A new biologically inspired color image descriptor

Jun Zhang[1,2], Youssef Barhomi[1], Thomas Serre[1]
[1]Brown University, [2]Hefei University of Technology

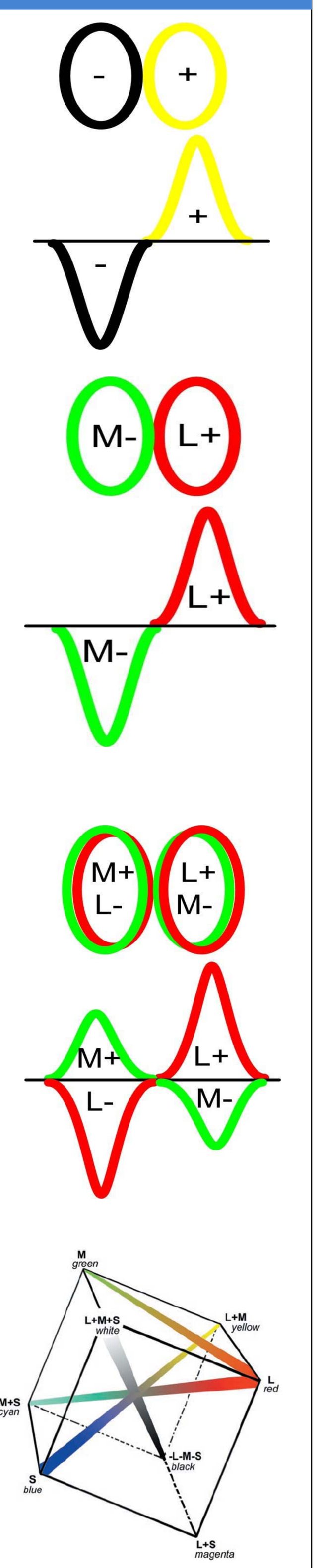
1. Motivation

Luminance neurons with "ON" and "OFF" subunit organization are selective to orientations (and high spatial frequency) thus contribute to edge detection.

Single-Opponent (SO) neurons (color-preferring) with "ON" and "OFF" subunit organization with different cone inputs exhibit little or weak orientation selectivity but strong selectivity to color regions.

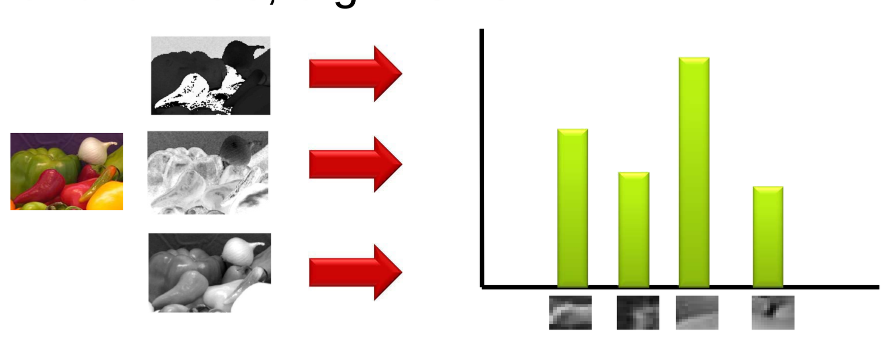
Double-Opponent (DO) neurons (color-luminance) with no defined "ON" and "OFF" subunit organization are selective for both color and orientations (and spatial frequency), and thought to influence the perception of form.

Consider chromatic axes as found in cortex.

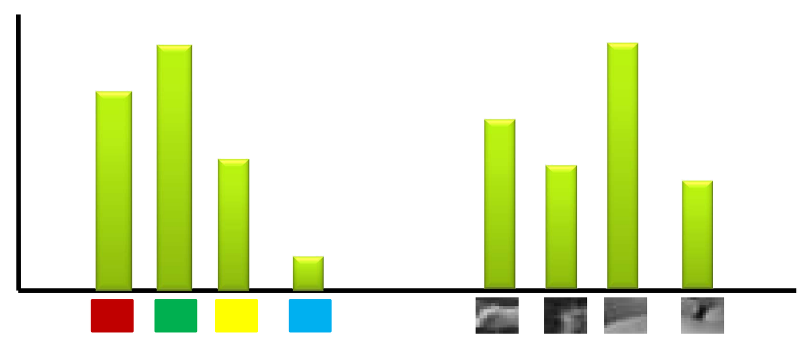


2. Computer vision approaches

- Applying shape-based descriptors on individual color channels, e.g. HSVSIFT.



