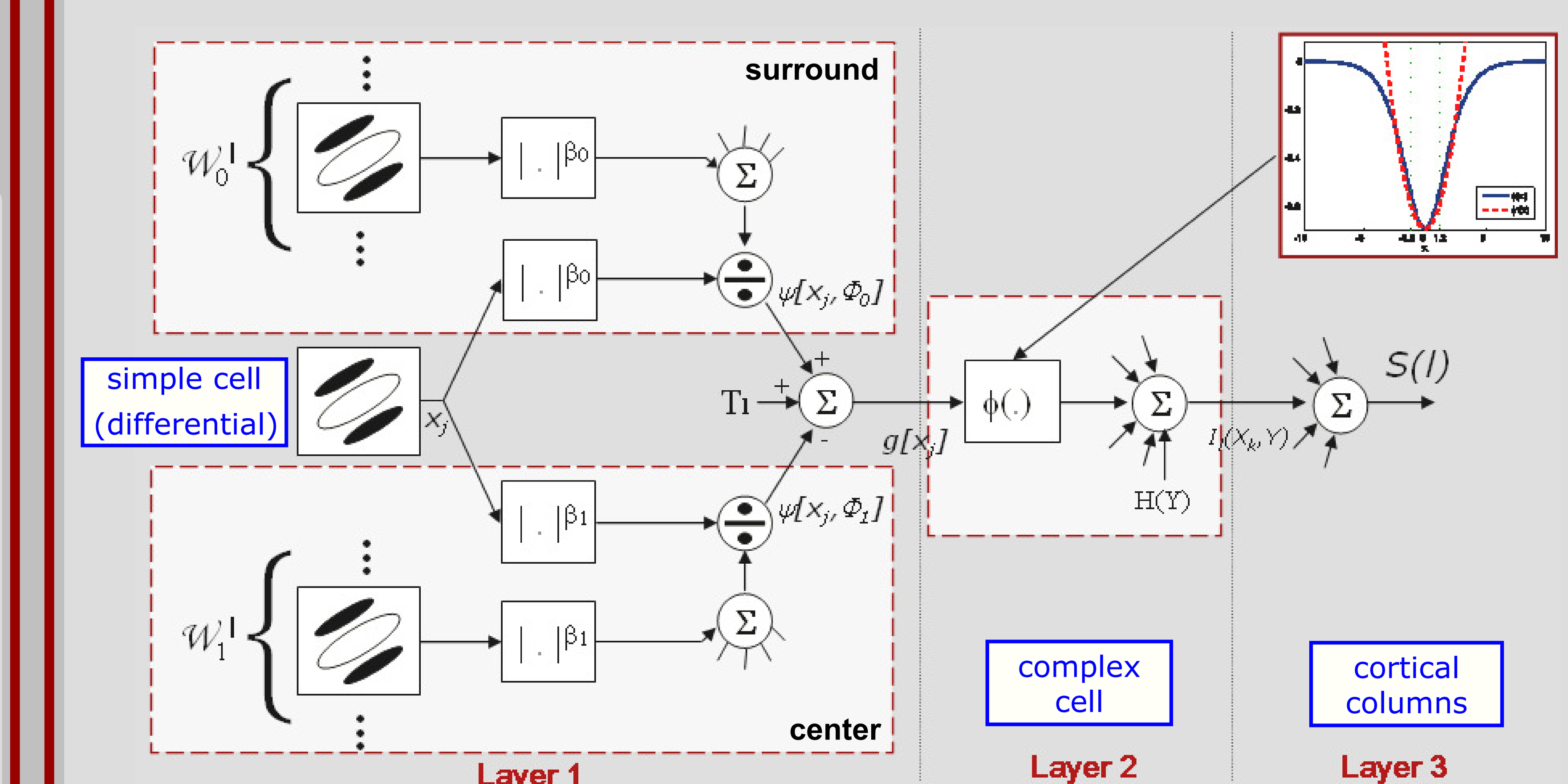
$$p(x;\alpha,\beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)}\exp\left(-(|x|/\alpha)^{\beta}\right)$$

α a *scale parameter*, β a *shape parameter*



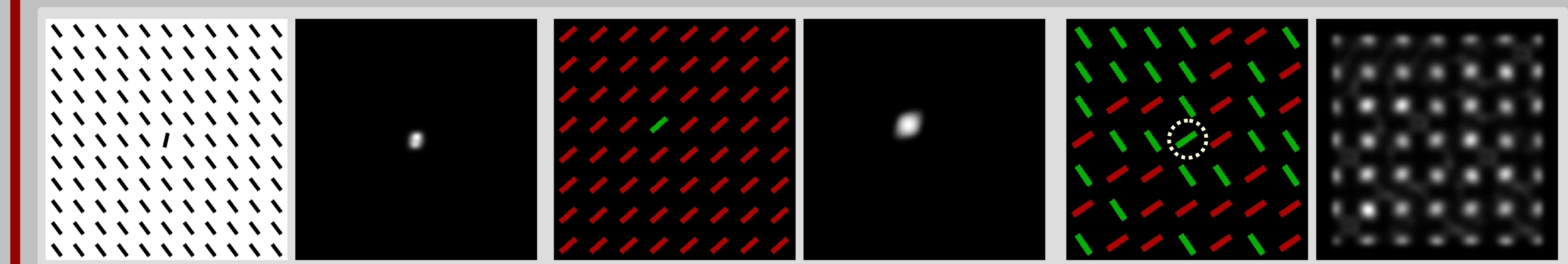
- GGD assumption enables a biologically plausible implementation.