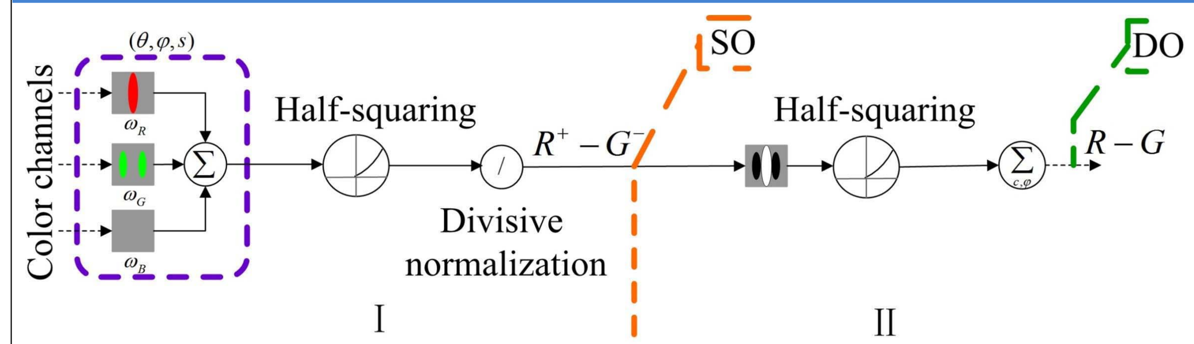
- Concatenation of shape-based descriptors with color histograms, e.g. HueSIFT.



3. Biological mechanisms

- Chromatic and spatial information should be represented **jointly** as done in the primate visual cortex.
- Neurons maintain **positive** firing rates, and neural circuits have been identified for **contrast gain controls.**

4. Approach overview



5. Approach details

Spatio-chromatic sensitivity function

$$f(x,y,\lambda)=w_R R(\lambda) f_R(x,y)+w_G G(\lambda) f_G(x,y)+w_B B(\lambda) f_B(x,y)$$

spatial sensitivity distribution — excitatory/inhibitory center/surround component

SO unit response: $s=I(x,y,\lambda)*f(x,y,\lambda)$

DO unit response: $d=s(x,y,c)*f(x,y)$

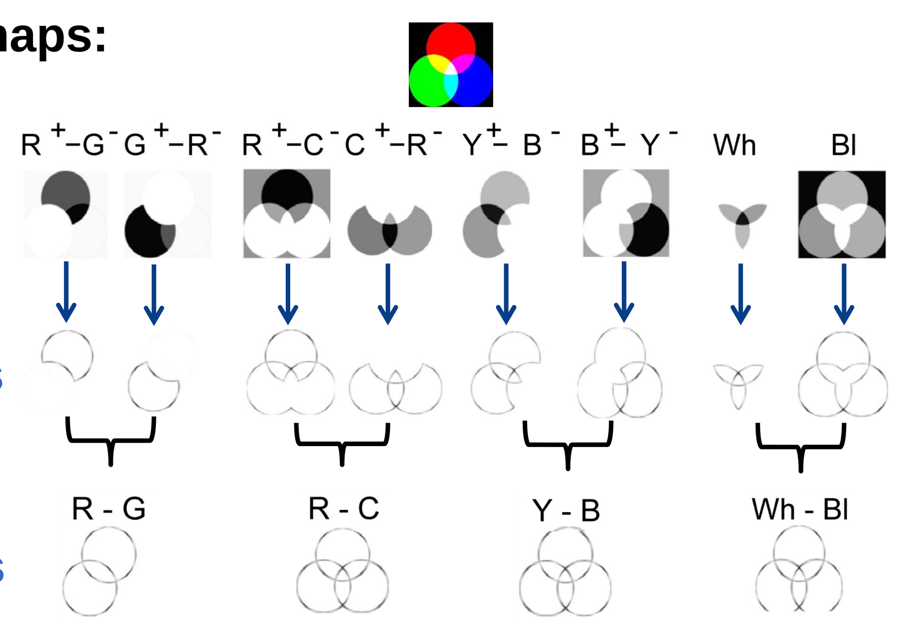
SO: opponent colors
DO: orientations

half-squared

Non-linear operations: $v(x,y,\zeta)=\dfrac{\sqrt{k\times u(x,y,\zeta)}}{\sigma^2+\sum u(x,y,\zeta)}$