Biologically Plausible Implementation

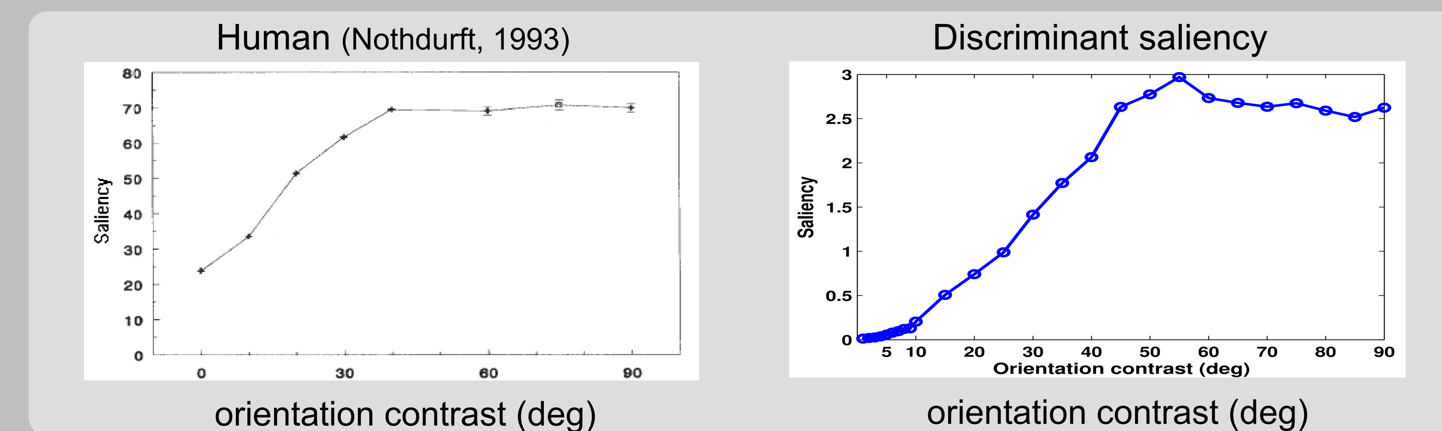


Consistency with Psychophysics

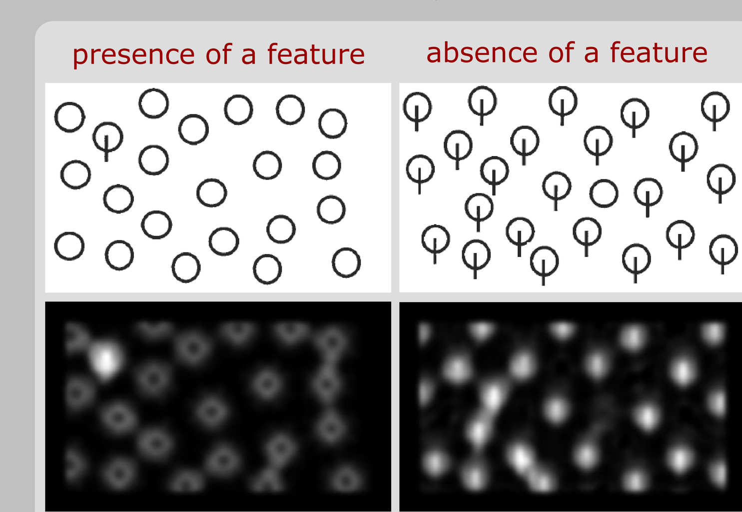
- Single and conjunctive feature search
  - consistent with Feature Integration Theory (Treisman & Gelade, 1980)
  - this is just what the image statistics suggests

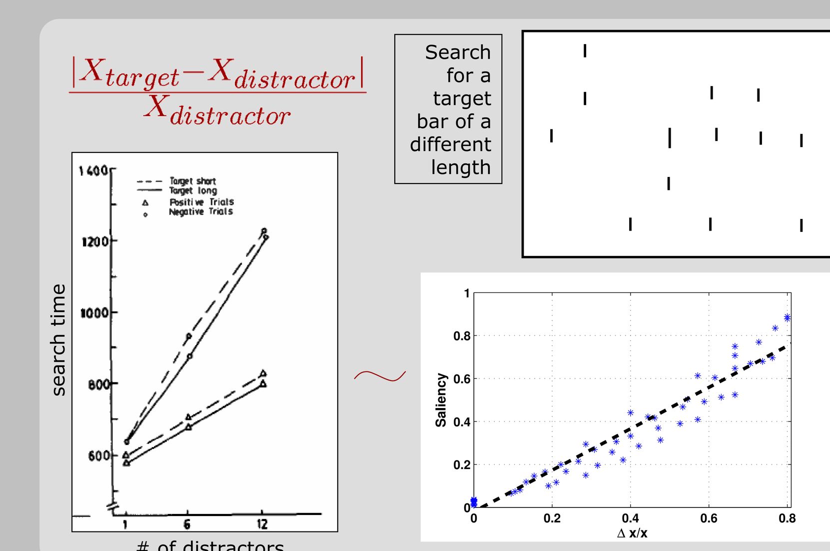


- Nonlinearity



- Visual search asymmetries and Weber's law (Treisman & Gormican, 1988)





Discussion

- Discriminant saliency connects a number of "disjoint" observations from neurophysiology and psychophysics
  - e.g., divisive normalization and saliency asymmetries
- A (unified) holistic functional justification for V1
  - V1 has the capability to optimally detects salient locations in the visual field in a decision-theoretic sense under certain approximations for the sake of computational parsimony
- Statistical inference in V1
  - probability inference, decision rule, and feature selection

cell type	function	description
simple	$-\log P_{X\|Y}[x\|c]$	negative log-likelihood
simple, differential	$\log\frac{P_{X\|Y}[x\|1]}{P_{X\|Y}[x\|0]}$	log likelihood ratio
complex	$I(X;Y)$	mutual information