divisive normalization pool

Response maps:



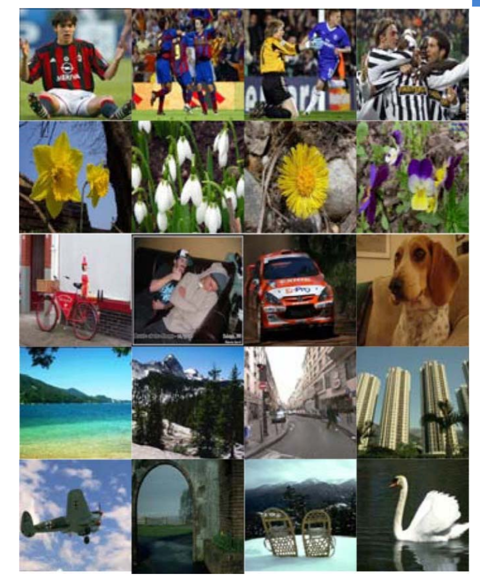
6. Color-gradient comparisons

7. System extensions

Different gradients / filters used in:

- SIFT-based bag-of-words approach
- HMAX model
- GIST algorithm for natural scene categorization
- BSDS500 for contour detection

Datasets



Soccer team
17-category flowers
PASCAL VOC 2007
8-category scenes
BSDS500

8. Results

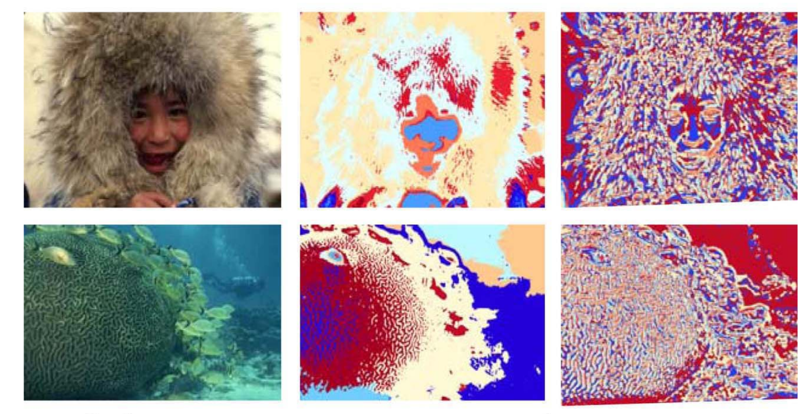
Shape predominant

			PASCAL VOC 2007			
Method	SIFT	HueSIFT	OpponentSIFT	CSIFT	**SODOSIFT**	**SODOHMAX**
AP	40(38.4)	41	43(42.5)	43(44.0)	**46.5**(33.3/39.8)	**46.8**(30.1/36.4)

Color predominant

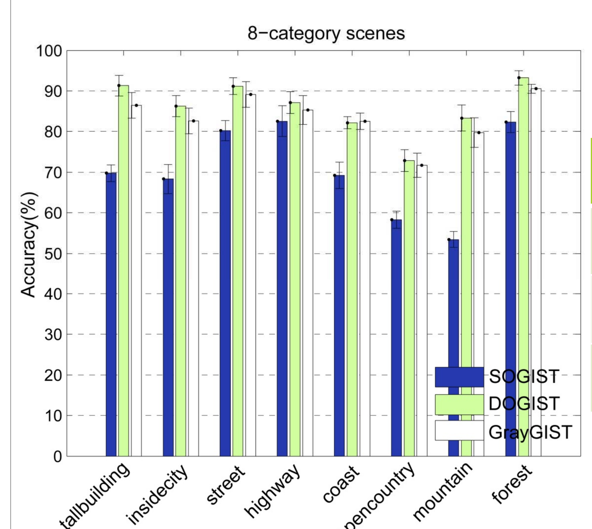
	Soccer team			Flower		
Method	Color	Shape	Both	Color	Shape	Both
Hue/SIFT	69(67)	43(43)	73(73)	58(40)	65(65)	77(79)
Opp/SIFT	69(65)	43(43)	74(72)	57(39)	65(65)	74(79)
SOSIFT/DOSIFT	**82**	**66**	**83**	**68**	**69**	**79**
SOHMAX/DOHMAX	**87**	**76**	**89**	**77**	**73**	**83**

BSDS500

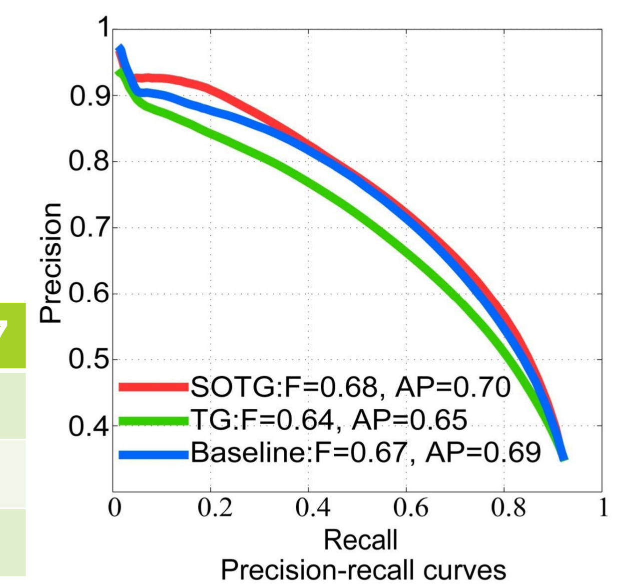


Color-texton map vs. grayscale texton map

			8-category scenes		
Method	GIST	RGBGIST	SOGIST	**DOGIST**	**SODOGIST**
Accuracy	83.5	84.1	70.5	**85.9**	**87.1**



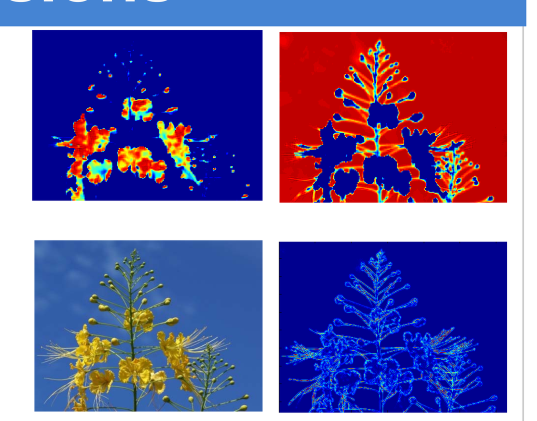
Method	Soccer team	PASCAL VOC 2007
Full model	**82.0/66.0**	**33.3/39.8**
Without half-squaring	62.0/60.0	30.3/36.7
Without normalization	70.0/53.0	32.9/40.7



Precision-recall curves

9. Conclusions

- SO encodes color regions
- DO encodes color edges



- Proposed descriptors perform on par or better than other color and shape descriptors.

10. References

[1] Conway, B.R.: The Journal of Neuroscience, 2007 [2] van de Weijer, J., Schmid, C. ICIP, 2007. [3] van de Sande et al. TPAMI, 2010. [4] Oliva, A., Torralba, A. IJCV, 2001. [5] Arbelaez, P et al. TPAMI, 2010.

11. Acknowledgments

This work was supported by DARPA (N10AP20013) and ONR (N000141110743), the Center for Computation and Visualization of Brown University,, and the Robert J. and Nancy D. Carney Fund for Scientific Innovation. J. Zhang was supported by China Scholarship Council (2010669020).

Software available online! **http://serre-lab.clps.brown.edu/projects/color**